The role of surface glycans in the attachment and invasion of *Neisseria meningitidis* into host epithelial cells

Stephanie Nicole Bartley BSc(Hons)

This thesis is presented for the degree of Doctor of Philosophy at The University of Western Australia.

School of Pathology and Laboratory Medicine.

Discipline of Microbiology and Immunology.

2012.
Abstract

The major surface carbohydrates of Neisseria meningitidis include the capsule, lipoooligosaccharide and the pilin glycan. Each of these structures are regulated and are variably distributed between strains of the species. Although alterations to the surface carbohydrates of the meningococcus are thought to aid in evasion of the host immune response, they also modulate virulence. As such the aims of this study were to examine the capsule, pilin glycan and lipoooligosaccharide structure and regulatory pathways in the attachment and invasion of host epithelial cells.

We compared exemplar strains from two hypervirulent clonal complexes, strain NMB-CDC from ST-8/11 cc and strain MC58 from ST-32/269 cc, in host cell attachment and invasion. Strain NMB-CDC attached to and invaded host cells at a significantly greater frequency than strain MC58. Type IV pili retained the primary role for initial attachment to host cells for both isolates regardless of pilin class and glycosylation pattern. In strain MC58, the serogroup B capsule was the major inhibitory determinant affecting both bacterial attachment to and invasion of host cells. Removal of terminal sialylation of lipoooligosaccharide (LOS) in the presence of capsule did not influence rates of attachment or invasion for strain MC58. However, removal of either serogroup B capsule or LOS sialylation in strain NMB-CDC increased bacterial attachment to host cells to the same extent. Although the level of inhibition of attachment by capsule was different between these strains, the regulation of the capsule synthesis locus by the two-component response regulator MisR, and the level of surface capsule determined by flow cytometry were not significantly different. However, the diplococci of strain NMB-CDC were shown to have a 1.89-fold greater surface area than strain MC58 by
flow cytometry. It was proposed that the increase in surface area without changing the amount of anchored glycolipid capsule in the outer membrane would result in a sparser capsule and increase surface hydrophobicity. Capsulation modulated the surface hydrophobicity of both isolates however encapsulated strain MC58 was more hydrophilic than strain NMB-CDC. In conclusion, strain MC58 adhered less to host cells and had a more hydrophilic surface than strain NMB-CDC. Improved levels of adherence of strain NMB-CDC to cell lines was associated with increased bacterial cell surface and surface hydrophobicity. This study shows that there is diversity in bacterial cell surface area and surface hydrophobicity within *N. meningitidis* and that this phenotype can influence steps in meningococcal pathogenesis.

*N. meningitidis* expresses lipooligosaccharide (LOS) consisting of an inner core of heptose residues (HepI and HepII) decorated with lacto-\textit{N}-neotetraose (LNT) on the α-chain, phosphoethanolamine (PEA) at positions \textit{O}-3 and \textit{O}-6 on HepII, glucose (Glc) at position \textit{O}-3 on HepII, and an \textit{O}-acyetyl group (OAc) at position \textit{O}-3 on the terminal \textit{N}-acyetyl glucosamine (GlcNAc) residue of the γ-chain. *N. meningitidis* strain NMB-CDC cannot phase vary the expression of LNT due to a phase locked LOS glycosyltransferase A (LgtA), but does possess the multiple genes required for the addition of \textit{O}-3 Glc, \textit{O}-3 OAc and \textit{O}-6 PEA groups to the LOS inner core. In the presence of LOS bearing LNT, the decoration of the LOS inner core with \textit{O}-6 PEA and \textit{O}-3 OAc or \textit{O}-3 Glc and \textit{O}-3 OAc or \textit{O}-3 OAc alone promoted meningococcal invasion into the host cells (18-fold, 4-fold and 9-fold increase compared to parental wild-type, respectively). The ability of the LOS inner core decorations to modulate the rate of meningococcal invasion into host cells was not inhibited by the presence of capsule unless the LOS structure contained an \textit{O}-3 PEA residue attached to HepII. The
expression of the LOS inner core structures with LNT promoted invasion into host cells in the absence of the Opa adhesins, suggesting that the LOS structure was acting as an independent ligand for binding a host cell receptor. Lastly, the decoration of the LOS inner core structure with O-6 PEA was up-regulated and terminal sialylation of LNT and capsule expression were down regulated in the presence of host cells. In summary, LOS inner core residues promote meningococcal invasion in the absence of the ability to remove LNT by phase variation in strain NMB-CDC.

*Neisseria gonorrhoeae* expresses lipooligosaccharide (LOS) terminated with N-acetyl D-galactosamine (GalNAc) which efficiently binds host receptors to mediate bacterial invasion into host cells. GalNAc is not present as a component of the LOS of the closely related pathogen, *Neisseria meningitidis*, but is a component of the polysaccharide capsules of serogroup 26E and Z strains. The ability to synthesise UDP-GalNAc has not been previously determined in either bacterial species. In other bacterial species UDP-galactose 4-epimerase (GalE) is a bi-functional enzyme capable of epimerising UDP-N-acetyl glucosamine (GlcNAc) to UDP-GalNAc. *N. gonorrhoeae* was shown to carry a GalE allele that is bi-functional and epimerises UDP-GlcNAc and UDP-Glc to UDP-GalNAc and UDP-Gal, respectively. The bi-functional gonococcal GalE can be converted to a mono-functional enzyme unable to epimerise GlcNAc by replacement of a Serine at position 299 in the substrate binding pocket. Conversely, a mono-functional meningococcal GalE was converted to a bi-functional enzyme by the replacement of the phenylalanine at position 300 with a serine residue in the binding site pocket. An analysis of the GalE alleles from both commensal and pathogenic *Neisseria sp.* indicates that the bi-functional GalE allele is found in all sequenced gonococci, *N.*
lactamica, N. polysaccharea, but N. meningitidis and N. cinerea carry either mono- and bi-functional alleles of GalE.
Declaration of Contribution

Results chapters 2, 3 and 4 are presented here as manuscripts which have either been submitted for publication (Chapter 2), or are intended to be submitted for publication (Chapters 3 and 4). The details of the authors and their contributions are detailed below.

Chapter 2. Revision submitted to PLoS One, 1/12/12

Attachment and invasion of *Neisseria meningitidis* to host cells is related to surface hydrophobicity, bacterial cell size and capsule.

Stephanie N. Bartley\(^1\), Yih-Ling Tzeng\(^2,3\), Kathryn Heel\(^4\), Chiang W. Lee\(^1\), Shakeel Mowlaboccus\(^1\), Torsten Seemann\(^5\), Wei Lu\(^1\), Ya-Hsun Lin\(^6\), Catherine S. Ryan\(^6\), Christopher Peacock\(^1\), David S. Stephens\(^2,3\), John K. Davies\(^6\), and Charlene M. Kahler\(^\ast\)

\(^1\) School of Pathology and Laboratory Medicine, The University of Western Australia, Perth, Australia.

\(^2\) Veterans Affairs Medical Center and

\(^3\) Emory University School of Medicine, Atlanta, Georgia USA.

\(^4\) Centre for Microscopy, Characterisation and Analysis, and Translational Cancer Pathology Laboratory, School of Pathology and Laboratory Medicine, The University of Western Australia, Perth, Australia.

\(^5\) Victorian Bioinformatics Consortium, Monash University, Melbourne, Australia.

\(^6\) Department of Microbiology, Monash University, Melbourne, Australia.

\(^\ast\) Corresponding author address: e-mail: Charlene.Kahler@uwa.edu.au
**Stephanie Bartley:**

Construction of all strains and plasmids in this study (i.e. those not otherwise referenced in Table 2.1).

Performed the bacterial attachment and invasion assays

Performed the immunoblots blots, and analysed capsule and lipoooligosaccharide by SDS-PAGE

Purified the capsule

Preparation of samples for flow cytometry and high resolution under flow using the Imagestream X (AMNIS)

Analysis of flow cytometry and Imagestream X data

Assessed viable counts of organisms

Sequence of pilC2 region and opaA/B/D/J loci

Prepared total RNA and performed qRT-PCR and analysis

Constructed figures 2.4-2.6, 2.8

Constructed Table 2.3, 2.4

Experimental design of the entire project with Charlene Kahler

Wrote complete first draft of manuscript

**Yih-Ling Tzeng:**

Performed Electrophoretic Mobility Shift assays

Constructed Figure 2.7

Designed qRT-PCR primers for the study and supervised Stephanie in her lab for two weeks

Experimental design
Kathryn Heel:
Operator of Flow cytometry and AMNIS
Analysis of flow cytometry and Imagestream\textsuperscript{X} data
Constructed Figures 2.2 and 2.3

Chiang lee:
Performed Hydrophobicity Interaction Chromatography
Preparation of samples for flow cytometry and Imagestream\textsuperscript{X}

Shakeel Mowlaboccus:
Performed Microbial Adhesion To Solvents assays
Constructed Table 2.5

Torsten Seemann:
Genome sequencing of strain NMB-CDC, genome assembly and annotation

Wei Lu:
Constructed Figure 2.1

Ya-Hsun Lin:
Constructed plasmid pJKD3341

Catherine Ryan:
Constructed plasmid pJKD3172
Supervisor of Ya-Hsun Lin
Christopher Peacock:
Supervisor of Wei Lu
Edited the manuscript

David Stephens:
Contributed to genome sequencing project
Edited the manuscript

John Davies:
Experimental design of the pilin switching strains
Supervisor of Ya-Hsun Lin
Edited the manuscript

Charlene Kahler:
Overall experimental design of the entire project
Analysis and interpretation of data with all authors
Constructed plasmid pJKD2728
Supervisor of Stephanie Bartley, Chiang Lee and Shakeel Mowlaboccus
Edited the manuscript
Corresponding author with the journal
Chapter 3. To be submitted to J. Bact pending additional experiments

**Contact-dependent regulation of lipooligosaccharide inner core structure determines meningococcal invasion into host cells.**

Stephanie N. Bartley¹, Jhuma Ganguly², Russell W. Carlson², Yih-Ling Tzeng³, Steven A. R. Webb⁴, and Charlene M. Kahler¹.

¹ School of Pathology and Laboratory Medicine, The University of Western Australia, Perth, Australia.
² Complex Carbohydrate Research Center, University of Georgia, Athens, USA.
³ Veterans Affairs Medical Center and Emory University School of Medicine, Atlanta, USA.
⁴ School of Medicine and Pharmacology, The University of Western Australia, Perth, Australia

**Stephanie Bartley:**

Construction of all strains and plasmid in this study (i.e. those not otherwise referenced in Table 3.1).

Performed the bacterial attachment and invasion assays

Performed all immunoblots and analysis of lipooligosaccharide (LOS) by SDS-PAGE

Performed all β-galactosidase assays

Analysis of genome sequences for the distribution of LOS biosynthetic genes (Section 3.4.1 and Table 3.2)
Construction of Figures 3.2, 3.4-3.7

Wrote complete first draft of manuscript

Experimental design of the entire project with Charlene Kahler

**Jhuma Ganguly:**

Preparation and MALDI-TOF MS analysis of LOS

Contributed data to Figure 3.3 and Table 3.3

**Russell Carlson:**

Supervisor of Jhuma Ganguly

Analysis and interpretation of MALDI-TOF MS data

Constructed Figure 3.3 and Table 3.3

Experimental design with Charlene Kahler

Edited the manuscript

**Yih-Ling Tzeng:**

Constructed plasmid pYT328

Technical advise with the β-galactosidase assays and supervised Stephanie for two weeks in her laboratory

**Steven Webb:**

Supervisor of Stephanie Bartley
Charlene Kahler:

Overall experimental design of the entire project

Analysis and interpretation of data with all authors

Supervisor of Stephanie Bartley

Edited the manuscript

Corresponding author with the journal
A bifunctional UDP-galactose 4-epimerase results in the biosynthesis of N-acetyl galactosamine for gonococcal LOS and the serogroup 29E and Z capsules of N. meningitidis

Stephanie N. Bartley¹, Keith Stubbs², Odile Harrison³, Martin Maiden³, Charlene M. Kahler¹

¹School of Pathology and Laboratory Medicine, University of Western Australia, Perth, Australia
²School of Chemistry and Biochemistry, University of Western Australia, Perth, Australia
³University of Oxford, Department of Zoology, South parks Road, Oxford OX1 3PS, United Kingdom

Stephanie Bartley:

Construction of all strains and plasmids listed in Tables 4.1 and 4.2 not otherwise referenced

Cloning expression and purification of the GalE proteins

Sequence analysis and phylogenetic tree construction

All bacterial attachment and invasion assays

All immunoblots and analysis of LOS by SDS-PAGE
Experimental design of the entire project with Charlene Kahler

Wrote complete first draft of manuscript

**Keith Stubbs:**

Technical advice with protein purification

Reaction and High Performance Liquid Chromatography analysis of GalE activity

Experimental design of the protein purification section with Stephanie Bartely

Edited the manuscript and contributed Table 2

**Odile Harrison:**

Contribution of *galE* sequences from all unpublished strains of *neisseria*

Contributed to the analysis of the phylogenetic data and interpretation

**Martin Maiden:**

Contribution of *galE* sequences from all unpublished strains

Supervisor of Odile Harrison

Contributed to the analysis of the phylogenetic data and interpretation

**Charlene Kahler:**

Overall experimental design of the entire project

Analysis and interpretation of data with all authors

Supervisor of Stephanie Bartley

Edited the manuscript

Corresponding author with the journal
<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stephanie Bartley</td>
<td>S. Bartley</td>
</tr>
<tr>
<td>Russell Carlson</td>
<td></td>
</tr>
<tr>
<td>John Davies</td>
<td>J. Davies</td>
</tr>
<tr>
<td>Jhuma Ganguly</td>
<td></td>
</tr>
<tr>
<td>Odile Harrison</td>
<td>O. Harrison</td>
</tr>
<tr>
<td>Kathryn Heel</td>
<td></td>
</tr>
<tr>
<td>Charlene Kahler</td>
<td>C. M. Kahler</td>
</tr>
<tr>
<td>Chiang Lee</td>
<td>L.</td>
</tr>
<tr>
<td>Ya-Hsun Lin</td>
<td></td>
</tr>
<tr>
<td>Wei Lu</td>
<td>W. Lu</td>
</tr>
<tr>
<td>Martin Maiden</td>
<td></td>
</tr>
<tr>
<td>Shakeel Mowlaboccus</td>
<td>S.</td>
</tr>
<tr>
<td>Christopher Peacock</td>
<td></td>
</tr>
<tr>
<td>Catherine Ryan</td>
<td>C. Ryan</td>
</tr>
<tr>
<td>Torsten Seemann</td>
<td></td>
</tr>
<tr>
<td>David Stephens</td>
<td></td>
</tr>
<tr>
<td>Keith Stubbs</td>
<td></td>
</tr>
<tr>
<td>Yih-Ling Tzeng</td>
<td></td>
</tr>
<tr>
<td>Steven Webb</td>
<td>S. Webb</td>
</tr>
</tbody>
</table>

I, Charlene Kahler, the coordinating supervisor of Stephanie Bartley, can indicate that the list of contributions made by Stephanie to her research project is correct.

C. M. Kahler.
Conference Presentations and Awards

Underline indicates presenter.

**Poster presentations**


5. **Stephanie Bell** and Charlene Kahler. (2009) “The lipooligosaccharide composition of Neisseria meningitidis affects invasion and is regulated in...
response to contact to epithelial cells.” Combined Biological Sciences Meeting.


Oral presentations


List of awards

2008 ASM poster prize CBSM, $250

2008 Convocation travel award, $2500

2009 ASM-AUS/ASM-USA Millis-Colwell postgraduate research award, $5000

The convocation travel award and the Millis-Colwell award were used to visit Assistant Professor Yih-Ling Tzeng’s laboratory at Emory University for two weeks to learn a variety of techniques and to attend and present at the American Society for Microbiology in Philadelphia, May 2009.

2009 ASM branch award, covered conference registration, approx $200

2010 ASM BD student award, Covered return flights to Sydney, conference registration, accommodation and an oral presentation, approx $1000
Acknowledgements

As the final piece to be written, to conclude this experience, it is humbling to consider all the friends, family and colleagues who have offered their support and counsel.

First, I must thank Charlene for providing me with this opportunity to complete my PhD in her laboratory. Without her support following the completion of honours in her laboratory, I would not have been able to commence doctoral studies, let alone complete them. Thank you for your understanding and the opportunity to prove my ability, to grow, for the challenge and support to be ambitious.

I must express my gratitude to the Amanda Young Foundation. Without their financial support I would not have been able to undertake a PhD at all. I thank many other researchers for their professional and personal advise, Harry Sakellaris, Geoff Stewart, Michael Wise, David Sutton, Yih-Ling Tzeng, Steve Webb, Paul Rigby, Kathy Heel and Keith Stubbs.

I am blessed with an incredible family. To my parents, thank you for your unconditional love, understanding and support. You made me the person I am today, and I hope I have made you are proud. Thank you for all you have done for me, and all that you have put up with from me.

My husband deserves the greatest thanks in the completion of my PhD. Blair has sacrificed so much for me to finish. His enduring love and patience means the world to
me. I love you with everything I have and all that I am. I look forward to more, less stressful, time together, and one day a little family of our own.

To the friends who have heard all about this journey, thanks for your friendship. Fayth, Jamie, Alex, Susannah, Naomi, Megan-your friendship means more than you will know.

Finally to all the people who have passed through Charlene’s lab over the years, thanks for the experience. Thanks to Susannah Piek for sharing this experience, I am grateful to have had someone here to walk through this with. Lea-Ann Kirkham, Jessica Scoullar and Chiang Lee, thanks for your knowledge and friendship during the course of my PhD studies.
Table of contents

Abstract ......................................................................................................................... iii

Declaration of Contribution ...................................................................................... vii

Conference Presentations and Awards ............................................................... xvii

Poster presentations ............................................................................................... xvii

Oral presentations .................................................................................................. xx

List of awards .......................................................................................................... xxi

Acknowledgements ............................................................................................... xxii

Table of contents ................................................................................................... xxiv

List of Figures .......................................................................................................... xxxi

List of Tables ............................................................................................................ xxxv

Chapter 1. Introduction ......................................................................................... 1

1.1 Neisseria meningitidis ....................................................................................... 2

1.1.1 Typing of N. meningitidis ........................................................................... 5

1.1.2 Mechanisms of variation of the meningococcus ....................................... 9

1.2 The interaction of N. meningitidis with the host .......................................... 13

1.2.1 Colonisation of the host epithelium ......................................................... 16

1.2.2 Invasion of host epithelial cells ............................................................... 17

1.2.3 Interaction with the host immune system ............................................. 18

1.2.4 Crossing the blood-brain barrier (BBB) ................................................ 20
1.3 The structure and regulation of virulence determinants of *N. meningitidis* .......................................................... 21

1.3.1 Pili............................................................................................................. 21

1.3.2 Outer-membrane adhesins .................................................................. 29

1.3.3 Lipooligosaccharide.............................................................................. 39

1.3.4 Capsule .................................................................................................. 45

1.5 Aims of my thesis .................................................................................... 56

1.6 References .............................................................................................. 57

Chapter 2. Attachment and invasion of *Neisseria meningitidis* to host cells is related to surface hydrophobicity, bacterial cell size and capsule. ............... 80

2.1 Abstract ....................................................................................................... 81

2.2 Introduction ................................................................................................ 82

2.3 Methods ...................................................................................................... 84

2.3.1 Bacterial strains and growth conditions ............................................. 84

2.3.2 Genome sequencing, assembly and annotation ................................. 85

2.3.3 Cell lines and culture ......................................................................... 88

2.3.4 Construction of mutants of strain NMB-CDC and MC58 .................. 88

2.3.5 Transformation protocols .................................................................... 90

2.3.6 Attachment and invasion assays .......................................................... 92

2.3.7 Determination of capsule, LOS, pilin and Opa phenotype .................. 93

2.3.8 qRT-PCR .............................................................................................. 94

2.3.9 Flow cytometry for the detection of meningococcal capsule and bacterial cell size using high throughput microscopy ........................................ 95
2.3.10 Electrophoretic mobility shift assay (EMSA) .............................................. 96
2.3.11 Capsule purification, electrophoresis and staining ..................................... 97
2.3.12 Hydrophobic interaction chromatography (HIC) ........................................ 97
2.3.13 Microbial adhesion to solvents (MATS) ..................................................... 97

2.4 Results ................................................................................................................... 98

2.4.1 Genetic comparison of N. meningitidis strain NMB-CDC (ST-8 cc)
with N. meningitidis strain MC58 (ST-32 cc) ............................................................ 98
2.4.2 Bacterial cell size differs between strain NMB-CDC and MC58. .............. 99
2.4.3 Strain NMB-CDC attaches to and invades human epithelial cells at
significantly higher rates than strain MC58. ......................................................... 102
2.4.4 Type IV pili, regardless of class and glycosylation phenotype, are
necessary for initial attachment by both strains to epithelial cells.............. 106
2.4.5 Surface expressed sialic acid inhibits bacterial association and
invasion of cell lines by strain MC58 to a greater extent than in
strain NMB-CDC. .................................................................................................. 107
2.4.6 Strain NMB-CDC and strain MC58 express the same amount of
surface capsule. .................................................................................................... 111
2.4.7 The surface of strain NMB-CDC is more hydrophobic than strain
MC58...................................................................................................................... 114

2.5 Discussion .......................................................................................................... 116

2.6 Acknowledgments .............................................................................................. 121

2.7 References .......................................................................................................... 122
Chapter 3. Contact-dependent regulation of lipooligosaccharide inner core structure determines meningococcal invasion into host cells................................. 130

3.1 Abstract ...................................................................................... 131

3.2 Introduction ................................................................................. 132

3.3 Methods ...................................................................................... 135

3.3.1 Bacterial strains and growth conditions........................................ 135

3.3.2 Attachment and invasion assays .................................................. 137

3.3.3 Construction of meningococcal mutants and LacZ reporter strains ...... 138

3.3.4 Transformation protocols .......................................................... 142

3.3.5 Determination of capsule, LOS, pilin and Opa phenotype .................. 142

3.3.6 Tricine-SDS-PAGE analysis of LOS preparations ........................... 143

3.3.7 β-galactosidase assay ................................................................. 144

3.3.8 Preparation of LOS and analysis. ................................................. 145

3.4 Results ....................................................................................... 145

3.4.1 Distribution of LOS inner core biosynthesis genes is associated with different lineages of N. meningitidis ........................................... 145

3.4.2 The construction of invasive LOS immunotype structures in strain NMB-CDC. .................................................................................. 147

3.4.3 Invasion but not attachment to Detroit 562 epithelial cells is moderated by the LOS inner core structure of N. meningitidis strain NMB-CDC. .................................................................................. 152

3.4.4 The presence of sialic acid masks the ability of LOS inner cores possessing O-3 PEA groups to initiate invasion into host cells. ............. 155
3.4.5 LOS can trigger meningococcal invasion into host cells in the absence of Opa adhesin ...................................................... 157

3.4.6 The LOS inner core transferases are regulated in response to host cell contact ................................................................. 160

3.5 Discussion .............................................................................. 160

3.6 Acknowledgments ................................................................. 166

3.7 References ............................................................................. 166

Chapter 4. A bi-functional UDP-galactose 4-epimerase results in the biosynthesis of N-acetyl galactosamine for gonococcal LOS and the serogroup 29E and Z capsules of N. meningitidis ................................................... 172

4.1 Abstract ................................................................................. 173

4.2 Introduction ............................................................................ 173

4.3 Materials and Methods .......................................................... 177

4.3.1 Bacterial strains and growth conditions .................................... 177

4.3.2 Cell lines and culture .......................................................... 178

4.3.3 Construction of expression strains .......................................... 178

4.3.4 Protein purification .............................................................. 180

4.3.5 HPLC.................................................................................. 181

4.3.6 Next generation sequencing and annotation of sequence data with BIGSDB ................................................................. 182

4.3.7 Sequence analysis and phylogenetic tree construction............... 183

4.3.8 Complementation of GalE using pCMK666 shuttle vector .......... 183

4.3.9 Transformation protocols ..................................................... 184
4.3.10 SDS-PAGE and western immunoblot ........................................184
4.3.11 Tricine-SDS-PAGE analysis of LOS preparations ....................185
4.3.12 Attachment and invasion assays ..............................................186
4.4 Results .......................................................................................187
  4.4.1 Neisseria sp. possess mono- and bi-functional epimerases which are
dependent on a single amino acid in the binding pocket ....................187
  4.4.2 Sequence analysis and distribution of galE alleles across Neisseria
genera ...............................................................................................189
  4.4.3 Characterisation of a putatively bi-functional GalE allele of
  N. meningitidis strain NMB-CDC .....................................................194
  4.4.4 Inactivation of the bi-functional GalE allele 85 in strain NMB-CDC
  results in increased pilin mobility and is complemented by the
  mono-functional GalE .........................................................................195
  4.4.5 Effect of GalNAc on attachment and invasion of NMB-CDC ....197
4.5 Discussion ....................................................................................198
4.6 Acknowledgments ..........................................................................202
4.7 References ....................................................................................203

Chapter 5. General Discussion ..........................................................208

  5.1. Implications to the model of attachment and invasion of
  N. meningitidis ..................................................................................209
  5.2 Variation in cell size and surface hydrophobicity in the
  interaction with the host ....................................................................213
  5.3 Synthesis of N-acetyl galactosamine in Neisseria .........................214

xxix
5.4 Conclusion ................................................................. 215
5.2 References ............................................................... 215

Appendix A. Abbreviations ................................................. 217
A(I) Nucleotides ............................................................. 217
A(II) Amino Acids ........................................................... 217
A(III) Units .................................................................. 218
A(IV) Symbols ............................................................... 218
A(V) Abbreviations ......................................................... 219

Appendix B. Sequence of galE alleles examined in
   Chapter 4 .................................................................... 223

Appendix C. Strains of Neisseria used to examine the
distribution of galE alleles .............................................. 259

Appendix D. Publication relating to thesis ......................... 264
D.1 ABSTRACT ................................................................. 265
D.2 INTRODUCTION .......................................................... 267
D.3 MATERIALS AND METHODS .................................. 268
D.4 RESULTS ................................................................. 271
D.5 DISCUSSION ............................................................ 276
D.6 ACKNOWLEDGEMENTS ......................................... 278
D.7 REFERENCES ........................................................... 278
List of Figures

Figure 1.1 Global incidence of meningococcal epidemics ...................................................... 3
Figure 1.2 Incidence of meningococcal disease by age .............................................................. 3
Figure 1.3 Mechanisms of phase variation in Neisseria meningitidis ....................................... 10
Figure 1.4 Mechanisms of antigenic variation in Neisseria meningitidis .................................. 12
Figure 1.5 Organisation of the human nasopharynx ................................................................. 14
Figure 1.6 Interaction of N. meningitidis with the host .............................................................. 15
Figure 1.7 Distribution of capsule types amongst carriage and disease isolates ...................... 19
Figure 1.8 The process of pilus biogenesis in Neisseria .............................................................. 23
Figure 1.9 Pilin glycan biosynthesis and process of glycosylation ........................................... 25
Figure 1.10 Structure and transport of LOS .............................................................................. 42
Figure 1.11 Global distribution of serogroups in disease .......................................................... 46
Figure 1.12 Genetic organisation of the capsule locus of N. meningitidis ................................... 48
Figure 1.13 Distribution of serogroup with respect to clonal complex ...................................... 55

Figure 2.1. Venn diagram showing the distribution of genes as orthologous groups (OG) across four genomes representing ST-8 cc (NMB-CDC, G2136), ST-11 cc (FAM18) and ST-32 cc (MC58) ................ 100

Figure 2.2. AMNIS ImageStreamX can be used to identify single diplococci in an NMB-CDC sample under flow ................................................................. 103
Figure 2.3. The surface area of strain NMB-CDC is larger than that of strain MC58…………………………………………………………………………………………………………………………104

Figure 2.4. Strain NMB-CDC associates with and invades epithelial cell lines at greater rates than strain MC58…………………………………………………………………………………………………………………………105

Figure 2.5. The role of Type IV pili is conserved in bacterial association with host cells…………………………………………………………………………………………………………………………108

Figure 2.6. The effect of capsule and total sialic acid on attachment and invasion of strains NMB-CDC and MC58………………………………………………………………………………………………………………………………………………110

Figure 2.7. MisR directly binds to the synA-ctrA promoter region…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………
Figure 3.6 LOS triggered invasion is repressed in a background devoid
of Opa proteins.................................................................159

Figure 3.7 Capsule, terminal LOS sialylation and LOS inner core
decorations are regulated in response to host cell contact
as determined by β-galactosidase reporter assay......................161

Figure 4.1 The structure of lipooligosaccharide of *N. meningitidis* and
*N. gonorrhoeae*.....................................................................................175

Figure 4.2 Alignment of mono-functional and bi-functional GalE sequences.....188

Figure 4.3 Phylogenetic tree of *galE* sequences from *Neisseria* spp........191

Figure 4.4 Phylogenetic tree of *galE* sequences from *N. meningitidis*.........193

Figure 4.5 Complementation of the GalE mutant of *N. meningitidis*
strain NMB-CDC with bi-functional GalE #85 and
mono-functional GalE #2...................................................................197

Figure 4.6 The rates of attachment and invasion of host
epithelial cells is hindered in the absence of a
UDP-Glu/GlcNAc 4-epimerase..........................................................198

Figure 5.1 The model of attachment and invasion of *N. meningitidis*............210

Figure 5.2 Proposed alternative model for attachment and invasion of
*N. meningitidis* ..................................................................................212

Figure D.S1: K-mer depth across assemblies containing variant
sequence segments.............................................................................286
| Figure D.S2 | ..................................................................................287 |
| Figure D.S3 | ..................................................................................289 |
| Figure D.S4 | ..................................................................................298 |
| Figure D.S5 | ..................................................................................301 |
| Figure D.S6 | ..................................................................................304 |
| Figure D.S7 | ..................................................................................306 |
| Figure D.S8 | ..................................................................................309 |
# List of Tables

Table 1.1. Characteristics of clonal complexes of *N. meningitidis*.................................8

Table 1.2 Structures of LOS immunotypes........................................................................40

Table 1.3 Summary of *N. meningitidis* capsule information........................................47

Table 2.1. Strains used in this study....................................................................................86

Table 2.2. Plasmids used in this study................................................................................91

Table 2.3. A comparison of virulence encoding genes between strain NMB-CDC and MC58.........................................................................................................................101

Table 2.4. qRT-PCR of gene expression in strain NMB-CDC and MC58 and their respective $\Delta$misR mutants.................................................................................................112

Table 2.5 Adhesion of bacterial strains to hexadecane in the MATS assay ..................116

Table 3.1. Strains and plasmids used in this study.............................................................136

Table 3.2. Summary of the distribution of LOS biosynthesis genes amongst clonal complexes of *N. meningitidis*.......................................................................................148

Table 3.3. Composition of the ions from the M7$\Delta$lgtG and M7$\Delta$lgtGΔlot.............152

Table 4.1 Strains used in this study....................................................................................178

Table 4.2 Plasmids used in this study................................................................................179

Table 4.3 Analysis of UDP-Glucose and UDP-GlcNAc 4-epimerase activity in various *Neisseria* strains.............................................................................................................188

Table D.1: Average k-mer depth of seq1, the sequence assembly that was identical to the reference sequence, for each experiment.........................................................280

Table D.2: Variant sequences detected in the first experiment with pilE from *N. gonorrhoeae* FA1090.................................................................................................................281
Table D.3: Variant sequences detected in the repeat experiment with pilE from *N. gonorrhoeae* FA1090......................................................282

Table D.4: Variant sequences detected in the first experiment with pilE from *N. gonorrhoeae* MS11..........................................................283

Table D.5: Variant sequences detected in the repeat experiment with pilE from *N. gonorrhoeae* MS11..........................................................284

Table D.6: Variation in the opaK gene of *N. gonorrhoeae* strain FA1090.........285

Table D.S1: Oligonucleotide primers used to amplify the genes of interest.........311

Table D.S2: Sequence quality data..........................................................312
Chapter 1. Introduction.
1.1 *Neisseria meningitidis*

The genus Neisseria comprises around 10 species, the majority of which are commensal organisms asymptotically colonising the epithelium of the upper respiratory tract. However, this Family contains two obligate human pathogens, *N. gonorrhoeae* and *N. meningitidis*. *N. gonorrhoeae* is never considered normal flora and infects the urogenital tract (148) while *N. meningitidis* can be present in the nasopharyngeal epithelium as normal flora (33). Although it is considered a commensal, *N. meningitidis* also causes meningococcal septicaemia when replication occurs in the bloodstream, and meningococcal meningitis when the bacteria cross the blood-brain barrier and replication occurs in the cerebral spinal fluid (CSF) leading to inflammation of the meninges (226).

*Neisseria meningitidis* is a Gram-negative diplococcus which solely inhabits *Homo sapiens* as the host with no other known vector or reservoir. Meningococcal disease has a worldwide incidence with the majority of occurrences being sporadic cases. Many epidemics have also occurred worldwide (Figure 1.1,(314)) particularly in Africa, Europe, Americas and Asia. Although this organism is considered a deadly pathogen, it is carried as a commensal organism of the human nasopharynx with approximately 10% of adults harbouring the bacteria in non-epidemic settings (43, 253). The rates of transmission and carriage increase with crowding, leading to the higher rates of carriage within households and in army barracks where carriage rates are much greater (99, 201).

The incidence of meningococcal disease is greatest in the 0-2 yr and 15-19 yr age groups (Figure 1.2,(157)). The reasons for this age distribution are based on risk factors
**Figure 1.1 Global incidence of meningococcal epidemics.**
Depicted above are the major epidemics (defined as “an unacceptable incidence rate requiring emergency control measures”-World Health Organisation) of meningococcal disease between the years of 1971 and 1997 worldwide. The sites of the epidemic are indicated with ● and the year/s of the outbreak are indicated. Taken from World Health Organisation, (314).

**Figure 1.2 Incidence of meningococcal disease by age in England (2006-11).**
The average number of cases per annum (2006/7-2010/11) are plotted on the y-axis against age on the x-axis by serogroup with serogroup B the dominant capsule type shown in blue. Taken from Ladhani et al., (157).
of meningococcal disease. In the first instance, the disease in the 0-2 yr age group, like that of individuals with complement deficiencies, is due to the ineffective innate immune response to bacterial challenge (109, 245). The increased incidence of meningococcal disease in the 15-19 yr age group is thought to be behavioural, resulting in increased crowding and smoking, hence elevated levels of transmission (277).

The majority of disease is caused by meningococci expressing one of five capsule polysaccharides (or serogroups) of the thirteen that exist, namely serogroup A, B, C, W-135 and Y (discussed in greater detail in section 1.3). Current vaccines to N. meningitidis are based on the capsular polysaccharide of the meningococcus conjugated to either CRM197, tetanus toxoid or diphtheria toxoid and can be monovalent or multivalent, covering serogroups A, C, W-135 and Y (examined in (203)). The conjugated vaccines have proved to be successful in reducing the burden of disease caused by the serogroups included in the vaccine but do not prevent disease caused by serogroup B capsule expressing strains. A serogroup B vaccine based on the capsular polysaccharide is not likely as the polymer is identical in chemical structure to human glycans, specifically the glycan found on the human neural cell adhesion molecule (N-CAM) (82). Therefore it is thought that the serogroup B polysaccharide may be recognised as a “self” antigen, and may elicit an autoimmune response although there is little scientific data to support this (95, 252). Serogroup B disease outbreaks have occurred in Cuba, Norway and New Zealand among others, however in these cases a “designer” vaccine was implemented, consisting of outer-membrane vesicle (OMV) preparations derived from the outbreak strain (reviewed in (130)). This vaccination approach is strain specific and does not result in cross-protection against other meningococcal isolates which are antigenically variable (discussed further in section
1.1.2.2. Current efforts in the development of a serogroup B vaccine are focused on the use of conserved outer-membrane proteins identified using a reverse vaccinology approach championed by Dr Rino Rappuoli, and may culminate in the formulation of a broadly protective serogroup B vaccine (reviewed in (243)).

1.1.1 Typing of *N. meningitidis*

*N. meningitidis* has the potential to cause epidemics of meningococcal disease and cause endemic sporadic cases. For this reason, it is important to know whether geographically related cases of disease are co-incidental sporadic cases, or the beginning of an outbreak (316). The characterisation or typing of strains has been used to differentiate related strains within a species where biochemical and biological testing cannot (140). Strains were initially typed using a serological method which yielded some information about the relatedness of strains, however further work has since demonstrated that the usefulness of this data is limited (47, 242). Electrophoretic typing was the next molecular method to type meningococci, however although this yielded better data for comparing strains, inherent difficulties in conducting this method in multiple diagnostic laboratories restricted widespread use (91). Currently genetic typing of strains based on the nucleotide sequence of specific alleles, known as Multi-Locus Sequence Typing (MLST) has enabled the creation of a worldwide reference library (170, 292)(PubMLST.org/neisseria).

1.1.1.1 Serological typing

Meningococci were originally typed by serological methods and were defined as serogroup, serotype, serosubtype and immunotype which we now know are measures of
capsule, Porin B (PorB), Porin A (PorA) and lipooligosaccharide (LOS)-type respectively (98). This method of typing meningococci relies on the interaction of monoclonal antibodies against each of the components and the comparison to known controls to assign variants. As an increasing number of isolates were examined for the variant of class 2 and 3 antigens, it became obvious that an extended panel of reagents was required and this soon became impractical (323).

The serogrouping of meningococcal strains enables the identification of the capsule serogroup possessed by the bacterium. Over the past 80 years, this method has been used to identify one of the 13 capsules of *N. meningitidis* and as no new capsular polysaccharides have been described since 1983, this method has remained useful (14, 79, 89, 110, 247, 298). Serotyping and serosubtyping has presented somewhat more of a challenge. These methods identify the allele of PorB and PorA respectively. Since PorA and PorB are antigenically variable, serosubtyping often resulted in the organism being categorised as “non-typable” (NT). Immunotyping identified 12 different LOS structures presented on the surface of meningococcal isolates (236). A major problem with the immunotyping of strains was that many strains demonstrated cross-reactivity with typing antibodies and had to be further distinguished by mass on SDS-PAGE gels. It is now known that the LOS structures of meningococci (particularly invasive strains) are similar and in addition to this, more than 12 structures are now known to exist due to variation in the LOS inner core (17, 145, 169, 315). This method of typing has limited usefulness as many strains can have the same serological type but the genotype is not necessarily related (140). This phenomenon can be largely explained due to horizontal gene exchange, antigenic variation and phase variation within *Neisseria* sp.
1.1.1.2 Multilocus Enzyme Electrophoretic (MLEE) typing

MLEE typing examines the electrophoretic mobility of enzymes of lysed bacterial cells on starch gels (reviewed in (316)). The method examines the mass of 7-15 enzymes with each unique combination of enzyme alleles designated an electrophoretic type (ET) (45, 65, 202). As this method can potentially differentiate single amino acid substitutions in key metabolic enzymes, it is a powerful method to identify phylogenetic relationships between strains and hence differentiate clonal strains globally (47). As this technique relies on the culture of strains, preparation of lysates as samples, SDS-PAGE and subsequent staining, this method does not result in reproducible results either between or within laboratories hence reducing the utility of the method (318).

1.1.1.3 MLST

Within the species of *N. meningitidis*, sequence variation of seven housekeeping genes has been used to define clonal groups. The genes are *abcZ* (a putative ABC transporter), *adk* (adenylate kinase), *aroE* (shikimate dehydrogenase), *fumC* (fumarate hydratase), *gdh* (glucose 6-phosphate dehydrogenase), *pdhC* (pyruvate dehydrogenase subunit) and *pgm* (phosphoglucomutase). These genes are highly conserved and the products are involved in central metabolic activities. This method has enabled the identification of over 9400 sequence types to date from over 96,000 typed strains.

Based on the sequence similarity of these housekeeping genes, *N. meningitidis* can be arranged into 37 groups of closely related sequence types (STs) termed clonal complexes (cc) which account for 61% of all strains collected (46) and correlate with ETs derived from MLEE (281). Some of these clonal complexes are more often
associated with carriage than disease (>1:1), while other clonal complexes have been deemed to be hyperinvasive as these strains are more commonly associated with invasive disease than carriage (generally >5:1) (44) (Table 1.1) and is discussed in section 1.4.

Table 1.1. Characteristics of clonal complexes of N. meningitidis. Adapted from (46)

<table>
<thead>
<tr>
<th>ST-complex</th>
<th>MLEE designation</th>
<th>No. isolates</th>
<th>No. STs</th>
<th>Dominant serogroup (%)</th>
<th>Disease/Carriage ratioa</th>
<th>Main Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-1</td>
<td>Subgroup I/II</td>
<td>204</td>
<td>49</td>
<td>A(97)</td>
<td>5.5</td>
<td>Russia, China</td>
</tr>
<tr>
<td>ST-5</td>
<td>Subgroup III</td>
<td>627</td>
<td>33</td>
<td>A(99)</td>
<td>19.5</td>
<td>Africa</td>
</tr>
<tr>
<td>ST-8</td>
<td>Cluster A4</td>
<td>283</td>
<td>107</td>
<td>B(51), C(35)</td>
<td>24.5</td>
<td>Europe</td>
</tr>
<tr>
<td>ST-11</td>
<td>ET-37 Complex</td>
<td>1142</td>
<td>239</td>
<td>C(57), W135(24), B(12)</td>
<td>6.6</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-18</td>
<td>Cluster J1</td>
<td>208</td>
<td>175</td>
<td>B(85)</td>
<td>5.5</td>
<td>Czech Republic, Poland</td>
</tr>
<tr>
<td>ST-22</td>
<td></td>
<td>363</td>
<td>243</td>
<td>W135(52), NG(25)</td>
<td>0.6</td>
<td>UK</td>
</tr>
<tr>
<td>ST-23</td>
<td>Cluster A3</td>
<td>385</td>
<td>154</td>
<td>Y(62), NG3(18)</td>
<td>0.8</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-32</td>
<td>ET-5 complex</td>
<td>1028</td>
<td>350</td>
<td>B(85)</td>
<td>3.5</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-35</td>
<td></td>
<td>329</td>
<td>214</td>
<td>B(59), NG(25)</td>
<td>0.5</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-41/44</td>
<td>Lineage 3</td>
<td>1796</td>
<td>1274</td>
<td>B(70)</td>
<td>1.2</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-53</td>
<td></td>
<td>272</td>
<td>93</td>
<td>NG(76)</td>
<td>&lt;0.1</td>
<td>UK</td>
</tr>
<tr>
<td>ST-60</td>
<td></td>
<td>225</td>
<td>148</td>
<td>B(30), 29E(22), NG(19)</td>
<td>0.7</td>
<td>Europe</td>
</tr>
<tr>
<td>ST-103</td>
<td></td>
<td>127</td>
<td>84</td>
<td>B(26), NG(22), C(16)</td>
<td>1.2</td>
<td>Worldwide (-Africa)</td>
</tr>
<tr>
<td>ST-162</td>
<td></td>
<td>140</td>
<td>63</td>
<td>B(74), NG(13)</td>
<td>0.8</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-167</td>
<td></td>
<td>201</td>
<td>144</td>
<td>Y(47), NG(36)</td>
<td>0.5</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-198</td>
<td></td>
<td>166</td>
<td>76</td>
<td>NG(76)</td>
<td>&lt;0.1</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-213</td>
<td></td>
<td>187</td>
<td>165</td>
<td>B(74), NG(16)</td>
<td>0.6</td>
<td>UK</td>
</tr>
<tr>
<td>ST-254</td>
<td></td>
<td>148</td>
<td>107</td>
<td>NG(35), B(24), 29E(12)</td>
<td>0.5</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-269</td>
<td></td>
<td>415</td>
<td>312</td>
<td>B(73)</td>
<td>2.8</td>
<td>Worldwide</td>
</tr>
<tr>
<td>ST-334</td>
<td></td>
<td>106</td>
<td>64</td>
<td>C(58), B(33)</td>
<td>5.7</td>
<td>UK</td>
</tr>
</tbody>
</table>

a: This ratio was calculated only from the information present in the pubMLST.org/Neisseria database at the time of compilation (6/2/09).

b: NG=Non-groupable and are considered unencapsulated
1.1.2 Mechanisms of variation of the meningococcus

Although the possession of different alleles of genes is an indication of genetic and antigenic diversity of strains of N. meningitidis, additional mechanisms of variation can occur within the same strain which leads to antigenic diversity. Phase variation and antigenic variation are two mechanisms by which variation in surface exposed antigens can be varied in the meningococcus (reviewed in (184)).

1.1.2.1 Phase variation

Phase variation is the process of reversible mutation that occurs at high frequency within gene sequences resulting in a reading frame shift (147) (Fig 1.3). This occurs through the change in the length of simple sequence repeats (SSR) or microsatellites and specifically refers to the on/off switching of open reading frames (ORFs) although the same mechanism can also result in alterations to the binding site of trans-acting transcriptional and translational regulators which results in “volume” control, the regulation of the rate of transcription or translation (297). This type of regulation of gene expression is not unique to Neisseria and is widely distributed amongst other genera such as Haemophilus, Helicobacter and Streptococcus to vary surface structures (23, 132, 175). SSRs are nucleotide repeats, ranging in length from 1 base to 15 bases (e.g. (N)$_n$ or (NNNNN)$_n$ etc.) and have been found throughout the genome. Alterations (or mutation) to the number (n) of repeat sequences in each of these locations occurs through a mechanism not yet well characterised, although it is known to occur in a RecA- independent manner and involve the recombinase XseB, the helicase UvrD, DNA mismatch repair proteins MutS and MutL, DNA polymerase DinB, and has been shown to involve the structural proteins PilP, and the ferric binding proteins FbpA and
Figure 1.3 Mechanisms of phase variation in *Neisseria meningitidis*.

**Panel A.** Translational phase variation. i) Slipped strand mispairing occurs when an alteration to the number of nucleotides present within a homopolymeric tract resulting in a change of frame mutation which leads to a premature stop codon. ii) Microsatellites (or simple sequence repeats) are repeats of nucleotides, ranging from 2 to 15 bases. Alteration in the number of repeats results in a change of reading frame. **Panel B.** Transcriptional phase variation. The same mechanisms demonstrated in Panel A, when present in the promoter region of a gene rather than the ORF lead to alterations in transcription. This can occur by i) altering the length between the -10 and -35 regions (blue boxes) changing the binding of the RNA polymerase (RNAP), or ii) altering the binding of *trans*-acting factors (TAF).

FbpB although their impact on phase variation is less clear (11, 12, 35, 174, 192, 222, 272). The rate of phase variation is influenced by the length of the repetitive tract and also the strain, which is possibly due to the allele of the above mentioned genes.
although this has been characterised only for MutS/L to date (227). The simplest mechanism of phase variation of on/off gene expression is through slipped-strand mispairing (reviewed in (164)). This RecA independent mechanism of phase variation is thought to occur due to the errors made in the DNA polymerase during cell replication when tracts of more than 8 repeating nucleotides (i.e. (N)$_8$) are present within a gene sequence, although repetitive sequences as few as 5 nucleotides have been observed to frame-shift although the frequency at which this occurs is reduced (66). Some of the genes which encode virulence determinants have homopolymeric tracts and these include $hmbR$, $pilC$ (see section 1.3.1.3), $pglA$ (see section 1.3.1.3), $lgtA$ (see section 1.3.3.2), $synD$ (see section 1.3.4.3), $lgtG$ (see sections 1.3.3.2 and 3.4.1) and $lot$ (see sections 1.3.3.2 and 3.4.1) (Reviewed in (210)). Alterations to the number of tandem repeats of simple sequences that are not multiples of three nucleotides can also result in frame-shift mutations. Examples of variation that occurs in this manner are the pentameric repeats in $opa$ (CTCTT)$_n$ (197, 256) and heptameric repeats in $pglE$ (CAACAAA)$_n$ (212), although the mechanism for alteration to the number of tandem repeats is not well characterised.

1.1.2.2 Antigenic variation

Antigenic variation is a RecA-dependent alteration of the coding sequence of the gene by homologous recombination which results in antigenically distinct products in a clonal population, with the source of the DNA encoded by the organism (29) (Fig 1.4). Antigenic variation is thought to be a mechanism by which surface expressed structures are varied as a method of immune evasion and examples of this in Neisseria include variation of the pilin gene $pilE$ and the potential adhesin $maf$. 
Silent partial copies of both \textit{pilE} (termed \textit{pilS}) and \textit{mafB} are encoded downstream of the active loci and are unidirectionally recombined into the active locus at measurable frequency to result in variation to the product which can result in a small change to the amino acid sequence, a large change to the amino acid sequence, or the formation of a mutation which can result in the loss of protein (22, 64, 299). Alterations to the constituents of the meningococcal surface through antigenic variation, either caused by the on/off switching of genes, volume control of genes or recombination between silent and active loci can alter the manner in which the organism interacts with the host. Although not considered the classical definition of antigenic variation, the horizontal transfer of genetic material between species of \textit{Neisseria} and also strains of \textit{N. meningitidis} leads to the diversification of alleles, best studied in variation of PorA (261).

\textbf{Figure 1.4 Mechanisms of antigenic variation in \textit{Neisseria meningitidis}.} The silent copy of the gene encoded on the chromosome of the organism, in this case \textit{pilS} (green) is the source of the donor sequence for recombination with the active locus \textit{pilE} (red). Recombination between the silent locus and the active locus can occur and give rise to many potential recombinants, a process termed gene conversion. This results in genes of hybrid sequence (both green and red segments), indicative of antigenic variation. Taken from Vink et al., (299).
1.2 The interaction of *N. meningitidis* with the host

*Neisseria meningitidis* specifically inhabits humans colonising the nasopharynx, the region of pharynx bordered by the nasal cavity and the soft palate (Fig 1.5a). The nasopharynx is covered with a complex epithelium consisting simple columnar cells and simple ciliated columnar cells (19). Dispersed throughout the columnar cells are the mucous secreting goblet cells, and the antigen sampling microfold cells (M cells) and dendritic cells (Fig 1.5b) (149, 152). The examination of the interaction between meningococci with the host epithelial cells through the use of organ culture models demonstrated that the bacteria will bind to non-ciliated columnar cells but not ciliated cells (254, 255). The entry of the bacteria into the host cells, a process required for disease to progress, only occurs through columnar cells (255), (Fig. 1.5c).

As *N. meningitidis* is a human specific pathogen, the use of animal models is limited both in relevance and application. Consequently, most studies investigating the molecular interactions between the meningococcus and host have made use of immortalised cell lines (300). This body of work has resulted in an understanding of some of the important virulence determinants of the meningococcus and some of the interactions that occur with the host. The virulence determinants are discussed in detail in section 1.3, while this section provides an overview of the interaction of those factors in the context of the host, and is summarised in Fig 1.6.
Figure 1.5 Organisation of the human nasopharynx.

Figure 1.6 Interaction of *N. meningitidis* with the host.
The meningococcus first adheres to the nasopharyngeal epithelium and invades the epithelial barrier to gain entry to the bloodstream. Here, the organism must survive the immune response before crossing the blood-brain barrier to enter the cerebral spinal fluid. Taken from Rosenstein *et al.*, (226).
1.2.1 Colonisation of the host epithelium

The initial interaction between the meningococcus and host occurs through the type IV pili (tfp) of the meningococcus. These long filamentous structures extend up to 3 μm from the bacterial cell surface and are approximately 6 nm in diameter (63, 205) and mediate the initial contact with the host. The structurally and functionally similar tfp of the closely related pathogen, *N. gonorrhoeae* also mediate the initial interaction of the bacteria with the host. The pili of the gonococcus have been demonstrated to bind to I-domain containing (IDC) integrins α1 and α2 and the blocking of these molecules results in decreased rates of gonococcal attachment (86). The gonococcal tfp is post-translationally glycosylated (similar to the meningococcal tfp, see section 1.3.1) and the glycan has been demonstrated to bind the host cell complement receptor 3 (CR3) as a ligand (137, 205). Despite the knowledge of the ligands of the gonococcal tfp, the expression of neither IDC integrins nor CR3 has been examined on nasopharyngeal epithelium or their involvement in meningococcal pathogenesis.

Following the initial tethering of the organism to the host cell membrane, the pilus retracts drawing the bacterial cell to the host cell where an intimate association occurs through the binding of the bacterial opacity associated outer-membrane proteins (Opa) with host cell receptors (reviewed in (181)). The host cell molecules which have been identified in interactions with meningococcal Opas include CD66, the carcinoembryonic antigen cell adhesion molecules (CEACAMs), (specifically CEACAM1, 3, 5 and 6) and heparan sulfate proteoglycans (HSPGs) (Reviewed in (232)). Once attachment to the host is achieved, the bacteria replicate on the cell surface leading to the formation of microcolonies. The bacteria within the microcolonies disperse, a process which is partially dependent on the expression of phosphoglycerol
on the pilin (48). The dispersion of the bacteria from the microcolony is thought to be a mechanism by which transmission of the bacteria from host to host occurs, while also permitting the colonisation of more epithelial cells of the same host (48).

### 1.2.2 Invasion of host epithelial cells

In the process of disease progression, some of the meningococci which have formed an intimate association with the host mucous membrane will proceed to invade the epithelial cells and gain access to the intracellular compartment (255). The mechanism by which the meningococci enter the host epithelial cell involves the meningococcal adhesion molecules Opa and Opc (290, 301, 303) and the epithelial cell response to the meningococcus is not yet well characterised although it is known to be different to the response observed in endothelial cells (158).

Tfp mediated attachment of the meningococcus to host epithelial cells results in cortical plaque formation which, as for endothelial cells, is the accumulation of ezrin, actin, CD44, EGFR and ICAM-1 under the colony (93, 180). This process on epithelial cells, unlike endothelial cells, does not result in the interruption to the tight junction proteins in any cell line examined (158, 216, 263). Following cortical plaque formation, actin polymerisation results in the engulfment and uptake of the bacteria into the interior of the host cell. In endothelial cells actin polymerisation is a Src tyrosine kinase dependent event which follows activation by the β2-adrenergic receptor/β-arrestin pathway and the ErbB2 receptor (58, 129). However, in epithelial cells this is not the case and polymerisation occurs in an as yet undefined manner. The invasion of host urethral epithelial cells by *N. gonorrhoeae* occurs in a phosphoinositide (PI) 3-kinase dependent
pathway which leads to uptake of the bacteria by endocytosis (87), a process indicative of a clathrin-dependent, receptor mediated event and indeed the receptor asialoglycoprotein receptor (ASGP-R) was identified as the receptor mediating this process (121).

The polysaccharide capsule expressed by the meningococcus, along with the lacto-N-neotetraose (LNT) α-chain of the lipooligosaccharide (LOS) (both described in greater detail in section 1.3) can inhibit the rates of attachment to host epithelial cells (190, 305). A major survey the carrier isolates has indicated that these isolates are largely unencapsulated and express a LOS structure devoid of LNT (142). Recently, the analysis of many more strains, collected from geographically and temporally distinct regions, has indicated that while the presence of unencapsulated strains is higher among carriers than in cases of disease, the lack of a capsule is not necessary for carriage as more than half of typed carriage strains were encapsulated (Fig 1.7, (44)).

1.2.3 Interaction with the host immune system

Although the expression of the capsular polysaccharide and the LNT containing LOS have both been demonstrated to inhibit the attachment and invasion of the meningococcus to both epithelial and endothelial cells, these molecules are important for the meningococcus to resist host defences when causing disease and are consistently expressed by disease isolates. The polysaccharide capsule of the meningococcus protects the bacterium from antibody and complement mediated lysis and inhibits
The complement component C3b was seen to bind the capsular polysaccharide of serogroups A, B, C, W-135 and Y and was rapidly inactivated to C3bi. Further to this, the meningococcal capsule can also act to inhibit the binding of C4 to the phosphoethanolamine (PEA) of the LOS inner core, to prevent membrane attack complex (MAC) formation. The LNT of the LOS, particularly when sialylated (contains a terminal sialic acid residue), mimics host cell glycans and hence is not immunogenic which results in another mechanism of immune escape. The LOS and capsule of *N. meningitidis* confers resistance to cationic antimicrobial peptides (CAMPs) such as polymyxin, colistin and LL-37 a major human secreted CAMP found in epithelial mucosal surfaces.
The meningococcal LOS (also known as endotoxin) is one of the most potent activators of the Toll-like Receptor 4 (TLR4) complex (324). The LOS of the meningococcus is bound by the LPS binding protein (LBP), a soluble protein of the blood and delivered to CD14 (238, 319). CD14, which can either be a soluble protein or anchored to the plasma membrane via a glycoposphatidylinositol (GPI) anchor, enhances the interaction of the LPS with the TLR4-MD-2 complex (287). The interaction occurs through the acyl chains of the lipid A moiety of the LOS molecule with the TLR4-MD-2 complex. The chemical composition of the lipid A determines how well it is recognized by TLR4 and consequently it determines the biological activity of the LPS. N. meningitidis has been reported to produce lipid A with six acyl chains, the optimal number for TLR4 recognition (195), while natural mutants have been observed which possess only five acyl chains and this change leads to a marked decrease in activation of the TLR4 complex and hence downstream signalling (96).

1.2.4 Crossing the blood-brain barrier (BBB)

The route of the meningococcus into the cerebral spinal fluid is thought to occur via the blood-brain barrier (BBB) (199). The BBB is the anatomical structure which physically separates the central nervous system from the bloodstream, facilitating the nutrient/waste exchange between the CSF and bloodstream. The barrier is formed by the endothelium of the brain capillaries and the choroid plexus which act together to tightly regulate permeability (163, 227). Following dissemination of N. meningitidis into the bloodstream, the bacteria bind host endothelial cells. Again, the initial interaction is mediated by the tfp, which triggers cortical plaque formation and
recruitment of tight junction proteins (58, 187). Intimate attachment is mediated not only by Opa proteins, but also Opc proteins which bind α-actinin and results in the loss of cell-cell integrity (231). Invasion of endothelial cells proceeds via a paracellular route (59, 237) and the bacteria are able to gain entry to the CSF in this manner.

1.3 The structure and regulation of virulence determinants of *N. meningitidis*

In the previous section, the interaction of *N. meningitidis* with the host was examined and the role of the virulence determinants in pathogenesis described. Here, the structure and the regulation of those virulence determinants are examined.

1.3.1 Pili

The pili of *N. meningitidis* are the primary adhesins, extending from the outer membrane of the organism, protruding through the capsule and are responsible for anchoring the organism to the host cell and mediating twitching motility across the cell surface. In addition to this function in host cell attachment, the pili also confer natural competence for DNA uptake (Reviewed in (62)).

1.3.1.1 Structure and synthesis of pili

The meningococcal pilus is a multi-protein structure. The pilus fibre is almost entirely composed of PilE subunits which oligomerise to form the shaft (63). Within the shaft, the minor pilins PilV, PilX and ComP are incorporated (34, 61, 123). PilC is essential
for pilus formation in wild-type strains of *Neisseria* and is thought to cap the emerging pilus fibre, acting as the tip adhesin (228). *N. meningitidis* possess two copies of the gene encoding PilC (*pilC1* and *pilC2*) in separate locations of the genome and although these genes are variable in sequence, antigenic variation of PilC has not been described. The rest of the pilus complex consists of approximately 23 different proteins, the majority of which are responsible for the assembly and anchorage of the pilus fibre to the bacterial membranes (41) (Fig 1.8a).

The pilin (*PilE, PilX, PilV* and *ComP*) subunits have an *N*-terminal signal sequence directing their translocation to the periplasm where the signal sequence is cleaved by PilD prior to disulfide bond donation by DsbA proteins (259, 275). During this process, the pilins are post-translationally modified by the addition of glycan and zwitterionic residues of PEA, phosphocholine (PC) and phosphoglycerol (PG) before they are assembled into a fibre (Fig 1.8b). Fibre assembly (pilus formation) and disassembly (pilus retraction) is thought to be controlled in part by the ATPases PilF and PilT respectively together with the accessory proteins PilM, PilN, PilO, PilG and PilP (100, 182, 311). The pilus fibre is secreted through the pore in the outer-membrane, formed by the secretin PilQ and stabilised by PilW (42). PilG, PilH, PilI, PilJ, PilK and PilW are essential for the maintenance of the pilus fibre in the presence of PilT, and are dispensable in the absence of PilT (41, 42).

Two alleles of *pilE* are present within the species of *N. meningitidis* and these are termed class I and class II alleles (5, 78, 209). Although the alleles are different in nucleotide and protein sequence with the alleles located in different chromosomal
locations, homology exists at the N-terminus and despite sequence variation, all variants can self-assemble to form the shaft of a functional pilus.

---

**Figure 1.8 The process of pilus biogenesis in Neisseria.**

Panel A shows the current model for pilus biogenesis, indicating all the known proteins involved in pilus fibre formation. Taken from Carbonnelle *et al.*, (41).

Panel B i) shows the structure of the PilE monomer with phosphoethanolamine (PE) and the glycan (Gal-DADDGlc) shown in green and the variable loop shown in purple, and ii) shows the assembled pilus fibre, complete with post-translational modifications observed on the surface. Adapted from Craig *et al.*, (63).
1.3.1.2 Post-translational modifications of the pilus

The meningococcal pilus is post-translationally modified by the addition of a glycan to Ser63, and phasoform modification of PEA or PC at Ser68 and/or Ser156 and PG at Ser93 and/or Ser69 of each of the PilE monomers (48, 49, 122, 258, 308). The pilin glycan of *N. meningitidis* is at most a trisaccharide comprised of either glyceramido acetamido 3-deoxyhexose [GATDH] (49) or 2,4-diacetamido-2,4,6-trIDEOXYHEXOSE [DATDH] (258) with either two galactose (258) or two glucose residues (30). The pilin glycan is synthesised on an undecaprenol pyrophosphate carrier to which DATDH is attached. DATDH is formed from a UDP-GlcNAc precursor which is dehydrated on carbons 4 and 6 by PglD, then carbon 4 is aminated by PglC which is subsequently acetylated by PglB (4, 120, 211). Strains which synthesise GADTH possess an alternative pglB ORF termed *pglB2* and an additional ORF termed *orf8* which presumably act together to glycerate the amine of C4, rather than acetylate this position as is observed in DATDH (49, 146). The acetylated amine on C2 of the precursor GlcNAc remains unchanged in both DATDH and GATDH (Fig 1.9a). The glycan is then elongated by the sequential addition of either two galactose residues by phase variable PglA and PglE respectively or two glucose residues by phase variable PglG and PglH respectively (2, 30, 212). The pilin glycan can be further modified by acetylation of the terminal hexose of the glycan (either glucose or galactose) by PglI which is also phase variable thus contributing to antigenic diversity of the pilus (Fig 1.9) (309). Following synthesis of the lipid-linked glycan on the cytoplasmic face of the inner membrane, the glycan is then translocated to the periplasmic face of the inner
membrane by the flippase PglF prior to ligation to the emerging PilE monomers by the ligase PglL (Fig 1.9b) (213).

![Diagram of pilin glycan biosynthesis and process of glycosylation](image)

**Figure 1.9 Pilin glycan biosynthesis and process of glycosylation.**

**Panel A.** The base sugar for pilin glycosylation (DATDH or GATDH) is synthesised from a UDP-GlcNAc precursor by the action of PglD, PglI and PglB/PglB2. Adapted from Abu-Qarn et al., (4). **Panel B.** The base glycan can be elongated by phase variable expression of PglA and PglE, prior to translocation of the glycan to the periplasm by PglF and ligation to the emerging PilE subunits by PglL. Taken from Power et al. (213).
In the absence of PptA neither PEA nor PC is found decorating the pilin suggesting that both are added to pilin by this transferase. However no biosynthetic pathways for phosphatidylcholine have been identified in the pathogenic Neisseria and phosphatidylcholine has only been detected once in *N. meningitidis* as a minor phospholipid constituent (260). Aas *et al.* have suggested that since the major phospholipid species is phosphatidylethanolamine, PEA may be added to the pilin and converted to phosphocholine by a methylase which has yet to be identified (1). PG is added to Ser69 and Ser93 of the emerging PilE monomers by the transferase PptB (48, 257). The addition of the glycan to the meningococcal pilin is not known to serve a purpose in the rate of attachment or invasion of the bacteria to host epithelial or endothelial cells (258) although it has been postulated that the addition of this variable sugar may contribute to the variation of antigens presented to the immune system and hence contribute to immune evasion (270). In contrast to this, the addition of PG to the meningococcal pilin has been demonstrated to have a biological role, involved in the dispersal of the meningococci from microcolonies and in the rate of invasion of the bacteria across the cellular monolayer, but having no role in the rate of attachment to host cells (48). The addition of PEA to the meningococcal (or gonococcal) pilin is not known to serve a purpose in the interaction of the bacteria with the host. PEA present on the meningococcal LOS will bind the complement component C4b and interaction with factor H results in rapid inactivation of this component hence serving to prevent bacterial lysis by the membrane attack complex (219). It is not known whether PEA of the pilin could assist in immune evasion in this way. The addition of PC to the LOS of *H. influenzae* mediates the binding of the platelet activating factor receptor (PAF-r) of host epithelial cells which leads to invasion of this organism into host cells and it is
thought that the PC of the meningococcal pilin could function in much the same fashion although data is limited (1, 269).

1.3.1.3 Regulation of pilus expression and modification

The regulation of most of the proteins in the tfp complex have not been examined, with the exception of pilE and pilC. The expression of the meningococcal tfp is dependant upon the coordinated expression of each of the components that comprise the structure. The expression of pilC is phase variable with each copy possessing a poly(G) tract generally 9-15 nucleotides in length and although two copies are encoded on the meningococcal chromosome, when both copies are phased off the formation of a pilus does not occur (143). Recombination of the silent pilS into the active pilE locus can result in the formation of a product which is out of frame and hence an active PilE (and hence pilus) cannot be formed (265).

At the transcriptional level, pilC1 expression is controlled by four promoters (271, 317) three of which (P2, 3 and 4) are regulated by CrgA, MisR and PilT respectively (72, 131, 317). In response to contact with epithelial cells, the expression of pilC1 was shown to increase through the binding of MisR and CrgA to a contact regulatory element of Neisseria (CREN or REP2) (72, 131, 193). Although the transcriptional regulation of pilE has been characterised in N. gonorrhoeae (106, 107, 124) it is less well understood for N. meningitidis expressing class I PilE (71, 204) while the transcriptional regulation of class II pilE has not been characterised at all.
Post-translational modifications of the meningococcal pilin are at least in part controlled by phase variation. Although no regulatory mechanism has been described for the synthesis and/or addition of the trideoxyhexose (either DATDH or GATDH), the extension of the glycan to a trisaccharide is dependent upon the phase variable expression of $pglA$, $pglE$, $pglG$ and $pglH$. $pglA$ has a poly(G) tract, $pglE$ has a heptameric repeat, $pglG$ and $pglH$ both have a poly(G) tract (212). The addition of PEA and PC by PptA is regulated by phase variation (138), while the addition of PG by PptB is regulated in a host cell contact dependent manner but is not phase variable (193).

1.3.1.4 Distribution of pilus biosynthesis and modification genes

Although the majority of the tfp machinery is conserved between all *Neisseria* spp., the distribution of $pilE$, $pilC$ and $pglB$ alleles differ between strains of *N. meningitidis*. Class I pilins are found in *N. gonorrhoeae* and most *N. meningitidis*, while class II pilins are found in commensal species and a minor population of *N. meningitidis* (10). The class I $pilE$ is found in the same genetic locus in all strains and is located upstream of approximately eight $pilS$ loci, while in contrast, the class II $pilE$ is also found in a conserved genetic locus, but is separate to the class I locus and is not located in proximity to other genes involved in pilin synthesis or modification. The strains which possess a class II $pilE$ also possess a truncated $pilS$ locus. While the class I $pilE$ has an antigenically variable loop (considered to aid in immune evasion), the class II $pilE$ lacks this region entirely. The class II $pilE$ is mainly found in the clonal complex 8 and 11 strains of *N. meningitidis* (9, 146), however as more genome sequences are available the class II $pilE$ locus has also been observed in one ST-1 strain, one ST-5 strain and one ST-4122 strain (Patric.org).
The proposed tip adhesin of PilC is thought to contribute to tissue tropism (151, 200). Nassif et al. have shown in strain 8013 that PilC1 confers adhesive phenotype while PilC2 will still facilitate pilus biogenesis however, does not mediate an adhesive phenotype toward the human umbilical vein endothelial cell line (HUVEC) (200). Both PilC alleles were later demonstrated to facilitate adherence to the epithelial cell line ME-180 of strains 8013 and FAM20. PilC conferred a different rate of attachment for each strain possibly as a consequence of variable protein sequences of pilC alleles (191).

The distribution of pilin glycosylation genes across different strains of N. meningitidis was examined by Power et al. (212). While the glycosyltransferases pglA and pglE are found across all genomes to date, the glycosyltransferases pglG and pglH are found in 67% of all strains examined, but in 90% of strains carrying the pglB2 allele while approximately half of strains carrying pglB also carry the pglGH locus (212). pglG and pglH are found in an operon downstream of pglB and form a genetic island which is deleted from some strains (Kahler et al., 2001). Approximately half of all strains of N. meningitidis carry the pglB2 allele, with all ST-8 cc, ST-11 cc and ST-41/44 cc strains carrying pglB2 and all ST-32 cc and ST-269 cc strains carrying pglB.

1.3.2 Outer-membrane adhesins

Neisseria meningitidis has many outer-membrane proteins, some of which are involved in the intimate association of the meningococcus with host cells. These proteins, although functionally important, also contribute to the variability between strains of
meningococci. The outer membrane adhesins include Opa, Opc, NadA and Nhh. Porins are a major class of outer-membrane protein, and while not specifically an adhesin, are discussed here as these proteins contribute to pathogenicity.

1.3.2.1 Opa

*N. meningitidis* possesses up to four phase variable (see section 1.1.2.1) copies of the *opa* gene located at distinct positions on the genome, while *N. gonorrhoeae* possesses eleven copies. The opacity proteins (Opa) of *Neisseria* were so named as their presence conferred an opaque colony phenotype when viewed on blood agar plates (264). Structurally, these proteins are predicted to form 8-strand beta-barrels embedded in the outer membrane with four surface exposed loops (67, 171). The *opa* loci of *N. meningitidis* have been termed *opaA*, *opaB*, *opaD* and *opaJ*. These designations do not reflect the sequence of the gene as substantial antigenic variation occurs within the surface exposed loops of *opa*, resulting in more than 330 alleles with the majority of the variation occurring in the surface exposed loops (57) http://neisseria.org/nm/typing/opa/. Three of the four surface exposed loops of Opa proteins are variable with one termed semi-variable (as variation only occurs in roughly half the sequence); while the other two loops are termed hyper-variable regions 1 and 2 respectively with the entire exposed loop varying between sequences.

The Opa proteins are thought to be the major adhesin of the meningococcus, with strains lacking Opa failing to adhere to host cells compared to Opa expressing variants (162, 301). Opa proteins have been demonstrated to interact not only with host molecules (described in section 1.2), but also with self-molecules (particularly
carbohydrates) including the LNT of LOS and serogroup A and B capsule polymers (190). The surface exposed loops of Opa have been identified as the substrate binding region of the protein as determined by mutation mapping, with the alteration of a single amino acid in loops two or three sufficient to obliterate binding of the protein to either CEACAM1 or CEACAM5 (68). As the surface exposed loops are also the hyper-variable regions of the proteins, this has prevented the design of a bioinformatics approach to determine the substrate of any given Opa allele.

1.3.2.1.1 Regulation of Opa.

The expression of Opa proteins is regulated at both the transcriptional and translational level. The most obvious form of regulation of the expression of Opa proteins occurs via the RecA-independent phase variation of the pentameric tract of CTCTT located in the N-terminal sequence of the protein, which regulates the expression of Opa at the level of translation (197, 256). An examination of the rate of phase variation of the Opa proteins of *N. gonorrhoeae* found that the level of transcription was proportional to the rate of phase variation, although a mechanism for this phenomenon has not been described (21). In *N. gonorrhoeae* all of the *opa* genes are regulated at the transcriptional level by the repressor Fur, however no *opa* gene is regulated by Fur in *N. meningitidis* (320). The transcription of at least one *opa* gene of *N. meningitidis* was shown to increase 1.3-fold following exposure to whole human blood (85) and the expression of at least one *opa* gene decreased by approximately 2.5-fold in the absence of the two-component system regulator MisR (285).
1.3.2.1.2 Distribution of \textit{opa}

A significant amount of diversity of \textit{opa} genes is observed within \textit{N. meningitidis} due to antigenic variation occurring both within and between strains, giving rise to more than 330 alleles of \textit{opa} at last count in 2006 (128), (http://neisseria.org/nm/typing/opa/). Despite this, specific alleles of \textit{opa} are more often associated with specific genetic lineages which may imply a role of specific Opa alleles in each of these genotypes (39).

1.3.2.2 Opc

Opc was originally termed a class 5 protein, along with the Opa proteins but later genetic and biochemical evidence demonstrated that although it was related to Opas, it was a distinct structure (6) consisting of a 10-sheet beta barrel with 5 surface exposed loops (179, 215). Opc is encoded by the gene \textit{opcA} and is a meningococcal specific gene, absent from \textit{N. gonorrhoeae} and all commensal strains. As such, it is a true meningococcal virulence determinant and is a factor involved in the invasion of host endothelial cells, leading to entry across the BBB (described in section 1.2) (301, 302). Opc has a similar binding repertoire to Opa, binding fibronectin and vitronectin which in turn mediate binding to integrins and HSPGs (70, 289, 304, 305). Unlike Opa, Opc mediates binding to $\alpha$-actinin which is a component of tight junctions and may be a mechanism by which meningococci can cross the BBB (231).

1.3.2.2.1 Regulation of Opc

The expression of Opc in \textit{N. meningitidis} is regulated at the transcriptional level by the phase variation of the poly(C) tract (discussed in section 1.1.2.1) located between the -10 and -35 sites of the promoter, alterations of which impact on the binding of the
sigma factor to the promoter and hence the level of transcription (233). No other mechanisms for the regulation of Opc have been described although the transcription of \( opcA \) increased 3-fold following exposure to human whole blood (85).

1.3.2.2.2 Distribution of \( opcA \)

The \( opcA \) gene is found in a large subset of \( N. meningitidis \), however it is absent from ST-8 cc and ST-11 cc strains (241). The coding sequence of \( opcA \) is highly conserved with very few polymorphisms at the nucleotide level, although more diversity exists in the flanking regions with 18 alleles identified (241). No function has been ascribed to the different alleles of \( opcA \) and despite the role of Opc in invasion of endothelial cells \textit{in vitro}, strains of ST-8 cc and ST-11 cc frequently cause meningococcal meningitis.

1.3.2.3 NadA

NadA is a trimeric autotransporter protein of \( Neisseria meningitidis \) belonging to the group of oligomeric coiled-coil adhesins. This adhesin was found to be expressed in \(~50\%\) of \( N. meningitidis \) strains isolated from patients but in only \(~5\%\) of strains from healthy individuals, and therefore it may be a risk factor for the development of meningococcal disease (56). The expression of NadA increases the rate of attachment with host epithelial cells, monocytes, macrophages, and monocyte-derived dendritic cells (40, 97, 177, 273) and this is thought to be through the interaction with the \( \beta1 \) component of integrins (198).
1.3.2.3.1 Regulation of NadA

The expression of NadA is highly regulated through many mechanisms. A complex interplay exists between classical mechanisms of transcriptional regulation and phase variation. The classical regulatory proteins integration host factor (IHF), Fur and the MarR-type regulator NadR (also known as FarR) bind sites located upstream and overlapping the RNA polymerase binding sites. The binding of these proteins and their subsequent activity is determined by the number of simple sequence repeats (in this case TAAA, discussed in section 1.1.2.1) present in this locus (173).

IHF is a heterodimer of IHFA and IHFB, two histone-like proteins similar to that found in other bacteria such as *E. coli*, which binds and bends the DNA 120° such that the binding of *cis*-* and trans*-acting factors such as RNA polymerase and regulatory proteins amongst others, is altered to result in either the activation or repression of transcription (reviewed in (101)). The expression of IHF in *N. meningitidis* is not currently characterised however IHF is involved in the regulation of genes including *farAB* and *pilE* of the related pathogen *N. gonorrhoeae* (125, 159).

Fur is a regulatory protein which is activated by the binding of iron (or other divalent cation) to result in dimerization and binding of recognition sequences (also known as Fur-box) to result in the differential regulation of transcription under different conditions (16, 69, 92). Fur is a global regulator of *N. meningitidis*, responsible for the differential expression of 233 genes (by direct and indirect mechanisms) under iron deplete versus iron replete conditions (113).
The transcription of \( \textit{nadA} \) is regulated by NadR, a MarR-type regulator with a typical DNA binding winged helix-turn-helix motif, which dimerises and binds inverted repeats often of a 9-1-9 organisation although the length of the repeat and of the gap can differ (235). Three NadR operator sequences are present upstream of the translation start point, with one overlapping the -10 region of the promoter and hence occludes the binding of the RNA polymerase. The number of TAAA repeats present is thought to alter the binding of IHF and hence the bending of the DNA such that NadR binding can be influenced between low and high level repression (183). The NadR repression of NadA may be de-repressed when in human saliva through the interaction of 4-hydroxyphenylacetic acid (4HPA) with NadR to prevent binding, which also alters the expression of the entire NadR regulon (94, 183).

Since the expression of \( \textit{nadA} \) was not altered in response to exposure to human blood (85), it appears these mechanisms of regulatory controls may be more pertinent during colonisation of mucosal surfaces or interaction with immune cells.

1.3.2.3.2 Distribution of \( \textit{nadA} \)

\( \textit{nadA} \) is present in approximately 50% of meningococcal isolates from disease. It is over-represented in hypervirulent lineages with 100% of ST-8 cc, ST-11 cc and ST-32c strains possessing \( \textit{nadA} \) while it is absent from ST-41/44 cc strains (55, 56). The carriage of \( \textit{nadA} \) is under-represented in carriage isolates with only 16% of strains possessing the gene (56) and it is absent from strains of \textit{N. gonorrhoeae}, \textit{N. cinerea} and \textit{N. lactamica} (55). In strains of the ET-15, a subset of the ST-11 clonal complex, an IS1301 insertion interrupts the \( \textit{nadA} \) ORF frequently (90). Although four alleles of
NadA exist and one allele lacks the leucine zipper, no functional differences have been ascribed as yet (55, 56).

1.3.2.4 NhhA

Neisseria Hia/Hsf homologue (NhhA) (similar to Haemophilus influenzae Hia/Hsf) is a trimeric autotransporter adhesin of *N. meningitidis*, located in the outer-membrane (208). NhhA mediates the attachment of the meningococcus to HSPGs and laminin of human epithelial cells (234), and binds activated human vitronectin which leads to an increase in serum resistance (115). In a murine model of meningococcal disease, NhhA was found to be essential for bacterial colonization of the nasopharyngeal mucosa, and it has also been shown to protect meningococci from phagocytosis and complement-mediated killing (246).

1.3.2.4.1 Regulation of nhhA

Different levels of NhhA have been observed, ranging from absent to highly expressed, however a mechanism by which differential expression occurs has not been examined. It has been suggested that the expression of NhhA increases in response to meningococcal attachment to host epithelial cells, however this data has not been published (246). NhhA was not differentially regulated in response to exposure to human blood (85) and no regulators of *nhhA* are known.

1.3.2.4.2 Distribution of *nhhA*

Although all examined strains of *N. meningitidis* possess the gene sequence of *nhhA*, the level of protein expression was variable between strains (84, 208). The sequence of
nhhA has been identified in *N. lactamica* and *N. polysaccharea*, however it is absent in genome sequences of *N. gonorrhoeae*. An examination of nhhA nucleotide sequences has revealed that there are four conserved regions of the gene and four variable regions (208), although different alleles were demonstrated to retain the function of binding activated vitronectin (115). A naturally occurring single point mutation results in the inability of the protein to trimerise in a subset of ST-41/44 cc strains and hence NhhA is found as a monomer in these strains (84).

### 1.3.2.5 Porins

Although porins have not been demonstrated to be involved in the attachment and invasion of *N. meningitidis* into host cells, these proteins contribute to the antigenic variation observed within the species (hence forming the original basis of typing (see section 1.1.1.1) and are a major and integral component of the bacterial cell (98, 280).

Meningococcal porins are proteins containing 16 β-pleated sheets, resulting in 8 surface exposed loops, which auto-aggregate to form a trimer in the outer membrane resulting in a voltage-gated channel through which small molecules may pass (74, 188). Two copies of porins are present within most meningococcal isolates however, each is distinct in sequence. Typically PorA is a ~45kDa protein encoded between a transcription elongation factor and aminopeptidase, while one of the two alleles of PorB (termed PorB2 and PorB3) is present in the locus bordered by the PEMK-like protein and the thiamine biosynthesis protein *thiC*. PorB2 is typically ~40kDa, while PorB3 is typically ~38kDa although hybrids have been observed (3, 18, 126, 196, 280).
PorA contains two variable regions (VR1 and VR2) and two semi-variable regions (SV1 And SV2) which correspond to the surface exposed loops 1 and 4, and loops 5 and 6 respectively (178). Although PorA is not considered to be antigenically variable in the classical sense, (i.e. no alternative locus for within-strain recombination to occur exists), a substantial amount of antigenic variation occurs by horizontal transfer and recombination between strains (see section 1.1.1.1) (261, 262).

The Neisserial porins have the capability to translocate from the bacterial cell membrane to the eukaryotic plasma membrane (168, 176, 310). The translocation of the gonococcal porin to the host epithelial cell membrane results in an alteration of the calcium flux and is thought to alter membrane potential, interfering with cell signalling pathways (15, 229, 288). The Neisserial porins can also translocate into the mitochondrial membrane which then binds ATP, leads to aberrations in signalling pathways and ultimately is thought to contribute to cell death (155).

1.3.2.5.1 Regulation of Porins

The expression of porA is regulated by phase variation, both transcriptionally by virtue of a poly(G) tract which is present between the -10 and -35 regions of the promoter (13), and translationally by virtue of a poly(A) tract in the coding sequence (293). In addition, the expression of porA was shown be up-regulated 6.7-fold in the presence of the phase-variable methyl transferase ModD found in ST-41/44 cc strains which is largely absent from other clonal complexes (240). The transcription of porA and porB was shown to decrease by 2-fold and 2.9-fold respectively in the absence of the Lrp/AsnC family regulator NMB0573, a situation thought to mimic nutrient deprivation (220).
1.3.3 Lipooligosaccharide

The meningococcal lipooligosaccharide (LOS) is a glycolipid constituting the major component of the outer leaflet of the outer membrane. The meningococcal LOS is termed an oligosaccharide as it lacks the repeating O-antigen observed in polysaccharides and consequently is a comparatively small endotoxin of approximately 5 kDa (114). Like other endotoxins, this molecule is tethered to the outer-membrane by the lipid A. The LOS/LPS of bacteria is an endotoxin as it is a heat-stable potent activator of TLR4 signalling pathways leading to a toxic effect as discussed in section 1.2 (324).

The meningococcal LOS was originally defined by reactivity with a panel of monoclonal antibodies which resulted in the identification of different LOS structures which were subsequently termed immunotypes (see section 1.1.1.1) (236). Further chemical analysis has revealed a greater diversity of these structures (Table 1.2).

1.3.3.1 Synthesis and transport of LOS

The meningococcal lipid A is synthesised on the cytoplasmic face of the inner-membrane in a similar manner to what is seen in E. coli (218). The UDP-N-acetylglucosamine precursor which is acylated with R-3-hydroxydodecanoyl by LpxA forming UDP-3-O-(R-3-hydroxydodecanoyl)-GlcNAc (217). This product is subsequently deacetylated by LpxC and acylated with N-linked R-3-hydroxytetradecanoyl by LpxD resulting in UDP-2-N-(R-3-hydroxytetradecanoyl)-3-O-(R-3-hydroxydodecanoyl)-GlcNAc which is dephosphorylated by LpxH, releasing UMP.
and resulting in Lipid X. Lipid X and its precursor are condensed by LpxB resulting in the release of UDP and the formation of 2-N-(R-3-hydroxytetradecanoyl)-3-O-(R-3-hydroxydodecanoyl)-GlcNAc β1’→6 2-N-(R-3-hydroxytetradecanoyl)-3-O-(R-3-hydroxydodecanoyl)-GlcNAc-1-PO₃ (156). The GlcNAc’ headgroup is then phosphorylated by LpxK using ATP, thus forming lipid IVₐ. The sequential addition of two KDO residues to C5’ of lipid IVₐ occurs next by the bi-functional enzyme KdtA (WaaA), which first makes a β1→5’addition to the lipid A headgroup, followed by β2→4 addition to the KDO of the KDOI-lipid IVₐ substrate (218). Following the addition of both KDO residues, the lipid IVₐ is further acylated first by the transfer of R-3-hydroxydodecanoyl to position 2 by LpxL2, and

<table>
<thead>
<tr>
<th>Immunotype</th>
<th>α-chain⁴</th>
<th>3-Hept</th>
<th>6-Hept</th>
<th>GlcNAc</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Glc-Gal-Gal +/-%NeuNAc</td>
<td>PEA</td>
<td>H</td>
<td>OAc</td>
<td>(77)</td>
</tr>
<tr>
<td>L2</td>
<td>Glc-Gal-GlcNAc-Gal +/-%NeuNAc</td>
<td>Glc</td>
<td>PEA</td>
<td>OAc</td>
<td>(108)</td>
</tr>
<tr>
<td>L3</td>
<td>Glc-Gal-GlcNAc-Gal-NeuNAc</td>
<td>PEA</td>
<td>H</td>
<td>H</td>
<td>(207)</td>
</tr>
<tr>
<td>L4</td>
<td>Glc-Gal-GlcNAc-Gal +/-%NeuNAc</td>
<td>H</td>
<td>PEA</td>
<td>OAc</td>
<td>(153)</td>
</tr>
<tr>
<td>L5</td>
<td>Glc-Glc-Gal-GlcNAc-Gal+/-%NeuNAc</td>
<td>Glc</td>
<td>H</td>
<td>OAc</td>
<td>(185)</td>
</tr>
<tr>
<td>L6</td>
<td>Glc-Gal-GlcNAc</td>
<td>H</td>
<td>PEA</td>
<td>OAc</td>
<td>(77)</td>
</tr>
<tr>
<td>L7</td>
<td>Glc-Gal-GlcNAc-Gal</td>
<td>PEA</td>
<td>H</td>
<td>H</td>
<td>(153)</td>
</tr>
<tr>
<td>L8</td>
<td>Glc-Gal</td>
<td>PEA</td>
<td>H</td>
<td>H</td>
<td>(214)</td>
</tr>
<tr>
<td>L9</td>
<td>Glc-Gal-GlcNAc-Gal</td>
<td>H</td>
<td>PEA</td>
<td>OAc</td>
<td>(134)</td>
</tr>
<tr>
<td>L10</td>
<td>Glc-Glc-Gal +/-%NeuNAc</td>
<td>PEA</td>
<td>PEA</td>
<td>OAc</td>
<td>(150)</td>
</tr>
<tr>
<td>L11</td>
<td>Glc-Gal-GlcNAc-Gal</td>
<td>H</td>
<td>PEA</td>
<td>OAc</td>
<td>(189)</td>
</tr>
</tbody>
</table>

⁴: see Appendix A(V) for abbreviations of sugars

Then R-3-hydroxydodecanoyl to position 2’ by LpxL1, resulting in KDO₂-lipid A (296). This hexa-acylated lipid A is the most common form observed in N. meningitidis however, naturally occurring mutants of lpxL1 are unable to add the second secondary acyl chain and hence this results in the formation of a penta-acylated lipid A structure which is associated with a reduced severity of disease (96).
The core residues of the LOS molecule are assembled next onto the KDO₂-lipid A precursor which is located on the cytoplasmic face of the inner-membrane. The first additions are two sequential Hep residues, first HepI is added to KDOI via an α₁→5 linkage by RfaC, then HepII is added to HepI by RfaF in an α₁→7 linkage (239, 321). The emerging LOS molecule is further extended by the addition of α₁→2 GlcNAc to HepII by RfaK, followed by β₁→4 Glc to HepI by LgtF (144) which results in the formation of the core LOS structure observed in all strains of *N. meningitidis*. Further variable glycan modifications are made to the LOS inner core including the addition of α₁→3 Glc to HepII by the LOS glycosyltransferase, LgtG when present. The α-chain extension of the LNT is formed by the sequential addition of β₁→4 Gal by the LOS glycosyltransferases LgtE, β₁→3 GlcNAc by LgtA and β₁→4 Gal by LgtB (306). Alternative α-chain additions can be formed including the α-chain of L1 and L5 immunotype LOS structures (Table 1.2 however, these structures are rarely observed in disease causing isolates (142). Lastly, the GlcNAc of the core structure can be further modified by the *O*-acetylation of C3 by the LOS O-acetyl transferase, Lot (145) (Fig 1.10a).

It is at this point in LOS biogenesis that the partial molecule is thought to be translocated across the inner-membrane by the ATP-binding cassette transporter MsbA (81, 322) to the periplasmic surface of the inner-membrane (Fig 1.10b). While at this location, PEA additions are made to the emerging LOS molecules. The PEA additions
to the lipid A headgroups and to two positions of the HepII are each mediated by separate enzymes, each using phosphatidylethanolamine as the substrate. The LOS PEA transferases are membrane bound enzymes which add PEA primarily to the C4’ of the headgroup (LptA), to the C3 position of HepII (Lpt3) and to the C6 position of HepII (Lpt6) (60, 169, 315). As PEA addition occurs after glycan addition, if the C3 position of HepII is occupied with Glc then PEA cannot be transferred to this position. Some studies have indicated that the O-acetylation of the core GlcNAc can also inhibit the
addition of PEA to C3, while others have demonstrated the occurrence of C3 PEA in the presence of O-acetylated GlcNAc in different strains (76, 145).

Following PEA modification, the LOS molecule is transported across the periplasm to the outer-membrane. The LOS is thought to be extracted from the inner-membrane by the ABC-transporter LptB/C/F/G complex (32, 230, 248), which then allows for the LPS transport protein H (a homologue to the E. coli LptA) to bind lipid A (276) and transport the molecule across the periplasm. Once the LOS molecule reaches the outer-membrane, the LptD/E complex (formerly known as Imp/RlpB) acts to translocate the molecule to the outer leaflet of the outer-membrane (81). As all of the Lpt proteins are found in a transenvelope complex, it is though that LOS transport most likely occurs via a protein bridge at a membrane contact site (31, 50).

Following translocation of the LOS molecule to the outer-membrane, the terminal Gal of the LOS molecule can be modified by the addition of α2,3 sialic acid (also known as N-acetyl neuraminic acid, (NANA) or NeuNAc), a process known as sialylation (172, 244). The addition of sialic acid by the LOS sialyl transferase (Lst) occurs on the outer-membrane of the organism, using CMP-NANA as the substrate which can come either from an endogenous source (in the case of serogroup B, C, W-135 and Y strains which synthesis sialic acid as part of the capsule, see section 1.3.4.2) or exogenous from the host as a component of host serum (244). The length of the α-chain of the LOS molecule affects whether or not the structure can be sialylated with L8 (Galβ1-4Glcβ1-4) being devoid of sialic acid, while L1 (Galα1-4Galβ1-4Glcβ1-4) and L10 (Galβ1-4Glcβ1-4Glcβ1-4R) are sialylated with α2,6 sialic acid or α2,3 sialic acid respectively (278). The sialylation of the L8 α-chain is thought to be inhibited by the proximity to
HepII and its decorations. The addition of α2,6 sialic acid to the α-chain of strain 126E is catalysed by an allele of Lst in which a change in the amino acid 168 from glycine to isoleucine has been demonstrated to facilitate bi-functionality and utilisation of a broader range of LOS α-chain acceptor molecules (278, 307).

1.3.3.2 Regulation of LOS

The biogenesis of the LOS molecule in *N. meningitidis* occurs via a process which is linked to the transport of the molecule to the outer-leaflet of the organism. Unlike in *E. coli*, the biosynthesis of the LOS molecule is reduced when a mutation is made which inhibits the correct localisation of the molecule to the outer-leaflet and results in severely reduced amounts of LOS (32, 274). Although a negative feedback system for the biosynthesis of LOS has been empirically observed in *N. meningitidis*, the mechanism by which this occurs remains elusive.

Phase variation is a mechanism used to vary the structure of LOS. The enzyme encoded by *lgtA* which is responsible for the elongation of the α-chain is phase variable in some strains (poly(G)5-15) while the gene encoding inner core glucosyl transferase (*lgtG*) is also phase variable (poly(G)9-13) (24) and *lot* has a poly(G)5-6 tract. The two component system regulator MisR/S (PhoP/Q) was shown to directly affect the transcription of *lgtG*, which is repressed under normal laboratory conditions (282). In the absence of MisR the transcription of *lgtG* increased as did the frequency of the inner core decoration with Glc. However, this structure was also devoid of PEA at C6 of the inner core, even though there was no significant change in the level of transcription of *lpt6* (282). The mechanism by which this alteration to LOS structure occurs is not known.
Since MisR is activated in response to contact with host epithelial cells (131), it is feasible to consider that the LOS structure may be altered in response to host cell contact although this has not been demonstrated yet.

**1.3.3.3 Distribution of LOS biosynthesis genes**

The distribution of LOS biosynthesis genes across different strains of *N. meningitidis* was examined by Berrington *et al.* (24). While the glycosyltransferase *lgtA* was present in all strains examined, the poly(G) tract differed in length from a consistent 5 nucleotides in ST-8 cc and ST-11 cc strains (cluster A4 and ET-37, respectively), to variable 5-11 nucleotides in ST-41/44 cc strains (Lineage 3), to 11-17 nucleotides in ST-32 cc strains (ET-5). As the length of the tract influences the rate at which phase variation occurs, it is has been proposed that the expression of the LNT in ST-32 cc strains is highly variable, while it is constant in ST-8 cc and ST-11 cc strains. The glucosyl transferase *lgtG* was seen to be present in ST-8 cc, ST-11 cc, ST-32 cc strains but absent from ST-41/44 cc strains (24). In all cases *lgtG* had a poly(G) tract of 9-13 nucleotides suggesting phase variation could occur at similar rates in the absence of other contributing factors (such as MutS/L (221)).

**1.3.4 Capsule**

*N. meningitidis*, unlike any other species of *Neisseria*, can produce polysaccharide capsules (52). There are currently 13 known capsule types identified by immunotyping and are termed serogroups, with the chemical structures of 12 being determined (Table 1.3) with the exception of serogroup D capsule (25, 26, 38, 135, 136, 166, 167, 186, 294, 295). Five capsule structures are commonly observed in cases of disease (A, B, C,
W-135 and Y) (226). However, the occurrence of disease caused by serogroup X strains is increasing since 1995 (80), being the causative agent of three large-scale epidemics in Niger, Togo and Burkina Faso with high morbidity and mortality (28, 73). The distribution of the meningococcal capsule type differs with respect to geographical location with serogroup A prevailing in Africa and Asia, and serogroup B and C prevailing in the western world (Fig1.11) (118). Biologically the capsule serves to protect the bacterium from the host, inhibiting the deposition of complement components and antibody recognition (112, 139, 194). The capsule helps promote survival of the organism within host cells although the mechanism is not known but is thought to increase resistance to oxidative stress induced by acidification of the vacuole (249).

Figure 1.11 Global distribution of serogroups in disease. N. meningitidis expressing serogroup B and C capsules are the major cause of disease in North America, South America, Europe and Australia. Disease caused by serogroup Y strains is observed in North America and the western regions of South America. While disease caused by serogroup A is greatest in Africa and Asia. Taken from Harrison et al., (118).
Table 1.3 Summary of *N. meningitidis* capsule information. Adapted from (104, 119)

<table>
<thead>
<tr>
<th>Serogroup</th>
<th>Structure*</th>
<th>Genes†</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>→6)-α-D-ManpNAc-(1→OPO₃→</td>
<td>sacA-sacD (mynA-mynD)</td>
<td>(38, 267)</td>
</tr>
<tr>
<td>B</td>
<td>→8)-α-D-Neup5Ac-(2→</td>
<td>siaA-siaD₂ (synA-synD)</td>
<td>(27, 105)</td>
</tr>
<tr>
<td>C</td>
<td>→9)-α-D-Neup5Ac-(2→</td>
<td>siaA-siaD₂ (synA-synD)</td>
<td>(27, 268)</td>
</tr>
<tr>
<td>D</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>29E</td>
<td>→3)-α-D-GalpNac-(1→7)-β-D-KDOp-(2→</td>
<td>cap29EA-H (AJ576117)</td>
<td>(25)</td>
</tr>
<tr>
<td>H</td>
<td>→4)-α-D-Galp-(1→2)-Gro-(3→OPO₃→</td>
<td>cpsZA-cpsHD</td>
<td>(295)</td>
</tr>
<tr>
<td>I</td>
<td>→4)-α-L-GulpNacA-(1→3)-β-D-ManpNacA(→</td>
<td>cpslKe-cpslKe</td>
<td>(186)</td>
</tr>
<tr>
<td>K</td>
<td>→3)-β-D-ManpNacA-(1→4)-β-D-ManpNacA(1→</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>→3)-β-D-GlcpNac-(1→3)-β-D-GlcpNac-(1→3)-α-D-GlcpNac-(1→OPO₃→</td>
<td>lcbA-lcbC (AF112478)</td>
<td>(135)</td>
</tr>
<tr>
<td>W-135</td>
<td>→4)-α-D-Neup5Ac-α-(2→6)-α-D-Gal-(1→</td>
<td>siaA-siaD₁₃₅ (synA-synD)</td>
<td>(26, 54, 161)</td>
</tr>
<tr>
<td>X</td>
<td>→4)-α-D-GlcpNac-(1→OPO₃→</td>
<td>xcbA-xcbC</td>
<td>(36-38, 286)</td>
</tr>
<tr>
<td>Y</td>
<td>→4)-α-D-Neup5Ac-α-(2→6)-α-D-Glc-(1→</td>
<td>siaA-siaD₂ (synA-synD)</td>
<td>(26, 54)</td>
</tr>
<tr>
<td>Z</td>
<td>→3)-α-D-GalpNacA-(1→1)-Gro-(3→OPO₃→</td>
<td>cpsZA-cpsZd (AJ744766)</td>
<td>(136)</td>
</tr>
</tbody>
</table>

*‡: Refer to Appendix A(V) for key to abbreviated sugars

†: Refer to text for gene function.

### 1.3.4.1 Synthesis and transport of capsule

The meningococcal capsule is a group II capsule presented on the outer-membrane, consisting of a diacylglycerol lipid anchor to which the carbohydrate polymer of 20-300 sugars is linked via a phosphodiester bridge (in at least serogroup B capsules) (111).

The chemical composition of the polysaccharide portion of twelve of the meningococcal capsules is currently known and is listed in Table 1.3. The genetic locus encoding the machinery for biogenesis of the meningococcal capsule is divided into 5 regions (A-E) (Fig 1.12), with only the synthesis locus (Region A) differing between serogroups (225). Regions B, C, D and E encodes the ABC transporter machinery and other factors required for capsule biogenesis, and is proposed to be conserved for all meningococcal capsule types (104).

The biochemical pathway for the synthesis of N-acetyl mannosamine (ManNAc) present in the serogroup A capsule is largely unknown. The first gene in the synthesis operon is *mynA* (*sacA*) which has homology to a GlcNAc 2-epimerase which is
Figure 1.12 Genetic organisation of the capsule locus of *N. meningitidis*.
Regions B, C, D, D’ and E are conserved between all serogroups (except the capsule null locus (cni)), with only the biosynthetic locus, region A, differing. Taken from Harrison et al., (119). Refer to text and Table 1.3 for gene names and function.
hypothesised to be responsible for the conversion of UDP-GlcNAc to UDP-ManNAc (267). The second gene in the operon is *mynB* (*sacB*) which has homology to proteins from Mycobacteria, is thought to be involved in the polymerisation of the ManNAc residues via a phosphodiester bond. The serogroup A capsule is *O*-acetylated, a requirement for capsule expression and this addition is thought to be mediated by the third ORF in the operon *mynC* (*sacC*) (116, 267). The final ORF *mynD* (*sacD*) has predicted trans-membrane domains, but little homology to other characterised proteins. The inactivation of this gene resulted in the reduction of capsule and is therefore thought to be involved in the transport of the capsule to the cell (267).

Sialic acid which is present in serogroup B, C, W-135 and Y capsules is synthesised from a GlcNAc-6-phosphate precursor which is epimerised to ManNAc-6-phosphate by the GlcNAc-6-P 2-epimerase SynA (*SiaA*). The phosphate is removed, presumably by a phosphatase which has not been identified, prior to condensation of ManNAc with phosphoenolpyruvate by SynC (*SiaC*) resulting in the formation of Neu5NAc. SynB (*SiaB*) then acts to activate Neu5NAc to CMP-Neu5NAc which is the substrate for the polysialyltransferase SynD (*SiaD*), which sequentially adds sugar monomers to the non-reducing end of the forming polymer (54, 88, 102, 105). The SynD for each serogroup is specific, mediating the different linkages observed (e.g. $\alpha2\rightarrow8$ for SynD$_B$ and $\alpha2\rightarrow9$ for SynD$_C$), while also mediating the disaccharide $\rightarrow4)-\alpha-D$-Neup5Ac-$\alpha-(2\rightarrow6)-\alpha-D$-Gal-(1$\rightarrow$ of serogroup W-135 and $\rightarrow4)-\alpha-D$-Neup5Ac-$\alpha-(2\rightarrow6)-\alpha-D$-Glc-(1$\rightarrow$ of serogroup Y (53, 54). Serogroups W-135, Y and H have been seen to carry an additional copy of *galU* which encodes the enzymes for the conversion of Glc-1-phosphate to UDP-Glc (83). The duplication of this gene, which is different in sequence to the chromosomally located copy, is likely to be carried with the capsule locus as a
means of satisfying the metabolic requirement for UDP-Glc or UDP-Gal for capsule expression (119). The polysaccharide polymers of at least serogroups C, W-135 and Y can be further modified by the O-acetylation of the sialic acid residue by SynE (Oat). Each O-acetyl transferase is specific for each capsule type with OatC for serogroup C and OatWY for serogroup W-135 and Y polymers (51). To date no modification of the serogroup B capsule has been observed, although the K1 polysaccharide capsule of *E. coli*, which is chemically identical to serogroup B polymers, can be O-acetylated by the phase variable O-acetyl transferase NeuO (75).

The serogroup E capsule (29E) consists of GalNAc and KDO (25). Although KDO synthesis is essential for the expression of all meningococcal capsules, the serogroup E capsule is the only capsule known to contain this sugar (283). A second and redundant KDO synthesis locus is located in region A of the capsule locus in strains of this serogroup (104). A further four ORFs are located in the synthesis region of the serogroup E strains however the function of these putative proteins is unknown.

The serogroup H and Z capsules both contain glycerol 3-phosphate residues which alternate with Gal and GalNAc residues respectively, and have similarity to Gram-positive teichoic acids (8, 136, 295). The first two genes in the operons of the serogroup H and Z synthesis region A have more than 90% amino acid identity to each other and are hypothesised to be involved in the synthesis of the glycerol 3-phosphate due to similarity in protein sequence to *cps2B* and *cps2C* from *Actinobacillus pleuropneumoniae* (119). The function of the other two ORFs has not been examined although the third putative proteins in the operons have similarity to teichoic acid synthase genes (*cps2D* and *cps9D* from *A. pleuropneumoniae*, serovars 2 and 9.
respectively) while the fourth putative proteins have similarity to a capsule synthesis protein from the same locus in *A. pleuropneumoniae* (119). Serogroups 29E and Z possess capsules containing GalNAc although no biosynthetic pathway for GalNAc synthesis has been described (160).

The synthesis operon located in region A of the capsule locus is considered to be near identical for serogroups I and K and are termed *csikA*-E (119). Although the synthesis locus appears identical, the chemical composition of the capsules differs with the serogroup I capsule consisting of O-acetylated →4)-α-L-GulpNAcA-(1→3)-β-D-ManpNAcA(1→ while the serogroup K capsule consists of O-acetylated →3)-β-D-ManpNAcA-(1→4)-β-D-ManpNAcA-(1→ (186, 294). The synthesis of *N*-acetylmannosaminuronic acid (ManNAcA) is most likely achieved through the genes *csikABE* which encodes enzymes with amino acid sequence similarity to capsule biosynthesis genes of *Mannheimia haemolytica* which makes a capsule composed of repeating disaccharide ManNAcA-ManNAc (7). The biosynthetic pathway resulting in the formation of GulNAcA is less well characterised and the pathway is as yet undefined in *N. meningitidis*. The genes *csikC* and *csikD* have been identified as putative glycosyltransferases, possibly involved in the polymerisation of the capsule polymer. The proteins CsiKC and CsiKD differ by only two amino acids which may be sufficient for differences in substrate specificity, similar to SynD<sub>W-135</sub> and SynD<sub>Y</sub> in which the difference in polymerase activity is dependent upon a single amino acid (53, 119).

The serogroup L and X capsules both consist of GlcNAc and phosphate, but are composed of different linkages with the serogroup L capsule being comprised of →3)-
β-D-GlcNAc-(1→3)-β-D-GlcNAc-(1→3)-α-D-GlcNAc-(1→OPO3→ and the serogroup X capsule consisting of →4)-α-D-GlcNAc-(1→ OPO3 → (38, 135). Both the serogroup L and X synthesis operons comprise three ORFs, however these genes have no similarity to each other. The first gene in the csl operon is predicted to encode a phosphotransferase, while the third gene is predicted to encode an acetyltransferase. The second gene in the operon cslB, is predicted to encode a glycosyltransferase (119). The first gene in the serogroup X capsule synthesis operon putatively encodes a phosphotransferase, however the other two ORFs do not share similarity with any characterised proteins (119).

The meningococcal capsule, like that of E. coli, is synthesised in the cytoplasm at the inner membrane (223, 313). The precursor activated sugars are sequentially added to the non-reducing end of the polymer however, the acceptor molecule is not yet known. In E. coli K5, the capsule is polymerised in the absence of a lipid anchor, while the polymerisation of the K1 polysialic acid capsule occurs in the presence of a lipidated precursor (312). It has been demonstrated in meningococci expressing sialic acid containing capsule (B/C/W-135/Y) that in the absence of KDO synthesis and transfer to lipid IV₅ in LOS biosynthesis the expression of capsule was also reduced, possibly indicating a role of KDO in the biosynthesis of meningococcal capsules (283). Work in E. coli on the K1 capsule has suggested that the polymerisation of the polysialic acid capsule occurs within its own compartment termed a “sialasome” (251).

Regardless of the order in which the assembly occurs, the activated sugar monomers are polymerised and are attached to a diacylglycerol phospholipid anchor which is then transported from the cytoplasmic face of the inner membrane to the extracellular surface.
of the outer-membrane. The translocation of the capsule polysaccharide occurs via the ABC transporter machinery termed Ctr (Capsule transport) and as the possession and organisation of these genes (regions B and C) are consistent between all serogroups of *N. meningitidis* their function is considered to be the same for all capsule types (104). In *Neisseria* there are currently seven recognised members of the Ctr system, CtrA-G with *ctrA*-*D* found in a single operon and comprising region C, *ctrE*-*F* comprising region B, and *ctrG* is found within the synthesis locus region A of sialic acid containing capsule loci (102, 127, 284). CtrA is located in the outer-membrane and contains eight trans-membrane domains and the N-terminus is hypothesised to protrude into the cytoplasm where it is thought to contact CtrB (103, 104). CtrB is an inner-membrane protein which has two trans-membrane domains at either end of the proteins with the termini located in the cytoplasm leaving the region between the trans-membrane domains in the periplasm (104). CtrC is an inner-membrane protein with six trans-membrane domains and is hypothesised to form the pore in the inner-membrane through which the capsule polymer is extruded (104). CrtD has homology to KpsT, and the ATP-binding region of the ABC-transporter and has been predicted to be located in the cytoplasm (102, 206). CtrD associates with CtrC, mediating the active transport of the capsule polymer to the outer-membrane (104). CtrE/F were originally called LipA/B due to protein similarity to the *E. coli*, and although the expression of these proteins is essential for export of the meningococcal capsule and mutants result in the accumulation of diacylgyerol terminated capsule polymers within the meningococcus, their role in the transport process has not been defined (284).
**1.3.4.2 Regulation of capsule**

Only the regulation of serogroup B and C capsules have been investigated to date. The intergenic region between the synthesis and transport operons (*ctrA-synA*) of each of the different serogroups is different, and as this is the region where *trans*-acting regulatory factors have been demonstrated to bind, this suggests that the regulatory mechanisms could differ between serogroups (119, 266). For serogroup B/C/W-135/Y strains which all possess the same intergenic region, it has been shown that the two-component system regulator MisR will bind to the divergent promoter and regulate the expression of both the synthesis and transport machinery (285). In addition, the LysR type regulator CrgA has also been demonstrated to regulate the transcription of the synthesis operon and this regulation occurs in response to contact with the host epithelial cells (71).

The synthesis locus (region A) of many different serogroups has been observed to contain insertion sequences. The insertion sequence IS1301 has been identified in serogroups W, E, Z, I and K. The insertion sequence IS1016 has been observed in serogroup X. Despite the insertion of these mobile elements, this rarely disrupts capsule expression (119, 286). Lastly, phase variation is another mechanism by which capsule expression is regulated. The gene *synD* encoding the polysialyltransferase has a poly(C) tract located within the coding region (117).

**1.3.4.3 Distribution of capsule types**

The capsule locus possessed by strains, in general, is consistent within a clonal complex (Fig 1.13). Strains of ST-1, -4 and -5 cc are typically serogroup A strains with >98% of
Figure 1.13 Distribution of serogroup with respect to clonal complex.
The frequency of capsule expression with respect to the clonal complex of the isolate is plotted. Data was obtained from pubmlst/neisseria. (This study).
capsule expressing strains possessing a serogroup A capsule. Similarly ST-750 cc strains typically express a serogroup X capsule with the remaining split between serogroup B (14%) and Z (5%). Capsule switching can occur between strains of different serogroups (268). As the genetic difference between the sialic acid containing capsules is minimal, and in the case of B/C differences is in the sialic acid polymerase, these switching events occur more frequently and require the recombination of only a small fragment of DNA with one study characterising a 2kb recombination event which included the sialic acid polymerase gene synD (224). Another characterised capsule switching event involved the change from serogroup B to W-135 which involved the recombination of a 45kb region including the entire capsule locus (20). Although capsule switching is thought to be a rare event, multiple capsule types are observed within each clonal complex of strains in the current library (PubMLST.org/neisseria).

1.5 Aims of my thesis

The major surface carbohydrates of Neisseria meningitidis include the capsule, lipoooligosaccharide and the pilin glycan. Each of these structures are regulated and are variably distributed between strains of the species. Although alterations to the surface carbohydrates of the meningococcus are thought to aid in evasion of the host immune response, the potential exists for these variable structures to act as virulence determinants in their own right.

The aims of this study were:

1. Compare the rates of attachment and invasion into host epithelial cells by strains of different genetic lineages
2. Examine the role of the variable LOS inner core structure in attachment and invasion of host epithelial cells

3. Determine the biosynthetic pathway of the sugar UDP-GalNAc which is present in serogroup Z and E capsules but for which the biosynthesis is uncharacterised

This study aims to determine the role of surface expressed carbohydrates and their role in mediating attachment and/or invasion into host epithelial cells, and regulatory pathways which may lead to alteration in the expression of these structures.

1.6 References


73


279. Tsai, C. M. 2001. Molecular mimicry of host structures by lipooligosaccharides of Neisseria meningitidis: characterization of sialylated and nonsialylated lacto-N-neotetraose (Galβ1-4GlcNAcβ1-3Galβ1-4Glc) structures in


Chapter 2. Attachment and invasion of Neisseria meningitidis to host cells is related to surface hydrophobicity, bacterial cell size and capsule.
2.1 Abstract

We compared exemplar strains from two hypervirulent clonal complexes, strain NMB-CDC from ST-8/11 cc and strain MC58 from ST-32/269 cc, in host cell attachment and invasion. Strain NMB-CDC attached to and invaded host cells at a significantly greater frequency than strain MC58. Type IV pili retained the primary role for initial attachment to host cells for both isolates regardless of pilin class and glycosylation pattern. In strain MC58, the serogroup B capsule was the major inhibitory determinant affecting both bacterial attachment to and invasion of host cells. Removal of terminal sialylation of lipooligosaccharide (LOS) in the presence of capsule did not influence rates of attachment or invasion for strain MC58. However, removal of either serogroup B capsule or LOS sialylation in strain NMB-CDC increased bacterial attachment to host cells to the same extent. Although the level of inhibition of attachment by capsule was different between these strains, the regulation of the capsule synthesis locus by the two-component response regulator MisR, and the level of surface capsule determined by flow cytometry were not significantly different. However, the diplococci of strain NMB-CDC were shown to have a 1.89-fold greater surface area than strain MC58 by flow cytometry. It was proposed that the increase in surface area without changing the amount of anchored glycolipid capsule in the outer membrane would result in a sparser capsule and increase surface hydrophobicity. Capsulation modulated the surface hydrophobicity of both isolates however encapsulated strain MC58 was more hydrophilic than strain NMB-CDC. In conclusion, strain MC58 adhered less to host cells and had a more hydrophilic surface than strain NMB-CDC. Improved levels of adherence of strain NMB-CDC to cell lines was associated with increased bacterial cell surface and surface hydrophobicity. This study shows that there is diversity in bacterial
cell surface area and surface hydrophobicity within *N. meningitidis* and that these phenotypes can influence steps in meningococcal pathogenesis.

### 2.2 Introduction

*Neisseria meningitidis* is a Gram-negative diplococcus which is asymptptomatically carried in the nasopharynx by approximately 10% of the adult population but is also the causative agent of epidemic septicaemia and meningitis which results in 5-20% case fatality rates (82). Strains of *N. meningitidis* isolated from carriage are usually unencapsulated while invasive isolates are encapsulated, with the identity of the polymer being the foundation of serogrouping (21, 26). Based on the sequence similarity of housekeeping genes, *N. meningitidis* can be arranged into thirty-seven lineages of closely related sequence types (STs) termed clonal complexes (cc) which account for 61% of all strains collected (15). Some of these clonal complexes are more often associated with carriage than disease (>2:1), while other clonal complexes have been deemed to be hyperinvasive as these strains are more commonly associated with invasive disease than carriage (generally >5:1) (14). Within the hyperinvasive lineages, distinct phylogenetic clades are evident for ST-8/11 cc, ST-32/269 cc, and ST-41/44 cc (10). Despite the existence of carriage and hyperinvasive clonal complexes, the only common virulence determinants for invasiveness of disease causing isolates is encapsulation (79) and the carriage of phages (5).

Representative strains of the ST-4, ST-18 and ST-32 hyperinvasive lineages have been used to examine the interaction of meningococci with host cells (17, 53, 62, 64, 65, 91, 100, 103). The initial interaction between the meningococcus and the host cell is
mediated through the meningococcal type IV pilus. The meningococcus then retracts the pilus bringing the bacterium into close contact with the host cell surface (13). Intimate association of the meningococcus with the host cell membrane occurs through the binding by the meningococcal outer-membrane opacity (Opa) and Opc adhesins to carcinoembryonic antigen cell adhesion molecules (CEACAMs), and to heparin sulfate proteoglycans and integrins through bridges of vitronectin and fibronectin (13). The attachment and invasion of host cells is inhibited in vitro by the expression of capsule and lacto-N-neotetraose (LNT) α-chain on lipooligosaccharide (LOS) (59, 105). The presence of capsule and LNT-bearing LOS both inhibit Opa function, presumably by physical occlusion or by acting as a ligand of Opa (59, 105). Therefore meningococcal invasion of epithelial cells is favoured when LNT and capsule expression has been phased off through translational frame-shift due to the expansion or contraction of polymeric nucleotide tracts within the relevant biosynthetic genes. Since these virulence determinants play various and sometimes opposing roles in the infection process, phase variable expression of virulence determinants such as pili, LNT-bearing LOS and capsule is essential for the success of meningococcal pathogenesis (25).

The prototype strain of ST-8 cc, NMB-CDC (77), has been extensively characterised with regards to LOS structure, capsule expression (34, 38-42, 85, 93-95, 97), interaction with human monocytes (73), pilin (43), biofilm formation (67) and the rate of phase variation (74), as has the prototype ST-32 cc strain MC58 (24, 29, 63, 64, 66, 98, 100, 102-106). Both isolates express a serogroup B capsule which is a membrane anchored glycolipid, consisting of a sugar polymer with an average length of approximately 300 α2→8 linked sialic acid residues covalently bound to a diacylglycerol lipid anchor which tethers the polymer to the outer membrane (34, 38-42, 85, 93-95, 97). The sugar
polymer is assembled at the cytoplasmic surface of the inner membrane, is moved to the lipid anchor and is translocated across the membranes *en bloc* to the bacterial cell surface (34, 38-42, 85, 93-95, 97). However, strains NMB-CDC and MC58 express different LOS structures (72) and type IV pilin classes (44). Strain MC58 expresses class I pili which undergo phase and antigenic variation (32, 74), while strain NMB-CDC expresses class II pili which do not undergo antigenic variation (32, 74), a phenotypic characteristic of strains from the ST-8/11 cc (10, 16, 44). Since the model for meningococcal attachment and invasion of host cells relies on phase variation for this process to succeed (99) and since strain NMB-CDC is known to phase vary the expression of genes at low frequency (74), we compared prototype strain NMB-CDC to the ST-32 prototype strain MC58. In testing the role of known virulence determinants that affect meningococcal attachment and invasion into host cells, we also tested whether bacterial cell size differed between isolates. Dalia and Weiser (23) have recently shown that minimization of bacterial cell size allows for complement evasion by *Streptococcus pneumoniae* and has proposed that long chain variants may be suited to a mucosal lifestyle, with invasive strains typically retaining a diplococcal form. Therefore, we examined whether meningococcal isolates varied in surface area and looked for correlates with the rates of attachment and invasion.

2.3 Methods

2.3.1 Bacterial strains and growth conditions

*N. meningitidis* strains NMB-CDC (ST-8, serogroup B, immunotype L2,4, PilEII), and MC58 (ST-32, serogroup B, immunotype L3,7,9, PilEI) were cultured under aerobic
conditions with 5% CO2 at 37°C on GC agar (GCA) or GC broth (GCB) (Oxoid) supplemented with 0.4% glucose, 0.01% glutamine, 0.2 mg of cocarboxylase per litre, and 5 mg of Fe(NO3)3 per litre. The wild-type strains and constructed mutants were piliated and expressed Opa, except where stated otherwise. The wild-type strains and constructed mutants used in this study are shown in Table 2.1. Antibiotic selection for meningococcal mutants was performed on GCA containing 100 µg/ml of kanamycin (sulfate salt), 60 µg/ml of spectinomycin, 5 µg/ml of tetracycline or 2 µg/ml of erythromycin (Sigma). *Escherichia coli* DH5α was used as a host for all DNA manipulations. It was routinely grown in Lysogeny broth (LB) or on Lysogeny agar (LA, Oxoid) which, where appropriate, were supplemented with antibiotics at the following concentrations: ampicillin at 100 µg/ml, spectinomycin at 50 µg/ml, kanamycin at 50 µg/ml, erythromycin at 300 µg/ml, tetracycline at 12.5 µg/ml and chloramphenicol at 30 µg/ml (Sigma).

### 2.3.2 Genome sequencing, assembly and annotation.

Sequencing of strain NMB-CDC was performed on a Roche 454 GS FLX and an Illumina GAIIx. The 454 yielded 92 Mbp (44 x) from a 3 Kbp mate pair library, average read length 403 bp. The GAIIx yielded 370 Mbp (176 x) of 36 bp paired-end reads from a 200bp fragment library. *De novo* assembly of the 454 reads was performed using gsAssembler 2.6.0, resulting in 158 contigs longer than 200 bp within 5 scaffolds totaling 2.1 Mbp. The GAIIx reads were aligned to the 454 assembly to correct two 454 homopolymer assembly errors using Nesoni 0.63 (http://vicbioinformatics.com). The draft genome was annotated using Prokka 1.01.
### Table 2.1. Strains used in this study.

<table>
<thead>
<tr>
<th>Strain Name</th>
<th>Genotype</th>
<th>*Major Phenotype</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMB</td>
<td>B:2B:P1.2,5:L2 (CDC8201085)</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>(83)</td>
</tr>
<tr>
<td>MC58</td>
<td>B:15:P1.7,16b:L3,7,9 (ATCCBAA-335)</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>(89)</td>
</tr>
<tr>
<td>M7</td>
<td>NMB∆synA::Tn916</td>
<td>Cap-, Opa+, NANA-LNT+, PilEII+, GLY+</td>
<td>(83)</td>
</tr>
<tr>
<td>JKD5106</td>
<td>NMB∆misR::tetM</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>(89)</td>
</tr>
<tr>
<td>JKD5114</td>
<td>NMBΔpgf::aphA-3</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY-</td>
<td>(43)</td>
</tr>
<tr>
<td>JKD5122</td>
<td>NMBΔpilE::aadA</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>(43)</td>
</tr>
<tr>
<td>CKNM367</td>
<td>NMB∆synD::aphA-3</td>
<td>Cap-, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM385</td>
<td>NMB∆lst::aphA-3</td>
<td>Cap+, Opa+, NANA-LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM389</td>
<td>NMBΔpilE::ermC</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM390</td>
<td>NMBΔiga::pilEI(MC58)-aadA</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM394</td>
<td>NMBΔiga::aadA</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM396</td>
<td>NMBΔiga::aadA pilE::ermC</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM397</td>
<td>NMBΔiga::pilEI(MC58)-aadA pilE::ermC</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM416</td>
<td>MC58Δpgf::aphA-3</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY-</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM417</td>
<td>MC58ΔpilE::aphA-3</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM419</td>
<td>MC58ΔsynD::aphA-3</td>
<td>Cap-, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM420</td>
<td>MC58ΔsynB::aadA</td>
<td>Cap+, Opa+, NANA-LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM421</td>
<td>MC58Δlst::aphA-3</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM423</td>
<td>MC58ΔmisR::tetM</td>
<td>Cap+, Opa+, NANA+LNT+, PilEII+, GLY+</td>
<td>This study</td>
</tr>
</tbody>
</table>

* Phenotypes: Capsule (Cap), Opacity proteins (Opa), N-acetyl neuraminic acid present on LOS (NANA), LOS α-chain of lacto-N-neotetraose (LNT), pilus expression (PilI), pilus class I or II (PilEI or PilEII), and the presence of glycosylation (GLY) which can occur on outer membrane proteins other than pilin.

(http://vicbioinformatics.com). Prokka uses Prodigal, Aragorn and RNAmer to identify CDS, tRNA, and rRNA features respectively. Gene function was predicted via HMMER3 profile searches against PFAM and CDD; and BLAST+ similarity searches against UniProtKB. All annotated protein sequences from published Neisseria genomes were clustered using OrthoMCL (52) with a cut-off of 1e -50 to identify any new genes. Venn diagrams were created using this website tool (http://creately.com/Draw-Venn-Diagrams-Online). GenBank Accession numbers for the genomes of the strains
analysed in this study are as follows: M6190 (AEQF0000000.1) and ES14902 (AEQI0000000.1), G2136 (CP002419.1), FAM18 (NC_008767.1), MC58 (NC_003112.2), 961-5945 (AEQK0000000.1), S0108 (ADWN0000000.1), K1207 (ADWM0000000.1) and H44/76 (CP002420.1).

The \textit{opaA} allele was amplified with primer pair KAP260 (5'\text{-}TACGCTGCAGAAAAATGAATCCAGCCCCC-3') and KAP261 (5'\text{-}ACATCGGAAATCCAAGTGTTGCC-3'), to generate a 3 kb amplicon. The \textit{opaB} allele was amplified with primer pair KAP230 (5'\text{-}AGGAGCAGTTCGCCTTGGAGGG-3') and KAP231 (5'\text{-}CAGACACCTGCAAGTATCCGC-3'), generating a 3 kb amplicon. The \textit{opaD} allele was amplified with primer pair KAP232 (5'\text{-}TAACGGGTAGGTTATCGGTCC-3') and KAP233 (5'\text{-}TGGAACCCAAATCGACGGAGGC-3'), generating a 5 kb amplicon. The \textit{opaJ} allele was amplified with primer pair KAP234 (5'\text{-}ACAACTTTGGGCTTTGGCACAAGC-3') and KAP235 (5'\text{-}ACGCGACAGAGGATGGCATAAGC-3'), resulting in a 4 kb amplicon. Each amplicon was used as a template for the PCR using primers KAP236 (5'\text{-}CCGTATTATGTGCAGGGATTAGC-3') and KAP237 (5'\text{-}TCCAAGCGTCCAGTTGGG-3') to confirm the presence of an \textit{opa} allele in each locus, followed by sequencing with KAP236, KAP237 and KAP238 (5'\text{-}AATCGTGGTAATCGCTTGGGC-3') to identify both the sequence of the allele and the phase status. The \textit{Neisseria meningitidis} Opa sequence database (http://neisseria.org/nm/typing/opa/), developed by Keith Jolley and Martin Callaghan and sited at the University of Oxford was used to identify the alleles.
The absence of pilC2 from the genome sequence of NMB-CDC was noted. The density of the reads of the pilC1 and pilC2 loci were assessed and indicated that a single copy of pilC was present in the genome. As both loci appeared to have good coverage, we predicted that pilC2 was absent from the genome of NMB-CDC. This was confirmed by PCR analysis using primer pair KAP509 (5’-CGTTTGTGGACGCACTGCTG-3’) and KAP510 (5’-CGACAAATTCCGCACAGGCAGC-3’) which resulted in an amplicon of 2.3 kb when amplified from NMB-CDC, compared to an amplicon of 8.7 kb when amplified from MC58 which possesses the pilC2 allele.

2.3.3 Cell lines and culture

The immortalised Detroit 562 (human pharyngeal carcinoma epithelial cells, ATCC CCL-138) and 16HBE14σ- (transformed human bronchial epithelial cells) (67) cell lines were grown to confluence in Minimal Essential Media (MEM) with Earles salts supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 1 mM sodium pyruvate and 1 x non-essential amino acids (Invitrogen) as per ATCC recommendations (68). At confluence the cells were lifted with 0.05% trypsin/EDTA (Invitrogen) and resuspended in cell media. Cells were counted and 1x10^5 cells were seeded into each of the wells of a 24-well plate (Nunc).

2.3.4 Construction of mutants of strain NMB-CDC and MC58

The plasmids constructed and used in this study are listed in Table 2.2.

To create capsule negative variants of strain NMB-CDC and MC58, synD was inactivated with a synD::aphA-3 cassette. A region containing the 5’ end of synD was
amplified from NMB-CDC gDNA with primer pair, KAP196 (5’-TATGCTTGCTAGGAGCAGTAGC-3’) and KAP366 (5’-GATCTGAGGATCTAGCTGATGAACTAATTGGC-3’). A region containing the 3’end of synD was amplified from NMB-CDC gDNA with primer pair, KAP369 (5’-CCTTGCGACATCCCATCCCCTTTGC-3’) and KAP367 (5’-CTAGGATCTGCAGATCTATTTGAGCTTCCTAGAAGCCC-3’). These first round PCR products were used as templates for a second round of PCR amplification with the nested primer pair KAP365 (5’-CTGCTTGCAATATTCGCAAAGGTGC-3’) and KAP368 (5’-GCATATTCAGGAAAGGGACATGC-3’) to create a fusion product with internal BamHI and PstI sites introduced by KAP366 and KAP367. This fragment was cloned into the BamHI and PstI sites of pHSG576 which were removed by treatment with T4 DNA polymerase, resulting in pCMK718. The aphA-3 cassette conferring kanamycin resistance was amplified from pUC18K with primer pair KAP128 and KAP129 and cloned into the BamHI site of pCMK718 resulting in pCMK727. Strain NMB-CDC and strain MC58 were transformed with pCMK727 and the transformants were identified by resistance to kanamycin then confirmed by PCR using primer pair KAP196 and KAP369, resulting in CKNM367 and CKNM419, respectively.

To construct an isolate of strain MC58 that does not synthesise sialic acid, the sialic acid synthetase, synB, was inactivated using a synB::aadA cassette. An internal region of synB was amplified from NMB-CDC gDNA with primer pair KAP371 (5’-GTTATACTTGCAGCAGCAAACAGC-3’) and KAP372 (5’-GTTCGTTGGTTGACCTACTGAACGA-3’). This fragment was cloned into the HindIII site of pHSG576, resulting in pCMK720. The cassette containing the
spectinomycin resistance gene *aadA* was obtained from pHP45Ω by digestion with *Hind*III and cloned into the *Hind*III site of pCMK720 resulting in pCMK723. Strain MC58 was transformed with pCMK723 and the transformants were identified by resistance to spectinomycin then confirmed by PCR using primer pair KAP370 (5’-TCCTGAAACGTGGAATGTTTCTGC-3’) and KAP373 (5’-ATAGAGACATCTGCATTGCCTGG-3’), resulting in CKNM420.

To create mutants in which LOS sialylation was absent, *lst*, encoding the LOS sialylation transferase, was inactivated using a *lst::aphA-3* cassette. The *aphA*-3 cassette conferring kanamycin resistance was amplified from pUC18K with primers KAP128 and KAP129 and cloned into the *Hinc*II site of pCK90 (Table 2.2) resulting in pCMK737. Strain NMB-CDC and MC58 were transformed with pCMK737 and the transformants were identified by resistance to kanamycin, then confirmed by PCR using primer pair KAP183 (5’- CAAAAGCCTGCACAATCGGCAGC-3’) and KAP184 (5’-GCAAATCCTGCCACGACAGTTTCC-3’), resulting in CKNM385 and CKNM421, respectively.

A mutant of strain MC58 that contained an inactivated *misR* was created by transformation with pJKD2539 (Table 2.2) to create CKNM423 as described previously (93).

### 2.3.5 Transformation protocols

Strain NMB-CDC was transformed via natural transformation as described previously (35). Natural transformation rates of strain MC58 are particularly low, therefore a modified chemical transformation procedure was used (6). Approximately 2 x 10⁹ cfu
Table 2.2. Plasmids used in this study.

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAermC’G</td>
<td>High copy vector</td>
<td>(109)</td>
</tr>
<tr>
<td>pHG45Ω</td>
<td>pUC18 carrying the aadA polar cassette</td>
<td>(70)</td>
</tr>
<tr>
<td>pHSG298</td>
<td>High copy cloning vector</td>
<td>(87)</td>
</tr>
<tr>
<td>pHSG576</td>
<td>Low copy cloning vector</td>
<td>(87)</td>
</tr>
<tr>
<td>pUC18K</td>
<td>pUC18 carrying aphA-3 non-polar cassette</td>
<td>(58)</td>
</tr>
<tr>
<td>pCK90</td>
<td>pHSG576+Ist</td>
<td>(42)</td>
</tr>
<tr>
<td>pCMK716</td>
<td>pHSG576+pilEi(MC58)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK718</td>
<td>pHSG576+synD</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK720</td>
<td>pHSG576+synB</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK723</td>
<td>pHSG576+synB::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK727</td>
<td>pHSG576+synD::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK737</td>
<td>pHSG576+Ist::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK739</td>
<td>pHSG576+pilEi(MC58)::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>pJKD2407</td>
<td>pHSG576+pglF::aphA-3</td>
<td>(43)</td>
</tr>
<tr>
<td>pJKD2425</td>
<td>pHSG298+pilEi(NMB)</td>
<td>(43)</td>
</tr>
<tr>
<td>pJKD2426</td>
<td>pHSG298+pilEi(NMB)::aadA</td>
<td>(43)</td>
</tr>
<tr>
<td>pJKD2539</td>
<td>pHSG298+misR::tetM</td>
<td>(93)</td>
</tr>
<tr>
<td>pJKD2580</td>
<td>pHSG576+NMBiga</td>
<td>This study</td>
</tr>
<tr>
<td>pJKD2581</td>
<td>pHSG576+NMBiga::aadA</td>
<td>This study</td>
</tr>
<tr>
<td>pJKD2728</td>
<td>pHSG298+pilEi(NMB)::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>pJKD3172</td>
<td>pUC18+MS11iga::aadA</td>
<td>C.S. Ryan</td>
</tr>
<tr>
<td>pJKD3341</td>
<td>pUC18+MS11iga::pilEi(MC58)-aadA::ermC</td>
<td>This study</td>
</tr>
</tbody>
</table>

were harvested after 12 hr growth on GCA and resuspended in 1 ml ice cold GC-TSB (transformation broth as previously described (20) made with GCB instead of LB). An aliquot of 200 µl was added to 200 ng of plasmid DNA and incubated on ice for 20 mins. A 1 ml aliquot of room temperature GCB was added to the bacteria/DNA mixture and incubated at 37°C with shaking for 1 hr. The bacteria were harvested via centrifugation and plated on GCA containing the appropriate antibiotics for the selection of transformants.
2.3.6 Attachment and invasion assays

Detroit 562 epithelial cells and 16HBE14σ- cells were grown to confluence in 24-well plates and utilised for invasion assays. Invasion assays were conducted as described previously (104) with the following modifications. The epithelial cells were inoculated with suspensions of meningococci at a multiplicity of infection (MOI) of 100:1 in epithelial cell media containing 2% FCS (103). The number of eukaryotic cells present in each well was estimated by sacrificing two wells for counting. To do this, the media was removed, the cell monolayer was washed with Dulbecco's PBS and the cells lifted with trypsin. The cells were enumerated using a haemocytometer and dead cells were excluded from this count by trypan blue staining. The average of the number of cells was used to estimate the number of eukaryotic cells present in each of the experimental wells for each experiment. Bacteria were harvested from a 14 hr agar plate culture, resuspended in GC broth and normalised to an OD560nm of 0.4. In the original experiment, the viable count was determined for OD560nm of 0.4. This standard was applied to each experiment thereafter to calculate the MOI of 100:1. For each experiment the actual number of bacteria was confirmed by viable count of the inoculum.

The epithelial cell monolayers were challenged with strains of N. meningitidis for either 1 hr or 6 hrs as indicated, and incubated in 5% CO₂ at 37°C. Non-adherent bacteria were then removed by washing the monolayers three times with Dulbecco’s PBS (Invitrogen). Following 1 hr incubation in either epithelial cell media containing 2% FCS (for enumeration of cell associated bacteria) or epithelial cell media containing 2% FCS with 100 μg/ml gentamycin (Sigma) (for enumeration of intracellular, protected bacteria), the monolayers were washed three times with Dulbecco’s PBS to remove
gentamycin and the epithelial cells were lysed with 1% saponin (Sigma) in epithelial cell media containing 2% FCS to release the intracellular bacteria. The number of cell associated bacteria and intracellular bacteria were enumerated by viable count. At least three biological repeats containing three technical repeats were performed for each strain. Attachment was determined as the proportion of the inoculum which had attached to epithelial cells, and invasion as the proportion of the attached bacteria which had invaded the epithelial cells. A Mann-Whitney t-test was used to determine statistically significant differences. All strains assayed were sensitive to gentamycin and resistant to 1% saponin as determined by viable count (data not shown).

2.3.7 Determination of capsule, LOS, pilin and Opa phenotype

Colony immunoblotting was utilised to identify transformants expressing class I pilin and to determine the rate of phase variation of synD and lgtA as described previously (83). Briefly, strains were plated at a dilution that obtained 30-300 colonies per GCA plate. The colonies were lifted using a nitrocellulose membrane (GE Healthcare) and the plates were re-incubated to allow the colonies to re-grow. The nitrocellulose discs were air dried, then fixed by UV exposure. The membranes were washed with TBS, then blocked with 2% BSA (Millipore) in TBS for 1 hr. Primary antibodies against class I pilin (MAb SM1 (101)), capsule (MAb 2-2-B (110)) and LNT (MAb 3F11 (1)) were used at concentrations of 1:100, 1:1000 and 1:10 in blocking buffer, respectively. Conjugated secondary antibodies (Santa Cruz Biotechnology) were used at a concentration of 1:1000 and the immunoblots were developed using NBT/BCIP was used as per the manufacturer’s instructions (Sigma Cat#N6876 and Cat#B8503 respectively). A negative control strain was included in each assay.
Opa expression was determined by immunoblotting. Whole cell lysates (750 ng) were separated by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by standard methods and transferred to nitrocellulose membranes. The membranes were blocked overnight with 2% BSA in TBS. The polyclonal rabbit anti-Opa primary antibody (86), was used at 1:1,000. A Horse radish peroxidase-conjugated anti-Rabbit IgG secondary antibody at a dilution of 1:1000 (Santa Cruz Biotechnology) was used and the membrane was developed with an ECL kit (GE Healthcare).

2.3.8 qRT-PCR

Total RNA was purified from bacteria using the High Pure RNA extraction kit (Roche). These RNA preparations were assayed for DNA contamination by PCR and nucleic acid concentration was determined by absorbance in a Nanodrop spectrophotometer (Thermo Scientific). cDNA was generated from 1 μg of total RNA as template and random hexamers using reverse transcriptase (New England Biolabs). A mock reaction lacking reverse transcriptase was also included and served as the negative control in the qRT-PCR reaction. qRT-PCR was performed using SYBR green I mastermix (Roche). Melt curve analysis confirmed the presence of a single product from each PCR reaction and the reverse transcriptase negative control was used to detect DNA contamination. The primer pairs used to assess the transcript levels were: \textit{lpxA} were DAP135 (5\textquotesingle- CGTTTTGGGCGCTACAC-3\textquotesingle) and DAP136 (5\textquotesingle- GGCGGTCATGGCGTAGTC-3\textquotesingle); \textit{synA} were KAP353 (5\textquotesingle- GGTCATGATTCACGGCGAC-3\textquotesingle) and KAP354 (5\textquotesingle- CCTACAGCTGCCTGCTAG-3\textquotesingle); \textit{ctrA} were KAP355 (5\textquotesingle- GTTTGGCGATGGTAAATGCTT-3\textquotesingle) and KAP356 (5\textquotesingle-
ACCAACTGCTCTGGCAACTT-3’); \( \text{lst} \) were KAP343 (5’-
CAGGTATTTGGCGATGAGT-3’) and KAP344 (5’-
GCCGTCGCTCCATAATGTTTT-3’); \( \text{lgtG} \) were KAP333 (5’-
GCTCGCTCAAACCAGAAAAA-3’) and KAP334 (5’-
GGAATGCGCCATAATCAGTT-3’); and \( \text{dsbD} \) were KAP357 (5’-
GTTGGGACAGCCTTCTTTCA-3’) and KAP358 (5’-
TTGCATAAGGAAAGGCAACC-3’). The relative changes in gene transcription were calculated using the comparative \( \Delta \text{CT} \) method (54), normalised to the level of \( \text{lpxA} \) transcript. Each set of qRT-PCRs was examined in triplicate and was repeated with at least three independent RNA preparations (96).

2.3.9 Flow cytometry for the detection of meningococcal capsule and bacterial cell size using high throughput microscopy

The procedure for the detection of capsule by flow cytometry was essentially that of Tzeng et al. (95). Meningococci grown to early stationary phase were collected and suspended in PBS. Cells were washed with 0.5 ml of PBSB buffer (0.5% [w/v] bovine serum albumin in PBS) once before the incubation with 100 µl of MAb 2-2-B (1:500 dilution) for 1 h at 37°C. After washing with 0.5 ml of PBSB buffer, the cells were further incubated with a 1:100 dilution of AF647-goat anti-mouse IgM (Invitrogen) for 1 hr at 37°C. Cell pellets were obtained after centrifugation and were resuspended in 0.5 ml of PBS and 0.5 ml of 5 µM SYTOX Green (Invitrogen), prepared in PBS from a 5 mM stock solution in dimethyl sulfoxide, to assess the membrane integrity of the bacteria (51, 76). The cells were then analysed for relative size and fluorescence labelling with a BD Influx cell sorter (BD Biosciences), in a biohazard containment
hood, utilising a 488 nm excitation laser and 520/15 nm emission filter for detection of SYTOX Green staining and 640 nm excitation laser and 670/30 nm emission filter for detection of the AF647 secondary antibody. All fluorescence emissions were acquired in log, and data was collected for 50,000 cells using BD Software V1.0.0.653 (BD Biosciences), the operating software on the BD Influx. Experimental data were analysed and geometric mean fluorescence was calculated using FlowJo V7.2.5 (Tree Star Inc.) flow cytometry analysis software. The surface area for spheres was calculated using: \(4\pi r^2\) (where “\(r\) is the radius) and multiplied by a factor of two for a diplococcus. For high throughput microscopy to determine cell size, ethanol fixed samples were prepared and stained with SYTOX Green as above. Images for 50,000 events were collected for each sample using the AMNIS ImageStream Imaging Cytometer (AMNIS Corporation) with INSPIRE V4.1 acquisition software. The bright field image was collected in channel 1 after white light illumination and the corresponding SYTOX Green signal collected in channel 2 (480-560 nm) after 488 nm laser excitation (12). Out of focus cells, debris and cell clumps were excluded from statistical analysis and cell area calculated using IDEAS V5.0 image analysis software (AMNIS). Cell area was calculated using an erode mask on the bright field image to separate the image pixels from the background pixel intensity. Pixel number is then converted to area in \(\mu\)m\(^2\).

### 2.3.10 Electrophoretic mobility shift assay (EMSA)

EMSA experiments were performed following previously reported procedures (96). The promoter fragment was obtained by PCR using primer pair LJ6 (5’-CATCCTACAATTTAAGTCCAC-3’) and JS44 (5’-GCTTGTTCATTGCTACCAAGTG-3’) and end-labelled with \([^{32}\text{P}]-\text{ATP}\) using T4

\[96\]
kinase (NEB). Competition with excess specific (unlabeled probes) and nonspecific competitors, a 593 bp internal coding sequence of misR obtained by PCR amplification using primer pair YT45 (5’-CGTAGATGACGATGCCCTGCTAACCG-3’) and YT46 (5’-GGCGGATGCTGGAGATGTGTACGTCG-3’), was performed to assess the specificity of the interaction.

2.3.11 Capsule purification, electrophoresis and staining

Capsule polysaccharides were purified from strains NMB-CDC and MC58 as per Hobb et al. (34). Capsule preparations were separated by deoxycholic acid (MP Biomedicals) polyacrylamide gel electrophoresis (DOC-PAGE) (71) and visualised by silver stain (Bio-Rad).

2.3.12 Hydrophobic interaction chromatography (HIC)

The procedure for the hydrophobic interaction chromatography was essentially that of Tzeng et al. (95) and Karlyshev et al. (46). Bacterial suspensions in PBS were loaded onto disposable plastic columns packed with 2 ml (bed volume) of octyl-Sephrose CL-4B resin (Sigma) and were washed with 5 ml of 0.2 M ammonium sulfate in 10 mM sodium phosphate buffer, pH 6.8. The proportion of bacteria cells that were eluted from the column was determined by optical density OD$_{600}$.

2.3.13 Microbial adhesion to solvents (MATS)

MATS was performed using the modified procedure of Ly et al. (56). N. meningitidis was harvested from overnight growth on agar plates, resuspended in GC broth and
centrifuged briefly at 60 x g to remove clumps. The supernatant was measured and the cell suspension equalised to an OD$_{560}$ of 0.4. The cells were pelleted by centrifugation at 3000 x g and resuspended in 0.01 M potassium phosphate buffer (PPB). The pellet was washed once again and resuspended in PPB and the OD$_{560}$ equilibrated to 0.4. Bacterial suspensions were mixed in a 6:1 ratio with hexadecane in glass McCartney bottles, vortexed for 30 sec to thoroughly mix the two phases and left to stand at room temperature for 30 min for the phases to separate. An aliquot of the aqueous phase was measured at an OD$_{560}$ nm. The percentage hydrophobicity was calculated using the formula (1-(Abs$_2$/Abs$_1$)) x 100%.

2.4 Results

2.4.1 Genetic comparison of *N. meningitidis* strain NMB-CDC (ST-8 cc) with *N. meningitidis* strain MC58 (ST-32 cc).

*N. meningitidis* strain NMB-CDC is the prototype ST-8 strain belonging to ST-8 cc while strain MC58 is the prototype ST-32 strain belonging to ST-32 cc. We sequenced and annotated the genome of strain NMB-CDC and used orthologous group (OG) clustering (OrthoMCL) analysis to show that this isolate is highly related to G2136 from the ST-8 cc (10) and FAM18 prototype strain of ST-11 cc but was least conserved with ST-32 cc representative strain MC58 (Figure 2.1). Phylogenetic trees from previous studies (eg. Budroni et al. (10)) have also shown that ST-8 cc and ST-11 cc isolates are highly related to one another, and are more distantly related to strains from ST-32 cc. Table 2.3 summarises the alleles of putative virulence determinants possessed by strain NMB-CDC and MC58. Of interest, was the observation that strain NMB-
CDC did not possess a phase variable *lgtA* indicating that expression of LNT-bearing LOS was not phase variable in this strain consistent with the sequenced genomes of strains from ST-11 cc (FAM18, M6190, ES14902, S0108 and K1207) and ST-8 cc (G2136 and 961-5945). In addition, both isolates had multiple copies of the meningococcal disease associated phage (MDA) (data not shown).

2.4.2 Bacterial cell size differs between strain NMB-CDC and MC58.

While determining the viable counts for the attachment and invasion assays, it was noticed that at the same OD (OD$_{560}$ of 0.4), the liquid cultures of strains NMB-CDC and MC58 yielded different viable counts. Strain NMB-CDC had 2.2 +/- 0.17x10$^8$ colony forming units (cfu) per ml, whilst strain MC58 had a 5.5 +/- 0.38 x10$^8$ cfu/ml which is a 2.5-fold difference (p<0.0001). Since optical density correlates with biomass (48), we postulated that as there is a direct proportional relationship between cell volume and mass, the cells of strain NMB-CDC must contain 2.5 times the dry mass and therefore have 2.5 times the volume of cells of strain MC58 (49). If the relative volume of strain NMB-CDC is 2.5 and for strain MC58 is 1, then using a standard calculation for
volume and surface area of a sphere, a single coccus of strain NMB-CDC was determined to possess a surface area 1.85 times larger than a single coccus of strain MC58. A comparison of the non-encapsulated variants of both strains, M7 and MC58ΔsynB, using live samples and analysed by flow cytometry revealed that the surface area of strain M7 was 1.6-fold larger than that of MC58ΔsynB (p<0.0001). Flow cytometry and high throughput microscopy using the AMNIS ImageStreamX confirmed the hypothesis that the cells of these isolates were of different sizes. Bacteria harvested from mid-log phase growth in liquid media, were fixed in 70% ethanol and then were analysed by AMNIS ImageStreamX which captures high resolution bright field and fluorescent images of the bacteria as they pass the detector. Fifty thousand events were captured and the in-focus images were gated into population R1 which was plotted
<table>
<thead>
<tr>
<th>Protein</th>
<th>*Gene designation</th>
<th>Function</th>
<th>**Phase variable tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>LgtA</td>
<td>NMB1929</td>
<td>LNT expression (45)</td>
<td>Yes</td>
</tr>
<tr>
<td>LgtG</td>
<td>NMB2032</td>
<td>N2-3 LOS glycosyl-transferase (2)</td>
<td>Yes</td>
</tr>
<tr>
<td>^OppA</td>
<td>NMB0442</td>
<td>Opacity protein/adhesin (11)</td>
<td>Yes</td>
</tr>
<tr>
<td>^OppB</td>
<td>NMB1636</td>
<td>Opacity protein/adhesin (11)</td>
<td>Yes</td>
</tr>
<tr>
<td>^OppD</td>
<td>NMB1465</td>
<td>Opacity protein/adhesin (11)</td>
<td>Yes</td>
</tr>
<tr>
<td>^OppJ</td>
<td>NMB0926</td>
<td>Opacity protein/adhesin (11)</td>
<td>Yes</td>
</tr>
<tr>
<td>^Oppc</td>
<td>NMB1053</td>
<td>Opacity protein/adhesin (103)</td>
<td>No</td>
</tr>
<tr>
<td>^++PiIC2</td>
<td>NMB0049</td>
<td>Pilus tip adhesion (64)</td>
<td>Yes</td>
</tr>
<tr>
<td>^+++PilE</td>
<td>NMB0018</td>
<td>Major pilin (16)</td>
<td>No (I)</td>
</tr>
<tr>
<td>PglA</td>
<td>NMB0218</td>
<td>Pilin glycan galactose transferase (18)</td>
<td>Yes</td>
</tr>
<tr>
<td>^+++PilB</td>
<td>NMB1820</td>
<td>Pilin glycan synthesis (18)</td>
<td>No</td>
</tr>
<tr>
<td>PglE</td>
<td>NMB0624</td>
<td>Pilin glycan galactose transferase (18)</td>
<td>Yes</td>
</tr>
<tr>
<td>^+++PglG</td>
<td>NMC0401</td>
<td>Pilin glycan glucose transferase (7)</td>
<td>Absent</td>
</tr>
<tr>
<td>^+++PglH</td>
<td>NMC0400</td>
<td>Pilin glycan glucose transferase (7)</td>
<td>Absent</td>
</tr>
<tr>
<td>Pgil</td>
<td>NMB1836</td>
<td>Pilin glycan O-acetyl transferase (107)</td>
<td>Yes</td>
</tr>
<tr>
<td>PptA</td>
<td>NMB0415</td>
<td>Pilin PEA transferase (61)</td>
<td>Yes</td>
</tr>
<tr>
<td>SynD</td>
<td>NMB0067</td>
<td>Sialic acid polymerase (30)</td>
<td>Yes</td>
</tr>
<tr>
<td>App</td>
<td>NMB1985</td>
<td>Neisserial penetration and adhesion protein (80)</td>
<td>No</td>
</tr>
<tr>
<td>MspA</td>
<td>NMB1998</td>
<td>Meningococcal serine protease A (92)</td>
<td>Yes</td>
</tr>
<tr>
<td>^++++NadA</td>
<td>NMB1994</td>
<td>Neisserial adhesin A (22)</td>
<td>No (1)</td>
</tr>
<tr>
<td>NhhA</td>
<td>NMB0992</td>
<td>Neisseria hsf/hia homologue (81)</td>
<td>No</td>
</tr>
</tbody>
</table>

* Gene designations as assigned to strain MC58 are NMBXXXX (88). NMCXXXX designations for open reading frames found in strain FAM18 but absent in strain MC58 are used where appropriate (4).

** The potential for phase variation was determined by the length of the homopolymeric (> 5 bp) or pentameric tracts (> 5 copies) within these genes.

^ Strain MC58 has the following alleles: allele 96 (OpA), allele 288 (OpB), allele 147 (OpD), allele 218 (OpJ). The Opa loci from strain NMB-CDC were amplified and sequenced (see Methods) and were determined to be: allele 246 (OpA), allele 277 (OpB), allele 161 (OpD) and allele 257 (OpJ). This is the typical allelic pattern for ST-8 cc isolates (11).

^ ^ Oppc and PiIC2 were absent from the sequenced genome of strain NMB-CDC and this was confirmed by PCR (see methods) (77).

*** The sequence and expression profile of the pilin loci and glycosylation status of strain NMB-CDC has been published previously (43, 88). (I) denotes Class I and (II) denotes Class II pilin expression (16).

**** The allele designations of NadA are described in Comanducci et al. (22).

---

according to scatter and bright field area (Figure 2.2). Examination of the images revealed that the sample could be divided into single diplococci (R2) and clumps (R3). Five hundred images were collected from each sample and the average pixel density was calculated for each sample set. Using this strategy, the surface area of the dehydrated cells of strain NMB-CDC was calculated to be 14.8µm² while MC58 had a
surface area of 10.0\(\mu\)m\(^2\), (p<0.0001), indicative of a 1.5-fold greater surface area for strain NMB-CDC. Lastly, live unfixed cells from the liquid mid-log phase cultures were processed by the BD Influx cell sorter with a small particle detector for analysis of bacterial cell size. The average forward scatter geometric mean of four independent experiments was calculated from 50,000 events. Since forward scatter is proportional to particle surface area (19) this data revealed that the cells of strain NMB-CDC had a surface area ~1.89 fold greater than cells of strain MC58 (p<0.0001, Figure 2.3). In summary, two measurements of live unfixed and one made on fixed cells indicated that strain NMB-CDC had a surface area that was 1.85-1.89-fold and 1.5-fold greater, respectively, than strain MC58. To determine if bacterial cell size had an effect on growth in liquid culture, equivalent cfu/ml were used to inoculate GC broth and the optical density and number of cfu by viable count were examined at hourly time points until stationary phase. This experiment did not detect any difference in the rate of bacterial cell division under these conditions (data not shown).

2.4.3 Strain NMB-CDC attaches to and invades human epithelial cells at significantly higher rates than strain MC58.

Since strain NMB-CDC and MC58 represent different genetic lineages and are different in bacterial cell surface area, we tested whether the strains interacted differently with host cells. The Detroit 562 pharyngeal carcinoma epithelial cell line and the transformed human bronchial epithelial cell line, 16HBE14\(\sigma\)- were challenged with suspensions of
Figure 2.2. AMNIS ImageStream™ can be used to identify single diplococci in an NMB-CDC sample under flow.

Panel A: Bivariate plot of the bright field and scatter channels for the particles in focus (R1 focus).

Panel B: Image gallery of hand tagged cells for analysis from the R2 region with bright field images displayed on the left and corresponding SYTOX Green image on the right.

Panel C: Image gallery of particles in the R3 region with brightfield images displayed on the left and corresponding SYTOX Green image on the right. The example given here is for strain NMB, but the same analysis was performed for preparations of strain MC58 (not shown).
wild-type NMB-CDC and MC58 at a multiplicity of infection of 100. On Detroit 562 epithelial cells, 38% of the inoculum of strain NMB-CDC versus 1.5% of the inoculum of strain MC58 associated with the cells, respectively (p <0.0001, Figure 2.4a). Similar results were obtained for attachment to 16HBE14σ- with 13.5% of the inoculum of strain NMB-CDC versus 0.18% of strain MC58 associated with host cells (p=0.0001, Figure 2B). Invasion was determined as the proportion of the adherent bacteria that were internalised at 6 hrs after the external bacteria were killed by treatment with gentamicin. Using this criterion, strain NMB-CDC invaded cells at significantly greater rates than MC58 in both cell lines. On Detroit 562 epithelial cells, 0.36% of attached NMB-CDC invaded and 0.01% of attached MC58 invaded (p=0.0004, Figure 2.4a). Invasion by strain MC58 into 16HBE14σ- cells could not be determined as it was below the limit of detection. Invasion by strain NMB-CDC of 16HBE14σ- was 1.9% of inoculum (Figure 2.4b). To examine whether the adherent and invaded bacteria in these assays represented phase variants from the original inoculum with regards to capsule
and LNT presentation, colony immunoblots were performed on the recovered isolates. In these assays the recovered colonies of NMB-CDC were found to have the same capsule and LOS phenotype as the inoculum, whereas a minor population of capsule and LOS phase variants were identified for strain MC58 as expected from previous studies (data not shown). Because invasion of MC58 into 16HBE14σ- cells could not be detected, the Detroit 562 cell line was used for the rest of the study.

![Diagram A](image1.png)

**Figure 2.4. Strain NMB-CDC associates with and invades epithelial cell lines at greater rates than strain MC58.**

The rates of association and invasion of strains NMB-CDC and MC58 into Detroit 562 epithelial cells (Panel A) and human bronchial epithelial cells 16HBE14σ- (Panel B) were assessed. Attachment as percent of the inoculum (black bars read off the left y-axis) and invasion as the percentage of the associated bacteria (white bars read off the right y-axis) is shown. The average rate (+/-SEM) from three biological repeats in triplicate following 6 hrs co-incubation is shown. *: p<0.001 determined by Mann-Whitney t-test compared to strain NMB-CDC, ND: Not Detected, viable counts were below the limit of detection.
2.4.4 Type IV pili, regardless of class and glycosylation phenotype, are necessary for initial attachment by both strains to epithelial cells.

As strain NMB-CDC associated with Detroit 562 cells better than strain MC58, the role of the type IV pili which are of different classes was tested. Type IV pilin mutants of both NMB-CDC and MC58 (JKD5122 and CKNM417, Opa+, LNT+, Cap+, Pil-, Table 2.1) were constructed. These strains each adhered at a rate of 0.2% of the inoculum indicating that the type IV pilus was the critical determining factor for attachment of both strain NMB-CDC and MC58 to this cell line (p<0.0001 and p=0.004 respectively, Figure 2.5a). As the rate of attachment of strains NMB-CDC and MC58 decreased from 38% and 1.5% respectively for the wild-type strains to 0.2% for their respective isogenic pilin mutants, we considered whether the differences in pilin class conferred different adhesive capabilities to the two strains. To determine whether the class of the pilin was contributing to the differences in host cell association between isolates, CKNM397 (Opa+, Cap+, LNT+, Table 2.1), which expresses the class I pilin from strain MC58 in a strain NMB PilEII-negative background, was constructed. There was no significant difference in the rate of association with host cells between CKNM397 and the control strain CKNM394 (Opa+, Cap+, LNT+, Pil+, Table 2.1), in which iga was insertionally inactivated (Δiga::aadA) in the presence of PilEII+ (p=0.053, Figure 2.5b). Both isolates express glycosylated pili, and the glycan structure has recently been shown by Jennings et al. (36) to initiate invasion into host cells by gonococci. To examine the role of the pilin glycan on attachment and invasion of strain NMB-CDC and strain MC58, a glycosylation (GLY) negative mutant (ΔpglF) was created in each strain in a Opa+, Pil+, LNT+, Cap+ background. Following a 6 hr incubation with Detroit 562 epithelial cells, no differences in attachment were observed when GLY was
removed from the pili of either strain (Figure 2.5c). However, removal of the glycan resulted in a 4-fold (p=0.0004) and 2-fold (p=0.0078) decrease in invasion of these mutants compared to their respective parental strains NMB-CDC and MC58. Therefore, the pilin glycan did affect bacterial invasion, but to a similar extent in both isolates.

2.4.5 Surface expressed sialic acid inhibits bacterial association and invasion of cell lines by strain MC58 to a greater extent than in strain NMB-CDC.

Previous work has demonstrated that the expression of capsule and sialic acid (N-acetyl neuraminic acid, NANA) decorated LOS inhibits attachment and invasion of \textit{N. meningitidis} into epithelial and endothelial cells (29). Terminal sialylation of the LNT-bearing LOS also blocks Opa mediated invasion (99). To determine the relative contribution of these structures to bacterial attachment and invasion, mutants in which the capsule was removed without affecting LOS sialylation (\(\Delta\text{synD}, \text{CAP-}, \text{NANA+}\)) and mutants in which all surface sialic acid was removed (\(\text{CAP-}, \text{NANA-}\)), were constructed. All strains used in the following assays were Pil+, Opa+ and LNT+ by colony and western immunoblot (see Table 2.1).

The removal of capsule alone (\(\Delta\text{synD}\) mutants, \(\text{CAP-}, \text{NANA+}\)) increased the rate of attachment of strain NMB-CDC by 2-fold and strain MC58 by 15-fold (p<0.002) (Figure 2.6a). The removal of capsule alone did not have a statistically significant effect on the invasion of strain NMB-CDC (p>0.2) however, the rate of invasion of strain
Figure 2.5. The role of Type IV pili is conserved in bacterial association with host cells.

Panel A. The ability of strains lacking pili (ΔpilE) to attach to Detroit 562 cells was compared to the parental wild-type. Strains NMB-CDC and MC58 and their respective pilE mutant strains are shown in black and white, respectively. All strains were Opa+, Cap+ and LNT+. The average rate (+/-SEM) of association as a percentage of the inoculum of three biological repeats in triplicate as determined by viable counts following 6 hr co-incubation is shown.

Panel B. The role of the class of pilin in association with host cells. Strain NMB-CDC was modified to express class I pilin from the iga locus in the absence of an intact pilEII locus (CKNM397). All strains were Opa+, Cap+ and LNT+. The ability of CKNM397 to associate with Detroit 562 cells was compared to parental wild-type strain NMB-CDC and CKNM394 containing an inactivated iga locus. The rate of association as a percentage of the inoculum of three biological repeats in triplicate was determined by viable counts following 1 hr co-incubation. The relative rate of association of each strain was normalised to strain NMB-CDC (value of 1) and is plotted as the average fold change (+/-SEM). *: p<0.005 determined by Mann-Whitney t-test.

Panel C. The expression of a pilin glycan containing either GAS (strain NMB-CDC) or DATD (strain MC58) retains the same role in bacterial invasion for both strains of meningococci. The biosynthesis of the pilin glycan was interrupted by insertional inactivation of pgIF. All strains were Pil+, Opa+, Cap+ and LNT+. The rate of association was determined by viable counts following 6 hr co-incubation. The relative rate of association of each strain was normalised to parental wild-type (value of 1) and is plotted as the average fold change (+/-SEM). Rates of attachment are shown in black bars (left y-axis) and invasion rates are shown in white bars (right y-axis). *: p<0.02 determined by Mann-Whitney t-test.

MC58 increased 35-fold (p<0.0001) (Figure 2.6b). Eighty percent of the strain NMBΔsynD inoculum was associated with epithelial cells in this assay. The removal of LOS sialylation alone (Δlst mutants, CAP+, NANA-) resulted in an increase in the rate of attachment (3.3-fold, p<0.0001) but not of invasion for strain NMB-CDC. In contrast, the rate of invasion increased slightly (1.7-fold, p=0.034) but in the absence of a detectable change in attachment in the strain MC58 background. The removal of LOS sialylation in the absence of capsule (CAP-, NANA-) resulted in no further gains in the rate of attachment of strain NMB-CDC to Detroit cells. However, the rate of invasion of NMB-CDC without sialic acid increased by 2.1-fold when compared to absence of
capsule alone (p=0.028) and 6-fold when compared to wild-type NMB-CDC (p<0.0001). In contrast, the removal of LOS sialylation in the absence of capsule (CAP-, NANA-) in strain MC58 background conferred a further 3.5-fold increase in attachment (p<0.0001) but no change in invasion (p=0.49) which is a 52-fold increase in attachment and 70-fold increase in invasion compared to wild-type MC58 (p<0.005).
Importantly, there was no significant difference in the attachment of MC58ΔsynB compared to M7, indicating parity between the strains was achievable when sialic acid biosynthesis was ablated, despite the multitude of variable protein alleles between the two strains. The difference in the rate of invasion of strains NMB-CDC and MC58 decreased from 30-fold in the presence of sialic acid to only 2.6-fold in the absence of capsule indicating the surface sialic acid had a major influence on this process but was not the only contributing factor to the invasion rate.

2.4.6 Strain NMB-CDC and strain MC58 express the same amount of surface capsule.

As the capsule of strain MC58 inhibited attachment and invasion to a greater extent than the capsule of NMB, we examined capsule expression in the two isolates. In serogroup B meningococci, capsule is synthesised via the synABCD locus which is divergently transcribed from an intergenic promoter region with the capsule transport locus of ctrABCD (27) and the entire region is termed the cps locus. The divergent promoter region in the cps locus is identical between strain NMB-CDC and MC58, and no IS1301 element was present (data not shown). We have previously shown that the two-component response regulator MisR represses the cps locus and de-repression by mutation of this regulator also results in hyper-encapsulation and serum resistance in strain NMB-CDC (96). To test whether MisR (which was identical in the two isolates) regulated the cps locus through a direct interaction with the intergenic promoter region, an EMSA was conducted (Figure 2.7a). MisR interacted with the cps intergenic probe containing both the synA and ctrA promoters in a dose-dependent manner and phosphorylated MisR generated by acetyl phosphate incubation showed a higher
affinity (compare lanes 2 and 5, Figure 2.7a). Further, the competition EMSA confirmed that this interaction was specific (Figure 2.7b). The transcription of both synA and ctrA increased in the misR mutants of both strains compared to the parental isolates as determined by qRT-PCR (Table 2.4). There were no significant differences in the levels of transcription of either synA or ctrA between the paired parental isolates or the paired ΔmisR mutants. The expression of lgtG and dsbD both of which are regulated by MisR (50, 93) and of lst, which is not part of the MisR regulon, were also consistent between the two isolates.

Table 2.4. qRT-PCR of gene expression in strain NMB-CDC and MC58 and their respective ΔmisR mutants.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Gene</th>
<th>wild-type²</th>
<th>ΔmisR²</th>
<th>Fold changeᵇ</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average/lpxA</td>
<td>st. dev.</td>
<td>Average/lpxA</td>
<td>st. dev.</td>
<td></td>
</tr>
<tr>
<td>NMB</td>
<td>lst</td>
<td>1.79</td>
<td>0.25</td>
<td>2.33</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>ctrA</td>
<td>4.48</td>
<td>0.84</td>
<td>17.30</td>
<td>5.37</td>
</tr>
<tr>
<td></td>
<td>synA</td>
<td>4.27</td>
<td>2.35</td>
<td>8.17</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>lgtG</td>
<td>0.05</td>
<td>0.03</td>
<td>0.68</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>dsbD</td>
<td>10.94</td>
<td>2.04</td>
<td>1.41</td>
<td>0.25</td>
</tr>
<tr>
<td>MC58</td>
<td>lst</td>
<td>1.73</td>
<td>0.76</td>
<td>3.10</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>ctrA</td>
<td>5.49</td>
<td>2.76</td>
<td>20.86</td>
<td>7.14</td>
</tr>
<tr>
<td></td>
<td>synA</td>
<td>5.16</td>
<td>2.31</td>
<td>11.47</td>
<td>4.13</td>
</tr>
<tr>
<td></td>
<td>lgtG</td>
<td>0.10</td>
<td>0.06</td>
<td>1.09</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>dsbD</td>
<td>15.10</td>
<td>7.09</td>
<td>2.29</td>
<td>1.04</td>
</tr>
</tbody>
</table>

²: The average relative level of gene expression of each gene listed under “Gene” normalised to lpxA, determined from three biological repeats conducted in triplicate, with standard deviation (st. dev.) indicated.

ᵇ: The fold change in relative gene expression between the wild-type and ΔmisR strains.

We examined the length of the capsule polymers of NMB-CDC and MC58 by separation of purified capsule by DOC-PAGE visualised by silver-staining. With the resolution of this analysis, it appears that both strains express a capsule polymer of similar length and modality (Figure 2.8a) (28).
The capsule of NMB-CDC and MC58 was examined by flow cytometry following labeling with anti-capsule antibody and this revealed that the geometric mean fluorescence (GMF) of strains NMB-CDC and MC58 were 24.03 and 20.76, respectively. Based on the transcriptional data and the quantitation of capsule polymer by FACs, it appears both strains express similar amounts of surface capsule.

![Image](image_url)

**Figure 2.7.** MisR directly binds to the synA-ctrA promoter region.

**Panel A.** MisR interacts with the *cps* intergenic probe containing both the synA and *ctrA* promoters. Phoshorylated MisR (MisR\(^{\text{P}}\), Lanes 2-4) and MisR (lanes 5-7) binds the probe in a dose-dependent manner. Lane 1 contains probe alone, lanes 2 and 5 contain 68 pmol of protein; lanes 3 and 6 contain 136 pmol of protein; lanes 4 and 7 contain 204 pmol of protein.

**Panel B.** Competition EMSA. The mobility shift of the labeled probe by MisR\(^{\text{P}}\) (136 pmol, lane 2) relative to labeled probe without protein (lane 1) was competed away by unlabelled probe (1 µg in lane 3 and 2 µg in lane 4). This interaction between MisR\(^{\text{P}}\) and the labeled probe was specific as the complex remained intact and could not be competed away when cold, unlabeled non-specific DNA was added (1 µg in lane 5 and 2 µg in lane 6).
2.4.7 The surface of strain NMB-CDC is more hydrophobic than strain MC58.

Previous studies have shown that hydrophobicity of bacterial surfaces is a function of capsule expression (31, 47, 95) and we reasoned that if there was a difference in capsule distribution on the bacterial surface of these two isolates this would be detected using hydrophobicity interaction chromatography (HIC, (95)) and microbial adhesion to solvents (MATS) assays. When using HIC, encapsulated isolates are more hydrophilic than non-encapsulated isogenic mutants, and thus the encapsulated bacteria will pass through the column in the effluent (95). The non-encapsulated control strains M7 and MC58ΔsynB were almost completely retained on the column and were not significantly different (p=0.06), indicating that the capacity of the column to bind the bacterial cells had not been saturated (Figure 2.8b). For strain NMB-CDC the OD of the effluent was 8-fold higher than that for M7 (p<0.0001). In comparison, the effluent of MC58 was 85-fold higher than that of MC58ΔsynB (p<0.0001). Since the effect of the loss of surface sialic acid from strain NMB-CDC was 10-fold less than that observed for strain MC58 using these columns, the surface of strain NMB-CDC was likely to be more hydrophobic than strain MC58. To further confirm that the hydrophobicity interaction columns could distinguish between hydrophilic surfaces with more or less capsule polymer, the effect of capsule over-expression was examined. A ΔmisR mutant of strain NMB-CDC was shown previously to express ~40% more capsule polymers than the parent strain, which led to an increase in resistance to normal human serum thus indicating that the capsule was on the bacterial cell surface (96). The OD of the effluent for the NMBΔmisR and MC58ΔmisR mutants was 1.7-fold higher than their respective
parents (p<0.001), consistent with the increased surface expression of capsule in the mutant strains and hence a decrease in surface hydrophobicity.

Since HIC was considered to be a qualitative assay, MATS was used to provide a quantitative measure of hydrophobicity for these isolates. In MATS, the level of adhesion of the bacteria to non-polar n-alkane, hexadecane, is measured and the percentage of the adhesion to the solvent (% hydrophobicity) is determined (75). In this
assay, the non-encapsulated strains NMB ΔsynA and MC58ΔsynB were 2.2-fold (p=0.004) and 4.1-fold (p=0.0002), than their respective parental encapsulated wild-type strains (Table 2.5). Although the hyper-encapsulated NMBΔmisR mutant was 2.16 fold less hydrophobic than the parental wild-type (p=0.005), no difference between strain MC58 and MC58ΔmisR was observed (p>0.1). No further decrease in hydrophobicity of the hyper-encapsulated MC58ΔmisR could be detected in buffers of increasing ionic strength (data not shown, (75)), suggesting that strain MC58 was already very hydrophilic and no further reductions in hydrophobicity could be measured in this assay. Overall, both the HIC and MATS assays correlated closely with each other and indicated that the contribution of capsule to surface hydrophobicity was less for strain NMB-CDC than that of strain MC58.

Table 2.5 Adhesion of bacterial strains to hexadecane in the MATS assay.

<table>
<thead>
<tr>
<th>Strain</th>
<th>% hydrophobicity*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMB</td>
<td>27.7±3.7</td>
</tr>
<tr>
<td>NMBΔsynA</td>
<td>62±9.5</td>
</tr>
<tr>
<td>NMBΔmisR</td>
<td>12.8±2.9</td>
</tr>
<tr>
<td>MC58</td>
<td>10.8±3.0</td>
</tr>
<tr>
<td>MC58ΔsynB</td>
<td>41.5±7.0</td>
</tr>
<tr>
<td>MC58ΔmisR</td>
<td>6.8±3.5</td>
</tr>
</tbody>
</table>

* The mean of three biological repeats each containing three technical repeats.

2.5 Discussion

Models investigating key steps in the pathogenesis of meningococcal attachment and invasion of human cells, have shown the importance of phase variable expression of pili, capsule, LOS, Opa and Opc (24, 29, 63, 64, 66, 68, 73, 78, 90, 98, 100, 102-106).
We hypothesised that since strain NMB-CDC has a significantly lower phase variation frequency and was unable to phase vary the LNT-bearing LOS to relieve inhibition of Opa adhesins, attachment and invasion of human cells would be comparatively less than strain MC58. Surprisingly, an encapsulated, LNT-bearing LOS, class II piliated and Opa expressing strain NMB-CDC was much more efficient than the encapsulated, LNT-bearing, class I piliated and Opa expressing strain MC58 in attachment and invasion on two model cell lines (Figure 2.4). The association and invasion data for strain MC58 derived from this model cell line is comparable with the outcomes described for this strain in previous work with other epithelial and endothelial cell lines (100, 102-106).

The current hypothesis generated from an examination of the genomes of non-disease and disease causing isolates of meningococci, predicts that differences in the combination of alleles of virulence factors, including type IV pili, capsule, and LOS structure, would mediate different interactions with the host cell. To test this hypothesis, the effect of pilin class and glycosylation, sialylation of LNT-bearing LOS and capsule expression by both strains was examined in greater detail. ST-8 cc strain NMB-CDC expresses a typical class II type IV pilus in comparison to ST-32 cc strain MC58 which expresses a class I variant (44). In both isolates the type IV pili had the same predominant effect on bacterial attachment to host cells (Figure 2.5). Exchanging the expression of class II pili to class I pili in strain NMB-CDC did not affect the performance of strain NMB-CDC to attach to host cells, indicating that the class of the pilin had no discriminatory effect in this process. Meningococcal type IV pili are variably glycosylated with a tri-saccharide (69). The glycan generally consists of two galactose residues linked to $\text{D}_{2,4}\text{-}\text{Acetamido-2,4,6\text{-}\text{Tri-deoxy-hexose}}$ (DATDH) (18), although a variant glycan containing $\text{Glyceramido\text{-Acetamido-2,4,6\text{-}\text{Tri-deoxy-hexose}}}$ (GATDH)
is produced by strains containing a variant synthesis gene, *pglB2* (43, 84). A similar glycan is present on class I type IV pili of gonococci and has been shown by Jennings et al. (36) to interact with complement receptor 3 to initiate invasion into *pex* cells. While both strain NMB-CDC and MC58 produce glycosylated pili, strain NMB-CDC expresses the glycan containing GATDH and strain MC58 expresses a glycan containing DATDH. Nevertheless, the removal of the glycan by inactivation of the synthesis gene, *pglF*, in both strains resulted in a decrease in rate of invasion without affecting bacterial association to host cells (17, 57). Lastly, the type IV pili contain antigenically variable PilC tip adhesins (60) which previous studies have implicated in tissue tropism. However since the differences in the ability of strain NMB-CDC and MC58 to attach to Detroit 562 cells reached parity when the amount of surface sialic acid was modulated on the bacteria, it was concluded that no significant differences in attachment could be attributed to the PilC alleles expressed by strain NMB-CDC and MC58 under the experimental conditions used in this study. Therefore, the enhanced performance of strain NMB-CDC to invade host cells when compared to strain MC58 was not attributable to the variant type IV pili expressed by these two isolates.

Previous work has clearly identified the capsule and terminal sialic acid of LNT-bearing LOS as two components which interfere with Opa function (99, 105). In strain MC58, the inhibitory effect of sialylated LNT-bearing LOS is relieved by the phase variable removal of the LNT moiety (105). However, since strain NMB-CDC cannot phase vary the expression of LNT as it lacks a phase variable *lgtA*, the role of LOS sialylation was examined in the presence of LNT-bearing LOS with and without capsule. In agreement with previous studies with strain MC58, capsule was the major determinant of inhibition of bacterial attachment and invasion of host cells while removal of sialylation
of LNT-bearing LOS resulted in no change in bacterial association or invasion unless capsule was also absent (105) (Figure 2.6). In contrast, in strain NMB-CDC, removal of sialylation of LNT-bearing LOS increased rates of attachment by 2-fold indicating that LOS itself could affect invasion even in the presence of capsule in this strain. Removal of capsule from strain NMB-CDC expressing LNT-bearing LOS resulted in only a 3-fold increase in bacterial association suggesting that both LOS sialylation and capsule had equivalent inhibitory roles, rather than the predominant effect of capsule alone that was seen in strain MC58 (Figure 2.6). Therefore, these data indicated that the serogroup B capsule of strain NMB-CDC was much less inhibitory to bacterial attachment and invasion than observed for strain MC58.

A number of theories were tested regarding the potential for variation in capsule presentation on the surface of these two isolates. Since the MisR-dependent regulation of capsule expression, and the measurement of surface capsule by flow cytometry, in addition to the length and modality of the polymers was similar for both strains (Table 2.4). Since meningococcal capsule is a glycolipid anchored to the surface of the bacterial outer membrane, it has a distribution profile per unit surface area. Therefore, as strain NMB-CDC had a larger surface area than strain MC58, the distribution of the anchored glycolipid polymer on the bacterial cell surface of NMB-CDC was hypothesised to be sparser than that of strain MC58. Capsule polysaccharide is a major determinant of surface hydrophobicity in numerous bacteria (31, 46, 47) due to the ability to retain polar water molecules close to the bacterial surface. However, surface hydrophobicity is a multi-factorial phenotype, resulting from the complex interplay of polar and apolar outer surface components, including major surface proteins in some bacterial species (55) and O-antigen of lipopolysaccharide (8). The capsule polymer of
strain NMB-CDC contributed to a smaller change in surface hydrophobicity (8-fold by HIC, and 2.2-fold by MATS) than in strain MC58 (85-fold by HIC and 4.4-fold by MATS), indicating that encapsulated strain NMB-CDC was more hydrophobic than strain MC58. Our model suggests that as the capsule polymers become sparser over a greater surface area, the overall surface of the isolate becomes more hydrophobic. Future work will address whether other factors contribute to surface hydrophobicity in the presence of capsule in meningococci.

Lastly, bacterial cell size has been shown to be controlled by a number of factors including nutrient availability (108) and cell cycle control by the initiation of replication protein, DnaA (33). Weart et al. (108) identified a metabolic sensor, UgtP, which localised at the cell division site in a nutrient dependent manner, inhibiting cell division and thus controlling cell size in the rod-shaped organism Bacillus sp. However, a search of the meningococcal genome did not find a homologue of this protein, suggesting other factors are needed for this phenotype in cocci. It is important to note that the Family Neisseriaceae contains both pathogenic and commensal bacterial species of variable cell dimensions encompassing diplococci, coccobacilli and rods (3). The difference in bacterial cell size between these two isolates suggests that their metabolism is different. Multiple studies of the conserved and accessory genomes of meningococcal isolates (37) and variation within the housekeeping genes used for multi-locus sequence typing (9), has recently led to the proposal that small changes in metabolism may lead to differences in virulence and transmission of isolates.

In conclusion, the two classes of type IV pili and pilin glycosylation with glycans containing either GATDH or DATDH, were shown to have conserved roles during
attachment and invasion into host cells by the two exemplar meningococcal isolates in this study. The surface of encapsulated strain MC58 was shown to be more hydrophilic and adhered less to host cells than strain NMB-CDC. Improved levels of adherence of strain NMB-CDC to cell lines correlated with increased bacterial cell surface and surface hydrophobicity. At this stage, future work will be required to understand the genetic and metabolic basis for the variability of meningococcal bacterial cell size and surface hydrophobicity, and to determine whether these phenotypes are shared by other isolates from the same ST or clonal complex.

### 2.6 Acknowledgments

This publication made use of the *Neisseria meningitidis* Opa sequence database (http://neisseria.org/nm/typing/opa/), developed by Dr Keith Jolley and Dr Martin Callaghan (University of Oxford). 16HBE14σ- cells and MAb 3F11 were generous gifts from Dr Mike Apicella (Iowa University). MAb 2-2-B was a generous gift from Dr Wendell Zollinger and Dr. Elizabeth Moran (Walter Reed Army Institute of Research, USA). MAb SM1 was a gift from Professor Mumtaz Virji (University of Bristol, UK). The polyclonal anti-Opa antibody was a gift from Dr. Hideyuki Takahashi (National Institute of Infectious Diseases, Japan). The authors acknowledge the facilities and scientific assistance of the Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy, Characterisation and Analysis, The University of Western Australia, a facility funded by the University, State and Commonwealth Governments.
2.7 References


27. Frosch, M., C. Weisgerber, and T. F. Meyer. 1989. Molecular characterization and expression in Escherichia coli of the gene complex


Chapter 3. Contact-dependent regulation of lipooligosaccharide inner core structure determines meningococcal invasion into host cells.
3.1 Abstract

*N. meningitidis* expresses lipooligosaccharide (LOS) consisting of an inner core of heptose residues (HepI and HepII) decorated with lacto-\(N\)-neotetraose (LNT) on the \(\alpha\)-chain, phosphoethanolamine (PEA) at positions \(O-3\) and \(O-6\) on HepII, glucose (Glc) at position \(O-3\) on HepII, and an \(O\)-acetyl group (OAc) at position \(O-3\) on the terminal \(N\)-acetyl glucosamine (GlcNAc) residue of the \(\gamma\)-chain. *N. meningitidis* strain NMB-CDC cannot phase vary the expression of LNT due to a phase locked LOS glycosyltransferase A (LgtA), but does possess the multiple genes required for the addition of \(O-3\) Glc, \(O-3\) OAc and \(O-6\) PEA groups to the LOS inner core. In the presence of LOS bearing LNT, the decoration of the LOS inner core with \(O-6\) PEA and \(O-3\) OAc or \(O-3\) Glc and \(O-3\) OAc or \(O-3\) OAc alone promoted meningococcal invasion into the host cells (18-fold, 4-fold and 9-fold increase compared to parental wild-type, respectively). The ability of the LOS inner core decorations to modulate the rate of meningococcal invasion into host cells was not inhibited by the presence of capsule unless the LOS structure contained an \(O-3\) PEA residue attached to HepII. The expression of the LOS inner core structures with LNT promoted invasion into host cells in the absence of the Opa adhesins, suggesting that the LOS structure was acting as an independent ligand for binding a host cell receptor. Lastly, the decoration of the LOS inner core structure with \(O-6\) PEA was up-regulated and terminal sialylation of LNT and capsule expression were down regulated in the presence of host cells. In summary, LOS inner core residues promote meningococcal invasion in the absence of the ability to remove LNT by phase variation in strain NMB-CDC.
3.2 Introduction

*Neisseria meningitidis* is a cause of epidemic rapidly fatal fulminant septicaemia in otherwise healthy individuals (3, 17). Although this organism is most often found asymptomatically colonising the nasopharynx of approximately 10% of the adult population, when it gains entry to the host circulatory system and evades the immune response, the vast amount of endotoxin released from the bacterial cell wall rapidly leads to septic shock (55). The LOS structures expressed by *N. meningitidis* are categorised into 12 immunotypes based on antibody reactivity (45), however only seven (L1, L2, L3, L4, L7, L8, L9) are commonly expressed by invasive isolates (45). Immunotype is dependent upon the variable composition of the α-chain and heterogeneous additions to the conserved inner core. The LOS of *N. meningitidis* is anchored to the outer membrane by lipid A which has a D-glucosamine disaccharide headgroup (30). A core oligosaccharide is attached to lipid A and is comprised of two residues of 2-keto D-manno octulosonic acid, two residues of L-glycero-D-manno-heptose (HepI and HepII), a terminal N-acetylglucosamine (GlcNAc) on HepII (γ-chain) and β1,4 glucose (β1,4 Glc) on HepI (reviewed in (29), Figure 3.1). The α-chain attached to HepI is commonly composed of lacto-N-neotetradose (LNT) comprised of Galβ(1→4)GlcNAc β(1→3)Gal β(1→4)Glc which is sequentially assembled by the addition of each residue by the LOS glycosyl transferases, LgtE, LgtA and LgtB respectively (21). The variable inner core decorations consist of the O-acetylation (O-3 OAc) of GlcNAc by the LOS O-acetyl transferase (Lot), the addition of glucose (Glc) to O-3 of HepII by LOS glycosyl transferase G (LgtG), and the addition of phosphoethanolamine (PEA) to O-3 and O-6 of HepII by the LOS phosphoethanolamine transferases, Lpt3 and Lpt6, respectively (2, 28, 32, 59). The
LOS molecule is synthesised on the cytoplasmic face of the inner membrane as a unit, translocated to the periplasmic face of the inner membrane by MsbA and transported across the periplasm by the LptB–H complex (5). The addition of PEA to the LOS structure occurs in the periplasm as the LOS transits this compartment and is inserted into the outer leaflet of the outer membrane. As a result the addition of glucose to the O-3 position of HepII by LgtG in the cytoplasm precedes PEA addition to this site. In addition, no detectable levels of O-3 PEA additions to the LOS inner core could be detected in the absence of O-3 linked glucose when sialylated LNT was attached to HepI and the terminal GlcNAc of the γ-chain was O-acetylated in strain NMB-CDC (28).

Figure 3.1 The structure of lipooligosaccharide of *N. meningitidis*. The core structure (shown in green) is conserved between all known structures of naturally occurring meningococcal LOS. The LNT α-chain, of Galβ(1→4)GlcNAc β(1→3)Gal β(1→4)Glc is shown in pink. The transferases and the linkages they catalyse for each residue is indicated. Annotations are: α2,3 LOS sialyltransferase (Lst); β1,4 LOS galactosyltransferase B (LgtB); β1,3 LOS N-acetylglucosamine transferase (LgtA); β1,4 LOS galactosyltransferase (LgtE); β1,4 LOS glucosyltransferase (LgtF); α1,5 LOS heptosyltransferase C (RfaC); α1,7 LOS heptosyltransferase F (RfaF); α1,2 N-acetylglucosamine transferase (RfaK); O-3 LOS O-acetylttransferase 3 (Lot3); α1,3 LOS glucosyltransferase (LgtG), O-3 LOS phosphoethanolamine transferase (Lpt3) and O-6 LOS phosphoethanolamine transferase (Lpt6).
Invasive meningococcal isolates are encapsulated and express LOS bearing LNT which is terminated with sialic acid by the LOS sialyltransferase, Lst (46). Binding studies suggest that sialylated LNT can interact with the Opa and Opc proteins hence inhibiting cognate interactions with host cell proteins such as CEACAMs, integrins or matrix components, thus inhibiting invasion of the bacteria into host cells (4, 8, 35, 57). Invasive isolates overcome this inhibitory effect by phasing “off” the expression of capsule and the LNT chain on the LOS (7, 24). lgtA which encodes the \( \beta 1,3 \) GlcNAc LOS glycosyltransferase is phase variable such that slipped-strand mispairing during bacterial replication of a homopolymeric tract within the gene will result in translational frameshifts which lead to the loss of protein expression (22). Similarly the polysialic acid sialyltransferase, SynD, is also phase variable and hence capsule expression can be lost via the same mechanism (13). In this manner, clones which have an adhesive phenotype emerge (Capsule and LNT devoid) which allows invasion of host cells. Similarly, clonal expansion of the bacteria within the host cell will result in phase variants in which capsule and LNT expression is restored and these clones are able to proceed to cause a systemic infection due to an ability to resist serum killing (57).

In the model of gonococcal invasion into host cells, the LOS structure is involved in triggering entry into female cervical cells, male urethral cells and human macrophages via complement receptor 3 (CR3), asialoglycoprotein R (ASPG-R), and macrophage galectin-like lectin (MGL), respectively (14, 15, 56). ASPG-R and MGL are lectins and along with galectin-3, bind the LNT \( \alpha \)-chain (23). In the model for meningococcal uptake into host cells, LOS has been considered to be less involved in this process due to the presence of the polysialic acid capsule which masks LOS and inhibits close interactions with host cells (58). However, we recently reported that strain NMB-CDC
had less dense capsule than other meningococcal isolates (See Chapter 2), and this has led to the hypothesis that LOS may act as a ligand for invasion into host cells by this isolate. Herein we demonstrate that meningococcal LOS in the presence of capsule in strain NMB-CDC can trigger invasion into host cells via a mechanism that is independent of Opa. We also show that substitution of the LOS inner core with different residues has an important modulatory role in this process and that some of the genes involved in modifying the LOS structure are regulated in response to contact with epithelial cells.

3.3 Methods

3.3.1 Bacterial strains and growth conditions

*Neisseria meningitidis* strains NMB-CDC (ST-8, serogroup B, immunotype L2,4,) and MC58 (ST-32, serogroup B, immunotype L3,7,9,) were cultured under aerobic conditions with 5% CO2 at 37°C on GC agar (GCA) or GC broth (GCB) (Oxoid) supplemented with 0.4% glucose, 0.01% glutamine, 0.2 mg of cocarboxylase per litre, and 5 mg of Fe(NO3)3 per litre. The wild-type strains and constructed mutants were piliated and expressed Opa, except where stated otherwise. The wild-type strains and constructed mutants used in this study are shown in Table 3.1. Antibiotic selection for meningococcal mutants was performed on GCA containing 100 µg/ml of kanamycin (sulfate salt), 60 µg/ml of spectinomycin, 5 µg/ml of tetracycline or 2 µg/ml of erythromycin (Sigma). *Escherichia coli* DH5α was used as a host for all DNA manipulations. It was routinely grown on Luria-Bertani broth (LBB) and agar (LBA, Oxoid) which, where appropriate, were supplemented with antibiotics at the following
concentrations: ampicillin at 100 µg/ml, spectinomycin at 50 µg/ml, kanamycin at 50 µg/ml, erythromycin at 300 µg/ml, tetracycline at 12.5 µg/ml and chloramphenicol at 30 µg/ml (Sigma).

Table 3.1. Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain Name</th>
<th>Genotype</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. meningitidis</td>
<td>NMBΔsynA::Tn916</td>
<td>(48)</td>
</tr>
<tr>
<td>MC58</td>
<td></td>
<td>(51)</td>
</tr>
<tr>
<td>M7</td>
<td>NMBopaD+</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM323</td>
<td>NMBopaD+ΔopaD::tetM</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM331</td>
<td>NMBΔgtG::aphA-3</td>
<td>(52)</td>
</tr>
<tr>
<td>JKD5161</td>
<td>NMBΔlpt6::ermC</td>
<td>(27)</td>
</tr>
<tr>
<td>JKD5172</td>
<td>NMBΔlot::tetM</td>
<td>(28)</td>
</tr>
<tr>
<td>JKD5170</td>
<td>NMBΔgtG::aphA-3Δlpt6::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>JKD5173</td>
<td>NMBΔgtG::aphA-3Δlot::tetM</td>
<td>(28)</td>
</tr>
<tr>
<td>CKNM333</td>
<td>NMBΔgtG::aphA-3Δlpt6::ermCΔlot::tetM</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM415</td>
<td>MC58ΔgtG::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM326</td>
<td>M7ΔgtG::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM328</td>
<td>M7Δlpt6::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM327</td>
<td>M7ΔgtG::aphA-3Δlpt6::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>CMK2</td>
<td>NMBΔgtF::aphA-3</td>
<td>(26)</td>
</tr>
<tr>
<td>CKNM422</td>
<td>MC58ΔgtF::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM349</td>
<td>NMBΔsynB::aadA</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM420</td>
<td>MC58ΔsynB::aadA</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>CKNM408</td>
<td>NMBΔsynB::aadAΔlgfF::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM428</td>
<td>MC58ΔsynB::aadAΔlgfF::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM359</td>
<td>NMBopaD::tetMΔgtG::aphA-3</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM360</td>
<td>NMBopaD::tetMΔlpt6::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM362</td>
<td>NMBopaD::tetMΔgtG::aphA-3Δlpt6::ermC</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM329</td>
<td>M7Δlot::aadA</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM330</td>
<td>M7ΔgtG::aphA-3Δlot::aadA</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM334</td>
<td>M7ΔgtG::aphA-3Δlpt6::ermCΔlot::aadA</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM449</td>
<td>NMB with promoterless::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM453</td>
<td>NMB with misR(−348/+104)::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM469</td>
<td>NMB with Pprep::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM461</td>
<td>NMB with Prep::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM497</td>
<td>NMB with Prep::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM485</td>
<td>NMB with Prep::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM489</td>
<td>NMB with Prep::lacZ fusion in the chromosomal 988 locus</td>
<td>This study</td>
</tr>
</tbody>
</table>
3.3.2 Attachment and invasion assays

The immortalised Detroit 562 (human pharyngeal carcinoma cells, ATCC CCL-138) cell line was grown to confluence in Minimal Essential Media (MEM+Earles salts) supplemented with 10% FCS, 2 mM L-glutamine, 1 mM sodium pyruvate and 1x non-essential amino acids (38) as per ATCC recommendations. At confluence the cells were
lifted with trypsin/EDTA (0.05%) and resuspended in cell media. 1x10⁵ cells were seeded into each of the wells of a 24-well plate (Nunc). The bacterial attachment and invasion assays were performed as previously described (See Chapter 2). This data was represented as fold changes compared to the parent, and the parent normalised to 1. All data from all experiments is included in attachment and invasion graphs. A Mann-Whitney T-test was used to determine statistical significance. All strains were sensitive to gentamycin and resistant to 1% saponin.

3.3.3 Construction of meningococcal mutants and LacZ reporter strains

To construct an isolate of strain NMB-CDC that does not synthesise sialic acid, the sialic acid synthetase, synB, was inactivated using a synB::aadA cassette. Strain NMB-CDC was transformed with pCMK723 and the transformants were identified by resistance to spectinomycin then confirmed by PCR using primer pair KAP370 (5’-TCCTGAAACGTGGAATGTTTCTGC-3’) and KAP373 (5’-ATAGAGACATCTGCATTGCCTGG-3’), resulting in CKNM349. Mutants of strain NMB-CDC that were deficient in both sialic acid synthesis and LNT addition of LOS were constructed by transforming CKNM420 and CKNM349 with pCK18 via chemical transformation and transformants identified by resistance to kanamycin and confirmed by tricine SDS-PAGE of LOS, resulting in CKNM428 and CKNM408 respectively.

To assess the role of LOS of NMB-CDC in invasion in the absence of opacity proteins (Opa), a clone of NMB-CDC in which only one Opa protein was phased “on” was identified by western immunoblot analysis, isolated and stocked as CKNM323. The
phase variable tracts of each of the *opa* genes were sequenced to identify which of the loci was expressed. *opaD* was found to be the only *opa* gene phased on. The 5’ flanking region of *opaD* was amplified with KAP232 (5’-TAACGGGTAGGGTATCGGTCGG-3’) and KAP329 (5’-GATATCGAGCTCTGATTCCCAGATACCGATGCC-3’), while the 3’ flanking region was amplified with KAP233 (5’-TGGAACCCAAATCGACGGAGGC-3’) and KAP330 (5’-GAGCTCGATATCGATTTTTGACTCCGTATTCCG-3’). The two PCR products were used as the template for spliced over-lapping extension PCR (SOE-PCR). The extended overlaps of KAP329 and KAP330 enabled the introduction of *EcoRV* and *SacI* sites. The second round PCR amplification was performed with the initial products as templates and the nested primer pair of KAP327 (5’-CGCCGTAGAGGAAATGATGCC-3’) and KAP328 (5’-CCAAACGCACCGCATCATCGG-3’). The T4 DNA polymerase (NEB) polished PCR amplicon was cloned into the *HincII* site of pHSG576 resulting in pCMK722. The *tetM* cassette from pUC18tetM (52) was cloned into *EcoRV* site of pCMK722 to create pCMK725. To inactivate *opaD* in strain NMB-CDC, pCMK725 was transformed into CKNM323 and the transformants were selected by tetracycline resistance. The correct transformants were identified by amplification of the insertion site with primer pair KAP232 and KAP233 and the lack of Opa expression for this strain confirmed by western immunoblot, thus resulting in CKNM351.

To construct strains in which *lgtG* was insertionally inactivated, M7, CKNM358 and CKNM360 were transformed with pCK49 (52) and transformants identified by resistance to kanamycin and confirmed by PCR using primer pair KAP51 (5’-GCAACCAACAACTTCAACACG-3’) and KAP52 (5’-
CGTTCATGACCTCTGTACAT-3'), resulting in CKNM415, CKNM326, CKNM359 and CKNM362 respectively.

To construct strains in which lpt6 was insertionally inactivated, CKNM331, CKNM326, M7 and CKNM358 were by transformation with pJKD2621 (27) utilising natural competency and transformants were identified by resistance to erythromycin and confirmed by PCR of the insertion site using primer pair DAP74 (5’-CTTCGTTCTGGTTTGTGGTG-3’) and DAP75 (5’-GCAGATAACGGTGCGAAACTTCC-3’), resulting in JKD5170, CKNM327, CKNM328 and CKNM360, respectively.

lot was inactivated in strain JKD5170 by transformation with pJKD2710 (28) utilising natural competency and transformants identified by resistance to tetracycline and confirmed by PCR of the insertion site using primer pair DAP74 and DAP75 resulting in CKNM333. The gene lot was inactivated in strains CKNM326 and CKNM327 by transformation with pCMK705 utilising natural competency and transformants identified by resistance to erythromycin and confirmed by PCR using primer pair DAP371 (5’-CGATTTTGTCGGGAAAGAAACCG-3’) and DAP372 (5’-GAAGCCAAAGCCAAATTGCTTGAGC-3’), resulting in CKNM330 and CKNM334.

To assess the role of the LOS inner core in attachment and invasion of host cells in the absence of sialic acid, inner core mutations were made in M7. As M7 is resistant to tetracycline, the antibiotic cassette had to be changed to aadA conferring resistance to spectinomycin. The aadA fragment was liberated from pHP45Ω by digestion with EcoRV and cloned into JKD2709 was digested with HincII and the resulting plasmid
identified by resistance to spectinomycin, restriction mapping and PCR with primers RP and UP. The resulting plasmid was called pCMK705.

To construct promoter::\textit{lacZ} reporter constructs for expression in \textit{N. meningitidis} strain NMB-CDC, the promoter region for each gene was amplified with primers tagged with 5’ \textit{Eco}RI (5’ \textit{Sma}I for the synA/ctrA intergenic region) and 3’ \textit{Bam}HI sequences, polished with T4 DNA polymerase and cloned into the \textit{Sma}I site of pHSG576. The primers used to amplify: $P_{\text{synA}}$ were KAP281 (5’-GACCCGGGCTGATGAAGGAATCGAGACG-3’) and KAP282 (5’-GAGGATCCGAAGTCGCTCTGGTACC-3’); $P_{\text{ctrA}}$ were KAP283 (5’-GACCCGGGCTGATGAAGGAATCGAGACG-3’) and KAP284 (5’-GAGGATCCGAAGTCGCTCTGGTACC-3’); $P_{\text{lst}}$ were KAP273 (5’-GAGAATTCCGTCGGTATGGGTATAAAAAATGG-3’) and KAP274 (5’-GAGGATCCGTCGGTATGGGTATAAAAAATGG-3’); $P_{\text{lpt6}}$ were KAP285 (5’-GAGAATTCCGTCGGTATGGGTATAAAAAATGG-3’) and KAP286 (5’-GAGGATCCGTCGGTATGGGTATAAAAAATGG-3’); and $P_{\text{lot}}$ were KAP275 (5’-GAGAATTCCGTCGGTATGGGTATAAAAAATGG-3’) and KAP276 (5’-GAGGATCCGTCGGTATGGGTATAAAAAATGG-3’). The cloned promoter in the resulting plasmids called pCMK831 pCMK829, pCMK826, pCMK830 and pCMK821 respectively were sequenced with primer RP (5’-AGCGGATAACAATTTCACACAGGA-3’) and UP (5’-GTTTTCCCAAGTACGAC-3’) to ensure no errors had occurred. The promoter fragment was liberated from the pHSG576 backbone by digestion with \textit{Eco}RI or \textit{Sma}I and \textit{Bam}HI and directionally cloned into the \textit{lacZ}-reporter plasmid, pYT328, which allows recombination of the expression cassette into the NMB0428 and NMB0430 locus [54]. Insertion was
identified by PCR using primers UP and promoter specific primers, resulting in pCMK853, pCMK854, pCMK846, pCMK848 and pCMK855 respectively. pYT328, pYT322, pYT343, pYT354, pCMK853, pCMK854, pCMK846, pCMK848 and pCMK855 were transformed into NMB-CDC and transformants identified by erythromycin resistance, confirmed by PCR of the insertion site using primer pair KAP188 (5'-CCTATGGTGATGGGTGTTGC-3’) and KAP189 (5'-CGATTCTTTTCACAGCTTGCG-3’), and checked for β-galactosidase activity by observation of a blue precipitate following growth on GCA containing 20 μg/ml 5-bromo-4-chloro-indolyl-β-D-galactopyranoside. The LacZ reporter strains were named CKNM449 (P-), CKNM453 (PmisR), CKNM461 (Pmtr), CKNM469 (PlgtG), CKNM485 (PsynA), CKNM489 (PctrA), CKNM473 (Plst), CKNM497 (Ppat6), and CKNM477 (Ptot).

3.3.4 Transformation protocols

Strain NMB was transformed via natural transformation as described previously (20).

3.3.5 Determination of capsule, LOS, pilin and Opa phenotype

Colony immunoblotting was utilised to confirm the expression of capsule in all clones examined in this study (48). Briefly, strains were plated for single colonies on GC agar from frozen stock. The plates were incubated until colonies were observed. The colonies were transferred to nitrocellulose, the plates were re-incubated to allow the colonies to grow again, and the nitrocellulose discs were air dried, then fixed by UV exposure. Colonies were washed off the nitrocellulose with TBS, then the nitrocellulose blocked with 2% BSA in TBS for 1 hr. Primary antibodies against capsule (2-2-B) were
used at a concentration of 1:1000 in blocking buffer. An alkaline phosphatase conjugated secondary antibody against 2-2-B was used for detection and then visualised via colourmetric development. A negative control was included in each assay.

Opa expression was determined by western blotting. Whole cell lysates (750 ng) were separated by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by standard methods and transferred to nitrocellulose membranes. The membranes were blocked overnight with 2% BSA in TBS. The polyclonal rabbit anti-Opa primary antibody (49), was used at 1:1,000. A HRP-conjugated anti-Rabbit IgG secondary antibody (Santa Cruz biotechnology) was used for detection and then visualised via chemilluminescence (GE healthcare).

3.3.6 Tricine-SDS-PAGE analysis of LOS preparations.

A mini Protean apparatus (Bio-Rad) was used for Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) performed according to the description of Schagger and von Jagow (44). Crude LOS preparations were prepared from whole-cell lysates of meningococcal growth suspended in distilled water. The protein concentrations of these preparations were approximated by the Bradford assay (Bio-Rad). Proteinase K digests consisted of 1 µg of protein in 2% SDS (total volume, 10 µl) to which 2 µl of 25 mg/ml proteinase K (Sigma) was added, and the mixture was incubated at 55°C for 30 min. A second aliquot of proteinase K was added and the digestion was repeated before the reaction was stopped with the addition of 25 µl of loading buffer (1 M Tris [pH 8.45], 10% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.05% bromophenol blue). A 4 µl aliquot of each sample was heated to 100°C for 4 min
before being loaded onto a 16% Tricine–SDS-PAGE gel. After electrophoresis, the gels were fixed in 40% ethanol–5% acetic acid overnight. The gels were silver stained according to the method of Hitchcock and Brown (16).

3.3.7 β-galactosidase assay

Suspensions of bacteria were added to wells of a 96-well plate containing confluent Detroit 562 cells at a MOI of 100, and a plate with no cells. Attachment was allowed to proceed for two hours. Wells were sacrificed to ascertain the percentage of adherent bacteria for each strain. Following attachment, non-adhered bacteria were removed by washing with D-PBS and LacZ expression assayed as per Gu et al. (12) The remaining adherent cells were lysed by the addition of 20 µl of permeabilisation solution (100 mM Na₂HPO₄, 20 mM KCl, 2 mM MgSO₄, 0.8 mg/ml cetyl trimethylammonium bromide (CTAB), 0.4 mg/ml sodium deoxycholate (DOC) and 5.4 µl/ml β-mercaptoethanol) incubated at room temperature for 30 mins to facilitate lysis, while bacteria collected from the media alone were enumerated by viable count, collected by centrifugation and resuspended in 20 µl lysis buffer and incubated in a 96-well plate as above. The optical density at 600 nm was recorded for each sample. To the lysed samples, 60 µl of substrate solution containing 60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 0.25 mg/ml ortho-Nitrophenyl-β-galactoside (ONPG) and 2.7 µl/ml β-mercaptoethanol was added. The plates were incubated at 37°C and once ONPG hydrolysis had occurred (as indicated by a yellow colour), 70 µl of stop solution (1 M Na₂CO₃) was added and the time recorded. The OD₄₂₀ was determined for each sample. Miller units were determined as 1000 x OD₄₂₀ / (OD₆₀₀ x volume (ml) x time (mins)) and adjusted for the average number of bacteria in each sample.
3.3.8 Preparation of LOS and analysis.

The LOS preparations, prepared as previously described (26), were washed three times with 9:1 ethanol/ water (v/v) mixture to remove contaminating phospholipids. The washed LOS preparations were suspended in water and lyophilized. Samples were then subjected to mild acid hydrolysis in 1% aqueous HOAc (v/v) for 2.5 hrs at 100°C with constant stirring. The lipid A precipitate that formed during hydrolysis was collected by centrifugation at 3000 rcf for 15 mins at 4°C, and supernatants containing the released oligosaccharides (OSs) were decanted and lyophilized.

Oligosaccharides were analyzed by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) using a 4700 Proteomics Analyzer instrument (Applied Biosystems) in reflectron mode. The OS samples were dissolved in water (1 mg/mL) and mixed in a 1:1 (v/v) ratio with 0.5 M 2,5-dihydroxybenzoic acid (DHB) in methanol matrix solution. Spectra were acquired in the positive acquisition mode.

3.4 Results

3.4.1 Distribution of LOS inner core biosynthesis genes is associated with different lineages of *N. meningitidis*

An examination of the publicly available genomes (28 meningococcal and 17 gonococcal (11)) for the LOS glycosyltransferases *lgtG, lptG, lot* and *lgtA* revealed that carriage of these genes is diverse (summarised in Table 3.2). The LOS *O*-
acetyltransferase, lot, is present in all genomes of N. meningitidis and N. gonorrhoeae. The ORF contains a homopolymeric tract of 5 guanine nucleotides in most genomes but in the three sequenced ST-32 clonal complex (cc) strains (MC58, H44/76 and CU385) 6 guanine nucleotides are present. However, the ORF is in-frame in all other available genomes, including ST-269 cc strains which are closely related to ST-32 cc (6).

The inner core transferases, lgtG and lpt6, are present in all available gonococcal genome sequences and the genomes of ST-8 cc and ST-11 cc strains of N. meningitidis (Table 3.2). Of the remaining 20 isolates, only two meningococcal isolates from unrelated clonal complexes (N1568 (ST-751 cc) and WUE2594 (ST-5 cc)) also possessed both LOS inner core transferases. Interestingly, the isolates from the ST-41/44 cc (5 strains) and ST-269 cc (2 strains) and four individual isolates representing different clonal complexes (α14 (ST-53 cc), 8013 (ST-18 cc), M01_240013 (ST-1159) and M01_240355 (ST-213)) did not possess either lgtG or lpt6. Five strains from ST-32 cc (3 strains), NS44 (ST-23 cc) and ATCC13091 (ST-7355) possessed lgtG alone without lpt6. Lastly, ST-4 cc strain Z2491 and ST-4821 strain 053442 possessed lpt6 alone without lgtG.

lgtA was potentially phase variable at high frequency (>10 guanine nucleotides in a single tract) in all strains of N. gonorrhoeae (17/17), and six meningococcal genomes [ST-32 cc (3/3 strains), ST-269 cc (1/2 strains), and ST-41/44 cc (2/5 strains)]. Of the remaining 22 meningococcal genomes, the lgtA gene contained less than 10 guanine nucleotides, with the majority of lgtA alleles from 20 genomes containing a 5 bp nucleotide tract. Recently, Alfsnes et al. (1) have shown that G tracts of 5 bp in length have no demonstrable levels of phase variation, leading to the prediction that these lgtA
alleles are not likely to phase vary. In summary, a non-phase variable \( lgtA \) allele is most commonly found in meningococcal disease isolates from the ST-8 cc and ST-11 cc, which correlates strongly with the possession of \( lgtG, lpt6, \) and \( lot. \)

3.4.2 The construction of invasive LOS immunotype structures in strain NMB-CDC.

To understand the contribution of the LOS inner core residues to meningococcal pathogenicity, we tested their role in bacterial attachment and invasion. A series of LOS inner core transferase mutants were constructed in strain NMB-CDC (ST-8 cc) which possesses functional \( lot, lpt6, lgtG \) and \( lpt3 \) genes (27) to create strains expressing seven different LOS structures (Figure 3.2). The LOS inner core structure was altered by insertional inactivation of \( lgtG \) (NMB\( \Delta lgtG \)), \( lpt6 \) (NMB\( \Delta lpt6 \)) and \( lot \) (NMB\( \Delta lot \)), alone and in combination (NMB\( \Delta lgtG\Delta lpt6 \), NMB\( \Delta lgtG\Delta lot \) and NMB\( \Delta lgtG\Delta lpt6\Delta lot \) as described previously (27, 28, 52). To aid clarity, the LOS inner core structures expressed by the different mutants were assigned structure designations (S1 through S7) and equivalent immunotype classifications have been indicated where possible (Figure 3.2A). Parental strain NMB-CDC expresses a mixed population of LOS structures, termed the L2 and L4 immunotypes and denoted here as the S1 and S2 inner core structures (40). The S1 structure is characterised by the addition of \( O-3 \) linked glucose (\( O-3 \) Glc) and \( O-6 \) linked phosphoethanolamine (\( O-6 \) PEA) to HepII and a \( O-6 \) linked \( O \)-acetyl group (\( O-3 \) OAc) on the terminal \( N \)-acetylglucosamine residue (GlcNAc) of the \( \gamma \)-chain. S1 LOS is the major species of LOS expressed by this isolate, while the minor LOS species has an inner core of the S2 structure in which the \( O-3 \) linked Glc is
absent. This effect is due to the transcriptional repression of the non-phase variable \( lgtG \) lipooligosaccharide \( O-3 \) glucosyltransferase by the response regulator (MisR), during normal culture conditions (52). The inactivation of \( lgtG \) in NMB-CDC results in a major population of S2 LOS (NMB\( \Delta lgtG \)) (46). Inactivation of \( lpt6 \), encoding the lipooligosaccharide \( O-6 \) phosphoethanolamine transferase, results in the expression of a mixture of LOS structures, with the major species having S3 inner core structure (\( O-3 \) Glc, \( O-6 \) H, \( O-3 \) OAc) and the minor species having S5 inner core structure (\( O-3 \) H, \( O-6 \) H, \( O-3 \) OAc) (27). Mutant strain NMB\( \Delta lgtG\Delta lpt6 \) expressed a pure population of S5 LOS (27).

### Table 3.2. Summary of the distribution of LOS biosynthesis genes amongst clonal complexes of *N. meningitidis*

<table>
<thead>
<tr>
<th>ST</th>
<th>n</th>
<th>lpt3</th>
<th>lpt6</th>
<th>lgtG</th>
<th>lgtA</th>
<th>lot1(nG)</th>
<th>lgtA1(nG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-4</td>
<td>1</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-5</td>
<td>1</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[8]</td>
</tr>
<tr>
<td>ST-8</td>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-11</td>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-18</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-23</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-32</td>
<td>3</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+_[6]</td>
<td>+_[4-1]</td>
</tr>
<tr>
<td>ST-41/44</td>
<td>5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5-11]</td>
</tr>
<tr>
<td>ST-53</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-213</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-269</td>
<td>2</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-751</td>
<td>1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-1159\a</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-4821</td>
<td>1</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
<tr>
<td>ST-7355\a</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+_[5]</td>
<td>+_[5]</td>
</tr>
</tbody>
</table>

\a: no clonal complex exists for this ST  
\b: All are phase variable with tracts of G>10, therefore on/off status cannot be determined due to accuracy of assembly  
\c: where n=5 the gene is considered to be “phase-locked”

Although strain NMB-CDC contains an active lipooligosaccharide \( O-3 \) phosphoethanolamine transferase, Lpt3, which can add \( O-3 \) linked PEA to HepII of the
inner core, previous structural studies indicated that O-3 linked PEA could not be added to the inner core until the OAc group was removed from the terminal GlcNAc (28). Therefore, inactivation of the lipooligosaccharide O-acetyl transferase (lot) in NMBΔlot results in the expression of a mixture of LOS structures with the major species having S4 inner core structure (O-3 Glc, O-6 PEA) and a minor population of S6 LOS inner core structures (O-3 PEA, O-6 PEA). Inactivation of lgtG in NMBΔlot (NMBΔlotΔlgtG) results in the expression of a pure population of S6 LOS. Inactivation of lpt6 in NMBΔlotΔlgtG (NMBΔlotΔlgtGΔlpt6) results in the expression of an S7 LOS
with only $O$-3 PEA on the LOS inner core and hence is equivalent in structure to the inner core of the L3 and L7 immunotypes (28).

Our previous studies on LOS had been conducted in serogroup B encapsulated and LOS sialylated meningococci. To examine whether encapsulation and LOS sialylation had an effect on the decoration of the LOS inner core, the LOS of strains M7ΔlgG and M7ΔlgGΔlot were analyzed by MALDI-TOF MS. The spectra and the proposed composition for the various ions are shown in Figure 3.3 and Table 3.3, respectively. The observed and calculated ions are monoisotopic masses for the various ionic forms (i.e. various sodiated species). The major species of LOS produced by M7ΔlgG produces oligosaccharides are consistent with the major S2 structure of the parental strain NMB-CDCΔlgG (Figure 3.2). The minor species of LOS produced by M7ΔlgG has an m/z = 1873.59, which is consistent with a structure containing two PEA groups; one of which is likely at $O$-6 of HepII and the second possibly at $O$-3 of HepII. In the case of M7ΔlgGΔlot, the oligosaccharide contains ions consistent with the loss of both a Hex and an OAc group from the terminal GlcNAc residue which would be consistent with the disruption of both lgtG and lot. The oligosaccharide preparation from M7ΔlgGΔlot also contained structures with two PEA groups to a greater extent than does the lgtG mutant, and may also contain structures containing a Gly residue. However, some other spectra (not shown) of the lgtG mutant oligosaccharide showed ions of low intensity consistent with the presence of a Gly residue. This analysis demonstrated that the appearance of the $O$-3 PEA group on the inner core of LOS of strain NMB-CDC can occur in the absence of terminal LOS sialylation and capsule.
Figure 3.3 The MALDI-TOF MS spectra of the LOS of M7ΔigtG and M7ΔigtGΔlot.
Panel A shows the spectra of M7ΔigtG and Panel B contains M7ΔigtGΔlot. The mass of the ions are indicated and the composition of these ions can be found in Table 3.3.
Table 3.3. Composition of the ions from the M7ΔltG and M7ΔltGΔlot.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Observed</th>
<th>Ion</th>
<th>Calculated</th>
<th>Proposed Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>M7ΔltG</td>
<td>1638.22</td>
<td>[M+H]^+</td>
<td>1638.38</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1662.57</td>
<td></td>
<td>[M+H]^+</td>
<td>1662.40</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)Ac(_1)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1678.21</td>
<td></td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>1706.59</td>
<td></td>
<td>[MNa(_2)+H]^+</td>
<td>1706.36</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)Ac(_1)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1728.58</td>
<td></td>
<td>[MNa(_2)+Na]^+</td>
<td>1728.36</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)Ac(_1)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1746.56</td>
<td></td>
<td>[MNa(_3)+Na]^+</td>
<td>1724.38</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)Ac(_1)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1765.58</td>
<td></td>
<td>[MNa+H]^+</td>
<td>1765.39</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1783.64</td>
<td></td>
<td>[MNa+H]^+</td>
<td>1783.41</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1873.59</td>
<td></td>
<td>[MNa(_3)+Na]^+</td>
<td>1873.39</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)Ac(_1)PEA(<em>2)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>M7ΔltGΔlot</td>
<td>1664.64</td>
<td>[MNa(_2)+H]^+</td>
<td>1664.32</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1686.65</td>
<td></td>
<td>[MNa(_2)+Na]^+</td>
<td>1686.32</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1743.68</td>
<td></td>
<td>[M+H]^+</td>
<td>1743.41</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1743.38</td>
<td></td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>1765.64</td>
<td></td>
<td>[MNa+H]^+</td>
<td>1765.39</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1831.66</td>
<td></td>
<td>[MNa(_3)+Na]^+</td>
<td>1831.36</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
<tr>
<td>1888.71</td>
<td></td>
<td>[MNa(_3)+Na]^+</td>
<td>1888.41</td>
<td>Gal(_2)GlcNAc(_2)Glc(_1)Hep(_2)PEA(_2)Gly(<em>1)Kdo(</em>{\text{anhydro}})</td>
</tr>
</tbody>
</table>

3.4.3 Invasion but not attachment to Detroit 562 epithelial cells is moderated by the LOS inner core structure of *N. meningitidis* strain NMB-CDC.

To examine the role of the LOS inner core in the attachment and invasion of epithelial cells by *N. meningitidis*, the parental wild type strain NMB-CDC was compared with each of the various LOS inner core mutants. All strains were shown by western immunoblot to express Opa, capsule and pili (data not shown).

The contribution of the different LOS inner core structures to the rate of attachment of epithelial cells indicated that all strains retained similar rates of attachment to parental...
wild-type strain NMB-CDC except for minimal increases of 1.7-fold (p=0.016) observed for NMBΔlpt6, and 1.5-fold (p=0.0116) for NMBΔlot (Figure 3.4A). The expression of three of the six different LOS inner core structures resulted in detectable increases in invasion of the epithelial cells compared to wild-type strain NMB-CDC (Figure 3.4B). Strain NMBΔlgtG in which the major LOS species has the S2 inner core structure, displayed the greatest change with an 18-fold increase in invasion over parental wild-type (p<0.0001). Strain NMBΔlgtGΔlpt6 which expresses a major LOS species with S5 inner core structure and strain NMBΔlpt6 which expresses a mixed population of LOS with inner core structures of S3 and S5, resulted in a 9-fold and 4-fold increase in invasion compared to the parental wild-type (p=0.0003 and 0.0098, respectively). Therefore, the removal of both O-3 linked Glc and O-6 PEA, individually or in combination with a terminal OAc group, from the LOS inner core led to an increase in invasion into host cells. Conversely, the presence of the O-6 PEA group on the LOS inner core correlated with the best rate of invasion, but this declined when the O-3 Glc residues were present (NMBΔlpt6 relative to NMBΔlgtGΔlpt6). The strains NMBΔlot, NMBΔlgtGΔlot and NMBΔlgtGΔlpt6Δlot which express major LOS species with O-3 PEA and lacked the terminal OAc group, did not have any changes in the rate of invasion relative to the parental wild-type. This data suggested that the replacement of the O-3 Glc with the O-3 PEA group on HepII of LOS had no effect on invasion of strain NMB-CDC in the presence of capsule. Therefore, the optimal LOS structure for contributing to improved invasion of strain NMB-CDC into host cells contained O-6 PEA attached to HepII and O-3 OAc on the GlcNAc of the γ-chain with a LNT α-chain.
Figure 3.4 The presentation of different LOS inner cores by strain NMB-CDC affects attachment and invasion into Detroit 562 epithelial cells.

Panel A. A phenotypic description of LOS structure, capsule, LNT and LOS sialylation is indicated for each strain used in Panel B and Panel C. All strains were encapsulated, piliated and Opa positive. Panel B shows the rates of association and Panel C show the rates of invasion of isogenic mutants of strain NMB-CDC expressing variable LOS inner core structures into Detroit 562 epithelial cells. Attachment as a percentage of the inoculum and invasion as the percentage of the associated bacteria is plotted, normalised to NMB-CDC (=1). The average rate (+/-SEM) from three biological repeats in triplicate following 6 hrs co-incubation is shown. *: p<0.02 determined by Mann-Whitney t-test compared to strain NMB-CDC.
3.4.4 The presence of sialic acid masks the ability of LOS inner cores possessing O-3 PEA groups to initiate invasion into host cells.

In the previous experiments, the LOS structure had an effect on triggering invasion of strain NMB-CDC into host cells even though this isolate remained encapsulated. This is unusual, since previous studies using the strain MC58 (ST-32 cc) which expresses the L3 immunotype (S7 in this report) had indicated that capsule was the predominant factor in preventing invasion into host cells and that LOS moderated invasion into host cells in the absence of encapsulation (58). To examine the effect of LOS on invasion of strain NMB-CDC into host cells, we examined whether terminal LOS sialylation (NANA) and capsule (CAP) affected invasion of the meningococcal LOS variants into host cells. To do this, we constructed mutations in *lgtG*, *lpt6* and *lot* alone and in combination in strain M7, a transposon mutant of NMB-CDC in which *synA* is interrupted and which has no NANA or CAP. All strains were shown by western immunoblot to express Opa, LNT and pili (data not shown).

We have previously shown that strain M7 attaches and invades Detroit 562 epithelial cells at significantly higher rates than the parental wild-type strain (See Chapter 2). M7Δ*lgtG* had a reduced rate of attachment (1.87-fold, p<0.0001) but increased rate of invasion (15-fold, p<0.0001) compared to parental M7 (Figure 3.5). Similarly, M7Δ*lpt6* and M7Δ*lgtGΔlpt6* demonstrated no change in the rate of association relative to the M7 parental strain, but had elevated rates of invasion of 3.4-fold and 7.7-fold respectively (p=0.0011 and p=0.0087 respectively). Therefore, the magnitude of increases in invasion for strains expressing S2, a mixture of S3/5 and S5 LOS structures were similar in both strain NMB-CDC and M7 backgrounds regardless of the NANA and
Figure 3.5 The LOS inner core of NMB affects attachment and invasion into Detroit 562 epithelial cells in the absence of sialic acid.

Panel A. A phenotypic description of LOS structure, capsule, LNT and LOS sialylation is indicated for each strain used in Panel B and Panel C. These strains were piliated and Opa positive. Panel B shows the rates of association and Panel C shows the rates of invasion of isogenic mutants of strain M7 into Detroit 562 epithelial cells. Attachment as a percentage of the inoculum and invasion as the percentage of the associated bacteria is plotted, normalised to M7 (= 1). The average rate (+/−SEM) from three biological repeats in triplicate following 6 hrs co-incubation is shown. *: p<0.01 determined by Mann-Whitney t-test compared to strain M7.
CAP profile. In addition, M7ΔlgtGΔlot which expressed a major population of the S6 structure was not significantly different in attachment and invasion from the parental wild-type M7 (p>0.06).

By comparison, M7Δlot demonstrated a significant increase in invasiveness (6.72-fold, p<0.0001) when compared to the parental M7, although no change in attachment was detected (p=0.8432). Strains M7ΔlgtGΔlot and M7ΔlgtGΔlpt6Δlot did not demonstrate any difference in the rates of attachment compared to the parental strain (p>0.06). However M7ΔlgtGΔlpt6Δlot invaded into epithelial cells 7.85-fold more than the parent strain M7 (p<0.0001). Since M7Δlot and M7ΔlgtGΔlpt6Δlot expressed LOS inner cores possessing O-3 PEA groups in the presence or absence of O-3 Glc (S4 and S7 structures respectively), these combinations of inner core resides were more successful at initiating invasion of meningococci into host cells than the combination of O-3 and O-6 PEA on the LOS inner core (S6 structures). The effect of these LOS structures on meningococcal invasion, however, were masked by CAP and sialylated LOS, since S4 and S7 structures did not influence invasion of encapsulated and LOS sialylated strains (Figure 3.4).

3.4.5 LOS can trigger meningococcal invasion into host cells in the absence of Opa adhesin.

A series of LOS inner core mutants were constructed in a strain of NMB-CDC devoid of Opa proteins to determine if the LOS structure could influence invasion in the absence of this major adhesin. NMB-CDCΔopa was shown to express no Opa proteins by western immunoblot but was still piliated and encapsulated by colony immunoblot
(data not shown). The LOS inner core structure was altered in strain NMBΔopa by insertional inactivation of \( lgtG \) and \( lpt6 \), alone and in combination. These strains were shown to possess LOS with a sialylated LNT as determined by SDS-PAGE and expressed no Opa proteins in the inocula, from bacteria attached to epithelial cells and bacteria recovered from inside epithelial cells by western immunoblot (data not shown). This indicated that although Opa expression is known to be phase variable in at least three of the four loci, phase variation of these genes was not detected throughout the assay and therefore did not affect the final interpretation of the results.

A comparison of the Opa deficient strains with different LOS inner cores revealed that the strain expressing an inner core with no substitutions except the terminal OAc residue (S5 in \( ΔlgtGΔlpt6 \)) attached more frequently than the Opa minus parent isolate expressing LOS inner cores with \( O-3 \) Glc and \( O-6 \) PEA (1.77-fold increase over the Opa minus parent, \( p=0.014 \)) (Figure 3.6A, B). Importantly, the ability of the LOS inner core variants to improve invasion remained in the absence of Opa. Thus, the strains expressing S2, S3 and S5 inner core structures (\( ΔlgtG, Δlpt6 \) and \( ΔlgtGΔlpt6 \)) retained elevated rates of invasion in an Opa deficient background compared to the Opa minus parent (\( p=0.0134, p=0.0418 \) and \( p=0.0400 \) respectively) (Figure 3.6C). The magnitude of the change diminished in the absence of Opa adhesins from 18-, 4- and 9- fold in the wild-type NMB-CDC parental background (Figure 3.4) to 8.4-, 4.1- and 2.6- fold respectively in the Opa mutant background (Figure 3.6). Therefore, while the LOS structures were capable of triggering invasion independently of Opa mediated adhesion, we hypothesise that there is a synergistic relationship between the intimate adhesion of the meningococcus mediated by Opa and the invasion mediated by the LOS structure.
Figure 3.6 LOS triggered invasion is repressed in a background devoid of Opa proteins.

Panel A. A phenotypic description of LOS structure, capsule, LNT and LOS sialylation is indicated for each strain used in Panel B and Panel C. These strains were pillated and encapsulated. Panel B shows the rates of association and Panel C shows the rates of invasion of strain NMB-CDCΔopa and the isogenic mutants lacking lgtG and lpt6 alone and in combination into Detroit 562 epithelial cells. Attachment as a percentage of the inoculum and invasion as the percentage of the associated bacteria is plotted, normalised to NMB-CDCΔopa (=1). The average rate (+/-SEM) from three biological repeats in triplicate following 6 hrs co-incubation is shown. *: p<0.05 determined by Mann-Whitney t-test compared to strain NMB-CDC.
3.4.6 The LOS inner core transferases are regulated in response to host cell contact.

Since the substitution of the LOS inner core is important in determining the rates of invasion into host cells, we examined whether the genes encoding the transferases, \( lgtG \), \( lpt6 \), \( lot \) and \( lst \) were regulated during attachment and invasion. The promoter for each LOS transferase gene (\( lot \), \( lpt6 \), \( lst \), \( lgtG \) and the alternative \( lgtG \) promoter \( mtr \)), capsule synthesis and transport operons (\( synA \) and \( ctrA \)) and \( misR \) was cloned upstream of a \( lacZ \) reporter gene and integrated into an intergenic region of the chromosome and LacZ expression measured as described by Tzeng et al. (54). In response to contact with epithelial cells for 2hrs, the expression of \( synA \) and \( ctrA \) declined by 35% (\( p<0.0001 \)) and 29% (\( p<0.0001 \)), respectively, compared to a control strain in which no promoter was cloned upstream of the \( lacZ \) reporter. In concert with the decline in the synthesis of sialic acid by the \( synA \) locus, \( lst \) expression also declined by 22% (\( p=0.0433 \)). Together, the data suggests that the amount of sialic acid synthesis and the ability to form capsule and sialylated LOS declined in response to contact with host cells thus improving meningococcal invasion into host cells (Figure 3.7). No change in the expression of \( P_{lot} \), \( P_{misR} \) and \( P_{mtr} \) which has been shown to be the MisR regulated promoter for \( lgtG \) (53), was detected at this time point. However, the expression from \( P_{lpt6} \) increased 27% (\( p=0.0124 \)) indicating that the substitution of the LOS inner core was regulated in response to contact with host cells.

3.5 Discussion

Strain NMB-CDC is an exemplar isolate of the ST-8 cc and has the typical phenotypic features associated with this lineage which includes the expression of Class II pili, the
opa alleles 257, 161, 277 and 246, and the absence of the Opc adhesin (43). This isolate has been shown by Richardson et al. (42) to have a slow rate of phase variation. In the previous studies conducted on the LOS profile of this isolate, LOS structures consistent with a phase variable truncated structure lacking LNT have never been detected (40). This has led to the hypothesis that this isolate did not have a phase variable α-chain which was substantiated by the sequence of the genome of this isolate which indicates that the lgtA allele has a short nucleotide poly-guanine sequence of 5 bp consistent with a gene that is intact but not phase variable. An examination of the publicly available genome sequences indicated that all isolates from the ST-8 cc and the closely related ST-11 cc, consistently possessed an intact lgtA allele with no phase variable tract. In the current model for meningococcal invasion into host cells, the LNT of LOS is considered to hinder the Opa adhesin and is removed by phase variation of lgtA so that invasion can proceed. However, the ST-8 strain NMB-CDC must effectively attach to and invade host cells while expressing LOS with an LNT α-chain which cannot be removed.

Figure 3.7 Capsule, terminal LOS sialylation and LOS inner core decorations are regulated in response to host cell contact as determined by β-galactosidase reporter assay.
Promoters were fused with lacZ and expressed from the chromosome of NMB-CDC as a single copy to construct reporter strains (Table 1). LacZ expression was measured from bacterial strains incubated in the absence of epithelial cells (white bars) and adherent to epithelial cells for 2 hrs (black bars) and are expressed as Miller units. The plotted values correspond to the mean of at least three independent experiments, each with triplicate measurements. * p<0.05.
by phase variation. Since the isolates from the ST-8 cc and ST-11 cc also possessed the
LOS inner core transferases which build various LOS inner core structures (S1-7), these
were individually tested for a role in enabling invasion into host cells in the presence of
the LNT α-chain.

The decorations of the LOS inner core were examined in both the serogroup B
encapsulated and LOS sialylated background in strain NMB and the non-encapsulated,
non-LOS sialylated background of the isogenic mutant M7. We have previously shown
that although strain NMB-CDC contained a functional Lpt3, the O-3 PEA group was
not detected on the inner core when a full length LNT with terminal OAc was present
on the LOS inner core (28). However, Mistretta et al. (34) reported that the O-3 PEA
group was present on LOS with LNT and terminal OAc in serogroup A strains which do
not express a sialic acid capsule and are unable to sialylate LOS during laboratory
culture unless they are supplied with an exogenous source of CMP-N-acetyl neuraminic
acid. To test the hypothesis that LOS sialylation of LNT had an effect on whether O-3
PEA substitution of the LOS inner core occurred in strain NMB, the LOS inner cores of
M7ΔlgtG and M7ΔlgtGΔlot were compared. MALDI-TOF MS spectra indicated that
the removal of the terminal OAc in the M7ΔlgtGΔlot resulted in the appearance of O-3
PEA on the inner core. However, since detectable levels of O-3 PEA on the LOS inner
core was also present in M7ΔlgtG this suggested that the lack of LNT sialylation alone
could allow the addition of this residue to the inner core.

The expression of different LOS inner core substitution profiles by strain NMB-CDC
affected invasion of this isolate into host cells in the presence of capsule. The most
effective LOS structures for triggering invasion for strain NMB-CDC were those which
contained $O$-$6$ PEA attached to HepII and $O$-$3$ OAc on the GlcNAc of the $\gamma$-chain in the presence of LNT $\alpha$-chain. Based upon the immunotyping scheme for LOS structures, this corresponds to the L4 immunotype (S2 in this report). The presence of capsule and LOS sialylation did not mask the ability of the S2 LOS structure to trigger invasion, a series of LOS structures (S6 and S7) that possessed $O$-$3$ PEA groups attached to HepII only affected invasion in the absence of capsule and LOS sialylation. Since the S1-S7 series of LOS structures did not change in overall length, but had variable additions to the LOS inner core, it is likely based upon structural modelling by Gidney et al. (10) that the conformation and flexibility of these structures differ. The conformation of the LOS inner core may affect the exposure of the PEA groups and hence their interactions with other macromolecules on the bacterial surface. Indeed, Ram et al. (41) has shown that LOS inner cores with $O$-$6$ PEA bind more complement factor C4b than strains with $O$-$3$ PEA groups on the LOS inner core and consequently are more sensitive to serum killing. In addition to this, amino groups on sugars have been shown to form hydrogen bonds with oxygen atoms of spatially proximal sugars of the E. coli lipopolysaccharide (37) which may indicate a more direct mechanism of masking of the residues on the LOS inner core by capsule. In this scenario the potential exists for hydrogen bonds to form between the amino groups of the PEA on the LOS inner cores and the carboxyl groups of the polysialic acid residues of the capsule polymers.

The effect of the LOS structure on invasion of meningococci into host cells was partially dependent upon a synergistic interaction with Opa, since the rate of invasion of the modified LOS structures compared to the wild-type LOS structure was lower in the encapsulated Opa-devoid mutant strain compared to the encapsulated Opa-competent wild-type strain. Moore et al (35) demonstrated that Opa proteins are lectins with the
ability to bind simple saccharides, in particular those terminating in sialic acid, at high affinity. Since the LOS structures were terminated with sialylated LNT, it appears that the changes to the decoration of the inner core of the LOS have affected the interaction of the LOS with the Opa proteins thus relieving inhibition and improving the interaction with host cell receptors such as CEACAM. Alternatively, the change in the LOS conformation may also have improved the ability of the LNT to act as a ligand with an unidentified host cell receptor as has been previously postulated by Lambotin et al. (31, 47). In both scenarios, the most pronounced improvement of the rates of invasion were governed by the addition of O-6 PEA to the LOS inner core in the presence of a non-phase variable LNT addition to the HepI of the LOS.

Previous work has shown that capsule expression is down-regulated in response to host cells as a mechanism to improve bacterial invasion of host cells (9). Since the expression of the O-6 PEA residue on the LOS inner core was shown to be important for the invasiveness of strain NMB-CDC, the expression of the genes involved in decorating the LOS inner core during invasion of host cells was analysed. lgtG can be phase variable in some meningococcal isolates (2) and is also under the transcriptional regulatory control of the two-component response regulator/histidine kinase system, MisRS. Since MisR also directly represses capsule expression by binding to the cps promoter region during broth culture in complex media (53) (See Chapter 2) , the transcription of misR, lgtG and the cps loci were assessed in strain NMB-CDC in response to contact with Detroit 562 epithelial cells. Although the assays indicated that contact dependent regulation of capsule expression was observed in strain NMB-CDC in the presence of Detroit 562 cells, the level of transcription from the promoters of misR and lgtG remained unchanged indicating no signalling through this pathway could
be detected under the conditions used in these assays. This may indicate a time dependency in the triggering of MisR, perhaps potentially in the very early events of host cell contact (19) whilst this assay was conducted at a much later time point. Previous studies have indicated that contact dependent regulation of the *cps* locus can occur via the regulatory protein CrgA (contact regulated gene A), a LysR-type regulator (9) although this has been challenged and may not be true for all meningococcal isolates (18). Morelle *et al.* (36) has shown that CrgA-dependent genes contain a CREN (contact regulatory element of Neisseria, formerly REP2) within the promoter region. However, this motif was not observed in the promoter region of strain NMB-CDC *cps* locus, potentially implicating a different regulatory system in this phenomenon.

*lpt6*, *lot*, *lpt3* and *lst* are not phase variable in strain NMB-CDC or other sequenced isolates (Table 3.2). Importantly, the transcription of both *lpt6* and *lst*, but not *lot* or *lpt3*, was altered in response to host cell contact in strain NMB-CDC on Detroit 562 cells at the 2 hr time point. There is no known regulatory mechanism for *lst* or *lpt6* expression in meningococci. *lst* transcription was down-regulated consistent with the concept that the terminal sialylation of the LNT of LOS is inhibitory to invasion. Conversely, *lpt6* was up-regulated consistent with the proposed model that the addition of *O*-6 PEA to the LOS inner core increases invasiveness of strain NMB-CDC. Neither *lst* nor the *lpt6* promoter fragments from strain NMB-CDC contained CREN (data not shown) consistent with previous analyses of in the genomes of strain MC58 or Z2491 (36). Future studies will be undertaken to identify the host contact dependent regulatory system controlling *lst* and *lpt6* expression in meningococci.
By virtue of an increased bacterial surface area, strain NMB-CDC possesses a capsule that is less dense than other meningococcal isolates (See Chapter 2). In this meningococcal isolate, the LOS inner core structure modulates the interaction of the bacterium with the host cell. In this setting, the inner core structure S2 (equivalent to L4 immunotype) LOS which contains an O-6 PEA group attached to the HepII, is O-acetylated on the γ-chain and contains LNT on the α-chain promoted the rate of invasion into host cells. Although the LOS structure triggered invasion independent of Opa proteins, there is synergism between Opa and LOS in promoting invasion into the host cell. Therefore, we hypothesise that the combination of Opa alleles and LOS inner core structures that a strain expresses has an impact on the optimal rate of bacterial invasion into host cells. Lastly, the LOS structure is not static, but actively re-modelled in response to contact with host cells suggesting that LOS has an important and dynamic role in meningococcal invasion into host cells.

### 3.6 Acknowledgments

MAb 2-2-B was a generous gift from Dr Wendell Zollinger and Dr. Elizabeth Moran (Walter Reed Army Institute of Research, USA). The polyclonal anti-Opa and anti-PilE antibodies were a gift from Dr. Hideyuki Takahashi (National Institute of Infectious Diseases, Japan).

### 3.7 References

**Neisseria meningitidis** model, p. 145, XIIIth international pathogenic **neisseria** conference Wurzburg, Germany.


Chapter 4. A bi-functional UDP-galactose 4-epimerase results in the biosynthesis of N-acetyl galactosamine for gonococcal LOS and the serogroup 29E and Z capsules of *N. meningitidis*
4.1 Abstract

*Neisseria gonorrhoeae* expresses lipooligosaccharide (LOS) terminated with *N*-acetyl D-galactosamine (GalNAc) which efficiently binds host receptors to mediate bacterial invasion into host cells. GalNAc is not present as a component of the LOS of the closely related pathogen, *Neisseria meningitidis*, but is a component of the polysaccharide capsules of serogroup 26E and Z strains. The ability to synthesise UDP-GalNAc has not been previously determined in either bacterial species. In other bacterial species UDP-galactose 4-epimerase (GalE) is a bi-functional enzyme capable of epimerising UDP-\(N\)-acetyl glucosamine (GlcNAc) to UDP-GalNAc. *N. gonorrhoeae* was shown to carry a GalE allele that is bi-functional and epimerises UDP-GlcNAc and UDP-Glc to UDP-GalNAc and UDP-Gal, respectively. The bi-functional gonococcal GalE can be converted to a mono-functional enzyme unable to epimerise GlcNAc by replacement of a Serine at position 299 in the substrate binding pocket. Conversely, a mono-functional meningococcal GalE was converted to a bi-functional enzyme by the replacement of the phenylalanine at position 300 with a serine residue in the binding site pocket. An analysis of the GalE alleles from both commensal and pathogenic *Neisseria* sp. indicates that the bi-functional GalE allele is found in all sequenced gonococci, *N. lactamica, N. polysaccharea*, but *N. meningitidis* and *N. cinerea* carry either mono- and bi-functional alleles of GalE.

4.2 Introduction

Of the genus *Neisseria*, *Neisseria gonorrhoeae* and *Neisseria meningitidis* are two pathogenic species which colonise the human as the sole host and cause disease. *N.
*Neisseria gonorrhoeae* colonises the mucosal surfaces of the urogenital tract in both males and females. In males a gonorrhoea infection usually results in acute inflammation of the urethra and the presence of a purulent discharge, while in females a gonorrhoea infection is usually asymptomatic. Higher morbidity is seen in women as untreated gonorrhoea can progress to pelvic inflammatory disease (PID) resulting in infertility in approximately 40% of patients (36) *Neisseria meningitidis* is often considered a commensal organism, asymptptomatically colonising the mucosa of the nasopharynx of approximately 10% of the adult population (8, 41). Occasionally this bacterium invades the host, entering the bloodstream where it can cause septicaemia, or crossing the blood-brain barrier to enter the cerebral spinal fluid where it can cause meningitis. Vaccines exist to four of the five common invasive serogroups (of 13 known), however serogroup B meningococci are not covered due to similarity of the capsule with human glycans (17).

The lipooligosaccharide (LOS) structures expressed by *N. gonorrhoeae* and *N. meningitidis* are very similar (Figure 4.1) and are important virulence determinants contributing to the pathogenesis of each organism. LOS is a glycolipid of the outer leaflet of the outer membrane and consists of a lipid A anchor attached to a conserved inner core containing heptose (Hep) and 2-keto D-manno octulsonic acid residues. The variable α-chain extension from HepI and a β-chain extension from HepII contribute to the diversity of these structures (Figure 4.1). The presence of a lacto-α-neotetraose (LNT) α-chain which can be terminally sialylated with N-acetyl neuraminic acid is essential for resistance to complement human serum in gonococci (35) and contributes to this phenotype in meningococci (30). In addition to this, the LOS has also been demonstrated to be important in triggering invasion of meningococci and gonococci.
into host cells (See Chapter 3, (16)). One interesting difference between the two pathogens is the ability of gonococci to terminate the LNT with the alternative residue, N-acetyl glucosamine (GalNAc).

Figure 4.1 The structure of lipooligosaccharide of *N. meningitidis* and *N. gonorrhoeae*.

The core structure (shown in green) is conserved between all known structures of naturally occurring meningococcal and gonococcal LOS. The pathogenic α-chain of LNT: Galβ(1→4)GlcNAcβ(1→3)Galβ(1→4)Glc is shown in pink (although alternatives exist). The transferases and the linkages they catalyse for each residue is indicated, with the gonococcal specific additions indicated in blue. Annotations are: α2,3 LOS sialyltransferase (Lst); β1,3 LOS galactosyltransferase D (LgtD); β1,4 LOS galactosyltransferase B (LgtB); β1,3 LOS N-acetylgalcosamine transferase (LgtA); β1,4 LOS galactosyltransferase (LgtE); β1,4 LOS glucosyltransferase (LgtF); α1,3 LOS heptosyltransferase C (RfaC); α1,3 LOS heptosyltransferase F (RfaF); α1,2 N-acetylgalcosamine transferase (RfaK); O-3 LOS O-acetylgalosamine transferase 3 (Lot3); α1,3 LOS glucosyltransferase (LgtG), O-3 LOS phosphoethanolamine transferase (Lpt3) and O-6 LOS phosphoethanolamine transferase (Lpt6).

The presence of GalNAc on the LOS of gonococcal strains was first suspected by Allen *et al.* (1) who showed that GalNAc-binding lectins recognised LOS. The presence of GalNAc was later confirmed by mass spectrometry and NMR which also confirmed the terminal location of GalNAc on the non-reducing end of LNT (12). The LNT chain of the LOS serves as an important ligand for triggering the invasion of *N. gonorrhoeae* into host urethral and macrophage cells through the binding of asialoglycoprotein
receptor (ASGP-R) and macrophage galectin-like lectin (MGL), respectively. The binding of these receptors is greatly enhanced when the LOS terminates with N-acetyl galactosamine (GalNAc) which results in increased rates of attachment and invasion (39, 55). The \textit{LOS} glycosyl transferase D (LgtD) is the phase variable transferase responsible for the addition of GalNAc to the terminal galactose (13). In \textit{N. gonorrhoeae}, the \textit{LOS} glycosyl transferase A (LgtA) which extends the LOS α-chain is also phase variable and hence for GalNAc to be present on the LOS of the gonococcus both LgtA and LgtD must be phased on. This is not the case for \textit{N. meningitidis} where GalNAc has never been identified as a component of the LOS and \textit{lgtD} is also usually absent (60). However GalNAc is a component of the meningococcal capsule polysaccharide of the “non-pathogenic” serogroup 29E and Z strains (6, 23). Although many of the biosynthetic pathways of \textit{Neisseria sp.} have been elucidated, the mechanism by which GalNAc is synthesised has not been proposed.

Microbes synthesize GalNAc either via a bi-functional UDP-galactose 4-epimerase (GalE) or a UDP-GalNAc 4-epimerase (GNE). Since \textit{Neisseria} sp. do not contain a GNE homologue, we investigated GalE as the potential source of UDP-GalNAc in \textit{Neisseria}. GalE epimerases are classified into three functional groups with group 1 enzymes specific for UDP-hexose epimerisation, group 2 enzymes epimerising both UDP-hexoses and UDP-hexNAc and group 3 enzymes specific for UDP-hexNAc epimerisation (21). We hypothesised that the neisserial GalE could be a group 2 bi-functional epimerase that synthesizes both UDP-galactose (UDP-Gal) and UDP-GalNAc from the substrates UDP-glucose (UDP-Glc) and UDP-\textit{N}-acetylglucosamine (UDP-GlcNAc), respectively. We examined the GalE of \textit{N. gonorrhoeae} and \textit{N. meningitidis} and found that biosynthesis of UDP-GalNAc was a product of a group 2
enzyme in gonococci. Most meningococci possessed a group 1 GalE enzyme and are unable to synthesise UDP-GalNAc via GalE mediated epimerisation of UDP-GlcNAc. We also demonstrated that substrate utilisation and hence functionality could be switched by changing a single amino acid in the binding pocket of the enzyme. Based upon this analysis, the functionality of the GalE of sequenced commensal and pathogenic *Neisseria sp.* were predicted to be either mono-functional or bi-functional with activity being restricted amongst Neisserial species.

### 4.3 Materials and Methods

#### 4.3.1 Bacterial strains and growth conditions

Meningococcal strains were cultured under aerobic conditions with 5% CO₂ at 37°C on GC agar (GCA) or GC broth (GCB) (Oxoid) supplemented with 0.4% glucose, 0.01% glutamine, 0.2 mg of cocarboxylase per litre, and 5 mg of Fe(NO₃)₃ per litre. The wild-type strains and constructed mutants used in this study are shown in Table 4.1. Antibiotic selection for meningococcal mutants was performed on GCA containing 100 µg/ml of kanamycin (sulfate salt), 60 µg/ml of spectinomycin, 5 µg/ml of tetracycline or 2 µg/ml of erythromycin (Sigma). *Escherichia coli* DH5α was used as a host for all DNA manipulations and BL21-DE3 rosetta as a host for all protein expression. *E. coli* were routinely grown on Luria-Bertani broth (LBB) and agar (LBA, Oxoid) which, where appropriate, was supplemented with antibiotics at the following concentrations: ampicillin at 100 µg/ml, spectinomycin at 50 µg/ml, kanamycin at 50 µg/ml, erythromycin at 300 µg/ml, tetracycline at 12.5 µg/ml and chloramphenicol at 30 µg/ml (Sigma).
4.3.2 Cell lines and culture

The immortalised Detroit 562 (human pharyngeal carcinoma epithelial cells, ATCC CCL-138) cell line was grown to confluence in Minimal Essential Media (MEM + Earles salts) supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 2 mM sodium pyruvate and 1x non-essential amino acids (Invitrogen) as per ATCC recommendations (37). At confluence the cells were lifted with 0.05% trypsin/EDTA (Invitrogen) and resuspended in cell media. Cells were counted and 1x10^5 cells were seeded into each of the wells of a 24-well plate (Nunc).

### Table 4.1 Strains used in this study.

<table>
<thead>
<tr>
<th>Strain Name</th>
<th>Genotype</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neisseria meningitidis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMB</td>
<td></td>
<td>(43)</td>
</tr>
<tr>
<td>MC58</td>
<td></td>
<td>(50)</td>
</tr>
<tr>
<td>SS3</td>
<td>NMB(^{galE})::Tn916</td>
<td>(42)</td>
</tr>
<tr>
<td>CKNM550</td>
<td>NMB+pCMK777</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM551</td>
<td>NMB+pCMK778</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM552</td>
<td>SS3+pCMK777</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM553</td>
<td>SS3+pCMK778</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM554</td>
<td>FAM18+pCMK777</td>
<td>This study</td>
</tr>
<tr>
<td>CKNM555</td>
<td>FAM18+pCMK778</td>
<td>This study</td>
</tr>
</tbody>
</table>

4.3.3 Construction of expression strains

The plasmids constructed and used in this study are listed in Table 4.2.

Genomic DNA was isolated from *N. gonorrhoeae* strain FA1090 and *N. meningitidis* strains NMB and MC58 using the Purelink™ genomic DNA purification kit (Invitrogen) according to manufacturer’s instructions. The gene *galE* was amplified with the primers KAP580 (5′-GGAATTCCATATGAAAAAAATTCTCGTTACC-3′) and KAP581 (5′-CGCGGATCCCTTAATCGTCGAGCCATTCC-3′) for MC58,
KAP582 (5’-GGAATTCCATATGACCGTCCTGATTACCG-3’) and KAP583 (5’-CGCGGATCCTTAATCCCCATATCTGCCG-3’) for FA1090 and KAP563 (5’-GGAATTCCATATGCCCTATACGGAAGATATG-3’) and KAP564 (5’-CGCGGATCCTTAATCCCCATATCCGTTGGG-3’) for NMB, such that there was a 5’ NdeI site overlapping the start codon and a 3’ BamHI site following the stop codon incorporated into the PCR fragment. Each PCR product was directionally cloned into the NdeI and BamHI sites of pET15b, resulting in the 5’ fusion of the open reading frame with a hexahistidine motif, resulting in plasmids pCMK730, pCMK729 and pCMK771. The cloned galE of N. gonorrhoeae and N. meningitidis were sequenced and their sequence found to be identical to the native gene.

Table 4.2 Plasmids used in this study

<table>
<thead>
<tr>
<th>plasmid</th>
<th>genotype</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET15b</td>
<td></td>
<td>Novagen</td>
</tr>
<tr>
<td>pHSG576</td>
<td></td>
<td>(47)</td>
</tr>
<tr>
<td>pCMK668</td>
<td></td>
<td>(33)</td>
</tr>
<tr>
<td>pCMK730</td>
<td>pET15b+galE(MC58)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK729</td>
<td>pET15b+galE(FA1090)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK771</td>
<td>pET15b+galE(NMB)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK733</td>
<td>pET15b+galE(MC58)F300S</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK734</td>
<td>pET15b+galE(FA1090)S299F</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK773</td>
<td>pHSG576+galE(MC58)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK774</td>
<td>pHSG576+galE(NMB)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK668</td>
<td>pHSG576+P_{ompR}:GFP</td>
<td>(33)</td>
</tr>
<tr>
<td>pCMK775</td>
<td>pCMK773+P_{ampR}</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK776</td>
<td>pCMK774+P_{ampR}</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK777</td>
<td>pCMK668+galE(MC58)</td>
<td>This study</td>
</tr>
<tr>
<td>pCMK778</td>
<td>pCMK668+galE(NMB)</td>
<td>This study</td>
</tr>
</tbody>
</table>

The cloned galE of N. gonorrhoeae strain FA1090 and N. meningitidis strain MC58 were mutagenised using site directed mutagenesis to change a single nucleotide. The primers KAP427 (5’-CGACTTGGCGTGTTTCTATGCGCCGACCC-3’) and KAP428 (5’-GGGTCGGCATAGAAACACCGCCAAGTGC-3’) to mutate the gonococcal galE, and KAP429 (5’-GGTGATTTGGCCTGCTCTATGCGCCGACCC-3’) and KAP430 (5’-GGGTCGGCATAGGAGCAGCCAAATCACC-3’) to mutate the meningococcal galE were used to amplify the plasmids pCMK729 and pCMK730 respectively using
the non-strand displacing polymerase phusion (NEB). The template DNA was digested with \textit{DpnI}, prior to transformation into DH5\textalpha. The resulting plasmids were sequenced and confirmed to contain only the introduced mutation. These plasmids were called pCMK733 and pCMK734 respectively. Expression strains were constructed by transforming the plasmids pCMK729, pCMK730, pCMK771, pCMK733 and pCMK734 into BL21-DE3 Rosetta. Inducible protein expression from these strains were confirmed by growing the strains to mid-log phase in 10 ml LBB, splitting the culture in two, and inducing one culture with 0.3 mM IPTG for 1 hr. The cells from 1 ml of each culture was collected by centrifugation and resuspended in 100 µl of distilled water to which 25 µl of 5x Laemlli loading buffer was added. The samples were heated to 100°C for 10 mins prior to separation on 12% SDS-PAGE at 150V, then stained with Coomassie Blue R250. Expression was determined by the presence of a band of increased intensity present in the induced sample compared to the uninduced sample at the correct molecular weight.

4.3.4 Protein purification

The strains expressing the meningococcal GalEs (both native and mutagenised) were grown overnight with shaking in 10 ml LBB containing 100 µg/ml ampicillin. An aliquot of 100 µl was used to inoculate 10 ml LBB containing 100 µg/ml ampicillin the following day and the strains grown to \textit{OD}_{600} 0.6-0.8 at which time protein expression was induced by the addition of IPTG to a final concentration of 0.3 mM. Induction occurred for 3 hr at 37°C prior to the purification of protein using the Wizard HisLink\textsuperscript{TM} Spin Protein Purification System (Promega) as per manufacturer’s
instructions. The gonococcal GalE could not be purified in this way due to the formation of inclusion bodies. The strains expressing the gonococcal GalE proteins were grown overnight with shaking in 10 ml LBB containing 100 µg/ml Ampicillin. An aliquot of 1 ml was used to inoculate 1 L LBB containing 100 µg/ml Ampicillin the following day and the strains grown to OD₆₀₀ 0.6-0.8 at which time protein expression was induced by the addition of IPTG to a final concentration of 0.3 mM. Induction occurred for 3 hr at 37°C prior to the collection of all bacterial cells by centrifugation at 3000 x g for 15 mins at 4°C. The pellet was resuspended in 50 ml binding buffer (20 mM sodium phosphate, 0.5 M NaCl and 20 mM imidazole, pH 7.4) and sonicated 1 min on, 1 min off for 40 mins. Cell debris was removed by centrifugation and the supernatant applied to a HisTrap™ FF column. The column was washed with 10 volumes of binding buffer and the protein eluted in 10 volumes of elution buffer (20 mM sodium phosphate, 0.5 M NaCl and 500 mM imidazole, pH 7.4).

The purified proteins were dialysed against 100 volumes of dialysis buffer (10% glycerol, 20 mM Tris/HCl, pH 7.9), changed three times. The gonococcal GalE proteins were concentrated to 100µl using a centricon-10 (millipore). Protein concentrations were determined by Bradford assay and glycerol was added to a final concentration of 50% to stabilize the enzyme.

4.3.5 HPLC

Activity of the gonococcal and meningococcal GalE proteins, and the site directed mutagenised proteins was done using methods described in Dong et al., (11). Briefly 1 mM sugar substrates (UDP-Glc, UDP-Gal, UDP-GlcNAc and UDP-GalNAc) were
incubated with 200 ng of purified protein in 20 mM Tris-HCl, 4 mM Mg$^{2+}$, and 1 mM NAD$^+$ (pH 8.0) in a final volume of 50 µl. The reactions were performed at 37°C for 2 hrs and were terminated by heating at 100°C for 5 mins. The products from the reaction were run on HPLC using a UV detector and compared to UDP-Glc, UDP-Gal, UDP-GlcNAc and UDP-GalNAc standards.

4.3.6 Next generation sequencing and annotation of sequence data with BIGSDB.

Sequencing, assembly, uploading and annotation of genomic sequences was performed as described previously (4). Briefly, genomic DNA was prepared following overnight growth on Columbia horse blood agar plates (Oxoid) using a Wizard Genomic DNA Purification kit (Promega). Pooled libraries of sheared genomic DNA were subject to paired end sequencing on an Illumina Genome Analyser II platform. Genome sequence data were assembled using VELVET version 1.0.10 with optimal parameters determined by the VELVETOPTIMISER.PL script within the software package (59) with the resultant contigs uploaded into a Bacterial Isolate Genome Sequence Database (BIGSDB) along with any available provenance data. Sequence definitions were generated for galE genes in the sequence definition database and seeded with corresponding reference nucleotide sequences. Iterative searches using progressively decreasing stringency settings by means of BLASTN algorithms (2) identified likely galE genes which were then tagged in the database enabling them to be extracted and exported for further analysis. Arbitrary allele numbers were assigned to each unique sequence for a given locus. Region D which contains the LPS synthesis genes including galE is duplicated among meningococci in another region known as Region D’.
galE ‘gene in Region D’ is however truncated and is characterised from other galE genes using flanking DNA sequences. Thus, the sequence definition for the pseudo galE’ gene is distinct. In addition, the location of either galE gene was checked before the genome was tagged as both genes flank separate regions of the capsule locus. Gene sequences were exported as XMFA files containing aligned sequence blocks and then converted to a fastA format for import into MEGA version 5.0 and Splitstree (20, 48). All galE sequences are accessible through the PubMLST database (http:pubmlst.org/neisseria), and are listed in Appendix B. In addition to the common gene name, this locus is assigned a value-free nomenclature (NEIS0048) following on from the FAM18 genome annotation but using the prefix NEIS instead of NMC.

4.3.7 Sequence analysis and phylogenetic tree construction

ClustalW2 was used to align GalE protein sequences of different species. The Neisserial sequences of galE were aligned using MEGA ver 4, and a phylogenetic tree constructed using minimum evolution parameters based on nucleotide sequences.

4.3.8 Complementation of GalE using pCMK666 shuttle vector

The plasmids pCMK730 and pCMK771 were digested with XbaI and BamHI to liberate the His-tagged galE sequences complete with a ribosome binding site. The fragment was polished with T4 DNA polymerase (NEB) and ligated into the HincII site of pHSG576, resulting in pCMK773 and pCMK774 respectively. The region containing the Neisserial ompR promoter (PompR) was amplified from pCMK668 using primer pair RP (5’-AGCGGATAACAATTTTCACACAGGA-3’) and UP (5’-
GTTTCCCAGTCACGAC-3’), and cloned into the BamHI and EcoRI digested pCMK773 and pCMK774, resulting in pCMK775 and pCMK776. The fragments containing P_{ompR:galE} were amplified from pCMK775 and pCMK776 using primer pair RP and UP and the polished amplicons were cloned into the HincII site of the shuttle vector pCMK666 as described previously (33). The insertion of the cassette was confirmed by PCR with primer pair RP and UP and directionality was confirmed by restriction digestion with EcoRV. Expression of GalE from the shuttle vector in *E. coli* was confirmed by western blot. pCMK777 and pCMK778 were transformed into NMB, SS3 and FAM18 and transformants identified by resistance to kanamycin, checked by PCR with primer pair RP and UP and confirmed by plasmid extraction. NMB+pCMK777 was called CKNM550, NMB+pCMK778 was called CKNM551, SS3+pCMK777 was called CKNM552, SS3+pCMK778 was called CKNM553, FAM18+pCMK777 was called CKNM554 and FAM18+pCMK778 was called CKNM555.

### 4.3.9 Transformation protocols

*E. coli* was transformed as described previously (10). Strains NMB, SS3 and FAM18 was transformed via natural transformation as described previously (22).

### 4.3.10 SDS-PAGE and western immunoblot

Detection of the expression of his-tagged GalE and PilE was determined by western immunoblotting. Whole cell lysates (750 ng) were separated by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by standard methods and
transferred to nitrocellulose membranes. The membranes were blocked overnight with 2% BSA in TBS. The polyclonal rabbit anti-PilE IgG (46) and the monoclonal mouse anti-His IgG (Sigma) primary antibodies were used at 1:1,000. Horse radish peroxidase-conjugated anti-Rabbit IgG and anti-Mouse IgG secondary antibodies (Santa Cruz Biotechnology) was used for detection and the membrane was developed with an ECL kit (GE Healthcare).

4.3.11 Tricine-SDS-PAGE analysis of LOS preparations.

A mini Protean apparatus (Bio-Rad) was used for Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) performed according to the description of Schagger and von Jagow (40). Crude LOS preparations were prepared from whole-cell lysates of meningococcal growth suspended in distilled water. The protein concentrations of these preparations were approximated by the Bradford assay (Bio-Rad). Proteinase K digests consisted of 1 µg of protein in 2% SDS (total volume, 10 µl) to which 2 µl of 25 mg/ml proteinase K (Sigma) was added, and the mixture was incubated at 55°C for 30 min. A second aliquot of proteinase K was added and the digestion was repeated before the reaction was stopped with the addition of 25 µl of loading buffer (1 M Tris [pH 8.45], 10% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.05% bromophenol blue). A 4 µl aliquot of each sample was heated to 100°C for 4 min before being loaded onto a 16% Tricine–SDS-PAGE gel. After electrophoresis, the gels were fixed in 40% ethanol–5% acetic acid overnight. The gels were silver stained according to the method of Hitchcock and Brown (18).
4.3.12 Attachment and invasion assays

Detroit 562 epithelial cells were grown to confluence in 24-well plates and utilised for invasion assays. Invasion assays were conducted as described previously (58) with the following modifications. The epithelial cells were inoculated with suspensions of meningococci at a multiplicity of infection (MOI) of 100:1 in epithelial cell media containing 2% FCS (57). The epithelial cell monolayers were challenged with strains of *N. meningitidis* for either 1 hr or 6 hrs as indicated, and incubated in 5% CO$_2$ at 37°C. Non-adherent bacteria were then removed by washing the monolayers three times with Dulbecco’s PBS (Invitrogen). Following 1 hr incubation in either epithelial cell media containing 2% FCS (for enumeration of cell associated bacteria) or epithelial cell media containing 2% FCS with 100 μg/ml gentamycin (Sigma) (for enumeration of intracellular, protected bacteria), the monolayers were washed three times with Dulbecco’s PBS to remove gentamycin and the epithelial cells were lysed with 1% saponin (Sigma) in epithelial cell media containing 2% FCS to release the intracellular bacteria. The number of cell associated bacteria and intracellular bacteria were enumerated by viable count. At least three biological repeats containing three technical repeats were performed for each strain. Attachment was determined as the proportion of the inoculum which had attached to epithelial cells, and invasion as the proportion of the attached bacteria which had invaded the epithelial cells. A Mann-Whitney t-test was used to determine statistically significant differences. All strains assayed were sensitive to gentamycin and resistant to 1% saponin as determined by viable count (data not shown).
4.4 Results

4.4.1 Neisseria sp. possess mono- and bi-functional epimerases which are dependent on a single amino acid in the binding pocket

The galE genes of *N. gonorrhoeae* strain FA1090 and *N. meningitidis* strain MC58 were cloned into pET15b to create NgGalE::His\textsubscript{6} and NmGalE::His\textsubscript{6}. The purified proteins were analysed for epimerase activity of the substrates UDP-Glc and UDP-GlcNAc and examined for conversion to UDP-Gal and UDP-GalNAc, respectively by HPLC (Table 4.3). NgGalE::His\textsubscript{6} could epimerise both UDP-Glc and UDP-GlcNAc indicating this protein is a bi-functional enzyme, with the equilibrium favouring the production of UDP-Gal and UDP-GalNAc at 61% and 68% respectively. In comparison, NmGalE::His\textsubscript{6} would only accept UDP-Glc as a substrate and could not epimerise UDP-GlcNAc indicating that this enzyme is mono-functional. The meningococcal GalE had an equilibrium favouring the formation of UDP-Glc at 63% of all products unlike the gonococcal enzyme.

Previous investigations of known bi-functional (group 2) and mono-functional (group 1) GalE proteins have shown that two amino acid residues in the enzyme can lead to changes in the activity of these enzymes (3, 49) (Figure 4.2). Mutation of Y299C in the *E. coli* GalE enables this enzyme to accept UDP-GlcNc as a substrate, while in *Y. enterocolitica* the mutation of either C297Y (analogous to position 299 in *E. coli*) or L136Y resulted in the loss of acceptance of UDP-GlcNAc as a substrate. *E. coli*, *N. gonorrhoeae* strain FA1090 and *N. meningitidis* strain MC58 all have a tyrosine residue at position 136 (137 for *N. meningitidis*) however *E. coli* and *N. meningitidis* have
Table 4.3 Analysis of UDP-Glucose and UDP-GlCNac 4-epimerase activity in various Neisseria strains.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>Equilibrium ratio UDP-Glucose:UDP-Gal</th>
<th>Equilibrium ratio UDP-GlCNac:UDP-GalNac</th>
</tr>
</thead>
<tbody>
<tr>
<td>NgFA1090GalE</td>
<td>UDP-Glucose</td>
<td>39:61</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>UDP-GlCNac</td>
<td>-</td>
<td>32:68</td>
</tr>
<tr>
<td>NgFA1090GalEES299F</td>
<td>UDP-Glucose</td>
<td>59:41</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>UDP-GlCNac</td>
<td>-</td>
<td>99.2:0.8</td>
</tr>
<tr>
<td>NmMC58GalE</td>
<td>UDP-Glucose</td>
<td>63:37</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>UDP-GlCNac</td>
<td>-</td>
<td>99.5:0.5</td>
</tr>
<tr>
<td>NmMC58GalEF300S</td>
<td>UDP-Glucose</td>
<td>67:33</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>UDP-GlCNac</td>
<td>-</td>
<td>73:27</td>
</tr>
<tr>
<td>NmNMB-CDC</td>
<td>UDP-Glucose</td>
<td>23:77</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>UDP-GlCNac</td>
<td>-</td>
<td>25:75</td>
</tr>
</tbody>
</table>

N. meningitidis  
Y. enterocolitica  
E. coli

The protein sequences of the known bi-functional GalE of Y. pseudotuberculosis, the known mono-functional GalE sequence of E. coli and the GalE sequences of N. gonorrhoeae strain FA1090 and N. meningitidis strain MC58 with unknown functionality were aligned using ClustalW2. The amino acids previously demonstrated to be association with substrate utilisation are indicated in yellow.

aromatic amino acids at the second location (Y299 for E. coli and F300 for N. meningitidis). Y. enterocolitica and N. gonorrhoeae have smaller amino acids at this position (C297 for Y. enterocolitica and S299 for N. gonorrhoeae). Based on these
observations we hypothesised that the S299 of GalE in *N. gonorrhoeae* and the F300 of GalE in *N. meningitidis* may determine substrate utilisation.

To test this hypothesis, we constructed NgGalES299F and NmGalEF300S using site directed mutagenesis and purified these recombinant enzymes. The purified proteins were analysed for substrate specificity using the substrates UDP-Glc and UDP-GlcNAc and analysed for the production of UDP-Gal and UDP-GalNAc by HPLC (Table 4.3). The mutation of NgGalES299F resulted in the complete loss of epimerase activity of UDP-GlcNAc while retaining the ability to convert UDP-Glc to UDP-Gal, although the equilibrium was altered to favour the formation of UDP-Glc compared to the wild-type enzyme (from 39% to 59%). In other words, this single site mutation converted the wild-type bi-functional enzyme to a mono-functional state and also altered the equilibrium state. The mutation of NmGalEF300S resulted in a gain of function with the enzyme now acting as an epimerase of both UDP-Glc and UDP-GlcNAc although the observed equilibrium resulted in less UDP-GalNAc for NmGalEF300S than NgGalE::His\textsubscript{x6} (68% UDP-GalNAc for NgGalE and 27% for NmGalEF300S) suggesting other amino acid changes which differ between the two proteins may impact on the equilibria states achieved by these enzymes.

### 4.4.2 Sequence analysis and distribution of *galE* alleles across *Neisseria* genera

We assessed the distribution of *galE* alleles across the genus Neisseria using all 162 available sequences comprising 105 strains of *N. meningitidis*, 17 strains of *N. gonorrhoeae*, 18 strains of *N. lactamica*, 5 strains of *N. cinerea*, 7 strains of *N. polysaccharea*, 3 strains of *N. mucosa*, 2 strains of *N. elongata* and 1 strain each of *N.
perflava, N. subflava, N. sicca, oral taxon 14 str. and N. flavescens (The sequence of the galE allele is available in Appendix B and the distribution by isolate is shown in Appendix C). A total of 105 nucleotide sequences were detected in this collection and a phylogenetic tree was constructed based on the nucleotide sequence of galE and plotted by the minimum evolution model (Figure 4.3). The galE sequences grouped into four distinct clusters \((d=0.12)\) with 98% of meningococcal sequences clustering in regions A (expanded in Figure 4.4) and D. The 9 isolates of N. mucosa, N. elongata, N. sicca, N. subflava, oral taxon 014 str. and N. perflava clustered in region B each with a unique allele. The three alleles of N. gonorrhoeae, 14 alleles of N. lactamica, 5 alleles of N. cinerea and 6 alleles of N. polysacchareae clustered in region C. Two meningococcal sequences also clustered in region C, one belonging to strain E26 (a serogroup X strain, ST-198) and the other strain \(\alpha_{14}\) (capsule null, ST-53) which are carriage isolates. The two alleles of galE which were most dissimilar to the others \((d=0.14)\) were alleles 14 and 15 of N. meningitidis found in region D. The strains carrying these alleles were \(\alpha_{153}\) (ST-60, serogroup 29E, carriage), \(\alpha_{707}\) (ST-254, serogroup 29E, carriage), 297-0 (ST-254, serogroup B, carriage) and 173 (ST-4443, non-groupable, unknown). An examination of the distribution of meningococcal clonal complexes and serogroup amongst the galE alleles generally did not reveal a correlation.

Each galE allele was translated and assigned a putative bi-functional or mono-functional role based upon the identity of the residue in the binding site pocket of the enzyme. Those GalE proteins containing a phenylalanine at the position correlating with F300 (as for NmGalE) were proposed to be mono-functional while those proteins with a serine residue in the analogous position were proposed to be bi-functional (Figure 4.3
Based upon this sorting strategy, 53 alleles were assigned mono-functional activity and 44 alleles were bi-functional. Eight GalE alleles had neither phenylalanine nor serine at position 300. The GalE alleles possessed by *N. mucosa* (allele 57, 70, 59), *Neisseria* oral taxon 014 str F0314 (allele 105) and *N. elongata* (alleles 32 and 71) had a valine at the position analogous with F300, while two outlier meningococcal sequences (alleles 14 and 15) had a cysteine residue. Since a cysteine residue is present at the analogous position in the bi-functional *C. jejuni* GalE sequence (5), and since serine is semi-conserved with cysteine, these GalE alleles were proposed to be bi-functional. This designation was supported by the observation that the GalE allele 14 which contained a cysteine residue was carried by strains α-153 (ST-60) and α-707 (ST-254) both of which possess a serogroup 29E capsule that contains GalNAc. Most meningococcal strains possess a mono-functional GalE (85/105) with the distribution having no correlation with MLST or serogroup. However, isolates from the ST-254 (2/2 strains) and ST-8 cc (8/10 strains) were more likely to possess a bi-functional *galE* allele, while clonal complex ST-11 (which is closely related to ST-8) did not have any bi-functional alleles present in this analysis (0 of 13 isolates). All strains of *N. meningitidis* that express a capsule containing GalNAc possessed a bi-functional GalE
(2/2 strains) however many of the strains possessing a bi-functional GalE did not express a GalNAc containing capsule (18/20) and the function of GalNAc in these strains is unknown.

4.4.3 Characterisation of a putatively bi-functional GalE allele of *N. meningitidis* strain NMB-CDC

The prototype strain of ST-8, NMB-CDC (38) is one of the eight ST-8 strains predicted to have a bi-functional GalE (allele 85) due to the serine present at amino acid 300. Although the sequence of GalE allele 85 was presumptively predicted to encode a bi-functional epimerase, GalNAc has not been observed as a component of any glycan expressed by this isolate (19, 26-30, 45, 51-54). Apart from the serine 300 residue in the active site pocket, this GalE allele is 84% identical and 90% similar to GalE from *N. gonorrhoeae* strain FA1090 which equates to 24 non-conserved amino acid positions (including the binding pocket residue). Therefore, the potential existed that the significant differences in amino acid identity may have masked other differences in enzyme architecture which could determine bi-functionality. To establish whether GalE
allele 85 encoded a bi-functional UDP-GlcNAc 4-epimerase, it was cloned from NMB-CDC into pET15b, purified and substrate usage examined as before (Table 4.3).

This analysis revealed that the NMB-CDCGalE had bi-functional activity and was able to accept both UDP-Glc and UDP-GlcNAc as substrates, epimerising them to UDP-Gal and UDP-GalNAc respectively. Additionally, the equilibrium of the NMB-CDC GalE favours the production of galactose based products (77% and 75% respectively) much like the gonococcal GalE (Table 4.3).

4.4.4 Inactivation of the bi-functional GalE allele 85 in strain NMB-CDC results in increased pilin mobility and is complemented by the mono-functional GalE. 

The pilin of *N. meningitidis* is glycosylated with at least a single trideoxyhexose (either glyceramido acetamido trideoxyhexose [GATDH] (9) or 2,4-diacetamido-2,4,6-trideoxyhexose [DATDH] (44)) and at most a trisaccharide consisting of the trideoxyhexose residue with either two galactose (44) or two glucose residues (7). Based upon the genotype, the glycan of NMB-CDC is proposed to consist of GATDH and two Gal residues. However the pilin glycan of a strain of *N. meningitidis* that can synthesise GalNAc has not been assessed. To examine whether the synthesis of UDP-GalNAc affected the pilin glycosylation pattern of strain NMB-CDC, we compared pilin migration of NMB-CDC and the *galE* mutant derivative SS3 (42). SS3 was complemented with pCMK777, a plasmid expressing *NmGalE::His*$_{6}$ (SS3+GalEMono) and with pCMK778, a plasmid expressing NMB-CDCGalE::His$_{6}$ (SS3+GalEBi). The pilin glycosylation profile of these strains was examined by SDS-PAGE and western
immunoblot (Figure 4.5). The expression of His-tagged GalE was assessed in each strain with a monoclonal anti-His antibody which showed that similar levels of GalE expression occurred in each isolate (Figure 4.5a). The wild-type glycosylated pilin of NMB-CDC migrated as a mixture of three bands of 16, 16.5, and 17 kDa whilst the SS3 pilin glycosylated with only GATDH migrated as a single band of 15 kDa indicating that the glycan addition had been lost. To determine if pilin migration patterns could be restored by the production of Gal or Gal and GalNAc, the pilin glycosylation profiles of SS3 were compared to SS3+GalEMono and SS3+GalEBi. Pilin migration was completely restored in SS3+GalEMono to express the three bands of 16, 16.5 and 17 kDa. SS3+GalEBi expressed the three bands found in the wild-type strain NMB-CDC, however, the truncated 15 kDa band was still present (Figure 4.5b). A similar phenomenon was noted for the LOS profiles of these strains. SS3 expresses a truncated LOS structure of 4 kDa while strain NMB-CDC expressed a wild-type structure of 6.5 kDa (Figure 4.5c). SS3+GalEMono had a restored wild-type LOS profile, while the SS3+GalEBi expressed a mixture of truncated and wild-type LOS. Although LOS is phase variable in many meningococcal isolates, this is not the case in strain NMB-CDC which does not possess a phase variable lgtA (see Chapter 3) (24). Therefore, the partial restoration of the expression of wild-type LOS (6.5 KDa) in SS3 complemented with the bi-functional GalE cannot be ascribed to a phase variation event in the LOS biosynthesis pathway. Therefore, the lack of complete restoration of the glycan profiles in the SS3 strain expressing the bi-functional GalE protein may be due to changes in the abundance of UDP-Glc, UDP-Gal and UDP-GlcNAc with alterations to the physiological concentrations of these sugars able to alter glycosylation profiles.
4.4.5 Effect of GalNAc on attachment and invasion of NMB-CDC

Gram-negative bacteria can produce a range of glycoproteins which is particularly evident in *Campylobacter jejuni* and *Helicobacter pylori* (31, 32). It is becoming increasingly evident that many other outer membrane proteins of *Neisseria spp.* (56) are also glycosylated. Strain SS3, lacking all Gal and GalNAc residues, attached at rate of 10% (p<0.0001) and invaded at a rate of 50% (p=0.0043) compared to wild-type NMB-
CDC (Figure 4.6). Complementation of SS3+GalEMono increased the rate of attachment to 50% of the wild-type strain, while the rate of invasion was restored to wild-type levels. Complementation of SS3+GalEBi resulted in the restoration of attachment and invasion rates to wild-type levels. This suggests that the synthesis of UDP-GalNAc in NMB-CDC may impact on attachment rates, but does not appear to be required for invasion.

![Graph showing attachment and invasion rates](image)

**Figure 4.6 The rates of attachment and invasion of host epithelial cells is hindered in the absence of a UDP-Glu/GlcNAc 4-epimerase.**

The rates of association and invasion of strains NMB, SS3, SS3+GalEMono and SS3+GalEBi into Detroit 562 epithelial cells were assessed. Attachment as the percentage of the inoculum (black bars read off the left y-axis) and invasion as the percentage of the associated bacteria (white bars read off the right y-axis) is shown. The average rate (+/−SEM) from three biological repeats in triplicate following 6 hrs co-incubation is shown. *: p<0.005 determined by Mann-Whitney t-test, ND: Not Detected, viable counts were below the limit of detection.

### 4.5 Discussion

GalE proteins can be classified into three groups based on substrate binding and utilisation (21). Group 1 epimerases catalyze the conversion between UDP-Glc and UDP-Gal, Group 2 epimerases can epimerise both UDP-Glc/UDP-Gal and UDP-
GlcNAc/UDP-GalNAc while Group 3 epimerases specifically convert between UDP-GlcNAc and UDP-GalNAc. Ishiyama et al. (21) found that these differences in the ability of each group to epimerise these substrates was caused by the ability of the epimerase to bind and rotate the relevant sugars in the binding pocket. The binding site pocket of the Group 2 and 3 epimerases allows the rotation of UDP-linked-$N$-acetylated sugars. However, Group 1 enzymes either exclude UDP-linked-$N$-acetylated sugars (Group 1A) or bind these sugars but not permit rotation within the binding pocket (Group 1B), a process required for epimerisation.

On examination, *N. gonorrhoeae* was shown to possess a Group 2 GalE epimerase that catalyses the conversion of UDP-GlcNAc to UDP-GalNAc. This was not unexpected since all strains of *N. gonorrhoeae* have the capacity to add GalNAc to the terminus of LOS when the GalNAc transferase, LgtD is expressed in the presence of LgtA. In comparison, the GalE allele (allele 2) from *N. meningitidis* strain MC58 was shown to be mono-functional, incapable of epimerising UDP-GlcNAc to UDP-GalNAc and thus belongs to Group 1. NmGalE could be converted to a bi-functional epimerase capable of accepting UDP-GlcNAc by exchanging the phenylalanine at position 300 with serine (Table 4.3). This is consistent with the concept advanced by Ishiyama et al. (21) that changes in the amino acid at this position of the binding site pocket would determine the ability of the UDP-linked $N$-acetyl sugar to rotate to undergo epimerisation. Thus it appears that the GalE allele 2 from *N. meningitidis* strain MC58 encodes a Group 1B enzyme.

Having determined that the GalE alleles can encode either Group 1B or Group 2 epimerases, the distribution of the GalE alleles in *Neisseria sp.* was examined. The
distribution of the galE alleles generally correlated with individual species however, four distinct phylogenetic groups: N. meningitidis (regions A and D, Figure 4.3 and 4.4), N. mucosa, N. elongate, N. sicca, N. subflava, N. perflava and Neisseria sp. oral taxon 14 in region B and N. cinerea, N. flavescens, N. lactamica, N. polysaccharea and N. gonorrhoeae in region C were observed. Although Region A contained most of the meningococcal strains, there were two outliers (allele 14 and 15) representing two meningococcal carriage isolates found in Region D. Each of the alleles were designated as either mono- or bi-functional based upon the identity of the amino acid at position 300. The bi-functional GalE alleles predominate in N. lactamica, N. gonorrhoeae and N. polysacchareae in region C. Although alleles of N. cinerea also clustered in Region C, alleles 53, 56, 55 and 30 were mono-functional and allele 54 was bi-functional. Region A and D containing N. meningitidis contained both mono-functional and bi-functional GalE alleles. Region D contained two alleles 14 and 15, which corresponded with the expression of serogroup 29E and Z polysaccharides both of which require GalNAc (14). The serogroup H and Z capsules are similar in composition consisting of glycerol-3-phosphate repeating with Gal and GalNAc, respectively. The single serogroup H isolate in this study possessed mono-functional GalE allele 12 while the serogroup Z strain(s) contained the bi-functional allele 15. However, the majority of putative bi-functional alleles in N. meningitidis did not correspond with the possession of a GalNAc containing capsule.

The putative bi-functional GalE alleles in N. meningitidis had other changes in the protein sequence when compared to the bi-functional enzymes from N. gonorrhoeae which may have affected their functionality. To confirm that the putative bi-functional activity of GalE alleles found in N. meningitidis without serogroup Z capsules, the GalE
allele 85 was cloned from *N. meningitidis* strain NMB-CDC and tested. The enzyme assay confirmed that GalE allele 85 was indeed bi-functional. This observation correlates with previous evidence on this strain acquired by Lee and co-workers (34) who detected UDP-GlcNAc epimerase activity in whole cell lysates of this strain and proved that this activity was lost following the mutagenesis of *galE*. However, Lee et al (34) complemented the GalE mutant SS3 with GalE allele 91 from strain FAM-20 and did not see restoration of the expression of UDP-GalNAc. However, our study shows that GalE allele 91 is mono-functional and therefore was unable to complement allele 85. Together, these data conclude that the bi-functional alleles of GalE in *N. meningitidis* that do not possess a serogroup Z capsule do epimerise UDP-GlcNAc to UDP-GalNAc *in vitro* and *in situ*.

Strain NMB-CDC has been extensively characterised for LOS and capsule biosynthesis (19, 26-30, 45, 51-54) and shown to have no GalNAc in these structures, hence we examined whether the pilin glycosylation pathway may be affected. The NMB-CDC GalE mutant SS3 was complemented with the bi-functional GalE allele 85 and the mono-functional allele 2 from strain MC58. The complementation of SS3 with the wild-type bi-functional GalE allele 85 resulted in partial restoration of pilin glycosylation and LOS α-chain synthesis while the presence of the mono-functional GalE allele 2 resulted in complete restoration of both glycans, thus it appears that UDP-GalNAc cannot serve as a substrate for pilin glycan biosynthesis even though UDP-GlcNAc can do so (15).

Since GalNAc attached to LOS in *N. gonorrhoeae* contributes to attachment and invasion into host cells, we examined whether or not the expression of the bi-functional GalE allele 85 and the mono-functional GalE allele 2 would lead to a difference in the
ability of *N. meningitidis* strain NMB-CDC to attach to and invade host cells. The GalE mutant, SS3, had significantly reduced rates of attachment and invasion compared to parental wild-type NMB-CDC. Interestingly, complementation by the bi-functional GalE allele 85 fully restored attachment and invasion rates to the SS3 mutant, even though the restoration of the LOS and pilin glycan profiles were partial. Conversely, complementation by the mono-functional GalE allele 2, partially restored attachment and fully restored invasion of this mutant strain. In all, this data would suggest that GalNAc on the surface of the bacterial cell may affect attachment to host cells. However, the site and location of GalNAc on the bacterial cell surface remains to be defined.

This study has confirmed the endogenous biosynthetic pathway of GalNAc in some *Neisseria* sp. While *N. gonorrhoeae* possesses a bi-functional enzyme from Group 2, *N. meningitidis* possesses both Group 2 and Group 1B mono-functional enzymes. Although meningococci express capsular serogroups containing GalNAc, many of the isolates with a bi-functional GalE allele did not possess these serogroups. Therefore, the purpose of UDP-GalNAc biosynthesis in these strains remains unclear. Although no definitive role for GalNAc was found, the biosynthesis of this molecule integral for LOS modification in *N. gonorrhoeae* and capsule synthesis of serogroup 29E and Z strains of *N. meningitidis* was determined.

### 4.6 Acknowledgments

This publication made use of the Neisseria Multi Locus Sequence Typing website ([http://pubmlst.org/neisseria/](http://pubmlst.org/neisseria/)) developed by Keith Jolley and sited at the University of
4.7 References


11. Dong, S., O. N. Chesnokova, C. L. Turnbough, Jr., and D. G. Pritchard. 2009. Identification of the UDP-N-acetylglucosamine 4-epimerase involved in


that delivers to the hepatic asialoglycoprotein receptor, but avoids receptor-mediated endocytosis. Pharm. Res. 19:1736-1744.


Chapter 5. General Discussion
5.1. Implications to the model of attachment and invasion of *N. meningitidis*

The model by which we understand the process of invasion of host epithelial cells by meningococci occurs in a series of steps, Figure 5.1 (7). First, the type IV pili anchor the meningococcus to the host cell. The pili then retract drawing the organism into close contact with the apical membrane of the eukaryotic cell, where an intimate adhesion is elicited, primarily between the Opa proteins of the meningococcus with CEACAMs of the host cell although other adhesin/receptor pairs are involved (1). This process of attachment is inhibited by the expression of capsule and lacto-\(N\)-neotetraose of LOS, and consequently variants of meningococci in which these structures are phased off are selected (11, 16). Bacterial cell division occurs on the mucosal surface resulting in the formation of micro-colonies which then disperse over the mucosal surface (3). Invasion proceeds and although it is known that the expression of capsule and lacto-\(N\)-neotetraose on LOS reduces the rate of invasion, the expression of both of these structures is imperative to the survival of the organism within the host through conferral of resistance to the host immune response (5).

Phase variation of the capsule (via the polysialyltransferase *synD*) and lacto-\(N\)-neotetraose (via the GlcNAc glycosyltransferase *lgtA*) is an essential component of attachment and invasion according to this model, which was primarily determined using strains of the ST-32 cc. In recent years, many more genome sequences have become available and analysis of this data (shown in this work and (8)) have demonstrated that a large proportion of all strains are incapable of phase variation of *lgtA*. In addition, the rate of phase variation differs between strains, ranging from fast to slow, with strain
NMB shown to phase vary at very low rates (13). Given phase variation is a key contributor to attachment and invasion in the current model of attachment and invasion, another mechanism must be utilised by strains which cannot phase vary these structures.

Strain NMB was shown to have a similar amount of surface capsule as strain MC58 however, as it was also shown to have a surface area nearly twice as large, this means that the concentration of the capsule per surface area is less in NMB than in MC58. Consequently the removal of the capsule affected the rates of attachment and invasion less in NMB than in MC58 and even in wild-type strains NMB attaches and invades host epithelial cells at higher rates than MC58. In response to contact with host cells, the expression of capsule synthesis and transport (synA and ctrA respectively) was

![Diagram of attachment and invasion of N. meningitidis](image)

**Figure 5.1 The model for attachment and invasion of N. meningitidis.**

The type IV pill of the meningococcus elicit the initial interaction with the host epithelial cells. The variants which have phased off capsule and the LNT of LOS (orange) attach and invade host epithelial cells at greater rates than the encapsulated and LNT+ (blue) variants through an Opa:CEACAM mechanism. The capsule and LNT are required for survival within the host. Adapted from (7).
reduced in strain NMB, which has also been seen in other isolates (6). Taken together, although NMB does not phase vary the expression of capsule, it does reduce the expression in response to contact with host cells and it has a lower density of capsule compared to MC58, an ST-32, a complex of strains observed to phase vary capsule expression (2).

The LNT of LOS is thought to inhibit attachment and invasion by the binding of particularly the LNT (particularly when sialylated) to the Opa proteins, hence phase variation of lgtA which results in the absence of the LNT of LOS relieves the inhibition of Opa binding (11). The LOS inner core is fixed in some strains and variable in others, dependent upon the carriage of genetic islands carrying the transferases lgtG and lpt6 and since this genetic island is not associated with any integrases, it appears these island are simply spread in the bacterial population via natural transformation. (this study, (8)). Alterations to the LOS inner core structure influence the rate of invasion into host epithelial cells but not attachment in strain NMB, specifically, the presence of the O-6 PEA attached to HepII improved invasion into host cells even in the presence of capsule. The expression of the LOS inner core transferase lpt6 and the LOS sialyl transferase lst are regulated in response to contact with host epithelial cells indicating a remodelling of the bacterial cell surface in response to host cell contact.

In the related pathogen *N. gonorrhoeae*, the α-chain of the LOS engages molecules on the epithelial cell surface and leads to receptor mediated endocytosis. We have not examined the mechanism by which alterations to the LOS inner core result in increased rates of invasion but hypothesise that changes in the inner core could alter the
presentation of the α-chain which could participate in receptor mediated uptake, similar to that observed for *N. gonorrhoeae*.

A new model for the understanding of meningococcal attachment and invasion is presented here, particularly for strains which lack the capability of phase variation of capsule and LNT of LOS. Although no changes in the role of pilin could be determined, in response to host cell contact the capsule is down regulated and LOS structure remodelled such that increased rates of attachment and invasion are achieved, Figure 5.2. These results have only been demonstrated with one strain and hence future work should investigate multiple strains of various complexes to extend these findings.

![Diagram](image)

**Figure 5.2 Proposed alternative model for attachment and invasion of *N. meningitidis*.

The type IV pili of the meningococcus elicit the initial interaction with the host epithelial cells. In response, the bacteria remodel the cell surface resulting in decreased capsule (decreased thickness of the red), a change in the LOS structure and a reduction in the amount of sialic acid on the LOS (from black to purple). Engagement of CEACAM by the meningococcal Opa proteins and of a putative LOS receptor leads to internalisation of the meningococci.
5.2 Variation in cell size and surface hydrophobicity in the interaction with the host

During this study two phenotypes that contribute to bacterial variation, surface hydrophobicity and bacterial cell size, were also examined. Surface hydrophobicity is a complex property of surfaces and is related to the relative concentration, distribution and configuration of polar and non-polar molecules (sugars, lipids and amino acids) on the bacterial outer surface (14). In many bacteria, surface hydrophobicity is related to the presence of capsule and outer membrane proteins (9, 15). The relative level of surface hydrophobicity has been implicated in diverse properties usually adhesion to various biotic and abiotic surfaces (10, 18). In this study of two isolates of meningococci, the two strains were shown to be different, with the less adherent strain MC58 having a more hydrophilic surface, while the hydrophobic isolate strain NMB was more adherent. The presence of capsule polysaccharide was shown to modulate surface hydrophobicity in both strains, but to a lesser degree in strain NMB than strain MC58, which again led to the hypothesis that the capsule was more sparse on the surface of strain NMB. However, further work is needed to establish the veracity of this model including whether or not other properties of the surface of strain NMB are involved in this phenotype.

The variation in cell size between strains of *N. meningitidis* is a novel observation. It was observed that metabolic alterations could result in differences in bacterial cell size within a strain of *Bacillus subtilis* (17), however no variation in size within a species has been reported. Although no metabolic differences between strains NMB and MC58 have been reported previously, this study has identified a difference in the biosynthesis
of UDP-GalNAc, with NMB capable of synthesising this amino sugar, while MC58 cannot. It has not yet been determined whether this alteration in metabolic pathways could influence cell size. The variation of bacterial cell size in *Streptococcus pneumoniae* has been reported by Dalia and Weiser (4) who characterised the role of chain length in uptake into neutrophils and complement deposition amongst others. In their study, they showed a positive correlation between chain length, complement deposition and uptake into neutrophils. The biological role and basis for differences in cell size in *N. meningitidis* is yet to be determined.

5.3 Synthesis of *N*-acetyl galactosamine in *Neisseria*

The meningococcal serogroup 29E and Z capsules contain the sugar GalNAc. This sugar, which has not been identified elsewhere in *N. meningitidis* is known to cap the gonococcal LOS and improve the interaction of the LOS with its receptors. Despite the knowledge of the presence of GalNAc, the biosynthesis has never been determined in gonococci nor meningococci. We identified the UDP-glucose 4-epimerase GalE as the bi-functional enzyme of *N. gonorrhoeae* responsible for the synthesis of GalNAc. Most strains of *N. meningitidis* possess an allele of GalE which is mono-functional and hence are incapable of synthesising GalNAc, however a subset of meningococcal strains can make GalNAc. The carriage of the bi-functional allele of *galE* did not correlate with clonal complex (except in the case of ST-8) or capsule type with the exception of all strains expressing a 29E or Z capsule possessed a bi-functional allele of GalE.

In strain NMB, which has a bi-functional GalE, GalNAc has never been identified as a component of LOS, capsule or pilin glycan and the role and purpose of GalNAc
biosynthesis in this strain remains elusive. *N. meningitidis* is a human specific pathogen and as such relies on the host to provide nutrients, including the carbon source. The possession of a bi-functional allele GalE may expand the carbon source range to include UDP-GalNAc, similar to that observed for *Leishmania major* (12).

### 5.4 Conclusion

In conclusion, although the expression of carbohydrates in *N. meningitidis* is variable between strains, the expression impacts on the interaction of the meningococcus with host epithelial cells. Further comparisons primarily between hypervirulent and commensal strains should be conducted to determine if glycan repertoire, role or regulation could explain variability in pathogenicity.

### 5.2 References


contact with epithelial cells is mediated by CrgA regulatory protein. Molecular microbiology 43:1555-1564.


Appendix A. Abbreviations

A(I) Nucleotides

A: Adenine  
T: Thymine  
G: Guanine  
C: Cytosine  
U: Uracil

A(II) Amino Acids

Small

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Molecular Structure</th>
<th>MW</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine (Gly, G)</td>
<td>H₂N-CH₂-COOH</td>
<td>75.06</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Alanine (Ala, A)</td>
<td>H₂N-CH₂-COOH</td>
<td>89.09</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Serine (Ser, S)</td>
<td>H₂N-CH₂-COOH</td>
<td>114.09</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Threonine (Thr, T)</td>
<td>H₂N-CH₂-COOH</td>
<td>117.11</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Cysteine (Cys, C)</td>
<td>H₂N-CH₂-COOH</td>
<td>121.15</td>
<td>H₂N-CH₂-COOH</td>
</tr>
</tbody>
</table>

Hydrophobic

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Molecular Structure</th>
<th>MW</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine (Val, V)</td>
<td>H₂N-CH₂-COOH</td>
<td>99.14</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Leucine (Leu, L)</td>
<td>H₂N-CH₂-COOH</td>
<td>113.16</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Isoleucine (Ile, I)</td>
<td>H₂N-CH₂-COOH</td>
<td>113.16</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Methionine (Met, M)</td>
<td>H₂N-CH₂-COOH</td>
<td>131.19</td>
<td>H₂N-CH₂-COOH</td>
</tr>
<tr>
<td>Proline (Pro, P)</td>
<td>H₂N-CH₂-COOH</td>
<td>97.12</td>
<td>H₂N-CH₂-COOH</td>
</tr>
</tbody>
</table>

Aromatic

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Molecular Structure</th>
<th>MW</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylalanine (Phe, F)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>147.18</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
<tr>
<td>Tyrosine (Tyr, Y)</td>
<td>H₂N-C₆H₄-OH</td>
<td>165.18</td>
<td>H₂N-C₆H₄-OH</td>
</tr>
<tr>
<td>Tryptophan (Trp, W)</td>
<td>H₂N-C₆H₅-NH-C₆H₄-COOH</td>
<td>186.21</td>
<td>H₂N-C₆H₅-NH-C₆H₄-COOH</td>
</tr>
<tr>
<td>Aspartic Acid (Asp, D)</td>
<td>H₂N-C₆H₄-COOH</td>
<td>113.09</td>
<td>H₂N-C₆H₄-COOH</td>
</tr>
<tr>
<td>Glutamic Acid (Glu, E)</td>
<td>H₂N-C₆H₄-COOH</td>
<td>129.12</td>
<td>H₂N-C₆H₄-COOH</td>
</tr>
</tbody>
</table>

Amide

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Molecular Structure</th>
<th>MW</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asparagine (Asn, N)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>114.11</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
<tr>
<td>Glutamine (Gln, Q)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>128.14</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
<tr>
<td>Histidine (His, H)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>137.14</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
<tr>
<td>Lysine (Lys, K)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>128.17</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
<tr>
<td>Arginine (Arg, R)</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
<td>156.19</td>
<td>H₂N-C₆H₄-CH₂-COOH</td>
</tr>
</tbody>
</table>

Taken from:


#.UMbK54bF3XQ
### A(III) Units

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Prefix</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pico</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>n</td>
<td>nano</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>μ</td>
<td>micro</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>k</td>
<td>kilo</td>
<td>$10^{3}$</td>
</tr>
<tr>
<td>M</td>
<td>mega</td>
<td>$10^{6}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>degree celsius</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>hr</td>
<td>hour</td>
</tr>
<tr>
<td>hrs</td>
<td>hours</td>
</tr>
<tr>
<td>L</td>
<td>litre</td>
</tr>
<tr>
<td>m</td>
<td>metre</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mins</td>
<td>minutes</td>
</tr>
<tr>
<td>mol</td>
<td>mole</td>
</tr>
<tr>
<td>v</td>
<td>volume</td>
</tr>
<tr>
<td>w</td>
<td>weight</td>
</tr>
</tbody>
</table>

### A(IV) Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>Δ</td>
<td>delta</td>
</tr>
<tr>
<td>λ</td>
<td>lambda</td>
</tr>
<tr>
<td>μ</td>
<td>mu</td>
</tr>
<tr>
<td>π</td>
<td>pi</td>
</tr>
<tr>
<td>Ω</td>
<td>omega</td>
</tr>
<tr>
<td>+</td>
<td>positive/present</td>
</tr>
<tr>
<td>-</td>
<td>negative/absent</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>III</td>
<td>3</td>
</tr>
<tr>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>V</td>
<td>5</td>
</tr>
<tr>
<td>²</td>
<td>squared</td>
</tr>
<tr>
<td>³</td>
<td>cubed</td>
</tr>
<tr>
<td>&lt;</td>
<td>less than</td>
</tr>
<tr>
<td>&gt;</td>
<td>greater than</td>
</tr>
<tr>
<td>=</td>
<td>equal to</td>
</tr>
<tr>
<td>%</td>
<td>percent</td>
</tr>
</tbody>
</table>
### A(V) Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{32}$P</td>
<td>radioactive phosphorous with weight of 32 amu</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>AF</td>
<td>alexafluor</td>
</tr>
<tr>
<td>amu</td>
<td>atomic mass units</td>
</tr>
<tr>
<td>Anhydro.</td>
<td>anhydrous</td>
</tr>
<tr>
<td>ASGP-r</td>
<td>asialoglycoprotein receptor</td>
</tr>
<tr>
<td>ATCC</td>
<td>American tissue culture collection</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine tri-phosphate</td>
</tr>
<tr>
<td>BBB</td>
<td>blood-brain barrier</td>
</tr>
<tr>
<td>BD</td>
<td>beckton Dickenson</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CAMP</td>
<td>cationic antimicrobial peptide</td>
</tr>
<tr>
<td>cap</td>
<td>capsule</td>
</tr>
<tr>
<td>cc</td>
<td>clonal complex</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CDS</td>
<td>coding sequences</td>
</tr>
<tr>
<td>cfu</td>
<td>colony forming unit</td>
</tr>
<tr>
<td>CMP</td>
<td>cytosine mono-phosphate</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>cps</td>
<td>capsule locus</td>
</tr>
<tr>
<td>CR3</td>
<td>complement receptor 3</td>
</tr>
<tr>
<td>CSF</td>
<td>cerebral spinal fluid</td>
</tr>
<tr>
<td>C$_T$</td>
<td>threshold cycle</td>
</tr>
<tr>
<td>CTAB</td>
<td>cetyl trimethylammonium bromide</td>
</tr>
<tr>
<td>DATDH</td>
<td>2,4-diacetamido-2,4,6-trideoxyhexose</td>
</tr>
<tr>
<td>DHB</td>
<td>2,5-dihydroxybenzoic acid</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DOC</td>
<td>deoxycholic acid</td>
</tr>
<tr>
<td>D-PBS</td>
<td>Dulbecco’s PBS</td>
</tr>
<tr>
<td>ECL</td>
<td>electrochemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamineeta-acetic acid</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>ET</td>
<td>electrophoretic type</td>
</tr>
<tr>
<td>FCS</td>
<td>fetal calf serum</td>
</tr>
<tr>
<td>Fe(NO$_3$)$_3$</td>
<td>iron (III) nitrate</td>
</tr>
<tr>
<td>FSC</td>
<td>forward scatter</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>N-acetyl galactosamine</td>
</tr>
<tr>
<td>GATDH</td>
<td>glyceramido acetamido trideoxyhexose</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>GC</td>
<td>gonococcal</td>
</tr>
<tr>
<td>GCA</td>
<td>gonococcal agar</td>
</tr>
<tr>
<td>GCB</td>
<td>gonococcal broth</td>
</tr>
<tr>
<td>gDNA</td>
<td>genomic DNA</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetyl glucosamine</td>
</tr>
<tr>
<td>Glu</td>
<td>glucose</td>
</tr>
<tr>
<td>Gly</td>
<td>glycine</td>
</tr>
<tr>
<td>GLY</td>
<td>glycosylated</td>
</tr>
<tr>
<td>GMF</td>
<td>geometric mean fluorescence</td>
</tr>
<tr>
<td>GulNAcA</td>
<td>N-acetyl-gulosaminic acid</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>Hep</td>
<td>L-glycero-D-manno-heptose</td>
</tr>
<tr>
<td>hex</td>
<td>hexose</td>
</tr>
<tr>
<td>HPLC</td>
<td>high pressure liquid chromatography</td>
</tr>
<tr>
<td>HRP</td>
<td>horse radish peroxidase</td>
</tr>
<tr>
<td>HSPG</td>
<td>heparin sulfate proteoglycan</td>
</tr>
<tr>
<td>IDC</td>
<td>I-domain containing</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IgM</td>
<td>Immunoglobulin M</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropylthio-1'-D-galactoside</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>KDO</td>
<td>2-keto D-manno octulosonic acid</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LBA</td>
<td>Luria-Bertani agar</td>
</tr>
<tr>
<td>LBB</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>LNT</td>
<td>Lacto-N-neotetraose</td>
</tr>
<tr>
<td>LOS</td>
<td>Lipooligosaccharide</td>
</tr>
<tr>
<td>LPS</td>
<td>lipopolysaccharide</td>
</tr>
<tr>
<td>MAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MALDI-TOF MS</td>
<td>Matrix assisted laser desorption ionization-time of flight mass spectrometry</td>
</tr>
<tr>
<td>ManNAc</td>
<td>N-acetyl mannosamine</td>
</tr>
<tr>
<td>ManNAcA</td>
<td>N-acetylmannosaminuronic acid</td>
</tr>
<tr>
<td>MDA</td>
<td>Meningococcal disease associated</td>
</tr>
<tr>
<td>MEM</td>
<td>minimal essential media</td>
</tr>
<tr>
<td>Mg2+</td>
<td>Magnesium divalent cation</td>
</tr>
<tr>
<td>MGL</td>
<td>macrophage galectin-like lectin</td>
</tr>
<tr>
<td>MgSO4</td>
<td>Magnesium sulfate</td>
</tr>
<tr>
<td>MLEE</td>
<td>multi-locus enzyme electrophoresis</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>n</td>
<td>number</td>
</tr>
<tr>
<td>Na2CO3</td>
<td>sodium carbonate</td>
</tr>
<tr>
<td>Na2HPO4</td>
<td>disodium hydrogen phosphate</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NAD+</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NaH2PO4</td>
<td>Sodium dihydrogen phosphate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>NANA</td>
<td>N-acetyl neuraminic acid (Neu5NAc, sialic acid)</td>
</tr>
<tr>
<td>Ncin</td>
<td>Neisseria cinerea</td>
</tr>
<tr>
<td>Ne</td>
<td>Neisseria elongata</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>Nf</td>
<td>Neisseria flavescens</td>
</tr>
<tr>
<td>Ng</td>
<td>Neisseria gonorrhoea</td>
</tr>
<tr>
<td>Nl</td>
<td>Neisseria lactamica</td>
</tr>
<tr>
<td>Nm</td>
<td>Neisseria meningitidis</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>Nmuc</td>
<td>Neisseria mucosa</td>
</tr>
<tr>
<td>Nper</td>
<td>Neisseria perflava</td>
</tr>
<tr>
<td>Npol</td>
<td>Neisseria polysaccharea</td>
</tr>
<tr>
<td>Nsic</td>
<td>Neisseria sicca</td>
</tr>
<tr>
<td>Nsub</td>
<td>Neisseria subflava</td>
</tr>
<tr>
<td>NT</td>
<td>Non-typable</td>
</tr>
<tr>
<td>OAc</td>
<td>O-acetylation</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OD&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Optical density at x nm</td>
</tr>
<tr>
<td>OG</td>
<td>Orthologous group</td>
</tr>
<tr>
<td>OMP</td>
<td>Outer-membrane protein</td>
</tr>
<tr>
<td>OMV</td>
<td>Outer membrane vesicle</td>
</tr>
<tr>
<td>ONPG</td>
<td>Ortho-Nitrophenyl-β-galactoside</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>OS</td>
<td>Oligosaccharide</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PBSB</td>
<td>PBS containing 0.5% (w/v) BSA</td>
</tr>
<tr>
<td>PC</td>
<td>Phosphorylcholine</td>
</tr>
<tr>
<td>PI</td>
<td>Phosphoinositide</td>
</tr>
<tr>
<td>PID</td>
<td>Pelvic inflammatory disease</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEA</td>
<td>Phosphoethanolamine</td>
</tr>
<tr>
<td>PG</td>
<td>Phosphoglycerol</td>
</tr>
<tr>
<td>pH</td>
<td>Power of Hydrogen</td>
</tr>
<tr>
<td>PO&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Phosphate</td>
</tr>
<tr>
<td>P&lt;sub&gt;geneX&lt;/sub&gt;</td>
<td>Promoter of geneX</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcriptase PCR</td>
</tr>
<tr>
<td>r</td>
<td>Radius</td>
</tr>
<tr>
<td>rcf</td>
<td>Relative centrifugal force</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SOE-PCR</td>
<td>Splicing with overlapping extension PCR</td>
</tr>
<tr>
<td>SSR</td>
<td>Simple sequence repeat</td>
</tr>
<tr>
<td>ST</td>
<td>Sequence Type</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>st. dev.</td>
<td>standard deviation</td>
</tr>
<tr>
<td>TBS</td>
<td>tris buffered saline</td>
</tr>
<tr>
<td>tfp</td>
<td>type IV pili</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>UDP</td>
<td>uracil di-phosphate</td>
</tr>
<tr>
<td>UMP</td>
<td>uracil mono-phosphate</td>
</tr>
<tr>
<td>UndP</td>
<td>undecaprenylphosphate</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
<tr>
<td>WHO</td>
<td>world health organisation</td>
</tr>
<tr>
<td>x</td>
<td>variable</td>
</tr>
<tr>
<td>x g</td>
<td>times gravity</td>
</tr>
</tbody>
</table>
Appendix B. Sequence of galE alleles examined in Chapter 4.

> 1
ATGAAAAAATTTCGTTAACCGGGCGGCACCGGTTTTATCGGCTCGCATACCGGTTTTGCTATT
GCGGCCCATACAGTCGATGGTTTTGTAATACTATGCAATTTCCAGCATCAAATATGCTGCCAGCTTGAAA
CGATTACCGGCCAAGAGATTCCGTTTTATCAAGGCGATATCCGCGATCGTGAGATTTTGCGCCGTATTT
TGCACCGAACATTGGCGGTATTTGGAAGAGGTGGGTGAAAGTTGCTGGATGCAGGCAAAAGTGCCTTATACG
GAAGGTGCAAGGCGGTATTTTTGCTGGATGGCGGAAGAAATGGCGCGTGCC

> 2
ATGAAAAAATTTCGTTAACCGGGCGGCACCGGTTTTATCGGCTCGCATACCGGTTTTGCTATT
GCGGCCCATACAGTCGATGGTTTTGTAATACTATGCAATTTCCAGCATCAAATATGCTGCCAGCTTGAAA
CGATTACCGGCCAAGAGATTCCGTTTTATCAAGGCGATATCCGCGATCGTGAGATTTTGCGCCGTATTT
TGCCGAAAACCGCATTGATTCGGTGATTCATTTTGCCGGCTTGAA
CGGTGGGTGAAAGTGTGGCCG

> 3
ATGAAAAAATTTCGTTAACCGGGCGGCACCGGTTTTATCGGCTCGCATACCGGTTTTGCTATT
GCGGCCCATACAGTCGATGGTTTTGTAATACTATGCAATTTCCAGCATCAAATATGCTGCCAGCTTGAAA
CGATTACCGGCCAAGAGATTCCGTTTTATCAAGGCGATATCCGCGATCGTGAGATTTTGCGCCGTATTT
TGCACCGAACATTGGCGGTATTTTGGCGCCATGATTTTTGCGCCGCTGGATGGCGGAAGAAATGGCGCGTGCC
CGCAGATGAAATATTATGCTATTGATGATTTTGCCGAGAAAGGCACGCTTGCAATTTTGCATGCTCTTAACG
CGGATGCGTGACTATATTCTATGCTGATGATTGATGATTTTGGCAGAAGGCGCATGATGCTCGCCGCTCAACG
CGGATGCGTGACTATATTCTATGCTGATGATTGATGATTTTGGCAGAAGGCGCATGATGCTCGCCGCTCAACG
CGGATGCGTGACTATATTCTATGCTGATGATTGATGATTTTGGCAGAAGGCGCATGATGCTCGCCGCTCAACG
CGGATGCGTGACTATATTCTATGCTGATGATTGATGATTTTGGCAGAAGGCGCATGATGCTCGCCGCTCAACG
TCACTGATATTCAAAAGCGGATCCGCGCCTGGAGGCTGATTGTTGCGCTATTTCACACCGGATATTGGGGCCGACAAAGCCGGCTTGATTTCCAGGATCCGCTGGAGCGTGATTGTGGCGCTATTTCAACCCGATTGGGGCGCACGAAGCGGCTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCAAGTGCTTCGGGCAGGCTGCCGCAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGTACGGGAAATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAGGCGGCGTTGCCGGCGTACATTTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGAAATCATCCGTGCCTTTTGAAGCGGCTTCGGGTTTAACCATTCCATTTGAAATCAAACCGCGCCGCGCCGGTGATTTGGCGTGTCCTTTGCCGACCCGTCCCATACCAAACAACAAACCGGCTGGGAAACCAAACGCGGCTTGCAGCAAATGATGGAAGATTCGTCGTTGGGTCAGCCGCAACCCCAACGGATATGGGGATTAA

>4

ATGAAATATTCTCGTTACCCTCGGTCGGCAACCGGCTTTTATTGGTTGCTGATCTGCTATTTCATTGTGGTGAATAGAGCCCAACCGGATCCGCTGGAGCGTGATTGTGGCGCTATTTCAACCCGATTGGGGCGCACGAAGCGGCTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCAAGTGCTTCGGGCAGGCTGCCGCAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGTACGGGAAATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAGGCGGCGTTGCCGGCGTACATTTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGAAATCATCCGTGCCTTTTGAAGCGGCTTCGGGTTTAACCATTCCATTTGAAATCAAACCGCGCCGCGCCGGTGATTTGGCGTGTCCTTTGCCGACCCGTCCCATACCAAACAACAAACCGGCTGGGAAACCAAACGCGGCTTGCAGCAAATGATGGAAGATTCGTCGTTGGGTCAGCCGCAACCCCAACGGATATGGGGATTAA

>5

ATGAAATATTCTCGTTACCCTCGGTCGGCAACCGGCTTTTATTGGTTGCTGATCTGCTATTTCATTGTGGTGAATAGAGCCCAACCGGATCCGCTGGAGCGTGATTGTGGCGCTATTTCAACCCGATTGGGGCGCACGAAGCGGCTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCAAGTGCTTCGGGCAGGCTGCCGCAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGTACGGGAAATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAGGCGGCGTTGCCGGCGTACATTTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGAAATCATCCGTGCCTTTTGAAGCGGCTTCGGGTTTAACCATTCCATTTGAAATCAAACCGCGCCGCGCCGGTGATTTGGCGTGTCCTTTGCCGACCCGTCCCATACCAAACAACAAACCGGCTGGGAAACCAAACGCGGCTTGCAGCAAATGATGGAAGATTCGTCGTTGGGTCAGCCGCAACCCCAACGGATATGGGGATTAA

>6

ATGAAATATTCTCGTTACCCTCGGTCGGCAACCGGCTTTTATTGGTTGCTGATCTGCTATTTCATTGTGGTGAATAGAGCCCAACCGGATCCGCTGGAGCGTGATTGTGGCGCTATTTCAACCCGATTGGGGCGCACGAAGCGGCTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCAAGTGCTTCGGGCAGGCTGCCGCAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGTACGGGAAATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAGGCGGCGTTGCCGGCGTACATTTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGAAATCATCCGTGCCTTTTGAAGCGGCTTCGGGTTTAACCATTCCATTTGAAATCAAACCGCGCCGCGCCGGTGATTTGGCGTGTCCTTTGCCGACCCGTCCCATACCAAACAACAAACCGGCTGGGAAACCAAACGCGGCTTGCAGCAAATGATGGAAGATTCGTCGTTGGGTCAGCCGCAACCCCAACGGATATGGGGATTAA
ACCGACATCCCCAAAAGCCGATCCGCGTTGGAGCGTGATTTTGTTGCGCTATTTCAACCCGATCGGCGC
TTGGAGCGTGATTTTGTTGCGCTATTTCAACCCGATCGGCGC
ACAAGAAAGCGGCTTGATCGGCGAACAGCCCAACGGCATCCCGACGGTACGG
GAATCCGCGGATCATTTATATTGATGGAATTGTTGGCAAGGCGCATTCCCAGCAGGATGACCGGAAAAG
GGCGCCCTTCCGCGTACATTGTTGTTCAACTGGTGTTGGCGACGCTATTCCGCTTACCTAAATCAACCCGCCGGCGCGGCGGCGGAGACCTCGCT
TGTTCTTTTGGGCCAGCCGTCCATACCAATCAAAACACGAGACGTGGGGAAGAACCGAGCCTTGGCACAAT
AGTATGGAAGATTTGGCTGGGCTTGGTCAGGCCGCAACCCCGACGGTAGCTAGGAT

>25
ATGAAAAAATTTCCTGTTACCAGCGGCGGCGACCGGTTTTATCGGCTCGCATAACCAGTTGTTTCTTCTTCTGCTGAAA
GCGGCCCATCAAGTCTGTTATTGGGATAAATCTGTAATCCAGCAGCTAATTTCTCTGGGAGGTTTGACG
CGATACCGGATATCACGATTTGGAATTCAGGAGTACCCAGTTACCCAGCCAACCGCCGAGGATCCGCTGTTT
TGTTCCTTTTGGGCCAGCCGTCCATACCAATCAAAACACGAGACGTGGGGAAGAACCGAGCCTTGGCACAAT
AGTATGGAAGATTTGGCTGGGCTTGGTCAGGCCGCAACCCCGACGGTAGCTAGGAT

>26
ATGAAAAAATTTCCTGTTACCAGCGGCGGCGACCGGTTTTATCGGCTCGCATAACCAGTTGTTTCTTCTTCTGCTGAAA
GCGGCCCATCAAGTCTGTTATTGGGATAAATCTGTAATCCAGCAGCTAATTTCTCTGGGAGGTTTGACG
CGATACCGGATATCACGATTTGGAATTCAGGAGTACCCAGTTACCCAGCCAACCGCCGAGGATCCGCTGTTT
TGTTCCTTTTGGGCCAGCCGTCCATACCAATCAAAACACGAGACGTGGGGAAGAACCGAGCCTTGGCACAAT
AGTATGGAAGATTTGGCTGGGCTTGGTCAGGCCGCAACCCCGACGGTAGCTAGGAT
AACCGACATCCAAAAAGCCGACCACCCCGCTGAGGACGTACTCCTGGTGCTATTTAATCCCGGATCCGCACGCAGACACGAGACGTGATCGGATTTGCTGAGAAAGATGACGATTAAGGATTTTATCGTTCTCTACAGTGGTTTTCCTGCTGGAATACAGACTCTTAAATGGGCTTCGATATCGGAGCCCGTTCGATTTTGCTGGAGGAGCTTCGCGGAGGCGTCCGCGGCGGAAACTCCCGGCTGATTTTGGATGCTCCGAGGCGGCGTATTTTGTTGCGGAAAACCGCATTGGATTCGGTGATTCATTTTGCCGGCTTGAAAGCGGTTGGGTGAAAGTGTGGCC
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCCGGCGTGATTTTAAAAATTGTGTTCAGCTCATCGGCTACGGTTTATGGCGATCCGGGCAAAGTGCCTTATACCGAGGATATGCCACTGGGCAGATACCACCAGCCCTTACGGCACCATCGAAATCAATGGTTGAGCGCATTC

>31
ATGACTATTCCTCTACTACCGGCCGCTACCGGTTTTTATCGTTCTCATACCGGTTTTCCTGCTGGAATACAGACTCTTAAATGGGCTTCGATATCGGAGCCCGTTCGATTTTGCTGGAGGAGCTTCGCGGAGGCGTCCGCGGCGGAAACTCCCGGCTGATTTTGGATGCTCCGAGGCGGCGTATTTTGTTGCGGAAAACCGCATTGGATTCGGTGATTCATTTTGCCGGCTTGAAAGCGGTTGGGTGAAAGTGTGGCC
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCCGGCGTGATTTTAAAAATTGTGTTCAGCTCATCGGCTACGGTTTATGGCGATCCGGGCAAAGTGCCTTATACCGAGGATATGCCACTGGGCAGATACCACCAGCCCTTACGGCACCATCGAAATCAATGGTTGAGCGCATTC

>32
ATGACTATTCCTCTACTACCGGCCGCTACCGGTTTTTATCGTTCTCATACCGGTTTTCCTGCTGGAATACAGACTCTTAAATGGGCTTCGATATCGGAGCCCGTTCGATTTTGCTGGAGGAGCTTCGCGGAGGCGTCCGCGGCGGAAACTCCCGGCTGATTTTGGATGCTCCGAGGCGGCGTATTTTGTTGCGGAAAACCGCATTGGATTCGGTGATTCATTTTGCCGGCTTGAAAGCGGTTGGGTGAAAGTGTGGCC
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCCGGCGTGATTTTAAAAATTGTGTTCAGCTCATCGGCTACGGTTTATGGCGATCCGGGCAAAGTGCCTTATACCGAGGATATGCCACTGGGCAGATACCACCAGCCCTTACGGCACCATCGAAATCAATGGTTGAGCGCATTC

>33
ATGAAAAAATTCTCGTTACCGGCCGCGAACCACGCTTTTATCGTTCTCGCATAACCGGTATTTTTTTGCAAAAACGCGCCATACAGCTGATTTTGGATAACCTATGCAACTCCGACTCATTACTTTGGGCAAAACTCCCGGCTGATTTTGGATGCTCCGAGGCGGCGTATTTTGTTGCGGAAAACCGCATTGGATTCGGTGATTCATTTTGCCGGCTTGAAAGCGGTTGGGTGAAAGTGTGGCC
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCCGGCGTGATTTTAAAAATTGTGTTCAGCTCATCGGCTACGGTTTATGGCGATCCGGGCAAAGTGCCTTATACCGAGGATATGCCACTGGGCAGATACCACCAGCCCTTACGGCACCATCGAAATCAATGGTTGAGCGCATTC
ATGAAAAAATTCTCGTTACCGGCGGCACCGGTTTTATCGGTCCACACCGTTTTTCTGCTGAAA
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGC
GTCGCGATGGCGAAGATATGCCACTGGGCGATACCACCAGCCCTTACGGCACATCGAAATCA
ATGGTTGAGCGCATTCG
CGGTGGGTGAAAGTGTGGCCGAGGCATGATTTTGTTGCGTTATTTCAATCCAATTGGCGC
GCATGAAAGCGGCTTGATTGGCGAGCAGCCAAACGGTATCCCAAATAATTTGTTGCCTTATAT
CTGCCAAGTGCGG

>40
ATGAAAAAATTCTCGTTACCGGCGGCACCGGTTTTATCGGTCCACACCGTTTTTCTGCTGAAA
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCG
GTCGCGATGGCGAAGATATGCCACTGGGCGATACCACCAGCCCTTACGGCACATCGAAATCA
ATGGTTGAGCGCATTCG
CGGTGGGTGAAAGTGTGGCCGAGGCATGATTTTGTTGCGTTATTTCAATCCAATTGGCGC
GCATGAAAGCGGCTTGATTGGCGAGCAGCCAAACGGTATCCCAAATAATTTGTTGCCTTATAT
CTGCCAAGTGCGG

>41
ATGAAAAAATTCTCGTTACCGGCGGCACCGGTTTTATCGGTCCACACCGTTTTTCTGCTGAAA
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCG
GTCGCGATGGCGAAGATATGCCACTGGGCGATACCACCAGCCCTTACGGCACATCGAAATCA
ATGGTTGAGCGCATTCG
CGGTGGGTGAAAGTGTGGCCGAGGCATGATTTTGTTGCGTTATTTCAATCCAATTGGCGC
GCATGAAAGCGGCTTGATTGGCGAGCAGCCAAACGGTATCCCAAATAATTTGTTGCCTTATAT
CTGCCAAGTGCGG

>42
ATGAAAAAATTCTCGTTACCGGCGGCACCGGTTTTATCGGTCCACACCGTTTTTCTGCTGAAA
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCG
GTCGCGATGGCGAAGATATGCCACTGGGCGATACCACCAGCCCTTACGGCACATCGAAATCA
ATGGTTGAGCGCATTCG
CGGTGGGTGAAAGTGTGGCCGAGGCATGATTTTGTTGCGTTATTTCAATCCAATTGGCGC
GCATGAAAGCGGCTTGATTGGCGAGCAGCCAAACGGTATCCCAAATAATTTGTTGCCTTATAT
CTGCCAAGTGCGG

TCGTTTCTACGCCGACCTTCTATCGAAAGCCGAGACCGGCTTGTTGGGAACCAAATGCAGATTTCGCGC
AGATGATGGAAGACTCATGGGCGTGGGTAGATAAATCCTGAAGATGGCTAGATGATTAA
TCGAGGATAGCAATTCGACCGCGGCACTGGGTGAGACGTATTTCTGGTGATTTTCTTTTGCATTTGGC
GAGCCGACAGTCCAAGAGCGGTGTTTCTGCTGAAA
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCG
GTCGCGATGGCGAAGATATGCCACTGGGCGATACCACCAGCCCTTACGGCACATCGAAATCA
ATGGTTGAGCGCATTCG
CGGTGGGTGAAAGTGTGGCCGAGGCATGATTTTGTTGCGTTATTTCAATCCAATTGGCGC
GCATGAAAGCGGCTTGATTGGCGAGCAGCCAAACGGTATCCCAAATAATTTGTTGCCTTATAT
CTGCCAAGTGCGG

236
TCACTGATATTTCAAAAAGCCGATCCGC

CGCAGAAAAGTGTTTTGATGGCGCACTGCACTGGATAAACGATTTCTGGCTTTTATCTGCC
AAAGTGGGCGACGCAACTTGCCGAGCAGGTTATCAGCGATCTATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

GCATGCGACTATATGCTGACTATTATAGGGTATAGAGGGGTTGGAAATCATCCGGCATGCTGCTTCTTGC

GAGCCGATGAAATATTATGATAATAATGTTTCGAGCTTTGGAATCAGTGATCTATTCTGGCAATGGCTATAGCC

CGCACGAAAGTGGTTTGTATTGGCAGAGCAGCCAAACGGCATCTCCCGACACCGTACCGCCCGGTGGATTTG

GTCGTTCTTCTGCAATCGGCGATTTCGCCGCTTTTATGCGCGGCGATCCGGGCAAAGTGCCTTATAC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

>43

ATGAAAAAAATCTCGTACCCGGCACCCCGGTTTTATCGGCTCGCATACTGGTTTTTGCTGAAAA

GCGGCACCATCAAGTCTGGATTTGGATAAATCTTCAAATTCGACCATCTATCTCCGACCGGGTTGAAAA

CGATTACCGGCAAGATTCGCTGATTTTGGGATTACCGATGACTACGTTTCTGTGAAAA

GCGGCCATCAAGTCGTGATTTTGGATAACCTATGCAATTCCAGCATCAATATCCTGCCACGCTTGAAAA

CGCATGAAAAGCGGCTTGATTGGCGAGCAGCCAAACGGCATCTCCCGACACCGTACCGCCCGGTGGATTTG

GTCGTTCTTCTGCAATCGGCGATTTCGCCGCTTTTATGCGCGGCGATCCGGGCAAAGTGCCTTATAC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

>44

ATGAAAAAAATCTCGTACCCGGCACCCCGGTTTTATCGGCTCGCATACTGGTTTTTGCTGAAAA

GCGGCACCATCAAGTCTGGATTTGGATAAATCTTCAAATTCGACCATCTATCTCCGACCGGGTTGAAAA

CGATTACCGGCAAGATTCGCTGATTTTGGGATTACCGATGACTACGTTTCTGTGAAAA

GCGGCCATCAAGTCGTGATTTTGGATAACCTATGCAATTCCAGCATCAATATCCTGCCACGCTTGAAAA

CGCATGAAAAGCGGCTTGATTGGCGAGCAGCCAAACGGCATCTCCCGACACCGTACCGCCCGGTGGATTTG

GTCGTTCTTCTGCAATCGGCGATTTCGCCGCTTTTATGCGCGGCGATCCGGGCAAAGTGCCTTATAC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

>45

ATGAAAAAAATCTCGTACCCGGCACCCCGGTTTTATCGGCTCGCATACTGGTTTTTGCTGAAAA

GCGGCACCATCAAGTCTGGATTTGGATAAATCTTCAAATTCGACCATCTATCTCCGACCGGGTTGAAAA

CGATTACCGGCAAGATTCGCTGATTTTGGGATTACCGATGACTACGTTTCTGTGAAAA

GCGGCCATCAAGTCGTGATTTTGGATAACCTATGCAATTCCAGCATCAATATCCTGCCACGCTTGAAAA

CGCATGAAAAGCGGCTTGATTGGCGAGCAGCCAAACGGCATCTCCCGACACCGTACCGCCCGGTGGATTTG

GTCGTTCTTCTGCAATCGGCGATTTCGCCGCTTTTATGCGCGGCGATCCGGGCAAAGTGCCTTATAC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

CGCGATGGAAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC

CGAGGATGGAAGACTCATGCGCTTGGGCTGACAAACACCAATGGCTACGAGCATAGCTAATCCGAGCCGCTTGGCACTTGGGCTGC
AACCGACATCCAAAAAGCGCAGTCCGCTGGGAGCGTAGTATTTCGTTGGCTATTTCGAATCCTCCGACGCGG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AGTGGCCCTCCGCCGAGCAGGCGTCCGCTGCTGCGCCGCCCGATGCAGGAG
GCACGGAAGGGTGTGACGCTGCTGCGGGAAGAGCCGACCGTCCGCTCTACACGG
AACCGACATCCAAAAAGCGATCCGCGTTGGAGC
GCTGATTTTGTTGCGCTATTTCAACCCGATCGGCGC
GCACGAAAGCGGTTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCA
AGTGGCTTCGGGCAGGCTGCCACAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGCACCG
GTATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCGAAAGGC
GGTGTTGCCGGCGTACA
**67**
ATGACCGTCTCGTATTACGCCGGCAGCCGGCTTCATCGGTTTCGCAACACCGCGTCTCGTCATCACTCC
GTTTACGATGCGGTATTCTGGATAATCTGTGAACTCTGTGCTTGGCGGCTTCACCCGCGCTTCGGCGA
TTACCGCAAAAACCTACCGTTTATACCTTCCTGCGGATACCCGCGCCATTITGGAGGCCAGATTTGGCC
GGCGAATAACGGCCGGCACACCCGAGCCGCGTGATCCGCGGAATGGCGGAAAGGGATGGCCGAAACGGCG
GCACGAAAGCGGTTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCA
AGTGGCTTCGGGCAGGCTGCCACAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGCACCG
GTATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCGAAAGGC
GGTGTTGCCGGCGTACA
**68**
ATGACCGTCTCGTATTACGCCGGCAGCCGGCTTCATCGGTTTCGCAACACCGCGTCTCGTCATCACTCC
GTTTACGATGCGGTATTCTGGATAATCTGTGAACTCTGTGCTTGGCGGCTTCACCCGCGCCATTITGGCC
GGCGAATAACGGCCGGCACACCCGAGCCGCGTGATCCGCGGAATGGCGGAAAGGGATGGCCGAAACGGCG
GCACGAAAGCGGTTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCA
AGTGGCTTCGGGCAGGCTGCCACAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGACCG
GTATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCGAAAGGC
GGTGTTGCCGGCGTACA
**69**
ATGACCGTCTCGTATTACGCCGGCAGCCGGCTTCATCGGTTTCGCAACACCGCGTCTCGTCATCACTCC
GTTTACGATGCGGTATTCTGGATAATCTGTGAACTCTGTGCTTGGCGGCTTCACCCGCGCCATTITGGCC
GGCGAATAACGGCCGGCACACCCGAGCCGCGTGATCCGCGGAATGGCGGAAAGGGATGGCCGAAACGGCG
GCACGAAAGCGGTTTGATCGGCGAACAGCCCAACGGTATTCCCAACAACCTTTTGCCCTATATCTGCCA
AGTGGCTTCGGGCAGGCTGCCACAACTGTCGGTATTCGGCGGCGATTATCCGACCCCCGACGGACCG
GTATGCGCGACTACATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCGAAAGGC
GGTGTTGCCGGCGTACA

AACCGACATCCAAAAAGCCGATCCGCGTTGAGCGTGATTGTTGCTATTTCACAACCGGATCGGC
GCAGCAAAATGTTGATCCGCGAACCGCCAACAGGTATTCCCAACAACCTTTTGCCCTATATCGCA
AGTGGCTTCGCGGCGGTACATTGTTCAACTTGGGTTTGGCCCGGCGATCCGCTG
TGTTTCCTTTGCCCGACCGCTCCATACAACAAAGACAGAGGATTTGGGAAAAACCCGGATCCTGCGACA
AATGATGGAAGATTCGGTGCTGGTTGGTACCGGCAGCAGGATATTGGGATTAA

>70
ATGACCATATTTCCTACGCGGCGTACGGGTTTATCGGGTGACACCTGGTATCTCTTTACTCAACG
GATTTGATTTGTTAATTTTGGGATATTTATGATAACCCGCGAAGCTCGCCGCTTACATCGAT
GCTGGAATACCTGTGCGCTTATGGAGGCGATATTGCTGAGGTTTTTTGCGGGGATTGGTTGCC
GAACATGATATGTTATCCTGTGATTTTCTCTGGGATGAAAGCTGATACGCTGGGCAAGCC
GCATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGACTATATTGCAACTACCTGCTCATTATTGGGATACCCGCGAATAACCTGCGCGGAAACCTTGG
CGATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGGAAATCATCCG
TGCTTTGGGCTATACCTGGTATTTTGGGATACCCGCGAATAACCTGCGCGGAAACCTTGG
CGATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG

>71
ATGACCATATTTCCTACGCGGCGTACGGGTTTATCGGGTGACACCTGGTATCTCTTTACTCAACG
GATTTGATTTGTTAATTTTGGGATATTTATGATAACCCGCGAAGCTCGCCGCTTACATCGAT
GCTGGAATACCTGTGCGCTTATGGAGGCGATATTGCTGAGGTTTTTTGCGGGGATTGGTTGCC
GAACATGATATGTTATCCTGTGATTTTCTCTGGGATGAAAGCTGATACGCTGGGCAAGCC
GCATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGGAAATCATCCG
TGCTTTGGGCTATACCTGGTATTTTGGGATACCCGCGAATAACCTGCGCGGAAACCTTGG
CGATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGGAAATCATCCG

>72
ATGACCATATTTCCTACGCGGCGTACGGGTTTATCGGGTGACACCTGGTATCTCTTTACTCAACG
GATTTGATTTGTTAATTTTGGGATATTTATGATAACCCGCGAAGCTCGCCGCTTACATCGAT
GCTGGAATACCTGTGCGCTTATGGAGGCGATATTGCTGAGGTTTTTTGCGGGGATTGGTTGCC
GAACATGATATGTTATCCTGTGATTTTCTCTGGGATGAAAGCTGATACGCTGGGCAAGCC
GCATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGGAAATCATCCG
TGCTTTGGGCTATACCTGGTATTTTGGGATACCCGCGAATAACCTGCGCGGAAACCTTGG
CGATGAAGTATTACGATAACAATGTGTCGGGTAGCTTGGTGTTGGCTGAAGAAATGGTGCGGGCGGG
TGTGTTCAACTTGGGTTCGGGGCGCGCCTATTCCGTTTTGGGAAATCATCCG
GGATATCCAAAAAGCCGATCCGCGTTGGAGCGCTGATTTTTGTTGCATACTATTCCGATCGCGCA
TGAAAGCGGGTGTTGACAGGCGACACCCGGCTTTAATGCTGGAATCTGCTCTGGCTCAGGCAAATC
GGCTTCAGGCAAGCTCGGCTCAAAATCCACCAAGCCGCTTGCTGAGTGTTGGAACCTTACGGGC
TCTACGGCAGATCCCGCTACACCGGCTATCCGCAAGATGCGCAGATG
GGTGTACCTATGGCCGTTGGGTCAGCAACAAACTCATCCATCGAGATGTTA
TTAACCGACATCCAAAAAGCCGATCCGCGTTGGGCAAGTGCATCTTTCATTCGCTGAAGCAGTCGAGACTGACGGAATCGCGACTACATCCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAA
GTAATGTAGCAGGC
>82
ATGAAAAAAATTTCGTTACCACCGGCAGCGCGACCGGTTTTATCGGCTCGATACCGTGTTTTTCTTTTGCTGAAAA
AGCGCCGATCGAAGTGGTTTCTTTTGCTGAAAA
GTAATGTAGCAGGC
>83
ATGAAAAAAATTTCGTTACCACCGGCAGCGCGACCGGTTTTATCGGCTCGATACCGTGTTTTTCTTTTGCTGAAAA
AGCGCCGATCGAAGTGGTTTCTTTTGCTGAAAA
GTAATGTAGCAGGC
>84
ATGAAAAAAATTTCGTTACCACCGGCAGCGCGACCGGTTTTATCGGCTCGATACCGTGTTTTTCTTTTGCTGAAAA
AGCGCCGATCGAAGTGGTTTCTTTTGCTGAAAA
GTAATGTAGCAGGC
}>250
TCACTGATATTCAAAAAGCCGATCCGCGCTGGAGCGTGATTTTGTGGCTATTTTCAACCCGATTGGGG
GGGATGCGTGACTATATTCATGTGATGGATTTGGCAGAAGGGCATATCGCGGCAATGAAGGCAAAAA
GTAATGAGGAGGATTCGTCGGGTGGGGTCAGCCGCAACCCCAACGGGATATGGGGATTA

>85
ATGAAAAAATTTCGTTACCGGCGGCACCGGTTTTTATCGGCTCGCATACCGTTGTTTTCTTTTGTGA
ATGCCGCCATCAAGTCGTCGATTTTTGATAACCTATGCAATTCCAGATCATAATCTGGCGTATGAAA
CGATTACCGCAGATTCGTGATTTTCCGATCCGATTCGATGCTGATTTTGGGCTGTTATTTCAATCCG
GGCAGATTTCTCAGGACGAGATTCTCTCGGGCGCTTCTTCGGTGTTGGAAATCATCCG
GTGCCTTTTCGGACCGCCGATCCGGCTTAAATATTCCCTATCAAATCCAACCCCGACGCCGGCGGAGACTTG
CGTGTTCCTTTTCGGACCGCCGATCCGGCTTAAATATTCCCTATCAAATCCAACCCCGACGCCGGCGGAGACTTG
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCA
GGGCGTGTTTAGCATTGTGTTCAGTTCTTCGGCGACGGTTTATGGCGATCCGGGCAAAGTGCCCTTATAC
CGAGGATATGTCCTGCAGGCACGCATTTGCTGAACTTAGGCTCCGGCCGCGCTTCTTCGGTGTTGGAAATCATCCG
GTGCCTTTTCGGACCGCCGATCCGGCTTAAATATTCCCTATCAAATCCAACCCCGACGCCGGCGGAGACTTG
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCA

>86
ATGAAAAAATTTCGTTACCGGCGGCACCGGTTTTTATCGGCTCGCATACCGTTGTTTTCTTTTGTGAA
ATGCCGCCATCAAGTCGTCGATTTTCCGATCCGATTCGATGCTGATTTTGGGCTGTTATTTCAATCCG
GGCAGATTTCTCAGGACGAGATTCTCTCGGGCGCTTCTTCGGTGTTGGAAATCATCCG
GTGCCTTTTCGGACCGCCGATCCGGCTTAAATATTCCCTATCAAATCCAACCCCGACGCCGGCGGAGACTTG
GAGCCGATGAAATATTATGATAATAATGTTTCCGGCAGCTTGGTGTTGGCGGAAGAAATGGCGCGTGCA

>87
ATGAAAAAATTTCGTTACCGGCGGCACCGGTTTTTATCGGCTCGCATACCGTTGTTTTCTTTTGTGAA
ATGCCGCCATCAAGTCGTCGATTTTCCGATCCGATTCGATGCTGATTTTGGGCTGTTATTTCAATCCG
GGCAGATTTCTCAGGACGAGATTCTCTCGGGCGCTTCTTCGGTGTTGGAAATCATCCG
GTGCCTTTTCGGACCGCCGATCCGGCTTAAATATTCCCTATCAAATCCAACCCCGACGCCGGCGGAGACTTG
GACATTCAAAAAGCCGATCCGCTGTTGAGACATGATTTTGTGGTATTTTCAATCCGATTGCGCGCAT
GAAAGCGGTTTGTATGGCAGAGCAACCGTATCCTAACAAATATTGGTCCCTATTTACTGCAAAGT
GCGGCAAGGCAAATCTCCGCAATTACCGTTATCGTGATGACTACCTACCCCTGATGATCGGACGGTAT
GCGGCATTATCATTGTGATGGGTATTTCGACAGATGATTTTGCCGAAGAAACTCAGTGATATTGGCGTC
TCTATTGGCCACCTCTCTATGCAAAACCGGCAAAATCCGGCTGCCGAAACCCAGCGTATTTAACCCAAATGA
TGGAAAGACTCATGGCCTGGGTTGAGTAATAATCCGAATGGTGACTGATTTTGAGAAAT

>103
ATGAAAAAAATTTCTCTTTACCCGGCGGCACCAGGGTTTATTTATGGGAATAACCTATGCAATTCGGCATCAATCTTCTGCACG
GATTACGCGCCAAGAGATTCGCTGGTATTGCTGATTTTCGACAGATGATTTTGCCGAAGAAACTCAGTGATATTGGCGTC
TCTATTGGCCACCTCTCTATGCAAAACCGGCAAAATCCGGCTGCCGAAACCCAGCGTATTTAACCCAAATGA
TGGAAAGACTCATGGCCTGGGTTGAGTAATAATCCGAATGGTGACTGATTTTGAGAAAT

>104
ATGAAAAAAATTTCTCTTTACCCGGCGGCACCAGGGTTTATTTATGGGAATAACCTATGCAATTCGGCATCAATCTTCTGCACG
GATTACGCGCCAAGAGATTCGCTGGTATTGCTGATTTTCGACAGATGATTTTGCCGAAGAAACTCAGTGATATTGGCGTC
TCTATTGGCCACCTCTCTATGCAAAACCGGCAAAATCCGGCTGCCGAAACCCAGCGTATTTAACCCAAATGA
TGGAAAGACTCATGGCCTGGGTTGAGTAATAATCCGAATGGTGACTGATTTTGAGAAAT

>105
ATGACCATATTCTCTTTACCCGGCGGCACCAGGGTTTATTTATGGGAATAACCTATGCAATTCGGCATCAATCTTCTGCACG
GATTACGCGCCAAGAGATTCGCTGGTATTGCTGATTTTCGACAGATGATTTTGCCGAAGAAACTCAGTGATATTGGCGTC
TCTATTGGCCACCTCTCTATGCAAAACCGGCAAAATCCGGCTGCCGAAACCCAGCGTATTTAACCCAAATGA
TGGAAAGACTCATGGCCTGGGTTGAGTAATAATCCGAATGGTGACTGATTTTGAGAAAT
CCGATATTCAGAAAGCCGACCCTC GTGGAGCGTGATTTTATTGCGTTATTTCAAC CCTATCGGTGC GCA TCAAAGCGGCTTAATCGGCGAGCA CCCCAAACGGGATTCCGAATAACCTGCTGCGTATATCTGCAAAGT GGCCGGCGGGAAAAC TACGC GTATCTTTTGCT ATTCGGCGGATGATTATCC AACTCTGCCGACCGTACCGGTAT GCGCGACTATTAT CATGTGGTGGATT TTGGCGGAAGGCCATG TTGCCGCAAT GAAGGCCAAAGCCAATG T TCCCGGC GTACACCTGCTGAAAT TTGGTTGCGGCGGCGGCGTATCGGT GAGATTGTCCTGTCGCA TTTGAAGCCGCTTCCGGGCT TAACG GCGCCGCAAAGCGGACTTGGCC GCTGT GTGTATGCCGATCCGACTTATACCAAACAGCAAACCGGCTGGGAAACCAAACGTGATTTGACTCAAAT GATAGAGACGCGT TGGTTGGTAAACCCGGAATCAAACGGGATACGCGGATTTGA
Appendix C. Strains of *Neisseria* used to examine the distribution of *galE* alleles.

<table>
<thead>
<tr>
<th>isolate</th>
<th><em>galE</em> allele</th>
<th>clonal complex (MLST)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N. meningitidis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120M</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>MCS8</td>
<td>2</td>
<td>ST-32 complex/ET-5 complex</td>
<td></td>
</tr>
<tr>
<td>FAM18</td>
<td>1</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>53442</td>
<td>3</td>
<td>ST-4821 complex</td>
<td></td>
</tr>
<tr>
<td>6748</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>129E</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>Alpha275</td>
<td>8</td>
<td>ST-22 complex</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>393</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>254</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>S5611</td>
<td>40</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>CN100</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3608</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha162</td>
<td>9</td>
<td>ST-23 complex/Cluster A3</td>
<td></td>
</tr>
<tr>
<td>Alpha707</td>
<td>14</td>
<td>ST-254 complex</td>
<td></td>
</tr>
<tr>
<td>171</td>
<td>10</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>Alpha388</td>
<td>11</td>
<td>ST-254 complex</td>
<td></td>
</tr>
<tr>
<td>172</td>
<td>10</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>173</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2027</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2029</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2028</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2025</td>
<td>6</td>
<td>ST-213 complex</td>
<td></td>
</tr>
<tr>
<td>196/87</td>
<td>41</td>
<td>ST-32 complex/ET-5 complex</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>38VI</td>
<td>42</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>M597</td>
<td>77</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>2059001</td>
<td>34</td>
<td>ST-4 complex/subgroup IV</td>
<td></td>
</tr>
<tr>
<td>BZ 133</td>
<td>43</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>BZ 147</td>
<td>44</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>BZ198</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>BZ 232</td>
<td>45</td>
<td>ST-37 complex</td>
<td></td>
</tr>
<tr>
<td>DK 24</td>
<td>46</td>
<td>ST-4240/6688 complex</td>
<td></td>
</tr>
<tr>
<td>NG 3/88</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NG 4/88</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>ST</td>
<td>Strain Description</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>---------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>NG 6/88</td>
<td>21</td>
<td>ST-269 complex</td>
<td></td>
</tr>
<tr>
<td>NG F26</td>
<td>48</td>
<td>ST-269 complex</td>
<td></td>
</tr>
<tr>
<td>NG H15</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>NG H41</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NG E31</td>
<td>49</td>
<td>ST-364 complex</td>
<td></td>
</tr>
<tr>
<td>NG E30</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>NG H36</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>NG 080</td>
<td>2</td>
<td>ST-32 complex/ET-5 complex</td>
<td></td>
</tr>
<tr>
<td>NG144/82</td>
<td>2</td>
<td>ST-32 complex/ET-5 complex</td>
<td></td>
</tr>
<tr>
<td>NG PB24</td>
<td>2</td>
<td>ST-32 complex/ET-5 complex</td>
<td></td>
</tr>
<tr>
<td>NG P20</td>
<td>1</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>297-0</td>
<td>14</td>
<td>ST-254 complex</td>
<td></td>
</tr>
<tr>
<td>3906</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>371</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>BRAZ10</td>
<td>42</td>
<td>ST-11 complex/ET-37 complex</td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>79128</td>
<td>79</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>322/85</td>
<td>39</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>79126</td>
<td>50</td>
<td>ST-1 complex/subgroup I/II</td>
<td></td>
</tr>
<tr>
<td>Z2491</td>
<td>13</td>
<td>ST-4 complex/subgroup IV</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>36</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>AK50</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>50/94</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>M40/94</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>931905</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>N45/96</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>91/40</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>88/03415</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>E26</td>
<td>37</td>
<td>ST-198 complex</td>
<td></td>
</tr>
<tr>
<td>A22</td>
<td>8</td>
<td>ST-22 complex</td>
<td></td>
</tr>
<tr>
<td>71/94</td>
<td>38</td>
<td>ST-23 complex/Cluster A3</td>
<td></td>
</tr>
<tr>
<td>860800</td>
<td>38</td>
<td>ST-167 complex</td>
<td></td>
</tr>
<tr>
<td>8013</td>
<td>27</td>
<td>ST-18 complex</td>
<td></td>
</tr>
<tr>
<td>ATCC13091</td>
<td>92</td>
<td>ST-7355</td>
<td></td>
</tr>
<tr>
<td>M01_240149</td>
<td>21</td>
<td>ST-41/44</td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M01_240355</td>
<td>97</td>
<td>ST-213</td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M01_240196</td>
<td>98</td>
<td>ST-269</td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NZ_05/33</td>
<td>21</td>
<td>ST-41/44</td>
<td></td>
</tr>
<tr>
<td>Alpha 14</td>
<td>16</td>
<td>cc53</td>
<td></td>
</tr>
<tr>
<td>Alpha 710</td>
<td>21</td>
<td>ST-41/44 complex/Lineage 3</td>
<td></td>
</tr>
<tr>
<td>NMB</td>
<td>85</td>
<td>ST-8 complex/cluster A4</td>
<td></td>
</tr>
<tr>
<td>NZD1-006-1</td>
<td>101</td>
<td>ST-22</td>
<td></td>
</tr>
<tr>
<td>NZ05/50</td>
<td>102</td>
<td>ST-11</td>
<td></td>
</tr>
<tr>
<td>B1940</td>
<td>103</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Year</td>
<td>YearStart</td>
<td>ST</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>630501</td>
<td>100</td>
<td>2019</td>
<td>ST-5</td>
</tr>
<tr>
<td>S10602</td>
<td>104</td>
<td>2019</td>
<td>ST-5</td>
</tr>
<tr>
<td>FAM-20</td>
<td>91</td>
<td>2019</td>
<td>ST-11</td>
</tr>
<tr>
<td>MA-1</td>
<td>39</td>
<td>2019</td>
<td>unknown</td>
</tr>
<tr>
<td>M0579</td>
<td>21</td>
<td>2020</td>
<td>ST-41/44</td>
</tr>
<tr>
<td>WUE2594</td>
<td>100</td>
<td>2020</td>
<td>ST-5</td>
</tr>
<tr>
<td>961-5945</td>
<td>82</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>Alpha153</td>
<td>14</td>
<td>2020</td>
<td>ST-60</td>
</tr>
<tr>
<td>ES14902</td>
<td>27</td>
<td>2020</td>
<td>ST-11</td>
</tr>
<tr>
<td>M6190</td>
<td>23</td>
<td>2020</td>
<td>ST-11</td>
</tr>
<tr>
<td>OX99.3034</td>
<td>21</td>
<td>2020</td>
<td>ST-41/44 complex/Lineage 3</td>
</tr>
<tr>
<td>N1568</td>
<td>96</td>
<td>2020</td>
<td>ST-751</td>
</tr>
<tr>
<td>H44/76</td>
<td>2</td>
<td>2020</td>
<td>ST-32 complex/ET-5 complex</td>
</tr>
<tr>
<td>G2136</td>
<td>84</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>B6116/77</td>
<td>83</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>SB25</td>
<td>88</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>94/155</td>
<td>87</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>312 901</td>
<td>89</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>AK22</td>
<td>86</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>BZ10</td>
<td>95</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>BZ163</td>
<td>94</td>
<td>2020</td>
<td>ST-8 complex/cluster A4</td>
</tr>
<tr>
<td>M01-240013</td>
<td>99</td>
<td>2020</td>
<td>ST-1159</td>
</tr>
<tr>
<td>M13399</td>
<td>93</td>
<td>2020</td>
<td>ST-269 complex</td>
</tr>
<tr>
<td>K1207</td>
<td>27</td>
<td>2020</td>
<td>ST-11 complex/ET-37 complex</td>
</tr>
<tr>
<td>S0108</td>
<td>27</td>
<td>2020</td>
<td>ST-11 complex/ET-37 complex</td>
</tr>
<tr>
<td>CU385</td>
<td>2</td>
<td>2020</td>
<td>ST-32 complex/ET-5 complex</td>
</tr>
</tbody>
</table>

**N. gonorrhoeae**

<table>
<thead>
<tr>
<th>Code</th>
<th>Year</th>
<th>YearStart</th>
<th>ST</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA1090</td>
<td>17</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCCP11945</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID24</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DGI18</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA6140</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1291</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35/02</td>
<td>81</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F62</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA19</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID1</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID18</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID332</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK-92-679</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DGI2</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK-93-1035</td>
<td>19</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCDC-NG8107</td>
<td>81</td>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>ATCC</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS-11</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**N. lactamica**

<table>
<thead>
<tr>
<th>Strain</th>
<th>ATCC</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>224</td>
<td>60</td>
<td>ST-624 complex (lactamica)</td>
</tr>
<tr>
<td>4116</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>09002S1</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>8206</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>004-12</td>
<td>63</td>
<td>ST-595 complex (lactamica)</td>
</tr>
<tr>
<td>005-12</td>
<td>64</td>
<td>ST-1494 complex (lactamica)</td>
</tr>
<tr>
<td>012-12</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>014-24</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>016-24</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>017-02</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>028-12</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>030-24</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>039-03</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>049-12</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>NS19</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Y92-1009</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>ATCC23970</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>020-06</td>
<td>18</td>
<td>ST-640 complex (lactamica)</td>
</tr>
</tbody>
</table>

**N. cinerea**

<table>
<thead>
<tr>
<th>Strain</th>
<th>ATCC</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCUG 346T</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>CCUG 5746</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>CCUG 25879</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>CCUG 27178 A</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>ATCC14685</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

**N. polysaccharea**

<table>
<thead>
<tr>
<th>Strain</th>
<th>ATCC</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCUG 18031</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>CCUG 24845</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>CCUG 24846</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>CCUG 27182</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>15883</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>NS342</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>ATCC43768</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

**N. mucosa**

<table>
<thead>
<tr>
<th>Strain</th>
<th>ATCC</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCUG 26878T</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>CCUG 10421</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>CCUG 804</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

**N. elongata**
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCUG 4554</strong></td>
<td>71</td>
</tr>
<tr>
<td><strong>ATCC29315</strong></td>
<td>32</td>
</tr>
<tr>
<td><strong>N. perflava</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CCUG 7826</strong></td>
<td>72</td>
</tr>
<tr>
<td><strong>N. sicca</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CCUG 24918</strong></td>
<td>74</td>
</tr>
<tr>
<td><strong>oral taxon 014 str.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>F0314</strong></td>
<td>105</td>
</tr>
<tr>
<td><strong>N. subflava</strong></td>
<td></td>
</tr>
<tr>
<td><strong>NJ9703</strong></td>
<td>31</td>
</tr>
<tr>
<td><strong>N. flavescens</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CCUG 28662</strong></td>
<td>76</td>
</tr>
</tbody>
</table>
Appendix D. Publication relating to thesis.

I am listed on an author on this publication submitted to PLoS ONE as I made the mutant of *N. meningitidis* strain NMB which is lacking *pilEII*, and instead expresses *pilEI* from the native promoter of MC58 in the *iga* locus using the plasmid constructed by Ya-Hsun Lin, under the supervision of John Davies and Catherine Ryan. This paper has demonstrated, using sensitive means, that antigenic variation of the class II *pilE* does not occur in strain ST-8 cc strain NMB nor ST-11 cc strain FAM18. In addition to this, although *pilEI* has been observed to antigenically vary in strains of *N. meningitidis* which naturally harbor this allele, the expression of *pilEI* in a *pilEII* strain does not result in observed antigenic variation.

**The use of high-throughput DNA sequencing in the investigation of antigenic variation: application to *Neisseria* species.**

John K. Davies¹, Paul F. Harrison², Ya-Hsun Lin¹, Stephanie Bartley³, Chen Ai Khoo¹, Torsten Seemann², Catherine S. Ryan¹, Charlene M. Kahler³ and Stuart A. Hill⁴.

¹ Department of Microbiology, Monash University, Clayton, VIC 3800, Australia.

² Victorian Bioinformatics Consortium, Monash University, Clayton, VIC 3800, Australia.
Antigenic variation occurs in a range of bacterial, protozoan and fungal pathogens. This process can occur when variant DNA is unidirectionally transferred from partial gene copies (or silent loci) into an expression locus, and resembles gene conversion. Previous studies of antigenic variation have involved the amplification and sequencing of individual genes from hundreds of colonies in order to detect the variant genes. Using the well-characterized pilE gene from Neisseria gonorrhoeae we have demonstrated that it is possible to use PCR amplification, followed by high-throughput DNA sequencing and a novel assembly process, to detect individual antigenic variation events. The ability to detect these events was orders of magnitude greater than has previously been possible. In accord with previous findings, the variant sequences were located at the 3’ end of the gene, their appearance was RecA-dependant, and the silent loci acting as donors are scattered around the chromosome. The pilE gene of N. gonorrhoeae and some strains of Neisseria meningitidis encode class I pilin, but other strains of N. meningitidis encode a class II pilin. The class II pili of meningococcal strain FAM18 have been reported to be non-variable. We have confirmed this, and shown that is also true for the class II pili of another strain NMB. In addition when a gene encoding class I pilin along with its native promoter, and a guanine quartet sequence that is needed for variation, were moved into a NMB background there was no evidence of antigenic variation. Finally we investigated several members of the opa gene family of N. gonorrhoeae, where it has been suggested that limited variation
occurs. Variation was detected in the *opaK* gene that is located close to *pilE*, but not at the *opaJ* gene located elsewhere. The approach described here promises to dramatically improve studies of the extent and nature of antigenic variation systems in a variety of species.
D.2 INTRODUCTION

Antigenic variation is a genetic process that leads to high-frequency changes in cell surface components in a wide range of species. To evaluate the experimental approach described here we initially investigated antigenic variation in a gene where that process has been well studied, the pilE gene of Neisseria gonorrhoeae. N. gonorrhoeae (the gonococcus), and the closely related N. meningitidis (the meningococcus), are both strict human pathogens and must continually evade the human immune system. What partly confounds the immune system is antigenic variation of the PilE pilin subunit that assembles into the type IV pilus. In this system the expressed gene (pilE) changes and the so-called silent loci (pilS) donate variant genetic information, but remain unchanged in the process. The variant sequences recombine into the distal two thirds of the pilE gene (5), and the process is dependent on the presence of the RecA protein (11). Mosaic proteins can sometimes be formed as multiple gene segments from the pilS loci are recombined into the expressed copy to yield a variant gene (5). Antigenic variation in this gene is a high-frequency event, with approximately 12% of randomly selected colonies containing a variant gene sequence (5). Two distinct classes of type IV pili are found in N. meningitidis: class I pili (also found in N. gonorrhoeae) antigenically vary, whereas at least some class II pili do not (7).

In addition to the type IV pilin system, other potential antigenic variation systems have been identified in the genome sequences of Neisseria species. The opa genes are a gene family encoding surface-exposed proteins, with the number of genes varying between strains. For N. gonorrhoeae strain MS11, nine intact opa genes have been cloned and sequenced allowing alignment of their gene sequences (1). From these alignments two hyper-variable regions can be observed in the middle of the genes as well as a semi-variable region located towards the 5’ end. Several studies have shown
DNA transformation-mediated horizontal transmission of chromosomal DNA where hyper-variable *opa* segments are exchanged between strains (2, 9, 10). However, limited data has been accrued as to whether *opa* genes also engage in antigenic variation as observed with the *pil* system. Anecdotal evidence suggests that limited variation does occur in at least some *opa* genes (1).

Various methods have been used to determine the extent and nature of antigenic variation within a population (5, 14). Recent studies have involved the amplification of the *pilE* gene from hundreds of randomly selected colonies, followed by sequencing of each amplicon using Sanger sequencing technology to detect the subset that contained variant sequences (5, 7, 8). Because of the cost and time involved the number of genes that were eventually sequenced was limiting, leading to conclusions being based on small data sets. For example, the extent of *pilE* gene variation in *N. gonorrhoeae* was based on analysis of 497 amplicons, and that for *N. meningitidis* on 260 amplicons (5, 7). We reasoned that high throughput sequencing of DNA extracted from a varying culture could be used to detect multiple gene variants simultaneously, especially those that might be occurring with a low frequency. In addition the coverage of the variant sequence, compared with the parental sequence, could be used to infer the frequency of individual antigenic variation events. This approach has dramatically increased the size of the data sets allowing a more detailed picture to emerge of the extent and nature of antigenic variation.

D.3 MATERIALS AND METHODS

**Bacterial strains and growth conditions.**

Initial experiments used gonococcal strains FA1090 and MS11 containing the *recA6* allele (15). This allele was first introduced into strain MS11, and subsequently genomic
DNA from MS11 recA6 was used to transform strain FA1090. In these strains the recA gene is transcribed from a promoter that is only operational in the presence of the inducer isopropyl thiogalactopyranoside (IPTG). Growing the strain in the absence of IPTG halts RecA production, and effectively “freezes” antigenic variation. The *N. meningitidis* strains used were NMB and FAM18, both of which produce class II pilin. The *N. meningitidis* strains used were NMB and FAM18, both of which produce class II pilin. The details of the construction of *N. meningitidis* strain CKNM397 will be published elsewhere, but briefly this is *N. meningitidis* strain NMB that now produces the class I pilin from *N. meningitidis* strain MC58. The pilE gene and upstream sequence containing the native promoter and a G4 sequence important for antigenic variation (3), was PCR amplified from MC58 and integrated into the iga gene of strain NMB. Subsequently the NMB pilE gene encoding class II pilin was insertionally inactivated. The solid and liquid media, and the growth conditions used have been described previously (6).

**Amplification and sequencing of genes of interest.**

A single colony of the relevant strain was picked from a GC agar plate and resuspended in 60-100 µl of GC broth. Neither the agar plate nor the broth contained IPTG, so in those strains containing the IPTG-inducible recA6 allele the recA gene was not being expressed, and antigenic variation was absent. A portion of this suspension was plated onto GC agar and incubated overnight. Strains that contained the recA6 allele were plated onto both GC agar and GC agar containing 2mM IPTG. The remainder of the original suspension was retained and used to generate a reference sequence by Sanger sequencing. The cells were harvested from the agar plates, and genomic DNA was extracted from both these samples and the reference sample using the GenElute™ Bacterial Genomic DNA Kit (Sigma). The genomic DNA was used as template in PCR with KOD Hot Start DNA Polymerase (Merck). The oligonucleotide primers used to PCR amplify the genes of interest are listed in Table S1 in the Supporting Information. These were designed such that they bound approximately 300
bp upstream and downstream of the gene of interest. The PCR conditions were as follows: 95°C for 2mins followed by 35 cycles of 95°C for 20 seconds, 55°C for 10 seconds and 70°C for 20 seconds. The PCR products were then gel-purified with QIAquick Gel Extraction Kit (QIAGEN). The PCR product from the reference sample was then subjected to Sanger sequencing using the same primers that were used for amplification. The amplicons were prepared for high-throughput sequencing using the Illumina Genomic DNA Sample Prep Kit as per the manufacturer’s instructions. Sequence data were obtained from an Illumina Genome Analyzer II using 36-cycle (Illumina 36C Sequencing Kit V2), or 72-cycle when it became available, paired-end chemistry. The raw sequence data from each sample have been deposited in the NCBI Sequence Read Archive, and the relevant accession numbers and sequence quality data are shown in Table S2 of the Supporting Information.

Assembly of sequence reads.

Reads were clipped where the quality fell below 20, and to remove Illumina adaptor sequences. Existing assembly programs, designed to handle moderate coverage of whole genome sequences, proved inadequate when faced with the extensive coverage of relatively short amplicon sequences used here. We therefore developed a novel assembler that was a variant on k-mer based assembly, where the term k-mer simply refers to a string of k nucleotides within a sequence read. Each sequence was assembled from a seed k-mer, with k=30 for 36-cycle and k=50 for 72-cycle chemistry. The seed for the initial assembly was the most common k-mer in each data set. In each subsequent assembly the seed was the most common k-mer that had not been a part of a previously assembled sequence. To assemble each sequence, the seed k-mer was progressively extended at either end. The tendency was therefore for each successive assembly to contain the next most common variant sequence. An extension was given a score based on the number of times the new k-mer was observed paired with k-mers already in the assembly, or with the reverse complement.
of k-mers already in the assembly. This score was in inverse proportion to the number of times the k-mer in the assembly occurred in the reads, so that evidence from the variant part of an assembly was not overwhelmed by evidence from k-mers present in most sequences. The variance of the score, under the assumption that counts were Poisson distributed, was also calculated. Extension halted when there was no extension that was better than the other options at that end by at least three standard deviations. We chose to assemble 50 sequences from each data set, on the assumption that the number of variant sequences present was unlikely to exceed this figure. Heat-maps of the 50 sequences were produced showing the number of occurrences of each pair of k-mers. This allowed checking for mis-assemblies, visualized as the unexpected absence of pairs of k-mers between two parts of an assembled sequence. The assembler can be obtained from http://www.vicbioinformatics.com/software.assemblet.shtml.

D.4 RESULTS

Analysis of sequence assemblies.

Each of the 50 sequence assemblies was used as a query sequence in a BLASTn search of the databases, and any assembly that did not relate to the gene and strain in question was removed from further consideration. These were mainly short assemblies with very low sequence coverage. The remaining assemblies were aligned to the reference sequence using ClustalW2. This occasionally identified additional mis-assemblies arising from the short Illumina sequence reads, and the presence of direct or inverted repeats in the amplified region. These were also removed from the analysis. The remaining assemblies were then manually searched for those that differed from the reference sequence by just one nucleotide, either a one nucleotide insertion or
deletion, or a nucleotide change. For the reasons outlined below, these assemblies were also removed from further analysis. The remaining assemblies were again aligned with the reference sequence using ClustalW2. In most cases the average k-mer depth exceeded $10^5$, and in some cases $10^6$ (Table 1). As expected, prominent amongst the remaining assemblies was an assembly (Seq 1) with high k-mer depth that was identical to the reference sequence obtained by Sanger sequencing of the amplicon. The frequency with which a particular variant appeared was estimated by dividing the average k-mer depth of the variant portion of the assembly by the average k-mer depth of same-sized regions immediately on either side of the variant sequence. This is illustrated for a particular assembly containing a variant sequence in Figure S1, Panel A. In some cases regions containing variant sequences were interrupted by region(s) of conserved sequence. An example of this is shown in Figure S1, Panel B. In these cases the spike in k-mer depth resulting from the presence of the internal conserved region would inflate the average k-mer depth if this was measured across the entire region, resulting in an artificially high frequency calculation. Therefore for these assemblies the frequency was calculated by dividing the k-mer depth of each individual variable segment by the average k-mer depth of same-sized regions immediately on either side of the entire region.

**Variation at the pilE gene encoding class I pilin in *N. gonorrhoeae* strain FA1090.**

The pilE of strain FA1090 recA6 grown in the absence of IPTG was amplified by PCR and sequenced using Sanger chemistry. An alignment with the FA1090 genome sequence (accession number AE004969) revealed sequence variation in the 3’ part of the gene (Figure S2). This variation can be explained by gene conversion using sequence from a specific silent locus, pilS1c1. In two independent experiments a culture grown in the absence of IPTG was split, with one half cultured in the absence of IPTG, whilst the other half was grown in the presence of IPTG, allowing antigenic variation. The pilE gene from both cultures was amplified by PCR, and subjected to
In both experiments there were multiple assemblies that differed from the reference sequence by just one nucleotide. Unlike the variant sequences described below, the single nucleotide changes were not concentrated in the 3’ end of the pilE gene, but were scattered throughout the pilE gene and the flanking sequences. In both experiments, only one assembly involved a single nucleotide change that occurred at higher frequency in the presence of RecA. The same change was detected in both experiments, and involved a single nucleotide change downstream of the pilE gene. It therefore seems that almost all of the assemblies involving single nucleotide changes were not the result of antigenic variation. Others have reported that antigenic variation can result in single nucleotide changes (5), but in our hands it seems more likely that these are the result of low frequency mutations occurring during amplification of the gene or (less likely) sequencing errors. Such changes were therefore not considered further.

In the first experiment 29 assemblies passed the assembly and screening process described above, while in the second experiment (maybe because of a much lower average k-mer depth in the RecA+ sample; Table 1) just six assemblies were detected. The results are shown in Table 2 and Table 3, for the first and second experiments, respectively. All of the variants were present at basal levels in the absence of RecA, and at a much higher frequency in the presence of RecA, suggesting active gene conversion during the experiment. In agreement with this suggestion, alignments of the assembled sequences are shown in Figures S3 and S4 and demonstrate that sequence variation was confined to the 3’ end of the pilE gene.

Of the 28 variant sequences detected in the first experiment, 21 were identical to a pilS locus in the FA1090 genome sequence (Table 2). A further 6 variant sequences differed from a pilS locus by just 1 or 2 nucleotides. In only one case (seq44) was a
mosaic sequence, derived from multiple silent loci, present. In three cases the variant sequence was identical to more than one silent locus, so it was not possible to unambiguously map the donor sequence. It was also evident that a particular silent locus can be involved in generating multiple variants. For instance, different portions of pilS3c3 were involved in generating the variant sequences in seq8, seq24, seq33 and perhaps seq47 (Table 2, Figure S3). In the repeat experiment all five variant sequences were identical to part of a silent locus (Table 3). Two variant sequences appeared in both experiments. Seq6 in the first experiment is identical to seq8 in the second experiment, and seq7 from the first experiment is identical to seq3 in the second experiment. A variety of silent loci served as the source of donor sequences, with 14 of the 19 silent copies potentially involved.

**Variation at the pilE gene encoding class I pilin in *N. gonorrhoeae* strain MS11.**

The pilE gene of gonococcal strain MS11 is also antigenically variable, but reportedly at a lower frequency than in strain FA1090 (8). In order to determine whether the approach outlined above could detect such differences, we again conducted two separate experiments, using strain MS11 recA6. In the first experiment 16 assemblies passed the screening procedure (Figure S5), while the second experiment yielded 8 assemblies (Figure S6). In both experiments seq1 was identical to both the reference sequence, and the MS11 pilE sequence deposited in the databases (Accession number K02078). In agreement with the earlier report (8) the frequency of antigenic variation observed was an order of magnitude lower than that seen in strain FA1090 (Table 4 and Table 5). Rather than being barely detectable in the RecA- culture, the variant sequences were present at a higher frequency than seen in FA1090, suggesting that despite single colony isolations a variant subpopulation was present. This occurred in both experiments and might suggest that the recA promoter is not as tightly controlled in the MS11 genetic background. In addition the presence of RecA only boosted the variant frequency approximately 10-fold in MS11, compared with
approximately 100-fold in FA1090 (Table 2, Table 3). Not all silent loci in strain MS11 have been sequenced and annotated. As a result in two cases in the first experiment, and three cases in the second experiment, it was not possible to assign a specific silent locus as the source of the donor sequence. In these cases the best pilE-related database match is given in Tables 4 and 5. Seq12 in the first experiment is identical to seq13 in the second experiment, and seq17 in both experiments are also identical.

**Lack of variation in the pilE gene of strains of *N. meningitidis* that express class II pilin.**

Meningococci also express type IV pili, with at least some of the class I pilin subunits also varying antigenically (16). However the class II pili of *N. meningitidis* FAM18 appear not to vary (7). We therefore searched for evidence of pilE variation in FAM18 and another meningococcal strain, NMB, which also produces class II pilin. In neither strain were we able to detect antigenic variation despite adequate k-mer depth (Table 1). We also looked for antigenic variation in meningococcal strain CKNM397. This strain is derived from strain NMB but produces the class I pilin from strain MC58 instead of the native class II pilin. In two separate experiments we were unable to detect any antigenic variants despite adequate k-mer depth (Table 1).

**Variation at opa genes.**

For the reasons outlined above, we also investigated whether there was any evidence of antigenic variation in the opa genes of *N. gonorrhoeae* strain FA1090. In an initial experiment we were unable to detect any variation in the FA1090 opaJ gene (data not shown). However it has been reported that the opaK gene, situated close to the pilE gene, is more variable than other opa loci (2). In two separate experiments, we were able to detect variant sequences in opaK (Figure S7 and Figure S8). In each case the variant portion of the assembly was identical to part of the opaD gene, which appeared to act as the donor sequence in these experiments. However in the first experiment the
variant sequences are clearly present in the RecA− sample and the presence of RecA boosts their levels only marginally (Table 6). In the second experiment both variants appeared much more frequently in the presence of RecA (Table 6).

**D.5 DISCUSSION**

Previous investigations of the extent and nature of antigenic variation in *Neisseria* have involved the amplification of genes from hundreds of individual colonies that were then sequenced by conventional Sanger sequencing technology in order to detect the subset that contained variant sequences (5, 7, 14). Such an approach was both time-consuming and expensive. The advent of affordable deep sequencing platforms has enabled an alternative approach to such studies. Here we have used PCR amplification of the genes of interest, followed by high throughput DNA sequencing, to detect variant sequences. This involved a single PCR amplification rather than hundreds, one sequencing reaction rather than hundreds, and a dramatically improved ability to detect low-frequency variants. The data shown in Table 2 and Table 3 suggest that this approach is indeed capable of detecting examples of antigenic variation, in that the variant sequences detected in these experiments are typical of antigenic variation events. They are (a) restricted to the 3’ end of the *pilE* gene, (b) detected at a high frequency only in the presence of RecA, and (c) in most cases identical to portion of a *pilS* copy. We were also able to detect antigenic variation in the *pilE* gene of *N. gonorrhoeae* strain MS11, and in agreement with an earlier report (8) this appeared to be occurring at a frequency lower than that observed for strain FA1090.

It has previously been reported that antigenic variation is undetectable in the *pilE* gene of *N. meningitidis* strain FAM18 expressing class II pilin (7). Despite potentially having the ability to detect variants at a much lower frequency than in the previous report, we
were also unable to detect variation in this gene, or in the equivalent gene from *N. meningitidis* strain NMB. Strains producing class II pili have only two silent loci, as opposed to eight *pilS* copies in other meningococcal strains. Also the *pilE* gene expressing class II pilin is located elsewhere on the chromosome, whereas in those meningococcal strains that produce class I pili the silent loci are adjacent to *pilE*. Both of these factors might adversely affect the ability of the *pilE* gene from class II-producing strains to undergo antigenic variation. In addition, a DNA structure in the *pilE* promoter region of *N. gonorrhoeae* that is necessary for pilin antigenic variation has recently been described (3). This guanine quartet structure is degenerate in meningococcal strains that produce class II pili, and this would also adversely affect the frequency of antigenic variation. However CKNM397 contains the *pilE* gene, and its native promoter, from strain MC58. The promoter region includes the sequences from MC58 that can form the guanine quartet structure. The absence of variation in this strain suggests that although this DNA structure is necessary for antigenic variation (3), by itself it is not sufficient to allow this process, at least in this genetic background. It seems that additional factor(s), present in class I-producing but not class II-producing strains, are needed for antigenic variation. Both FAM18 and NMB are disease-causing, rather than carriage, isolates. It therefore seems that antigenic variation of *pilE* is not essential for virulence, despite the fact that this process can be observed in all *N. gonorrhoeae* and many *N. meningitidis* strains.

We have also investigated another gene family where it has been suggested that antigenic variation might be occurring. In initial experiments with the *opaJ* gene of *N. gonorrhoeae* strain FA1090, we were unable to detect any variation. However it has been reported that the *opaK* gene, located close to *pilE*, is more variable than other *opa* genes (2). In two separate experiments with *opaK* we were able to detect sequence variants. However only in the second experiment were they clearly more frequent when the RecA protein was present. There were also distinct differences from
the variation observed in \textit{pilE}. Firstly the number of different sequence variants detected was much smaller than seen in \textit{pilE}. Secondly, compared with \textit{pilE}, there was a more distinct bias in the source of the donor sequences. In every case the variant portion of the sequence was identical to part of the \textit{opaD} gene, suggesting that this gene alone was donating sequence to yield variants of \textit{opaK}.

Here we have demonstrated that PCR amplification of the genes of interest, followed by high throughput DNA sequencing, can be used to investigate antigenic variation. This approach could therefore be applied to a wide range of antigenic variation systems at a level that has not previously been possible. For example \textit{Borrelia burgdorferi}, the causative agent of Lyme disease has a surface-exposed lipoprotein, VlsE, that undergoes antigenic variation (4, 17). The Msp2 and Msp3 systems of \textit{Anaplasma marginale} (12) and the VlhA system of \textit{Mycoplasma synoviae} (13) are additional examples of antigenic variation of surface components that could be explored using this approach. These genes all contain both conserved and variable segments. Given the data depth that can be achieved using this approach, it should therefore be possible to undertake more systematic searches for conserved segments in these antigenically variable genes.

D.6 ACKNOWLEDGEMENTS

We thank S. Coutts and Micromon for assistance with the high-throughput sequencing.

D.7 REFERENCES


Table D.1: Average k-mer depth of seq1, the sequence assembly that was identical to the reference sequence, for each experiment

<table>
<thead>
<tr>
<th>Strain</th>
<th>Gene</th>
<th>Experiment</th>
<th>Average depth (x10^3)</th>
<th>k-mer RecA&lt;sup&gt;*&lt;/sup&gt;</th>
<th>Average depth (x10^3)</th>
<th>k-mer RecA&lt;sup&gt;*&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N. gonorrhoeae</em> FA1090</td>
<td>pilE</td>
<td>1</td>
<td>250</td>
<td>1070</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. gonorrhoeae</em> FA1090</td>
<td>pilE</td>
<td>2</td>
<td>248</td>
<td>158</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. gonorrhoeae</em> MS11</td>
<td>pilE</td>
<td>1</td>
<td>240</td>
<td>202</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. gonorrhoeae</em> MS11</td>
<td>pilE</td>
<td>2</td>
<td>264</td>
<td>278</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. meningitidis</em> FAM18</td>
<td>pilE</td>
<td>1</td>
<td>-</td>
<td>1540</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. meningitidis</em> FAM18</td>
<td>pilE</td>
<td>2</td>
<td>-</td>
<td>1062</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. meningitidis</em> NMB</td>
<td>pilE</td>
<td>1</td>
<td>-</td>
<td>669</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. meningitidis</em> CKNM397</td>
<td>pilE</td>
<td>1</td>
<td>-</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. meningitidis</em> CKNM397</td>
<td>pilE</td>
<td>2</td>
<td>-</td>
<td>220</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. gonorrhoeae</em> FA1090</td>
<td>opaK</td>
<td>1</td>
<td>159</td>
<td>208</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>N. gonorrhoeae</em> FA1090</td>
<td>opaK</td>
<td>2</td>
<td>87</td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assembly</td>
<td>Donor sequence</td>
<td>% identity</td>
<td>RecA- frequency (x10^{-3})^a</td>
<td>RecA+ frequency (x10^{-3})^a</td>
<td>RecA+/RecA^- ratio^b</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td>Seq1</td>
<td>Reference</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Seq3</td>
<td>pilS1c4 or pilS3c1</td>
<td>100</td>
<td>0.062</td>
<td>10.4</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td>Seq6</td>
<td>pilS6c1</td>
<td>100</td>
<td>0.021</td>
<td>7.19</td>
<td>342</td>
<td></td>
</tr>
<tr>
<td>Seq7</td>
<td>pilS2c4</td>
<td>100</td>
<td>0.070</td>
<td>6.42</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Seq8</td>
<td>pilS3c3</td>
<td>99</td>
<td>0.104</td>
<td>3.78</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Seq10</td>
<td>pilS3c1</td>
<td>100</td>
<td>0.027</td>
<td>15.1</td>
<td>559</td>
<td></td>
</tr>
<tr>
<td>Seq11</td>
<td>pilS1c2</td>
<td>100</td>
<td>0.088</td>
<td>4.10</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Seq12</td>
<td>pilS7c1</td>
<td>100</td>
<td>0.018</td>
<td>2.68</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>Seq14</td>
<td>pilS7c1 or pilS6c2</td>
<td>100</td>
<td>0.019</td>
<td>5.91</td>
<td>311</td>
<td></td>
</tr>
<tr>
<td>Seq15</td>
<td>pilS6c1</td>
<td>100</td>
<td>0.038</td>
<td>7.50</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>Seq16</td>
<td>pilS6c2</td>
<td>100</td>
<td>0.013</td>
<td>2.86</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>Seq18</td>
<td>pilS1c1</td>
<td>100</td>
<td>0.034</td>
<td>5.96</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Seq19</td>
<td>pilS6c2</td>
<td>100</td>
<td>0.027</td>
<td>4.66</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>Seq22</td>
<td>pilS2c1</td>
<td>99</td>
<td>0.023</td>
<td>4.19</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>Seq23</td>
<td>pilS1c5</td>
<td>100</td>
<td>0.057</td>
<td>2.36</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Seq24</td>
<td>pilS3c3</td>
<td>100</td>
<td>0.104</td>
<td>3.48</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Seq25</td>
<td>pilS1c4</td>
<td>100</td>
<td>0.011</td>
<td>3.67</td>
<td>334</td>
<td></td>
</tr>
<tr>
<td>Seq30</td>
<td>pilS1c4</td>
<td>100</td>
<td>0.075</td>
<td>0.92</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Seq33</td>
<td>pilS3c3</td>
<td>100</td>
<td>0.010</td>
<td>2.31</td>
<td>231</td>
<td></td>
</tr>
<tr>
<td>Seq34</td>
<td>pilS7c1</td>
<td>100</td>
<td>0.024</td>
<td>1.78</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Seq35</td>
<td>pilS2c4</td>
<td>100</td>
<td>0.009</td>
<td>0.60</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Seq40</td>
<td>pilS7c1</td>
<td>100</td>
<td>0.015</td>
<td>4.76</td>
<td>317</td>
<td></td>
</tr>
<tr>
<td>Seq41</td>
<td>pilS1c1</td>
<td>99</td>
<td>0.043</td>
<td>5.25</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Seq43</td>
<td>pilS1c5</td>
<td>96</td>
<td>0.003</td>
<td>0.81</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>Seq44</td>
<td>mosaic</td>
<td>-</td>
<td>0.014</td>
<td>3.23</td>
<td>231</td>
<td></td>
</tr>
<tr>
<td>Seq47</td>
<td>pilS3c3 or pilS6c3</td>
<td>97</td>
<td>0.019</td>
<td>3.02</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>Seq48</td>
<td>pilE2</td>
<td>100</td>
<td>0.031</td>
<td>1.22</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>Seq49</td>
<td>pilS2c5</td>
<td>97</td>
<td>0.043</td>
<td>5.50</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>Seq50</td>
<td>pilS1c2</td>
<td>100</td>
<td>0.011</td>
<td>4.97</td>
<td>452</td>
<td></td>
</tr>
</tbody>
</table>

^a The average k-mer depth of the variant portion of the assembly divided by the average k-mer depth of same-sized regions immediately on either side of the variant sequence.  
^b The frequency in the presence of RecA divided by the frequency in the absence of RecA.
Table D.3: Variant sequences detected in the repeat experiment with *pilE* from *N. gonorrhoeae* FA1090.

<table>
<thead>
<tr>
<th>Assembly</th>
<th>Donor sequence</th>
<th>% identity</th>
<th>RecA⁻ frequency (x10^-3)ᵃ</th>
<th>RecA⁺ frequency (x10^-3)ᵃ</th>
<th>RecA⁺/RecA⁻ ratioᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq1</td>
<td>Reference</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Seq3</td>
<td>pilS2c4</td>
<td>100</td>
<td>0.10</td>
<td>12.8</td>
<td>128</td>
</tr>
<tr>
<td>Seq8</td>
<td>pilS6c1</td>
<td>100</td>
<td>0.027</td>
<td>6.11</td>
<td>226</td>
</tr>
<tr>
<td>Seq9</td>
<td>pilS6c2</td>
<td>100</td>
<td>0.075</td>
<td>3.91</td>
<td>52.1</td>
</tr>
<tr>
<td>Seq12</td>
<td>pilS1c2</td>
<td>100</td>
<td>0.018</td>
<td>3.43</td>
<td>191</td>
</tr>
<tr>
<td>Seq17</td>
<td>pilS7c1</td>
<td>100</td>
<td>0.244</td>
<td>1.60</td>
<td>6.56</td>
</tr>
</tbody>
</table>

ᵃ The average k-mer depth of the variant portion of the assembly divided by the average k-mer depth of same-sized regions immediately on either side of the variant sequence.

ᵇ The frequency in the presence of RecA divided by the frequency in the absence of RecA.
Table D.4: Variant sequences detected in the first experiment with *pilE* from *N. gonorrhoeae* MS11.

<table>
<thead>
<tr>
<th>Assembly</th>
<th>Donor sequence</th>
<th>Accession number</th>
<th>% identity</th>
<th>RecA⁻ frequency (x10⁻³)ᵃ</th>
<th>RecA⁺ frequency (x10⁻³)ᵃ</th>
<th>RecA⁺/RecA⁻ ratioᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq1</td>
<td>Reference</td>
<td>K02078</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Seq3</td>
<td>pilS1c2</td>
<td>M11663</td>
<td>100</td>
<td>0.40</td>
<td>2.26</td>
<td>5.65</td>
</tr>
<tr>
<td>Seq4</td>
<td>Uncertainc</td>
<td>X58404</td>
<td>100</td>
<td>0.45</td>
<td>2.56</td>
<td>5.69</td>
</tr>
<tr>
<td>Seq6</td>
<td>pilS5c1</td>
<td>X60748</td>
<td>100</td>
<td>0.34</td>
<td>1.70</td>
<td>5.00</td>
</tr>
<tr>
<td>Seq8</td>
<td>pilS1c1</td>
<td>M11663</td>
<td>100</td>
<td>0.22</td>
<td>1.73</td>
<td>7.86</td>
</tr>
<tr>
<td>Seq9</td>
<td>pilS1c4</td>
<td>M11663</td>
<td>100</td>
<td>0.27</td>
<td>1.12</td>
<td>4.15</td>
</tr>
<tr>
<td>Seq10</td>
<td>pilS7c1</td>
<td>X60750</td>
<td>100</td>
<td>1.22</td>
<td>5.48</td>
<td>4.49</td>
</tr>
<tr>
<td>Seq11</td>
<td>pilS1c1 or pilS1c3</td>
<td>M11663</td>
<td>100</td>
<td>0.70</td>
<td>1.25</td>
<td>1.79</td>
</tr>
<tr>
<td>Seq12</td>
<td>Uncertainc</td>
<td>M18245</td>
<td>98</td>
<td>0.16</td>
<td>1.45</td>
<td>9.06</td>
</tr>
<tr>
<td>Seq13</td>
<td>pilS7c1</td>
<td>X60750</td>
<td>100</td>
<td>0.32</td>
<td>1.68</td>
<td>5.25</td>
</tr>
<tr>
<td>Seq17</td>
<td>pilS6c1</td>
<td>X60749</td>
<td>100</td>
<td>0.17</td>
<td>1.52</td>
<td>8.94</td>
</tr>
<tr>
<td>Seq19</td>
<td>pilS1c4</td>
<td>M11663</td>
<td>100</td>
<td>0.21</td>
<td>1.01</td>
<td>4.81</td>
</tr>
<tr>
<td>Seq22</td>
<td>pilS5c1</td>
<td>X60748</td>
<td>100</td>
<td>0.19</td>
<td>1.03</td>
<td>5.42</td>
</tr>
<tr>
<td>Seq28</td>
<td>pilS1c2</td>
<td>M11663</td>
<td>100</td>
<td>0.21</td>
<td>0.82</td>
<td>3.90</td>
</tr>
<tr>
<td>Seq42</td>
<td>pilS7c1</td>
<td>X60750</td>
<td>100</td>
<td>0.40</td>
<td>0.43</td>
<td>1.08</td>
</tr>
<tr>
<td>Seq48</td>
<td>pilS1c6</td>
<td>M11663</td>
<td>97</td>
<td>0.28</td>
<td>0.43</td>
<td>1.54</td>
</tr>
</tbody>
</table>

ᵃ The average k-mer depth of the variant portion of the assembly divided by the average k-mer depth of same-sized regions immediately on either side of the variant sequence.

ᵇ The frequency in the presence of RecA divided by the frequency in the absence of RecA.

c Not all silent loci in strain MS11 have been sequenced and annotated. The best database match is given.
Table D.5: Variant sequences detected in the repeat experiment with *pilE* from *N. gonorrhoeae* MS11.

<table>
<thead>
<tr>
<th>Assembly</th>
<th>Donor sequence</th>
<th>Accession number</th>
<th>% identity</th>
<th>RecA⁻ frequency (x10⁻³)ᵃ</th>
<th>RecA⁺ frequency (x10⁻³)ᵃ</th>
<th>RecA⁺/RecA⁻ ratioᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq1</td>
<td>Reference</td>
<td>K02078</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Seq7</td>
<td><em>pilS5c1</em></td>
<td>X60748</td>
<td>100</td>
<td>0.44</td>
<td>1.96</td>
<td>4.45</td>
</tr>
<tr>
<td>Seq9</td>
<td>Uncertainᶜ</td>
<td>X66827</td>
<td>100</td>
<td>0.55</td>
<td>2.81</td>
<td>5.11</td>
</tr>
<tr>
<td>Seq11</td>
<td>Uncertainᶜ</td>
<td>X58404</td>
<td>100</td>
<td>0.42</td>
<td>2.23</td>
<td>5.31</td>
</tr>
<tr>
<td>Seq13</td>
<td>Uncertainᶜ</td>
<td>M18245</td>
<td>98</td>
<td>0.18</td>
<td>1.63</td>
<td>9.06</td>
</tr>
<tr>
<td>Seq17</td>
<td><em>pilS6c1</em></td>
<td>X60749</td>
<td>100</td>
<td>0.12</td>
<td>1.23</td>
<td>10.3</td>
</tr>
<tr>
<td>Seq28</td>
<td><em>pilS7c1</em></td>
<td>X60750</td>
<td>100</td>
<td>0.16</td>
<td>1.90</td>
<td>11.9</td>
</tr>
<tr>
<td>Seq49</td>
<td><em>pilS6c1</em></td>
<td>X60749</td>
<td>100</td>
<td>0.022</td>
<td>0.61</td>
<td>27.7</td>
</tr>
</tbody>
</table>

ᵃ The average k-mer depth of the variant portion of the assembly divided by the average k-mer depth of same-sized regions immediately on either side of the variant sequence.

ᵇ The frequency in the presence of RecA divided by the frequency in the absence of RecA.

ᶜ Not all silent loci in strain MS11 have been sequenced and annotated. The best database match is given.
Table D.6: Variation in the opaK gene of *N. gonorrhoeae* strain FA1090

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Assembly</th>
<th>RecA frequency (x10⁻³)ᵃ</th>
<th>RecA⁺ frequency (x10⁻³)ᵃ</th>
<th>RecA⁺/RecA⁻ ratioᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Seq1ᶜ</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>Seq14</td>
<td>6.21</td>
<td>10.24</td>
<td>1.65</td>
</tr>
<tr>
<td>1</td>
<td>Seq39</td>
<td>3.94</td>
<td>6.50</td>
<td>1.65</td>
</tr>
<tr>
<td>1</td>
<td>Seq44</td>
<td>4.06</td>
<td>6.56</td>
<td>1.62</td>
</tr>
<tr>
<td>2</td>
<td>Seq1ᶜ</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Seq15</td>
<td>0.35</td>
<td>9.13</td>
<td>26.16</td>
</tr>
<tr>
<td>2</td>
<td>Seq46</td>
<td>1.05</td>
<td>5.85</td>
<td>5.57</td>
</tr>
</tbody>
</table>

ᵃ The average k-mer depth of the variant portion of the assembly divided by the average k-mer depth of same-sized regions immediately on either side of the variant sequence.

ᵇ The frequency in the presence of RecA divided by the frequency in the absence of RecA.

ᶜ Seq1 was identical to the reference sequence.
Figure S1: K-mer depth across assemblies containing variant sequence segments. The green trace depicts the k-mer depth in an amplicon derived from a culture grown in the presence of RecA, and therefore antigenic variation. The blue trace shows the k-mer depth for an amplicon obtained from a culture grown in the absence of RecA and therefore antigenic variation. The vertical pale green bar shows the variant segment of the assembly across which the k-mer depth was averaged, and the vertical grey bars the same-sized conserved sequences on either side across which the k-mer depth was averaged, for the frequency calculation. A. An assembly containing a single variant segment. B. An assembly where a conserved region of more than k bases interrupts a variant segment, resulting in a spike of k-mer depth (marked with a downward arrow) that would result in an artificially high frequency calculation if the k-mer depth was averaged across the entire region.
**Figure D.S2**: Alignment of the sequence of the pilE gene from the stock of *N. gonorrhoeae* strain FA1090 used in these experiments (top), and the FA1090 genome sequence (bottom). Blue text indicates sequence flanking the pilE gene (black text). Sequence differences are highlighted in yellow.
| seq6  | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq22 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq12 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq33 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq16 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq35 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq25 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq10 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq50 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq1  | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq48 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq14 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq40 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq9  | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq43 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq30 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq34 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq11 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq7  | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq49 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq15 | TATTCTAACGGTAAATTCAAAATTCTCAAAATTCTCAATGGCCAATTACAAACCACAGTACC 60 |
| seq35 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq16 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq33 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq12 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq22 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq6  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq3  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq15 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq44 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq23 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq41 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq18 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq24 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq8  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq40 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq14 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq48 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq3  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq11 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq34 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq19 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq43 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq49 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq10 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq50 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq30 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq3  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq1  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq50 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq35 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq16 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq33 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq12 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq22 | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq6  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
| seq3  | CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC 120 |
Figure D.S3: Alignment of the variant sequences detected in the first experiment with pilE in N. gonorrhoeae FA1090. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the pilE gene (black text). Sequence differences are highlighted in yellow.
seq1  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60
seq8  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60
seq3  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60
seq12  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60
seq9  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60
seq17  TATTCTAACGCGTAAATTCAAAAATCTCAAATTCCGACCCAATCAACACACCCGATACCC  60

********************

seq1  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC  120
seq8  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC  120
seq3  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAAC
seq12  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC  120
seq9  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC  120
seq17  CATGCCAATAAAAAAGTAACGAAAATCGGCACTAAAACTGACAATTTTCGACACTGCCGC  120

**************************************

seq1  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180
seq8  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180
seq3  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180
seq12  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180
seq9  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180
seq17  CCCCTACTTCCGCAAACCACACCCACCTAAAAGAAAATACAAAATAAAAACAATTATATA  180

************************************************************

seq1  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240
seq8  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240
seq3  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240
seq12  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240
seq9  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240
seq17  GAGATAAACGCATAAAATTTCACCTCAAAACATAAAATCGGCACGAATCTTGCTTTATAA  240

************************************************************

seq1  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG
seq8  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG
seq3  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG
seq12  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG
seq9  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG
seq17  TACGCAGTTGTCGCAACAAAAAACCGATGGTTAAATACATTG

************************************************************

seq1  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360
seq8  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360
seq3  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360
seq12  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360
seq9  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360
seq17  AGCCTGAGGCATTTCCCCTTTCAATTAGGAGTAATTTTATGAATACCCTTCAAAAAGGCT  360

*************

seq1  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420
seq8  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420
seq3  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420
seq12  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420
seq9  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420
seq17  TTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCC  420

************************************************************

seq1  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480
seq8  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480
seq3  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480
seq12  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480
seq9  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480
seq17  TTCCCGCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCG  480

************************************************************

seq1  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540
seq8  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540
seq3  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540
seq12  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540
seq9  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540
seq17  AAGGTCAAAAATCAGCCGTTACCGGGTATTACCTGAATCACGGCATATGGCCGGAAGACA  540

************************************************************

seq1  ACACCTTTCGCGCGTGGCATCCCCCCTCCGACATCAAAGGCAAATATGTTCAAAGCGT  600
seq8  ACACCTTTCGCGCGTGGCATCCCCCCTCCGACATCAAAGGCAAATATGTTCAAAGCGT  600
seq3  ACACCTTTCGCGCGTGGCATCCCCCCTCCGACATCAAAGGCAAATATGTTCAAAGCGT  600

**Figure D.S4:** Alignment of the variant sequences detected in the repeat experiment with pilE in N. gonorrhoeae FA1090. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the pilE gene (black text). Sequence differences are highlighted in yellow.

<table>
<thead>
<tr>
<th>Seq</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>seq1</td>
<td>CCGCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1175</td>
</tr>
<tr>
<td>seq8</td>
<td>CCGCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1181</td>
</tr>
<tr>
<td>seq3</td>
<td>CC GCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1175</td>
</tr>
<tr>
<td>seq12</td>
<td>CCGCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1170</td>
</tr>
<tr>
<td>seq9</td>
<td>CCGCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1190</td>
</tr>
<tr>
<td>seq17</td>
<td>CCGCCCTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAA</td>
<td>1193</td>
</tr>
</tbody>
</table>
Figure D.S5: Alignment of the variant sequences detected in the first experiment with pilE in N. gonorrhoeae MS11. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the pilE gene (black text). Sequence differences are highlighted in turquoise.
Figure D.S6: Alignment of the variant sequences detected in the repeat experiment with pilE in N. gonorrhoeae MS11. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the pilE gene (black text). Sequence differences are highlighted in turquoise.
seq1  AAAAAATCTAAGGTCAAAGCGTGTACCGCTCGCGGAAAATGAAACCAAGC 60
seq44  AAAAAATCTAAGGTCAAAGCGTGTACCGCTCGCGGAAAATGAAACCAAGC 60
seq14  AAAAAATCTAAGGTCAAAGCGTGTACCGCTCGCGGAAAATGAAACCAAGC 60
seq39  AAAAAATCTAAGGTCAAAGCGTGTACCGCTCGCGGAAAATGAAACCAAGC 60

************************************************************
seq1    GGCGTAAACAAAGAAATCAAAGGCAAAAAACTCTCCCTGTGGGCCAAGCGTGAAGACGGT 120
seq44   GGCGTAAACAAAGAAATCAAAGGCAAAAAACTCTCCCTGTGGGCCAAGCGTGAAGACGGT 120
seq14   GGCGTAAACAAAGAAATCAAAGGCAAAAAACTCTCCCTGTGGGCCAAGCGTGAAGACGGT 120
seq39   GGCGTAAACAAAGAAATCAAAGGCAAAAAACTCTCCCTGTGGGCCAAGCGTGAAGACGGT 120

************************************************************
seq1    TCGGTAAAATGGTTCTGCGGACAGCCGGTTAAGCGCGACGCCGGCGCCAAAGCCGACGAC 180
seq44   TCGGTAAAATGGTTCTGCGGACAGCCGGTTAAGCGCGACGCCGGCGCCAAAGCCGACGAC 180
seq14   TCGGTAAAATGGTTCTGCGGACAGCCGGTTAAGCGCGACGCCGGCGCCAAAGCCGACGAC 180
seq39   TCGGTAAAATGGTTCTGCGGACAGCCGGTTAAGCGCGACGCCGGCGCCAAAGCCGACGAC 180

************************************************************
seq1    GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240
seq44   GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240
seq14   GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240
seq39   GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240

************************************************************
seq1    GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300
seq44   GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300
seq14   GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300
seq39   GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300

************************************************************
seq1    GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360
seq44   GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360
seq14   GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360
seq39   GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360

************************************************************
seq1    GGCGGGGTCGTCCGTTCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420
seq44   GGCGGGGTCGTCCGTTCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420
seq14   GGCGGGGTCGTCCGTTCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420
seq39   GGCGGGGTCGTCCGTTCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

************************************************************
seq1    GCCTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAATCCGCCCCCTC 480
seq44   GCCTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAATCCGCCCCCTC 480
seq14   GCCTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAATCCGCCCCCTC 480
seq39   GCCTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAATCCGCCCCCTC 480

************************************************************
seq1    ATTTGTTAATCCGCCATATTGTATTGAAACACCGCCCGGAACCCGATATAATCCGCCCTT 540
seq44   ATTTGTTAATCCGCCATATTGTATTGAAACACCGCCCGGAACCCGATATAATCCGCCCTT 540
seq14   ATTTGTTAATCCGCCATATTGTATTGAAACACCGCCCGGAACCCGATATAATCCGCCCTT 540
seq39   ATTTGTTAATCCGCCATATTGTATTGAAACACCGCCCGGAACCCGATATAATCCGCCCTT 540

************************************************************
seq1    CAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATC 600
seq44   CAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATC 600
seq14   CAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATC 600
seq39   CAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATC 600

************************************************************
seq1    CTTCTCTTCGGCAGCGCAGGCGGCAAGTGGAAGGCAATGGCCGCGGCCCGTATGTGCAGGC 720
seq44   CTTCTCTTCGGCAGCGCAGGCGGCAAGTGGAAGGCAATGGCCGCGGCCCGTATGTGCAGGC 720
seq14   CTTCTCTTCGGCAGCGCAGGCGGCAAGTGGAAGGCAATGGCCGCGGCCCGTATGTGCAGGC 720
seq39   CTTCTCTTCGGCAGCGCAGGCGGCAAGTGGAAGGCAATGGCCGCGGCCCGTATGTGCAGGC 720

************************************************************
seq1    GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 780
seq44   GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 780
seq14   GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 780
seq39   GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 780
**Figure D.S7:** Alignment of the variant sequences detected in the first experiment with opaK in N. gonorrhoeae FA1090. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the opaK gene (black text). Sequence differences are highlighted in yellow.
seq1
AAATATGTTCAAGCCTGCTGAAAACGCCGCTGTTACCGCGAATGAAACCGAAGC 60

seq46
AAATATGTTCAAAGCGTTACGGTCGCAAACGGCGTCGTTACCGCCGAAATGAAACCAAGC 60

seq15
AAATATGTTCAAGCCTGCTGAAAACGCCGCTGTTACCGCGAATGAAACCGAAGC 60

seq1
GGCGTAACAAAGGAAATCAAAGGCAAAAACTCTCCCTTGGCAAGCGTGAAGACGGT 120

seq46
GGCGTAACAAAGGAAATCAAAGGCAAAAACTCTCCCTTGGCAAGCGTGAAGACGGT 120

seq15
GGCGTAACAAAGGAAATCAAAGGCAAAAACTCTCCCTTGGCAAGCGTGAAGACGGT 120

seq1
TCGGTAATAAGTCTCTGGCGACAGCGGTATAAGGCGACGCGCGCCAAAGCAGGAC 180

seq46
TCGGTAATAAGTCTCTGGCGACAGCGGTATAAGGCGACGCGCGCCAAAGCAGGAC 180

seq15
TCGGTAATAAGTCTCTGGCGACAGCGGTATAAGGCGACGCGCGCCAAAGCAGGAC 180

seq1
GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240

seq46
GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240

seq15
GTCAAAGCCGACGCCGCCAACGCCATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGAT 240

seq1
GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300

seq46
GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300

seq15
GAATCATCTGCCACCTAAGGCAAATTAGGCCTTAAATTTTAAATAAATCAAGCGGTAAGT 300

seq1
GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360

seq46
GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360

seq15
GATTTCCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCG 360

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq46
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq15
GGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGT 420

seq1
GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 775

seq46
GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 775

seq15
GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGCTCCAGG 775
Figure D.S8: Alignment of the variant sequences detected in the repeat experiment with opaK in N. gonorrhoeae FA1090. The seq1 assembly is identical to the reference sequence obtained by Sanger sequencing of the amplicon. Blue text indicates sequence flanking the opaK gene (black text). Sequence differences are highlighted in yellow.
<table>
<thead>
<tr>
<th>Target gene</th>
<th>Oligonucleotide number</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pilE</td>
<td>DAP2008</td>
<td>CCCTATTCTAACCGCGTAAATTTC</td>
</tr>
<tr>
<td></td>
<td>DAP2005</td>
<td>CATTTTCGGCTCCTTTATTCGG</td>
</tr>
<tr>
<td>opaJ (NGO1922)</td>
<td>DAP2006</td>
<td>ACCATATCAAAGCAGATTGGGAA</td>
</tr>
<tr>
<td></td>
<td>DAP2007</td>
<td>CGTTATTTATACAAGACCTGCAAG</td>
</tr>
<tr>
<td>opaK (NGO2132)</td>
<td>DAP2066</td>
<td>GCCAATATGTTCAAAACGCTT</td>
</tr>
<tr>
<td></td>
<td>DAP2067</td>
<td>CGTTTGGAGGACGACTGCTG</td>
</tr>
<tr>
<td>N. gonorrhoeae MS11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pilE</td>
<td>DAP2113</td>
<td>TCCCCCTTTCAATTTAGGAGT</td>
</tr>
<tr>
<td></td>
<td>DAP2114</td>
<td>TCGATATATATTCCACCC</td>
</tr>
<tr>
<td>N. meningitidis NMB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pilE (class II)</td>
<td>DAP2092</td>
<td>GTCACAACGCAAGAAAGC</td>
</tr>
<tr>
<td></td>
<td>DAP2094</td>
<td>CTTGTTTTCAGGCGC</td>
</tr>
<tr>
<td>pilE (MC58 class I)</td>
<td>DAP2095</td>
<td>AATTACTCAAACCCAGC</td>
</tr>
<tr>
<td></td>
<td>DAP2096</td>
<td>GACTTTTTGGAACACATCC</td>
</tr>
<tr>
<td>N. meningitidis FAM18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pilE (class II)</td>
<td>DAP2092</td>
<td>GTCACAACGCAAGAAAGC</td>
</tr>
<tr>
<td></td>
<td>DAP2093</td>
<td>CGGAATCCCAAACGGTTTC</td>
</tr>
</tbody>
</table>

Table D.S1: Oligonucleotide primers used to amplify the genes of interest.
Table D.S2. Sequence quality data.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Gene</th>
<th>Experiment</th>
<th>Condition</th>
<th>k used</th>
<th>Raw pair count</th>
<th>Read length</th>
<th>Q10 average length</th>
<th>Q20 average length</th>
<th>Q30 average length</th>
<th>Pairs longer than k after clipping</th>
<th>Single reads longer than k after clipping</th>
<th>Average read length of read longer than k after clipping</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA1090</td>
<td>pilE</td>
<td>1</td>
<td>RecA−</td>
<td>30</td>
<td>3570393</td>
<td>101</td>
<td>86.7</td>
<td>73.9</td>
<td>52.1</td>
<td>2871938</td>
<td>391521</td>
<td>76.8</td>
</tr>
<tr>
<td>FA1090</td>
<td>pilE</td>
<td>1</td>
<td>RecA+</td>
<td>30</td>
<td>15423940</td>
<td>101</td>
<td>86.5</td>
<td>73.6</td>
<td>51.6</td>
<td>12414274</td>
<td>1674442</td>
<td>76.8</td>
</tr>
<tr>
<td>FA1090</td>
<td>pilE</td>
<td>2</td>
<td>RecA−</td>
<td>50</td>
<td>8231485</td>
<td>101</td>
<td>62.7</td>
<td>49.9</td>
<td>34.3</td>
<td>5614198</td>
<td>411232</td>
<td>75.5</td>
</tr>
<tr>
<td>FA1090</td>
<td>pilE</td>
<td>2</td>
<td>RecA+</td>
<td>50</td>
<td>5393387</td>
<td>101</td>
<td>61.9</td>
<td>48.8</td>
<td>33.7</td>
<td>3593646</td>
<td>314371</td>
<td>75.3</td>
</tr>
<tr>
<td>MS11</td>
<td>pilE</td>
<td>1</td>
<td>RecA−</td>
<td>30</td>
<td>4841714</td>
<td>76</td>
<td>51.7</td>
<td>45.7</td>
<td>34.5</td>
<td>3749972</td>
<td>420492</td>
<td>49.0</td>
</tr>
<tr>
<td>MS11</td>
<td>pilE</td>
<td>1</td>
<td>RecA+</td>
<td>30</td>
<td>3868814</td>
<td>76</td>
<td>52.4</td>
<td>46.4</td>
<td>35.0</td>
<td>3668367</td>
<td>343934</td>
<td>49.6</td>
</tr>
<tr>
<td>MS11</td>
<td>pilE</td>
<td>2</td>
<td>RecA−</td>
<td>30</td>
<td>7684482</td>
<td>79</td>
<td>44.4</td>
<td>28.7</td>
<td>11.0</td>
<td>4139825</td>
<td>146614</td>
<td>48.2</td>
</tr>
<tr>
<td>MS11</td>
<td>pilE</td>
<td>2</td>
<td>RecA+</td>
<td>30</td>
<td>7691690</td>
<td>79</td>
<td>45.2</td>
<td>29.2</td>
<td>11.2</td>
<td>4303951</td>
<td>1445555</td>
<td>48.6</td>
</tr>
<tr>
<td>FAM18</td>
<td>pilE</td>
<td>1</td>
<td>RecA+</td>
<td>30</td>
<td>22239209</td>
<td>75</td>
<td>52.8</td>
<td>46.0</td>
<td>35.0</td>
<td>13278485</td>
<td>4733138</td>
<td>51.6</td>
</tr>
<tr>
<td>FAM18</td>
<td>pilE</td>
<td>2</td>
<td>RecA+</td>
<td>30</td>
<td>12089013</td>
<td>101</td>
<td>86.4</td>
<td>70.9</td>
<td>46.4</td>
<td>9679915</td>
<td>1104181</td>
<td>75.5</td>
</tr>
<tr>
<td>NMB</td>
<td>pilE</td>
<td>1</td>
<td>RecA+</td>
<td>30</td>
<td>11509861</td>
<td>101</td>
<td>82.9</td>
<td>67.1</td>
<td>43.1</td>
<td>8704569</td>
<td>1270132</td>
<td>73.2</td>
</tr>
<tr>
<td>CKNM397</td>
<td>pilE</td>
<td>1</td>
<td>RecA+</td>
<td>50</td>
<td>7237777</td>
<td>101</td>
<td>66.0</td>
<td>53.2</td>
<td>36.8</td>
<td>5232278</td>
<td>325200</td>
<td>77.5</td>
</tr>
<tr>
<td>CKNM397</td>
<td>pilE</td>
<td>2</td>
<td>RecA+</td>
<td>50</td>
<td>5562633</td>
<td>101</td>
<td>65.0</td>
<td>51.8</td>
<td>35.5</td>
<td>3939626</td>
<td>254958</td>
<td>76.9</td>
</tr>
<tr>
<td>FA1090</td>
<td>opaK</td>
<td>1</td>
<td>RecA−</td>
<td>50</td>
<td>12125942</td>
<td>101</td>
<td>68.0</td>
<td>55.2</td>
<td>39.7</td>
<td>6091224</td>
<td>877613</td>
<td>74.3</td>
</tr>
<tr>
<td>FA1090</td>
<td>opaK</td>
<td>1</td>
<td>RecA+</td>
<td>50</td>
<td>14154924</td>
<td>101</td>
<td>69.4</td>
<td>56.4</td>
<td>40.6</td>
<td>7572874</td>
<td>1052849</td>
<td>74.4</td>
</tr>
<tr>
<td>FA1090</td>
<td>opaK</td>
<td>2</td>
<td>RecA−</td>
<td>50</td>
<td>5369370</td>
<td>101</td>
<td>59.1</td>
<td>46.0</td>
<td>31.2</td>
<td>3381187</td>
<td>280238</td>
<td>72.7</td>
</tr>
<tr>
<td>FA1090</td>
<td>opaK</td>
<td>2</td>
<td>RecA+</td>
<td>50</td>
<td>6921549</td>
<td>101</td>
<td>59.2</td>
<td>46.5</td>
<td>31.8</td>
<td>4380718</td>
<td>354293</td>
<td>73.1</td>
</tr>
</tbody>
</table>

a the average length before the first base with quality less than 10 (10% chance of miscalled base) is encountered.
b the average length before the first base with quality less than 20 (1% chance of miscalled base) is encountered.
c the average length before the first base with quality less than 30 (0.1% chance of miscalled base) is encountered.