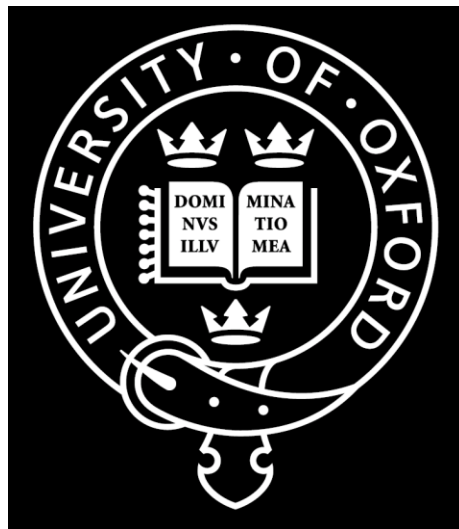


Interactions of *Neisseria meningitidis* with the Human Immune System

Rachel Harding

Sir William Dunn School of Pathology

Wadham College



A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy at the University of Oxford

Trinity Term 2014

This thesis is dedicated to my loving husband Timothy Harding, who has offered me encouragement and support through each step of my DPhil.

Interactions of *Neisseria meningitidis* with the Human Immune System

Rachel Harding

DPhil Pathology, Trinity Term 2014

Sir William Dunn School of Pathology, Wadham College

Abstract

Neisseria meningitidis is an obligate human pathogen causing over 1000 cases of meningococcal disease within the U.K., 10 % of which result in long-term disability or fatality. 10-70 % of the population carry *N. meningitidis* in their nasopharynx, the natural reservoir of this bacterium, as a commensal. The host-pathogen interactions of this species are complex and a greater understanding of the molecular mechanisms involved in pathogenesis and immune evasion is required. Three aspects of *N. meningitidis* pathogenesis were explored in this study.

One mechanism of immune evasion which promotes serum resistance of *N. meningitidis* is recruitment of complement factor H through domains 6 and 7 (fH₆₇) by factor H binding protein (fHbp). In this study, mouse fH₆₇ was recombinantly expressed and purified. fHbp did not bind mouse fH₆₇ at physiologically relevant protein concentrations. The structure of mouse fH₆₇ was solved, showing differences in domain orientation and surface chemistry compared to the human version of this protein, potentially accounting for the host specificity of this interaction.

Type IV pili (T4P) are crucial adhesins of *N. meningitidis*, the fibre of which is composed of thousands of copies of PilE. A method was developed to recombinantly produce large quantities of this protein from a variety of meningococcal strains and the structure was solved of one PilE protein. Subsequent analysis was performed with the PilE proteins investigating their interaction with the putative pilus receptor CD46 and human epithelia as well as their immunogenicity. A method was also established to produce PilC, the proposed tip-associated adhesin of T4P.

ZapE has recently been identified as an important protein in pathogen colonisation, functioning as an ATPase linked to Z-ring formation in bacterial cell fission. Both *N. meningitidis* and *E. coli* ZapE were recombinantly produced. The domain boundaries were mapped and ATPase activity was confirmed. No interaction was seen with FtsZ but DNA binding and modulation was observed by shift assays, the exact function of which remains to be elucidated in future studies.

Acknowledgements

Firstly I would like to thank my supervisor, Professor Susan Lea. Throughout both my undergraduate and graduate studies, Susan has been inspiring, encouraging and a great role model. I am so grateful for the opportunities she has given me as well as the consistent support she has shown me for the duration of my studies and research. Thanks to Susan, I am now very excited to pursue a career in biophysics and structural biology, I hope to do her proud.

All of the members of the Lea lab, both past and present, must be thanked for creating such an open and happy work environment, where no question was too ridiculous to ask and everyone was always on hand with words of advice or encouragement. In particular, Steven Johnson, whose critique of my work I always highly valued as well as his ideas for experiments to try or tweaks to make. His ability to absorb copious amounts of information and then create models or new hypotheses will always be something I aspire to. Patrizia Abrusci must also be thanked for being on hand to cure any cloning or protein purification problem with her seemingly endless knowledge and experience. Thanks also to Joseph Caesar, who supervised me in my early years in the lab, taught me most of what I know about X-ray crystallography and offered invaluable pieces of information during our chats in the office. The company of my fellow DPhil students James Lillington, Mel McDowell, Sarah Rollauer, Matthew Cottee, Paul Chappell and Violet Feng as well as the other members of the lab, has made lab work great fun, especially over the background murmurs of BBC 6 Music.

I have also been extremely fortunate to have been given the opportunity to work alongside Rachel Exley, as an honorary member of Chris Tang's lab. Rachel is an excellent teacher and has patiently taught me no end of new techniques as well as offering fantastic insight into so many aspects of Neisserial biology. Thanks goes to both Chris and Rachel for letting me be a part of "Team Pilin"!

My family have been supportive throughout my Oxford studies and have always been interested in my research. The love and pride they have shown towards me has been a great motivator and I hope to continue to make them proud in the future.

Finally, I would like to thank my husband Timothy Harding, who now knows more about X-ray crystallography than any other bicycle expert you will ever meet. His reassurance and help throughout my graduate studies has been invaluable and I am looking forward to the next stage of our lives together.

Publication resulting from this study

Design and Evaluation of Meningococcal Vaccines through Structure-Based Modification of Host and Pathogen Molecules

Steven Johnson,[#] Lionel Tan,[#] Stijn van der Veen,[#] Joseph Caesar,[#] Elena Goicoechea De Jorge,[#] Rachel J. Harding,[#] Xilian Bai, Rachel M. Exley, Philip N. Ward, Nicola Ruivo, Kaushali Trivedi, Elspeth Cumber, Rhian Jones, Luke Newham, David Staunton, Rafael Ufret-Vincenty, Ray Borrow, Matthew C. Pickering, Susan M. Lea, and Christoph M. Tang (2012) PLoS Pathogens, 8(10)e1002981 DOI: 10.1371/journal.ppat.1002981

[#] These authors contributed equally to this work

Abbreviations

APBS	Adaptive Poisson-Boltzmann Solver
CCP4	Collaborative Computational Project number 4
CD	Circular Dichroism
C _p	Specific Heat Capacity
CV	Column Volume
DAPI	4',6-diamidino-2-phenylindole
DATDH	2,4-diacetamido 2,4,6-trideoxy α -D-hexose
DNase	Deoxyribonuclease
DSC	Differential Scanning Calorimetry
DTT	Dithiothreitol
EDC	Ethyl(dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assays
EM	Electron Microscopy
EMSA	Electrophoretic Mobility Shift Assay
FFAS	Fold and Function Assignment Server
fH	Factor H
fHbp	Factor H Binding Protein
GAG	Glycosaminoglycan
GATDH	2-acetamido 4-glyceramido 2,4,6-trideoxy- α -D-hexose
GST	Glutathione S-Transferase
h-fH	Human Factor H
HSQC	Heteronuclear Single Quantum Coherence
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Luria Bertani
LDH	Lactate Dehydrogenase
LLG	Log-Likelihood Gain
LPS	Lipopolysaccharide
m-fH	Mouse Factor H
MAC	Membrane Attack Complex
MALLS	Multiangle Laser Light Scattering
MBL	Mannose Binding Lectin

MCS	Multiple Cloning Site
MLST	Multilocus Sequence Typing
MR	Molecular Replacement
MWCO	Molecular Weight Cut Off
NHS	N-hydroxysuccinimide
NMR	Nuclear Magnetic Resonance
NMWL	Nominal Molecular Weight Limit
OD ₆₀₀	Optical Density at 600 nm
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline, 0.05% (v/v) Tween-20
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PFA	Paraformaldehyde
PGA	Poly-γ-glutamic acid
PilE	Pilin expression
PilC	Pilus associated protein C
RFZ	Rotation Function Z-score
RMSD	Root Mean Squared Deviation
ROS	Reactive Oxygen Species
RNS	Reactive Nitrogen Species
SBA	Serum Bactericidal Assay
SCR	Short Consensus Repeat
SDS	Sodium Dodecyl Sulphate
SEC	Size Exclusion Chromatography
SPR	Surface Plasmon Resonance
SSM	Secondary Structure Matching
TEER	Trans-epithelial Electrical Resistance
TFZ	Translation Function Z-score
Tris	Tris(hydroxymethyl)aminomethane
T4P	Type IV Pili
ZapE	Z-ring Associated Protein E

Contents

1. Introduction	1
1.1 Bacterial Pathogenicity.....	1
1.2 Studying Bacterial Pathogenesis using Biophysical Methods	4
1.3 <i>Neisseria meningitidis</i>	7
1.4 Immune Evasion and Host Specificity of <i>N. meningitidis</i>	10
1.4.1 Evasion of the Complement System.....	11
1.4.2 Host Specificity of <i>N. meningitidis</i>	14
1.4.3 THESIS AIMS:	15
1.5 Adhesion by <i>N. meningitidis</i>	16
1.5.1 An Introduction to Type IV Pili	17
1.5.2 Type IV Pilus Mediated Adhesion.....	25
1.5.3 Insights into Pile Function	27
1.5.4 THESIS AIMS:	28
1.6 Colonisation and Bacterial Cell Division in the Context of Infection by <i>Neisseria meningitidis</i>	30
1.6.1 THESIS AIMS:	31
 2. Investigations into Host Specificity of <i>N. meningitidis</i>	 32
2.1 Functional Analysis of m-fH ₆₇	34
2.1.1 Expression and Purification of m-fH ₆₇	34
2.1.2 Surface Plasmon Resonance Analysis.....	36
2.2 Structural Analysis of m-fH ₆₇	40
2.2.1 Crystallisation of m-fH ₆₇	40
2.2.2 Data Collection and Processing	42
2.2.3 Characterisation of the Asymmetric Unit.....	45
2.2.4 Phase Determination, Refinement and Structure Validation	45
2.2.5 Model Building and Refinement.....	47
2.2.6 The Structure of m-fH ₆₇	49
2.2.7 Structural Analysis of m-fH ₆₇	50
2.3 Perspectives.....	58

3. Structural Analysis of Meningococcal PilE	60
3.1 Expression, Purification and Characterisation of PilE from <i>N. meningitidis</i>	62
3.1.1 Construction of PilE Expression Vectors	65
3.1.2 Optimising Expression of the pETM- <i>pilE</i> Constructs	69
3.1.3 Large Scale Expression and Purification of PilE	71
3.2 Structural Analysis of PilE	76
3.2.1 Crystallisation of PilE	76
3.2.2 Data Collection and Processing	77
3.2.3 Phase Determination, Refinement and Structure Validation	79
3.2.4 Model Building and Refinement	79
3.2.5 The Structure of <i>N. meningitidis</i> PilE.....	80
3.2.6 Pursuit of a Class II Pilin Structure.....	83
3.2.7 Modelling of a Class II Pilin	84
3.2.8 Crystal Packing of PilE_1.MC58 ΔN24	88
3.2.9 Comparison with the Gonococcal MS11PilE Structure	91
3.3 Perspectives	97
 4. Functional Analysis of Meningococcal PilE	 101
4.0.1 Adhesion to host cell surfaces.....	101
4.0.2 Host cellular effects upon bacterial adhesion.....	102
4.0.3 Immunogenicity of PilE proteins	102
4.1 Investigations into the Interaction between PilE and CD46	104
4.1.1 Production of PilE and CD46	104
4.1.2 ¹ H ¹⁵ N-HSQC analysis of the Interaction Between PilE and CD46	106
4.1.3 SPR analysis of the Interaction Between PilE and CD46	108
4.1.4 Conclusions of PilE-CD46 Interaction Studies	109
4.2 Investigations into the Effect of PilE Incubation with Host Epithelia.....	111
4.2.1 Preparation of Protein Samples for Human Epithelial Cell Studies	111
4.2.2 Human Epithelial Cell Culture	113
4.2.3 Investigating the Effects of Purified Recombinant PilE on Epithelial Cell Monolayers by Fluorescence Microscopy.....	114
4.2.4 Investigating the Effects of Purified Recombinant PilE on Epithelial Cells Lysis	117
4.2.5 Investigating the Effects of Purified Recombinant PilE on Epithelial Cells by Cytokine Release	119

4.2.6 Investigating the Effects of Purified Recombinant PilE on Polarised Monolayers or Epithelial Cells by Transepithelial Electrical Resistance Assay	119
4.2.7 Conclusions of Investigations into the Role of PilE in Meningococcal Induced Cell Responses.....	121
4.3 Investigations into the Immunogenicity of PilE.....	123
4.3.1 Preparation of PilE Samples for Immunisation and Immunisation Schedule.....	123
4.3.2 Analysis of the Reactivity of Sera Raised Against PilE	124
4.4 Investigations into the Functional Qualities of Sera Raised against PilE.....	127
4.5 Perspectives.....	130
5. Expression and Purification of PilC	134
5.1 Attempts to Express and Purify PilC using the PilY1 Protocol	135
5.1.1 Sequence Analysis of PilC compared to PilY1.....	135
5.1.2 Expression of PilC-C using the PilY1 Protocol	137
5.2 Development of a Successful Protocol for PilC Expression and Purification.....	139
5.2.1 Expression in Disulphide Promoting Cell Lines.....	139
5.2.2 Construction of the pET32b-PilC Vectors for both Alleles	140
5.2.3 Optimising Expression of the pET32b-PilC-C Constructs.....	141
5.2.4 Large-Scale Expression and Purification of pET32b- PilC ₁₄₈₉₋₁₀₄₈	143
5.2.5 Mutagenesis of the Enterokinase Cleavage site to a 3C Cleavage Site in the pET32b- PilC1-C Construct.....	145
5.2.6 Optimisation of Buffer Conditions for pET32b- PilC ₁₄₈₉₋₁₀₄₈ Purification	146
5.3 Perspectives.....	153
6. Functional Characterisation of the Novel Cell Division Protein ZapE	156
6.1 Purification and Characterisation of ZapE from <i>E. coli</i> and <i>N. meningitidis</i>	158
6.1.1 Construction of pET28b Full-Length ZapE for both <i>E. coli</i> and <i>N. meningitidis</i>	159
6.1.2 Optimising Expression of the pET28b-ZapE Constructs	161
6.1.3 Large Scale Expression and Purification of <i>E. coli</i> and <i>N. meningitidis</i> ZapE.....	163
6.1.4 Mutagenesis of ZapE Walker Box Motif and Subsequent Purification of the Mutant Proteins	167
6.2 Attempts Towards Structural Characterisation of ZapE.....	169
6.2.1 Expression and Purification of <i>N. meningitidis</i> ZapE Selenomethionine Derivative	169
6.2.2 Crystallisation Attempts of Selenomethionine Derivatised <i>N. meningitidis</i> ZapE	172

6.2.3 FFAS analysis of the Protein Sequence	176
6.2.4 Limited Proteolysis of <i>N. meningitidis</i> ZapE	177
6.3 Functional Analysis of ZapE	185
6.3.1 Analysis of the ATPase Activity of ZapE	185
6.3.2 Investigations into FtsZ Binding by ZapE	190
6.3.3 Investigations into DNA Binding by ZapE	194
6.4 Perspectives	202
7. Thesis Summary	205
8. Appendix	207
8.1 Standard Laboratory Procedures	207
8.1.1 Standard Method of PCR	207
8.1.2 Site-Directed Mutagenesis Protocol	207
8.1.3 Transformation Protocol	208
8.1.4 TAE Agarose Gel Buffer System	208
8.1.5 SDS-PAGE Buffer System	208
9. References	210

List of Figures

1. Introduction

Figure 1.1 Outline of the complement cascade.....	12
Figure 1.2 The interaction between fH and fHbp.....	13
Figure 1.3 Components of the outer membrane of <i>N. meningitidis</i> involved in adhesion.....	17
Figure 1.4 Model of the T4P with current understanding of assembly.....	18
Figure 1.5 The structure of PilE from <i>N. gonorrhoeae</i> PDB ID:2PIL	20
Figure 1.6 Pseudoatomic resolution pilus model.....	21
Figure 1.7 Class I and class II pilE loci structure.....	22

2. Investigations into Host Specificity of *N. meningitidis*

Figure 2.1 Size-exclusion chromatography and SDS-PAGE analysis of the purified m-fH ₆₇	36
Figure 2.2 – SPR results of m-fH ₆₇ with fHbp.....	39
Figure 2.3 – m-fH ₆₇ crystal grown in Proplex condition A9.....	42
Figure 2.4 Systematic absences of the h00, 0k0 and 00l reflections suggest the crystal belongs to spacegroup P2 ₁ 2 ₁ 2 ₁	44
Figure 2.5 – Sequence alignment of h-fH ₆₇ and m-fH ₆₇	46
Figure 2.6 Structure of m-fH ₆₇	49
Figure 2.7 Crystal packing of m-fH ₆₇ in the P2 ₁ 2 ₁ 2 ₁ crystal form.....	50
Figure 2.8 Superposition of the m-fH ₆₇ and h-fH ₆₇ structures.....	51
Figure 2.9 – Reorientation of SCRs 6 and 7 into a fHbp “binding” conformation reveals residue clashes across the binding interface	52
Figure 2.10 – m-fH ₆₇ overlaid onto h-fH ₆₇ in complex with fHbp.....	53
Figure 2.11 Consurf analysis of m-fH ₆₇	55
Figure 2.12 Surface charge analysis of m-fH ₆₇ compared with h-fH ₆₇	56

3. Structural Analysis of Meningococcal PilE

Figure 3.1 Alignment of PilE sequences of <i>N. gonorrhoeae</i> MS11 and <i>N. meningitidis</i> MC58	61
Figure 3.2 Structure of gonococcal PilE (PDB ID: 1AY2) with alternative start residues highlighted	63
Figure 3.3 Alignment of full-length PilE sequences of strains to be investigated.....	64
Figure 3.4 pETM vectors used for expression of PilE.....	65
Figure 3.5 pilE genes successfully amplified as shown by analysis on 1% agarose TAE gel.....	67

Figure 3.6 Digestion of the pile inserts and pETM vectors with NcoI and XhoI.....	67
Figure 3.7 Example SDS-PAGE analysis of PilE expression trials.....	70
Figure 3.8 Size-exclusion profiles and SDS-PAGE analysis of PilE_1.MC58.....	74
Figure 3.9 Biophysical analysis of the pETM-14 PilE_1.MC58 PilE ΔN24 protein sample.....	75
Figure 3.10 Crystals grown of MC58 ST-32 PilE ΔN24.....	77
Figure 3.11 The structure of PilE_1.MC58 ΔN24 solved to 1.4Å resolution.....	81
Figure 3.12 Topology diagram of PilE_1.MC58 ΔN24 structure.....	83
Figure 3.13 Homology modelling of class II pilins.....	85
Figure 3.14 Consurf analysis of PilE using class I and class II PilE sequences.....	86
Figure 3.15 Consurf analysis of PilE using class II PilE sequences.....	88
Figure 3.16 Crystal packing arrangement of PilE_1.MC58 ΔN24.....	89
Figure 3.17 P6 ₃ Crystal packing of TcpA of <i>V. cholerae</i>	90
Figure 3.18 Superposition of PilE of MC58 ST-32 PilE ΔN24 with the <i>N. gonorrhoeae</i> MS11 PilE....	92
Figure 3.19 Superposition of one molecule of PilE_1.MC58 ΔN24 onto one gonococcal pilin subunit in the EM reconstruction of the pilus fibre (PDB ID: 2HIL).....	93
Figure 3.20 Surface charge analysis of pilus models and the residues responsible.....	94

4. Functional Analysis of Meningococcal PilE

Figure 4.1 Size-exclusion and SDS-PAGE analysis of CD4634 purification.....	105
Figure 4.2 Investigating the putative interaction between PilE and CD46 by ¹ H ¹⁵ N-HSQC.....	108
Figure 4.3 Investigating the putative interaction between PilE and CD46 by SPR.....	109
Figure 4.4 Fluorescent microscopy analysis of epithelial cells with PilE.....	116
Figure 4.5 LDH activity of supernatants from epithelial cells incubated with PilE.....	118
Figure 4.6 TEER analysis of epithelial cells incubated with PilE.....	120
Figure 4.7 ELISA analysis of homologous recognition of PilE sera.....	125
Figure 4.8 Western blot analysis of PilE sera cross-reactivity.....	126
Figure 4.9 SBA conducted with anti-class I PilE and anti-class II PilE sera.....	127
Figure 4.10 Adhesion blocking assay conducted with anti-class I PilE and anti-class II PilE sera....	128

5. Expression and Purification of PilC

Figure 5.1 Alignment of the C-terminal protein sequences <i>N. meningitidis</i> FAM18 PilC1 and PilC2	135
Figure 5.2 Alignment of the C-terminal protein sequences of <i>Pseudomonas</i> PAK PilY1 with <i>N. meningitidis</i> FAM18 PilC1 and PilC2.....	136

Figure 5.3 Analysis of the regions of common sequence on the PilY1 β -propeller structure.....	137
Figure 5.4 Digestion of pETM14-PilC-C vectors and pET32b.....	140
Figure 5.5 pET32b-PilC-C expression construct.....	141
Figure 5.6 Gels showing the relative expression levels of thioredoxin-tagged PilC ₁₄₈₉₋₁₀₄₈ and PilC2-C ₅₀₅₋₁₀₃₈ in Origami B cells in a range of expression conditions.....	142
Figure 5.7 Size-exclusion chromatography traces of undigested and digested PilC ₁₄₈₉₋₁₀₄₈	145
Figure 5.8 Gels showing the purification of pET32b- PilC ₁₄₈₉₋₁₀₄₈ with the three different buffer systems	148
Figure 5.9 Purification of pure PilC ₁₄₈₉₋₁₀₄₈	150
Figure 5.10 Purification of pure monomeric PilC ₁₄₈₉₋₁₀₄₈	151

6. Functional Characterisation of the Novel Cell Division Protein ZapE

Figure 6.1 Alignment of the full-length ZapE protein sequences of <i>E. coli</i> and <i>N. meningitidis</i>	158
Figure 6.2 Codon-optimised ZapE genes were ordered from GeneArt.....	159
Figure 6.3 Digestion of GeneArt vectors and pET28b.....	159
Figure 6.4 Gels showing relative expression levels of <i>N. meningitidis</i> and <i>E. coli</i> ZapE.....	162
Figure 6.5 Size-exclusion chromatograms and SDS-PAGE analysis of the expression and purification of ZapE from <i>E. coli</i> and <i>N. meningitidis</i>	164
Figure 6.6 Mass spectrometry results of ZapE protein samples.....	165
Figure 6.7 DSC of ZapE protein samples.....	166
Figure 6.8 CD analysis of the ZapE protein samples.....	167
Figure 6.9 Gels of the final ZapE protein, wildtype and Walker box A mutant from <i>E. coli</i> and <i>N. meningitidis</i>	168
Figure 6.10 Size-exclusion profiles and SDS-PAGE analysis of the expression and purification of selenomethionine derivatised ZapE from <i>N. meningitidis</i>	170
Figure 6.11 Mass spectrometry results of the selenomethionine derivatised ZapE from <i>N. meningitidis</i>	171
Figure 6.12 MALS analysis of <i>N. meningitidis</i> ZapE.....	172
Figure 6.13 Sequence alignment of the ZapE sequences from <i>N. meningitidis</i> and <i>E. coli</i> with FFAS regions of similarity highlighted	177
Figure 6.14 Limited proteolysis of <i>N. meningitidis</i> ZapE.....	178
Figure 6.15 4-20% SDS-PAGE of limited proteolysis products for N-terminal sequence analysis..	179
Figure 6.16 Mass spectrometry analysis of the proteolysed fragments.....	180
Figure 6.17 Comparable proteolysis site on FFAS top hits.....	182

Figure 6.18 Size-exclusion profiles of digested and undigested ZapE.....	183
Figure 6.19 Predicted domains of ZapE.....	184
Figure 6.20 A_{650} at a range of ATP concentrations with wild type and mutant <i>E. coli</i> and <i>N. meningitidis</i> ZapE, compared to background ATP hydrolysis.....	186
Figure 6.21 Standard curve for free phosphate and A_{650}	187
Figure 6.22 Michaelis-Menten fits of the specific change in A_{650} against ATP concentration.....	188
Figure 6.23 Size-exclusion profile and SDS-PAGE analysis of the expression and purification of FtsZ from <i>E. coli</i>	192
Figure 6.24 SPR trace of FtsZ “binding” ZapE.....	193
Figure 6.25 Initial EMSA of <i>E. coli</i> ZapE with pET28b <i>E. coli</i> ZapE plasmid.....	195
Figure 6.26 EMSAs conducted to investigated the DNA binding properties of <i>E. coli</i> ZapE.....	197
Figure 6.27 Alignment of the vector origins of pCDF, pCOLA, pACYC and pET.....	200
Figure 6.28 EMSA conducted to investigated the DNA binding properties of <i>N. meningitidis</i> ZapE	201

List of Tables

1. Introduction

Table 1.1 Evidence for and against the role of CD46 in Neisserial adhesion to host cells.....	27
---	----

2. Investigations into Host Specificity of *N. meningitidis*

Table 2.1 Expressed sequence of pET15b-m-fH ₆₇ construct.....	34
Table 2.2 Crystallisation screens (all Molecular Dimensions) trialled in the crystallisation of m-fH ₆₇ and their relative properties.....	41
Table 2.3 Data quality statistics from processing.....	43
Table 2.4 - Data collection and structure solution statistics.....	44
Table 2.5 Calculation of the number of molecules per asymmetric unit.....	45
Table 2.6 Molecular replacement solutions for m-fH ₆₇ from PHASER.....	47
Table 2.7 – Refinement statistics.....	48

3. Structural Analysis of Meningococcal PilE

Table 3.1 Primers used to amplify the full-length pilE sequences from meningococcal genomic DNA using PCR.....	66
Table 3.2 Primers used to create N-terminal deletion constructs of the pETM-pilE vectors.....	68
Table 3.3 Primers to introduce the E130G and V55I mutations into PilE_4 constructs to give the sequence for PilE_3	68
Table 3.4 Conditions screened for optimal soluble expression in <i>E. coli</i> expression strains	69
Table 3.5 Expressed sequences of pETM-14 PilE ΔN24 constructs.....	73
Table 3.6 Conditions in which MC58 ST-32 PilE ΔN24 crystals grew.....	77
Table 3.7 - Data collection and structure solution statistics.....	78
Table 3.8 Molecular replacement solution for MC58 ST-32 PilE ΔN24 from PHASER.....	79
Table 3.9 – Refinement statistics.....	80
Table 3.10 Primers used to create N-terminal deletion constructs of the pETM-PilE vectors.....	84

4. Functional Analysis of Meningococcal PilE

Table 4.1 Components of minimal growth medium.....	106
Table 4.2 Cells lines used to study the cytotoxic effects of PilE.....	113

5. Expression and Purification of PilC

Table 5.1 Conditions screened for optimal soluble expression in <i>E. coli</i> expression strains.....	141
--	-----

Table 5.2 Expressed sequence of pET32b- PilC1 ₄₈₉₋₁₀₄₈	144
Table 5.3 Primers for mutagenesis of enterokinase cleavage site to a 3C cleavage site in the pET32b- PilC1 ₄₈₉₋₁₀₄₈ construct.....	146
Table 5.4 Buffers used to improve solubility of PilC1 ₄₈₉₋₁₀₄₈ in solution.....	146
Table 5.5 A summary of the methods trialled for PilC purification and their respective outcomes	152

6. Functional Characterisation of the Novel Cell Division Protein ZapE

Table 6.1 Expressed sequences of ZapE constructs.....	160
Table 6.2 Conditions screened for optimal soluble expression in <i>E. coli</i> expression strains.....	161
Table 6.3 Primers for mutagenesis of Walker Box A lysine to alanine.....	168
Table 6.4 The ATP analogues used, their relative features and a protocol for use.....	174
Table 6.5 All crystallisation trials conducted.....	175
Table 6.6 Proteases used to characterise ZapE structure and generate protein products more amenable to crystallisation	178
Table 6.7 N-terminal sequencing results of the <i>N. meningitidis</i> ZapE proteolysed fragments.....	180
Table 6.8 Sequences of the dominant proteolytic fragments as determined by mass spectrometry and N-terminal sequencing.....	181
Table 6.9 Kinetic parameters of ZapE ATPase activity	189
Table 6.10 Expressed sequences of FtsZ constructs.....	191

7. Appendix

Table 7.1 General PCR protocol.....	207
Table 7.2 Site-directed mutagenesis protocol.....	207
Table 7.3 Transformation protocol.....	208
Table 7.4 SDS-PAGE loading dye and gel components.....	209

1. Introduction

1.1 Bacterial Pathogenicity

A pathogen is a disease causing organism, a bacterium, virus, protozoa, prion or fungus (Greenwood, 2012). This work is solely concerned with bacterial pathogens and more specifically, *Neisseria meningitidis*. The virulence of a bacterial species is the degree of pathogenicity caused by the organism within its specific host which may be quantified by the 50 % lethal dose (LD_{50}), the number of organisms needed to kill half the hosts, or by the 50 % infectious dose (ID_{50}), the number of organisms needed to cause infection in half the hosts (Brooks and Jawetz, 2013). The level of virulence is determined by a number of factors including genetic, biochemical and structural features of both the pathogen and the host which affect the disease progression and outcome. The host-pathogen relationship is extremely dynamic and is characterised by the interplay of the degree of virulence of the pathogen and as well as the relative resistance or susceptibility of the host towards that specific pathogen, which may modulate throughout the progression of infection (Hawkey and Lewis, 2004). When this balance shifts in favour of the pathogen over the host, this will result in the disease state. Strains of pathogens which cause severe disease states, often in combination with epidemics due to high infectivity rates, are commonly referred to as hypervirulent (Schoen et al., 2008).

The underlying mechanism of bacterial pathogenicity can be divided into two stages; invasiveness and toxigenesis (Hawkey and Lewis, 2004). Invasiveness encompasses the colonisation of the host by the pathogen, through adherence to a particular host niche and in many cases expansion through cell division, followed by the production of invasins to aid dissemination and also evasion of the host defence mechanisms. Toxigenesis is predominantly characterised by the production of toxins, two broad categories of which are endotoxins and exotoxins.

Following transmission of the pathogen into the host, colonisation begins (Greenwood, 2012). Adherence within a given niche is fundamental for many pathogens. Most pathogens express adhesins which allow specific binding interactions with host receptors. Receptor binding specificity can confer host species specificity, tissue tropism as well as strain or isolate specificity. Adherence to host surfaces may also be non-specific and rely on avidity effects rather than high affinity receptor-based interactions (Hawkey and Lewis, 2004). Pili and capsules are both important features for pathogens which require adherence to host surfaces, especially mucous membranes. Self-adherence is also important for bacteria which form biofilms, which can help protect from host defences (Brooks and Jawetz, 2013). Cell division by pathogens during and after this point is important for expansion of the pathogen. Invasins can be produced which aid dissemination of the bacteria within the host. The pathogen may also produce toxins such as lipopolysaccharide (LPS) or a protein toxin (Greenwood, 2012). Toxins may elicit their effect on a target cell or interact with cells of the immune system resulting in the release of immunological mediators which can cause pathophysiological effects. Whilst damaging to the host, the effects of toxins can be beneficial to pathogen proliferation and dissemination as well as continuation or progression of the disease state. Following this, the host immune response, both innate and acquired, will be elicited. The final stage of pathogenesis is resolution or progression of the disease, decided largely by whether the virulence and number of the pathogens outweigh the host resistance to the pathogen (Brooks and Jawetz, 2013).

Evasion of host immune responses is one of the most important stages in bacterial pathogenesis and is often ongoing throughout the entire process of infection. Evolution of bacterial species is often rapid due to the comparatively frequent rate of cell division of bacterial cells compared to host cells. Bacteria are also able to alter their genetic material through point mutation, horizontal DNA transfer and phase variation (Lemaire et al., 2012). The quick evolution of bacterial species means that most avenues of host immune evasion have been explored and exploited (Hawkey and Lewis, 2004). Most pathogenic bacterial species combat the host immune system in a

complex and multi-factorial manner (Reddick and Alto, 2014). Resistance to host immune systems can follow a number of strategies. Bacteria can be inherently resistant to host bactericidal responses including complement, one of the first lines of host defence, and phagocytosis. Immunological tolerance, antigenic disguise and host structure mimicry are also common methods of immune evasion. Other pathogens may induce immunosuppression or persist at sites inaccessible to immune responses to avoid any host defence mechanisms. Some pathogens present antigens which induce production of ineffective antibodies or may secrete soluble antigens to adsorb pathogen-specific antibodies. Finally, mechanisms which allow bacteria to rapidly change their genetic material are common in pathogens allowing antigenic variation and swapping, through mutation, horizontal gene transfer and phase variation as mentioned previously (Greenwood, 2012).

1.2 Studying Bacterial Pathogenesis using Biophysical Methods

Research into bacterial pathogenesis centres around identifying and characterising the biochemical moieties involved in virulence, the interactions which are crucial for pathogenesis, the sequence of events in pathogenesis and the mechanisms of action as well as the evolutionary relationships between closely related pathogens or systems (Hawkey and Lewis, 2004).

Biophysical methods allow a thorough exploration of the molecular mechanisms of bacterial pathogenesis and can give results which provide a detailed insight into the role of individual protein molecules. The dramatic improvement in genetic methods in recent years and the growing wealth of information now freely available, with respect to libraries of annotated bacterial genomes as well as associated functional analyses, provides a good platform from which such studies can begin. Compared to cell biology or bacteriology techniques, biophysical methods tend to use purified protein samples, allowing dissection of pathogenesis at the level of individual molecules rather than complex structures, systems or whole cells (Van Holde et al., 2006).

Following protein production and purification of a homogenous sample, protein characterisation can be conducted by an array of techniques. The secondary structural elements of a protein can be evaluated by circular dichroism (CD), which in combination with differential scanning calorimetry (DSC) can also provide a melting temperature for the protein in question, providing further information regarding the fold and stability of the sample. The oligomeric state can be assessed by size-exclusion chromatography either alone or for more precise analysis, in tandem with multi-angle laser-light scattering (MALLS) (Van Holde et al., 2006).

Subsequent identification of binding partners of the protein being studied, allows characterisation of the interaction both in terms of the kinetics and thermodynamics by techniques such as nuclear magnetic resonance (NMR) heteronuclear single quantum coherence (HSQC) spectroscopy, surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC).

Analysis of the binding interface can be conducted through point-mutation analysis and then reanalysis of the interaction parameters by the afore-mentioned methods (Fu, 2004).

Finally, one of the most powerful tools for thorough characterisation of bacterial pathogenesis and validation of results obtained through various biophysical methods is structural analysis of the proteins involved. High resolution protein structure solution by NMR or X-ray crystallography can allow analysis of structure-function relationships involved in bacterial pathogenesis.

Structures of protein complexes can guide future experiments, for example, informed mutagenesis analysis of an interaction interface to assess residue-specific contributions towards the interaction. Comparison of homologous structures can illuminate evolutionary relationships and disparities. Solution of protein structures also allows model building of larger assemblies or complexes, which are difficult to analyse with current high resolution techniques and so can provide further information about the system being studied. In combination with lower resolution structural methods such as cryo-electron microscopy (cryo-EM), small-angle X-ray scattering (SAXS) or atomic force microscopy (AFM), pseudo-atomic resolution models may be built of large macromolecular complexes (Rupp and Kantardjieff, 2010).

A detailed understanding of bacterial pathogenesis is important and relevant to modern day medical research and clinical applications. Firstly, such research is significant to this field from an academic perspective. Building a greater understanding of the evolution of various bacterial species and understanding the links between homologous systems as well as the details of host-pathogen interactions permits a greater level of information to be inferred and extrapolated in other systems (Greenwood, 2012). Secondly, from a practical perspective, research in the field of bacterial pathogenesis allows accurate diagnosis due to a greater understanding of disease progression and mechanisms. There is also the opportunity for the development of better therapeutic agents, either for disease intervention in the form of preventative vaccines or other

prophylactic agents, or for disease treatment through design of specific antimicrobial drugs (Brooks and Jawetz, 2013).

1.3 *Neisseria meningitidis*

Nearly all work completed in this thesis centres around one bacterial pathogen, *Neisseria meningitidis*. This diplococcal Gram-negative bacterial species is the predominant causative agent of a range of diseases encompassed by the term invasive meningococcal disease (Stephens et al., 2007). Meningitis is defined as the infection of the meninges, membranes which envelope the brain and the spinal cord, though this is not always a symptom of infection by *N. meningitidis* (Greenwood, 2012). The most common disease states are acute septicaemia (meningococemia), which may present with or without meningitis, or purulent meningitis, characterised by the pupuric rash. The fulminant form of the disease is characterised by multisystem involvement and a high mortality rate (Cartwright, 1995). *N. meningitidis* is an obligate human pathogen causing over 1000 cases of meningitis in the U.K. every year, 10% of which result in a long-term disability or fatality (Agency for Meningitis Research, 2013). *N. meningitidis* is however predominantly a commensal and a large proportion of the population can carry this bacterium asymptotically, up to 70% in an epidemic (Sivonen et al., 1978, Imrey et al., 1995).

Meningococcal strains are historically categorised on the basis of the antigenic variation of the polysaccharides in their capsule, which are the key components. By this method, 13 different serogroups have been identified, five of which are the most common disease causing subsets of strains: A, B, C, Y and W-135, whilst D, H, I, K, L, 29E, X and Y are rarely associated with disease (Cartwright, 1995, Harrison et al., 2013). Meningococci can be further divided into serotypes as well as serosubtypes and defined by multilocus sequence typing (MLST) analysis of seven house-keeping genes (Brehony et al., 2007, Maiden et al., 1998). More recently, whole genome sequencing of isolates is allowing more thorough and precise categorisation of different *N. meningitidis* strains (Bentley et al., 2007). Strains of *N. meningitidis* are highly diverse due to mutation, phase variation and horizontal gene transfer mechanisms which constantly alter the compact yet dynamic genome of this species (Katz et al., 2011).

The precise pathogenic factors which make particular hypervirulent strains dominate disease outbreaks in certain regions are unclear. Some host factors which increase meningococcal disease susceptibility have been identified. Complement deficiencies are also commonly linked to Neisserial infection (Mathew and Overturf, 2006), (Figueroa et al., 1993). Two thirds of all cases occur within the first five years of life, with a peak at one year old (Zahlanie et al., 2014). Adult subjects who are long term asymptomatic carriers of the bacterium are thought to have serum bactericidal activity against the pathogen, explaining why this disease is confined to younger infants (Harrison, 2010). Active or passive smoking is also an associated risk factor with meningococcal disease.

Also from the *Neisseriaceae* family is *Neisseria gonorrhoeae* which causes the sexually transmitted disease gonorrhoea. The gonococcus and meningococcus are similar pathogens, not only belonging to the same genus but being Gram-negative oval-shaped diplococci with pili (Brooks and Jawetz, 2013). The genomes of both species are highly homologous at the protein level though each has several hundred unique genes (Stabler et al., 2005). Both are human-specific pathogens, but they reside in different niches, meningococcus in the posterior nasopharynx and gonococcus in the urogenital tract. Many studies infer information about one species from the other due to the high degree of similarity of these two pathogens although this is not always reliable given the fundamental differences in host niche and disease progression between these two species. There are many other members of the *Neisseria* genus and some are common commensals of the respiratory tract. However most are of low pathogenicity, causing disease almost solely in immunocompromised hosts, so rarely are comparisons extended to species such as *Neisseria lactamica* or *Neisseria polysacchareae* (Brooks and Jawetz, 2013).

Disease progression in cases of *N. meningitidis* is rapid and symptoms are not always easy to recognise or differentiate from less dangerous infections, making diagnosis of this disease extremely difficult. At the time of starting my DPhil research, there were no vaccines which were

effective against all strains of *N. meningitidis*, in particular serogroup B strains, which have an exact mimic of a modification on a neural cell adhesion molecule in their capsule (Finne et al., 1983). This renders the capsule a poor antigen in humans, unlike other serogroups, and inoculation with such an antigen raises concerns over generating a potential autoimmune response. Therefore, a thorough characterisation of disease progression is important to illuminate aspects of the bacterial pathogenesis which may be exploited to inform clearer diagnosis and the development of a broad spectrum vaccine. A year ago, a vaccine called Bexsero®, which provides protection against serogroup B strains of *N. meningitidis*, was brought to market by the pharmaceutical company Novartis International AG (Giuliani et al., 2006). The vaccine contains four outer membrane components; fHbp (factor H binding protein), PorA (an outer membrane porin), NadA (Neisserial adhesin) and NHBA (Neisserial heparin binding antigen). Research for a vaccine providing complete protection against all strains of *N. meningitidis*, which is effective and long-lasting, continues both in academic and pharmaceutical research fields.

This thesis describes three main projects which look at different aspects of the pathogenesis of *N. meningitidis*. The background for each project as well as the premise and aims for each project is described in **1.4**, **1.5** and **1.6**.

1.4 Immune Evasion and Host Specificity of *N. meningitidis*

N. meningitidis is specifically adapted for survival within the human host. One of the key determinants of host specificity are the adaptations of this pathogen to combat the human immune system at many stages of pathogenesis during invasive disease, as well as commensal residence in the nasopharynx (Jarvis, 1995, Yazdankhah and Caugant, 2004). The human immune system is extremely complex and confronts infection by the meningococcus with an array of different strategies to eliminate this bacterium. However, the rapidly evolving *N. meningitidis* has evolved mechanisms to overcome attempts to eliminate the bacteria and evade the human immune response (Lo et al., 2009).

The meningococcus is able to combat host produced antimicrobial peptides and proteins, part of the evolutionary ancient innate immune system (Kolls et al., 2008). Phosphoethanolamine adducts to lipid A by *N. meningitidis*, present in the outer membrane, increases meningococcal resistance against the human microbicidal peptide LL-37 (Tzeng et al., 2005). The operon *mtrCDE* encodes an efflux pump which has been shown to be important for antimicrobial peptide resistance (Tzeng et al., 2005). Human serum lactoferrin and transferrin, which normally scavenge and chelate surplus free iron, an important growth factor for most pathogens, are specifically targeted by *N. meningitidis* (Schryvers and Morris, 1988) which can remove iron directly from the host protein and transport it inside the bacterial cell through a TonB-dependent transport system (Perkins-Balding et al., 2004, Noinaj et al., 2012).

Neutrophils and macrophages are both highly adapted professional phagocytes with roles in defence as well as homeostasis and immune regulation (Kantari et al., 2008). Both produce reactive oxygen species (ROS) and reactive nitrogen species (RNS) (Fang, 2004). The oxidative stress caused by ROS and RNS has a toxic effect on bacterial proteins, lipids and genetic material. To directly deal with such species, *N. meningitidis* expresses both catalase and superoxide dismutase to neutralise ROS and RNS from phagocytes (Archibald and Duong, 1986), (Seib et al.,

2004). This pathogen is able to acquire L-glutamate from the local environment which is then converted to glutathione to help maintain the redox potential when dealing with ROS and RNS (Pomposiello and Demple, 2002). Additionally, denitrification enzymes of *N. meningitidis* allow RNS to be converted to nitrous oxide which can then be used for growth (Anjum et al., 2002) whilst base excision repair mechanisms allow reparation of DNA after ROS and RNS treatment (Davidsen et al., 2007).

The capsule has been shown to be one of the most important virulence factors of this species, as acapsulate bacteria have significantly reduced serum resistance compared to encapsulated forms (Hammerschmidt et al., 1994, Kahler et al., 1998). Certain aspects of the human immune system are specifically down-regulated by a number of highly adapted means by the meningococcal capsule, including complement, which is discussed in more detail in the next section. The capsule also acts against other aspects of the human immune system, conferring resistance to antimicrobial peptides and defensins (Spinosa et al., 2007), reducing adherence to dendritic cells and masking antigenic surface expressed proteins to reduce phagocytosis (Unkmeir et al., 2002a).

1.4.1 Evasion of the Complement System

A key branch of human innate immunity is the complement system which is able to rapidly eliminate invading pathogens either independently or in concert with antibody responses and phagocytosis (Walport, 2001a, Walport, 2001b). This proteolytic cascade of approximately 40 plasma and cell surface-bound proteins can be initiated by one of three pathways; the classical pathway (initiated via antibody-antigen interactions), the lectin pathway (initiated by mannose-binding lectin (MBL) binding to microbial carbohydrates) or the alternative pathway (which is constitutively active and constitutes the amplification arm of the system). All pathways converge on one of two C3 convertase enzymes which cleave C3 to C3b. C3b has numerous functions including opsonisation, accentuating inflammation, chemotaxis and lysis via membrane attack complex (MAC) formation, see **Figure 1.1**.

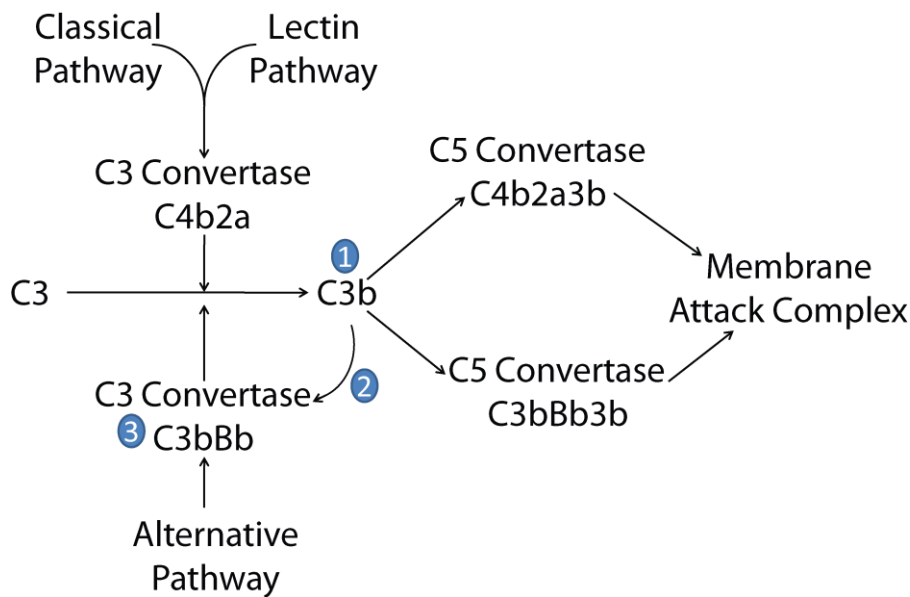


Figure 1.1 Outline of the complement cascade

The three points at which fH regulates the complement system are highlighted on the cascade: 1 - cofactor activity for fI mediated cleavage of C3b, 2 - compete with fB for binding of C3b, 3 - accelerate dissociation of C3bBb C3 convertase

Given the prolific nature of complement, regulation is essential throughout the system to protect host cells and tissues from non-specific complement attack. The alternative pathway is constitutively active, so regulation is particularly essential to prevent inappropriate complement activity via this route. Factor H is one such regulator of the alternative pathway, functioning as a cofactor for the factor I (fI) mediated cleavage of C3b, accelerating dissociation of the C3 convertase C3bBb and also competing with factor B (fB) for binding of C3b, reducing the amount of C3bBb formed (Pangburn et al., 2000). fH is functional in these respects in both the fluid plasma phase and when bound to cell surfaces through specific interactions with host cell surface-exposed polyanions and glycosaminoglycans (GAGs) targeting complement down-regulation directly to host cell and tissue surfaces (de Cordoba et al., 2004). Sialic acid in the meningococcal capsule, present as a molecular mimic, reduces antibody recognition of the capsule and hence complement activation against the bacterium via the classical pathway (Walport, 2001a).

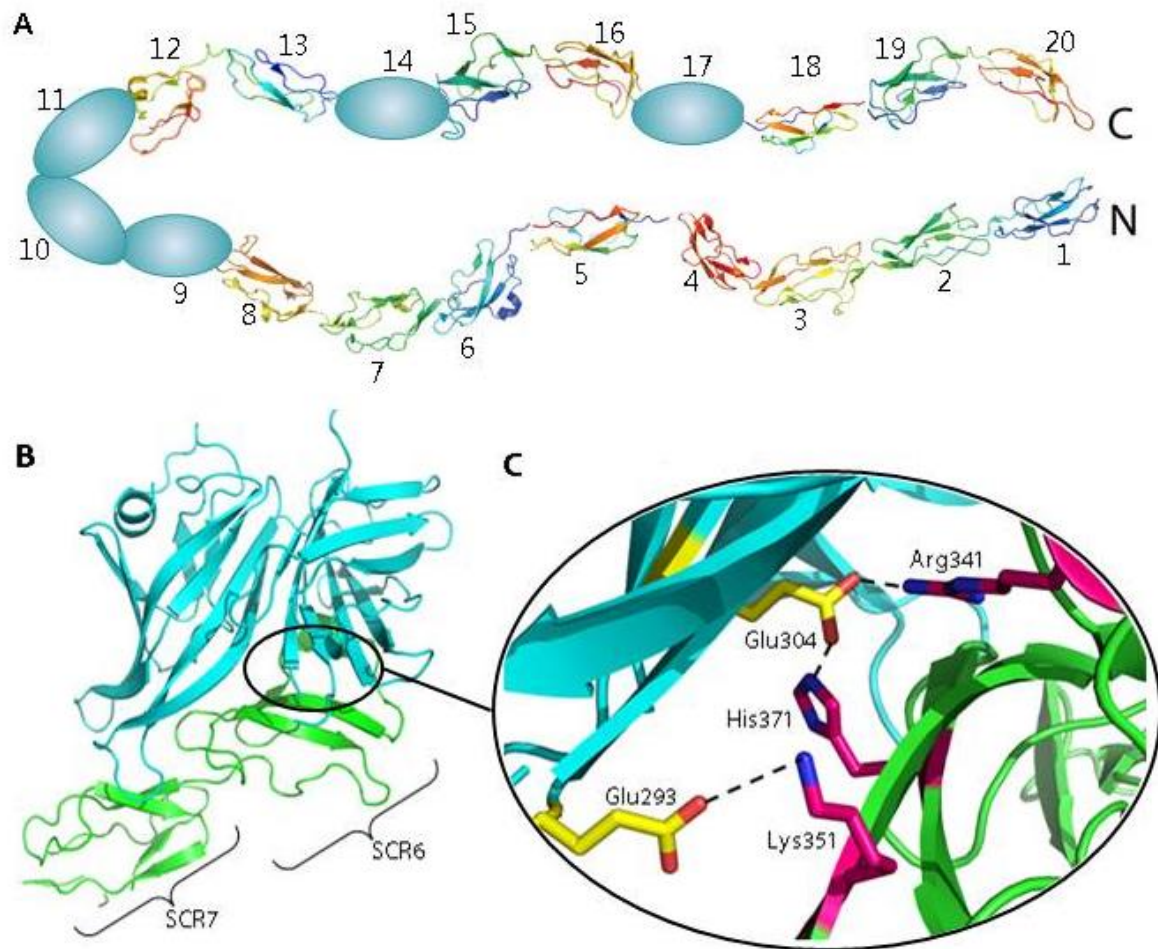


Figure 1.2 The interaction between fH and fHbp

Panel A shows a schematic representation of fH with domains for which structures have been solved shown in cartoon format PDB IDs: 1HCC, 1HFH, 2BZM, 2JGX, 2IC4, 2QFG, 2QFH and 2KMS (Barlow et al., 1993, Herbert et al., 2006, Jokiranta et al., 2006, Okemefuna et al., 2008, Prosser et al., 2007a, Schneider et al., 2009, Morgan et al., 2012, Makou et al., 2012, Barlow et al., 1991, Herbert et al., 2007, Fernando et al., 2007, Schmidt et al., 2010). Panel B shows the complex of fH-fHbp PDB ID: 2W80, both proteins shown in cartoon format, fH in green and fHbp in blue. Panel C shows some of the crucial charge based interactions at the fH-fHbp interface at SCR6.

Many disease causing serogroups contain sialic acid within their capsule, a molecule which within the human circulatory system is found on the surface of endothelial cells to recruit fH, a negative regulator of the alternative pathway. The presence of sialic acid in the meningococcal capsule has been shown to down regulate the alternative pathway of the complement system (Jarvis and Vedros, 1987). The capsule is able to physically hinder MAC formation in the outer membrane of

the *N. meningitidis* (Geoffroy et al., 2003) as well as reducing binding of MBL to avoid complement activation by this route (Jack et al., 2001).

N. meningitidis is able to bind fH through a specific high-affinity interaction with fH binding protein (fHbp), see **Figure 1.2**. fHbp, a 27kDa surface-exposed lipoprotein also called GNA1870 or LP 2086, is expressed by nearly all strains of *N. meningitidis* (Seib et al., 2009). This key virulence factor recruits fH and thus enhances pathogen serum resistance (Schneider et al., 2006, Madico et al., 2006). The structure of this complex has been solved by X-ray crystallography to 2.35 Å resolution and the interaction analysed by SPR, giving a calculated dissociation constant (K_D) of ~5nM (Schneider et al., 2009). fHbp is a highly variable protein which can be classified on the basis of variable and non-variable modules into three groups (Beemink and Granoff, 2009). Variation of the fHbp gene is much higher within serogroup B compared with serogroups A, C, Y or W-135 (Mothibeli et al.), (Brehony et al., 2009). Irrespective of sequence variations, all meningococcal fHbps consist of two independent beta barrels, connected to the outer membrane via a 20 amino acid anchor, allowing full exposure of this protein to the external environment of the cell and promoting fH recruitment. The short consensus repeats (SCRs) 6 and 7 of fH are sufficient to confer the high affinity interaction with fHbp. The interaction interface, as defined by the solved complex structure, is large and has an area of $2,860 \pm 177$ Å² (Schneider et al., 2009).

1.4.2 Host Specificity of *N. meningitidis*

The host specificity of *N. meningitidis* renders *in vivo* studies of this pathogen difficult. Neisseria infection has been studied in rhesus macaque, though only for native Neisserial species, not *N. meningitidis* (Weyand et al., 2013). To overcome this problem, rodent models subject to infection analysis, are loaded with high bacterial titres and large amounts of free-iron to promote rapid bacterial growth (Johansson et al., 2003). Such methods allow the murine model immune system to be overwhelmed and produce an infection model. However, given that in native situations, such an infection is very unlikely to occur; these models are limited in the pursuit of information

regarding the pathogenesis of *N. meningitidis* in humans. However, the search for therapeutic agents against *N. meningitidis* does require animal models in which to validate the safety of such drugs. Therefore, research into the host specificity of this pathogen which could aid development of better models through genetic manipulation is important to the field.

Previous studies have shown that Neisserial fHbp recruitment is specific for human fH (h-fH) with much weaker binding seen for chimpanzee fH and none at all for lower primates such as rhesus macaque or baboon fH (Granoff et al., 2009b). Administration of h-fH to meningococcal challenge models increases the serum levels of the bacteria in a dose-dependent manner. The hypothesis which may be inferred from such studies is that the fH-fHbp interaction is one of the factors which confer host specificity of *N. meningitidis* to humans. Sequence disparity of fH SCRs 6 and 7 (fH₆₇) across species could mean that the structure of this region of fH has a different surface topology and chemistry and fHbp of *N. meningitidis* has evolved to bind that of the h-fH₆₇. If this is the case, then this would be one of the molecular mechanisms by which *N. meningitidis* is a pathogen specifically adapted only to humans.

1.4.3 THESIS AIMS:

- To analyse the interaction between mouse fH₆₇ (m-fH₆₇) and fHbp by SPR
- To elucidate whether sequence differences in fH₆₇ translate to differences in the structure of this protein in humans compared to mice, a species commonly used for initial vaccine investigations, and analyse what this indicates in the context of the SPR results

1.5 Adhesion by *N. meningitidis*

As a predominantly commensal species, *N. meningitidis* colonises the human nasopharynx through adhesion to the mucosal surfaces in this niche (Greenwood, 2012). The interaction between the meningococcus and the nasopharyngeal epithelium is mediated by a number of surface expressed adhesins, see **Figure 1.3**.

The opacity proteins Opa and Opc are so named due to the opaque phenotype they impart on agar-grown colonies (Aho et al., 1991). Between three and four loci for Opa are present in most meningococcal strains. Opa proteins consist of a conserved eight transmembrane strand β -barrel with four sequence variable surface exposed loops. Phase variation of the different alleles is thought to be important in immune evasion, though certain versions of Opa dominate with respect to adhesion and virulence phenotypes (Stern and Meyer, 1987, Malorny et al., 1998). A single locus for Opc exists in *N. meningitidis* and the structure has been solved of the encoded protein, revealing an eight-stranded transmembrane β -barrel but with five surface exposed loops (Prince et al., 2002). Opa proteins have been shown to interact with numerous members of the carcinoembryonic antigen-related cell adhesion molecule family (CEACAMs) as well as heparin sulphate proteoglycans (HSPGs) (Virji et al., 1996a, Hammarstrom, 1999, de Jonge et al., 2003). Opc mediates adhesion through a trimolecular complex involving either serum vitronectin or fibronectin and their corresponding integrin receptors in addition to HSPGs (Unkmeir et al., 2002b, de Vries et al., 1998).

Neisserial adhesin A (NadA), an oligomeric coiled coil is a phase variable minor adhesin (Capecci et al., 2005). Neisserial hia/hsf homologue A (NhhA) (Sjolinder et al., 2008), adhesion penetration protein (App) (Serruto et al., 2003), meningococcal serine protease A (MspA) (Turner et al., 2006) and porin B (PorB) (Zeth et al., 2013) are all also thought to have a role in meningococcal adhesion. However, all of these minor adhesins are largely uncharacterised and many of their receptors have not been identified to date.

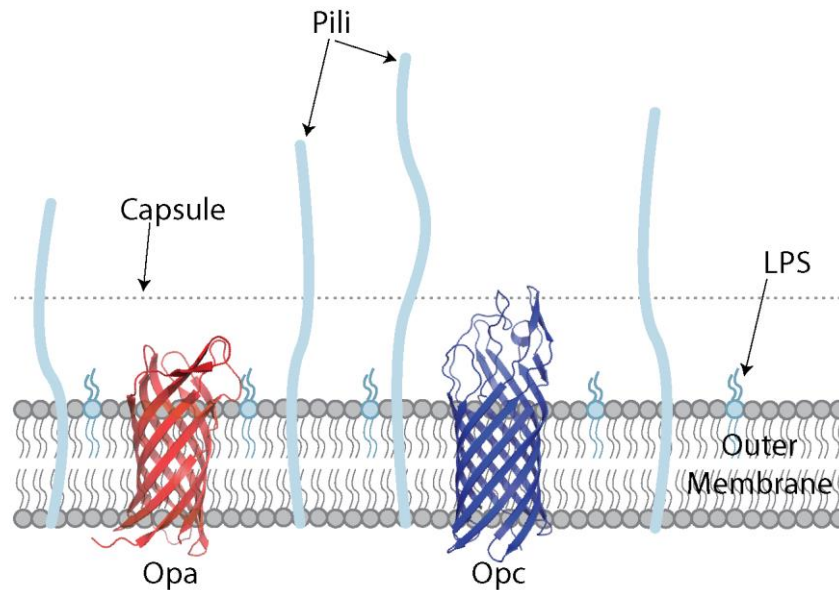


Figure 1.3 Components of the outer membrane of *N. meningitidis* involved in adhesion

Many adhesins reside in the outer membrane of *N. meningitidis*. The schematic shows the extended type IV pili traversing the membrane and capsule, whilst Opa and Opc (PDB ID: 1K24) reside in the outer membrane. The structural model shown for Opa is that of neisserial surface protein A (NspA), which like Opa is an eight-stranded β -barrel molecule (PDB ID: 1P4T).

1.5.1 An Introduction to Type IV Pili

The predominant adhesin of *N. meningitidis* is the type IV pilus (T4P). T4P are macromolecular complexes common to many species of bacteria, including *Pseudomonas aeruginosa* and *Vibrio cholerae*, and also have roles in twitching motility, microcolony and biofilm formation, DNA uptake by natural transformation, cell-signalling and phage attachment in addition to their roles in host adhesion. T4P can be rapidly assembled and disassembled using powerful molecular motors which hydrolyse ATP. Nearly all T4P functions can be attributed to either pilus extension or retraction or pilus fibre surface mediated interactions (Craig et al., 2004). Given the range of functions, T4P are crucial virulence factors and non-piliated strains have markedly reduced pathogenicity.

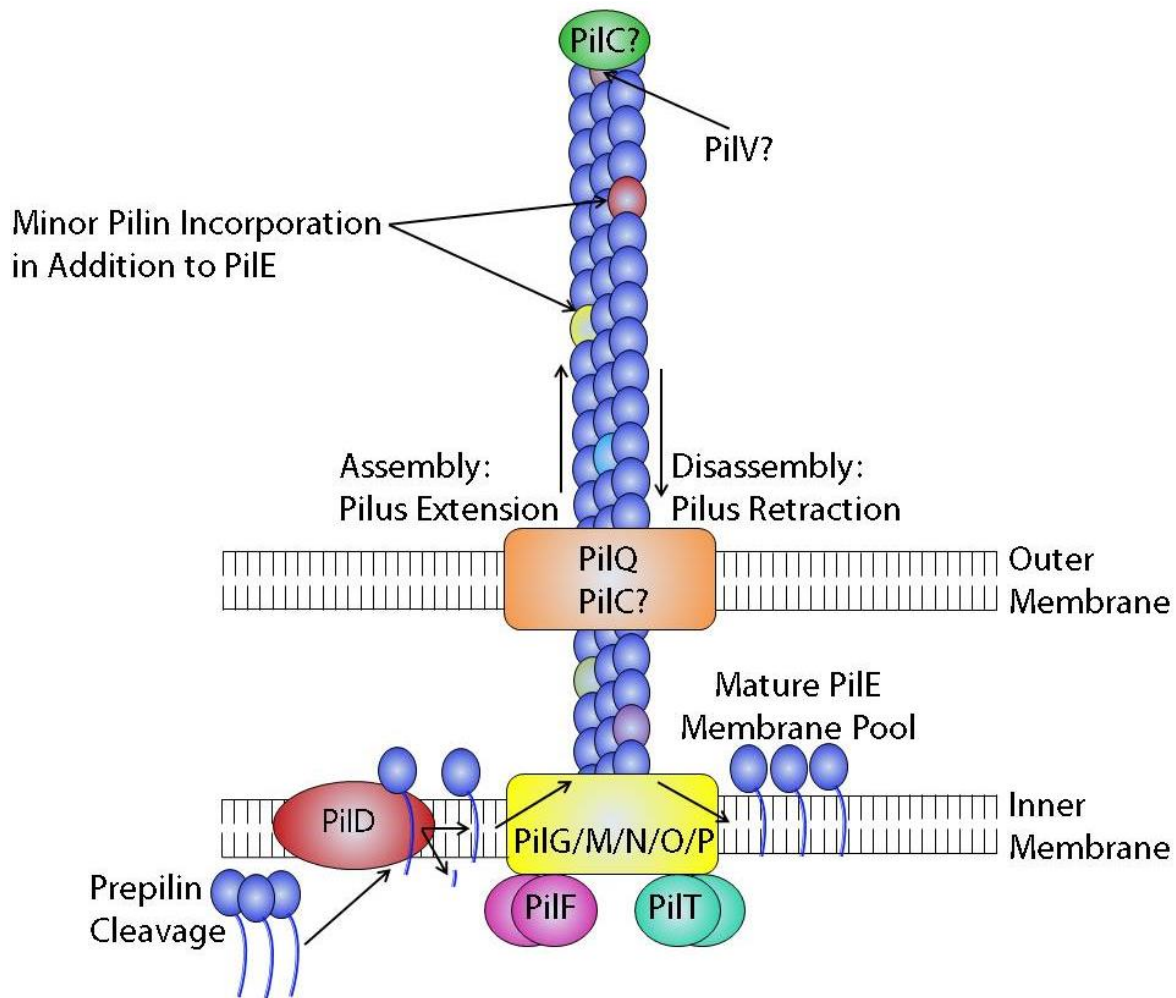


Figure 1.4 Model of the T4P with current understanding of assembly

This model was created using the current understanding of pilus components as described in Georgiadou et al (2012) (Georgiadou et al., 2012)

T4P are strong but flexible filaments approximately 6nm in diameter but up to micrometres in length and are able to withstand forces of up to 100pN (Maier et al., 2002). The overall assembly of T4P from *N. meningitidis* involves approximately 15 proteins in pilus biogenesis and 23 overall (Bardy et al., 2003, Georgiadou et al., 2012). The principal components are conserved across the archaeal flagellar and type II secretion systems. In the case of the T4P of *N. meningitidis*, these include: the major pilin subunit, PilE; an inner membrane prepilin peptidase which cleaves the signal sequence of prepilin to form mature pilin, PilD; the assembly ATPase, PilF; the retraction ATPase, PilT; the inner membrane protein which recruits the cytosolic ATPases, PilG; an outer membrane secretin, PilQ. Also within the pilus fibre are the less abundant minor pilins PilX, PilV

and ComP. PilM, PilN, PilO and PilP are implicated in the early stages of pilus synthesis whilst PilC, PilH, PilI, PilJ, PilK and PilW are necessary for pilus maturation and biogenesis though these roles are not fully characterised. A model of the current understanding of pilus components in the assembly can be seen in the **Figure 1.4**.

The pilus fibre is made up of thousands of copies of the major pilin PilE. Although at the time of starting this project there was no structure of meningococcal PilE, there were numerous PilE structures available for *N. gonorrhoeae* which is highly similar in sequence, sharing 86% sequence identity, PDB IDs: 1AY2 (2.6 Å resolution), 2PIL (2.6 Å resolution but revealing phosphoserine at residue 68) and 2HI2 (2.4 Å resolution) (Parge et al., 1995, Li et al., 2012, Forest et al., 1999). Like other type IV pilin structures, this broadly represents a ladle with a globular domain and an extended α -helix protruding from this structure, see **Figure 1.5**. Gonococcal PilE has a highly conserved N-terminal α -helix which is very hydrophobic, the first 28 amino acids of which protrude from the C-terminal globular head whilst residues 29-52, which are more amphipathic in nature, act as a scaffold for the rest of the structure. This helix is shown to be S-shaped in the full-length structure due to kinks at Pro22 and Gly42. The lack of rigidity in this region is thought to contribute towards fibre flexibility in the assembled pilus. Next is an extended $\alpha\beta$ loop which, in the structures solved of gonococcal derived PilE, has a sugar moiety bound to it. This leads into a four-stranded antiparallel β -sheet with extended loops. After this is a disulphide bound region known as the D-region, which contains two more β -strands forming a β -hairpin, followed by a short unstructured strand which leads to the C-terminus.



Figure 1.5 The structure of PilE from *N. gonorrhoeae* PDB ID:2PIL

The gonococcal PilE structure is shown in cartoon format and is coloured to highlight the different structural regions: the hydrophobic N-terminal α -helix (**red**), the C-terminal globular head composed mainly of a 4-strand (**blue**) and a 2-strand (**orange**) β -sheet, the hypervariable D-region (**cyan** and **orange**) at extreme C-terminus flanked by two cysteine residues which form a disulphide bond (**yellow**). The remainder is largely unstructured composed of just a few turns of α -helix and extended loops (**green**). Post-translational modifications (**magenta**) present are the methylated phenylalanine at residue 1, the O-linked 2,4-diacetamido 2,4,6-trideoxy α -D-hexose (DATDH) sugar moiety at residue 63 and the phosphorylated serine at residue 68.

Once the PilE protein has been synthesised, it is targeted to the inner membrane due to the N-terminal class III signal sequence (Szabo et al., 2007). The SEC translocase system inserts PilE into the inner membrane and the N-terminal α -helix becomes the transmembrane domain. The PilD prepilin peptidase then cleaves the signal sequence and methylates the new N-terminal phenylalanine, a conserved post-translational modification across all Neisserial PilE proteins (Strom et al., 1993). All pilin sequences throughout this work nominate the N-terminal phenylalanine residue as residue 1 in the sequence numbering. A pool of mature pilin resides in the membrane before assembly into the pilus fibre though the mechanism that is largely uncharacterised in *N. meningitidis* although a model has been proposed for *Thermus thermophilus*

(Karuppiyah et al., 2013). Pooling PilE in the membrane is postulated to allow rapid fibre assembly and recycling of these protein subunits upon successive rounds of retraction and extension (Burrows, 2005).

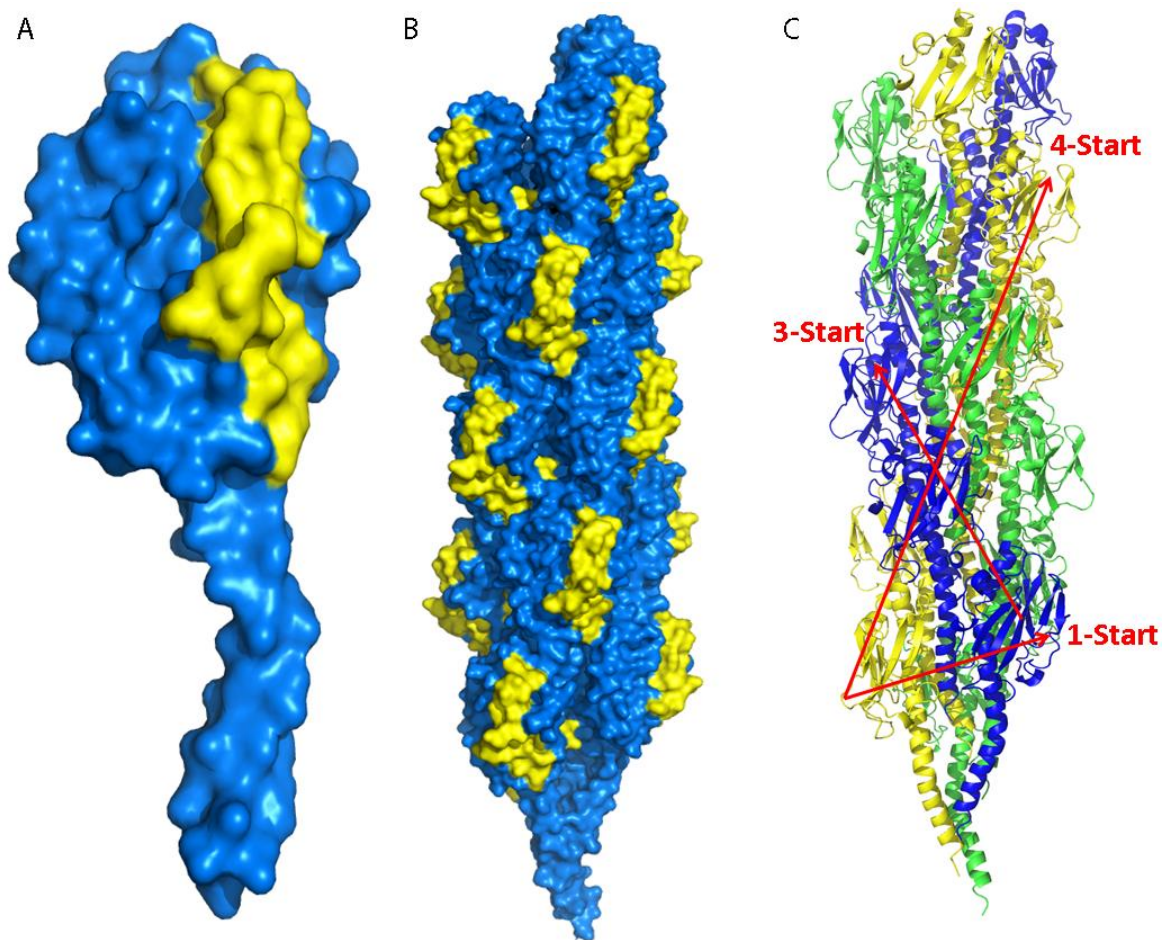


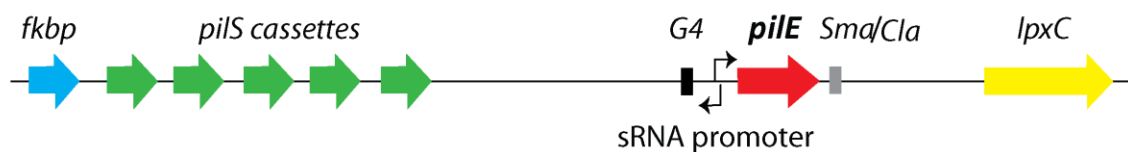
Figure 1.6 Pseudoatomic resolution pilus model

Panel A shows a surface representation of the *N. gonorrhoeae* PilE with the D-region highlighted in yellow. Panel B shows the pseudoatomic resolution model generated by fitting the high resolution crystal structure of PilE into the EM electron density map. Panel C shows the three different helical symmetry operators present in the modelled pilus assembly (Craig et al., 2006b).

Cryo-EM analysis of gonococcal pili produced a 12.5Å resolution model of the pilus fibre, into which gonococcal PilE was modelled to generate a pseudo-atomic resolution model for the pilus (Craig et al., 2006a), see **Figure 1.6**. PilE proteins are arranged in a helical array to form the pilus fibre. Adjacent pilin monomers are related by a rise of 10.5Å and a rotation of 100.8° along a

right-handed 1-start helix, though 3- and 4-start helical symmetry is present in the arrangement. The C-terminal head groups of the structure pack outwards into the external environment whilst the N-terminal hydrophobic α -helices form a coiled coil bundle in the centre. The D-region is entirely surface-exposed in this assembly. Such an arrangement rationalises the high tensile strength of T4P, due to the strength of the hydrophobic interactions in the middle of the assembly.

***pilE* region *N. meningitidis* 8013 - Class I**



***pilE* region *N. meningitidis* FAM18 - Class II**

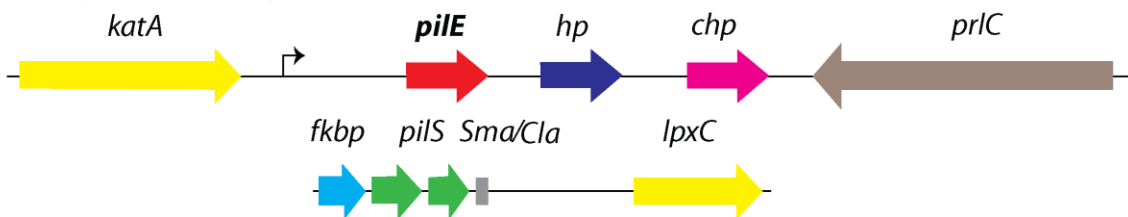


Figure 1.7 Class I and class II *pilE* loci structure

pilE loci in *N. meningitidis* strains can be categorised as either class I or class II (Wormann et al., 2014), see **Figure 1.7**. PilE class is normally determined by the cross-reactivity of PilE to the monoclonal antibody SM1, which reacts to an epitope found on class I PilE protein but not class II (Virji and Heckels, 1983). *N. meningitidis* strains with class II *pilE* loci maintain the same sequence throughout hundreds of generations and the same *pilE* allele can be found in many different strains. For example, isolates derived from a case of *N. meningitidis* infection in the U.S.A in 1964 were shown to have the same *pilE* sequence as a case 45 years later in the U.K. (Wormann et al., 2014), illustrating the conservation of *pilE* sequences in meningococcal strains with class II *pilE* loci. Strains with class I loci however, have varied *pilE* sequences, both between and within isolates (Cehovin et al., 2010). Within the class I loci are multiple cassettes encoding alternative

pilE sequences, called *pilS* cassettes. Transcription of the G4-associated sRNA, forms a guanine quartet allowing proteins involved in recombination to create a new variant *pilE* sequence by recombination with the adjacent *pilS* cassettes by a process known as gene conversion.

Recombination events produce varied *pilE* sequences and this intra-strain sequence switching is known as antigenic variation. Sequence variation is generally localised to the $\alpha\beta$ loop and the D-region (Helm and Seifert, 2010, Davies et al., 2014). Both of these regions of the PilE structure are prominently outward facing in the pilus model and therefore sequence alterations in these regions could affect the pilus surface chemistry (Craig et al., 2003a).

PilE structure can also be varied due to post-translational modification of PilE. Similarly to sequence variation, variation of the post-translational modifications occurs not only between strains but within a strain (Aas et al., 2006a). The commonly used laboratory strain of *N. meningitidis* 8013, which expresses class I PilE, was shown to express two different PilE proteoforms in a 4:1 ratio, whilst patient strain 278534, which also expresses class I PilE, expressed four different proteoforms in a 1:1:1:1 ratio (Gault et al., 2014). In addition to the conserved N-terminal methylated phenylalanine, PilE can have the unusual glycosylation modifications of 2,4-diacetamido 2,4,6-trideoxy α -D-hexose (DATDH) or 2-acetamido 4-glyceramido 2,4,6-trideoxy- α -D-hexose (GATDH) which in turn may be modified further with the addition of up to two galactose or glucose subunits and may also be O-acetylated (Stimson et al., 1995, Parge et al., 1995). Numerous phosphate based modifications have also been described which include the modification of residues by the addition of phosphate, phosphoethanolamine, phosphocholine and phosphoglycerol, originally revealed through high resolution X-ray crystal structures, PDB ID: 2PIL (Forest et al., 1999), but more recently with an approach combining mass profiling and tandem mass spectrometry (Gault et al., 2014). Similarly to sequence variation of PilE, all post-translational modifications except the N-terminal phenylalanine methylation are surface exposed and therefore alter the chemistry of the pilus surface (Craig et al., 2003a).

Minor pilins are generally conserved in sequence (Cehovin et al., 2010). Post-translational modifications of these proteins have not been fully characterised. Although dispensable for pilus biogenesis, roles are attributed to PilX, PilV and ComP in T4P pilus functions including pilus-mediated bacterial aggregation, adherence to human cells and competence for natural transformation respectively (Cehovin et al., 2013, Winther-Larsen et al., 2001, Brissac et al., 2012). Controversy surrounds the incorporation of minor pilins into the T4P fibre, regarding the number of copies present in the fibre, their location or grouping and even whether they are present in the fibre at all (Imhaus and Dumenil, 2014, Giltner et al., 2010, Korotkov and Hol, 2008).

Studies investigating hypervariability in PilE sequence and post-translational modification, have given insight into the function of PilE, aside from acting as a molecular scaffold for the minor pilins and PilC. PilE chemistry varies with such sequence and modification changes and this will likely impact the whole pilus and the surface chemistry of T4P fibres (Miller et al., 2014a). All functions associated with interactions involving the pilus fibre, including host cell adhesion as well as pili-pili interactions which are important in biofilm and microcolony formation, are likely to be affected by pilus surface chemistry alterations. Subtle changes in PilE sequence, as seen in subvariants within a strain, can confer large changes in adherence. Such examples are the well characterised SA and SB variants of *N. meningitidis* 8013, SA has a low adherence phenotype and SB has a high adherence phenotype (Nassif et al., 1994). Additionally, the SA subvariant does not form bacterial clumps or microcolonies as well as SB. A total of 19 residues differ between these subvariants, all of which can be found in the $\alpha\beta$ loop and the D-region, attributing pilus surface chemistry changes to be responsible for these changes of T4P phenotype.

The role of the hypervariation of the PilE sequence has often been attributed to immune escape within the human host (Hubert et al., 2012) and hypervariability does modulate a variety of pilus functions (Nassif et al., 1994, Virji et al., 1996b). However, as most studies investigating T4P

function use class I PilE strains of *N. meningitidis*, very little is known about class II PilE, which does not appear to undergo hypervariation, despite the fact class II PilE expressing strains represent many of those implicated in the annual meningitis epidemics in sub-Saharan Africa (Stephens et al., 2007). The difference between class I and class II PilE expressing strains of *N. meningitidis* is in need of further investigation in order to further understand the role of PilE hypervariability and to provide a better understanding of their biochemical and immunological properties and class I and class II pili.

1.5.2 Type IV Pilus Mediated Adhesion

Adhesion is thought to be primarily conferred through PilE and PilC (Scheuerpflug et al., 1999, Rudel et al., 1992). PilC is a poorly studied component of the meningococcal T4P, despite being implicated in pilus biogenesis and adhesion by genetic analysis (Rudel et al., 1995a, Rudel et al., 1995c). This is possibly due to a lack of good biochemical reagents such as antibodies which specifically target PilC as well as problems synthesising PilC recombinantly. Two alleles of *pilC* are present in most strains of *N. meningitidis*, although the role of the two PilC proteins is controversial. PilC1 and PilC2 have been proposed to confer different adhesive properties (Morand et al., 2001) or initiate different cell signalling pathways (Porsch et al., 2013) whilst some studies have reported one of the alleles to be non-functional (Morand et al., 2001). Studies have also shown that the two halves of this protein have different functions with the C-terminal domain being implicated in cell signalling through a Ca^{2+} -based mechanism and the N-terminal domain involved in adhesion (Rudel et al., 1995b, Ryll et al., 1997). The lack of published information regarding PilC in the Neisserial T4P field, means that more distant species must be analysed for comparison. A structure has been solved for the C-terminal domain of PilY1, the *Pseudomonas aeruginosa* homologue of Neisserial PilC, revealing a 7-fold β -propellor structure with calcium binding sites (Johnson et al., 2011, Orans et al., 2010). However, very little is still known about the exact location of PilC within the pilus. Early immuno-gold labelling experiments appear to suggest it is resident within the outer membrane as well as the pilus tip, two extremely

different environments for a protein to be located (Rudel et al., 1995c). If PilC is the T4P tip adhesin, then the cell signalling roles are unlikely to fit in with this model without an additional step linking PilC-dependent adhesion to signal transduction across the outer membrane of the bacteria. Further investigation is necessary to rationalise all of these findings.

The mechanism of PilE and PilC mediated adhesion as well as which receptor or host molecule they interact with on human cells is poorly understood. CD46, the membrane bound complement regulator, was credited with the role of pilus receptor (Kallstrom et al., 1997). However, a lack of reproducibility of the experiments implicating CD46 in *N. meningitidis* adhesion, as well as no published studies conducting protein-protein interaction studies between pili or PilE and CD46 have made this an area of controversy within the field (Johansson et al., 2003, Kallstrom et al., 2001a, Gill and Atkinson, 2004, Gill et al., 2003, Koransky et al., 1975b), see **Table 1.1**. Models have suggested that PilC mediates specific interactions with a host cell receptor and PilE subunits mediate secondary, less specific interactions with the surrounding host cell surface, whilst other models suggest that PilE is the receptor-specific T4P adhesin (Nassif et al., 1997, Rudel et al., 1992).

The minor pilin PilV is often attributed to mediating T4P-associated adhesion but studies have shown that this is due to *N. meningitidis* PilV mutants having reduced levels of overall piliation as well as PilC expression (Winther-Larsen et al., 2001). Conclusions were therefore drawn that PilV plays a role in the presentation of PilC as an adhesin and pilus biogenesis. However, given the confusing role PilC plays in both pilus biogenesis and adhesion, further investigation is required to clarify the interaction of T4P with CD46.

Experimental evidence in support of CD46 as the pilus receptor	Experimental evidence against CD46 as the pilus receptor
• Anti-CD46 antibodies inhibit bacterial binding to target cells (Kallstrom et al., 1997) • Meningococci specifically bind CD46 expressing CHO cells (Kallstrom et al., 1997) • Purified CD46 inhibits bacterial adhesion (Kallstrom et al., 1997) • Piliated <i>Neisseria</i> co-aggregate with <i>Staphylococcus aureus</i> coated with CD46 (Kallstrom et al., 1997) • Purified pili bind protein same size as CD46 (Kallstrom et al., 1997) • Transgenic mice expressing CD46 are susceptible to meningococcal disease (Johansson et al., 2003) • Soluble (S) pilin interacts with CD46 as shown by immunoblotting (Rytönen et al., 2001)	• Adherence by bacteria does not correlate with levels of expressed CD46 (Kallstrom et al., 2001b, Tobiason and Seifert, 2001) • No direct interaction between purified CD46 and pili ever shown • Agglutination of erythrocytes with gonococcus is observed despite the fact these host cells do not express CD46 (McNearney et al., 1989, Koransky et al., 1975a)

Table 1.1 Evidence for and against the role of CD46 in *Neisserial* adhesion to host cells

1.5.3 Insights into PilE Function

T4P-mediated adhesion is associated with host cell signalling events. *N. gonorrhoeae* colonisation was shown to reduce apoptosis in a fallopian tube infection model in an adhesion dependent manner (Reyes et al., 2007). In an earlier study, reports alluded to the fact that meningococcal pili are toxic towards the vascular endothelium (Dunn et al., 1995, Melican et al., 2013). Piliation was shown to work synergistically with LPS release to cause cytotoxic damage to endothelial surfaces to which meningococci are adhered. A correlation was observed between strains expressing PilE sequences which conferred increased bacterial adherence and increased host cytotoxicity suggesting the exact sequence of PilE may be important (Miller et al., 2014a). T4P-dependent adhesion of *N. meningitidis* was also shown to induce Ca^{2+} signalling pathways in the host

epithelia ME180 model through casein kinase II (Kallstrom et al., 1998). More recently, the hypervariable D-region of PilE has been implicated in modulating host cell specificities and signalling responses (Miller et al., 2014b), the molecular mechanisms for which are not completely clear.

Pili are known to be immunogenic based on early mouse experiments as well as the presence of anti-pili antibodies in patient sera (Schoolnik et al., 1984). Pilin proteins were suggested as vaccine candidates due to the fact that they are surface exposed and expressed by most virulent strains of *N. meningitidis* (Forest et al., 1996). However, given the hypervariable nature of class I PilE genes, efforts were focussed on conserved PilE epitopes and minor pilin proteins (Cehovin et al., 2011), which were conserved irrespective of whether the strain underwent hypervariation of its PilE sequence or not, despite the D-region being reported as the immunodominant epitope (Schmidt et al., 2010). Such efforts have so far been fruitless. Pilin proteins were also implicated as being immunosilent in one study which used chimeric pilin proteins lacking a D-region, mimics of class II PilE proteins to some extent, implying antibodies could not be raised against these proteins, irrespective of any sequence variation issues and so pursuit of a PilE-based vaccine halted (Hansen et al., 2007). No work has been published investigating whether recombinant PilE proteins constitute good vaccine candidates in the case of *N. meningitidis*.

1.5.4 THESIS AIMS:

- Develop a method for large scale PilE expression and purification to provide PilE from a variety of strains, both class I and class II sequences
- Attempt to solve the structure of meningococcal PilE, class I and class II
- Investigate the putative interaction between CD46 and PilE by biophysical analysis
- Investigate host cell responses caused by class I and class II PilE proteins
- Investigate the immunogenicity of class I and class II PilE proteins

- Develop a method for expression of meningococcal PilC and begin functional analysis if successful

1.6 Colonisation and Bacterial Cell Division in the Context of Infection by *Neisseria meningitidis*

N. meningitidis divides by cellular fission like other bacterial species. This process is the splitting of the mother cell into two new daughter cells in a spatially regulated manner which also faithfully partitions the genetic material. FtsZ is a homologue of tubulin, a component of the eukaryotic cytoskeleton, found in most prokaryotes (Bi and Lutkenhaus, 1991). This GTPase self associates into protofilaments which develop into the Z-ring (de Boer, 2010). FtsZ is thought to be the first protein to localise at the site of future bacterial cell division. Many cell division accessory proteins are recruited by FtsZ which tether this protein to the membrane and regulate the progression through cell division, forming a mature divisome (Lutkenhaus et al., 2012). Other Z-ring associated proteins are involved in modulating Z-ring stability and dynamics (Galli and Gerdes, 2010), (Durand-Heredia et al., 2012), (Hale et al., 2011). It is thought that as the Z-ring constricts, the membrane is pulled with the Z-ring and the mother cell separates into two daughter cells.

Very little is known about bacterial cell division under conditions of low-oxygen despite the fact that many pathogens thrive in anaerobic conditions, particularly in mammalian hosts (Marteyn et al., 2011). A recent report which screened a library of *Shigella flexneri* mutants, identified mutants defective for gastrointestinal tract colonisation (Marteyn et al., 2010). One of the mutations was mapped to an unannotated open reading frame. Further characterisation of the protein encoded indicated that this novel ATPase is recruited to the Z-ring at the late stages of cell division (Marteyn et al., 2014). Retrospectively this protein was designated ZapE (Z-ring associated protein E). Overexpression or inactivation of ZapE results in a bacterial cell elongation phenotype, attributed to modulation of Z-ring dynamics. Under conditions of low oxygen, a temperature dependent elongation phenotype is observed in the knock out mutant.

N. meningitidis is a strict aerobe, though is capable of growing in conditions of limited oxygen as this pathogen is able to use a denitrification pathway, reduction of nitrite to nitrous oxide via

nitric oxide, to supplement growth (Rock et al., 2005). Such low oxygen conditions may be encountered by *N. meningitidis* during disease due to blood vessel damage causing poor perfusion at tissues due to a lack of oxygenated blood flow. However, the relevance of ZapE in Z-ring modulation solely under conditions of low oxygen is called into question as this protein is conserved in most Gram-negative pathogenic species, not all of which are able to grow anaerobically.

Given the compact and ever-evolving nature of the *N. meningitidis* genome, it is perhaps unlikely that this gene would have been conserved if it played no role in promoting colonisation and virulence through bacterial cell division. With a role in bacterial cell division, this protein could be an interesting target for novel antibiotic therapeutics. Having only been recently discovered, no *in vivo* data is available as to the potential role in virulence this protein plays in *N. meningitidis* and further characterisation is required.

1.6.1 THESIS AIMS:

- Characterise ZapE from *N. meningitidis* and *E. coli* through biophysical, structural and functional analyses

2. Investigations into Host Specificity of *N. meningitidis*

Despite advances in specific serogroup vaccinations over the past few decades, meningitis remains a major concern with over 3000 cases of infection reported each year in the U.K. (Agency), 2013). For over 90% of these incidents, *Neisseria meningitidis* serogroup B is the causative agent for which, at the time of completing this project, there was no effective vaccine available on the market. fHbp is a promising antigen, eliciting a serum bactericidal response against a broad range of *N. meningitidis* strains (Welsch et al., 2008). The human fH-fHbp interaction can be knocked out by mutating two glutamate residues to alanine (Schneider et al., 2009), which would constitute a better vaccine candidate as antibodies would not be made against the complex but fHbp alone and thus reduce the chance of an autoimmune response.

Testing such vaccines however is problematic given that *N. meningitidis* is a bacterial species highly adapted to its human host, the nasopharynx of which is its sole natural reservoir.

Commonly used animal models are inappropriate to investigate meningococcal disease, relying on over-loading the model's immune system with physiologically irrelevant titres of bacteria in combination with a high dose of iron to promote rapid bacterial growth within the blood stream (Johansson et al., 2003), thus not representing the true and natural pathogenic progression of meningitis *in vivo*.

Given the prominent role the fHbp-fH interaction has in enhancing the serum resistance of *N. meningitidis*, this implies that this interaction is specific for human fH (Granoff et al., 2009b). As such, it may be hypothesised that fH from different species may have a different binding interface in terms of shape or charge distribution, due to the known sequence differences across species. Therefore, these different factor H molecules could not be bound by fHbp with the same high affinity as in humans and therefore aid evasion of host immune responses in sera.

This work aimed to determine the molecular basis for meningococcal resistance in mice at the level of fH through the recombinant production of mouse fH SCRs 6 and 7 (m-fH₆₇), structural

analysis of this protein and SPR analysis with fHbp to determine the binding characteristics of this interaction.

2.1 Functional Analysis of m-fH₆₇

2.1.1 Expression and Purification of m-fH₆₇

The gene fragment corresponding to m-fH₆₇ was subcloned into pET15b and a protocol devised for expression and purification of this protein in work previously completed as my Part II research project, part of my undergraduate degree work. See **Table 2.1**.

Construct Name	Amino Acid Sequence of Open Reading Frame	Molecular Weight (Da)	Theoretical pI
m-fH ₆₇ pET15b	MALKPCEFPQFKYGRLYEESLRPNFPVSIGNKYSYK CDNGFSPPSGYSWDYLRCTAQGWEPEVPCVRKCVF HYVENGDSAYWEKVYVQGQSLKVQCYNQYSLQNG QDTMTCTENGWSPPPKCI	14346.1 (14241.9 without start Met)	6.54

Table 2.1 Expressed sequence of pET15b-m-fH₆₇ construct

The start methionine is the only additional amino acid in the expressed protein sequence. Previous mass spectrometry analysis showed that this N-terminal methionine was cleaved after expression of the protein.

The pET15b-m-fH₆₇ expression vector was transformed into *E. coli* B834 (DE3) cells for expression.

Overnight cultures were diluted 1:100 to allow growth of 4 x 1 l LB culture with 100 mg/ml ampicillin. Cultures were grown at 37 °C until OD₆₀₀ reached 0.6 and were then induced by addition of 1mM IPTG. After 3 hours further growth at 37 °C, the cells were harvested. The cell pellets were resuspended in a total of 40 ml PBS supplemented with 1 mg/ml lysozyme (Sigma), 400 U/μl DNase I (Sigma) and 1 mM EDTA (resuspension buffer). After homogenisation at 15,000 psi (Emulsiflex C.5, Avestin), 20 μl Tween-20 was added to the sample and gently mixed for 10 minutes at 4 °C before the cell lysate was clarified by centrifugation at 33,000 xg. The pellet was then resuspended in resuspension buffer supplemented with 0.05 % v/v Tween-20. After gentle mixing for 1 hour at 4 °C, the sample was centrifuged at 33,000 xg. The now clean pellet of inclusion bodies was denatured and reduced in 40 ml solubilisation buffer consisting of 8 M urea, 50 mM Tris, 1 mM EDTA, 25 mM DTT, pH 8.0 and rocked for 2 hours at 4°C. The protein solution

was then acidified using 37 % HCl to bring the pH to 3-4. The solution was then centrifuged again at 33,000 xg before dialysis overnight using snakeskin molecular weight cut-off (MWCO) 10,000 (Pierce) in a buffer of 6 M urea pH3-4 at 4 °C. After regular changes in the dialysis buffer, the protein was then added drop-wise to a degassed refold buffer (500 mM arginine, 20 mM ethanolamine, 2 mM cysteine, 1 mM cysteine, 1 mM EDTA, pH 11) at 4 °C and stirred gently overnight. The next day, the refold mixture was concentrated to 50 ml using Vivaflow MWCO 10,000 (Sartorius) and then dialysed into size-exclusion chromatography buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl). The different refold species were separated using size-exclusion chromatography with a Superdex S75 16/60 column (GE Life Sciences) and the peak corresponding to correctly folded m-fH₆₇ was assayed by SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) using a 15% acrylamide gel. It was shown to be a single species consistent with the predicted mass of 14.1 kDa. See **Figure 2.1**.

The over-expression of this protein in inclusion bodies yields large amounts of protein and the preparation process renders the sample very pure by the point of size-exclusion chromatography. The different species seen in the size-exclusion trace represent differently folded species of the refolded protein. The reason monomeric, correctly folded m-fH₆₇ elutes from the column at 2.5 CV is due to its intrinsic function as a sugar binding protein (Prosser et al., 2007a). *In vivo*, glycosaminoglycans on the surface of host cells are the natural ligand of fH and the superdex column matrix is composed of sugars of similar structure.

Under the conditions used for protein purification, the refolding of m-fH₆₇ was relatively inefficient yielding 2.5 mg of correctly folded protein per 4l of culture. Approximately 200-300 mg of protein is purified at the stage of inclusion body solubilisation giving a refold efficiency of approximately 1 %. Correct disulphide bond formation of m-fH₆₇ in this peak was previously shown, through work conducted during my Part II project, to be correct by matrix assisted laser

desorption/ionisation mass spectrometry (MALDI-MS) of trypsinised samples with and without reduction.

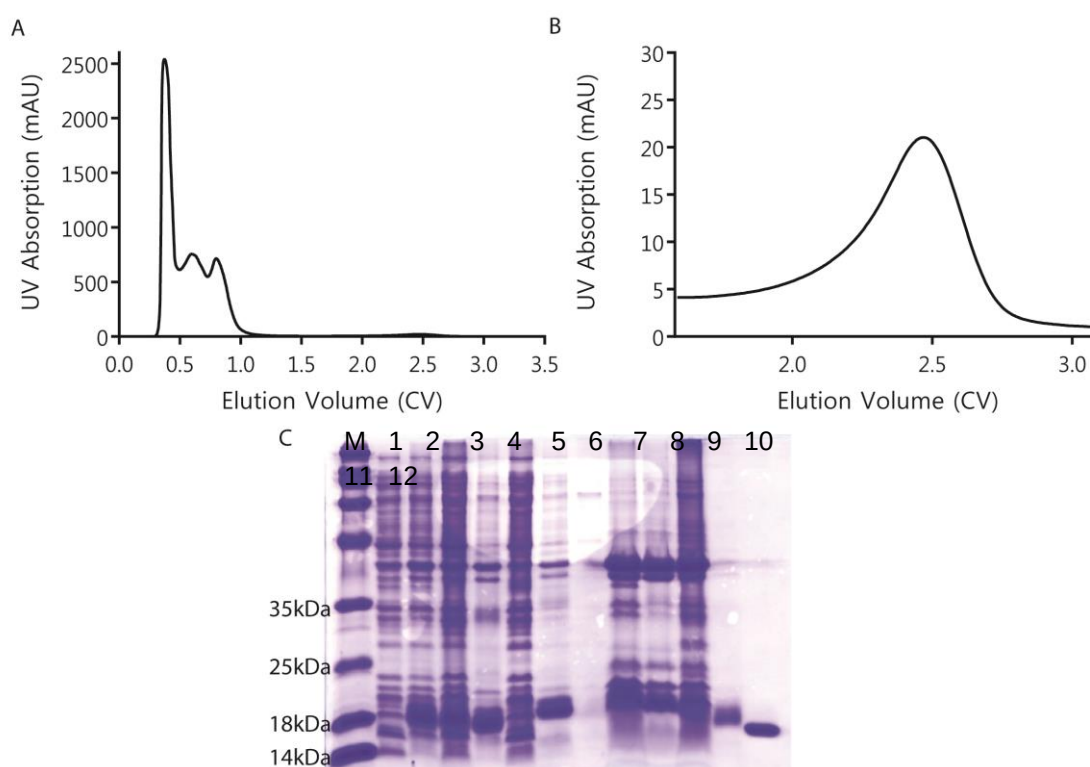


Figure 2.1 Size-exclusion chromatography and SDS-PAGE analysis of the purified m-fH₆₇

Panel A shows the elution trace of the purified m-fH₆₇ sample. Panel B shows the peak corresponding to the elution of m-fH₆₇ from the column. Panel C shows the non-reducing SDS-PAGE analysis of samples taken throughout the expression and purification of m-fH₆₇: M - molecular weight markers, 1 - pre-induction cells, 2 - post-induction cells, 3 - homogenate, 4 - resuspended pellet, 5 - supernatant after inclusion body harvesting by centrifugation, 6 - solubilised inclusion bodies, 7 - pellet after dialysis in 6 M urea pH3-4, 8 - supernatant after dialysis into 6 M urea pH3-4, 9 - sample mid way through concentrating with Vivaflow, 10 - sample after concentrating with Vivaflow, 11 - sample after dialysis into 50 mM Tris, 150 mM NaCl, pH 7.5, 12 - sample from size-exclusion chromatography peak. The inconsistent position of the m-fH₆₇ band is due to the non-reduced nature of the gel and also the extremely varied buffer conditions of the samples taken throughout the purification which affects gel migration.

2.1.2 Surface Plasmon Resonance Analysis

To investigate whether m-fH₆₇ bound fHbp, a series of SPR experiments were conducted. All experiments were conducted using a Biacore 3000 machine at 25 °C. After the CM5 chip (GE Life Sciences) was docked and primed in running buffer of 10 mM HEPES pH7.5, 150 mM NaCl, 1 mM

EDTA, 0.005 % Tween-20, 80 μ l 50 mM NaOH was injected over all 4 flow channels followed by 120 μ l 1:1 mix of NHS:EDC (N-hydroxysuccinimide: ethyl(dimethylaminopropyl) carbodiimide). Approximately 1500 RU of fHbp variants 1, 2 and 3 were coupled to the sensor surface in channels 1, 2 and 3 respectively. In each case, a 20 μ M stock of fHbp was diluted 1:10 with 10 mM sodium acetate pH 4.5 before being injected onto the relevant flow channel. By analysing the potential interaction with a representative fHbp from each class, it was possible to rule out any bias in binding to a particular subset of strains of *N. meningitidis*. All fHbp samples were produced and supplied by Dr Joseph Caesar, University of Oxford. Flow channel 4 was left as a reference channel. Next 120 μ l 500mM ethanolamine pH 8.5 was injected over all four channels to quench any remaining activated carboxyl groups. Following equilibration of the chip in running buffer, m-fH₆₇ or h-fH₆₇ (supplied by Dr Joseph Caesar, University of Oxford) was then flowed in the fluid phase at a flow rate of 40 μ l/min. Each injection was 160 μ l. After each injection, a dissociation time of 600 seconds was allowed and the chip was regenerated with three 20 μ l injections of glycine pH 3.0. The data were analysed using the Biaevaluation software package. All injection series were repeated three times.

Principally, the interaction with fHbp variant 1 was investigated as most published kinetic analysis of the h-fH₆₇ interaction with fHbp had been conducted with this variant (Schneider et al., 2009). A one-to-one stoichiometry was expected if binding was seen given the kinetics and structure of h-fH₆₇ with fHbp. An example series of curves can be seen in **Figure 2.2A**. However, no kinetic data could be fitted from the Biacore analysis of m-fH₆₇ with fHbp variant 1 using any of the models available within this programme. Instead, the equilibrium response given when m-fH₆₇ was flowed over the chip can be plotted against the concentration of m-fH₆₇ in that instance (see **Figure 2.2**), though this analysis is flawed in that not all of the curves reached exact equilibrium. However, this was the only basic analysis which could be conducted given the constraints of the data. The binding curve for 8 μ M m-fH₆₇ being flowed over the chip has a different shape compared to those at other concentrations. This anomaly could be due to an error in the injection

of the sample or some other error in the way the experiment was run, as all samples were made from a serial dilution of one stock of protein, or due to problems in the base line subtraction. It can also be seen on nearly all the binding curves that the response level does not drop back down to zero straight after the injection was finished. This implies that there are two modes of interaction, a high probability event which is weak and has a fast off rate, as well as a low probability event with a slower off rate.

Data analysis by this method produced a straight line graph (see **Figure 2.2B**), indicating that the concentrations of m-fH₆₇ tested are well below that of the K_D of this interaction. As such we can presume that the K_D is far greater than the highest concentration of m-fH₆₇ tested of 16 µM. As the predicted K_D value of the interaction of m-fH₆₇ is greater than 16 µM, but the calculated physiological concentration of fH in serum is 5 µM (Hakobyan et al., 2008), this interaction can be deemed irrelevant as it is too weak to have any significant effect on fH sequestration *in vivo*. This implies that the mouse complement system does not allow serum resistance of *N. meningitidis* by this means. Another interpretation of this data is a biphasic fit of the data in **Figure 2.2B**, the initial phase representing the interaction between m-fH₆₇ and fHbp and the latter phase representing build up of aggregation of the m-fH₆₇ protein on the chips surface..

Data shown in **Figure 2.2C** show that for all three variants with fH₆₇ at physiological concentration (5 µM), binding is reduced for the m-fH₆₇ protein compared to h-fH₆₇. This indicates similar binding properties for all three fHbp variants.

Overall, binding of fHbp by m-fH₆₇ has been shown to be irrelevant *in vivo*, implying that serum resistance in the mouse can not be conferred to *N. meningitidis* by this mechanism.

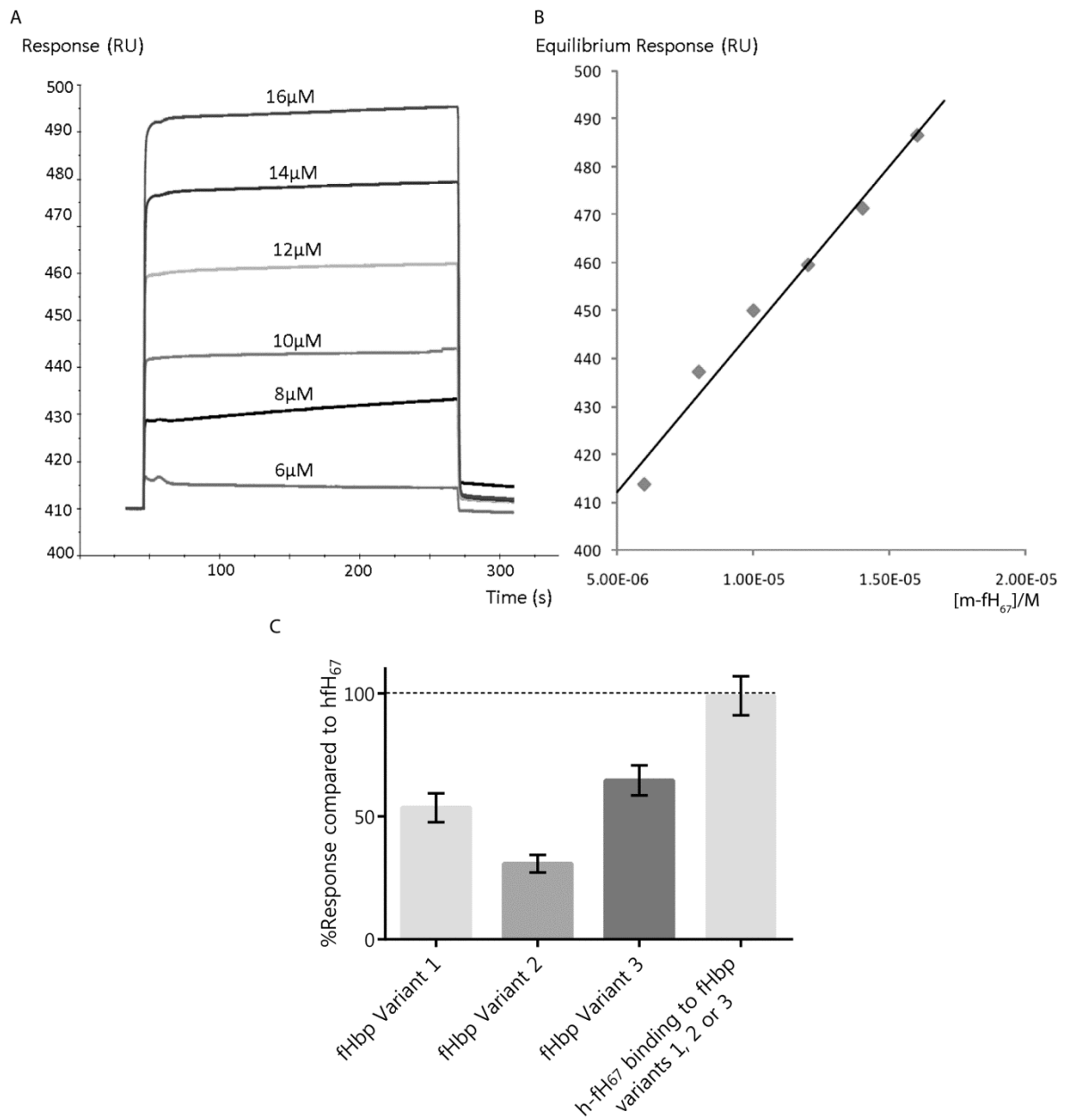


Figure 2.2 – SPR results of m-fH₆₇ binding to fHbp

Panel A shows the equilibrium binding response of m-fH₆₇ to fHbp variant 1. Panel B shows the equilibrium responses plotted against fHbp concentration. Panel C shows the relative binding response of all three fHbp variants as a % of the response seen for binding h-fH₆₇ when at physiological concentration for the h-fH₆₇ interaction with fHbp, the error bars represent the standard deviation of these measurements from the three repeat experiments.

2.2 Structural Analysis of m-fH₆₇

2.2.1 Crystallisation of m-fH₆₇

The sample produced by the method described in **2.1.1** is homogenous and as such is a good candidate for structure solution by X-ray crystallography. The m-fH₆₇ protein sample was concentrated to 11.9 mg/ml using Amicon Ultra-15 and -4 Centrifugal Filter Units NMWL (Nominal Molecular Weight Limit) of 10,000 Da and used to set up sparse-matrix crystallisation screen experiments. A concentrated protein sample will produce crystals when the protein molecules slowly precipitate out of solution and form an ordered lattice structure (Rupp and Kantardjieff, 2010). Whilst the properties of the protein molecule in question contribute significantly to its ability to form protein crystals, such as the presence of less ordered domains and the overall stability of the structure, the chemical properties of the mother liquor and protein buffer are also very important. As such, high throughput crystallisation trials were completed which used pre-made 96-well crystallisation screens to vary the pH, salt concentration, nature and concentration of organic precipitates and various additives within the solution of the crystallisation drop.

Crystallisation trials were set up using a Hydra 96 (Alpha Biotech Ltd) which transferred 75 µl mother liquor from the 96-well deep-well block of the crystallisation screen to the reservoirs of a MRC 2-drop crystallisation plate (Molecular Dimensions). All crystals were grown using the sitting-drop vapour-diffusion technique using an Oryx Nano-crystallisation robot (Douglas Instruments, U.K.). A small drop of concentrated protein solution is mixed with mother liquor at two desired ratios and dispensed in two small wells immediately adjacent to the 75 µl reservoir of mother liquor. Over time, the precipitant containing reservoir promotes the gradual dehydration of the drop, in an attempt to promote crystallisation of the protein within the drop. Drops contained 0.1 µl reservoir solution and 0.1 µl protein solution (11.9 mg/ml m-fH₆₇, 50 mM Tris, 150 mM NaCl, pH 7.5), corresponding to the 50 % dilution, or 0.15 µl reservoir solution and 0.05 µl protein solution, corresponding to the 25 % dilution. Plates were then sealed using a StarSeal Advanced Polyolefin

Film (STARLAB) and incubated at 21 °C. Six screens were set up using these conditions: Structure screen I + II, PACT *premier*, JCSG-*plus*, Proplex™, The Stura Footprint Combination and Morpheus® (all Molecular Dimensions), see **Table 2.2**.

Screen	Properties
Structure screen I + II	Sparse matrix screen based on previously successful crystallisation conditions – wide range of precipitants, salts, buffers and pH
PACT <i>premier</i>	Systematic screen of anions, cations, pH and PEGs
JCSG- <i>plus</i>	Sparse matrix screen based on previously successful crystallisation conditions – wide range of precipitants, salts, buffers and pH
Proplex™	Sparse matrix screen based on previously successful crystallisation conditions for growing crystals of protein complexes
The Stura Footprint Combination	Systematic screen allowing investigation of the protein solubility limit through variation of the pH, PEGs and salts
Morpheus®	Sparse matrix cryoprotecting screen with low molecular weight ligands intended to promote crystal contact formation as well as a variety of pH, PEGs and salts
The PGA Screen™	Systematic screen using poly-γ-glutamic acid (PGA) polymer with variable salts, buffer, pH and additional precipitant

Table 2.2 Crystallisation screens (all Molecular Dimensions) trialled in the crystallisation of *m-fH₆₇* and their relative properties

The crystals grown from such a protein preparation were predominantly found in the 50 % dilution of the protein solution and in conditions containing a polyethylene glycol (PEG) precipitant. Generally only one, fairly large crystal was found per drop in hit conditions. The crystals took over six months to grow and were all of a similar morphology with some seemingly curved edges. The best diffracting crystal, was found in Proplex condition A9 (0.2M NaCl, 0.1M MES, pH 6.0, 20 % w/v PEG2000MME) at 50 % dilution see **Figure 2.3**.

The crystals were then mounted using CrystalCap nylon Cryoloops (Hampton Research) and then cryo-protected before flash cooling in liquid nitrogen. Crystals were cryo-protected in a 1:5 ratio of ethylene glycol: mother liquor. This condition was selected using the condition optimised cryoprotectant previously determined (McFerrin and Snell, 2002).

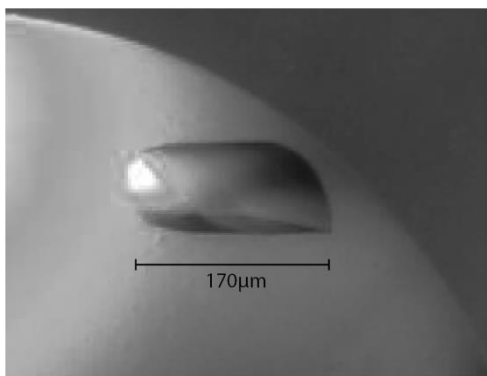


Figure 2.3 – $m\text{-fH}_{67}$ crystal grown in Proplex condition A9

2.2.2 Data Collection and Processing

Data collection was performed by Professor Susan Lea at beamline I03 of the Diamond Light Source, Harwell, U.K. Diffraction data sets were collected at 100K to reduce radiation damage to the sample (Garman and McSweeney, 2007). The crystal was rotated whilst in a $0.9795\text{\AA}$ wavelength X-ray beam through 180° along a single axis using a goniometer. With an oscillation angle of 0.5° , 360 diffraction images were collected, constituting a complete data set.

After collection at the beam line, diffraction images were processed. Images were indexed and integrated using Xia2 (Winter, 2010) in 3d mode using XDS (Kabsch, 2010) and then scaled and merged using SCALA (Evans, 2006, Kabsch, 2010, Winn et al., 2011a). Data processing allows assessment of the quality of the data collection through extraction of the parameters needed for subsequent structure solution and then analysis of various features of the data. First, the reflections are identified on each image. The unit cell parameters can then be calculated and the Laue group determined from the repeating distances and symmetry between reflections. Next, the reflections are integrated to determine the intensity (I_{hkl}) and the intensity standard deviation (σ_{hkl}) of every reflection. The integration of the pixel area is over a 3-dimensional pixel area for fine-sliced data such as these. By measuring the counts outside of the reflection area, the background can be determined and thus I_{hkl} and σ_{hkl} calculated. Next the data are scaled. As the same reflection may be captured over many images, especially in finely sliced data collection,

these partial reflections must be combined into a single reflection. In addition, intensities of symmetry related reflections must also be brought onto the same scale to account for fluctuations in beam intensity, anisotropic absorption, radiation damage, noise and crystal orientation deviations (Evans, 2006). Finally the data are merged so that symmetry related reflections and multiply measured reflections are combined to give a single I_{hkl} value for each unique reflection.

The quality of the processed data were analysed by the statistics detailed in **Table 2.3**. Processing statistics are given in **Table 2.4**. The crystal gave a diffraction pattern compatible with a primitive orthorhombic lattice. Systematic absence of the $h00$, $0k0$ and $00l$ reflections, suggests that the crystal belongs to spacegroup $P2_12_12_1$ (see **Figure 2.4**).

The data quality statistics over the whole data set were favourable. As such the data set was kept in its entirety as this gave an almost complete set of reflections (99.8%) and gave a high multiplicity. Given the I_{hkl}/σ_{hkl} showed that there was measureable signal up to $1.58\text{\AA}$, this was used as the highest resolution shell. The R_{merge} and R_{pim} are both higher than expected in the outer shell and confusingly the R_{pim} is not lower than the R_{merge} . This could indicate some twinning or other issue with the crystal diffraction. However, given that both values are lower than the benchmark 0.6 cut off and subsequent data analysis proved successful this was not readdressed.

$\langle I /\sigma(I) \rangle = \frac{1}{N} \sum_h \frac{ I_{(h)} }{\sigma(I_{(h)})}$	Signal to noise ratio given as the average I_{hkl}/σ_{hkl} as defined for N reflections in a resolution shell.
$R_{\text{merge}} = \frac{\sum_h \sum_{i=1}^N I_{(h)i} - \bar{I}_{(h)} }{\sum_h \sum_{i=1}^N I_{(h)i}}$	Measurement of the consistency of the data after having been scaled and merged. Calculated as the average deviation of redundant reflection h from the intensity to which it has been scaled to (I_h). If the data has a high level of multiplicity this will increase.
$R_{\text{pim}} = \frac{\sum_h \left(\frac{1}{N-1}\right)^{1/2} \sum_{i=1}^N I_{(h)i} - \bar{I}_{(h)} }{\sum_h \sum_{i=1}^N I_{(h)i}}$	By introducing the term $\left(\frac{1}{N-1}\right)^{1/2}$ into the R_{merge} calculation, multiplicity of the data can be taken into account as this improves the precision of the data.

Table 2.3 Data quality statistics from processing

Diffraction source	I03 - Diamond Light Source (Harwell, UK)
Wavelength	0.9795Å
Detector	ADSC CCD
Temperature (K)	100
Space group	P2 ₁ 2 ₁ 2 ₁
Z	4
Unit cell parameters (Å)	a = 29.6, b = 34.6, c = 107.4
Solvent content (%)	35.7
Resolution (Å)	35.79-1.58 (1.58-1.58)
No. unique reflections (criteria for reflection)	15809 (2037)
No. observed reflections	102788 (9846)
R_{merge}	0.035 (0.48)
R_{pim}	0.042(0.50)
[I/σ(I)]	28.7 (3.2)
Completeness (%)	99.8 (99.7)
Multiplicity	6.5 (6.4)
CC1/2	0.99 (0.88)
Matthews coefficient V_M (Å³ Da⁻¹)	1.91
Data processing software	Xia2, XDS and SCALA

Table 2.4 - Data collection and structure solution statistics

Values for the outer shell are given in brackets.

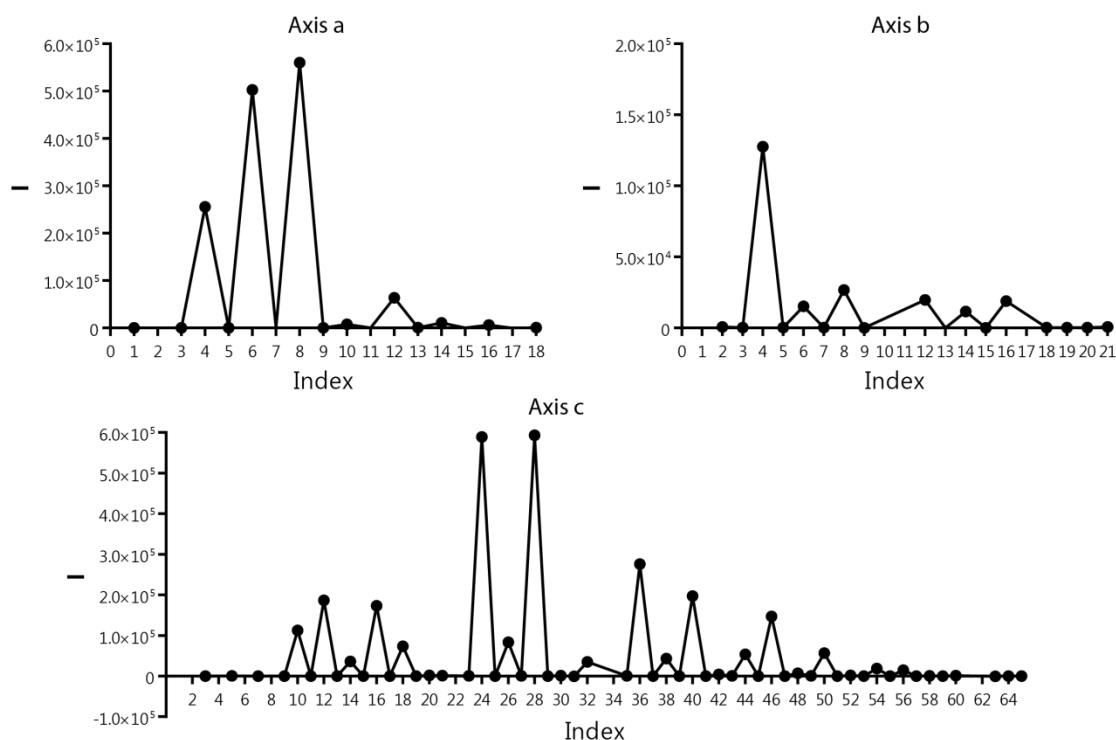


Figure 2.4 Systematic absences of the $h00$, $0k0$ and $00l$ reflections suggest the crystal belongs to spacegroup P2₁2₁2₁

2.2.3 Characterisation of the Asymmetric Unit

Multiple copies of the protein molecule related through non-crystallographic symmetry functions may constitute the asymmetric unit. The Matthew's coefficient (V_m) (Matthews, 1968, Kantardjieff and Rupp, 2003) allows calculation of the number of molecules of m-fH₆₇ within the asymmetric unit in relation to the solvent content of the crystal (V_s):

$$V_m = \frac{\text{Volume of unit cell}}{\text{MW protein (Da)} \times Z \times \text{No. molecules in the asymmetric unit}}$$

$$V_s = 1 - \left(\frac{1.23}{V_m} \right)$$

Within the CCP4 software suite (Winn et al., 2011b), the Matthews programme (Kantardjieff and Rupp, 2003) was run showing that unequivocally only one copy of the m-fH₆₇ protein molecule was present within the asymmetric unit. The solvent content of the crystal from which the data set was collected is very low, possibly due to dehydration from poor tray sealing over the extensive period of time the crystals took to grow (6 months).

Number of m-fH ₆₇ molecules per asymmetric unit (MW/Da)	V_m / (Å ³ /Da)	V_s / %	Probability across all resolution ranges
1 (14214.9 Da)	1.91	34.6	1.00

Table 2.5 Calculation of the number of molecules per asymmetric unit

2.2.4 Phase Determination, Refinement and Structure Validation

The phase problem must be solved in order to create an electron density map of m-fH₆₇. The I_{hkl} values in the processed data output correspond only to the amplitudes of the scattered waves but an estimation of the phase angles of these waves is required in order to proceed in the structure solution. As m-fH₆₇ shares 55% sequence identity with the published structure of human factor H (**Figure 2.5**) SCRs 6 and 7 (h-fH₆₇) (PDB code: 2W80), this was used as a model for molecular replacement.

PHASER (Mccoy et al., 2007, Read et al., 2007) within the CCP4 software suite, was used to perform the molecular replacement. This programme uses the maximum likelihood functions to first rotate and then translate the search model within the unit cell until agreement is made between the observed and calculated structure factor amplitudes. To monitor the success of the molecular replacement, the different solutions are assessed by both the log-likelihood gain score which indicates the extent to which the model used for molecular replacement gives a better prediction of the data than a random array of atoms. Z-scores are generated for the rotation and translation functions (RFZ and TFZ respectively). Clash scores are also reported by PHASER.

```

m-fH67      ALKPCEFPQFKYGRLYYEESLRPNFPVSINKYSYKCDNGFSPPSGYSWDYLRCTAQGWE
h-fH67      TLKPCDYPDIKHGGLYHENMRRPYFPVAVGKYYSYCDEHFETPSGSYWDHIHCTQDGWS
              : * * * : : * * : * * : * : * * * : : * : * * * : * . . * * * * : * * : * * .

m-fH67      PEVPCVRKCVFHYVENGDSAYWEKVYVQGQSLKVQCYNGYSLQNGQDTMTCTENGWSPPP
h-fH67      PAVPCLRKCYFPYLENGYNQNYGRKFVQGSIDVACHPGYALPKAQTTVTCTMENGWSPTP
              * * * * : * * * * * : * * * * . : : : * * * : * . * : * * : * : . * : * * * * * * * . *

m-fH67      KCI
h-fH67      RCI
              : * *

```

Figure 2.5 – Sequence alignment of h-fH₆₇ and m-fH₆₇

The sequence of the recombinant m-fH₆₇ protein produced can be seen in the alignment above (m-fH amino acids 339-462, h-fH amino acids 321-444). The start methionine was shown in previous work to be cleaved when this protein is expressed in *E. coli*. Residues R341, K351 and H371 which are important in the h-fH₆₇-fHbp interaction are highlighted in red.

Using CHAINSAW (Stein, 2008), a model of h-fH₆₇ was generated in which the side chains had been pruned back to the last common atom of each amino acid on the basis of the sequence alignment of h-fH₆₇ with m-fH₆₇. Molecular replacement using PHASER with this model generated just one solution. However, the log-likelihood gain score was low given the expected conservation of the structures which share 55 % sequence identity. As such, the molecular replacement was conducted searching sequentially for SCR6 and then SCR7 by splitting the CHAINSAW modelled-h-fH₆₇ into two pdb files corresponding to its respective domains. Completing the search with these

two models gave vastly improved scores for the solution determined indicating that this was likely the correct solution.

Search model (s)	Number of Solutions	LLG	RFZ	TFZ	Number of clashes <3Å
h-fH ₆₇ CHAINSAW	1	11.9	5.2	3.2	0
h-fH ₆ CHAINSAW	1	65.4	5.2	9.1	0
h-fH ₇ CHAINSAW after placing h-fH ₆ CHAINSAW	1	619	16.8	28.8	0

Table 2.6 Molecular replacement solutions for m-fH₆₇ from PHASER

Molecular replacement with PHASER using the sequential search models of h-fH₆ and then h-fH₇ produced an initial model for the structure of m-fH₆₇ and a corresponding electron density map. Analysis of the initial model of m-fH₆₇ and h-fH₆₇ revealed a difference in the orientation of the two SCR domains relative to one another, indicating why searching for each domain sequentially gave the best molecular replacement solution.

2.2.5 Model Building and Refinement

By analysing the model of m-fH₆₇ constructed by PHASER within the electron density map, it was evident that the structure solution was correct.

As CHAINSAW was used to generate the molecular replacement search model, parts of the m-fH₆₇ structure were missing, primarily the side chains. Using Coot (Emsley and Cowtan, 2004), analysis of the difference map, calculated by the subtraction of the model derived structure factor amplitudes (F_c) from the experimentally derived amplitudes (F_o), revealed side chain density along the pruned polypeptide chain model. ARP/wARP (Perrakis et al., 2001) was then used to auto-build in these side chains using the sequence of m-fH₆₇ and the model was then refined using AutoBUSTER version 1.7.5 (Blanc et al., 2004). Further rounds of iterative model building and refinement generated a model of m-fH₆₇ which better fit the electron density. As the phases used to calculate F_o were taken from the m-fH₆₇ model, a $2F_o - F_c$ map was visualised in order to emphasise the experimental data and reduce model bias. Model adjustment was completed in

Coot whilst refinement was completed using a combination of Refmac and AutoBUSTER (Blanc et al., 2004, Winn et al., 2003). This process brings the experimentally and model derived structure factor amplitudes into close alignment, whilst also imposing restraints on bond lengths and angles as well as interatomic distances normally expected in protein chemistry.

Progress through the model building and refinement cycles is monitored through two statistical values. The R factor assesses the discrepancy between the experimental and model derived structure factor amplitudes using the following equation:

$$R = \frac{\sum_h ||F_o| - |F_c||}{\sum_h |F_o|}$$

If refinement is successful, then the R factor should reduce with each subsequent round, as the F_c agrees more closely with F_o .

In order to minimise model bias, 5 % of reflections are not used in the refinement, allowing calculation of a separate R-factor against these called R_{free} . The final structure was validated within Coot (Emsley and Cowtan, 2004) and online using Molprobit (Chen et al., 2010) to assess the chemistry of the final model. Refinement and validation statistics can be seen in **Table 2.7**.

Resolution (Å)	1.58-35.79 (1.58-1.69)
Number of reflections	15758 (2594)
Number of reflections for Rfree	820 (168)
R/Rfree	0.19/0.21 (0.21/0.25)
No. atoms – Protein	1950
No. atoms – Heterogens	133 - Ethylene Glycol
No. atoms – Water	129
RMS deviation bond lengths (Å)	0.009
RMS deviation bond angles (°)	1.07
Residues in Ramachandran plot favoured regions	122
Residues in Ramachandran plot unfavoured regions	0
Molprobit Score	1.84
Molprobit Score Percentile	83 rd

Table 2.7 – Refinement statistics

Values for the outer shell are given in brackets.

2.2.6 The Structure of m-fH₆₇

The structure of m-fH₆₇ (see **Figure 2.6**) shows two classic SCR domains as was predicted. These domains of approximately 60 amino acids have the two signature disulphide bonds which are responsible for maintaining the overall structure of the protein. These disulphide bonds are formed in the expected manner as predicted from the h-fH₆₇ structure. There are very few elements of secondary structure otherwise, accounting for the splayed nature of the SCR domains.

The exact nature of the high resolution data obtained from the m-fH₆₇ is illustrated in the “zoomed in” part of the image, where a clear hole in the electron density map is present in the centre of the aromatic ring of the tyrosine residue seen.

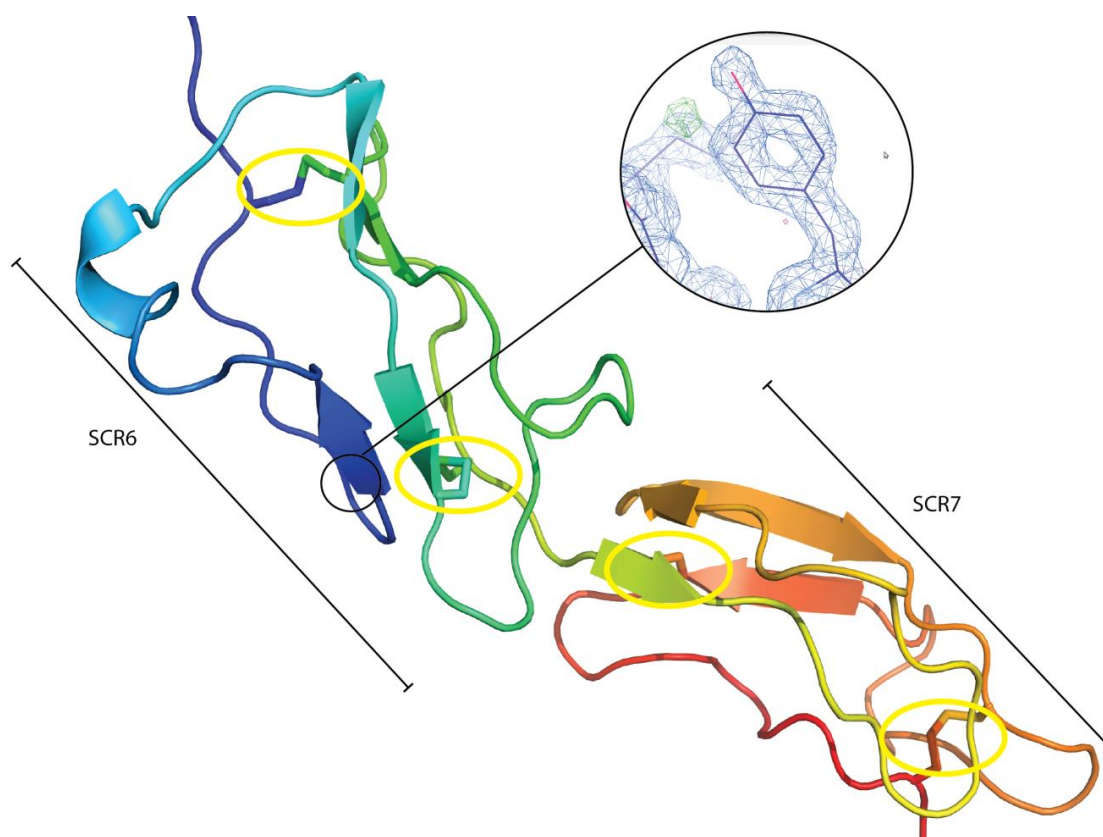


Figure 2.6 Structure of m-fH₆₇

The structure of m-fH₆₇ is shown in cartoon format coloured by rainbow from blue to red from N-terminus to C-terminus. The four disulphide bonds are highlighted by the yellow rings in the figure. The illustration of the model of m-fH₆₇ within the calculated electron density map ($2F_o - F_c$) is contoured to 1.8σ and was visualised in Coot.

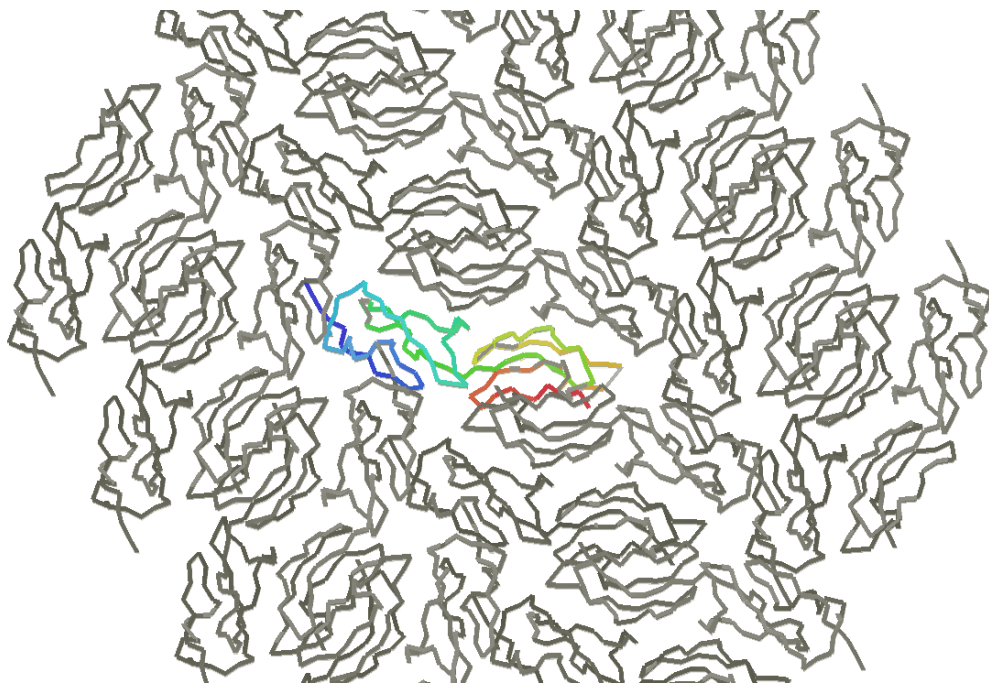


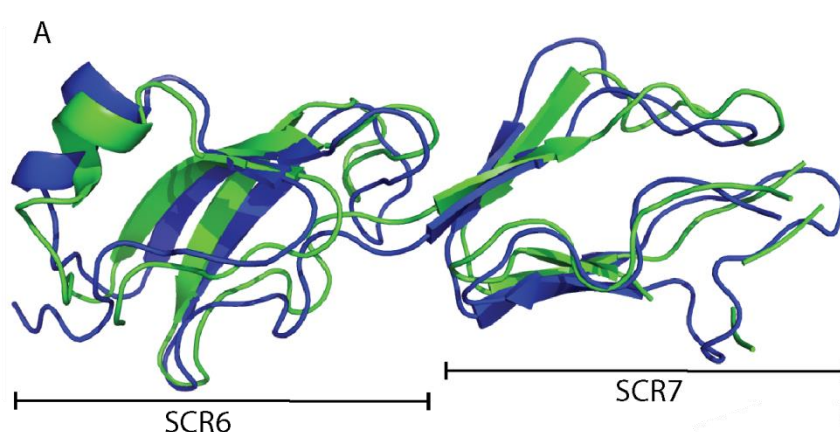
Figure 2.7 Crystal packing of *m-fH₆₇* in the *P2₁2₁2₁* crystal form

The low solvent content of this crystal form, as calculated using the Matthews software, becomes apparent when the crystal packing is examined using Coot. The molecules of *m-fH₆₇* are tightly packed in according to the *P2₁2₁2₁* symmetry operators with very little space for any solvent. See **Figure 2.7**.

2.2.7 Structural Analysis of *m-fH₆₇*

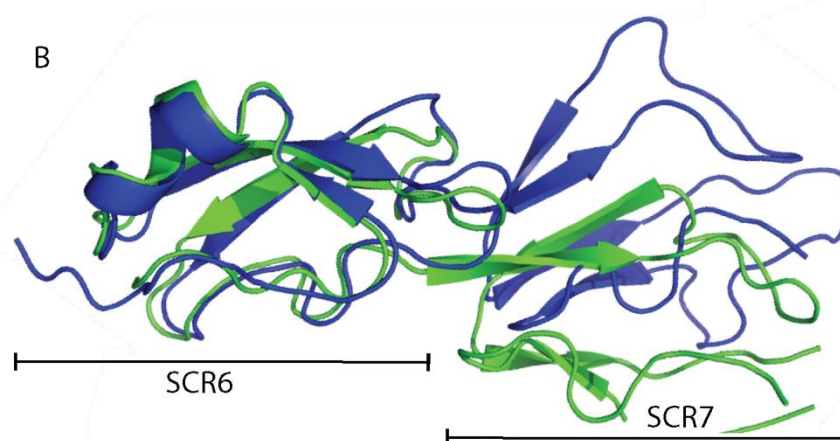
At first glance, the structure looks highly similar to that of *h-fH₆₇*. Superposition analysis conducted of the two structures using Superpose (Maiti et al., 2004). The SSM algorithm was run which matches structures through common secondary structural elements. This superposition analysis does reveal some subtle differences between the two structures, see **Figure 2.8**.

The superpositioning of the two structures can be measured by the root mean square deviation (RMSD) of the polypeptide backbone of each protein relative to each other. For the superpositions over one domain, the RMSD is 3.75Å for both for that particular domain, whilst over the whole structure, the overall RMSD is 4.51Å.



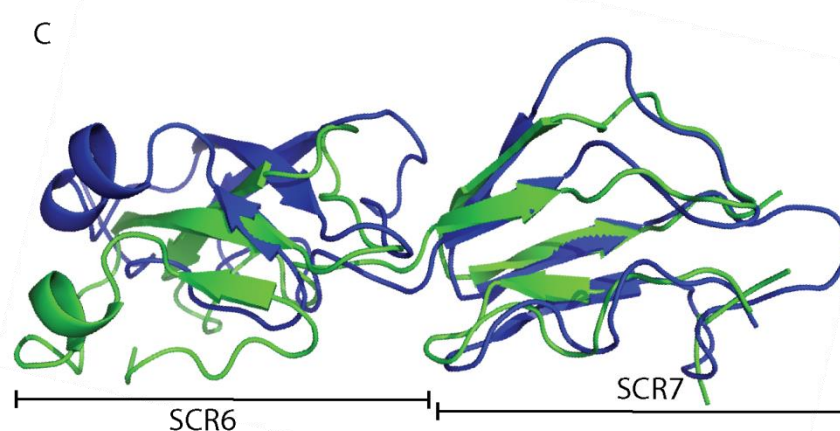
Superposition SCRs 6-7

Alignment over 119 residues: m-fH₆₇ 342-461, h-fH₆₇ 325-444.
RMSD = 4.51 Å overall



Superposition SCR 6

Alignment over 62 residues: m-fH₆₇ 342-404, h-fH₆₇ 325-387.
RMSD = 0.767 Å over SCR6 and 3.75 Å overall.



Superposition SCR 7

Alignment over 57 residues: m-fH₆₇ 404-461, h-fH₆₇ 387-444.
RMSD = 1.405 Å over SCR7 and 3.75 Å overall.

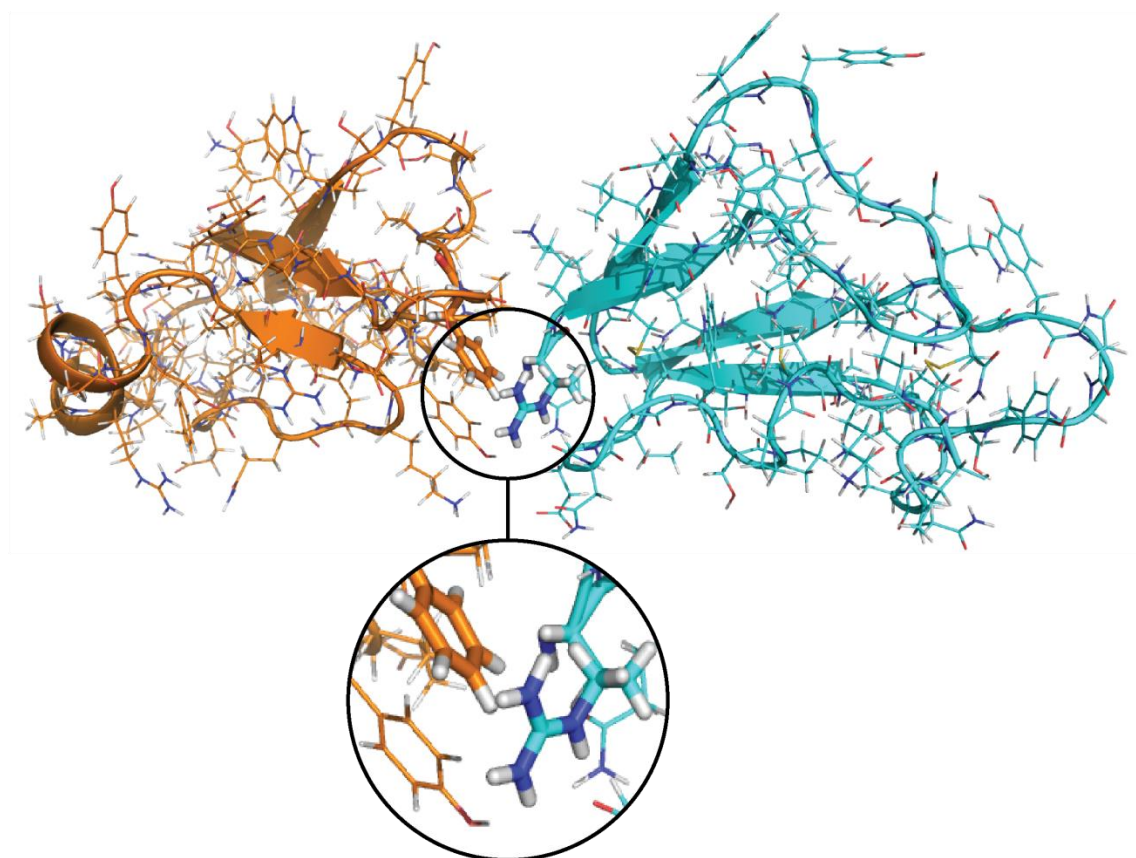
Figure 2.8 Superposition of the m-fH₆₇ and h-fH₆₇ structures

In all three panels are the superpositions of the structures of m-fH₆₇ in dark blue and h-fH₆₇ in green. Panel A shows the superposition if you overlay SCR6 and 7, panel B if you overlay SCR6 and panel C if you overlay SCR7.

These results show that in the instance of the individual domains, the structures overlay very well.

However, when the whole structure is considered the superposition is not as good. This is because the orientation of the domains relative to each other is quite different when you

compare these two structures. As such, it would be expected that the binding interface of m-fH₆₇ relative to h-fH₆₇ would be quite different in shape.



Residue	Atom	Residue	Atom	Atom Overlap/Å	Clashscore
Arg67	CZ	Phe41	HZ	0.88	46
Arg67	CZ	Phe41	CZ	0.69	46
Arg67	NH1	Phe41	CZ	0.75	30
Arg67	HA	Phe41	CE1	0.55	21
Arg67	NH1	Phe41	CE2	0.42	30

Figure 2.9 – Reorientation of SCRs 6 and 7 into a fHbp “binding” conformation reveals residue clashes across the binding interface

Superposition of m-fH SCRs 6 and 7 individually over h-fH₆₇ aligns them in the “human” conformation. Such an orientation introduces clashes into the structure as shown above between phenylalanine-41 and arginine-67. SCR 6 is shown in orange and SCR7 is shown in cyan. The atomic overlaps and clashscores, as calculated in Molprobit, are detailed in the table below.

However, it is possible to imagine that SCR 6 and 7 of m-fH could have a flexible linker connecting the two domains together allowing them to pivot and therefore adopt a conformation more

adept at binding fHbp. By superimposing SCR6 and SCR7 of the m-fH₆₇ structure separately onto h-fH₆₇, this “binding” orientation reveals an increased number of steric clashes across the inter-SCR domain interface as confirmed by analysis using Molprobability, see **Figure 2.9**. Molprobability defines a clash as an overlap of non-donor-acceptor atoms ≥ 0.4 Å. Phenylalanine-41 and arginine-67 in particular now clash in this new conformation, indicating that this angle between m-fH SCRs 6 and 7 is likely to be energetically unfavourable. The orientation of SCRs 6 and 7 of m-fH₆₇ is therefore unlikely to be an artefact of crystallisation.

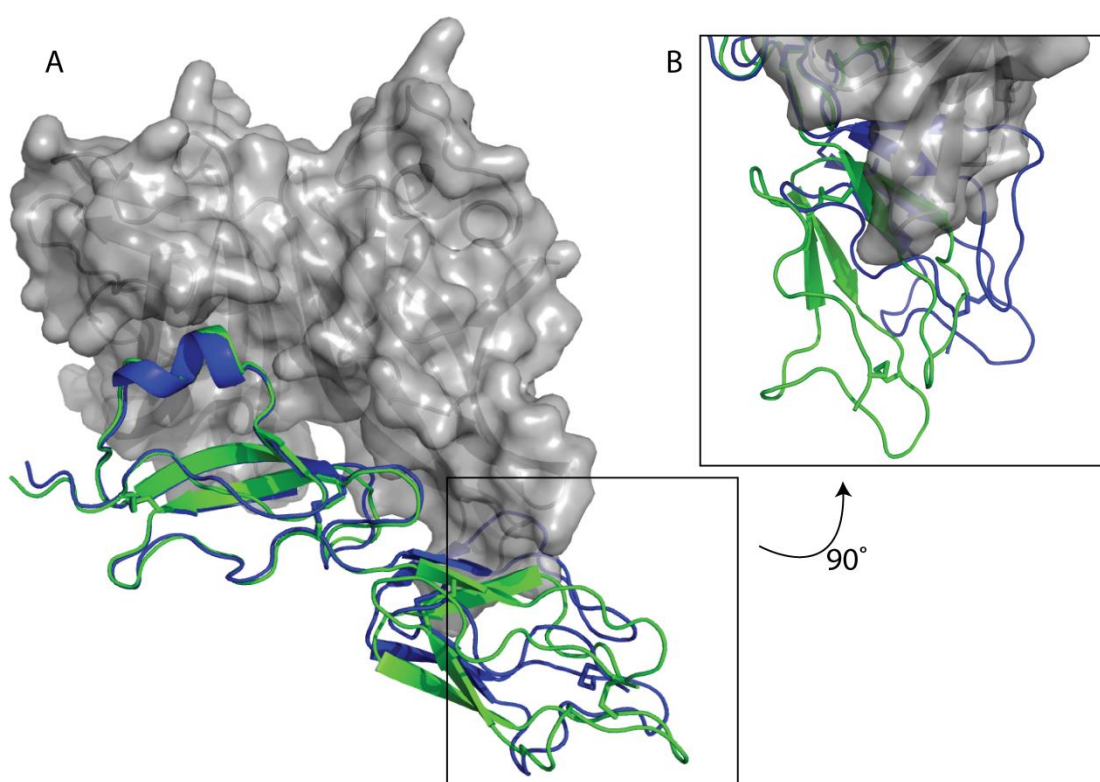


Figure 2.10 – m-fH₆₇ overlaid onto h-fH₆₇ in complex with fHbp

m-fH₆₇ is shown in blue and h-fH₆₇ is shown in green, both in cartoon format. fHbp is shown in as grey surface. Whilst SCR6 of m-fH₆₇ can be overlaid well onto h-fH₆₇ and would appear to dock into fHbp in a similar manner to the human form of this protein, SCR7 is sterically occluded from binding in this manner as highlighted in the diagram above.

If the structure of m-fH₆₇ is overlaid with h-fH₆₇ in complex with fHbp, we can see that this difference in SCR orientation is likely to sterically hinder binding with fHbp, see **Figure 2.10**. With

a difference in SCR orientation to the point of fHbp steric occlusion, this would account for the reduced affinity of m-fH₆₇ for fHbp compared with h-fH₆₇.

However, it is not only the orientation of the two SCR domains which is likely to contribute to the reduced affinity of m-fH₆₇ for fHbp compared with h-fH₆₇. Most of the binding interface of the h-fH₆₇-fHbp complex is within SCR6 whilst SCR7 of h-fH₆₇ was shown to be orientated slightly differently in different crystal forms of the h-fH₆₇-fHbp complex (Schneider et al., 2009). Therefore another explanation for the reduced binding affinity of m-fH₆₇ for fHbp compared to the human form of this protein, could be differences in the surface chemistry of this domain.

Further analysis of the structure was conducted using the programme Consurf (Ashkenazy et al., 2010), which maps sequence conservation onto the structure model using a multiple-sequence alignment of hundreds of different factor H sequences from a variety of species. The analysis revealed a conserved and a variable interface of fH₆₇ on opposing sides of the fH₆₇ structure,

Figure 2.11A and B. The three residues of h-fH₆₇ which were shown to be important for the interaction with fHbp; Arg341, Lys351 and His371 (Schneider et al., 2009)(see **Figure 1.2**) are found within the less conserved interface. Sequence variation across this interface is likely to disrupt the interaction as shown previously with mutation analysis of the interface (Schneider et al., 2009). This result supports the finding that the host specificity of the *N. meningitidis* fHbp interaction is predominantly for the human form of fH, as previously found (Granoff et al., 2009a).

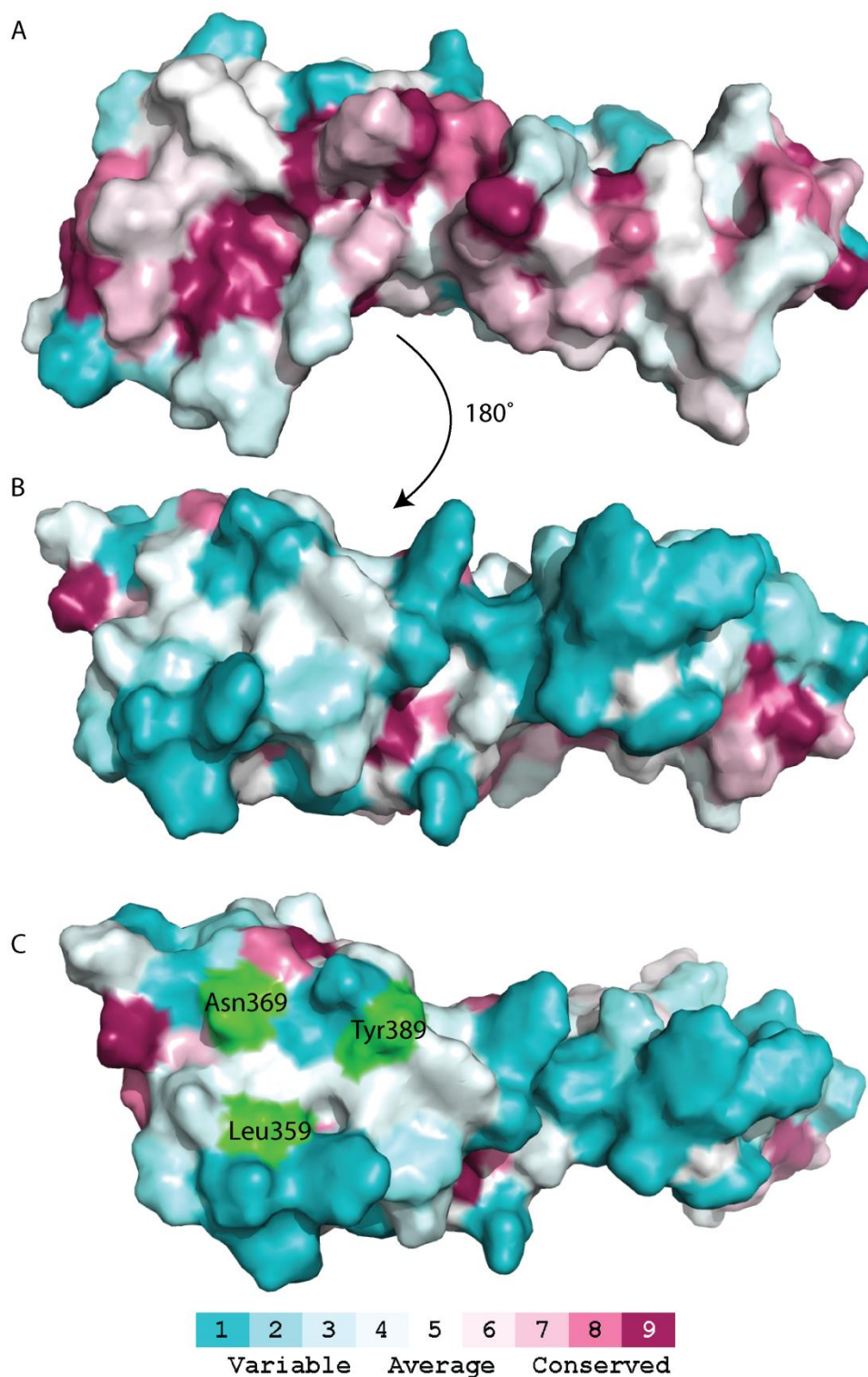


Figure 2.11 Consurf analysis of m-fH₆₇

The conservation of fH₆₇ is mapped onto a surface representation of the m-fH₆₇ structure, the colour key for which is shown at the bottom of the figure. Panels A and B show the model rotated by 180° within the horizontal plane of the page, showing a generally conserved and a variable interface on opposing sides of the fH₆₇ structure. In panel C, three m-fH₆₇ residues are highlighted in green which correspond to the residues of h-fH₆₇, shown to be important for the interaction with fHbp: Arg341 (Leu359), Lys351 (Asn369) and His371 (Tyr389).

APBS (Unni et al., 2011) analysis which calculates a surface charge map of the two structures reveals that the distribution of charge is actually quite similar, despite the lack of conservation as shown by the Consurf analysis. However, the surface topology of the two proteins is quite different which could potentially mean the exact location of charged and hydrophobic residues important in contributing to the interaction interface are not orientated very well in the mouse form of the protein, see **Figure 2.12**.

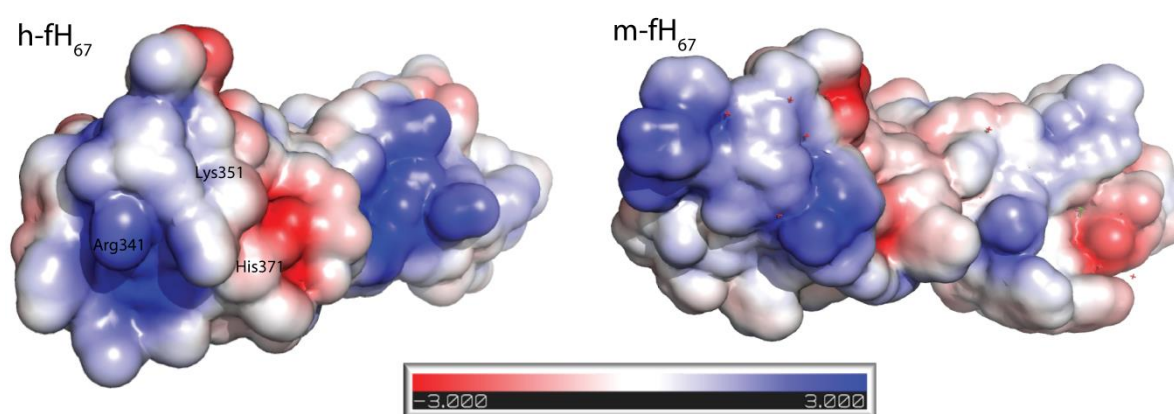


Figure 2.12 Surface charge analysis of m-fH₆₇ compared with h-fH₆₇

Both structures are orientated in the same manner. The prominent positively charged patch on each of the structure, highlighted in the case of h-fH₆₇, corresponds to the portion of the structure involved in crucial charge-based interactions with fHbp. The three key amino acids for fHbp binding by h-fH₆₇ are highlighted on the structure.

As mentioned previously, three residues of h-fH₆₇ were shown by site-directed mutagenesis analysis in previous studies to be crucial for the binding of fHbp (Schneider et al., 2009). Mutation of each residue significantly reduced the affinity of the interaction, whilst double mutants would obliterate the interaction almost completely by comparison to the wild type protein. In all three cases these residues are not conserved: Arg341Leu, Lys351Asn and His371Tyr with respect to the human to mouse sequence difference. Without these important SCR6 residues, this could be a further reason as to why m-fH₆₇ does not bind fHbp with the same affinity as the human form of this protein, as even if the two SCR domains were able to orientate themselves into a

conformation more amenable to binding fHbp, the surface architecture would not promote high affinity binding.

2.3 Perspectives

The results of the m-fH₆₇ project reveal that the binding of this protein to fHbp from variant 1 *N. meningitidis* (MC58) is likely to be minimal *in vivo* with an estimated K_D greater than 16 μ M, considerably higher than the serum concentration of fH which at its highest estimate is approximately 5 μ M (800 μ g/ml) (Hakobyan et al., 2008). Additionally, all variants of fHbp are bound by m-fH₆₇ with reduced affinity compared to that of h-fH₆₇. These findings can now be rationalised with the structure. The decreased binding possibly reflects deviation of the m-fH₆₇ surface structure away from that of h-fH₆₇ both in terms of domain orientation and sequence difference mapped across the binding interface. This supports the SPR data collected and the conclusions drawn from this work. As such, it can be concluded that binding of *N. meningitidis* to fH in mice *in vivo* is physiologically irrelevant.

Using this information to generate chimeric mouse models with “humanised” fH is likely to provide much more relevant answers when studying meningitis vaccines as well as disease progression and pathology. As the interface has such a different shape between the human and mouse forms of this protein, single amino acid mutations which humanise m-fH₆₇ will not improve binding to fHbp. To develop mouse models to study meningitis, chimeric knock-ins containing human SCRs 6 and 7 of the fH gene should display more human-like binding characteristics with *N. meningitidis* and therefore are likely to provide a model which has a humanised pathogenesis with this disease.

Mice with chimeric fH in which SCRs 6-8 are exchanged for the human sequence have since been developed (Johnson et al., 2012). SCR8 was also mutated to the human form due to its role in glycosaminoglycan (GAG) binding with SCRs 6 and 7 (Prosser et al., 2007b). This chimeric animal model was shown to be physiologically relevant due to the normal levels of C3 in sera. fHbp variant 1 proteins which had mutations to obliterate or reduce fH binding were trialled for vaccine candidacy in these mice compared to wildtype fHbp variant 1. The mutated fHbp proteins were

shown to be just as antigenic as the wildtype by enzyme-linked immunosorbent assay (ELISA) analysis as well as producing bactericidal antibodies as shown by serum bactericidal assays (SBAs).

Since completing this project, a new meningitis vaccine made by Novartis International AG and targeted specifically at meningitis serogroup B strains has been approved and is now on the market (Esposito and Principi, 2014). The vaccine contains four key components which are fHbp, PorA (an outer membrane porin), NadA (Neisserial adhesion) and NHBA (Neisserial heparin binding antigen). The fHbp protein used in this multi-component vaccine is a wildtype variant 1 version. Non-functional fHbp molecules are the next stage in broad spectrum meningitis vaccine development. The high affinity interaction between fH and fHbp implies that there is a risk if wildtype fHbp molecules are used for vaccination due to the potential for production of anti-fH antibodies. This risk is null and void when vaccinating with non-functional fHbp molecules.

The work conducted in this project has informed chimeric mouse development, the human-SCR678 model which has since been validated with vaccination studies using non-functional fHbp proteins. This work has also been published (Johnson et al., 2012).

3. Structural Analysis of Meningococcal PilE

T4P are important multifunctional extracellular structures with roles in adhesion, twitching motility, natural transformation, microcolony and biofilm formation. These membrane anchored, long, thin and flexible filaments are composed of thousands of copies of the major pilin protein and are able to retract and extend through dynamic assembly and disassembly from a reservoir of pilin subunits in the inner membrane (Craig et al., 2004).

Many aspects of pilus function are poorly understood including the precise mechanism for extension and retraction of the pilus fibre, the location and frequency of minor pilins within the main pilus structure as well as the role the major pilin plays in performing various pilus-associated functions. Recently however, the major pilin has been increasingly viewed as more than just the structural scaffold for the functional minor pilins and pilus-associated PilC. Hypervariability of the D-region sequence of PilE can have a dramatic effect on the overall function of the T4P and hence the bacterial phenotype (Miller et al., 2014a). The relationship between sequence changes in the D-region of the pilin and overall T4P function is incompletely understood, as small changes of just a few amino acids can have a dramatic impact of T4P function (Nassif et al., 1994), whilst diverse pilin sequences can function very similarly (Cehovin et al., 2010). Another poorly characterised aspect of pilin sequence variation is that class II pilins do not undergo any form of hypervariation. Sequence modulation is supposedly important for immune evasion, yet class II strains represent some of the most pathogenic strains of *N. meningitidis* (Wormann et al., 2014).

Some of the earliest and most thorough work elucidating pilus structure-function relationships, focussed on examining the T4P of *N. gonorrhoeae*. The sequence identity between the major pilin of *N. gonorrhoeae* and *N. meningitidis*, in the C-terminal globular head group, is 85% with most differences found in the D-region. With increased understanding of gonococcal and meningococcal biology and disease processes, extrapolating ideas from one species to the other may not be appropriate given they live in different niches and interact with different epithelial

surfaces. Against this background, this project set out to solve the structure of the major pilin from *N. meningitidis*, in the process, generating a protocol for the expression and purification of large amounts of meningococcal PilE proteins from different strains. Such resources could be used in the functional analysis of this protein, described and discussed later in chapter 3, as well as elucidating further the structure-function relationships of PilE.

<i>N. meningitidis</i> MC58	FTLIELMIVIAIVGILAAVALPAYQDYTARAQVSEAILLAEGQKSAVTEYYLNHGEWPGN
<i>N. gonorrhoeae</i> MS11	FTLIELMIVIAIVGILAAVALPAYQDYTARAQVSEAILLAEGQKSAVTEYYLNHGKWPEN
	*****;***
<i>N. meningitidis</i> MC58	NTSAGVATS-SEIKGKYVKSVEVKNGVVTQAQMASSNVNNEIKGKKLSLWAKRQNGSVKWF
<i>N. gonorrhoeae</i> MS11	NTSAGVASSPTDIKGKYVKEVEVKNGVVTATMLSSGVNNEIKGKKLSLWARRENGSVKWF
	*****:*:*****.***** ** *****;*:*****
<i>N. meningitidis</i> MC58	CGQPVTTRDKAKAANDDVTA AAAANGKKIDTKHLPSTCRDASDAS
<i>N. gonorrhoeae</i> MS11	CGQPVTTRTDD-----DTVADAKDGKEIDTKHLPSTCRDNFDAK
	***** . ,* * :*:***** ** ,

Figure 3.1 Alignment of PilE sequences of *N. gonorrhoeae* MS11 and *N. meningitidis* MC58

The conserved N-terminal α -helix is highlighted with the blue box whilst the D-region is highlighted with the yellow box.

3.1 Expression, Purification and Characterisation of PilE from *N. meningitidis*

Whilst point mutations or extensive sequence differences in the D-region sequence of PilE can affect function, the overall structure of the major pilin is largely conserved. The kinked and protruding hydrophobic N-terminal α -helix is followed by an $\alpha\beta$ structural scaffold, which contains a disulphide bond, the two cysteines of which flank the hypervariable sequence section, the D-region. Previously, full-length structures of the PilE and PAK, from *N. gonorrhoeae* and *Pseudomonas aeruginosa* respectively, were solved by X-ray crystallography using pilin purified directly from native pili of the bacterial species in question (Parge et al., 1995, Craig et al., 2003b). However, given the safety issues of working with large scale cultures of live *N. meningitidis*, PilE was produced recombinantly. The main limitation of this is the lack of post-translational modifications on the protein produced.

The hydrophobic nature of the N-terminal α -helix makes expression in *E. coli* difficult. N-terminal deletion of constructs allows increased levels of expression. Analysis of the structure of the *N. gonorrhoeae* full-length major pilin (Parge et al., 1995, Jokiranta et al., 2006), as well as available N-terminally deleted structures (Helaine et al., 2007), suggested logical deletions to residue 14 (Δ N14), where the kink in the helix occurs, and to residue 24 (Δ N24), where the helix meets the globular head structure, see **Figure 3.2**. For complete analysis, initial constructs were cloned as full-length, Δ N14 or Δ N24 pilin sequences into various expression vectors.

Given the sequence variability of the PilE, it was important that multiple sequences were investigated in this study as well as ensuring that both class I and class II pilin sequences were explored, especially as no class II pilin sequences have been studied previously. No nomenclature system has yet been implemented for describing different PilE sequences, therefore the following scheme was used. For all class I PilE sequences, the PilE allele number would be followed by the name of the strain from which this sequence could be found, as multiple alleles can be found in a single strain. 8013-ST18 and MC58-ST32 are commonly studied strains of *N. meningitidis*

expressing class I pilins. Most analysis of T4P assembly and function has been conducted with 8013. Both PilE sequences were studied in this project and the sequences were designated PilE_6.8013 and PilE_1.MC58 respectively. The remainder of the PilE sequences studied were class II pilins. For all class II PilE sequences, the sequence is designated only with the allele number as single alleles are present in multiple different strains due to the conservation of class II sequences. Alleles PilE_15, PilE_22, PilE_4 and PilE_3 were all studied. Representative strains for these alleles are M03 240910, M00 243158, M01 0240374 and FAM18. M03 240910 is one of the strains responsible for most meningococcal epidemics in sub-Saharan Africa whilst M01 0240374 causes outbreaks globally.

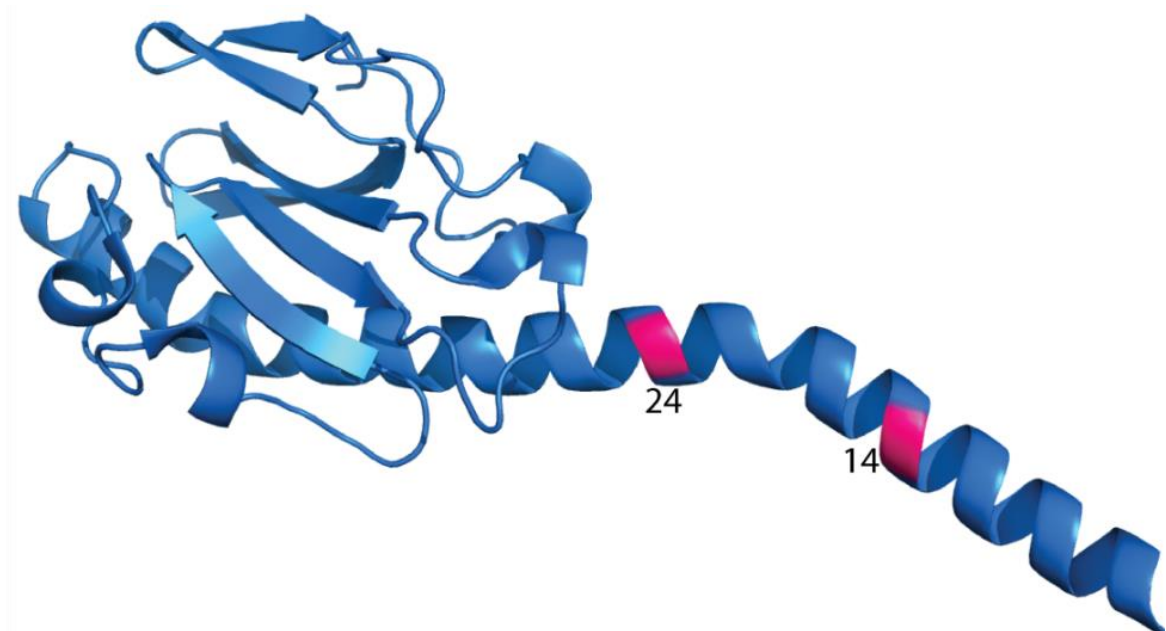


Figure 3.2 Structure of gonococcal PilE (PDB ID: 1AY2) with alternative start residues highlighted
Residues 14 and 24 were chosen as points to which truncations could be made to promote soluble protein expression.

It is important to note that there is more sequence disparity between the class I and class II sequences chosen here than there is between *N. gonorrhoeae* MS11 and *N. meningitidis* MC58, which are 87% identical, see **Figure 3.3**.

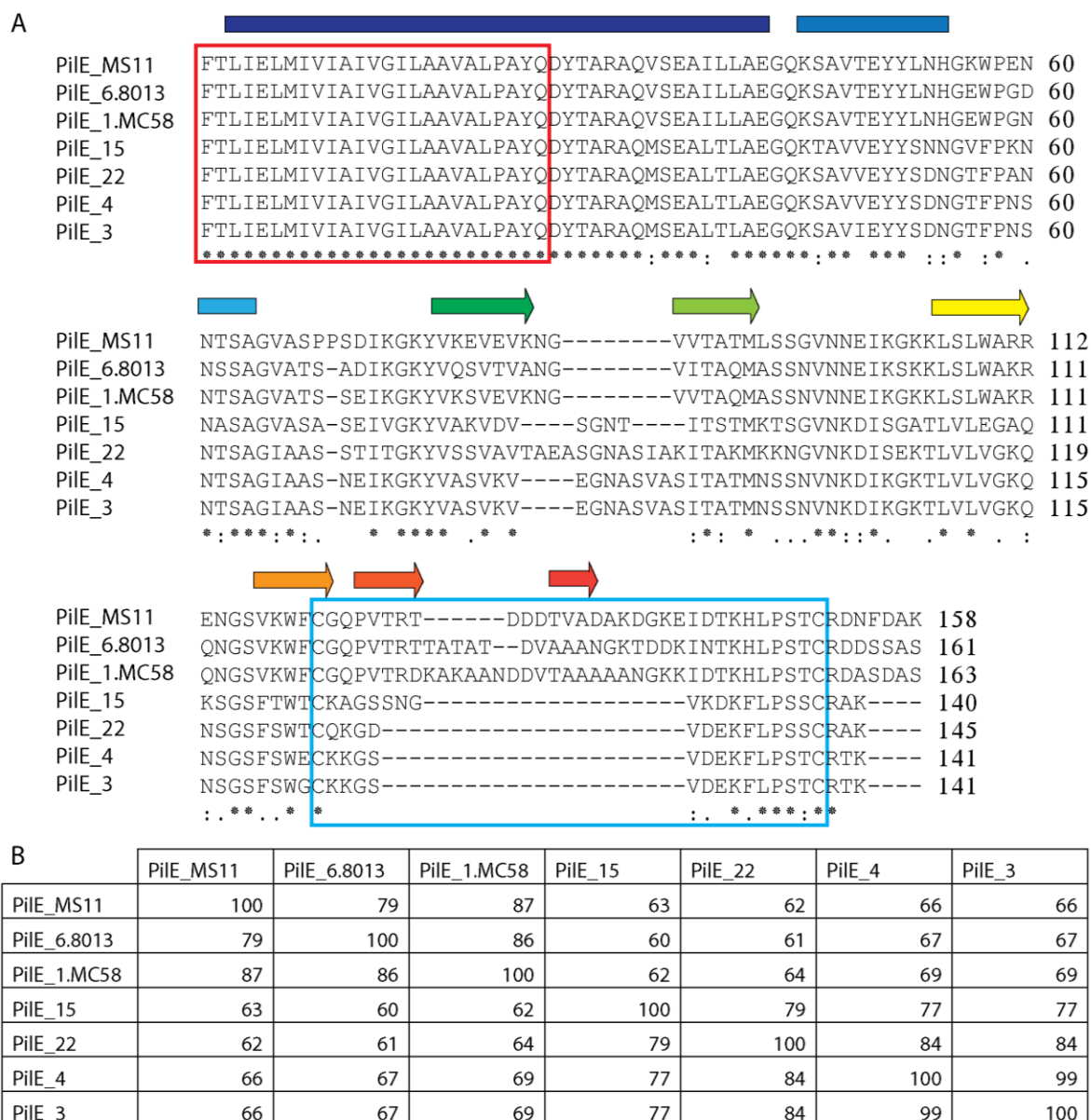


Figure 3.3 Alignment of full-length PiIE sequences of strains to be investigated

Panel A shows the alignment of the full-length sequences of the pilins chosen for this study as well as the *N. gonorrhoeae* MS11 sequence. Identical residues are indicated with a •, a : (colon) indicates conservation between side chains with strongly similar properties, a . (period) indicates conservation between side chains with weakly similar properties. The red box highlights the conserved N-terminal 24 amino acids which form the hydrophobic α -helix. The blue box highlights the D-region of the protein sequence. The secondary structure of PiIE is shown above the sequence alignment, as seen in the *N. gonorrhoeae* MS11 structure (PDB ID: 1AY2). α -helices are shown as blocks whilst β -strands are shown as arrows. Panel B shows the percentage sequence identity of each pilin sequence with each other.

3.1.1 Construction of PilE Expression Vectors

Pilin is a challenging protein to produce recombinantly in *E. coli*. It contains a disulphide bond, self associates and is very insoluble. Previous published methods of producing pilin used an expression system which exported an MBP-tagged version of the protein into the periplasm from where it was then purified using affinity chromatography and size-exclusion chromatography (Helaine et al., 2007, Berry et al., 2013). However, this project aimed to generate a simpler method of protein purification to allow rapid expression and purification of large quantities of meningococcal major pilin, so a new approach was taken.

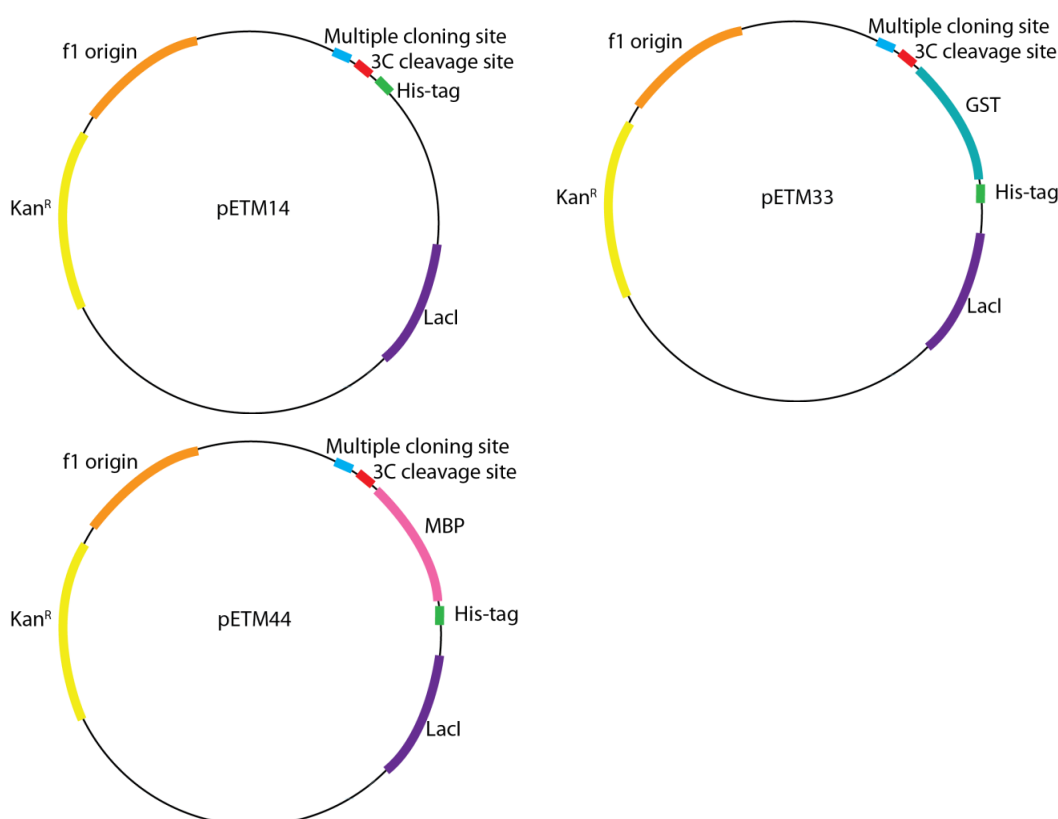


Figure 3.4 pETM vectors used for expression of PilE

This series of vectors allows practical high-throughput cloning due to the fact the vectors only differ in the N-terminal tag and have a common multiple cloning site.

Constructs were produced which were full-length, Δ N14 or Δ N24 with either an N-terminal hexa-his-tag, hexa-his-tag with a GST tag or hexa-his tag with an MBP tag. It was envisaged that MBP and GST tags could improve solubility of the protein whilst providing an additional means of

affinity chromatography for protein purification. Expression vectors from the pETM series were used to facilitate high-throughput cloning methodology, as they share a common multiple cloning site.

Full length pilin sequences were generated through PCR amplification from genomic DNA (provided by Dr Rachel Exley, University of Oxford) of *N. meningitidis* from strains MC58, M03 240910 and M00 243158 for PilE sequence PilE_1.MC58, PilE_15 and PilE_22 respectively using the primers detailed in **Table 3.1**.

Primer Name	Sequence	Tm (°C)
PilE_1.MC58 Forward	5'GGGCCCATGGGT TTTACCCTTATCGAGCTGATGATTGTGATTGCCATC GTCGGC -3'	63
PilE_1.MC58 Reverse	5'GGTGCTCGAG TTAGCTGGCATCACTGCGTCGCGGCAGGTTGAC -3'	64
PilE_15 Forward	5'GGGCCCATGGGT TTCACTCTGATCGAGCTGATGATCGTCATCGCCATC GTCGG -3'	65
PilE_15 Reverse	5'GGTGCTCGAG TTATTTGGCGCGGCAGGAAGACGGCAGGAATTTGTC -3'	63
PilE_22 Forward	5'GGGCCCATGGGT TTACCCTGATCGAGCTGATGATCGTCATCGCCAT CGTCGG -3'	66
PilE_22 Reverse	5'GGTGCTCGAG TTATTTGGCGCGGCAGGAAGACGGCAGGAATTTTC G -3'	63

Table 3.1 Primers used to amplify the full-length pilE sequences from meningococcal genomic DNA using PCR

The sequence of the primer complementary to the pilin gene is highlighted in blue.

Amplicons were generated using the PCR protocol described in **7.1**. The PCR was successful in producing a single species of the expected length by analysis, see **Figure 3.5**. 1 µg of each PCR product and 3 µg of each pETM vector were both separately double-digested by incubation with *Nco*I FastDigest, *Xho*I FastDigest (both enzymes present at 1 U per 1µg plasmid) and 1X FastDigest Buffer (all Fermentas) overnight at 37°C, see **Figure 3.6**. Linearised pETM vectors were treated with Fast Alkaline Phosphatase (Fermentas) and purified with the QIAquick PCR Purification Kit (Qiagen) though using a Qiagen Miniprep cartridge. Linearised vectors were eluted with 50 µl EB at 79.6 ng/µl, 104.7 ng/µl and 39.8 ng/µl for pETM-14, pETM-33 and pETM-44 respectively. The

fragments corresponding to the *pilE* genes were treated with *DpnI* and *ExoI* and then further purified using the PCR Purification kit (Qiagen) and eluted with 25µl buffer EB at 22.6 ng/µl, 23.5 ng/µl and 22.9 ng/µl for the *PilE_1.MC58*, *PilE_15* and *PilE_22* fragments respectively. Each combination of vector and insert was mixed at a 1:3 molar ratio, ligated using the Ligafast kit (Promega) and transformed into DH5α *E. coli* competent cells. Presence of the correct insert in purified pETM-*pilE* constructs was verified by sequencing with the T7F and T7R primers (Source Bioscience).

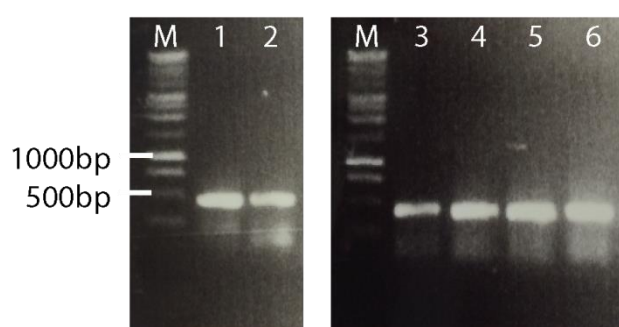


Figure 3.5 *pilE* genes successfully amplified as shown by analysis on 1% agarose TAE gel

Lanes M show the DNA ladder (1kb ladder - Promega), lane 1 and 2 are *PilE_1.MC58* amplicon, lanes 3 and 4 are *PilE_15* amplicon, lanes 5 and 6 are *PilE_22* amplicon. Lanes 1, 3 and 5 show PCR product from reactions run with 1.5mM MgCl₂, whilst lanes 2, 4 and 6 show PCR product from reactions run with 3mM MgCl₂.

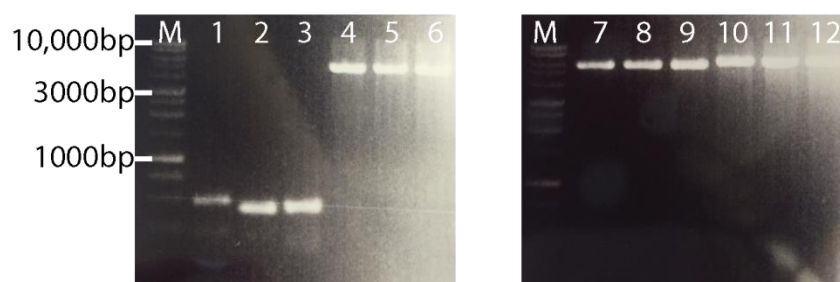


Figure 3.6 Digestion of the *pilE* inserts and pETM vectors with *NcoI* and *XhoI*

All three vectors and three inserts were digested with both *NcoI* and *XhoI*. M corresponds to markers. Lanes 1, 2 and 3 show *pilE* *PilE_1.MC58*, *PilE_15* and *PilE_22* fragments digested with both enzymes. Lanes 4, 5 and 6 show digested pETM-14; 7, 8 and 9 digested pETM-33; 10, 11 and 12 digested pETM-44. Lanes 4, 7 and 10 are digested only with *NcoI*; lanes 5, 8 and 11 just *XhoI*; lanes 6, 9 and 12 with both enzymes.

The full-length constructs were N-terminally truncated by 14 or 25 amino acids, using the Quikchange mutagenesis kit (Qiagen) and the primers described in **Table 3.2**. Purified pETM-*pile* deletion constructs were checked using T7F and T7R sequencing (Source Bioscience).

Primer name	Sequence	Tm (°C)
PiE_1.MC58 ΔN14 Forward	5'CTGTTCCAGGGGCCCCATGATTTTGGCGGCAGTCGCCCTTCC-3'	70
PiE_1.MC58 ΔN14 Reverse	5'GGAAGGGCGACTGCCGCCAAAATCATGGGCCCTGGAACAG-3'	70
PiE_1.MC58 ΔN24 Forward	5'CTGTTCCAGGGGCCCCATGCAAGACTACACAGCCC GCGCAC-3'	72
PiE_1.MC58 ΔN24 Reverse	5'GTGCGCGGGCTGTGTAGTCTTGCATGGGCCCTGGAACAG-3'	72
PiE_15 ΔN14 Forward	5'GTTCTGTTCCAGGGGCCCCATGATCTTGGCAGCCGTCGCCCTG-3'	71
PiE_15 ΔN14 Reverse	5'CAGGGCGACGGCTGCCAAGATCATGGGCCCTGGAACAGAAC-3'	71
PiE_15 ΔN24 Forward	5'GTTCTGTTCCAGGGGCCCCATGGACTACACCGCGCGTGCC-3'	71
PiE_15 ΔN24 Reverse	5'GGCACGCGCGGTGTAGTCCATGGGCCCTGGAACAGAAC-3'	71
PiE_22 ΔN14 Forward	5'GTTCTGTTCCAGGGGCCCCATGATCTTGGCAGCCGTCGCCCTG-3'	72
PiE_22 ΔN14 Reverse	5'CAGGGCGACGGCTGCCAAGATCATGGGCCCTGGAACAGAAC-3'	72
PiE_22 ΔN24 Forward	5'GTTCTGTTCCAGGGGCCCCATGGACTACACCGCGCGCGC-3'	70
PiE_22 ΔN24 Reverse	5'GCGCGCGCGGTGTAGTCCATGGGCCCTGGAACAGAAC-3'	70

Table 3.2 Primers used to create N-terminal deletion constructs of the pETM-*pile* vectors

Primers Name	Sequence	Tm (°C)
PiE_4 E130G Forward	5'CCGGCTCGTTCTCTTGGGGATGTAAAAAAGTTCTGTG-3'	82
PiE_4 E130G Reverse	5'CACAGAACCTTTTTTACATCCCCAAGAGAACGAGCCGG-3'	82
PiE_4 V55I Forward	5'GGTCAAAAATCCGCAGTGATCGAGTATTATCCGACAAC-3'	79
PiE_4 V55I Reverse	5'GTTGTCGGAATAATACTCGATCACTGCGGATTTTGTACC-3'	79

Table 3.3 Primers to introduce the E130G and V55I mutations into PiE_4 constructs to give the sequence for PiE_3

Equivalent constructs were made by collaborators (Dr Rachel Exley, Felicia Tan and Mikaila Bendorra, all University of Oxford) for the PiE_4 and PiE_6.8013 sequences. Next, the PiE_4

constructs were mutated by Quikchange mutagenesis to match the sequence for Pile_3, just two amino acids different (E130G and V55I) using the primers shown in **Table 3.3**.

3.1.2 Optimising Expression of the pETM-*pile* Constructs

High-throughput expression trials were conducted, to determine the optimal conditions for expression of these constructs. The conditions shown in **Table 3.4** were investigated for all constructs.

Variable	Conditions Tested
<i>E. coli</i> strain	BL21 (DE3) of B834 (DE3)
Concentration of IPTG for induction upon reaching a cell density where OD ₆₀₀ ~ 0.6	1 mM or 0.5 mM
Growth medium	TB or LB
Conditions of growth after induction with IPTG	3 hours at 37 °C or overnight at 21 °C

Table 3.4 Conditions screened for optimal soluble expression in *E. coli* expression strains

The total number of different condition combinations was 16 per construct tested.

5ml media were inoculated with 50µl overnight culture and grown to OD₆₀₀ ~ 0.6 before induction with IPTG. After the cultures had grown at their designated time and temperature, the cells were harvested by centrifugation. Cell pellets were resuspended in 210 µl PBS supplemented with 1 % v/v Tween-20, 1 mg/ml lysozyme (Sigma), 400 U/µl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) and then placed on an orbital shaker at 1000 rpm until lysed. After clarification of the lysate, Ni-affinity chromatography was performed using NiNTA beads (Qiagen) which were incubated with clarified lysate at room temperature for 30 minutes. The beads were washed using PBS with 20 mM imidazole and then the protein was eluted with 50 µl PBS with 250 mM imidazole. To assess expression, two samples for each condition were generated by mixing 5 µl cell lysate with 25 µl 3X SDS-PAGE loading dye and 45 µl ddH₂O and also by adding 25 µl 3X

SDS-PAGE loading dye to the total elution of the NiNTA beads. After boiling for 5 minutes, 30 μ l of each sample was loaded on 15 % acrylamide gels for SDS-PAGE analysis.

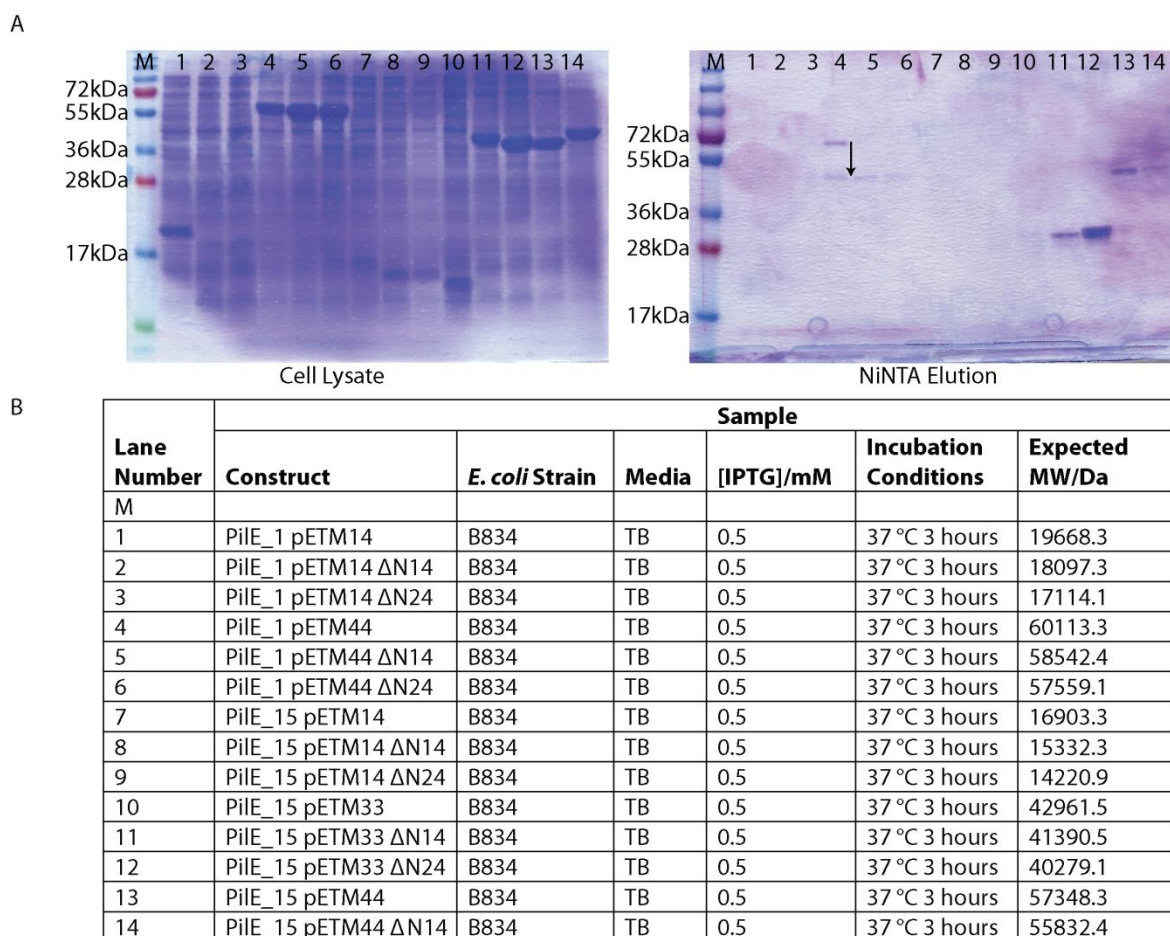


Figure 3.7 Example SDS-PAGE analysis of PilE expression trials

Panel A shows acrylamide gels for the cell lysate and NiNTA elution of the same conditions described in the table of panel B. In most cases the dominant band seen on the gel is the correct size for the overexpressed PilE with its associated tag for the gel showing the cell lysate. However, the bands seen for the NiNTA elution of each condition generally correspond to the tag alone. The arrow in lane 4 shows the degradation of the MBP-tagged construct PilE_1 pETM44 from the high molecular weight tagged protein to just the MBP tag.

In nearly all conditions investigated, SDS-PAGE analysis of cell lysate samples showed over-expressed pilin as large bands on the gels, all of which are the size predicted from the sequence of the constructs. However, when compared to the equivalent analysis of samples taken from the NiNTA elution, very few of the conditions and variables trialled, enabled extraction of soluble his-

tagged fusion proteins. Many protein bands seen, correspond to the correct molecular weight of his-tagged MBP or GST, suggesting that the pilin protein was cleaved and then degraded to leave just the tag. In cases where protein was present in whole cell lysates but not in the soluble fraction, it was likely that the protein is expressed largely as inclusion bodies and hence could not be extracted from the soluble fraction of the cell lysate. See **Figure 3.7**.

3.1.3 Large Scale Expression and Purification of PilE

The only proteins of the correct expected molecular weight were the MBP fusion proteins, both full-length and truncated. This suggested that the extremely soluble MBP successfully draws the pilin proteins into solution therefore allowing soluble overexpression, irrespective of the truncation at the N-terminus. However, large scale expression and purification of these constructs revealed that the pilin protein, once cleaved with 3C protease to remove the MBP tag, was rendered insoluble and unfolded. Therefore, large scale expression was performed using “gentle” expression conditions of the simple pETM-14 *pilE* Δ N24 constructs as this was the approach most likely to generate material for subsequent analyses.

Overnight cultures of *E.coli* BL21 (DE3) transformed with the pETM-14 *pilE* Δ N24 were used to inoculate 6 x 1l LB each in a dilution of 1:100. Cultures were grown at 37°C until OD₆₀₀ reached approximately 0.6. Cultures were then induced with 0.5mM IPTG and grown overnight at 21°C. After cells were harvested by centrifugation, pellets were resuspended in 50 mM Tris pH7.5, 150 mM NaCl, 50 mM imidazole supplemented with 1 mg/ml lysozyme (Sigma), 400 U/ μ l DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) before cell disruption by homogenisation at 15,000 psi (Emulsiflex C.5 Avestin). The lysate was clarified by centrifugation at 26,000 $\times g$ for 20 minutes before applying the supernatant to a 1 ml NiNTA Superflow Cartridge (Qiagen) at 1 ml/minute, pre-equilibrated in resuspension buffer. After a 5 CV wash with resuspension buffer, the protein was eluted in 50 mM Tris pH7.5, 150 mM NaCl, 500 mM imidazole over 8 fractions of 1 ml each. Analysis of the eluted fractions by measuring A₂₈₀ using a

Nanodrop ND-1000 (Labtech) to determine protein concentration, showed fractions 3, 4 and 5 to contain the most protein. 1 mg 3C protease was added to these pooled fractions and digestion of the N-terminal tag was performed over night at 4°C in dialysis against 50 mM Tris pH7.5, 150 mM NaCl using snakeskin MWCO 10,000 (Pierce). The dialyte was then subject to reverse Ni-affinity chromatography with a 1 ml NiNTA Superflow Cartridge in tandem with a 1 ml GSTrap HP (GE Life Sciences Health Care). This step removed any Ni-affinity chromatography contaminants, common with protein preparations of low yield, as well as the cleaved hexa-his-tag, uncleaved pilin protein and 3C protease. The flow-through from this step was then subject to a second round of dialysis overnight at 4°C, again using snakeskin MWCO 10,000 but this time against a buffer of 50 mM Tris pH9.0, 150 mM NaCl, 4 mM cystine, 2 mM cysteine. This form of “refold dialysis” promotes disulphide formation in simple cases, such as this, where there is only one disulphide bond. The dialyte was concentrated to 0.5 ml using an Amicon Ultra-4 Centrifugal Filter Unit NMWL of 10,000 Da. The concentrated protein was then subject to size-exclusion chromatography using an Akta Purifier (GE Healthcare Life Sciences) in a buffer of 50 mM Tris pH7.5, 150 mM NaCl. The sample was run on an S75 10/300 GL column at 1 ml/minute. Fractions of the peak corresponding to purified PiIE ΔN24 were pooled and concentrated to 1 ml.

All six PiIE ΔN24 proteins were expressed and purified in this manner with varying yields. PiIE_1.MC58 ΔN24 and PiIE_6.8013 ΔN24 gave the highest yields of around 1.5-2.5 mg protein per 6l culture, whilst PiIE_4 ΔN24, PiIE_15 ΔN24 and PiIE_22 ΔN24 gave just 0.2-0.8 mg protein per 6 l culture grown. Additionally, the proteins which were produced at lower yield, were more inconsistent in yield from one preparation to the next. Interestingly, PiIE_3 ΔN24 could be produced consistently with at least twice the yield as that of PiIE_4 ΔN24, approximately 1.5 mg per 6 l, despite only being two amino acids different. Given the generally close sequence identity of all of the pilin proteins made, it was very interesting that such limited changes in sequence have such a pronounced effect on yield. Details of all constructs expressed can be seen in **Table**

3.5.

Construct Name	Amino Acid Sequence of Open Reading Frame	Molecular Weight (Da)	Theoretical pI
pETM-14 PiIE_6.8013 PiIE ΔN24	GPMDY TARAQVSEAILLAEGQKSAVTEYYLNHGEW PGDNSSAGVATSADIKGKYVQSVTVANGVITAQMAS SNVNNEIKSKKLSLWAKRQNGSVKWFCGQPVTRTTA TATDVAAANGKTDDKINTKHLPTCRDDSSAS	14710.3	8.57
pETM-14 PiIE_1.MC58 PiIE ΔN24	GPMDY TARAQVSEAILLAEGQKSAVTEYYLNHGEW PGNNTSAGVATSSEIKGKYVKSVEVKNGVVTAQMASS NVNNEIKGKKLSLWAKRQNGSVKWFCGQPVTRDKA KAANDDVTA AAAANGKKIDTKHLPTCRDASDAS	15086.8	9.23
pETM-14 PiIE_15 PiIE ΔN24	GPMDY TARAQMSEALTLAEGQKTAVVEYYSNNGV FPKNNASAGVASASEIVGKYVAKVDVSGNTITSTMKTS GVNKDISGATLVLEGAQKSGSFTWTCKAGSSNGVKDK FLPSSCRAK	12193.6	9.15
pETM-14 PiIE_22 PiIE ΔN24	GPMDY TARAQMSEALTLAEGQKSAVVEYYSDNGTF PANNTSAGIAASSTITGKYVSSVAVTAEASGNASIAKIT AKMKKNGVKNKDISEKTLVLVGKQNSGSFSWTCQKGD VDEKFLPSSCRAK	12837.3	8.89
pETM-14 PiIE_4 PiIE ΔN24	GPMDY TARAQMSEALTLAEGQKSAVVEYYSDNGTF PNSNTSAGIAASNEIKGKYVASVKVEGNASVASITATM NSSNVNKDIKGKTLVLVGKQNSGSFSWECKKGSVDEK FLPSTCRCK	12552.0	8.89
pETM-14 PiIE_3 PiIE ΔN24	GPMDY TARAQMSEALTLAEGQKSAVIEYYSDNGTF PNSNTSAGIAASNEIKGKYVASVKVEGNASVASITATM NSSNVNKDIKGKTLVLVGKQNSGSFSWGCKKGSVDEK FLPSTCRCK	12494.0	9.12

Table 3.5 Expressed sequences of pETM-14 PiIE ΔN24 constructs

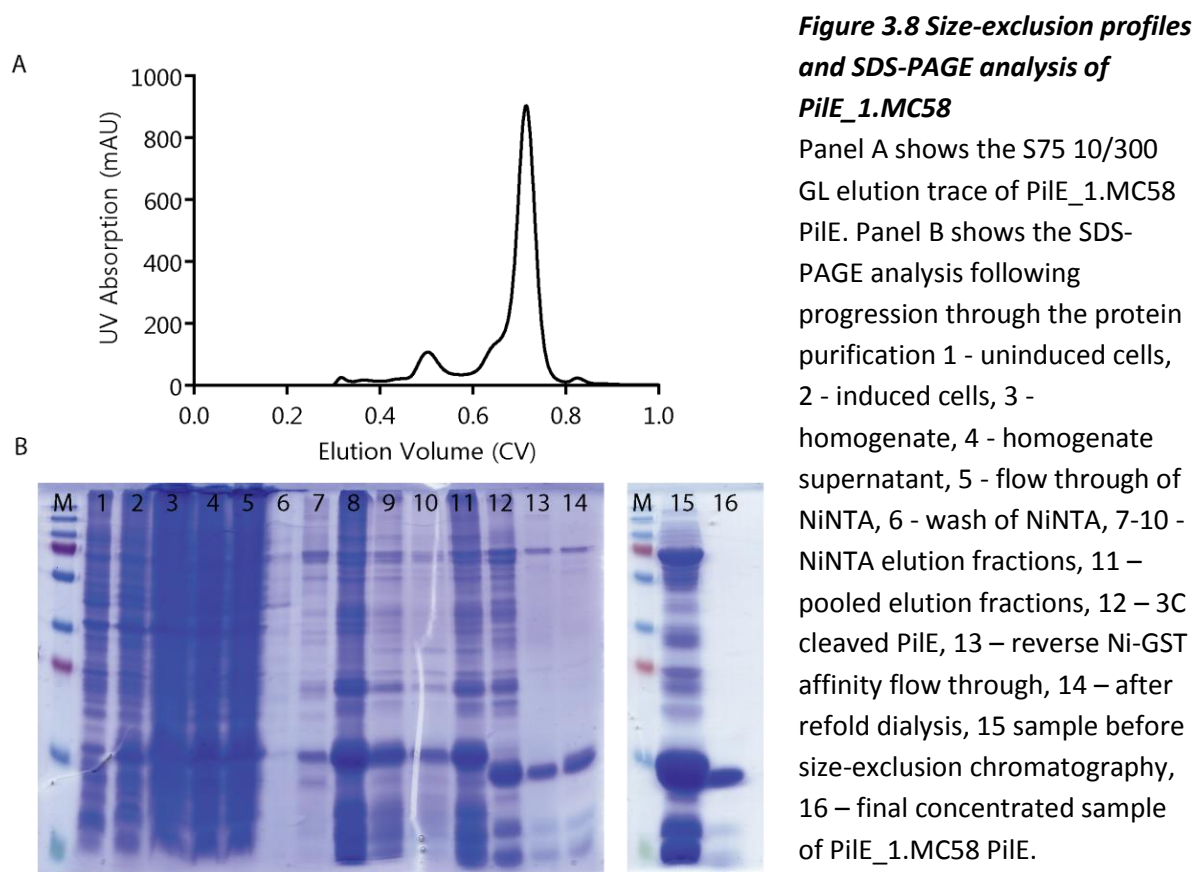
This table details the names of each of the constructs, the sequence of the protein after 3C cleavage, the molecular weight of the protein product as well as its pI, both calculated by ExPASy ProtParam. The additional sequence of the N-terminal tag, derived from the pETM-14 vector, is highlighted in blue.

PiIE_1.MC58 ΔN24 gave the highest and most consistent yield of protein. The size-exclusion elution profile of the final sample and SDS-PAGE analysis of such a preparation can be seen in

Figure 3.8.

As with all pilin proteins produced, the PiIE_1.MC58 ΔN24 sample was analysed by mass spectrometry, performed by Dr David Staunton, University of Oxford, which showed it to be of the correct expected mass for the 3C cleaved version of this construct, see **Figure 3.9A**. A 0.1 mg/ml sample of this protein in 10 mM HEPES pH7.4, 20 mM NaCl at room temperature was analysed by circular dichroism to assess the secondary structure of the protein, the results of which can be

seen in **Figure 3.9B**. The CD spectra was analysed by K2D2 (Perez-Iratxeta and Andrade-Navarro, 2008), and revealed the secondary structure prediction to be 48 % α -helical and 8 % β -strand. However, the equivalent *N. gonorrhoeae* structure (when analysing just the head group) has 30 % α -helical and 24 % β -strand (Parge et al., 1995). This disparity suggested that the CD analysis by K2D2 was unreliable.



Differential scanning calorimetry (DSC) was performed by Dr David Staunton, University of Oxford, to assess the stability of *PilE_1.MC58* Δ N24 in 50 mM Tris-HCl pH7.5, 150 mM NaCl, see **Figure 3.9C**. This technique measures the excess specific heat capacity (C_p) as a function of increasing temperature. A sharp endothermic peak occurs at the T_m of the protein in solution. The low melting temperature calculated for the *PilE_1.MC58* Δ N24 by DSC analysis was not surprising. Assuming a similar content of secondary structure as the *N. gonorrhoeae* structure, the predicted melting temperature of the protein should be less than 50 °C. The calculated melting

temperature is 35 °C for PiIE_1.MC58 Δ N24. A low melting temperature for a protein which is present physiologically as a compact multimer, such as PiIE in the T4P helical fibre, is not unreasonable, as *in vivo* it would be expected that subunit-subunit interactions would contribute to the stability of the individual subunits.

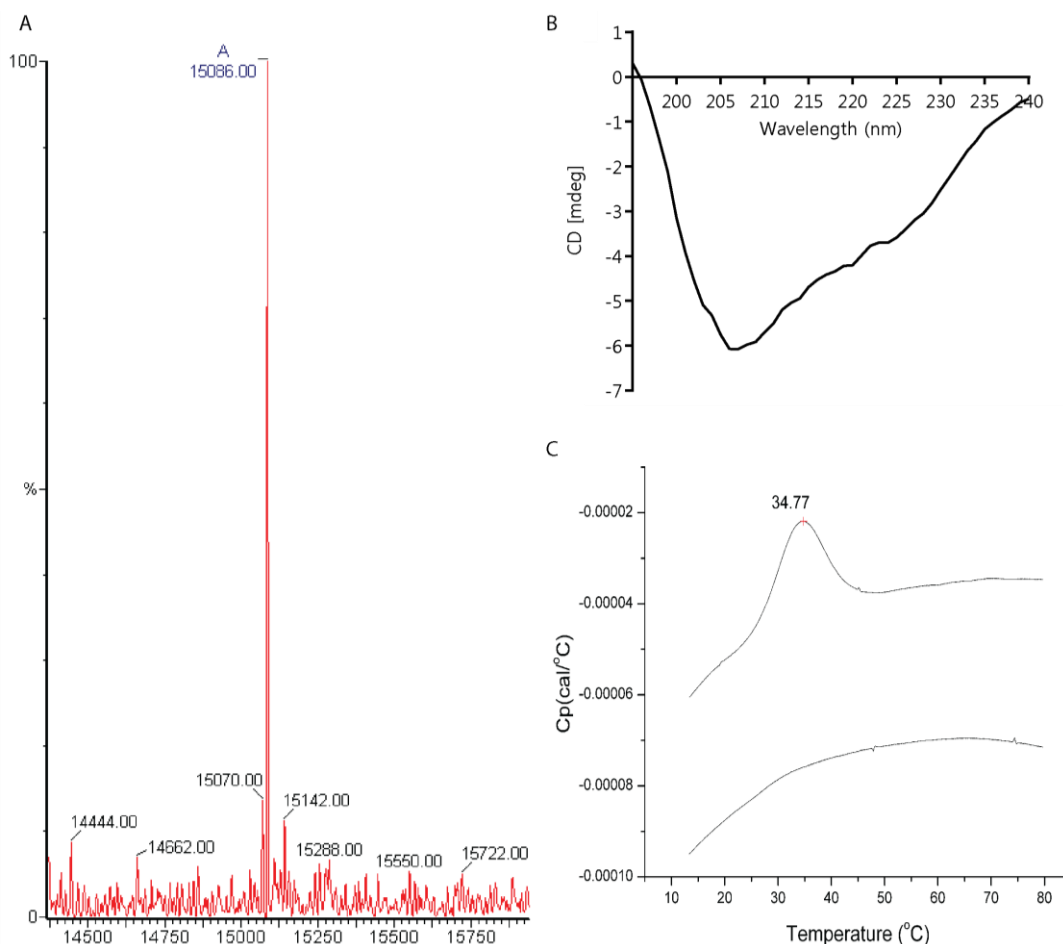


Figure 3.9 Biophysical analysis of the pETM-14 PiIE_1.MC58 PiIE Δ N24 protein sample

Panel A shows the mass spectra of the protein sample which is only 0.6 Da different from the calculated mass (**Table 3.6**). Panel B shows the CD spectra for this sample. Panel C shows the DSC protein melting curve and the calculated melting temperature of the sample.

Given the purity, homogeneity and stability of the PiIE_1.MC58 Δ N24 sample produced, this protein was used in subsequent crystallisation trials.

3.2 Structural Analysis of PilE

3.2.1 Crystallisation of PilE

Whilst the expression levels of many different PilE Δ N24 proteins varied a lot between strains and yields were unreliable from one purification to the next, PilE_1.MC58 Δ N24 was produced consistently at a high yield, was homogenous and therefore was considered a good candidate for structure determination by X-ray crystallography. Purified protein was concentrated to 10 mg/ml or 15.1 mg/ml using Amicon Ultra-15 and -4 Centrifugal Filter Units NMWL of 10,000Da and used in high-throughput sparse-matrix crystallisation screen experiments.

Crystallisation trials were set up as described in 2.2.1. Drops contained 0.2 μ l reservoir solution and 0.2 μ l protein solution (10 or 15.1mg/ml MC58 ST-32 PilE Δ N24, 50 mM Tris, 150 mM NaCl, pH 7.5) (50% dilution), or 0.3 μ l reservoir solution and 0.1 μ l protein solution (25% dilution). Plates were then sealed and incubated at 21°C. Six screens were set up at these two protein concentrations using these conditions: Structure screen I + II, PACT *premier*, JCSG-*plus*, Proplex™, The Stura Footprint Combination and Morpheus® (all Molecular Dimensions).

Crystals were found in the 10 mg/ml trays. The largest and most regular-shaped crystals were produced in two almost identical conditions, consisting of HEPES buffer and a high concentration of a low molecular weight PEG. Many small shard-like crystals grew in the drops and took over six months to form. The conditions are described in **Table 3.6** and the crystals are shown in **Figure 3.10**.

Crystals were mounted using CrystalCap nylon Cryoloops (Hampton Research) and then cryo-protected in a 1:10 ratio of ethylene glycol: mother liquor before flash cooling in liquid nitrogen. This strategy was adapted from the condition optimised cryoprotectant previously determined (McFerrin and Snell, 2002).

A and C – Stura Macrosol C2 (50% and 25% dilution respectively)	0.1M Na-HEPES pH8.2, 40% v/v PEG550 MME
B and D – Stura Macrosol C8 (50% and 25% dilution respectively)	0.1M Na-HEPES pH7.5, 45% v/v PEG600

Table 3.6 Conditions in which MC58 ST-32 Pile Δ N24 crystals grew

A, B, C and D refer to the images in **Figure 3.10**

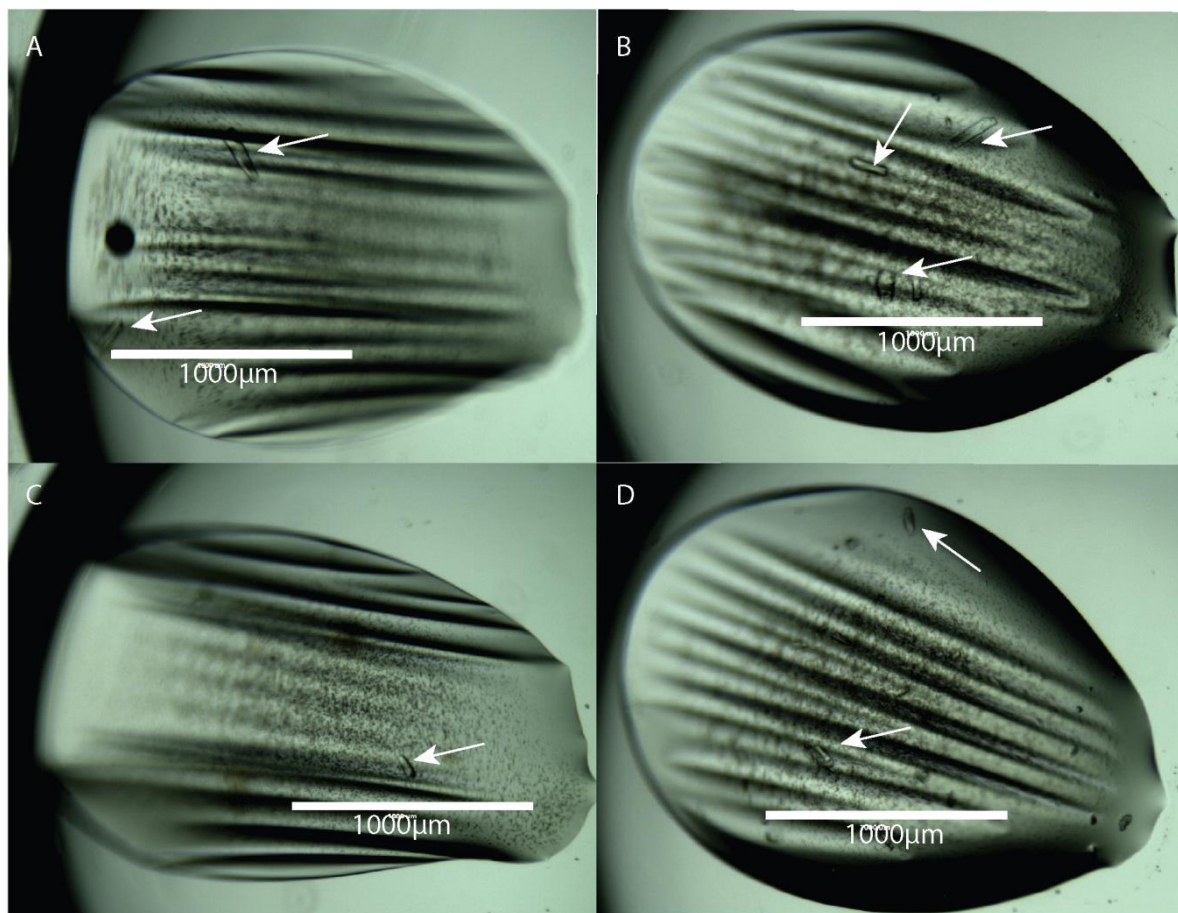


Figure 3.10 Crystals grown of MC58 ST-32 Pile Δ N24

3.2.2 Data Collection and Processing

Data collection was performed by Professor Susan Lea and myself at beamline I04 of the Diamond Light Source, Harwell, U.K. Diffraction data sets were collected at 100K to minimise radiation damage to the sample (Garman and McSweeney, 2007). The crystal was rotated whilst in a 0.984

Å wavelength X-ray beam through 120° along a single axis using a goniometer. With an oscillation angle of 0.2 °, 600 diffraction images were collected constituting a full data set.

After data collection, diffraction images were processed. Images were indexed and integrated in Xia2 (Winter, 2010) with XDS (Kabsch, 2010) and then scaled and merged using SCALA (Evans, 2006), both from the CCP4 suite (Winn et al., 2011b). The crystal gave a diffraction pattern compatible with a primitive hexagonal lattice. Processing determined the spacegroup to be P6. Analysis using MATTHEWS COEF (Kantardjieff and Rupp, 2003) in CCP4 showed there was only one protein molecule in the asymmetric unit. Processing statistics are given in **Table 3.7**.

Diffraction source	I04 - Diamond Light Source (Harwell, UK)
Wavelength	0.984Å
Detector	ADSC CCD
Temperature (K)	100
Space group	P6
Z	6
Unit cell parameters (Å)	a = 105.06, b = 105.06, c = 23.06
Solvent content (%)	56.5
Resolution (Å)	23.06-1.47 (1.43-1.43)
No. unique reflections (criteria for reflection)	27504 (1976)
No. observed reflections	187917 (8455)
R_{merge}	0.049 (0.595)
R_{pim}	0.022(0.358)
[I/σ(I)]	19.3 (2.5)
Completeness (%)	99.9 (99.9)
Multiplicity	6.8 (4.3)
CC1/2	0.99 (0.72)
Matthews coefficient V_M (Å³ Da⁻¹)	2.83
Data processing software	Xia2, XDS and SCALA

Table 3.7 - Data collection and structure solution statistics

Values for the outer shell are given in brackets.

The data quality statistics over the whole data set were favourable. Correspondingly, the data set was kept in its entirety as this gave an almost complete set of reflections (99.9%). Given the I_{hkl}/σ_{hkl} showed that there was measureable signal up to 1.43Å, this was used as the highest

resolution shell. The R_{merge} and R_{pim} are both low in the outer shell suggesting the data are of good quality.

3.2.3 Phase Determination, Refinement and Structure Validation

The phase problem must be solved in order to create an electron density map of PiE_1.MC58 Δ N24. PiE_1.MC58 Δ N24 shares 83% sequence identity with the equivalent sequence from *N. gonorrhoeae* for which there are published structures. The highest resolution structure (2.3Å compared to 2.6Å from the original 1995 publication) was used as a model for molecular replacement (PDB ID: 2HI2) (Parge et al., 1995, Craig et al., 2006b). PHASER (McCoy et al., 2007), within the CCP4 software suite, was used to perform the molecular replacement.

Using CHAINSAW (Stein, 2008), a model of gonococcal MS11 PiE, which had been truncated by the first 24 residues, was generated in which the side chains had been pruned back to the most common atom of each amino acid on the basis of the sequence alignment of MS11 PiE Δ N24 with PiE_1.MC58 Δ N24. Molecular replacement using PHASER with this model generated just one solution with acceptable scores, indicating that this was likely the correct solution.

Search model	No. Solutions	LLG	RFZ	TFZ	Number of clashes <3Å
MS11 PiE Δ N24 CHAINSAW	1	2435	26.3	55.4	0

Table 3.8 Molecular replacement solution for MC58 ST-32 PiE Δ N24 from PHASER

By analysing the model of PiE_1.MC58 Δ N24 constructed by PHASER within the electron density map using Coot (Emsley and Cowtan, 2004), it was evident that the structure solution was correct. No solvent flattening or density modification was required before model building as the initial R factor and R free were 0.379 and 0.385 respectively.

3.2.4 Model Building and Refinement

ARP/wARP (Perrakis et al., 2001) was used to build in the side chains of the initial model. Analysis of the difference map in Coot, allowed modification of the model of PiE_1. MC58 Δ N24 further to

and created a better model to fit the electron density. As the phases used to calculate F_o were taken from the PiLE_1.MC58 Δ N24 model, a $2F_o - F_c$ map was visualised in order to emphasise the experimental data and reduce model bias. Model adjustment was completed in Coot whilst refinement was completed using Refmac (Winn et al., 2003). The final structure was validated within Coot and online using Molprobit (Chen et al., 2010) to assess the chemistry of the final model. Refinement and validation statistics can be seen in **Table 3.9**.

Resolution (Å)	1.43-23.06 (1.43-1.47)
Number of reflections	26116 (1723)
Number of reflections for Rfree	1383 (121)
R/Rfree	0.1402/0.1759 (0.198/0.252)
No. atoms – Protein	2202
No. atoms – Heterogens	1 – chlorine
No. atoms – Water	102
RMS deviation bond lengths (Å)	0.013
RMS deviation bond angles (°)	0.859
Residues in Ramachandran plot favoured regions	140
Residues in Ramachandran plot unfavoured regions	2
Molprobit Score	1.15
Molprobit Score Percentile	98 th

Table 3.9 – Refinement statistics

3.2.5 The Structure of *N. meningitidis* PiLE

The PiLE_1.MC58 Δ N24 model showed a classic pilin fold. The hydrophobic N-terminal α -helix acts as a scaffold for the rest of the structure. This is followed by an extended loop with limited features of secondary structure which, *in vivo* would have a disaccharide or trisaccharide group bound (Aas et al., 2006a, Schoen et al., 2008). Next, two β -hairpins form a four-stranded antiparallel β -sheet, the last strand of which contains the first cysteine of the disulphide bond. A long extended loop region comprises the D-region which is then completed by the presence of the second cysteine of the disulphide bond. Part of this C-terminal extended loop is not built in the structure model due to poor electron density in this region. Instead, lysine-130 has been modelled

only to the β -carbon in the side chain as no density was resolved for the side chain and the adjacent alanine-131 is missing.

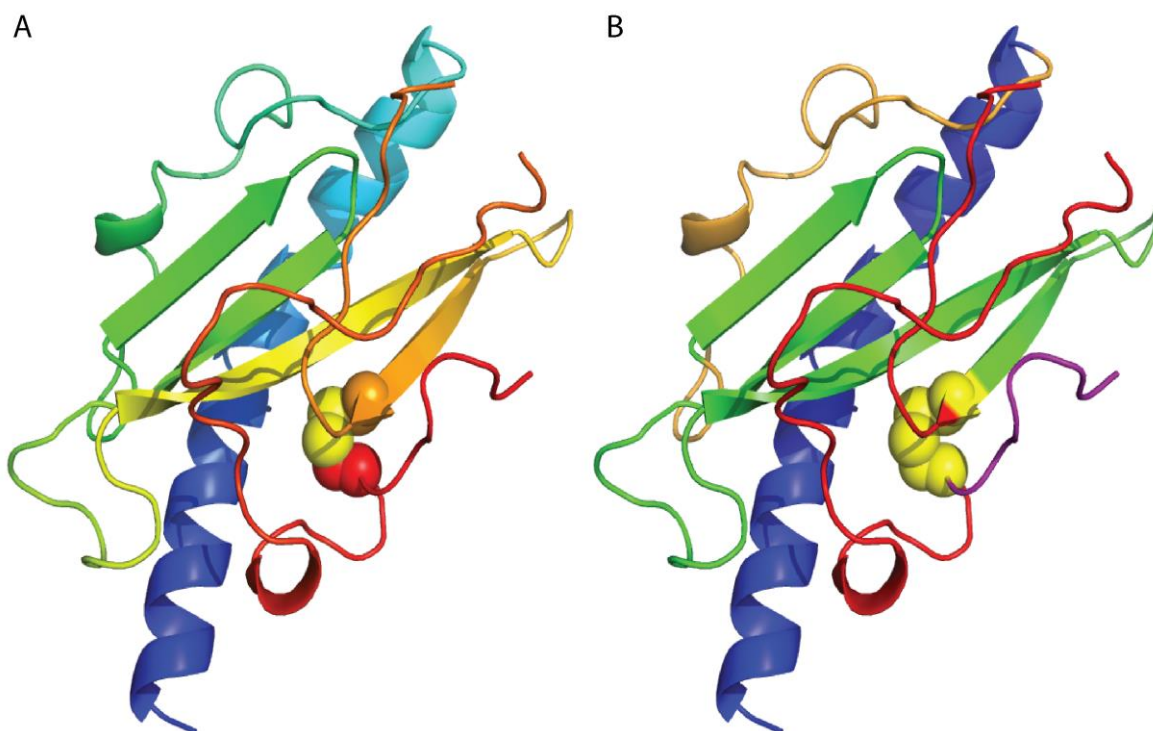


Figure 3.11 The structure of Pile_1.MC58 Δ N24 solved to 1.4Å resolution

Panel A shows the structure of Pile_1.MC58 Δ N24 in cartoon format, coloured rainbow from N-terminus (blue) to C-terminus (red). Panel B shows the structure again in the same cartoon format coloured by feature of the structure: the N-terminal α -helix in dark blue, the extended loop which normally has a disaccharide or trisaccharide bound is shown in orange, the β -sheet region in green, the disulphide in yellow, the D-region in red and the extreme C-terminus in purple.

The lack of secondary structural elements is quite striking in a number of regions. Bar the long N-terminal α -helix and the four β -strands, the rest of the structure is composed of long extended loop regions. The D-region in particular contains only one small turn of α -helix. A run of five alanine residues in the D-region (138-142) is surprisingly ordered given the potential flexibility of such a sequence. However, analysis of the crystal packing reveals that this region is involved in making crystal contacts with the adjacent Pile_1.MC58 Δ N24 molecule. The lack of large numbers of secondary structural elements correlates well with the low calculated temperature determined

from DSC analysis. The structure also confirmed that the CD spectra calculation by K2D2 was incorrect, as the secondary structure content is 26% α -helix and 19% β -sheet, not 48% and 8% respectively as predicted.

Another interesting point of note is that the N-terminal α -helix is not classified as a continuous α -helix, instead a series of smaller ones with breaks at certain residues. This is shown more clearly in the topology diagram of **Figure 3.12**. Depending on the programme being used to analyse the secondary structure in this region (Pymol (DeLano, 2004) or Pro-origami (Stivala et al., 2011)), the helix is broken up at different intervals, splitting into smaller helices. In the previous *N. gonorrhoeae* pilin structures, there is just one break at glycine-42 in the head group region and a second in the truncated region at proline-22. The extra breaks in the meningococcal structure will likely introduce some extra curvature or flexibility into this helix compared to the equivalent gonococcal pilin structure for which this is not seen using either programme. A full discussion of the comparison of this structure and the previous gonococcal structure can be found later in this section, **3.2.9**.

Overall, if the secondary structure is analysed as defined by Pymol, which uses geometric and hydrogen-bonding constraints, fewer elements are seen compared to when the same PDB coordinate file is analysed by the DSSP server (Chen et al., 2006), which uses the same type of constraints but with less strict definitions. As such, many of these border-line regions of secondary structure will be classified as true elements of secondary structure by DSSP-based servers but not by the “stricter” Pymol-based software. Analysis of a delta-phi/delta-psi plot of the superposed structures using LSQMAN (Yusuf et al., 2008) revealed no significant changes in dihedral angle at these “break points” indicating that the difference between the two structures is very slight.

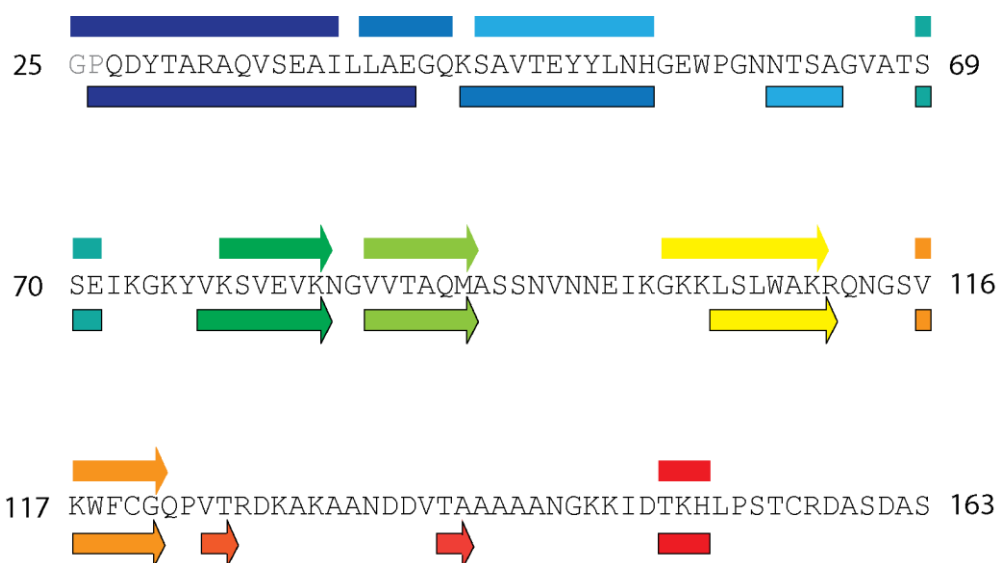


Figure 3.12 Topology diagram of PilE_1.MC58 ΔN24 structure

This diagram shows the basic secondary structure elements found in the structure of PilE_1.MC58 ΔN24. The additional Gly-Pro residues present from the 3C cleavage sequence are in grey and are not included in the numbering system for the protein sequence. α -helices are shown as rectangles and β -sheets as arrows. The secondary structure assignment generated by Pymol are shown above the sequence whilst the assignment from Pro-origami which runs using DSSP, is shown below the sequence and the objects are outlined in black.

3.2.6 Pursuit of a Class II Pilin Structure

The structure of PilE_1.MC58 ΔN24 reveals that the N-terminus protrudes beyond the C-terminal globular head, especially with the additional glycine-proline residues derived from the 3C-cleavage sequence of the vector, see **Table 3.5**. Although, this seems to be important for the crystal contacts in the case of PilE_1.MC58 ΔN24, generally a more globular structure without such an extension is likely to be more amenable to crystallisation. Therefore, the pETM-14 PilE ΔN24 constructs for the class II pilins PilE_15, PilE_22, PilE_4 and PilE_3, were further mutated to remove a total of 29 amino acids from N-terminus. This was achieved using the Quikchange method and the primers detailed in **Table 3.10**.

Primers Name	Sequence	Tm (°C)
ΔN-ter29 Forward	5' GGGCCCATGGACGCCCAAATGTCC -3'	80
ΔN-ter29 Reverse	5' GGACATTTGGGCGTCCATGGGCCC -3'	80

Table 3.10 Primers used to create N-terminal deletion constructs of the pETM-PilE vectors

The modification improved the yields of the pilin produced but despite attempts to crystallise these proteins, no crystals were grown. As such, reliance on models such as those generated in **3.2.7**, must be continued until such a structure is solved.

3.2.7 Modelling of a Class II Pilin

The structure solved of PilE_1.MC58 ΔN24 is that of a class I PilE protein as defined by the extended D-region, the previously established cross-reactivity with the monoclonal antibody against PilE SM1 and the known diversity of this hypervariable sequence within the MC58 strain of *N. meningitidis*. To attempt to extrapolate from the structure of a class I pilin, exactly what a class II pilin would resemble is potentially unreliable. However, attempts were made to model class II pilins using SWISSMODEL (Bordoli et al., 2009) by submitting the sequence of PilE_4 ΔN24.

A comparison of the two structures in **Figure 3.12** shows us that the class II PilE sequences are compatible with the class I PilE fold. The D-region is confined to a smaller loop region in the class II pilin which extends down towards the N-terminus and then back again, therefore the structure is far more compact. The D-region is the surface-exposed region of the pilin in the context of the pilus fibre, correspondingly the more compact structure of the class II pilin could change the topology of the pilus fibre by reducing the number of structural protrusions on the pilus surface as well as the relative pilus diameter. This is also likely to impact on the surface chemistry of the pilus as the side chains of the residues forming the β -sheet will now be exposed to the extracellular environment. In conclusion, the difference between class I and class II PilE sequence is likely to have a significant effect on T4P biology.

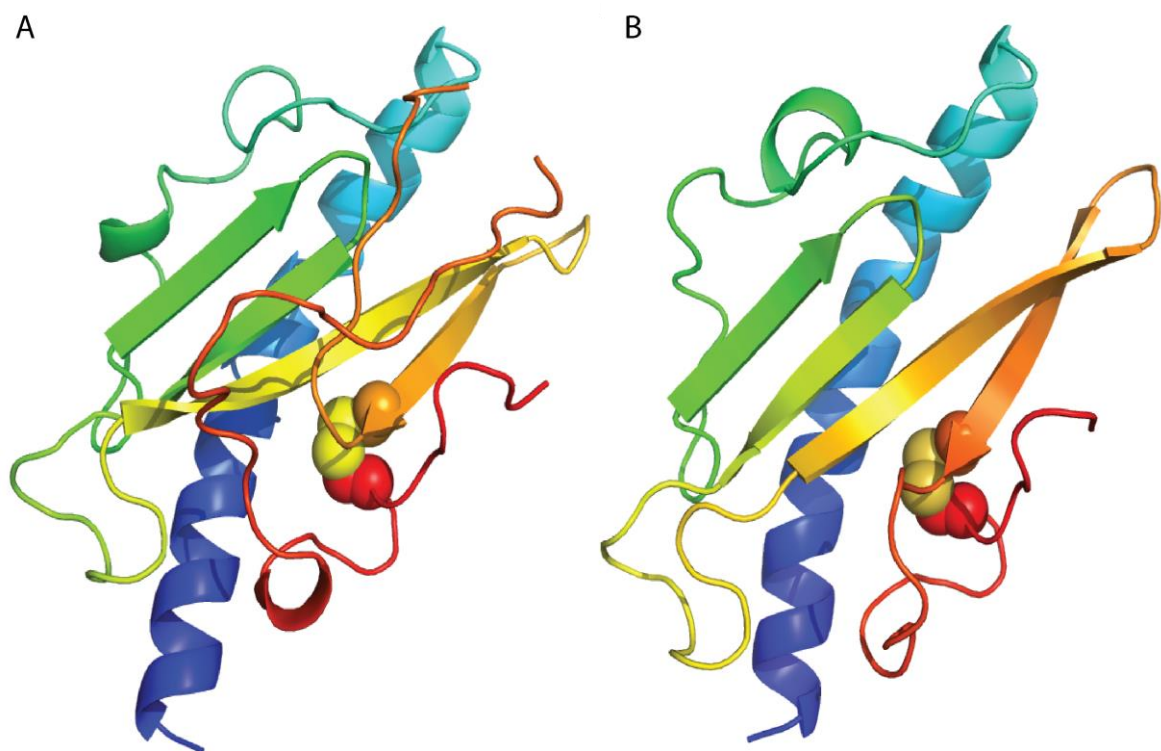


Figure 3.13 Homology modelling of class II pilins

Panel A shows the structure of PilE_1.MC58 ΔN24, a class I pilin. Panel B shows the homology model of PilE_4 ΔN24, a class II pilin, built using SWISS MODEL.

Consurf analysis of the structure was conducted (Ashkenazy et al., 2010). Two multiple sequence alignments were used in this analysis, one of 100 representative class I PilE sequences and the other of the 23 known class II PilE sequences. As such, the conservation of sequence for each class of PilE can be mapped onto the structure. The results can be seen in **Figure 3.14**.

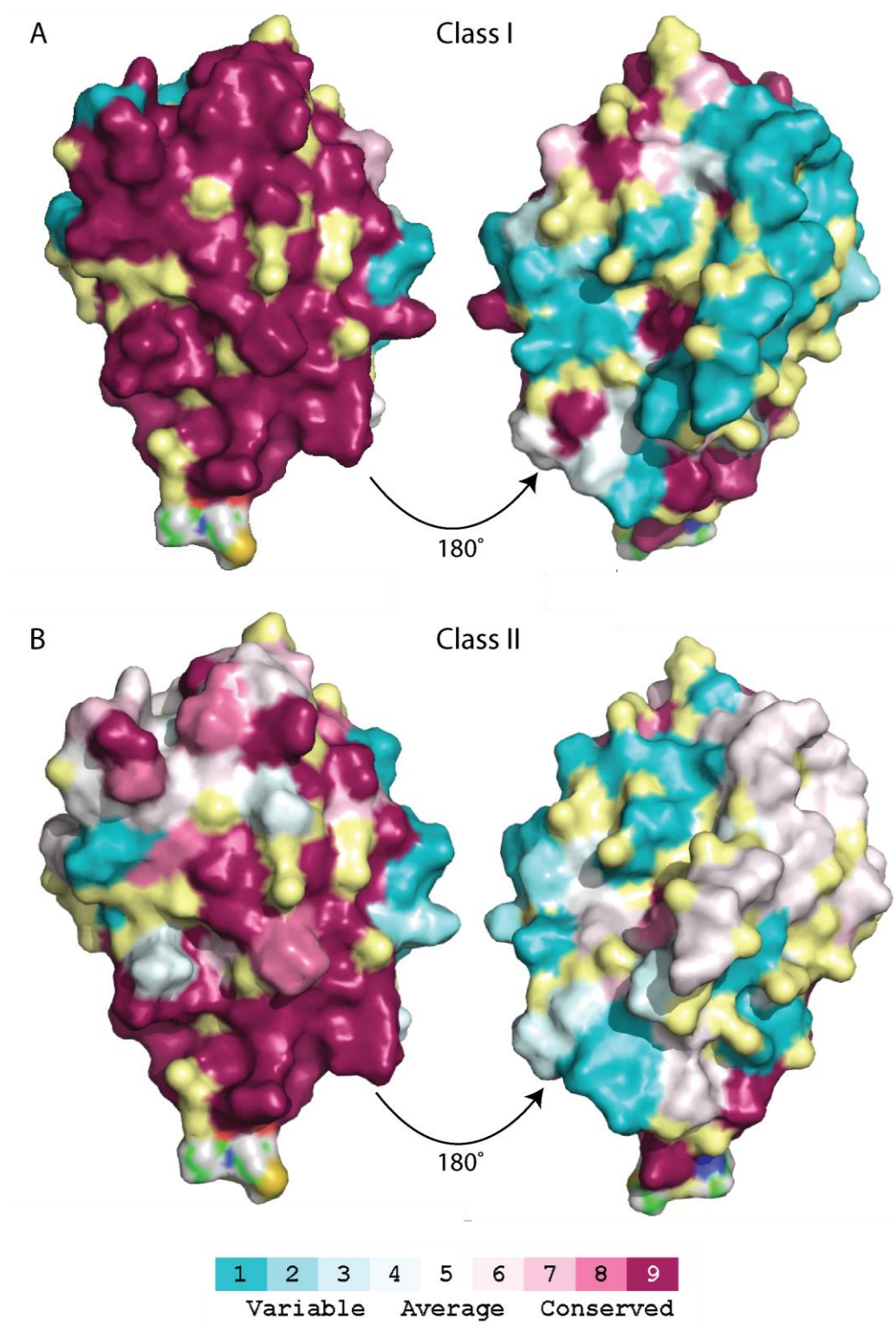


Figure 3.14 Consurf analysis of PiIE using class I and class II PiIE sequences

In both panels, the left hand side shows the back surface of the PiIE structure which has the hydrophobic backbone α -helix, whilst the right hand side shows the front surface of the PiIE structure containing the D-region which would be exposed in the context of the pilus fibre. For residues where there is insufficient data to make a statistically significant decision as to its level of conservation, these are shown in yellow.

In both panels, the left hand side shows the back surface of the PiE structure which has the N-terminal hydrophobic backbone α -helix, whilst the right hand side shows the front surface of the PiE structure containing the D-region which would be exposed in the context of the pilus fibre. There is strong conservation of sequence on the back face of the PiE for class I PiE sequences whilst the front face shows much greater variation, as expected. However, for class II PiE sequences, the back face is less conserved, particularly at the top of the N-terminal backbone α -helix, the epitope of the PiE class defining SM1 antibody. Class II PiE sequences also show reduced variation on the front face compared to class I PiE sequences. This may be explained by poorer modelling of the class II sequences on the PiE_1.MC58 Δ N24 structure due to the short D-regions of these PiE sequences. Sequence variation within class II PiE sequences is generally lower than it is within class I PiE sequences as exemplified in the sequences used in this study, the alignment of which can be seen in **Figure 3.1B**.

Perhaps a better model of class II PiE sequence conservation can be seen in **Figure 3.15** where the same consurf analysis is applied to the SWISSMODEL of class II PiE generated earlier and shown in **Figure 3.13**.

The conservation on the back face of PiE now shows much greater conservation in sequence compared to the previous model in **Figure 3.14B**. However the front exposed PiE face still does not show as much variation as that of class I PiE sequences. This can most likely be explained by the smaller D-region which defines the class II PiE sequences. Drawing conclusions from this analysis must be limited as the structure used is a model. Confirmation of these findings if and when a class II PiE protein structure is determined will be necessary.

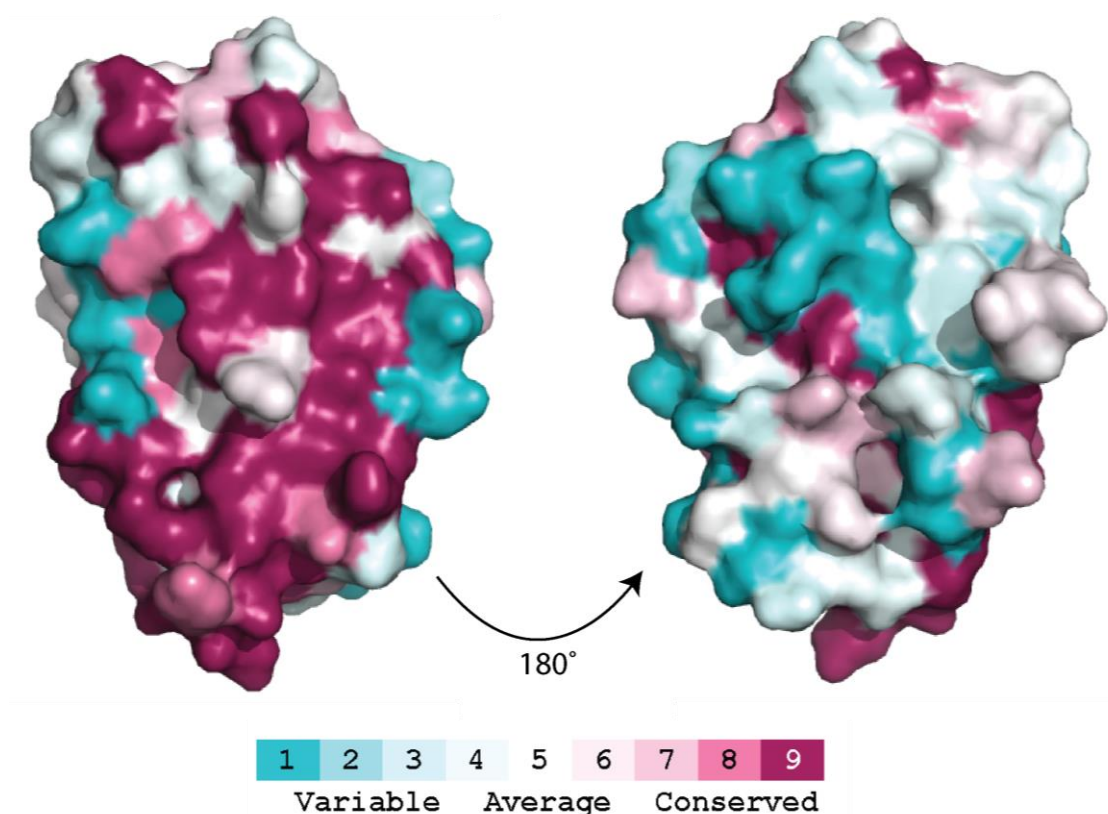


Figure 3.15 Consurf analysis of PilE using class II PilE sequences

The left hand side shows the back surface of the PilE structure which has the hydrophobic backbone α -helix, whilst the right hand side shows the front surface of the PilE structure containing the D-region which would be exposed in the context of the pilus fibre. For residues where there is insufficient data to make a statistically significant decision as to its level of conservation, these are shown in yellow.

3.2.8 Crystal Packing of PilE_1.MC58 Δ N24

The crystal packing of PilE_1.MC58 Δ N24 revealed rings consisting of six protein subunits arranged in a head-to-tail fashion which are stacked in perfect register to form long tubes. Such an arrangement accounts for the high solvent content of this crystal form (56.5%), see **Figure 3.16**.

The extreme N-termini in this arrangement are packed into the adjacent protein molecule, indicating that this form of packing is not physiologically relevant as there is no room for the truncated N-terminal α -helix to be accommodated. However, this packing is not dissimilar to that seen in the TcpA crystal form (Craig et al., 2003b), the homologous PilE from *Vibrio cholerae*, see **Figure 3.17**. Again in this P6₃ crystal form, rings of TcpA were formed of six protein molecules

which stacked together into tubes. However, in this case, the N-termini are pointing towards the central gap of the tube and the subunits are positioned as a three-start helical arrangement rather than flat rings of protein molecules. The paper which reported this structure and crystal form concluded this as being physiologically relevant as the crystals were grown in “near physiological” conditions, there is space for the truncated N-terminal α -helix to occupy the centre of the tube, the orientation expected of the N-terminal helices, if fitted, is similar to that expected of the TCP filaments and the diameter of the tubes is close to that measured by electron microscopy experiments of TcpA fibres. This crystal form is thought to represent the “relaxed” fibre given the lack of N-terminal α -helices to constrain the structure and confer its stability.

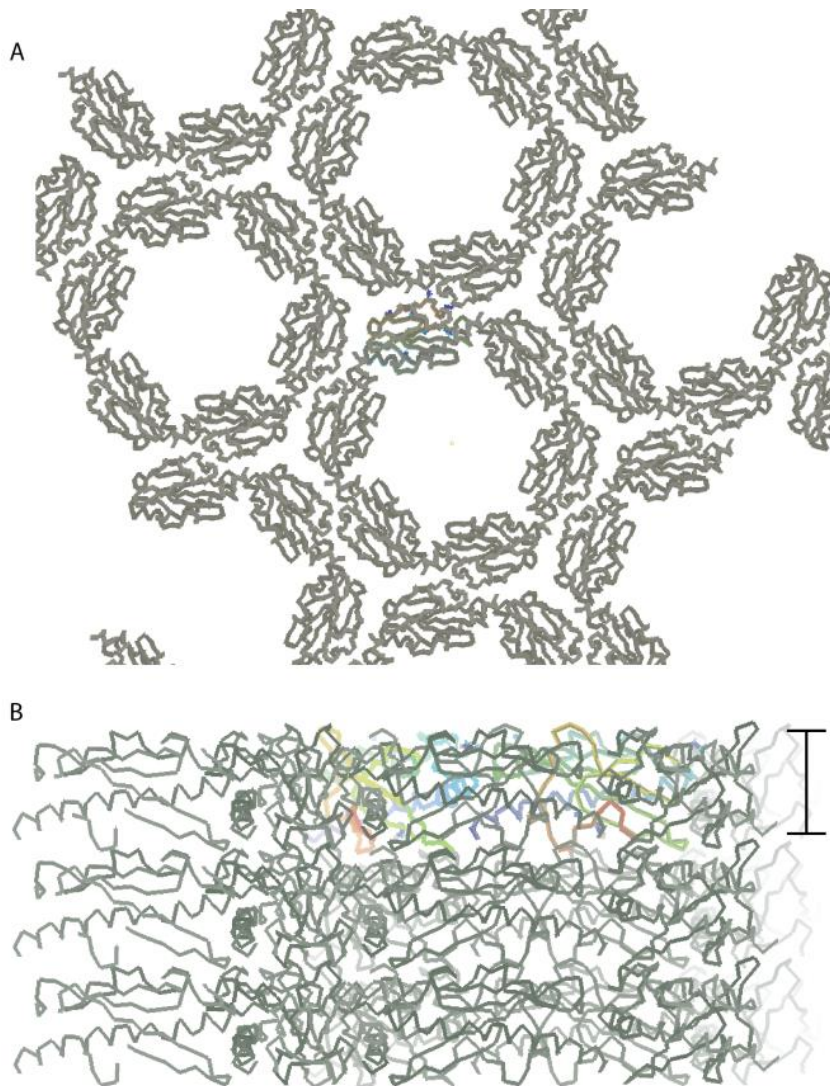


Figure 3.16 Crystal packing arrangement of PilE_1.MC58 Δ N24

Panel A shows the hexameric rings of PilE_1.MC58 Δ N24. Two protein molecules are coloured in the same rainbow fashion as the structure in **Figure 3.11**, from N-terminus (blue) to C-terminus (red), illustrating the head-to-tail arrangement of the pilin head groups in the ring. Panel B shows how the rings in this crystal packing stack in perfect register on top of each other to form flat planes of helical ring assemblies, with the height of each layer indicated.

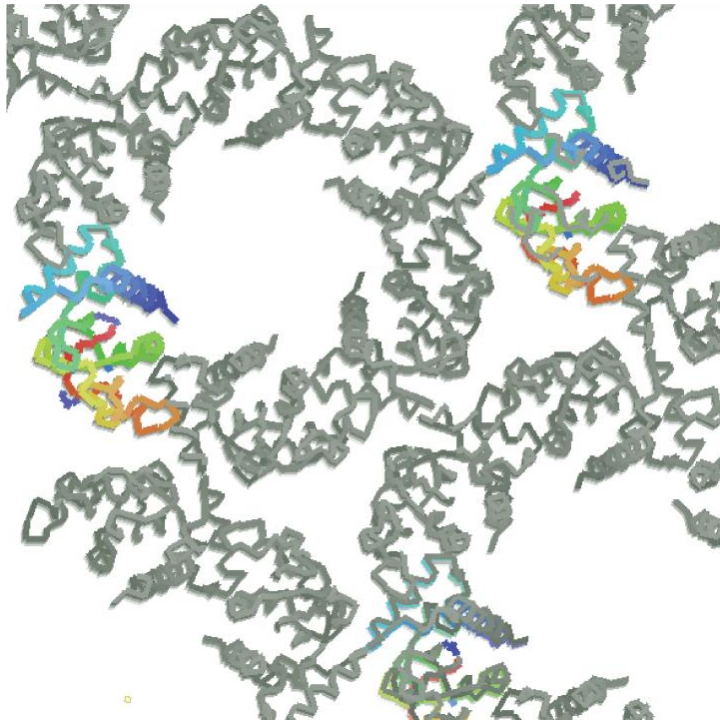


Figure 3.17 $P6_3$ Crystal packing of TcpA of *V. cholerae*

Although having a similar crystal packing to MC58 ST-32 PilE $\Delta N24$, the assembly of TcpA is in fact helical and close to the expected physiological packing of TCP fibres due to the tilted orientation of the protein molecules which would accommodate the N-terminal α -helix within the assembly core.

The reason PilE_1.MC58 $\Delta N24$ may form such rings, is not immediately obvious beyond this arrangement forming favourable crystal contacts. However, the assembly of pilin monomers *in vitro* into such protein nanotubes has also been described for *P. aeruginosa* (Petrov et al., 2013, Audette et al., 2005). The nanotubes formed differed slightly from pilus fibres in terms of their helical pitch and diameter but retained much of the tensile strength of pilus fibres and showed promise for DNA binding and other nanotechnological applications. Anecdotal evidence of EM analysis of pilus preparation in which whole T4P are extracted by a shearing process from live *N. meningitidis*, indicates that nanotubes of pilin are consistently present in these samples. These tubes are hypothesised to form of dissociated pilin. The packing of the PilE monomers would be expected to conform more to the $P6_3$ arrangement of the TcpA crystal packing, a relaxed pilus fibre, due to the hydrophobic N-terminal α -helices present on these PilE proteins, not found in the protein produced recombinantly.

3.2.9 Comparison with the Gonococcal MS11PilE Structure

The most closely related PilE structure available was that of the *N. gonorrhoeae* MS11 PilE which was used for the molecular replacement (PDB ID: 2HI2). The head groups of the two structures share 83% sequence identity, 85% over the entire length of the protein sequence. The key regions of disparity in the sequence, as seen by the alignment in **Figure 3.1**, are within the D-region, which in the PilE_1.MC58 Δ N24 form is much longer. Superposition of the PilE_1.MC58 Δ N24 with the *N. gonorrhoeae* MS11 PilE full-length structure (PDB ID: 2HI2) can be seen in **Figure 3.18**.

Panel A shows how in the majority of the structure, the peptide backbones overlay very well with an RMSD of 1.03Å over 129 residues. However, rotation of the superposed structures by 90° in panel B reveals some disparity between the two structures, particularly in the loop regions. The curvature of the N-terminal α -helices of both structures is very similar despite some differences in the completeness of this helix in PilE_1.MC58 Δ N24.

Panel C shows the three key regions of the head group which are most different, all are flexible loop regions. Region 1 most likely corresponds to the intrinsic flexibility of this region of the structure as it extends away from the rest of the pilin head group. Region 2 corresponds to the D-region where the sequence of PilE_1.MC58 Δ N24 is longer. This region, as discussed previously is largely unstructured and as such, the PilE_1.MC58 Δ N24 has two fewer β -strands compared to MS11 PilE when the secondary structure is analysed by Pymol. Region 3 contains the run of five alanines and as such is inherently more flexible, so again it is unsurprising that this loop is extended. These five alanine residues are within the hypervariable loop region of the D-region and are not conserved in other class I PilE alleles. Therefore it is not expected that the extension of this loop region is conserved across other class I PilE proteins.

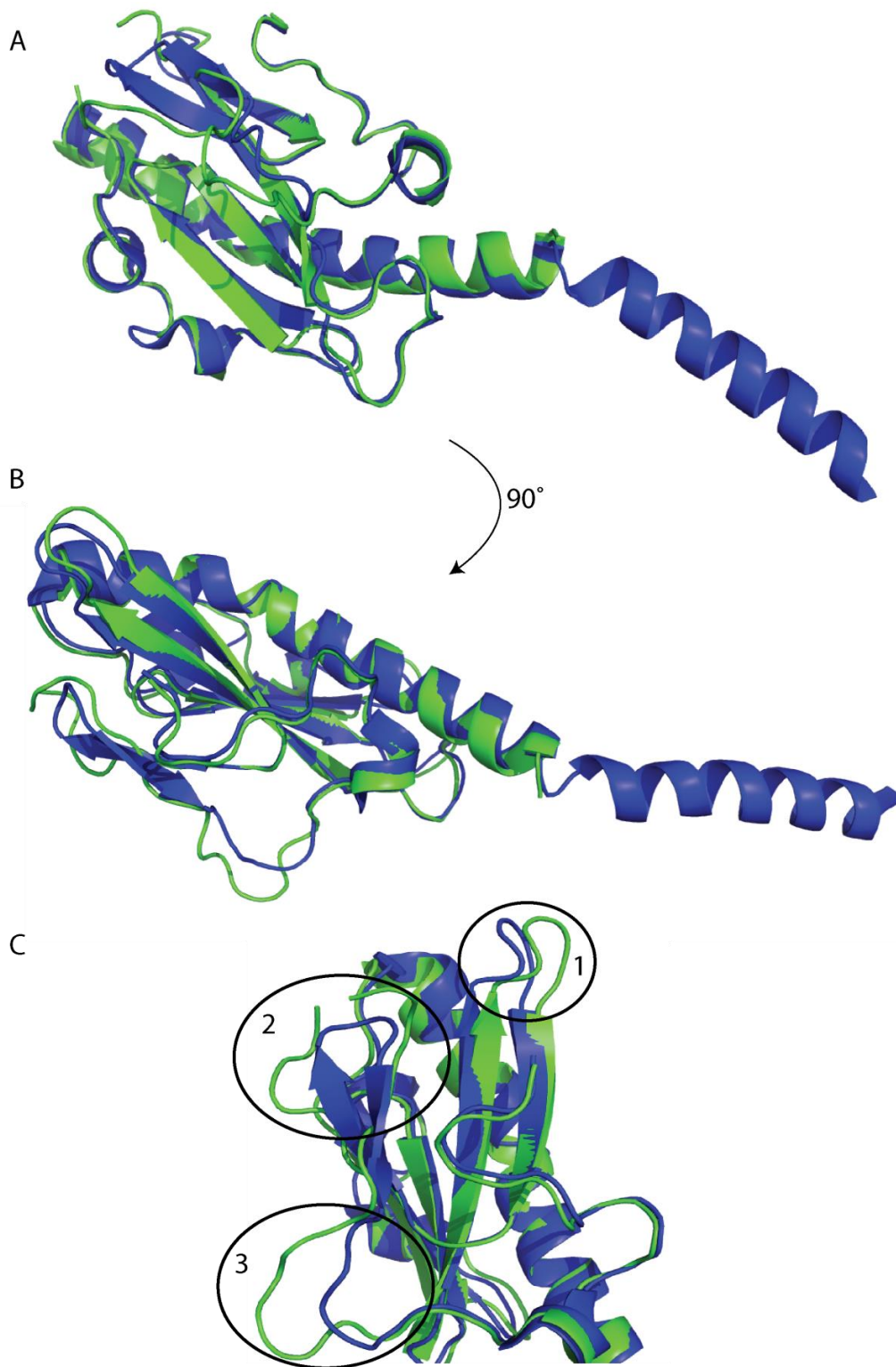


Figure 3.18 Superposition of PiE of MC58 ST-32 PiE Δ N24 with the *N. gonorrhoeae* MS11 PiE
 Throughout, the meningococcal PiE is shown in green, the gonococcal PiE (PDB ID: 2HI2) in blue. Panel A shows the superposition of these two structures. Panel B shows the same superposition rotated 90°C. Panel C focusses on the three key areas of major difference in the head group region.

The pseudo-atomic resolution EM reconstruction structure of the *N. gonorrhoeae* MS11 pilus fibre allowed modelling of PilE_1.MC58 ΔN24 pilin in the context of the pilus. **Figure 3.19** shows superposition of one molecule of PilE_1.MC58 ΔN24 the reconstruction (PDB ID: 2HIL). In this way the changes in pilus surface topology, due to the extended and loosely structured D-region of the PilE_1.MC58 ΔN24, are clearly illustrated in **Figure 3.19C**.

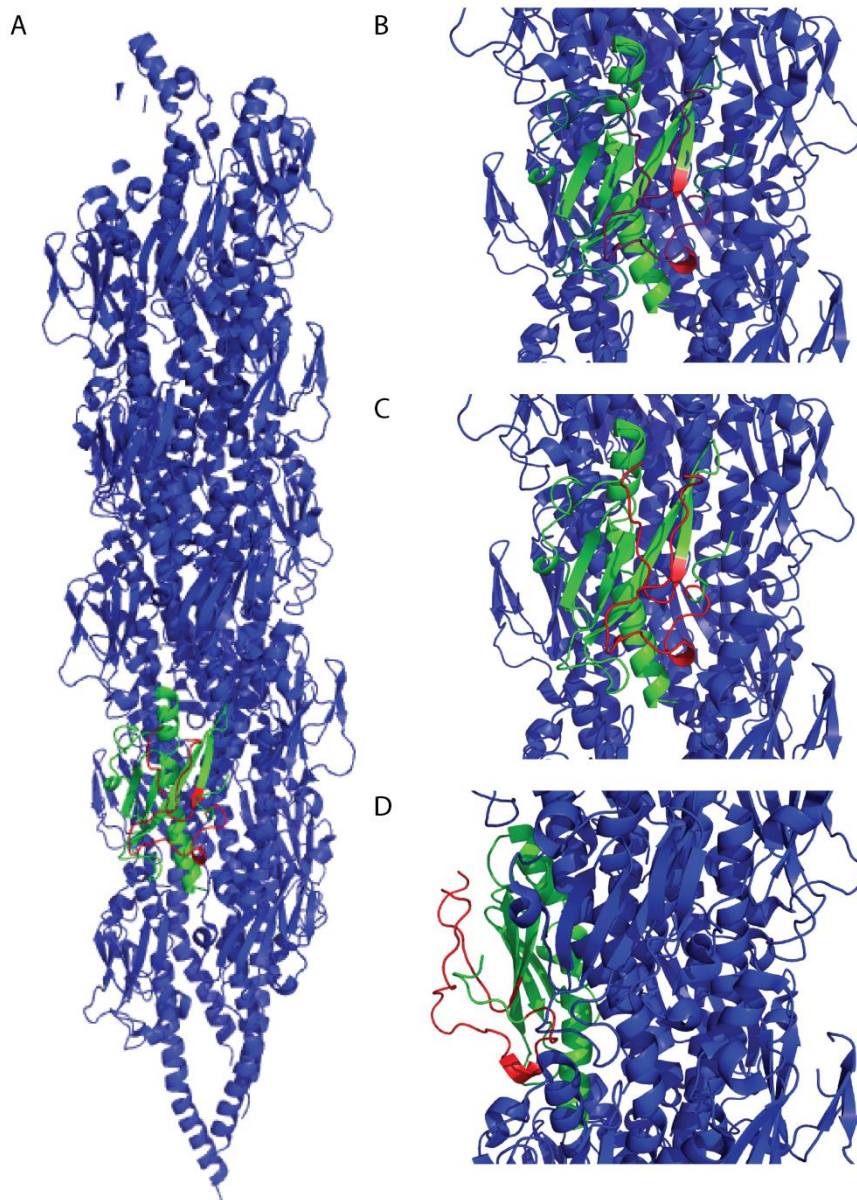


Figure 3.19
Superposition of one molecule of PilE_1.MC58 ΔN24 onto one gonococcal pilin subunit in the EM reconstruction of the pilus fibre (PDB ID: 2HIL)

Throughout, PilE_1.MC58 ΔN24 is shown in green with the D-region in red, the gonococcal MS11 PilE subunits in blue. Panel A shows the PilE_1.MC58 ΔN24 molecule in the context of the fibre. Panel B shows the superposition of the two protein molecules within the fibre. Panel C and D show the meningococcal protein in the context of the fibre without its gonococcal superposition partner at two different angles.

Models of the pilus structure for *N. meningitidis* MC58 can be built by two means, the end results of which are identical. Firstly this model can be achieved by superposition of every single pilin subunit within the EM reconstruction. Secondly a model can be built using a script which copies

the pilin subunit PDB coordinates, rotates and translates them by the same distance (100.8Å) and angle (10.5°) as adjacent molecules in the pilus fibre, generating a helical assembly. Both methods were used successfully. Analysis of the *N. meningitidis* pilus model in Molprobit revealed very few clashes.

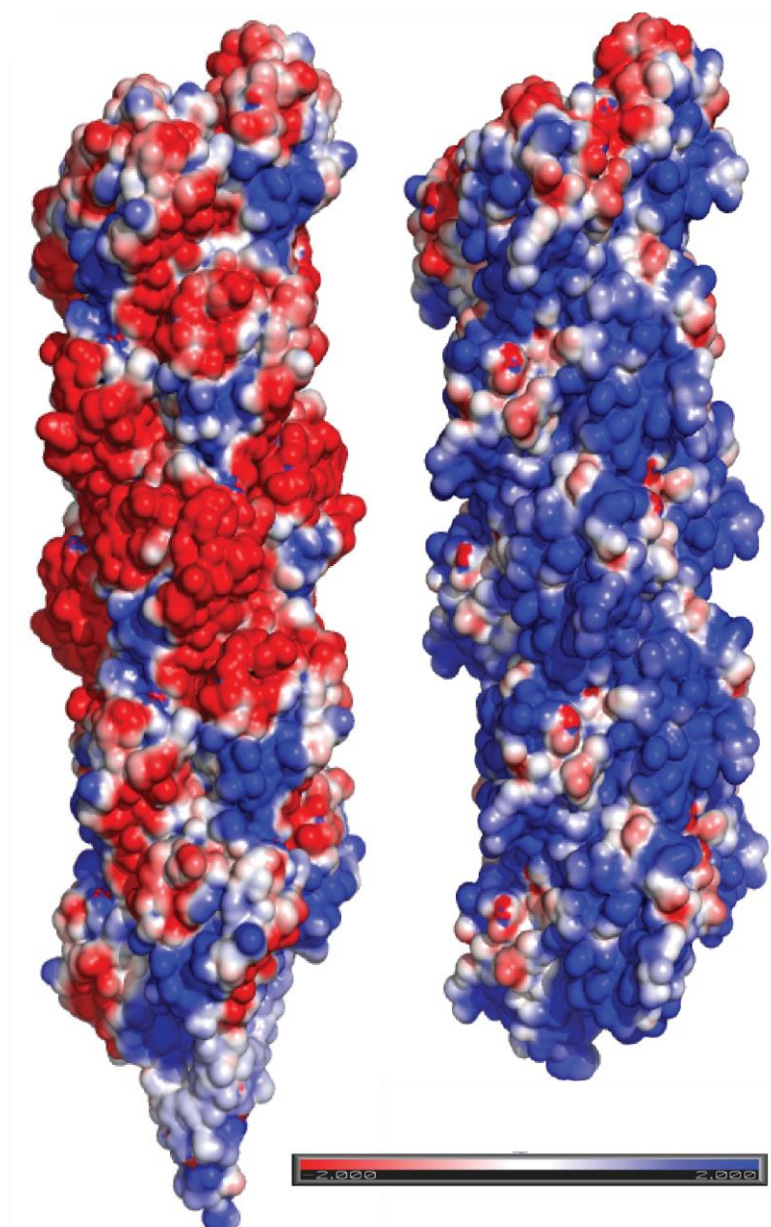


Figure 3.20 Surface charge analysis of pilus models and the residues responsible

Surface charge models of the gonococcal (left hand side) and meningococcal (right hand side) pilus fibre models.

Given the expected differences in surface chemistry of the *N. meningitidis* MC58 compared to the *N. gonorrhoeae* MS11 pilus fibres based on sequence differences in known surface exposed regions, the gonococcal MS11EM reconstruction (PDB ID:2HIL) and meningococcal MC58model were analysed by APBS through Pymol (Unni et al., 2011). This software assesses the electrostatics of a molecules surface using the Poisson-Boltzmann equation in conjunction with the finite element tool kit, a library of element modelling classes. The results are striking, see

Figure 3.20

The gonococcal MS11 pilus fibre is generally negatively charged, but has bands of positive charge across its structure. This positively charged groove was reported to be suitable for DNA binding in natural transformation (Craig et al., 2006b). The alternating bands of positive and negative charge running across the pilus fibre surface were also proposed to be important in terms of microcolony formation as it was suggested adjacent pili could bind the alternate and complementary regions of charge in order to link bacteria together.

However, the meningococcal MC58 pilus fibre model is almost entirely positively charged. If the proposed functionality of the charged patches of the gonococcal pilus fibre were correct, then this would suggest that the *N. meningitidis* MC58 pilus fibre would have very different properties in terms of its contribution to natural transformation, cellular adhesion and microcolony formation.

It is important to note that these models do not include any post-translational modifications in the APBS calculations as the software is unable to analyse such large structures with any non-protein moieties. PilE_1.MC58 Δ N24 was produced recombinantly so to model in such modifications would be unreliable and the exact nature of the modifications is unknown. PilE of both species has a sugar group and a number of serine residues which are modified with negatively charged groups such as phosphate. As all of these modifications are predicted to be surface exposed, they will not only alter the surface topology of the pilus fibre, but also distort the surface charge distribution to be increasingly negative. Whether these modifications would affect

DNA binding by the pilus or pilus-pilus interactions has not been explored in the context of these strains yet.

3.3 Perspectives

The work described in this chapter has developed and defined a successful protocol to produce N-terminally truncated meningococcal PilE proteins in large quantities. Both class I and class II pilin proteins were successfully expressed and purified using this protocol, allowing further analysis of PilE described in **Chapter 4**. The structure of PilE_1.MC58ΔN24 was solved to 1.4Å resolution by molecular replacement with the gonococcal PilE protein structure. The structure reveals disparities between the PilE proteins of these two closely related species which contribute towards differences in the topologies and surface charge of the pilus fibre models. The models do not take into account the presence of post-translational modifications or the presence of minor pilins in the pilus fibre. However, the models built do provide the first comparison of major pilin structures from *Neisseria* species and a good starting point from which to discuss PilE sequence variation in the context of pilus structure-function relationships when comparing PilE proteins of different sequences.

T4P are long polymers of PilE which also contain the minor pilins and are associated with PilC. T4P functions are generally attributed to the minor pilins and pilus-associated proteins; DNA binding to ComP (Berry et al., 2013), aggregation to PilX (Brissac et al., 2012) and cell signalling to PilV and adhesion to PilC (Cehovin et al., 2013, Brissac et al., 2012, Winther-Larsen et al., 2001, Rudel et al., 1995c). Elucidating the role of PilE itself is difficult due to its crucial role in pilus assembly and structure. Previous research in the field of *Neisseria* T4P has shown that subtle sequence differences in the PilE protein can alter various pilus functions including adhesion and cell signalling (Nassif et al., 1994). Despite this, the driving force behind PilE sequence variation is often attributed to host immune evasion with the observed alterations in adhesion a secondary effect of this mechanism (Hubert et al., 2012). More recent studies have implicated PilE in these T4P functions. Suggestions have also been made that minor pilins are actually located in the periplasm where they regulate T4P biogenesis, their location within the fibre not being necessary

for their function (Imhaus and Dumenil, 2014). Structural analysis in this field is important to glean insight into the complicated structure-function relationships of T4P proteins which are not always evident from the primary sequence.

In this study, PilE sequence changes have been shown through the detailed comparative analysis of the meningococcal and gonococcal structures described here to affect the overall chemistry of the pilus (Craig et al., 2003a). Due to the hypervariable nature of the PilE gene within class I pilin strains, the sequence is not only extremely different between strains but is also undergoing variation within a strain itself. As such, a “one rule for all” method is a very simplistic approach to understanding the sequence-function relationships of PilE. Differences in surface charge may be considered to represent adaptation to the different host niches of these two species. However, given that the *N. meningitidis* MC58 and *N. gonorrhoeae* MS11 PilE sequences are more closely related than *N. meningitidis* class I and class II PilE sequences, this is a weakened argument. If PilE sequence variation is so crucial to variation in pili surface chemistry, then it is difficult to rationalise why some of the most virulent strains of *N. meningitidis* which cause epidemic disease in sub-Saharan Africa express conserved class II pilins. Such strains may have other virulence factors which dominate over PilE in the pathogenesis of the disease, though limited research has been carried out investigating the sequence variability of multiple virulence factors and disease severity or infectiousness for *N. meningitidis*.

A further level of variation is added to PilE chemistry, though not investigated in this work, by the changing variety and frequency of post-translational modifications present on the surface of these proteins (Aas et al., 2006b, Craig et al., 2006a, Virji et al., 1996b, Aas et al., 2006a). Most of the known modifications are surface exposed and will likely influence pilus chemistry. Recently, high-throughput mass spectrometry analysis of meningococcal isolates revealed that within a single isolate, multiple forms of the major pilin co-exist, representing proteins with different sets of post-translational modifications (Gault et al., 2014).

In the context of natural transformation, the minor pilin ComP has been shown to be important in the binding of DNA (Cehovin et al., 2013). However, natural transformation rates vary between different strains of *N. meningitidis* despite the ComP sequence being largely conserved within the species (Cehovin et al., 2010). Other factors must account for this difference in natural transformation efficiency. Like all minor pilins, it is unknown how many copies of ComP are present in the pilus fibre nor whether they are clustered or spread throughout the structure, or even if this varies strain to strain. It could be postulated that whilst ComP has a high affinity for DNA, in particular for the DNA uptake sequence (DUS) regions (Cehovin et al., 2013), non-specific binding of DNA by the T4P contributes to tethering the DNA to the pilus fibre. Such non-specific binding could be mediated by the PilE protein. Variation of the PilE sequence could then account for changes in the surface charge and topology and hence the affinity of the pilus for DNA, therefore having a direct impact on natural transformation efficiency. Likewise similar changes in pilin sequence could then be accountable for differences in ability to form microcolonies through pilus-pilus interactions.

Other minor pilins are also likely to influence the overall surface chemistry of pili. However, given the lack of information regarding their frequency and location with the T4P, it is risky to speculate regarding the impact these proteins might have on the pilus structure.

Obtaining a *N. meningitidis* class II PilE structure is important due to the differences between class I and class II PilE proteins from this species (Wormann et al., 2014). Class II PilE proteins are always much smaller than those of class I and in particular have much smaller D-regions, typically 13-18 amino acids compared to 20-40 amino acids of class I PilE. Also they do not vary, at least not by gene conversion events, and are likely to have different post-translational modifications compared to class I pilins. A structure of such a PilE molecule would help our understanding of these differences between class I and class II PilE proteins.

To further understand the relevance of the surface chemistry of pili expressing different PilE sequences, analysis of the functional differences of strains expressing different sequences should provide an insight into the precise role PilE plays in the functional activity of pili. Pilin-swap mutants, in which whole PilE genes are swapped in different strains should allow characterisation of exactly what different functions different pilin sequences convey. Such work is currently being conducted by Dr Mirka Wormann in Professor Christoph Tang's laboratory, the provisional results of which are already confirming that PilE plays more than just a structural role in the context of T4P.

Overall the intricacies of PilE structure-function relationships remain elusive. Comprehensive analysis of pilus functions in a large range of PilE swap mutants, should allow elucidation of PilE function in the context of variable sequence.

4. Functional Analysis of Meningococcal PilE

PilE is often considered as a structural component in its role within the T4P, accommodating the functional minor pilins within the pilus fibre and the tip-associated PilC to which most T4P functions are generally assigned. However, several reports suggest that PilE plays more than just a structural role, with amino acid sequence changes and alternative post-translational modifications of PilE affecting whole T4P functions (Miller et al., 2014a). With the development of a protocol to produce recombinant PilE Δ N24 proteins from a range of different *N. meningitidis* strains, validated by the structure of *N. meningitidis* PilE_1.MC58 Δ N24, properties and functions of pilins with different amino acid sequences could now be explored in more detail, including a comparison of class I and class II PilE expressing strains.

4.0.1 Adhesion to host cell surfaces

Pili are known to be very important for adhesion and colonisation of host surfaces. However, the interaction between pili and host cells is poorly characterised. Both PilE and PilC are implicated as being directly involved in the interaction with the human nasopharyngeal epithelium (Rudel et al., 1992), the niche which *N. meningitidis* colonises, acting as direct adhesins, whilst minor pilins PilX and PilV are indirectly involved in adhesion of T4P to host epithelia (Winther-Larsen et al., 2001, Helaine et al., 2005). However it is unclear if both of these proteins bind the same cellular receptor or not, or even what the putative receptor for either protein could be. CD46 has been reported as the pilus receptor (Kallstrom et al., 1997) though this has not been universally accepted by the Neisserial field due to the contradictory published evidence (Johansson et al., 2003), see **Table 1.1**. Until this point, no work has been published detailing a direct interaction between PilE and CD46 at the protein-protein level, for example, through biophysical analyses such as SPR.

4.0.2 Host cellular effects upon bacterial adhesion

Bacterial adhesion to host epithelia, in general, is known to potentially induce certain cellular effects including cell signalling pathways, barrier disruption and cytotoxicity (Brooks and Jawetz, 2013). *N. gonorrhoeae* colonisation was shown to reduce apoptosis in a fallopian tube infection model in an adhesion dependent manner (Reyes et al., 2007). In an earlier study, reports alluded to the fact that meningococcal pili are toxic towards the vascular endothelium (Melican et al., 2013). Piliation was shown to work synergistically with LPS release to cause cytotoxic damage to this endothelial surface to which meningococci are adhered (Dunn et al., 1995). A correlation was observed between strains expressing PilE sequences which conferred increased bacterial adherence and increased host cytotoxicity suggesting the exact sequence of PilE may be important. T4P-dependent adhesion of *N. meningitidis* was also shown to induce Ca^{2+} signalling pathways in the host epithelia ME180 model through casein kinase II (Kallstrom et al., 1998). Overall however, the interplay between bacterial adhesion, the consequential effects of Neisserial colonisation of host epithelia and the role PilE sequence has in this complex host-pathogen interaction have been poorly characterised at the molecular level, with much generalisation of the adhesion systems of the gonococcus and meningococcus.

4.0.3 Immunogenicity of PilE proteins

Piliation is a requirement of all strains of *N. meningitidis* for adhesion and colonisation of host epithelia and endothelia, a pre-requisite to fulminant meningococcal infection. Pili are known to be immunogenic based on early mouse experiments as well as the presence of anti-pili antibodies in patient sera (Schoolnik et al., 1984, Forest et al., 1996). It was therefore suggested that pilin represented an excellent vaccine candidate due to these immunogenic properties and its surface exposed location. However, given the hypervariable nature of PilE in many strains, efforts were focussed on raising antibodies against minor pilins or PilE epitopes which were conserved irrespective of whether the strain underwent hypervariation of its PilE sequence or not (Cehovin et al., 2011), despite the D-region being reported as the immunodominant epitope. Other studies

reported that chimeric pilin proteins lacking a D-region, mimics of class II PilE proteins, are in fact immunosilent implying antibodies could not be raised against these proteins, irrespective of any sequence variation issues (Hansen et al., 2007). No work has been published investigating whether recombinant PilE proteins constitute good vaccine candidates in the case of *N. meningitidis*.

The work in this chapter set out to characterise the properties of recombinant PilE proteins and describe the molecular events in three functional aspects; the putative epithelial cellular binding receptor CD46, the gross effects on incubating PilE with human epithelial cells and finally its immunogenic properties. All three functional areas were explored with respect to the role of PilE sequence differences and class I compared to class II PilE proteins.

4.1 Investigations into the Interaction between PiLE and CD46

The true identity of the meningococcal T4P receptor has long remained elusive. The membrane bound complement regulator CD46 was initially proposed almost 20 years ago, but this remains controversial within the field. Nearly all studies investigating the role CD46 has in Neisserial T4P adhesion, use cell biology based methods to study the interaction, highlighted in **Table 1.1**. The work discussed in the previous chapter describes the development of a protocol for soluble cytosolic expression and purification of recombinant PiLE, validated by the structure of PiLE_1.MC58 ΔN24. A method for recombinant production of the SCR domains of CD46 had already been established and validated. This study intended to assess the interaction between PiLE and CD46 at the molecular level through biophysical analyses. Whilst writing this thesis, CD147 was described as the putative T4P receptor. A discussion of this work can be found in **4.5**.

4.1.1 Production of PiLE and CD46

The four extracellular SCR domains of CD46 may be produced by over-expression of the protein in *E. coli* into inclusion bodies, purification of these inclusion bodies, refolding the protein and then purifying the correctly refolded species. Two, three and four SCR domain gene fragments were previously cloned in the pET14b expression vector by Dr Kirstin Leath, University of Oxford. The method used was identical to that for purification of m-fH₆₇ described in **2.1.1**, though as CD46 does not interact with glycosaminoglycans in the way fH does, the protein elutes from the size-exclusion column within 1 CV. The highest yield of CD46 was produced from two SCR-domain constructs, most probably because as the number of disulphide bonds increases, there is greater potential to form incorrect disulphide bonds within the protein molecule. A previous study implicated SCRs 3 and 4 of CD46 in the binding of the meningococcal pilus. Therefore CD46₃₄ was expressed and purified, a construct corresponding to SCR domains 3 and 4. See **Figure 4.1**.

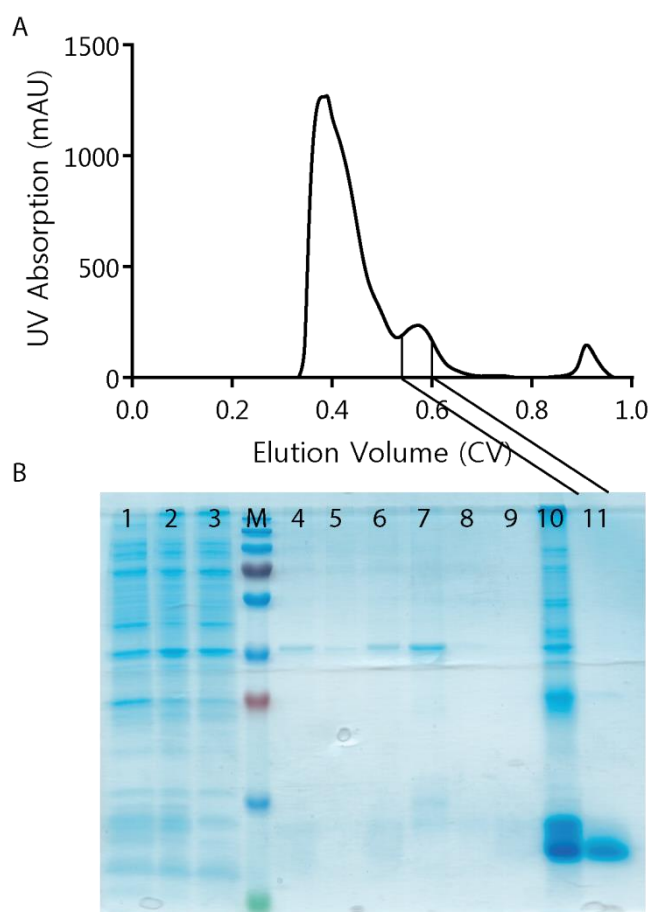


Figure 4.1 Size-exclusion and SDS-PAGE analysis of CD46₃₄ purification

Panel A shows the size-exclusion elution trace of the purified CD46₃₄ sample.

Panel B shows the non-reducing SDS-PAGE analysis of samples taken throughout the expression and purification of CD46₃₄: M - molecular weight markers, 1 - pre-induction cells, 2 - post-induction cells, 3 - homogenate, 4 - supernatant after inclusion body harvesting by centrifugation, 5 - solubilised inclusion bodies, 6 - pellet after dialysis in 6 M urea pH3-4, 7 - supernatant after dialysis into 6 M urea pH3-4, 8 - Vivaflow flow through, 9 - flow through after concentrating with spin concentrator, 10 - sample after dialysis into 50 mM Tris, 150 mM NaCl, pH 7.5, 11 - sample from size-exclusion chromatography peak

For the SPR experiments conducted with the CD46 protein (**4.1.3**), native protein samples were sufficient. However for HSQC analysis of the interaction between PiE and CD46, labelled samples were produced, as the natural abundance of ¹⁵N is not high enough to generate signal. ¹⁵N - isotopic enrichment is necessary to provide an appropriate sample for NMR spectroscopy. It was achieved by ensuring that ¹⁵NH₄Cl was the only nitrogen source available to the cells whilst protein expression occurred. Overnight cultures grown in LB were diluted 1 in 100 to allow growth of 2 x 1l cultures in TB media. Cells were grown at 37 °C until OD₆₀₀ reached approximately 0.7-0.9. Cells were harvested by centrifugation and the cell pellet was resuspended in 50 ml 10X M9 salts and then made up to 500 ml with autoclaved water and the additional media components. medium components including ¹⁵NH₄Cl, seen in **Table 4.1**. The cells were rested by incubating at 21 °C for 20 minutes. Protein expression was then induced by adding 1 mM IPTG and growing overnight by shaking the flasks at 21°C. Cells were harvested by centrifugation and the CD46 was

purified using the inclusion-body protein purification method described in **2.1.1**. By this method, 500 µl of a 500 µM ^{15}N -CD46₃₄ protein sample in 25 mM acetate pH5.5 supplemented with 5% (v/v) D₂O was produced.

Component		Concentration
M9 salts:	Na ₂ HPO ₄	42 mM
	KH ₂ PO ₄	22 mM
	NaCl	8 mM
	Buffered to pH7.4	
Trace Elements:	EDTA	270 µM
	FeCl ₃	48 µM
	ZnCl ₂	3 µM
	CuCl ₂	1 µM
	CoCl ₂	1 µM
	H ₃ BO ₃	2.5 µM
	MgSO ₄	1 mM
	CaCl ₂	0.1 mM
Other components:	Glucose	0.5 %
	Thiamine	1 mM
	Ampicillin	100 µg/ml
	$^{15}\text{NH}_4\text{Cl}$	1 µg/ml

Table 4.1 Components of minimal growth medium

PiLE_1.MC58 ΔN24 and PiLE_15 ΔN24 proteins were produced as described in **3.1.3**. All proteins produced were checked for contamination and homogeneity by SDS-PAGE analysis.

4.1.2 ^1H - ^{15}N -HSQC analysis of the Interaction Between PiLE and CD46

Heteronuclear single quantum coherence (HSQC) is a common type of NMR experiment used in protein biochemistry. The basic scheme for this experiment is the transfer of magnetisation from the proton (^1H) to the second nucleus (in this case ^{15}N) via an insensitive nuclei enhanced by polarisation transfer (INEPT) step. A retro-INEPT step transfers the magnetisation back after time delay t_1 and the signal is recorded. The ^1H signal is detected in the directly measured dimension in each experiment, while the chemical shift of ^{15}N is recorded in the indirect dimension which is formed from the series of experiments. HSQC spectra collected of the same protein in different

environments allow you to see which amino acids of the protein are affected by these different conditions.

All NMR experiments conducted were performed on a Bruker Avance II 500 MHz spectrometer at 25 °C by Dr Devon Sheppard. All spectra were processed within the Bruker TopSpin software and subsequently analysed with Sparky.

Spectra were collected of each ^{15}N -labelled deuterated CD46₃₄ protein samples at 200 μM in 25 mM acetate pH 5.5. The spectra showed distinct and dispersed peaks in both dimensions. The central area of the spectrum does have some peak overlap. However, for observing the peak shifts of exposed side chains, this did not hinder subsequent analysis. The sample was diluted 1:1 with a 200 μM PiE ΔN24 protein in 25 mM sodium acetate pH 5.5 solution and after thorough mixing, the spectra were collected again. Both protein samples were dialysed in the same buffer to ensure any shifts observed related directly to protein-protein interactions and not the buffer environment. Two experiments were completed for CD46₃₄ in combination with PiE_1.MC58 ΔN24 and PiE_15 ΔN24 . These two pilins were studied as they represent class I and class II pilins exhibiting stability at high protein concentrations enabling thorough investigation of the putative interaction.

Figure 4.2 shows an example spectrum of ^{15}N -labelled CD46 with and without PiE ΔN24 present in green and red respectively. All peaks overlapped extremely well, and any visible shifts were very slight indicating that there is no interaction between PiE and CD46 under these conditions. The spectra looked identical irrespective of which PiE protein was added.

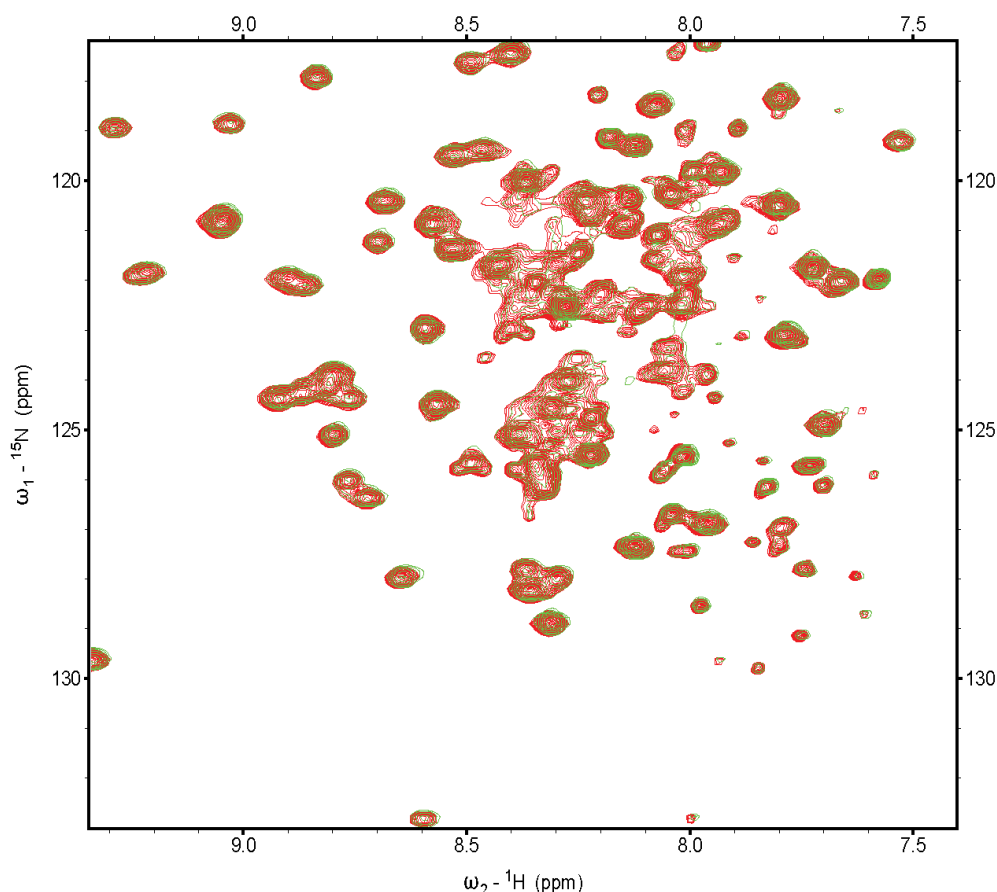


Figure 4.2 Investigating the putative interaction between PiLE and CD46 by $^1\text{H}^{15}\text{N}$ -HSQC

An example $^1\text{H}^{15}\text{N}$ -HSQC spectra for CD46 and PiLE. This spectra depicts ^{15}N -labelled CD46₃₄ alone in red, and in combination with PiLE₁₅ ΔN24 in green. The spectra for all combinations of PiLE ΔN24 and CD46 interaction analysis showed no significant peak shifts.

4.1.3 SPR analysis of the Interaction Between PiLE and CD46

The putative PiLE:CD46 interaction was investigated further by SPR. All experiments were conducted at 25 °C in running buffer of 10 mM HEPES pH7.5, 150 mM NaCl, 1 mM EDTA and 0.005 % Tween-20. The methodology used was the same as previous SPR experiments described in 2.1.2. 25 μM PiLE₁.MC58 ΔN24 and PiLE₁₅ ΔN24 samples were diluted 1:10 with acetate pH 4.5 and amine-coupled to CM5 chips so 1000 RU, 2000 RU and 4000 RU was coupled to flow channels 2,3 and 4 respectively, leaving channel 1 as a reference channel. CD46₃₄ binding was investigated, with a concentration range of 1 μM to 200 μM in each case. Under no conditions was binding observed that gave a measurable response above background. At the highest concentrations of

CD46, responses of 20RU could be observed on the channel with the highest density of PiLE compared to the reference channel (4000 RU), a ratio of 1:200. However, this is likely to be due to the sheer bulk of protein in solution and represent non-specific interactions between the proteins especially as in some cases the response in the control channel exceeded that in the channels in which pilin was bound.

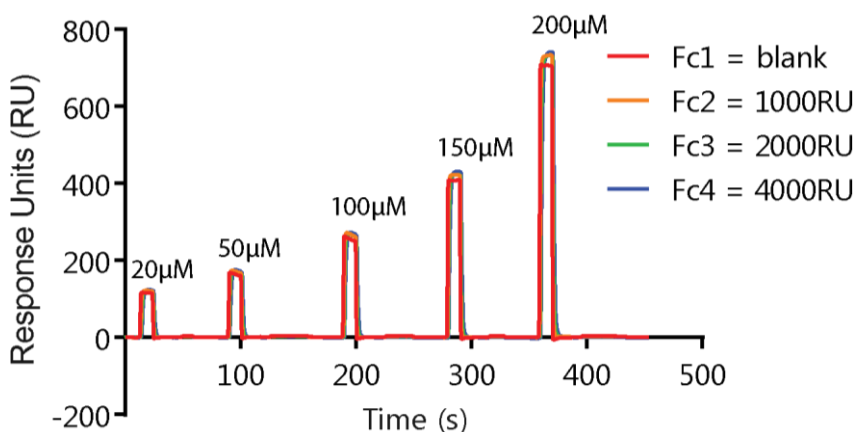


Figure 4.3 Investigating the putative interaction between PiLE and CD46 by SPR

The SPR trace of the CD46₃₄ and PiLE₁₅ ΔN24 is shown, though this is representative of all CD46-PiLE ΔN24 SPR experiments. CD46₃₄ injections are shown at five different concentrations.

4.1.4 Conclusions of PiLE-CD46 Interaction Studies

For the first time, the interaction between PiLE and CD46 has been investigated at the molecular level using recombinant proteins. Previous studies, both in support and against CD46 as the putative pilus receptor, have used adhesion blocking assays with anti-CD46 antibodies as well as soluble CD46, immunoblot analysis and bacterial binding assays with cell lines knocked out for expression of CD46 compared to wild type (Gill and Atkinson, 2004). Such analyses do not reveal anything about the molecular mechanism of this interaction. However, no interaction could be seen between PiLE and CD46 by these biophysical methods.

It could be postulated that if pili do interact with CD46 *in vivo*, then the CD46 molecule on the host cell surface is likely to bind multiple copies of PiLE. As such the interaction would be

dominated by avidity effects rather than the affinity of the individual Pile subunits for CD46. Therefore, when conducting HSQC analysis, we would expect to be investigating an interaction with an especially high K_D . HSQC analysis can be used to detect interactions with K_D values up to 10 mM in some cases. This does however rely on having extremely soluble protein components in order to actually quantify such an interaction. Nonetheless, even at concentrations of protein significantly below the K_D , some small peak shifts on the HSQC spectra would be expected between bound and unbound CD46 samples. Similarly with SPR, although the detection limit for low affinity interactions is not as good as HSQC, it would be expected that a consistent response above the reference channel should be seen, particularly at the higher concentrations. The responses observed were marginal and thought to represent non-specific binding between the two proteins. Both interaction study experiments conducted here are limited as neither includes a positive control for CD46 binding or pilus binding as neither of these controls is available. Therefore the experimental set up may not be correct for characterising this interaction.

T4P are complex structures and adhesion to host cells is likely to be multi-factorial and complex. The results of this study are not overly surprising given the lack of understanding of the precise meningococcal adhesion mechanism. Almost identical experimental procedures have produced contrasting results, which support and negate the role of CD46 in Neisserial adhesion.

Assumptions are made throughout such studies that the gonococcus and meningococcus can be viewed synonymously and that all strains of these species will adhere to host epithelia and endothelia by the same mechanisms. A further discussion of adhesion can be found in **4.5**.

4.2 Investigations into the Effect of PilE Incubation with Host Epithelia

The interaction between *N. meningitidis* and host epithelia in the nasopharynx has long been implicated as the key interaction by which colonisation of this largely commensal bacterium can evolve into the initiation of meningococcal infection in the case of some strains. Adhesion of *Neisseria* and the interaction between T4P and host cell surfaces have been implicated in the result of numerous cellular outcomes including cell signalling and cytotoxicity. The precise molecular mechanisms by which such effects are mediated are poorly characterised and most experimental investigations in this area have been conducted with whole bacteria or purified T4P which make it difficult to resolve the exact molecules involved in such responses.

The close association of bacterial adhesion and PilE sequence, lead to the idea that perhaps different PilE sequences or class I compared to class II PilE proteins would influence host epithelial cell outcomes. Investigation into the role of PilE in influencing human epithelial cells and resolving the role of this protein and sequence differences in the previously described effects could glean insight into the differences between predominantly pathogenic and commensal strains at a biochemical level. This work set out to characterise various markers of host cell response to bacterial adhesion specifically mediated by PilE by incubating human respiratory derived epithelial cells with different recombinant pilin proteins.

4.2.1 Preparation of Protein Samples for Human Epithelial Cell Studies

This study needed to distinguish between the host cell effects caused by the PilE proteins and contaminating LPS which is present in most protein preparations derived from *E. coli*. As such, all samples had to be LPS free. LPS is a potent endotoxin, a biologically active and prevalent component of the outer membrane of Gram negative bacteria. LPS triggers an extensive cellular response through toll-like receptor 4, thus making it difficult to differentiate between the effectors present within a complex sample. As all PilE proteins were produced recombinantly by

over-expression in *E. coli*, they were heavily contaminated with LPS, even after extensive dialysis, dilution and reconcentration to obtain the final protein sample.

PilE proteins were prepared as previously described in **3.1.3**. A control protein, unrelated to PilE but prepared by the same methods from *E. coli* was needed. The control protein used in these assays is MxiGN, the cytoplasmic N-terminal domain of *Shigella flexneri* protein MxiG, a structural component of the type III secretion system which was solubly expressed in *E. coli* and purified by Ni-affinity chromatography and size-exclusion chromatography. MxiGN is located within the C-ring of the type III secretion system, it is not surface exposed and unlikely to adhere to human epithelial surfaces, so constituted a credible control protein for these assays. The final size-exclusion chromatography buffer in all cases was PBS. Samples were concentrated to approximately 1 ml in volume after the final size-exclusion chromatography step using an Amicon Ultra-4 Centrifugal Filter Unit NMWL of 10,000 Da.

To remove the LPS from each protein sample, a 1 ml Pierce High-Capacity Endotoxin Removal Spin Column (Thermo Scientific, Pierce) was prepared according to the manufacturer's instructions and equilibrated with 3 CV of sterile PBS. The 1 ml protein sample was then added to the column and incubated at 4 °C by gentle rocking for 1 hour before removal of the protein solution from the column resin by centrifugation at 500 *xg* for 1 minute.

To verify efficient removal of LPS from the sample, a Toxin Sensor Kit (Genscript) was used to measure the levels of LPS spectrophotometrically, by the limulus amoebocyte lysate assay method. This confirmed that for all protein samples, including the MxiGN control sample, the LPS content was less than 0.5 EU/ml, a satisfactory level at which a very limited effect was seen upon incubation with human epithelial cells according to FDA approved studies.

4.2.2 Human Epithelial Cell Culture

In order to investigate the putative effects of Pile on epithelial cells, robust models of the human respiratory epithelial barrier were required. The cell lines used and their media requirements are described in **Table 4.2**. All cells were adherent and grown in 75 cm², rectangular, canted neck cell culture flasks with vented caps (Costar). Cell lines derived from human nasopharyngeal or lung epithelial cancer isolates, were grown to confluence forming tight monolayers of cells closely resembling the barrier of the nasopharyngeal epithelium *in vivo*.

Cell Line	Derivation	Media for Growth
A549	Human adenocarcinomic alveolar basal epithelial cells	DMEM (Sigma) with 10 % v/v foetal calf serum (Sigma)
Calu-3	Human adenocarcinomic bronchial epithelial cells	DMEM/F-12 with 10 % v/v foetal calf serum
Cor-L23	Marrow trephine of poorly-differentiated squamous cell carcinoma of patient with lung cancer	RPMI 1640 (Sigma) with 2 mM Glutamine (Invitrogen) and 10 % v/v foetal calf serum
Detroit562	Human pharyngeal carcinoma epithelial cells	EMEM (EBSS) + 2 mM glutamine + 1 % non-essential amino acids (Invitrogen) + 1 mM sodium pyruvate (Sigma) + 0.1 % lactalbumin hydrolysate (Sigma) + 10 % FCS

Table 4.2 Cells lines used to study the cytotoxic effects of Pile

Using standard tissue culture methods, the cells were passaged and expanded for all experimental requirements. Briefly, once cells had grown to form a confluent monolayer in each flask, the media was removed and the adherent cell layer washed with 10 ml sterile PBS three times. Cells were incubated with 1 ml Trypsin Triple Express solution (Invitrogen) at 37 °C for two minutes or until no longer adherent. Once all cells were in solution, 9 ml of the appropriate media was added and the 10 ml cellular suspension transferred to a 50 ml tube. Cells were counted by the Trypan blue method by diluting 20 µl cell resuspension 1:5 with 80 µl HyClone Trypan Blue solution (Thermo Scientific) and counting with a haemocytometer (Bright Line) under a light microscope to assess viable cell density. The cell resuspension was then centrifuged at 500 xg for 5 minutes to pellet the cells. The pellet was resuspended in 10 ml fresh media and the cells seeded in a fresh

75 cm² flask at a ratio of 1:3, or approximately 5×10^5 cells/ml by dilution in media to give a total volume of 15 ml per flask, before continuing cell growth by incubation at 37 °C, 5 % CO₂ and 90 % humidity. All aspects of this process were carried out in a laminar flow hood and using aseptic technique under strictly sterile conditions where possible. Cell growth was monitored and cells passaged every 3-4 days dependent on growth rates and the confluency of the cells.

4.2.3 Investigating the Effects of Purified Recombinant PilE on Epithelial Cell Monolayers by Fluorescence Microscopy

Observing cell shape and size by microscopy is a simple method by which confluent monolayers can be examined. Cells which are no longer adherent and have dissociated from the flask surface can be observed by monitoring the number of bare or sparse patches within a previously confluent region of cells in the trial conditions compared to the control, as described previously during initial investigations into the phenomenon of meningococcal-induced cytotoxicity (Dunn et al., 1995).

Initial experiments began by seeding 24-well flat bottomed plates with 1×10^5 cells per well in the appropriate medium in a total volume of 1 ml per well. Cells were seeded onto sterile glass coverslips placed within the plate wells. After incubation for two days, the coverslips were confluent with a dense monolayer of cells. Cells were washed three times by 1 ml fresh media and then finally adding 0.5 ml fresh media. Pilin proteins were diluted with sterile PBS to 100 µg/ml. 2 µl, 4 µl, 6 µl or 8 µl of each pilin was added directly to the media to give final concentrations of approximately 400 ng/ml, 800 ng/ml, 1.6 µg/ml and 3.2 µg/ml respectively of each pilin. For the control experiments, 8 µl sterile PBS or 8 µl of 100 µg/ml MxiGN were added to cells. Cells were incubated at 37 °C for 1 hour, 7 hours or 24 hours. The media was removed from each well and stored at -80 °C for later analysis. Each well was then washed three times with 1 ml sterile PBS. The cells were fixed by treatment with 0.5 ml 3 % paraformaldehyde (PFA) in PBS per well and

incubation for 20 minutes at room temperature. Coverslips were then washed and stored in PBS at 4 °C until required for staining.

It was difficult to determine whether the concentration of PilE added to the epithelial cells was comparable to bacterial challenge experiments. Assuming ~5000 copies of PilE per pilus and approximately 10 pili per bacterium, the highest concentration used was equivalent to adding 10^9 bacteria per well. This is higher than that added in previous *Neisseria* adhesion assays but the concentration range covers standard bacterial titres used. Therefore the concentration range chosen represented conditions under which any host cell effects induced by PilE which occur should be observed.

4',6-diamidino-2-phenylindole (DAPI) staining of DNA of the cells allows visualisation of the nuclear material. This dye is able to penetrate fixed cells and bind AT-rich genetic material inside the cell without disrupting the cellular structure. Phalloidin is able to bind actin fibrils within the cell, again with limited disruption of the cellular morphology. Fluorescent actin conjugates allow visualisation of cellular shapes and overall morphology.

Coverslips washed three times in PBS-Tween 0.05% (v/v) (PBS-T) before blocking for 30 minutes at room temperature in 2% (w/v) BSA PBS-T. Coverslips were then washed three times in PBS-T before incubation with Alexafluor-647-Phalloidin (Molecular Probes – Life Technologies) diluted 1:40 with PBS-T for 45 minutes at room temperature. Coverslips were washed again three times in PBS-Tween 0.05% (v/v) before incubation with DAPI (Molecular Probes – Life Technologies) diluted 1:500 with PBS-T for 20 minutes at room temperature before washing five times with PBS-T and finally three times with ddH₂O. After blotting away all excess liquid, coverslips were mounted cell side down onto glass slides using mounting solution (Sigma) and were then left to dry for an hour before visualising. Imaging was undertaken using a CCD2 Upright Zeiss Light Microscope.

Evidence for pilin induced changes to epithelial monolayers was sought by examining gross cell morphology by brightfield microscopy as well as fluorescent staining for actin and nucleus staining. Images were assessed in a double-blind manner by Hayley Lavender. The cell size, shape and adherence appeared to be indistinguishable under all conditions. In all cases, there were no obvious sparse patches on the coverslips from which cells has dissociated, as may be expected if the PiIE, or the control reactions, had induced any cellular changes on the growing epithelial cells. Although initial experiments trialled all four cell lines, A549 cells readily formed flat monolayers and were therefore imaged with greater ease within just one plane of focus. Images for A549 cells incubated with the class I PiIE_6.8013 Δ N24 and class II PiIE_15 Δ N24 are shown in **Figure 4.4** and are representative of images collected for A549 cells incubated with all other PiIE Δ N24 proteins. Under all conditions tested, there was no substantial visual difference between cells incubated with any of the pilin proteins compared to the control MxiGN or PBS.

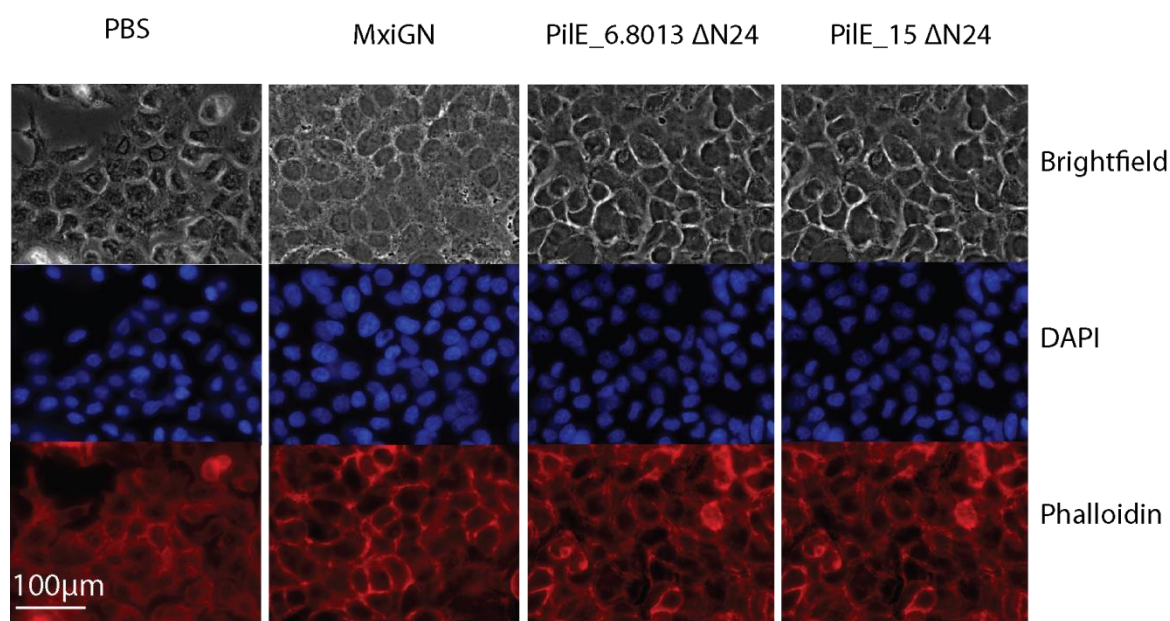


Figure 4.4 Fluorescent microscopy analysis of epithelial cells with PiIE

The following images are at 25x magnification of A549 cells incubated with the highest concentration (3.2 μ g/ml) of each PiIE. Two different PiIE experiments are shown here PiIE_6.8013 Δ N24 (class I) and PiIE_15 Δ N24 (class II). The scale for each image is the same.

4.2.4 Investigating the Effects of Purified Recombinant PilE on Epithelial Cells Lysis

Both the class I and class II pilins tested gave no obvious cytotoxic phenotype compared to the control protein when observed by microscopy. To investigate whether purified PilE proteins could affect cell viability we assayed for lactate dehydrogenase (LDH) release. This normally cytosolic enzyme is released into the external environment upon cell lysis. Cell viability can be assessed by the release of this enzyme into the extracellular environment via a colorimetric assay for the enzyme activity. The LDH-Cytotoxicity Assay Kit II (Abcam) was used for this experiment. LDH released into the external environment reacts with lactate to form NADH which in turn reacts with the WST substrate to form a yellow product, detectable by measuring the absorbance at 450 nm of the reaction solution.

10 μ l media of each condition from the microscopy experiment conducted with Calu-3 cells, as well as mock supernatants in which no cells were grown, were incubated in an optically clear plate (Nunc) with 100 μ l LDH reaction mix (assay buffer containing the WST substrate) at room temperature for 30 minutes. LDH release was quantified by measuring at A_{450} using a SpectraMax M5 (Molecular Devices). Controls were devised as follows. An equivalent number of epithelial cells as those incubated with PilE were resuspended in 0.5 ml media and then lysed by freeze-thaw at -20 °C. The lysate was then clarified by centrifugation and 10 μ l of this solution was used as the positive control. A second positive control using 1 unit of purified LDH resuspended in 10 μ l of PBS was employed to verify the efficacy of the reagents and methodology used. All experiments were repeated in triplicate from the same media. **Figure 4.5** shows representative data of all cell lines and PilE Δ N24 experiments and shows the results of the LDH assay using supernatants from the Calu-3 cells incubated with four different PilE Δ N24.

At all three time points, cells incubated with PilE proteins had equivalent LDH activity, directly proportional to A_{450} , to control cells incubated with PBS or MxiGN. As such, the presence of purified meningococcal pilins can be concluded to have no significant effect upon epithelial cell

viability above that of the control protein and buffer reactions. Control readings taken of media, ddH₂O, PBS and empty wells were lower than the reactions of the supernatants from cells. However, this is most likely due to additional components in the media after incubation with cells which have been excreted by the cells, most likely affecting the pH of the media and therefore its colour due to the phenol red indicator in the media. A small percentage of the epithelial cells would be expected to die within this time frame. This media was also stored at -80°C so may have been altered by this storage process compared to fresh media. The reactions run with lysed cell supernatant and purified LDH showed that the components of the assay were functional and also that cell lysis correlated to increased LDH activity.

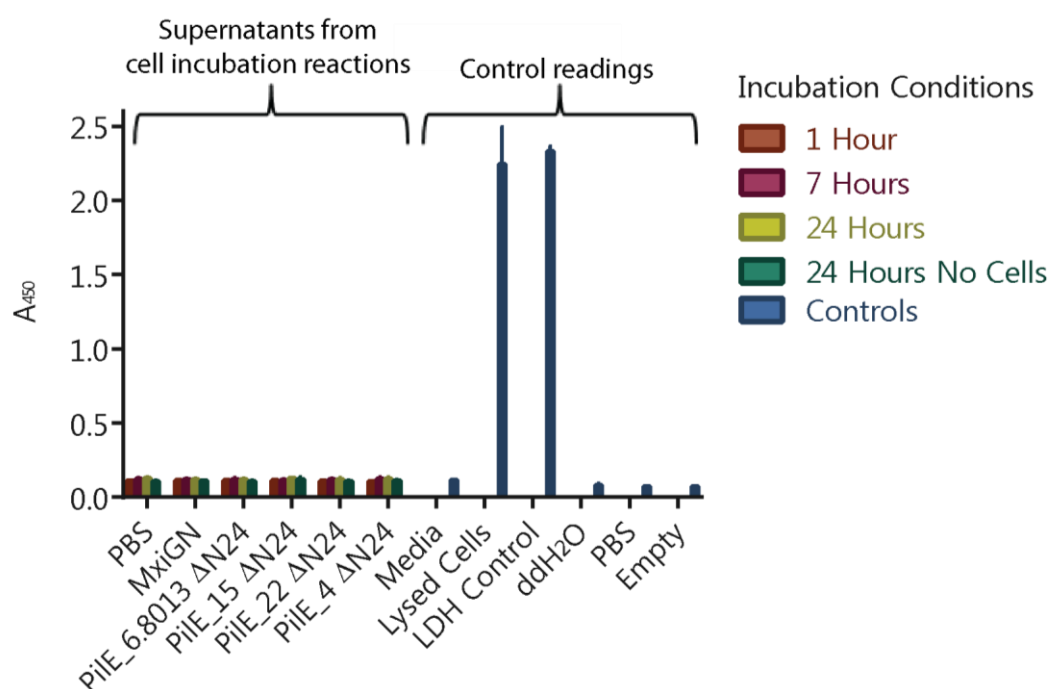


Figure 4.5 LDH activity of supernatants from epithelial cells incubated with PiIE

Results of the LDH assay using supernatants from the Calu-3-PiIE ΔN24 incubation experiments.

4.2.5 Investigating the Effects of Purified Recombinant PiE on Epithelial Cells by Cytokine

Release

As discussed previously, LDH release is a marker of cell death and hence cell viability. However, cellular responses induced by PiE may be more subtle than outright cell death resulting in cell lysis. Cytokine release is often a marker of changes in cellular state. The supernatants used in the LDH assay were also subject to Luminex assay. The 17-plex Human Cytokine Assay Kit (Biorad) allows relative quantification of 17 different cytokines, chemokines and growth factors commonly released by cells upon interaction with extracellular components or in situations of stress: IL-1 β , IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12 (70), IL-13, IL-17, G-CSF, GM-CSF, IFN- γ , MCP-1 (MCAF), MIP-1 β and TNF- α . This system allows detection of very low levels of each of these signaling molecules through antibody coupled magnetic beads and subsequent fluorescence detection in a multi-plex fashion, allowing high throughput analysis of low volume samples using a Luminex machine (Biorad).

In a preliminary experiment, 50 μ l of each supernatant from the Calu-3-PiE Δ N24 incubation experiments was subject to analysis by this system. Only subtle increases in the levels of IL-6 and IL-8 were detected relative to cells incubated with PBS. The changes were statistically insignificant. Such changes most likely represented the epithelial cell response to the minimal amounts of LPS left in the protein samples. Human cells have a low level response to LPS levels as low as 0.02EU/ml so although the protein samples did not contain large levels of LPS there were likely to be responses due to the trace amounts of LPS, which is extremely difficult to remove and indistinguishable from possible effects of PiE.

4.2.6 Investigating the Effects of Purified Recombinant PiE on Polarised Monolayers or Epithelial Cells by Transepithelial Electrical Resistance Assay

To further investigate the potential effects of PiE on the integrity of epithelial barriers, transepithelial electrical resistance (TEER) assays were conducted. Meningococcal traversal of the

monolayer of cells in this *in vitro* system has previously been shown to be via the transcellular route (through cells in the monolayer barrier), this has previously been shown to not disrupt the electrical resistance across the membrane. However, if PiIE causes a cytotoxic effect or reorganizes cellular junctions which would in turn increase the membrane permeability, then this should be measurable by this method.

Transwell filters, which allow the permeation of small molecules but hinder the transition of cells, create a second chamber within 24-well tissue culture plates by which the electrical resistance can be measured across cell monolayers. When the membrane is not coated in a confluent monolayer of cells or the cellular barrier has been disrupted in some way, the lack of barrier or loss of barrier integrity can be measured by the lack of resistance across the membrane, whilst a fully confluent monolayer of cells on the membrane would confer a greater electrical resistance across the two compartments. As such, this binary read-out system made it possible to measure if PiIE was able to disrupt the epithelial cell barrier.

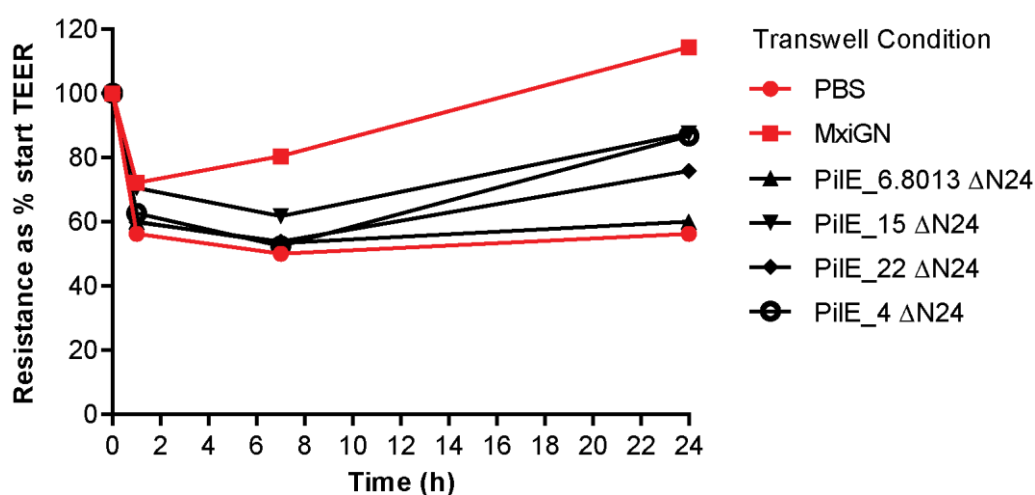


Figure 4.6 TEER analysis of epithelial cells incubated with PiIE
Controls are shown in red.

Calu-3 cells were used in this assay as they had been previously validated to form a monolayer of polarised cells. The cells were seeded within the transwell compartments at 5×10^5 cells in 1 ml with an additional 1.5 ml media in the well surrounding the transwell. Cells were then grown at 37 °C until the resistance measurements showed that the monolayer was confluent. Resistance measurements were made in comparison to cell-free transwells and upon plateauing of the resistance, it was assumed a confluent monolayer of cells had grown. The media was changed every day during this growth period and electrical resistance measurements were made using an epithelial voltohmmeter (EVOM – World Precision Instruments) showed that the resistance would drop by 30-40 % of the initial value within the first hour. Resistance would however recover over the next few hours, indicating that this change was not due to barrier disruption but alteration in the chemical composition of the media or disruption of the equilibration of the system as this was also seen in the cell-free transwells. Prior to introducing PilE to the polarised monolayers, old media was removed and then 1 µg PilE in 1 ml fresh media was added to each transwell, resistance measurements were made after 1 hour, 7 hours and 24 hours.

We observed a characteristic resistance drop initially, due to the change in media composition as described previously. However the resistance values recovered and there was no observable pilin-induced decrease in the electrical resistance of the epithelial monolayer. This indicated that treatment with PilE protein was not sufficient to disrupt the permeability of epithelial monolayers, consistent with the previous experiments described.

4.2.7 Conclusions of Investigations into the Role of PilE in Meningococcal Induced Cell

Responses

Overall, this analysis of the effects of pilin proteins on epithelial cell layers showed that PilE, in its truncated and monomeric form had no obvious effect on respiratory epithelial cells grown *in vitro*. However, this does not mean that pili are not sufficient to induce any of these events. Most likely, the ability to confer adhesiveness by T4P brings host cells into close enough proximity to

meningococci that the LPS released can then act in a cytotoxic way, rather than pili being cytotoxic *per se*. Additionally other cell surface factors may operate in such a manner. It must also be considered that the synonymous approach of host-pathogen interactions of the gonococcus and meningococcus in the literature in many cases, simplifies a very complex system. PilE from different strains may induce different cell responses when presented in the context of the T4P. The approaches used in this study are hampered by the lack of appropriate positive controls in many of the assays, but this was a limitation of the experimental systems used rather than design.

4.3 Investigations into the Immunogenicity of PilE

The immunogenicity of T4P has long been studied, initially with *N. gonorrhoeae* using whole pili as a vaccine candidate (Schoolnik et al., 1984, Forest et al., 1996). However, antigenic variation of PilE reduced the efficacy of this vaccine candidate as the D-region was shown to be the dominant epitope. Previous studies investigating the immunogenicity of recombinant Neisserial pilin proteins, studied *N. gonorrhoeae* MS11 PilE compared to the same pilin protein with the D-region replaced by a short Asp-Glu linker, as seen in the *P. aeruginosa* strain K (PAK) pilin structure (Hansen et al., 2007). This chimeric pilin consisted mainly of conserved protein sequence compared to the hypervariable D-region sequence of MS11 PilE used in the comparison. The study determined that the conserved pilin was immunosilent whilst antibodies were raised against the native PilE protein, again implicating the hypervariable D-region of PilE as crucial to its immunogenic properties.

Therefore, to compare the immunogenicity of class I and class II pilin proteins, mice were immunised with different recombinant *N. meningitidis* PilE proteins to study the effects of the presence of the hypervariable D-region. This study allowed investigation of antibody responses to truncated recombinant meningococcal PilE proteins compared to the chimeric analysis completed in the previous work. All PilE protein reagents were synthesised by myself but the experiments to generate and analyses the sera were devised and conducted by the persons credited throughout this section.

4.3.1 Preparation of PilE Samples for Immunisation and Immunisation Schedule

N. meningitidis PilE Δ N24 samples were prepared as described in 3.1.3 and then treated to remove LPS as described in 4.2.1. PilE_6.8013 Δ N24, PilE_1.MC58 Δ N24, PilE_15 Δ N24, PilE_22 Δ N24, PilE_4 Δ N24 and PilE_3 Δ N24 were all used in the immunisation experiment.

Antigen formulation for immunisation was completed by Dr Rachel Exley, University of Oxford. Each mouse was subcutaneously inoculated with 20 μ g of the appropriate pilin in PBS

supplemented with 10mM histidine hydrochloride with a final concentration of NaCl of 150 mM in a total volume of 200µl. Alhydrogel (Invitrogen) was used as an adjuvant at a final concentration of 3mg/ml.

Female Balb-C mice aged 13-14 weeks were inoculated with three doses of PiE on day 0, 21 and 35. On day 49, the experiment was terminated and the mice were bled. Within each group of 10 mice, nine were inoculated with the PiE antigen and one was inoculated with PBS as a control. Blood was pooled from the groups of nine mice which had been inoculated with the same antigen, but the control blood was kept separately. All animal work was completed by the PSB Animal House, University of Oxford, in line with all Home Office procedures.

The blood was allowed to coagulate at room temperature for 2 hours before harvesting the sera by low-speed centrifugation at 4°C. The sera were stored in 100 µl aliquots at -80°C.

Immunisation experiments were repeated three times with PiE proteins, independently prepared. The first immunisation investigated PiE_6.8013 ΔN24, PiE_15 ΔN24, PiE_22 ΔN24 and PiE_4 ΔN24 whilst the final two rounds of immunisation included PiE_1.MC58 and PiE_3 ΔN24 as well.

4.3.2 Analysis of the Reactivity of Sera Raised Against PiE

Sera were checked for specificity against the homologous PiE antigens by ELISAs run by Dr Rachel Exley, University of Oxford, see **Figure 4.7**. This series of experiments showed that all sera raised against PiE antigens, contained antibodies which recognised the homologous PiE protein immobilised on the surface of the ELISA plate. No immunosilence was seen for sera raised against class II PiE proteins which not only have a D-region of reduced size, the major immunodominant epitope, as may have been predicted from previously published data (Hansen et al., 2007).

To further investigate the cross-reactivity of the sera raised towards heterologous PiE proteins, western blots were run by Dr Rachel Exley, University of Oxford, see **Figure 4.8**. This showed that

sera raised against class I PilE antigens could recognise other class I antigens and similarly so for class II sera. However, class I sera did not recognise the class II pilins as well by this method and the class II sera did not crossreact with class I pilins.

Errors in loading the cell lysate samples or exact sera dilutions could alter the level of cross-reactivity observed. Additionally, the extent of piliation or PilE expression was not measured to ensure this was consistent across all six strains. A lot of background staining can be seen in some cases indicating a lack of specificity of these antibodies, especially in the case of anti-PilE₁₅.

It can be concluded that the sera raised against class I pilins and class II pilins both contain antibodies which bind at sequences or structures unique to the class. Given the key differences in structure between the two pilin classes are the D-region, it is likely that the antibodies were raised against this region in both cases.

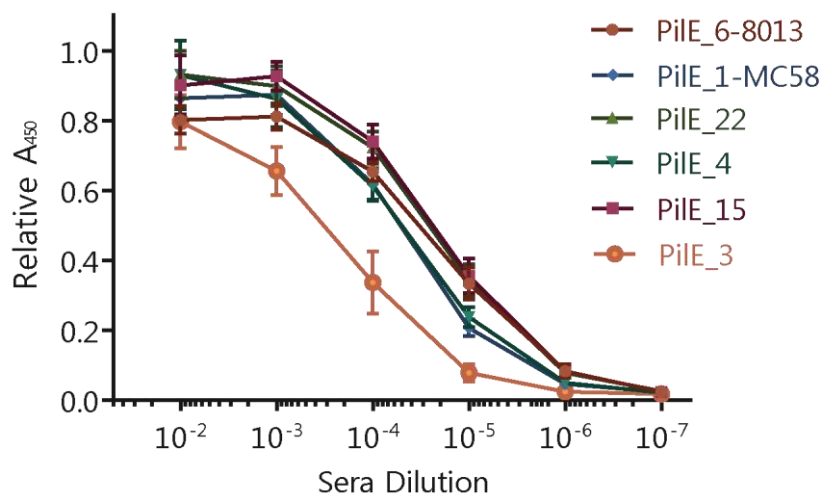


Figure 4.7 ELISA analysis of homologous recognition of PilE sera

The panel of sera displayed cross reactivity patterns against the two classes of pilin which suggested that the D-region in each case is the immunogenic epitope to which the antibodies are raised in both cases.

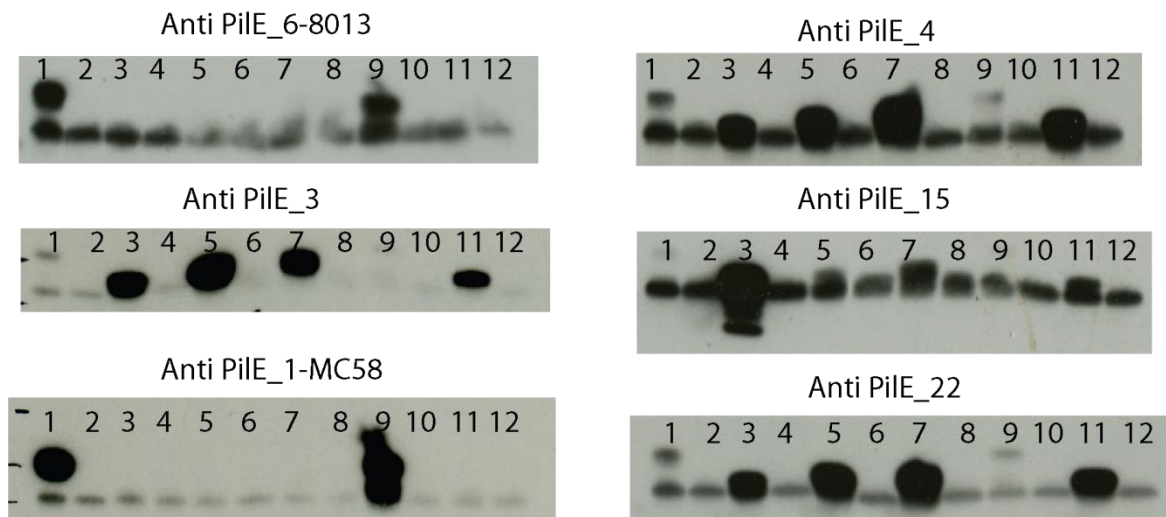


Figure 4.8 Western blot analysis of *PilE* sera cross-reactivity

Cell lysates of the following strains of *N. meningitidis* were loaded as follows: lane 1 – PilE_6-8013, lane 2 - PilE_6-8013 Δ PilE, lane 3 – PilE_15, lane 4 - PilE_15 Δ PilE, lane 5 – PilE_22, lane 6 - PilE_22 Δ PilE, lane 7 – PilE_4, lane 8 - PilE_4 Δ PilE, lane 9 – PilE_1-MC58, lane 10 - PilE_1-MC58 Δ PilE, lane 11 – PilE_3 and lane 12 - PilE_3 Δ PilE. In nearly all blots the gel front cross-reacts with the sera giving rise to the fainter bands seen in every lane.

4.4 Investigations into the Functional Qualities of Sera Raised against PiIE

Given the unexpected cross-reactivity of the sera generated, PiIE could represent a good vaccine candidate and the sera may also be a useful biological reagent. The ability of sera raised against PiIE to confer serum bactericidal activity (SBA) and also blocking of *N. meningitidis* adhesion to human respiratory epithelia were investigated.

SBA test whether the antibodies within the sera are able to recruit various factors from the complement system, to form the membrane attack complex (MAC) which has bactericidal activity by initiating cell lysis following formation of a membrane spanning pore. Exogenous complement is added to each reaction to standardise complement recruitment. Formation of MAC is crucial in human immunity against Neisserial strains as individuals with abnormalities in their genes encoding components of the complement system, are more vulnerable to Neisserial infection than normal individuals. The assay compares bacterial titres before and after incubation with serum at a range of different dilutions with a set amount of complement present.

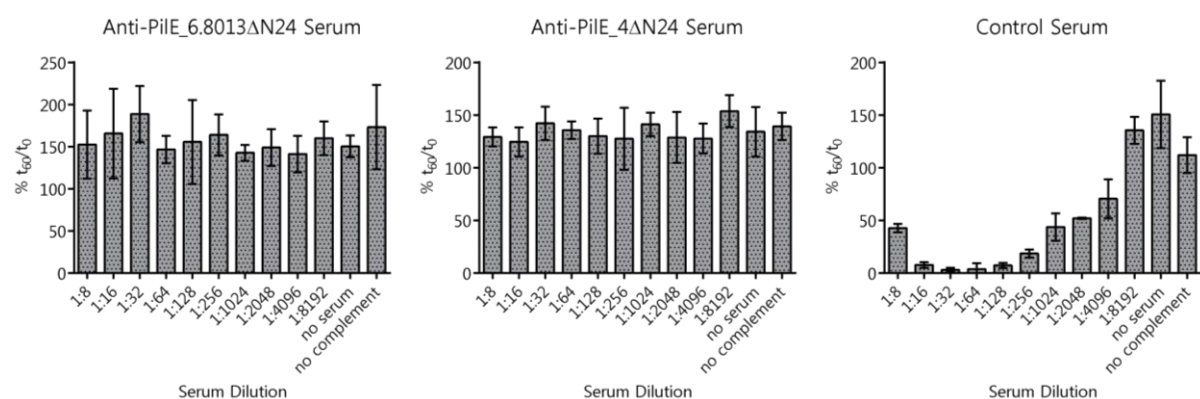


Figure 4.9 SBA conducted with anti-class I PiIE and anti-class II PiIE sera

The %t₆₀/t₀ shows the percentage bacteria surviving after 1 hour. Reactions run with no sera or complement are used as controls. Positive control SBA assay with human anti-capsule sera shows strong SBA compared to the anti-PiIE sera.

The SBAs conducted by Dr Rachel Exley, University of Oxford, for the sera raised against the class I PiIE_6.8013 ΔN24 and class II PiIE_4 ΔN24, show that neither of the sera at any dilution have SBA as seen by the lack of reduction in bacterial survival, compared to the control human sera known

to elicit a bactericidal response, see **Figure 4.9**. Despite these antibodies demonstrating good recognition of the pilin proteins, they were unable to recruit the necessary complement components to initiate MAC formation to mediate bacterial lysis.

Assays investigating the ability of anti-pilin antibodies to block bacterial interaction with human cells compare adhesion levels of bacteria to human epithelial cell monolayers following preincubation with PiE sera, control sera or PBS. In the case of the class I PiE_6.8013 Δ N24 and class II PiE_4 Δ N24 generated sera, no statistically significant drop in adhesion was seen in either case, see **Figure 4.10**. The adhesion which was taking place in the assay by the *N. meningitidis* was shown to be T4P dependent by comparison with the relevant PiE knock out mutant which showed very limited adhesion. This result could be because the anti-pilin antibodies in the sera do not bind pilin in the folded state or when assembled into T4P as this has yet to be demonstrated. Another explanation is that with thousands of copies of PiE within the pilus structure, a very large number of antibodies would be required to cover a large proportion of the pilus surface to prevent PiE exposure and hence adhesion.

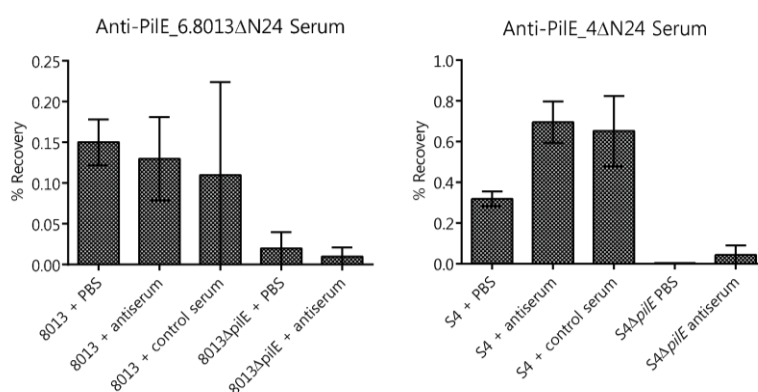


Figure 4.10 Adhesion blocking assay conducted with anti-class I PiE and anti-class II PiE sera

By these two methods, the antibodies generated against the recombinant PiE proteins lack properties which would render them interesting as vaccine candidates or useful reagents beyond

recognition of PiE by ELISA or Western blot analysis though investigations into their binding properties of PiE assembled into T4P would be interesting.

4.5 Perspectives

During these investigations into the functional properties of PilE, a series of novel conclusions can be drawn. It was confirmed that no interaction can be seen between CD46 and monomeric PilE as determined by HSQC and SPR analysis using purified recombinant proteins. Very recently, Bernard *et al* (2014) reported that CD147 is in fact the pilus receptor (Bernard et al., 2014). This immunoglobulin membrane-bound receptor has been shown to bind PilE and PilV by a series of methods in this recent study including fluorescent microscopy, competition assays and atomic force microscopy. Although SPR experiments were conducted to investigate the interaction with monomeric PilE proteins and CD147, no binding was seen. Similarly to the work completed in this investigation with CD46, this was reported to be due to the fact that the postulated interaction would rely more on avidity effects, rather than a high affinity interaction between these protein molecules.

To verify the assertion that either CD147 or CD46 is in fact the pilus receptor, SPR experiments should be conducted using whole sheared pili coupled to the chip surface, where CD147 or CD46 is flowed over the top at a number of different concentrations. This will potentially confirm binding if the interaction is not involved with PilE avidity effects but rather a binding site consisting of adjacent PilE monomers assembled in the shaft. Experiments should also use pili from class I and class II PilE expressing strains as only class I strains were investigated in this study. Additionally, until this point, the minor pilin PilV was only ever implicated in adhesion in its association and presentation with PilC, not as an adhesin itself *per se* (Winther-Larsen et al., 2001). However, given the experiments which support CD147 as the pilus receptor are comparable to those conducted to show that CD46 was the pilus receptor (Johansson et al., 2003); significant proof will be needed from multiple research groups within the field before this candidate receptor is accepted as being the true pilus adhesion receptor.

Given the highly charged nature of the surface of the pili (Craig et al., 2006a) and the extensive number of subunits within each pilus fibre, perhaps the avidity effects of binding normal host cells surfaces, covered in many different types of sugars and proteins is sufficient to confer adherence to the bacteria. Whilst *N. meningitidis* are known to specifically colonise the human nasopharynx, this preference may be driven by successful immune evasion strategies specifically targeted at the human immune system in this physiological niche rather than pilus binding specificity for a specialised type of epithelia. Binding specificity may perhaps be conferred through other surface or pilus associated proteins, for example the tip-associated adhesin PilC, whilst “adhesiveness” may be increased by subtle changes in PilE sequence which have a bulk effect on pilus surface chemistry. PilE sequence may also affect bundling of adjacent pili, a phenotype known to increase adhesiveness of strains. Given the variation of PilE sequences known to be expressed by adhesive meningococcal isolates, a more general mechanism such as this, seems far more likely. This was originally suggested by Nassif *et al* in 1997 (Nassif et al., 1997) but has not been followed up due to a series of technical issues conducting the experiments necessary to support this idea, including producing and characterising PilC *in vitro*.

The analysis of human epithelial monolayers incubated with PilE also conclusively showed that monomeric recombinant PilE Δ N24 is not sufficient to impart any cellular response measurable by the methods used. Whilst pili may be able to induce some cell responses (Kallstrom et al., 1998), what is more likely in the case of meningococcal disease, is that bacteria which are piliated are in closer contact to the host epithelium. As such, any chemicals or proteins which are released from the bacteria or are surface exposed which do induce a response from host cells, are more intimately associated with the epithelial surface and therefore trigger a greater reaction. After completion of this work, a paper was published which shows the hypervariable region has a role in modulating endothelial cell responses in the context of whole strains (Miller et al., 2014a). However, this paper did not take into account the other factors which are likely to be different across strains so attributing the complete difference in phenotype to just the D-region sequence,

is an over-interpretation of these results. Repeating this series of experiments with purified pili would validate this hypothesis further.

PilE proteins were shown to be immunogenic though the sera raised against the different PilE classes did not have the same cross-reactivity. Pilin proteins lacking a D-region were previously shown to be immunosilent by Hansen *et al* (2007) (Hansen et al., 2007). However the sera raised against class II PilE proteins not only cross-reacted with its homologous pilin protein but also with all other pilin proteins tested, indicating that the sera specificity was most likely for a conserved epitope within the pilin protein. Sera raised against class I PilE contained antibodies most likely targeted against a sequence within the D-region as no cross-reactivity was seen with any of the class II pilin proteins tested. This is consistent with the previous hypothesis that the hypervariable D-region is the immunodominant epitope of PilE (Forest et al., 1996, Schoolnik et al., 1984). None of the sera raised had any activity in terms of SBA or ability to block adhesion. One limitation of this work is that the ability of the anti-pilin antibodies to bind whole T4P has not been verified, if this is not the case then a lack of adhesion blocking would be plausible. As such, the PilE would not constitute a good vaccine candidate and the anti-pilin sera has limited therapeutic potential other than diagnostic use for pilin detection. It must be borne in mind that the sera was generated in mice, a species which is not susceptible to meningococcal infection. However, previous meningococcal vaccine candidates such as fHbp have been used in such experiments to generate sera with SBA (Seib et al., 2009).

Hypervariability of PilE sequences has often been attributed as a mechanism of host immune escape. However, it has now been shown that antibodies can be generated against both class I and class II PilE proteins, although these antibodies do not illicit a SBA response. In the case of class II PilE, it may be hypothesised that one reason the sequence does not vary is due to a lack of pressure from the host immune system as only non-functional antibodies are generated against this antigen. This may explain why class I and class II PilE expressing strains may be equivalently

virulent, as PilE sequence hypervariability is not involved in evasion of host immune responses but perhaps in modulation of adhesion and host epithelia and endothelia response modulation. For the case of both classes of PilE protein, having the immunodominant D-region epitope against which non-functional antibodies are generated, could be considered a novel immune evasion strategy and this may explain why *Neisseria* release S-pilin within the host to deflect the host immune system. However, given the sera was raised in mice, validation of these findings with human sera for this hypothesis would be essential.

5. Expression and Purification of PilC

Despite being described as an important protein involved in pilus biogenesis in the literature over 20 years ago (Rudel et al., 1992, Nassif et al., 1994), PilC remains largely uncharacterised.

Controversy surrounds the location of this protein in the type IV pilus (T4P) assembly, with some reports suggesting it is located at the T4P tip, others describing it as an integral membrane protein and the remainder reporting both locations to be true (Rudel et al., 1995a, Rudel et al., 1995c).

PilC has been shown to be important in the T4P-dependent adhesion of *N. meningitidis* to host epithelium (Nassif et al., 1994, Rudel et al., 1995c, Scheuerpflug et al., 1999, Morand et al., 2001).

Whether it performs this function as the tip-associated adhesin or by modulating T4P structure or extension is unclear. As no specific pilus receptor has been correctly identified, no interaction studies can be performed with PilC to clarify this problem. Studies show that the N- and C-terminal halves of the protein have different functions, with the N-terminal domain playing a role in adhesion and the C-terminal domain involved in pilus biogenesis. Given the presumption that any protein involved in pilus biogenesis would be expected to function within the cytoplasm whilst an adhesin could only operate in the extracellular environment, it remains to be elucidated how these seemingly contradictory functions can occur simultaneously.

More recently, the structure of the C-terminal domain of PilY1, the PilC homologue in *Pseudomonas aeruginosa*, was published (Orans et al., 2010). The 2.1Å resolution structure highlights a specific region within the modified β -propeller structure which shows an EF-hand-like calcium binding site (PDB ID: 3HX6). Disruption of this site affects twitching motility of pili *in vivo* and indicates that pilus extension and retraction are controlled through the binding and release of calcium by PilC.

This publication provided a good starting point to begin investigating PilC from *N. meningitidis*. It was decided that using the domain boundaries defined by this paper, attempts would be made to express, purify and characterise the C-terminal domain of *N. meningitidis* from FAM18.

5.1 Attempts to Express and Purify PilC using the PilY1 Protocol

To add further uncertainty to our understanding of PilC, two isoforms are present in certain strains of some species with T4P. However, poor genome annotation of strains less commonly studied in the laboratory as well as the high divergency of *pilC* sequences means that many *pilC* genes have simply been annotated as open-reading frames in some genomes. The FAM18 strain was chosen for study for this project as collaborators have the necessary *in vivo* mutants in this strain to perform complementary functional studies with live *N. meningitidis*. Both isoforms were initially investigated in this work.

5.1.1 Sequence Analysis of PilC compared to PilY1

The different PilC isoforms from the same strain are generally of high sequence identity in their C-terminal domain sequence (greater than 75%) whilst the N-terminal domain sequence is far more divergent. In the case of FAM18, the sequence identity of the C-terminal domains of these proteins is 89%, whilst the N-terminal domains share just 59% sequence identity, see **Figure 5.1**.

N. meningitidis FAM18 PilC1	PKYSQRYRIRENGNNGKKDLGDIVNSPIVAVGGYLATSandGMVHIFKQSGGDKRGYNLK
N. meningitidis FAM18 PilC2	PKYSQRYSRDNGKH-ERNLGDIVNSPIVAVGGYLATSandGMVHIFKQSGGDKRSYNLK ***** : * : * . : *
N. meningitidis FAM18 PilC1	LSYIPGTMPRKNIENNDS TLAKELRAFAEKGYGVDRYGVDGGFVLRRITDDQDKQHFFM
N. meningitidis FAM18 PilC2	LSYIPGTMPRKDQIQTSTLAKELRAFAEKS YVGDRYGVDGGFVLRKVER--NGKDHVF M ***** : * : * . : *
N. meningitidis FAM18 PilC1	FGAMGFGGRGAYALDLTKADDNDP--TKASLFDVKDGNNGNNNGNRVELGYTVGTPQIG
N. meningitidis FAM18 PilC2	FGAMGFGGRGAYALDSL KIDSGNGNLADVS LFDVKHD----KNGNNGVKLG YTVGTPQIG ***** : * : * . : *
N. meningitidis FAM18 PilC1	KTHNGKYAAFLASGYATKKIDDPTNK TALYVYDLENNNGTPIAKIEVPNGKGGLSSPTLV
N. meningitidis FAM18 PilC2	KTHDGKYAAFLASGYATK DITSGDNKTALYVYDLES-SGT LIKKIEVPNGKGGLSPTLV ***** : * : * . : *
N. meningitidis FAM18 PilC1	DKLDLGTVDIAIAGDRGNGMYRFDLSNNDP TKWSVRTIFKGT PDKPITSAPAVSKLKDKR
N. meningitidis FAM18 PilC2	DKLDLGTVDIAIAGDRGNGMYRFDLSNNDP TKWSVRTIFKGT PDKPITSAPAVSKLKDKR
N. meningitidis FAM18 PilC1	VVIFGTGSDLSEDDVKDDIQSIYGIFDNDTGTDVAEEGQGKG LLEQHLTQEDKTLFLT D
N. meningitidis FAM18 PilC2	VVIFGTGSDLSEDDVKDDIQSIYGIFDNDTGTDVAEEGQGKG LLEQHLTQEDKTLFLT D ***** : * : * . : *
N. meningitidis FAM18 PilC1	YKRSDGSGDKGWVVKLEAGQRVT VKPTVVLRTAFVTIRKYN DGGCGAETA ILGIN TADGG
N. meningitidis FAM18 PilC2	YKRSDGSGDKGWVVKLEAGQRVT VKPTVVLRTAFVTIRKYN DGGCGAETA ILGIN TADGG ***** : * : * . : *
N. meningitidis FAM18 PilC1	KLTKKSARPIVPEANTAVAQYSGHKQTAKGKS IPIGCMWKNNETVC PNGYVYDKPVNVRY
N. meningitidis FAM18 PilC2	KLTKKSARPIVPEANTAVAQYSGHKQTAKGKS IPIGCMWKNNETVC PNGYVYDKPVNVRY
N. meningitidis FAM18 PilC1	LDEKKTTDDFPVTADG DAGSGSTGFKEGKKPARNNRCFS GKGVRTLMMNDLSLDITGPTCG
N. meningitidis FAM18 PilC2	LDEKKTDGFTTADG DAGSGSIDPA GKRSGKNNRCS QKGVRTLMMNDLSLDITGPTCG ***** : * : * . : *
N. meningitidis FAM18 PilC1	MKRISWREIFY
N. meningitidis FAM18 PilC2	MKRISWREVFY *****

Figure 5.1 Alignment of the C-terminal protein sequences *N. meningitidis* FAM18 PilC1 and PilC2

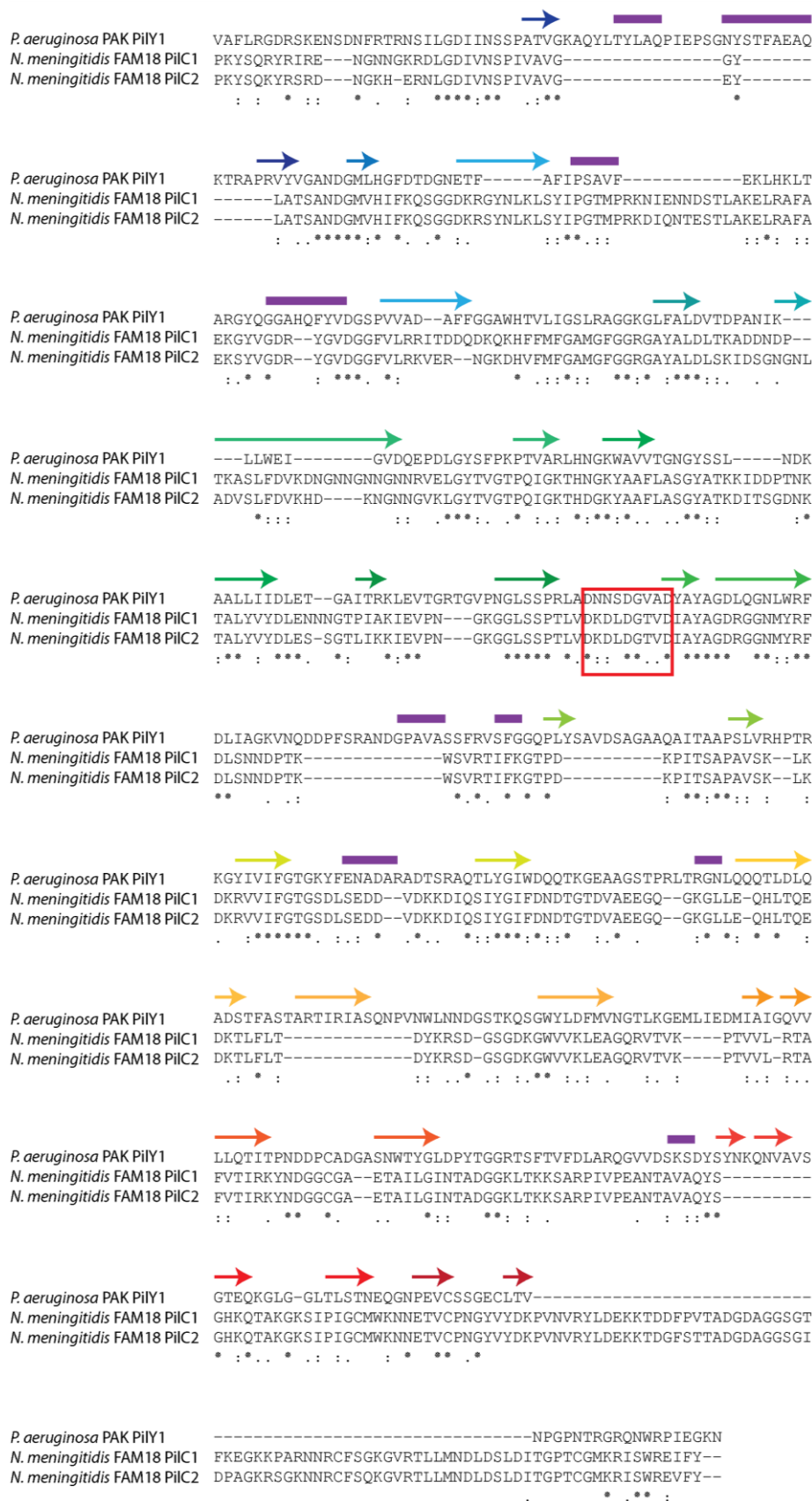


Figure 5.2
Alignment of the
C-terminal
protein
sequences of
Pseudomonas
PAK PilY1 with
N. meningitidis
FAM18 PilC1 and
PilC2

The C-terminal PilY1 sequence shares 30% sequence identity with each of the C-terminal PilC sequences, but only 24% identity across the full-length protein sequence. The DNNSDGVAD sequence of the EF-hand-like domain involved in calcium binding is highlighted with the red box, DKDLGTV D. The residues in bold are those which coordinate the calcium ion. Secondary structure is shown with β -strands as arrows from blue to red (N- to C-terminus) and α -helices as purple blocks.

Alignment of the C-terminal amino acid sequence of PilC1 and PilC2 with PilY1 shows that the sequences only share 30% sequence identity. The regions in which there is stronger conservation include the calcium binding motifs, highlighted in **Figure 5.2**, as well as loop regions extended away from the overall β -propeller structure of PilY1. Mapping all of the conserved sequences between PilC1/PilC2 and PilY1 onto the published structure reveals that they are distributed primarily in the first 300 amino acids of the C-terminal domain along the β -strands, (**Figure 5.3**). Phyre analysis of this PilC sequence indicates that a β -propellor structure is the likely outcome, similar to PilY1, despite only showing good alignment within the first half of the protein sequence (Kelley and Sternberg, 2009).

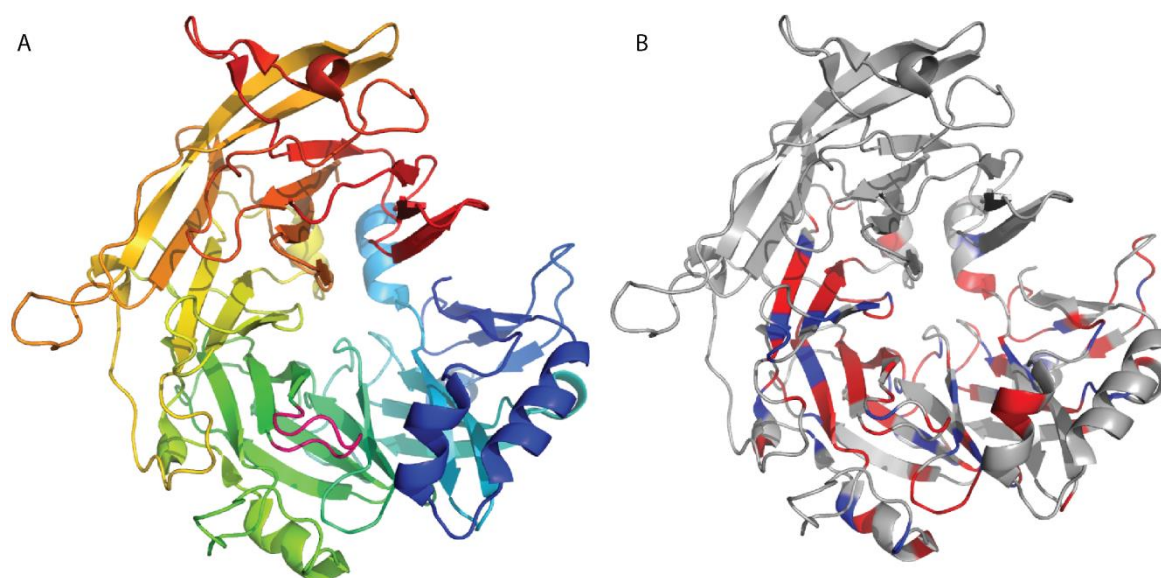


Figure 5.3 Analysis of the regions of common sequence on the PilY1 β -propeller structure

Panel A shows the PilY1 structure (PDB ID: 3HX6) coloured in rainbow blue to red from N- to C-terminus. The calcium binding loop is shown in pink. In panel B, PilY1 is depicted in grey with conserved amino acids between PilY1 and PilC are highlighted in red whilst related amino acids are highlighted in blue. These are all primarily in one half of the structure.

5.1.2 Expression of PilC-C using the PilY1 Protocol

Despite the limited sequence identity between PilC1/PilC2 and PilY1, attempts were initially made to express and purify the meningococcal proteins using the method described for PilY1. PilC1₄₈₉₋₁₀₄₈ and PilC2₄₈₉₋₁₀₃₈ were cloned into the pETM14 expression vector. Over expression in *E. coli*

BL21 (DE3) cells was followed by Ni-affinity chromatography and size-exclusion chromatography to attempt to purify both proteins as per the described method. No protein was visible in the elution trace nor from analysis of the eluted fractions by measuring A_{280} using a Nanodrop ND-1000 (Labtech) to determine protein concentration.

The failure of this protein expression and purification is not completely unsurprising. The PilY1 and PilC proteins are not extremely close in sequence and the PilC proteins contain five cysteines, compared to the two of PilY1 which form a disulphide. Having an odd number of surface exposed cysteines can promote aggregation of protein and expression into inclusion bodies. As such, expression and purification of these proteins required a more complex expression system.

5.2 Development of a Successful Protocol for PilC Expression and Purification

The presence of five cysteines in both the PilC₁₄₈₉₋₁₀₄₈ and PilC₂₄₈₉₋₁₀₃₈ sequences likely makes expression of these proteins within the *E. coli* cytoplasm very difficult given the reducing nature of this environment. There are a number of means by which this can be overcome which are discussed in this next part of the chapter.

5.2.1 Expression in Disulphide Promoting Cell Lines

E. coli expression strains thought to promote the formation of disulphide bonds were trialled. The Origami 2 strain is derived from K12 and has mutations in the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which enhance disulfide bond formation whilst Shuffle cells are knocked out for the *trxB* and *gor* genes and also constitutively express DsbC, a disulphide bond isomerase which again promotes disulphide bond formation in the *trxB* and *gor* knockout background. Expression trials in both of these cell lines, investigating different types of media, IPTG concentrations for induction as well as incubation times and temperatures post-induction, showed no expression of either protein under any conditions studied.

Next the pET32b-Origami B expression system was trialled. This system has previously been shown to allow expression of difficult protein targets. The pET32b vector expresses the protein with an N-terminal thioredoxin tag, cleavable by enterokinase. The thioredoxin tag not only improves the solubility of expressed proteins but also can catalyse disulphide bond formation in a background of *trxB* and *gor* mutants such as the Origami cell lines. Origami B cells are used in preference to Origami 2 cells as this cell line is derived from the Tuner cell line. The Tuner cell line is a *lacZY* mutant of the expression line BL21 allowing precise control of expression levels through subtle modulation of the IPTG concentration. The mutations in the *trxb* and *gor* genes are the same as in the Origami 2 cells but in this background cell line the expression level is supposed to be generally higher. Attempts were made to use the Origami B cells in combination with PilC₁₄₈₉₋₁₀₄₈ and PilC₂₄₈₉₋₁₀₃₈ in pETM14, allowing the potential expression of the his-tagged version of both

of these proteins. Expression trials in this cell lines, investigating different types of media, IPTG concentrations for induction as well as incubation times and temperatures post-induction, showed no expression of either protein under any conditions studied. Therefore investigations using PilC1₄₈₉₋₁₀₄₈ and PilC2₄₈₉₋₁₀₃₈ cloned into the pET32b vector and then expression in Origami B cells was then continued.

5.2.2 Construction of the pET32b-PilC Vectors for both Alleles

The *pilC*-C fragments were cloned into pET32b using “cut and paste” cloning. 3 µg of each of the pETM14-PilC-C vectors and pET32b were digested with NcoI and XhoI (Fermentas) in a ratio of 1 U each enzyme to 1 µg plasmid DNA in 1X Fermentas FastDigest Green Buffer.

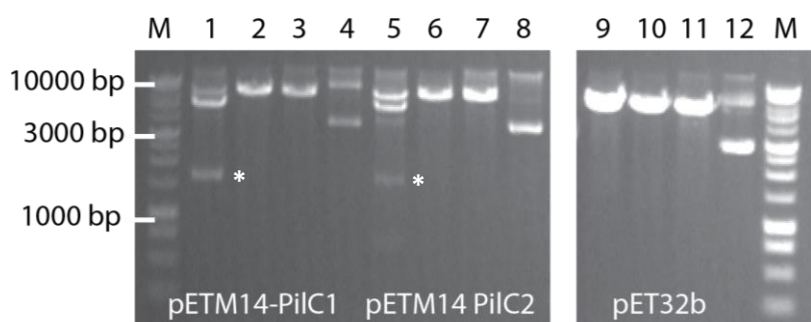


Figure 5.4 Digestion of pETM14-PilC-C vectors and pET32b

All three vectors were digested with both NcoI and XhoI. M corresponds to markers. Lanes 1, 5 and 9 show vector digested with both enzymes; lanes 2, 6 and 10 show the digest reaction with just NcoI; lanes 3, 7 and 11 show the digest reaction with just XhoI; lanes 4, 8 and 12 show the undigested vectors. The excised *pilC*-C fragments are highlighted with an adjacent *.

Linearised pET32b was treated with Fast Alkaline Phosphatase (Fermentas) and purified with QIAquick PCR Purification Kit (Qiagen) though using a Qiagen Miniprep cartridge. Linearised pET32b was eluted with 50 µl EB at 80 ng/µl. The fragments corresponding to the two *pilC*-C genes were gel extracted using the QIAquick Gel Extraction Kit (Qiagen) and eluted with 25 µl buffer EB at 20 ng/µl and 17.5 ng/µl for the *pilC1*-C and *pilC2*-C fragments respectively. The vector was ligated with each of the insert fragments using the Ligafast kit (Promega) and transformed

into DH5 α *E. coli* competent cells. Presence of the correct insert in purified pET32b-PilC-C constructs was verified by sequencing with the T7F and T7R primers (Source Bioscience).

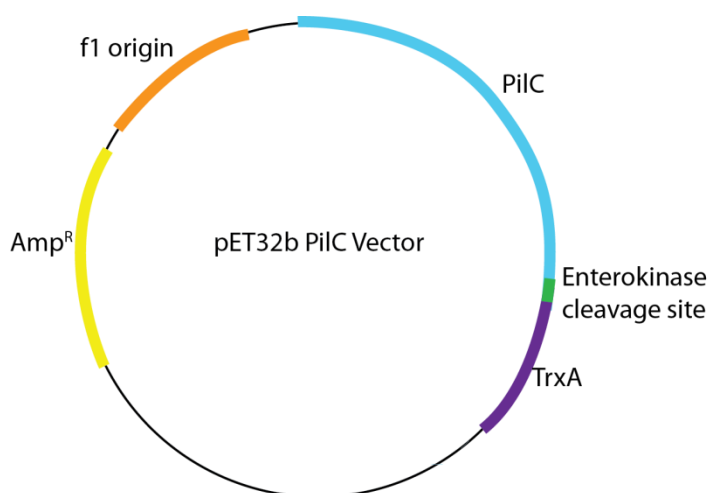


Figure 5.5 pET32b-PilC-C expression construct

5.2.3 Optimising Expression of the pET32b-PilC-C Constructs

To find the optimal conditions for expression of these constructs, high-throughput expression trials were conducted. In the case of both constructs, the conditions seen in **Table 5.1** were investigated.

Variable	Conditions Tested
Cell type	Origami B
Concentration of IPTG for induction upon reaching a cell density where OD ₆₀₀ ~ 0.6	1 mM, 0.5 mM, 0.2 mM or 0.1 mM
Media for growth	TB or LB
Conditions of growth after induction with IPTG	3 hours at 37 °C, overnight at 25 °C, overnight at 21 °C or overnight at 16 °C

Table 5.1 Conditions screened for optimal soluble expression in *E. coli* expression strains

The total number of different condition combinations comes to 32 per construct trialled

5 ml media were inoculated with 50 μ l overnight culture and grown to OD₆₀₀ ~ 0.6 before induction with IPTG. After the cultures had grown at their designated time and temperature, the cells were harvested by centrifugation. Cell pellets were resuspended in 210 μ l PBS supplemented

with 1 % v/v Tween-20, 1 mg/ml lysozyme (Sigma), 400 U/μl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) and then placed on an orbital shaker at 1000 rpm until lysed. After clarification of the lysate by centrifugation, Ni-affinity chromatography was performed using NiNTA beads (Qiagen) which were incubated with clarified lysate at room temperature for 30 minutes. The beads were washed using PBS and 20 mM imidazole and then the protein was eluted with 50 μl PBS and 250 mM imidazole. 25 μl 3X SDS-PAGE loading dye was added to the elution. After boiling, 30 μl of this sample was loaded on a 4-20 % acrylamide gel for SDS-PAGE analysis. The results of this trial can be seen in **Figure 5.6**.

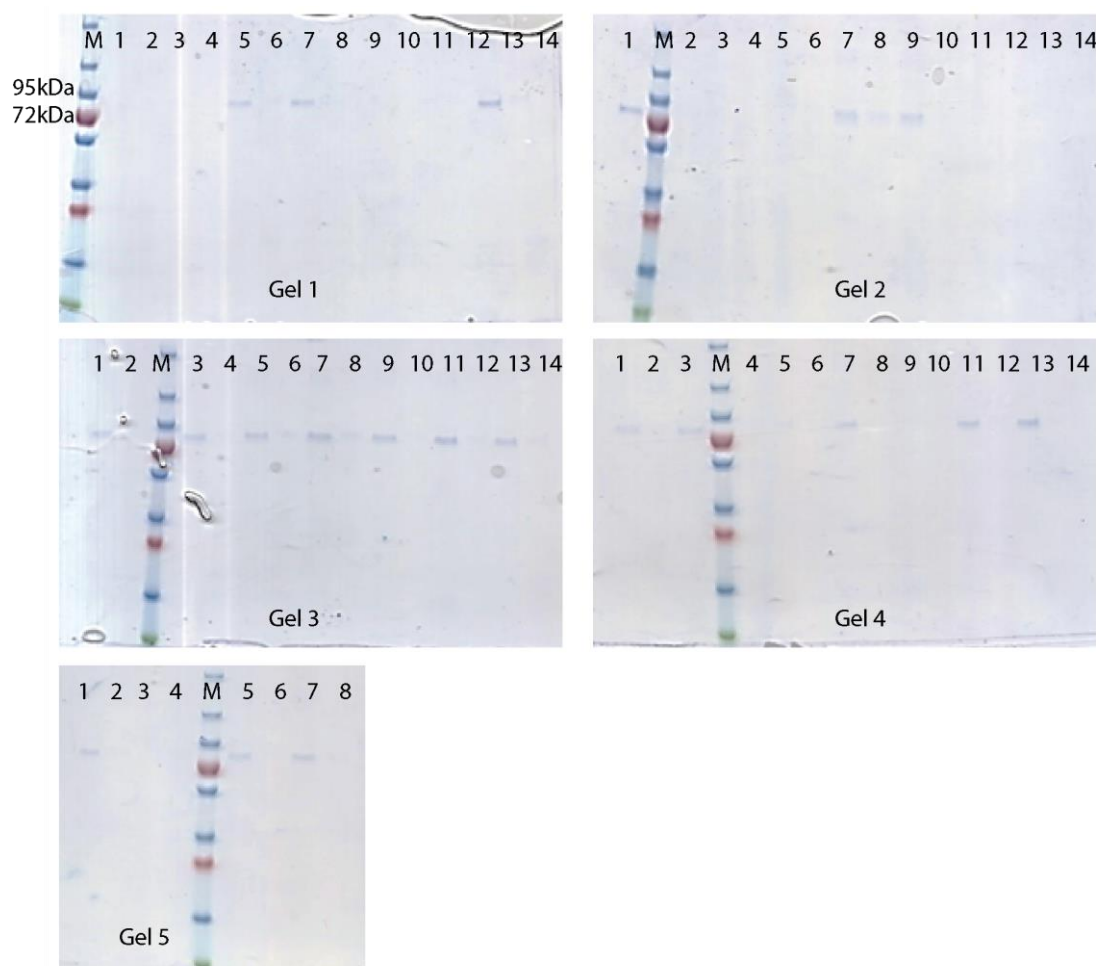


Figure 5.6 Gels showing the relative expression levels of thioredoxin-tagged *PilC1*₄₈₉₋₁₀₄₈ and *PilC2-C*₅₀₅₋₁₀₃₈ in *Origami B* cells in a range of expression conditions. Each lane represents a different growth condition as detailed in **Table 5.1**.

The faint bands seen correspond to expression of thioredoxin-tagged PilC1₄₈₉₋₁₀₄₈.

Given the low expression levels seen, “gentle” conditions were chosen for latter large scale expression using nutrient rich TB media, 0.1 mM IPTG to induce expression and growth over night at 16 °C after induction with IPTG. As PilC1₄₈₉₋₁₀₄₈ shows much greater levels of expression than PilC2-C₅₀₅₋₁₀₃₈ in all conditions trialled, only this construct was optimised further.

5.2.4 Large-Scale Expression and Purification of pET32b- PilC1₄₈₉₋₁₀₄₈

Given the results of the expression trial of the pET32b- PilC1₄₈₉₋₁₀₄₈, the following expression protocol was followed. Overnight cultures of these Origami B transformed with the pET32b- PilC1₄₈₉₋₁₀₄₈ constructs were used to inoculate 2 x 1 l TB each in a ratio of 1:100. Details of the protein sequence can be seen in **Table 5.2**. Cultures were grown at 37 °C until OD₆₀₀ reached approximately 0.6. Cultures were then induced with 0.1 mM IPTG and grown overnight at 16 °C. After cells were harvested by centrifugation, pellets were resuspended in 50 mM MOPS pH 6.5, 150 mM NaCl, 50 mM imidazole which was supplemented with 1 mg/ml lysozyme (Sigma), 400 U/μl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) before homogenisation at 15,000 psi (Emulsiflex 0.5 Avestin). The homogenate contained a large quantity of precipitated protein, indicating the buffer conditions were not optimal for purification of this protein. The lysate was clarified by centrifugation at 26,000 *xg* for 20 minutes before applying the supernatant to a 1 ml NiNTA Superflow Cartridge (Qiagen) at 1 ml/min, pre-equilibrated in resuspension buffer. After a 5 CV wash with resuspension buffer, the protein was eluted in 50 mM MOPS pH 6.5, 150 mM NaCl, 500 mM imidazole over 8 fractions of 1 ml each. Analysis of the eluted fractions by measuring A₂₈₀ using a Nanodrop ND-1000 (Labtech) to determine protein concentration, showed fraction 3 to contain the most protein. As such this fraction was concentrated to 1 ml using an Amicon Ultra-4 Centrifugal Filter Unit NMWL of 50,000 Da. The concentrated protein was then subject to size-exclusion chromatography using an Akta Purifier (GE Healthcare Life Sciences) in a buffer of 50 mM MOPS pH6.5, 150 mM NaCl. The sample was run on an S200 10/300 GL column at 1 ml/min.

After the main peak from the size-exclusion had been pooled, 1 U enterokinase (Novagen) was added per mg of protein in solution, calculated from A₂₈₀ analysis, then the sample was left to digest overnight in size-exclusion buffer at 4°C. Following collection of the flow through of reverse Ni-affinity chromatography using a 1ml NiNTA Superflow Cartridge (Qiagen), the protein sample was concentrated and repurified using a S200 10/300 GL, see **Figure 5.9**.

Construct Name	Amino Acid Sequence of Open Reading Frame	Molecular Weight (Da)	Theoretical pI
PilC1	MSDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPCKMIAPILDE IADEYQGKLTVAKLNIQNPQTAPKYGIRGIPTLLLFKNGEVAATK VGALSKGQLKEFLDANLAGSGSGHMHSHHSSGLVPRGSGM KETAAAKFERQHMDSPDLGTDGDDKAMGKPKYSQRYRIRENGN NGKRDLDIVNSPIVAVGGYLATSANDGMVHIFKQSGGDKRGY NLKLSYIPGTMPRKNIENNDSTLAKELRAFAEKGYVGDRYGVDDGG FVLRITDDQDKQKHFFMFGAMGFGGRGAYALDLTKADDNDPT KASLFDVKDNGNNGNNGNRRVELGYTVGTPQIGKTHNGKYAAF LASGYATKKIDDPTNKTALYVYDLENNNGTPIAKIEVPNGKGLSS PTLVKDLDDGTVDIAYAGDRGGNMYRFDLSNNDPTKWSVRTIFK GTPDKPITSAPAVSKLKDKRVIIFGTGSDLEDDVDKDIQSIYGIF DNDTGTDVAAEEGQGKGLLEQLTQEDKTLFLTDYKRSDGSGDKG WVVKLEAGQRVTVKPTVVLRTAFVTIRKYNDGGCGAETAILGINT ADGGKLTKKARPVPEANTAVAQYSGHKQTAKGKSIPIGCMWK NNETVCPNGYVYDKPVNVRYLDEKKTDDFPVTADGDAGGSGTF KEGKKPARNNRCFSGKGVRTLLMNDLDSLITGPTCGMKRISWR EIFY-	77169.6	7.58

Table 5.2 Expressed sequence of pET32b- PilC1₄₈₉₋₁₀₄₈

This table details the sequence expressed, the molecular weight of that protein product as well as its pI, both calculated by ExPASy ProtParam. The thioredoxin is highlighted in red, the hexa-his tag in blue, the thrombin cleavage sequence in teal, the S-tag in green, the enterokinase cleavage sequence in purple and all additional sequence in orange.

The traces reveal that very little protein is left after enterokinase cleavage and reverse nickel-affinity chromatography. This can be for one of two reasons. Firstly, the enterokinase is very poor at cleaving the PilC1₄₈₉₋₁₀₄₈ from the thioredoxin tag so the protein was removed in the reverse

nickel-affinity stage. Secondly, the resultant PilC₄₈₉₋₁₀₄₈ is very insoluble without the thioredoxin tag in MOPS pH6.5.

To overcome the apparent problems of poor cleavage and solubility, the enterokinase cleavage site was mutated to a 3C cleavage site and the subsequent protein constructs purified in buffers over a range of different pH values.

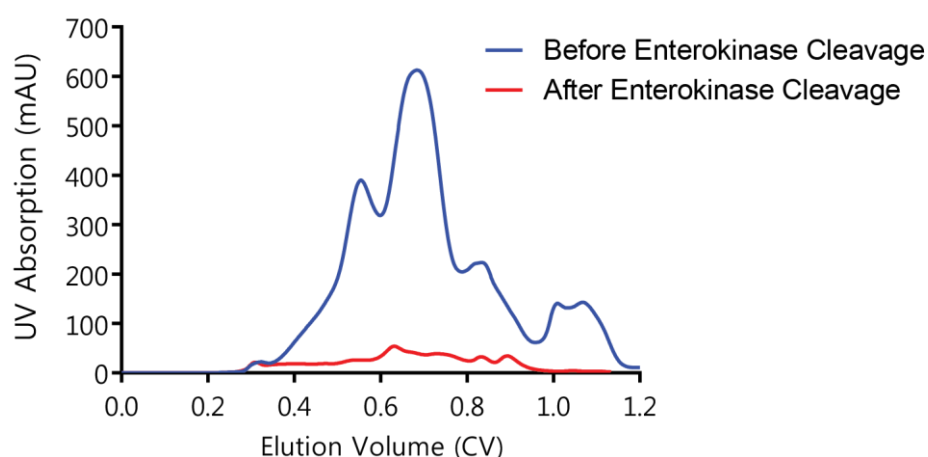


Figure 5.7 Size-exclusion chromatography traces of undigested and digested PilC₄₈₉₋₁₀₄₈

5.2.5 Mutagenesis of the Enterokinase Cleavage site to a 3C Cleavage Site in the pET32b-PilC1-C Construct

Within the N-terminal tag of the pET32b- PilC₄₈₉₋₁₀₄₈ construct, the DDDDK/X cleavage site of enterokinase, was mutated to the LEVLFG/GP cleavage site of 3C. Unlike enterokinase which is a human protein normally expressed in the duodenum containing six disulphide bonds, 3C is derived from the human rhinovirus and has no disulphide bonds so can be easily recombinantly expressed and purified from *E. coli*. Therefore, despite offering similar levels of cleavage specificity, inefficiency can be overcome without great financial cost by just increasing the ratio of 3C:target protein. This sequence change was achieved using the Quikchange mutagenesis kit

(Stratagene) and the primers described in **Table 5.3**. Purified pET32b- PilC₄₈₉₋₁₀₄₈ constructs were checked using T7F and T7R sequencing (Source Bioscience).

Primer	Primer Sequence	T _m /°C
pET32b- PilC ₄₈₉₋₁₀₄₈ Enterokinase to 3C Forward	CCCAGATCTGGGTACCCCTGGAAGTTCTGTTCAGGGGCC CGCCATGGGCAAACCG	51.7
pET32b- PilC ₄₈₉₋₁₀₄₈ Enterokinase to 3C Reverse	CGGTTTGCCCATGGCGGGCCCCTGGAACAGAACTTCCAG GGTACCCAGATCTGGG	

Table 5.3 Primers for mutagenesis of enterokinase cleavage site to a 3C cleavage site in the pET32b- PilC₄₈₉₋₁₀₄₈ construct

The region in each primer corresponding to the 3C sequence is highlighted in blue. The low melting temperature of these primers is due to the large region of non-complementary sequence within the primer.

5.2.6 Optimisation of Buffer Conditions for pET32b- PilC₄₈₉₋₁₀₄₈ Purification

To attempt to improve protein solubility before and after cleavage of the N-terminal tag, parallel purification experiments were run with three different buffer systems; PBS pH 6.5, HEPES pH 7.5 and bicine pH 8.5. Using the protein purification method outlined in section 5.2.4, a cell pellet harvested from 1 l of culture was purified using each buffer system. Samples from each purification were analysed by SDS-PAGE using 4-20 % acrylamide gels. The buffers used are outlined in **Table 5.4**.

Buffer	Resuspension	Wash	Elution
PBS pH 6.5	1 x PBS, 1 mM EDTA, 5% Glycerol, pH6.5	1 x PBS, 1 mM EDTA 20 mM imidazole, pH 6.5	1 x PBS, 1 mM EDTA 500 mM imidazole, pH 6.5
HEPES pH 7.5	20 mM HEPES, 1 mM EDTA, 5 % Glycerol, pH 7.5	20 mM HEPES, 1 mM EDTA, 20 mM imidazole, pH 7.5	20 mM HEPES, 1 mM EDTA, 500 mM imidazole pH 7.5
Bicine pH 8.5	50 mM bicine, 1 mM EDTA, 5 % glycerol, pH 8.5	50 mM bicine, 1 mM EDTA, 20 mM imidazole pH 8.5	50 mM bicine, 1 mM EDTA, 500 mM imidazole pH 8.5

Table 5.4 Buffers used to improve solubility of PilC₄₈₉₋₁₀₄₈ in solution

To improve solubility at the stage of homogenisation, 5% glycerol was added to all resuspension buffers. The resuspension buffers were also supplemented with 1mg/ml lysozyme (Sigma), 400U/μl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) directly before use.

Despite the bicine pH 8.5 buffer system giving the least contaminated nickel-affinity purified protein sample, the stringency of the resuspension buffer and wash buffer was further increased for future purification experiments by including 40 mM imidazole in each and increasing the salt concentration in the wash buffer to 300 mM. Both of these modifications should decrease non-specific binding for the NiNTA resin. When the purification was repeated again with a cell pellet harvested from 1 l culture with the buffers of increased stringency, the eluted protein from the nickel-affinity chromatography step was highly pure. The protein sample was then incubated with 1 mg 3C protease overnight whilst dialysing in 50 mM bicine pH8.5, 150 mM NaCl, 1 mM EDTA. SDS-PAGE analysis showed that the cleavage was successful, see **Figure 5.8**.

The protein produced by this method is pure though higher order disulphide linked multimers are present in the purified protein sample. As such, the dialysis/3C cleavage buffer was modified to contain 1 mM TCEP. This reducing agent has a long half life and it will not disrupt buried disulphides, only maintain free single cysteines in their reduced state. The SDS-PAGE analysis monitoring progress through such a purification with these modifications can be seen in **Figure 5.9**. The purification of PilC₁₄₈₉₋₁₀₄₈ has produced a pure protein sample. Comparison of the molecular weight of the PilC₁₄₈₉₋₁₀₄₈ bands in **Figure 5.9A** and **Figure 5.9C** shows that 3C cleavage has been successful as we observe a drop in molecular weight from 77 kDa to 60 kDa. The larger species seen on size-exclusion chromatography correspond to higher order disulphide linked protein multimers or other forms of self-associating PilC₁₄₈₉₋₁₀₄₈. The process of forming such higher order structure is highly dynamic as can be seen by the redistribution of the peak in **Figure 5.9B** indicated by a * in the elution trace seen in **Figure 5.9D**.

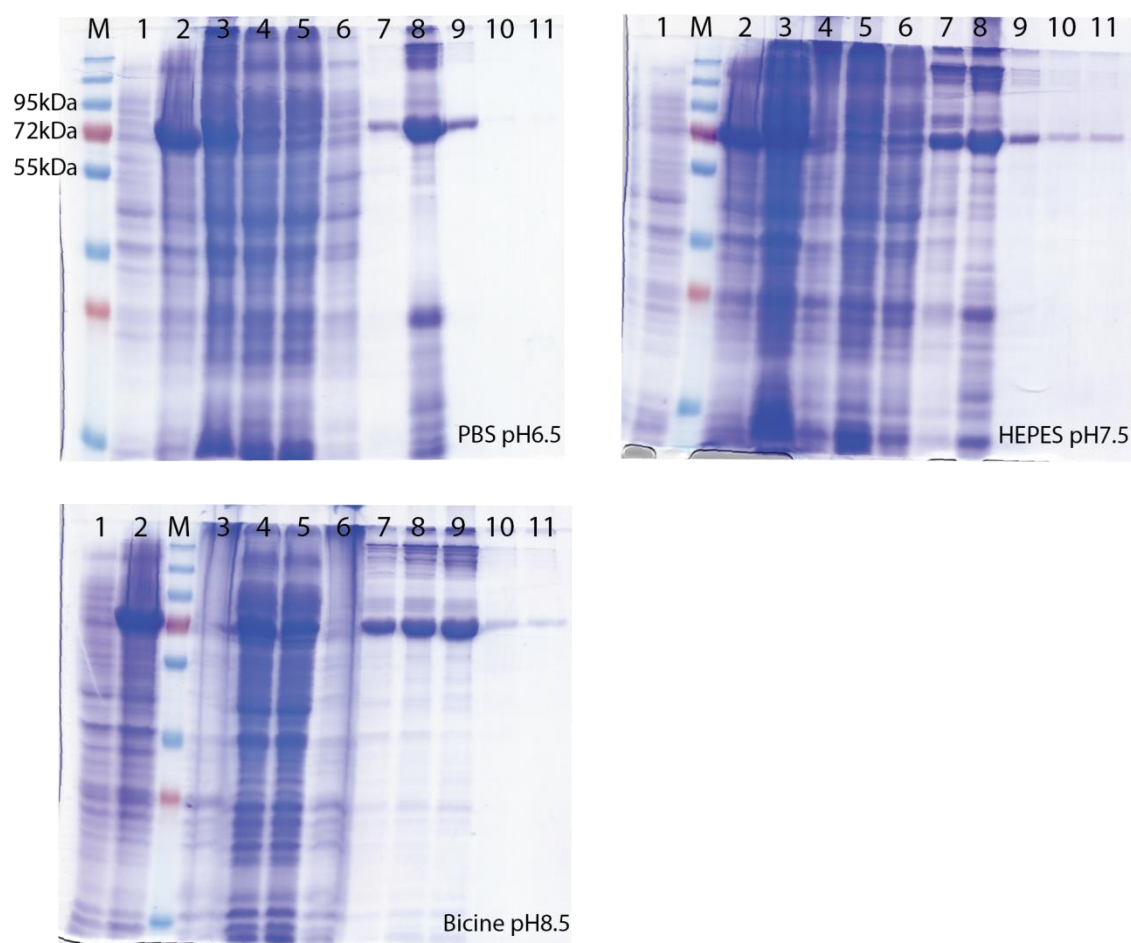


Figure 5.8 Gels showing the purification of pET32b- PilC₄₈₉₋₁₀₄₈ with the three different buffer systems

In each case, 1 - uninduced cells, 2 - induced cells, 3 - homogenate, 4 - homogenate supernatant, 5 - flow through of NiNTA, 6 - wash of NiNTA, 7-11 - NiNTA elution fractions.

This purification method still leads to the formation of many different PilC₄₈₉₋₁₀₄₈ species. To try and reduce the aggregation and multimerisation of the protein sample into these higher order species, the sample was subject to methylation. By methylating primary amines on the protein surface, present on exposed lysine residues, the surface chemistry of the PilC₄₈₉₋₁₀₄₈ protein can be altered. Such a modification can often help reduce self-association or promote crystallisation. Briefly, the fractions corresponding to the monomeric PilC₄₈₉₋₁₀₄₈ protein were pooled and mixed before realiquotting into 1 ml samples in light protecting tubes. To each aliquot 20 μ l 1 M dimethylaminoborane (DMAB) and 40 μ l 1 M formaldehyde were added. The samples were then

rocked at 4 °C for 2 hours. 20 µl 1 M DMAB and 40 µl 1 M formaldehyde were added to the aliquots again and rocked at 4 °C for a further 2 hours. This step was then repeated. Finally, just 10 µl DMAB was added to the samples followed by overnight rocking at 4 °C. The reaction was then quenched by addition of 25 µl 1 M Tris pH 8.5 to each aliquot. The protein sample was then concentrated using Amicon Ultra-4 Centrifugal Filter Unit NMWL of 50,000 Da to 1 ml and run on S200 10/300 GL. As **Figure 5.10 D** shows, despite concentrating, only the monomeric form of the protein remains after concentrating, the redistribution to the various multimeric forms as seen in **Figure 5.9 D** is no longer observed.

Variations in yield between different purifications of PilC1₄₈₉₋₁₀₄₈ were extremely large. Given that the peak of monomerised PilC1₄₈₉₋₁₀₄₈ is still not a normal distribution on the size-exclusion elution trace, there is clearly some lack of homogeneity in the sample. Further to this, despite numerous attempts, no mass spectrometry data could be collected for this protein, native or methylated, possibly due to the intrinsic properties of the protein preventing transition to the gas phase successfully. No further work was conducted to improve the preparation of this protein any further given these restrictions in sample validation and a poor likely outcome of protein crystallisation for subsequent structural investigations.

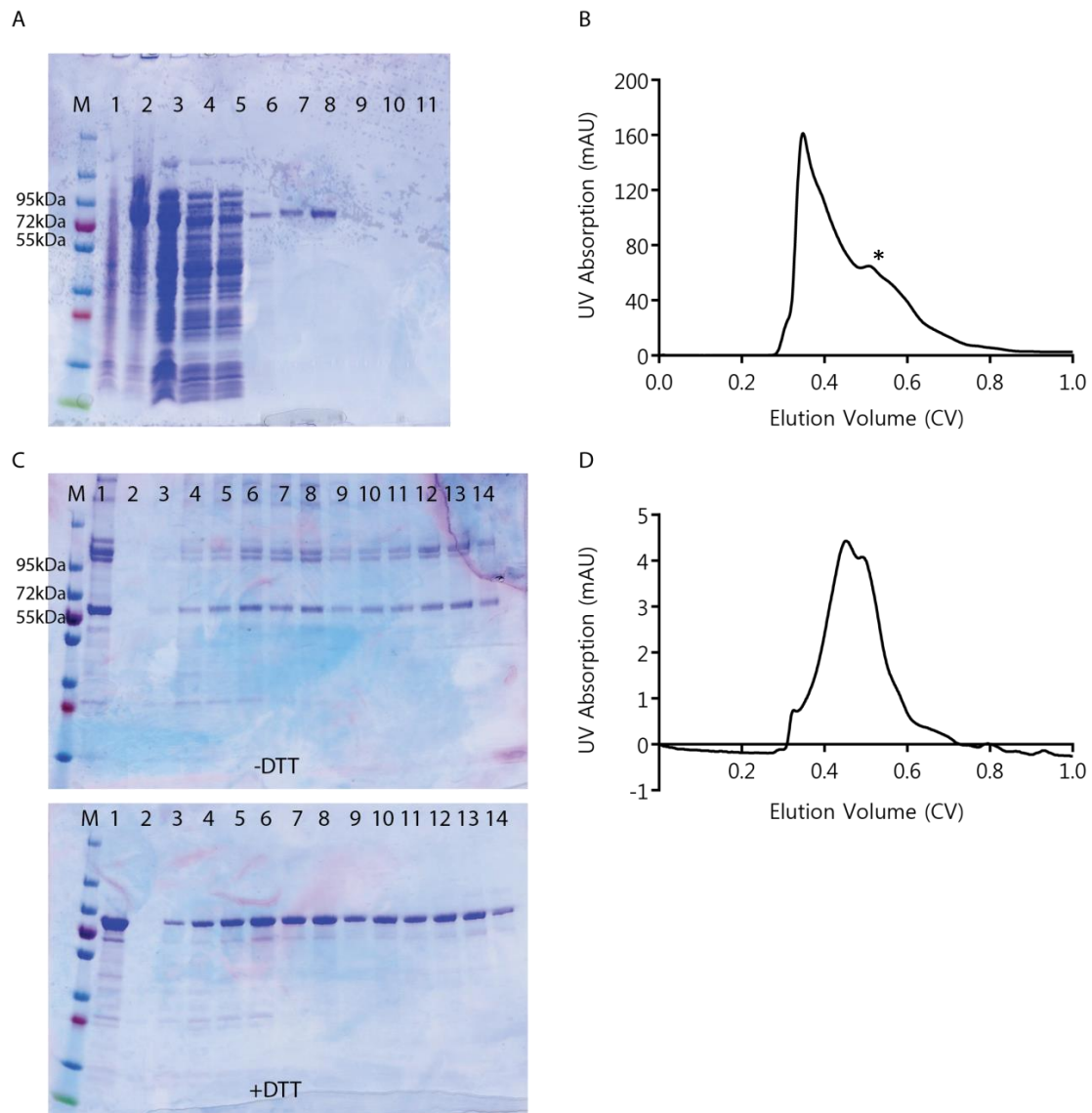


Figure 5.9 Purification of pure *PilC1*₄₈₉₋₁₀₄₈

Panels A and C show the 4-20% acrylamide gel analysis of the purification of *PilC1*₄₈₉₋₁₀₄₈. In panel A 1 - uninduced cells, 2 - induced cells, 3 - homogenate, 4 - homogenate supernatant, 5 - flow through of NiNTA, 6 - wash of NiNTA, 7-11 - NiNTA elution fractions. Panel C shows the size-exclusion chromatography fractions run with or without DTT from 0.3-0.8 CVs of the elution shown in panel B. Panel D shows the re-running of the 1 ml fraction from the size-exclusion chromatography elution indicated with a * on the same S200 10/300 GL column.

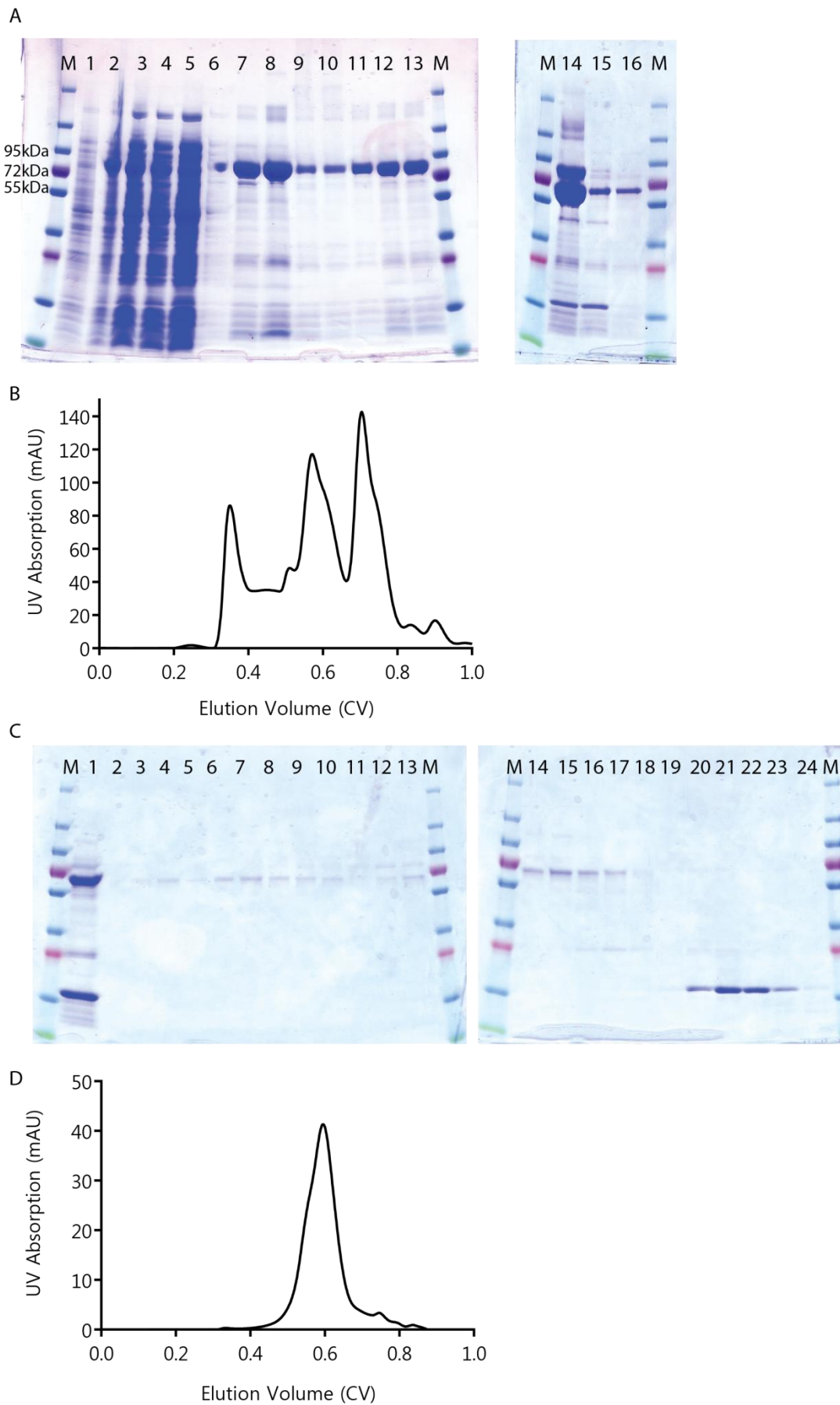


Figure 5.10 Purification of pure monomeric PilC₁₄₈₉₋₁₀₄₈

Figure 5.10 Purification of pure monomeric PilC₁₄₈₉₋₁₀₄₈ - continued

Panels A and C show the 4-20 % acrylamide gel analysis of the purification of PilC-C. In panel A 1 - uninduced cells, 2 - induced cells, 3 - homogenate, 4 - homogenate supernatant, 5 - flow through of NiNTA, 6 - wash of NiNTA, 7-13 - NiNTA elution fractions, 14 – after 3C digest, 15 – filtered digested protein, 16 – after reverse Ni affinity chromatography. Panel B shows the size-exclusion chromatography trace of the reverse Ni-affinity chromatography flow through. Panel C shows the size-exclusion chromatography fractions from 0.3-0.8 CVs of the elution shown in panel B. Panel D shows the size-exclusion trace of the concentrated methylated PilC₁₄₈₉₋₁₀₄₈ sample, now a stable monomer. Fractions 11-13 show some uncleaved PilC₁₄₈₉₋₁₀₄₈, present as a higher molecular weight band on the gel. Fractions 20-24 show residual thioredoxin tag which has not been successfully removed from the sample by reverse Ni affinity chromatography. As such, only fractions 14-17 were used for the subsequent methylation of the sample.

In summary, extensive attempts were made to express and purify the C-terminal domain of PilC1 and PilC2 from *N. meningitidis* FAM18. All methods trialled are summarised in **Table 5.5**.

Method trialled to express and purify the C-terminal domain of PilC from <i>N. meningitidis</i> FAM18	Outcome of method trialled
PilY1 protocol – expression in BL21 <i>E. coli</i> cells, Ni-affinity chromatography followed by size-exclusion chromatography	Failed
Expression in three disulphide promoting <i>E. coli</i> cell lines Origami 2, Origami B and Shuffle, followed by Ni-affinity chromatography	Failed
Expression in Origami B in combination with constructs in pET32b	Protein expressed but heavily contaminated
Optimise buffer conditions for initial purification and cleavage of tag	Contamination reduced and improved solubility of protein sample
Methylation of cleaved protein	Aggregation minimised and stable monomer produced

Table 5.5 A summary of the methods trialled for PilC purification and their respective outcomes

5.3 Perspectives

During this project, PilC₁₄₈₉₋₁₀₄₈ from *N. meningitidis* FAM18, was successfully cloned, expressed and purified in a monomeric form after extensive work trialling different expression systems and optimising the protocol for expression and purification as well as subsequent protein modification by methylation. Despite this, inconsistent yields from this protocol and a lack of evidence to validate the sample produced, ended this project due to potential spurious results in future functional and structural work which would have otherwise have been completed.

The difficulty in producing PilC which is not only pure but of a defined oligomeric state makes elucidating the function and properties of this protein highly elusive. Since the publication of the structure of PilY1 five years ago (Orans et al., 2010), only three further papers have been published detailing brief characterisation of homologous proteins, none of which concerned *N. meningitidis*. Due the competitive nature of the T4P field, especially concerning pathogenic species such as *N. meningitidis*, it would otherwise be expected that more progress should have been made in the understanding of PilC, especially given its crucial roles in pilus biogenesis and cellular adhesion. Clearly this protein is an extremely challenging target.

During and after working on this project, three papers were published from the same research group who solved the structure of PilY1. These PilC follow up studies in *P. aeruginosa*, *Neisseria gonorrhoeae* and *Kingella kingae* further investigated the role of the calcium binding sites using mutation analysis (Johnson et al., 2011, Porsch et al., 2013, Cheng et al., 2013). PilC from *P. aeruginosa* was shown to have an additional calcium binding site, N-terminal of the sequence corresponding to that in the crystal structure. This additional site as well as a conserved RGD motif were shown to be crucial to T4P mediated binding of this bacteria to integrins. Neither the sequence of the additional calcium binding site nor the RGD motif are conserved in other species indicating that this mechanism of adhesion may be specific to *Pseudomonas* only. The studies in *N. gonorrhoeae* and *K. kingae* show that mutation of the C-terminal PilC calcium binding site

impacts adhesion and transformation efficiency but not levels of piliation or twitching motility as in *P. aeruginosa*. These species specific differences in mutation phenotype further complicate our understanding of the functional role of PilC and rationalisation of all of the currently available data on PilC is difficult. The sequence of PilC from *N. gonorrhoeae* is the closest evolutionarily to that of *N. meningitidis*. Unlike the *P. aeruginosa* and *K. kingae* proteins which were expressed in BL21 cells, this was produced using the Origami 2 system, very similar to the Origami B system used for the *N. meningitidis* FAM18 PilC1-C. Interestingly, the PilC protein expressed is from the FA1090 strain which is less commonly used in the laboratory as a standard strain, indicating screening of PilC sequences from different strains may have been completed in order to find the most optimally expressed and purified sequence.

Crucially no paper published on PilC or any of its homologues has addressed the key problem of where this protein resides within the T4P structure, since early work which pointed towards the protein residing in both the membrane and at the tip of the pilus structure (Rudel et al., 1995c). Without this vital piece of information, it will be very difficult to design meaningful constructs for the expression and purification of this protein or truly understand what any functional data means.

The pET32b-Origami B system was essential in making any steps forward in expressing and purifying this protein. To really investigate PilC from *N. meningitidis*, I think it would be necessary to screen for the expression of C-terminal domains from a variety of strains as the sequence variability is very high. Hopefully such an approach would allow identification of a strain which allowed better expression of a protein of a defined oligomerisation state in a consistent manner.

Inconsistent purification quality and yields of PilC in this study called into question the reliability of any conclusions drawn based on such preparations. The exact concentration of the protein at the first stage of size-exclusion chromatography was crucial for producing any monomeric PilC and it was shown that once the protein had formed higher order species, it could not revert to the

monomeric state. This made the purification process extremely technically challenging to consistently reproduce the same sample despite the large number of variables which would undoubtedly affect this.

Even the monomeric protein produced through the purification protocol which was developed was not homogeneous as determined by size-exclusion chromatography. Antibodies against PilC available commercially are extremely poor and have broad cross-reactivity. As such identification of the protein by Western blot analysis, which was attempted, is generally unsuccessful, even with a purified sample. Without confirmation by mass spectrometry as to the identity of the protein, no further studies were conducted in order to focus on more productive projects.

Hopefully in the future, identification of the location of PilC in the T4P structure as well as further genetics analysis to distinguish the exact roles of the various domains of the protein will inform more productive investigations into the protein biochemistry of PilC in the context of its structure and function.

6. Functional Characterisation of the Novel Cell Division Protein ZapE

Accurate bacterial cell division is dependent on controlled cell enlargement as well as precise timing and localisation of cytokinesis (Lutkenhaus et al., 2012). FtsZ, a GTPase structurally homologous to eukaryotic tubulin, polymerises at the mid-cell to form the Z-ring (Bi and Lutkenhaus, 1991). Constriction of the Z-ring is followed by formation of the new daughter cells, though the exact mechanism by which the FtsZ polymers achieve this is poorly understood. Approximately 10 essential factors are recruited by FtsZ, whilst many non-essential accessory proteins have roles in modulating Z-ring dynamics (Lutkenhaus et al., 2012). The elucidation of the exact interplay of all FtsZ controlling or modulating factors is ongoing.

In a recent screen of *Shigella flexneri* mutants defective for host colonisation in the gastrointestinal tract, *zapE* was described as being essential for host colonisation (Marteyn et al., 2010, Marteyn et al., 2014). ZapE is reported as a novel non-essential putative ATPase involved in Gram negative bacterial cell division, with roles in Z-ring modulation in conditions of elevated temperature and anaerobia. Such conditions are akin to those encountered by Gram negative pathogens such as *Shigella flexneri* and *E. coli* upon colonisation of the human host. Growth and cell division mechanisms in such environments are poorly understood.

E. coli ZapE has been shown to bind FtsZ through pull-down analysis and colocalisation studies. Both *E. coli* mutants which are knocked out for ZapE as well as those which constitutively express ZapE have filamentous cells and multiple disorganised Z-ring structures. This phenotype is not observed for the ZapE K84A Walker box A mutant. In the context of overexpressing the wild type ZapE, addition of ATP caused these large helical FtsZ assemblies to dismantle.

The *zapE* gene was shown to be conserved across a broad range of bacterial pathogens despite the fact that very few colonise host niches with environmental conditions similar to that of *S. flexneri*. *N. meningitidis* is one of the species in which the *zapE* gene is conserved though the

putative role of ZapE in this species is open to broad speculation currently due to limited characterisation of this protein.

To increase our understanding of this novel protein, further functional characterisation of ZapE *in vitro* is required, assessing both its ATPase activity and FtsZ binding functions using a variety of biophysical techniques. This work aimed to provide some insight into the mechanistic details of the role ZapE plays in Gram-negative bacterial cell division.

6.1 Purification and Characterisation of ZapE from *E. coli* and *N. meningitidis*

The previous study investigating ZapE conducted limited characterisation of ZapE at the protein level and instead focussed mainly on cell biology and microscopy techniques to elucidate the function of this protein. In order to fully characterise this protein in more detail, in terms of both its biophysical properties and functions, it was decided that full-length *E. coli* ZapE (to replicate work completed by Marteyn *et al*) and *N. meningitidis* ZapE (the pathogen model of choice for this work) would be investigated in parallel.

<i>N. meningitidis</i> ZapE	MSNRDQLFKAPPFENHSPLTWYQAASQLPNFIRDDAQAAAIEHLDRLWTE	50
<i>E. coli</i> ZapE	-----MQSVTPTSQYLKALNEGSHQPDDEVKEAVSRLEIIYQE	38
	: : * : * * : . . * . * * : : * :	
<i>N. meningitidis</i> ZapE	LMMFK-----RKRNRFLGRSLRSPQVP-KGLYFYGGVGRGKSF	89
<i>E. coli</i> ZapE	LINSTPPAPRTSGLMARVGKLWGKREDTKHTPVRGLYMWGGVGRGKT	88
	* : . : : : * : : : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	DAFFGCLPYRRKRRVHFHAFMAEIHQRLKTLKSESNPLKSVAEIAKETR	139
<i>E. coli</i> ZapE	DLFYQSLPGERKQRLHFHRFMLRVHEELTALQGQTDPLEIIADRFEKETD	138
	* * : . * . * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	VLCFDEFHVSDIADAMILGRLLLENLLNEGVLVATSNYAPSELYPQGQNR	189
<i>E. coli</i> ZapE	VLCFDEFFVSDITDAMLLGGLMKALFARGITLVATSNIPPDELYRNGLQR	188
	* : * : * : * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	SSFLPTIALIESSLTVLNVDGGEDYRLRTLRLPAEIFFTPANEENEAKLAK	239
<i>E. coli</i> ZapE	ARFLPAIDAIKQHCDVMNVDAGVDYRLRTLTLQAHLWLSPLHDETRAQMDK	238
	: * : * : * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	LFKEMTGITDLNPGISTIHGREIPKAESEGRAIWFDLFRALCFGRPSQSDY	289
<i>E. coli</i> ZapE	LWLALAGGKRENSPTLEINHRPLATMGVENQTLAVSFTTLCDVARSQHDY	288
	* : : * . * : * : . . : : . . * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	LYLAEHYEMVFISGLEQLSPQEKAEARRLTWLDVLYDFRVKLCATGAVD	339
<i>E. coli</i> ZapE	IALSRLFHTVMLFDVPVMTLMESEARRFIALVDEFYERHVKLVVSAEVP	338
	: * : . : * : : : : : : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	VNHIYTEGDFAEFTRTASRMVEMQSEVYLEQPHLTLSPKASGG	383
<i>E. coli</i> ZapE	LYEIIYQGDRLKFEFQRCLSRQEMQSEYLLKREHLA-----	375
	: . * . : * * * * : * : * : * : * : * :	

Figure 6.1 Alignment of the full-length ZapE protein sequences of *E. coli* and *N. meningitidis*

Identical residues are indicated by ●, a colon indicates conservation between side chains with strongly similar properties, a period indicates conservation between side chains with weakly similar properties. The sequence identity of the two proteins is 38%. The Walker A motif is highlighted with the red box whilst the Walker B-like motif is highlighted with the blue box.

6.1.1 Construction of pET28b Full-Length ZapE for both *E. coli* and *N. meningitidis*

E. coli and *N. meningitidis* ZapE sequences were ordered, the latter of which was codon-optimised for expression in *E. coli* (GeneArt, Invitrogen) with flanking restriction sites, to allow “cut and paste” cloning into the pET28b expression vector (Novagen), a vector allowing expression of the ZapE proteins with a N-terminal hexa-his tag.

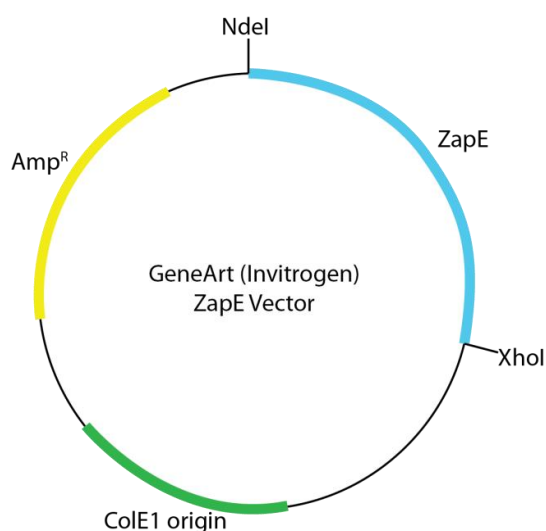


Figure 6.2 Codon-optimised ZapE genes were ordered from GeneArt

Gene sequences for both species were codon optimised for expression in *E. coli* and designed with flanking 5' NdeI and 3' XhoI restriction sites.

3 µg of each of these GeneArt vectors and pET28b were digested with NdeI and XhoI (Fermentas) in a ratio of 1 U of each enzyme to 1 µg plasmid DNA in 1X Fermentas FastDigest Green Buffer.

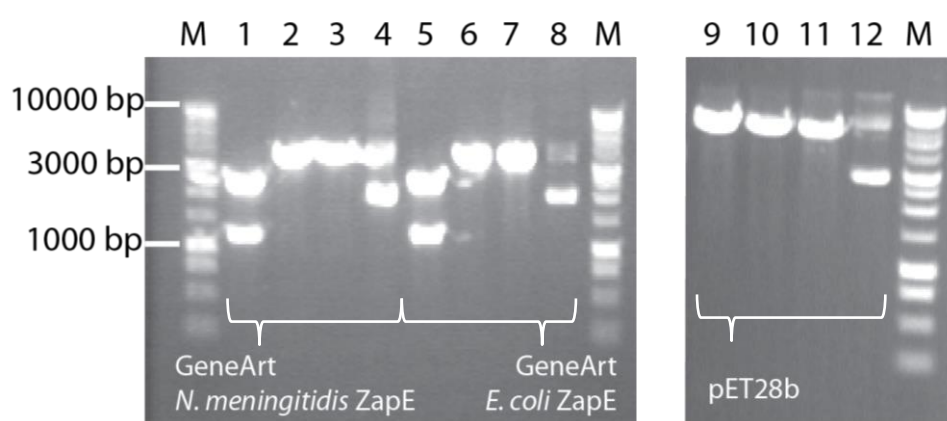


Figure 6.3 Digestion of GeneArt vectors and pET28b

All three vectors were digested with both NdeI and XhoI. M corresponds to markers. Lanes 1, 5 and 9 show vector digested with both enzymes; lanes 2, 6 and 10 show the digest reaction with just NdeI; lanes 3, 7 and 11 show the digest reaction with just XhoI; lanes 4, 8 and 12 show the undigested vectors.

Linearised pET28b was treated with Fast Alkaline Phosphatase (Fermentas) and purified with QIAquick PCR Purification Kit (Qiagen), though using a Qiagen Miniprep cartridge. Linearised pET28b was eluted with 50 µl EB at 43 ng/µl. The fragments corresponding to the two *zapE* genes were gel extracted using the QIAquick Gel Extraction Kit (Qiagen) and eluted with 25 µl buffer EB at 33 ng/µl and 30 ng/µl for the *E. coli* and *N. meningitidis zapE* fragments respectively. The vector was ligated with each of the insert fragments using the Ligafast kit (Promega) and transformed into DH5α *E. coli* competent cells. Presence of the correct insert in purified pET28b ZapE constructs was verified by sequencing with the T7F and T7R primers (Source Bioscience). The details of the proteins expressed by these constructs can be seen in **Table 6.2**.

Construct Name	Amino Acid Sequence of Open Reading Frame	Molecular Weight (Da)	Theoretical pI
<i>E. coli</i> ZapE WT pET28b	MGSSHHHHHSSGLVPRGSHMQSVTPTSQYLKAL NEGSHQPDDVQKEAVSRLEIIYQELINSTPPAPRTSG LMARVGKLWGKREDTKHTPVRLYMWGGVGRGK TWLMDLFYQSLPGERKQRLHFHFRMLRVHEELTAL QGQTDPLEIIADRFAETDVLCFDEFFVSDITDAMLL GGLMKALFARGITLVATSNIPPDELYRNLQRRARFLP AIDAIKQHCDVMNVDAVDYRLRLTLQAHLWLSPL HDETRAQMDKLWLALAGGKRENSPTLEINHRPLAT MGVENQTLAVSFTTLCVDARSQHDYIALSRLFHTVM LFDVPVMTRLMESEARRFIALVDEFYERHVKLVSVAE VPLYEIQGDRLKFEFQRCLSRQEMQSEYLRKREHL AG	45227	6.7
<i>N. meningitidis</i> ZapE WT pET28b	MGSSHHHHHSSGLVPRGSHMSNRDQLFKAPPFE NHSPLTWYQAASQLPNFIRDDAQAAIEHLRLWT ELMMFKRKRNRFLGRSLRSPQVPKGLYFYGGVGRG KSFLMDAFFGCLPYRRKRRVHFHAFMAEIHQRLKTL KSESNPLKSVAEIAKETRVLCFDEFHVSDIADAMILG RLLENLLNEGVLVATSNYAPSELYPQGQNRSSFLPTI ALIESSLTVLNDGGEDYRLRLRPAEIFFTPANEENE AKLAKLFKEMTGITDLNPGISTIHGREIPKHAESGRAI WDFRALCFGPRSQSDYLYLAEHYEMVFISGLEQLSP QEKAEARRLTWLIDVLYDFRVKLCATGAVDVNHIYT EGDFAEEFTRTASRMVEMQSEVYLEQPHLTLSPKAS GG	45978	7.03

Table 6.1 Expressed sequences of ZapE constructs

This table details the names of each of the constructs, the sequence expressed, the molecular weight of that protein product as well as its pI, both calculated by ExPASy ProtParam. The additional sequence of the N-terminal tag, derived from the pET28b vector, is highlighted in blue. This contains the hexa-histidine affinity tag as well as the thrombin cleavage sequence: LVPR/GS.

6.1.2 Optimising Expression of the pET28b-ZapE Constructs

To find the optimal conditions for expression of these constructs, high-throughput expression trials were conducted. In the case of both constructs, the conditions seen in **Table 6.2** were investigated.

Variable	Conditions Tested
<i>E. coli</i> strain	BL21 (DE3) or B834 (DE3)
Concentration of IPTG for induction upon reaching a cell density where $OD_{600} \sim 0.6$	1mM, 0.5mM, 0.2mM or 0.1mM
Media for growth	TB or LB
Conditions of growth after induction with IPTG	3 hours at 37°C, overnight at 25°C, overnight at 21°C or overnight at 16°C

Table 6.2 Conditions screened for optimal soluble expression in *E. coli* expression strains

The total number of different condition combinations comes to 64 per construct trialled

5 ml media were inoculated with 50 μ l overnight culture and grown to $OD_{600} \sim 0.6$ before induction with IPTG. After the cultures had grown at their designated time and temperature, the cells were harvested by centrifugation. Cell pellets were resuspended in 210 μ l PBS supplemented with 1 % v/v Tween-20, 1 mg/ml lysozyme (Sigma), 400 U/ μ l DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) and then placed on an orbital shaker at 1000 rpm until lysed. After clarification of the lysate by centrifugation, Ni-affinity chromatography was performed using NiNTA magnetic beads (Qiagen) which were incubated with clarified lysate at room temperature for 30 minutes. The beads were washed using PBS with 20 mM imidazole and then the protein was eluted with 50 μ l PBS with 250 mM imidazole. 25 μ l 3X SDS-PAGE loading dye was added to the elution. After boiling, 30 μ l of this sample was loaded on 15 % acrylamide gels for SDS-PAGE analysis. The results of this trial can be seen in **Figure 6.4**.

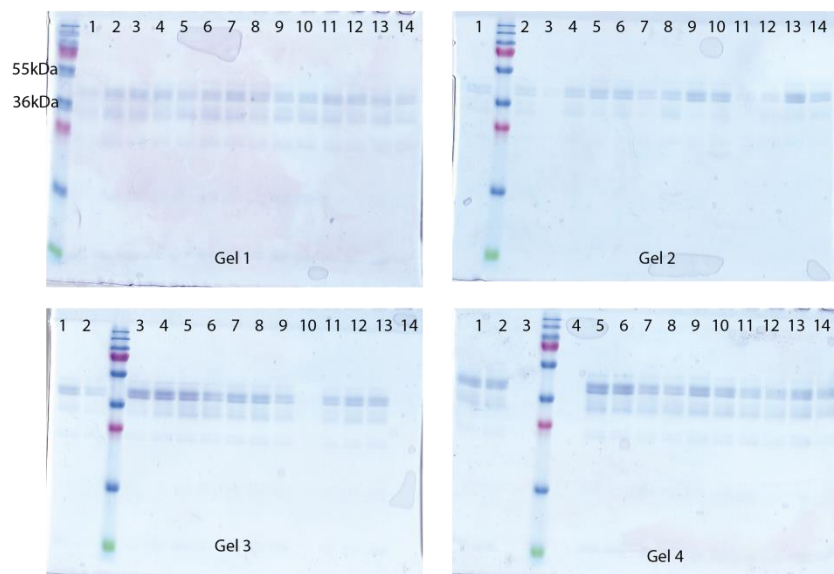
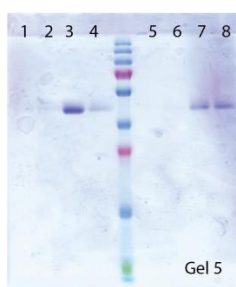
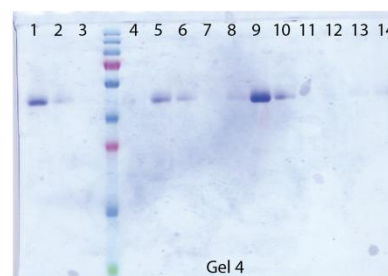
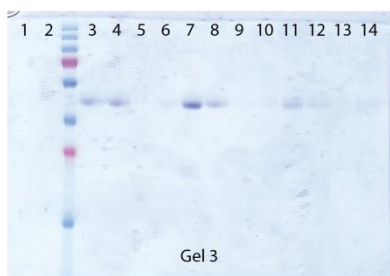
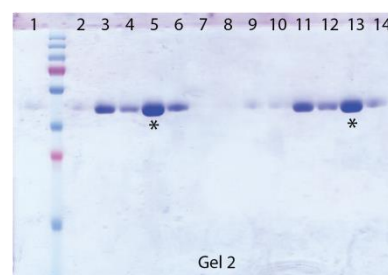
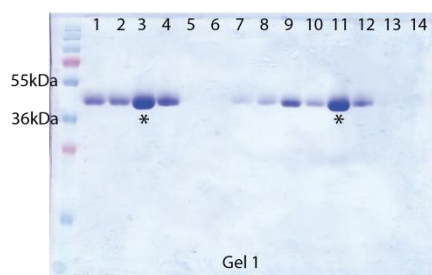


Figure 6.4 Gels showing relative expression levels of *N. meningitidis* and *E. coli* ZapE

Highest levels of expression for *N. meningitidis* ZapE are seen under the conditions with an asterisk. These correspond to purification of protein from cultures of *E. coli* BL21 (DE3) cells in TB grown at 37°C for three hours after induction with any concentration of IPTG. Degradation is evident for the *E. coli* ZapE protein. The purification is completed at room temperature which would potentially account for this degradation, despite the presence of protease inhibitor tablets in the lysis buffer.



A - *E. coli* ZapE



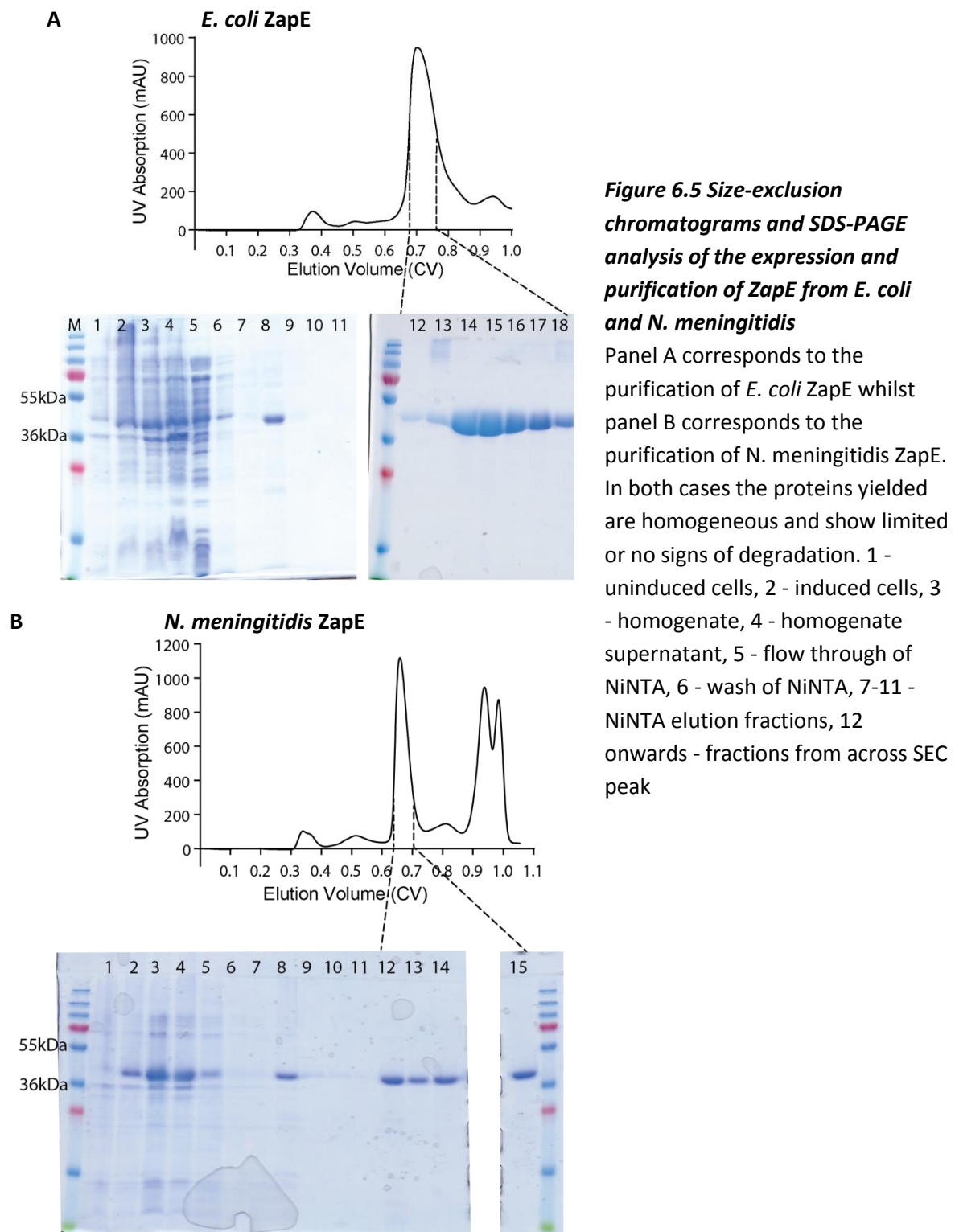
B - *N. meningitidis* ZapE

The gels indicate that *N. meningitidis* ZapE is overexpressed well when cultures of *E. coli* BL21 (DE3) cells are grown in TB at 37°C after induction with any concentration of IPTG. The protein also appears to be much more stable relative to the *E. coli* version which shows obvious degradation in nearly all conditions. As such, it was decided that *E. coli* ZapE would be grown under the optimised conditions for *N. meningitidis* ZapE in the first instance.

6.1.3 Large Scale Expression and Purification of *E. coli* and *N. meningitidis* ZapE

4 x 1 l TB cultures of *E. coli* BL21 (DE3) pET28b ZapE were grown for each of the *E. coli* and *N. meningitidis* version of this protein as described in the expression trial: 37 °C until OD₆₀₀ reached 0.6, then induced each culture with 1 mM IPTG and grown for a further three hours at 37 °C. The harvested cell pellets were resuspended in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 50 mM imidazole (resuspension buffer) which was supplemented with 1 mg/ml lysozyme (Sigma), 400 U/μl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) before homogenisation at 15,000 psi (Emulsiflex C.5 Avestin). The lysate was clarified by centrifugation at 26,000 *xg* for 20 minutes before applying the supernatant to a 5 ml NiNTA Superflow Cartridge (Qiagen) at 1 ml/min, pre-equilibrated in resuspension buffer. After a 5 CV wash with resuspension buffer, the protein was eluted in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 500 mM imidazole over 5 fractions of 5 ml each. Fractions were assessed by measuring the sample A₂₈₀ using a Nanodrop ND-1000 (Labtech) to determine protein concentration. Those which contained the eluted protein, namely fractions 2 and 3, were then concentrated to 5 ml using an Amicon Ultra-15 Centrifugal Filter Unit NMWL of 30,000 Da. The concentrated protein was then subjected to size-exclusion chromatography using an Akta Purifier (GE Healthcare Life Sciences) in a buffer of Tris-HCl pH 7.5, 150 mM NaCl. In the case of the meningococcal protein, the sample was applied at 3 ml/min to a Superdex 200 26/60 column (GE Healthcare Life Sciences), and in the case of the *E. coli* protein, the sample was applied at 1 ml/min to a Superdex 200 16/60. SDS-PAGE analysis of the eluted peak of ZapE showed that in both cases, the protein was pure and had limited, if no, degradation. See **Figure 6.5**. Limited degradation for *E. coli* ZapE is observed compared to the high throughput

expression and purification, as all stages were performed at 4 °C where possible. The asymmetric peak shape does imply some oligomeric heterogeneity or flexibility in the sample.



Yields for both proteins were approximately 10 mg per litre of culture grown. Samples sent for mass spectrometry (conducted by Dr David Staunton, University of Oxford) gave the expected mass for each protein, though in both cases were slightly smaller than the expected mass, *E. coli* ZapE by 127 Da and *N. meningitidis* ZapE by 126 Da. In both cases, this can probably be attributed to loss of the start methionine which would give expected masses of 45096 Da and 45847 Da for *E. coli* and *N. meningitidis* respectively.

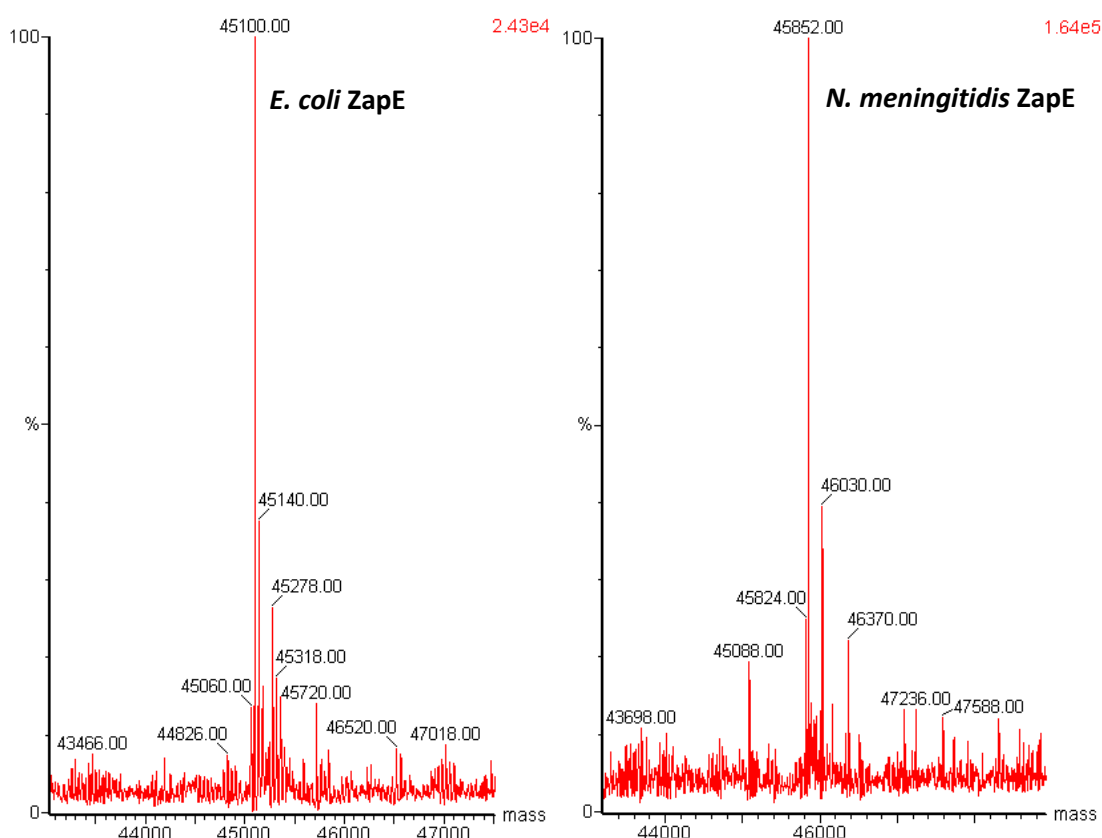


Figure 6.6 Mass spectrometry results of ZapE protein samples

The spectra above indicate clean protein samples of the expected mass if you allow for loss of the start methionine in the protein sequence in each case, with a 4Da and 5Da discrepancy for *E. coli* ZapE and *N. meningitidis* ZapE respectively.

To investigate whether the protein samples had secondary structure and overall fold, both samples were subject to DSC analysis, conducted by Dr David Staunton, University of Oxford. In both cases the melting temperature was shown to be approximately 50 °C indicating that the proteins are folded.

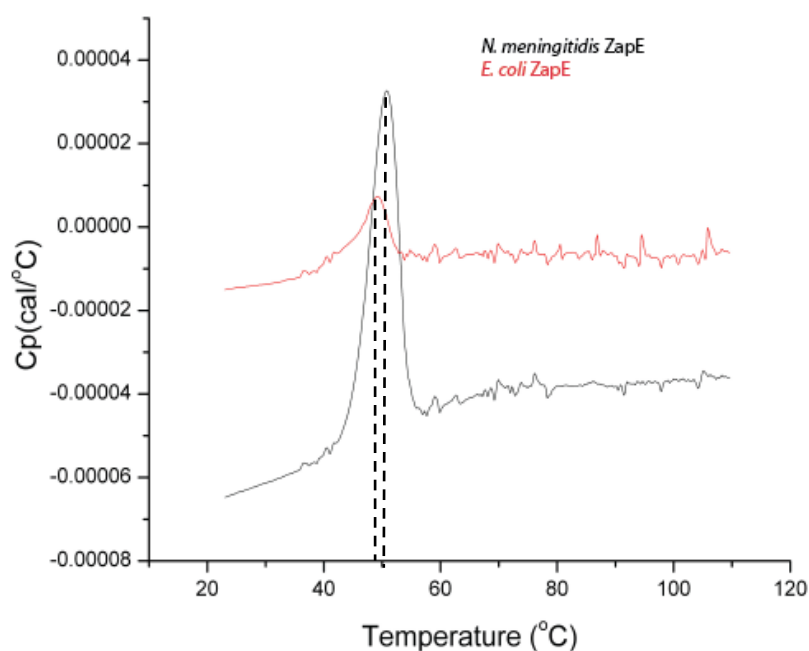


Figure 6.7 DSC of ZapE protein samples

Both samples, at 1mg/ml in 50 mM Tris-HCl pH 7.5, 150 mM NaCl have a melting temperature of approximately 50 °C although the signal:noise of the *E. coli* ZapE protein is much lower than that of the *N. meningitidis* ZapE protein

This finding was confirmed by complementary circular dichroism (CD) analysis of each protein, conducted with 0.1 mg/ml samples in 10 mM HEPES pH 7.4, 20 mM NaCl at room temperature. In both cases K2D2 server analysis of the spectra indicated that there was a large α -helical content with a small amount of β -strand. CD melt analysis, measuring the CD spectra over a temperature range of 20-80 °C, indicated a melting temperature of between 45-55 °C depending on which wavelength in the CD scan you used to monitor the change in ellipticity. This is in line with the DSC results. As the temperature dropped back to room temperature, the curve seen previously at 20 °C did not reform, indicating that the melting process was irreversible, as confirmed by the large amount of protein precipitate visible in the CD quartz cuvette. The folding of ZapE within the *E. coli* cell takes place in the presence of ATP, but the “refolding” upon heating during the CD experiment, takes place in the absence of ATP. The reason ZapE is unable to refold as the temperature slowly drops back from 80 °C to 20 °C may be due to the lack of ATP which could be needed for correct folding.

The change in Cp is low for *E. coli* ZapE and the transition between folded and unfolded protein as observed by the CD experiment, is not very sharp. This could imply that the protein is not folded

in the same manner as the *N. meningitidis* ZapE and may contain elements of molten-globule protein.

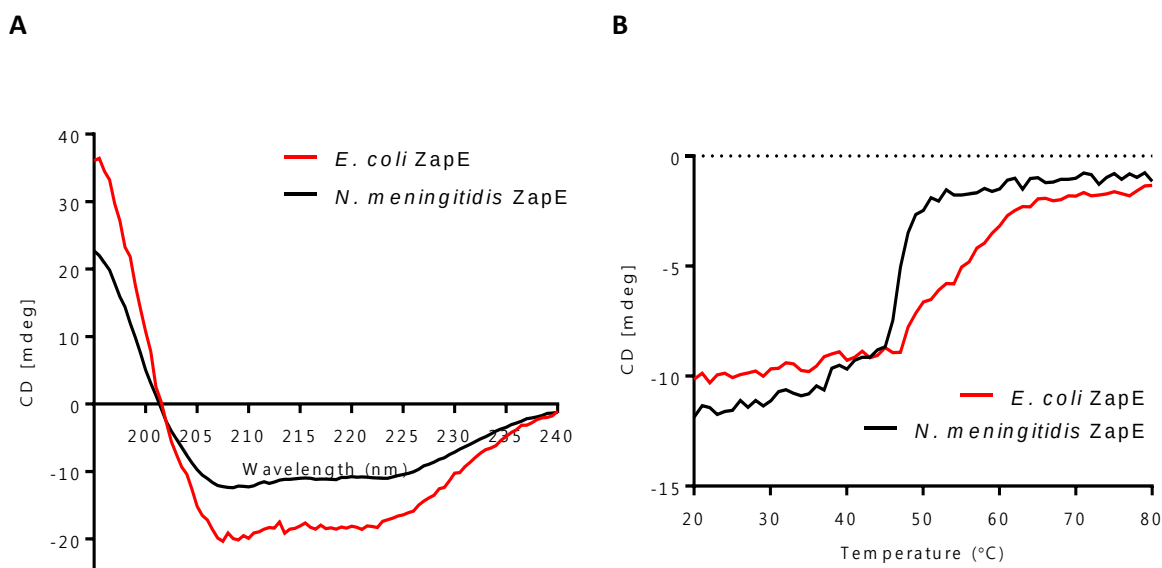


Figure 6.8 CD analysis of the ZapE protein samples

A – the ellipticity scan at constant temperature (21°C) for both the *E. coli* and *N. meningitidis* ZapE proteins from 190 to 240 nm. **B** – the change in ellipticity at 208 nm over a temperature range of 20 to 80 °C for both the *E. coli* and *N. meningitidis* ZapE proteins

6.1.4 Mutagenesis of ZapE Walker Box Motif and Subsequent Purification of the Mutant Proteins

To investigate the role of the Walker Box A motif in the function of ZapE, it was important to generate a mutant which would oblate the ATPase activity of the protein. As described previously (Winter, 2010), mutation of the lysine in the Walker A motif to an alanine is sufficient to achieve this. As such, site-directed mutagenesis was performed to generate the corresponding K84A mutant of *E. coli* ZapE and the K85A mutant of *N. meningitidis*. This was achieved using the Quikchange mutagenesis kit (Stratagene) and the primers described in **Table 6.3**. Purified pET28b ZapE mutant constructs were checked using T7F and T7R sequencing primers (Source Bioscience).

Primer Name	Sequence	T _m (°C)
<i>E. coli</i> ZapE K84A Forward	5'-GTGGTGTGGTCGTGGT GCA ACCTGGCTGATGGACC-3'	82.2
<i>E. coli</i> ZapE K84A Reverse	5'-GTCCATCAGCCAGGT TGC ACCACGACCAACACCAC-3'	82.2
<i>N. meningitidis</i> ZapE K85A Forward	5'-GGGGTCGGACGCGGC GCA AGCTTTCTGATGGACG-3'	89.4
<i>N. meningitidis</i> ZapE K85A Reverse	5'-CGTCCATCAGAAAGCT TGC GCCGCGTCCGACCCC-3'	89.4

Table 6.3 Primers for mutagenesis of Walker Box A lysine to alanine

The region in each primer corresponding to the lysine to alanine mutation is highlighted in blue.

Mutant proteins were produced by the same protocol as the wild-type ZapE proteins as described previously in **6.1.3**. Yields were comparable to the wild type proteins with highly similar size-exclusion chromatography profiles. Again, the final protein products were homogeneous and showed limited or no degradation, see **Figure 6.9**. Interestingly, despite two conserved cysteine residues present in both proteins, no reducing agent was required throughout the protein purification to reduce aggregation or diminish the presence of higher order structures, possibly suggesting that both cysteine residues are buried in the globular structure of both proteins.

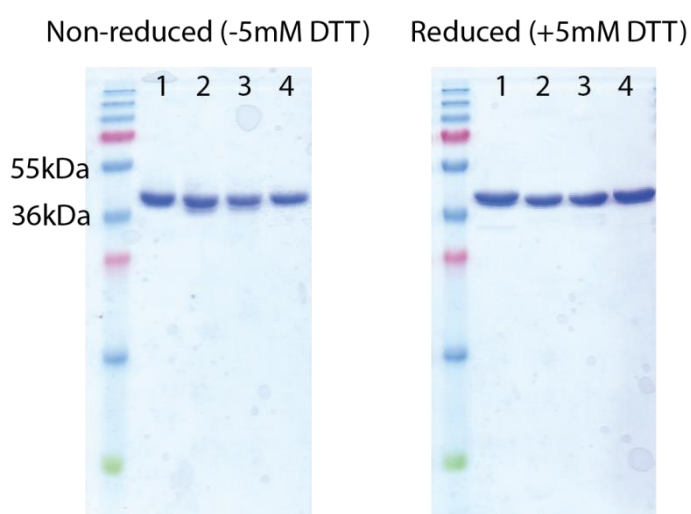


Figure 6.9 Gels of the final ZapE protein, wildtype and Walker box A mutant from *E. coli* and *N. meningitidis*

1 - *E. coli* ZapE WT, 2 - *E. coli* ZapE K84A, 3 - *N. meningitidis* ZapE WT, 4 - *N. meningitidis* ZapE K85A

6.2 Attempts Towards Structural Characterisation of ZapE

The high throughput expression trials of *E. coli* ZapE indicate that this version of the protein is prone to degradation when left at room temperature. This property renders the *E. coli* ZapE not amenable to crystallisation experiments where protein may be incubated in solution in a variety of buffer conditions at room temperature for anytime from days to years. As such, attempts were focussed on the meningococcal version of the protein. FFAS analysis of the protein sequence indicated no close structural homologues, the highest hit with a score of -14.5 was a bacterial primosome component (PDB ID: 2W58) with 20 % sequence identity over a 142 residue region of amino acids 73-215 of the sequence. Given the general lack of structural information, it was decided that purification and crystallisation of the selenomethionine derivative of the protein to allow MAD phasing would be the optimal strategy.

6.2.1 Expression and Purification of *N. meningitidis* ZapE Selenomethionine Derivative

In order to ensure the protein produced for crystallisation contained only selenomethionine instead of native methionine, the pET28b *N. meningitidis* ZapE construct was transformed into the B834 (DE3) *E. coli* expression strain, a methionine auxotroph entirely dependent on environmental methionine for protein synthesis. Overnight cultures were grown in LB and then washed twice through harvesting the cell pellet by centrifugation and then gently resuspending the cells in sterile PBS of the same volume as the original culture. This cell resuspension was then diluted 1 in 100 into glucose-free M9 minimal media (Molecular Dimensions) supplemented with the Selenomethionine Medium Nutrient Mix (Molecular Dimensions), 40 µg/ml selenomethionine (Fisher) and 0.4 % glycerol to allow growth of 4 x 1 l of culture.

Cultures were then grown at 37 °C until OD₆₀₀ ~ 0.6. Induction of the cultures with 1 mM IPTG was followed by overnight incubation at 21 °C. Despite the fact that in rich media ZapE can be expressed at 37 °C, in minimal media more gentle growth conditions generally promote better expression levels. After cell harvesting, the purification of the protein was completed using the

protocol outlined in **6.1.3** though the final size-exclusion buffer contained 1 mM TCEP, to maintain a mildly reduced environment to ensure all selenium atoms are of an equivalent reduced nature. This is important for the subsequent data collection from a crystal derived from this protein sample to ensure that the anomalous signal is as strong as possible and that the protein molecules within the crystal lattice are homogenous.

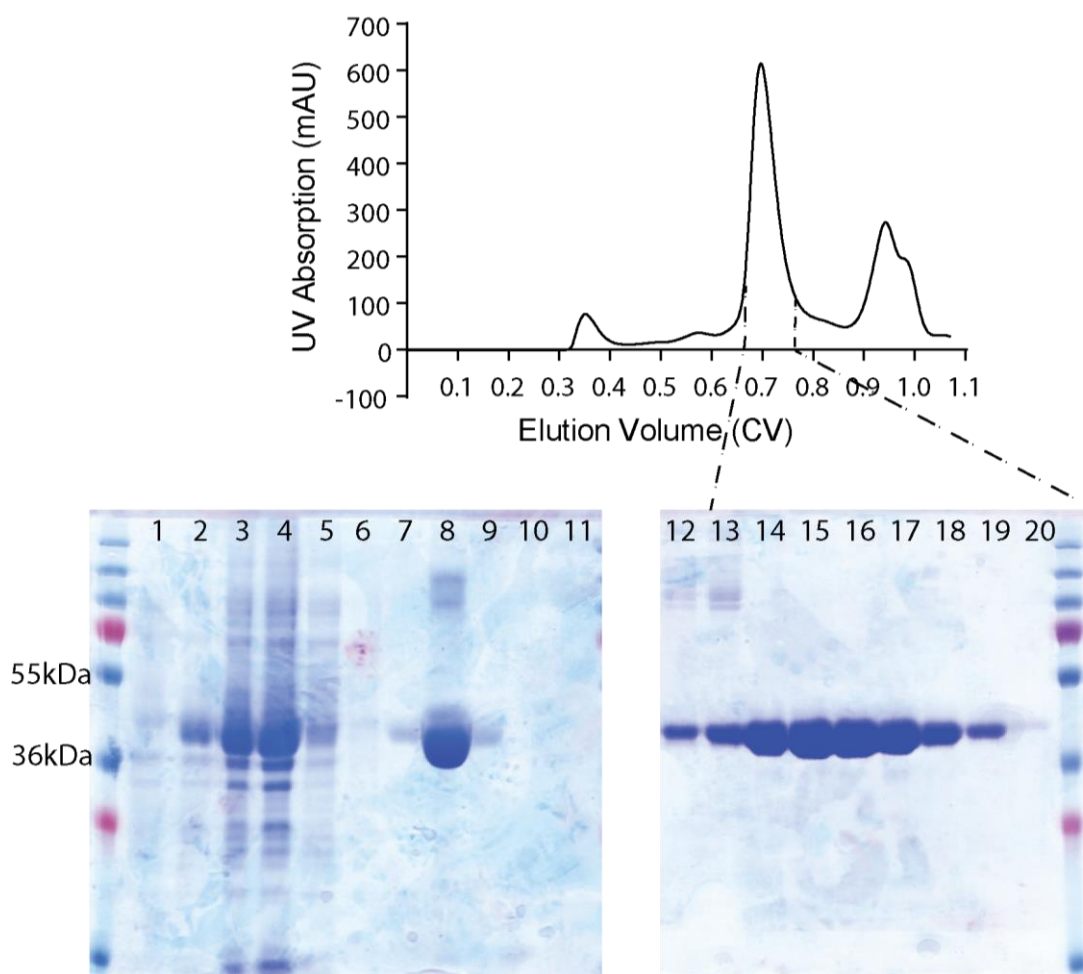


Figure 6.10 Size-exclusion profiles and SDS-PAGE analysis of the expression and purification of selenomethionine derivatised ZapE from *N. meningitidis*

1 - uninduced cells, 2 - induced cells, 3 - homogenate, 4 - homogenate supernatant, 5 - flow through of NiNTA, 6 - wash of NiNTA, 7-11 - NiNTA elution fractions, 12-20 - fractions from across SEC peak. Fractions 12 and 13 show higher molecular weight contaminants are present. As such only fractions 14-19 were pooled and subsequently concentrated for crystallisation experiments.

Once fractions were pooled and the protein concentrated with an Amicon Ultra-15 Centrifugal Filter Unit NMWL of 30,000 Da, approximately 1 ml of protein at 20 mg/ml could be obtained. Such a sample was stable when kept at 4 °C for up to three days as verified by SDS-PAGE analysis.

To verify the selenomethionine incorporation into the ZapE protein, a sample was sent for analysis by mass spectrometry, conducted by Dr David Staunton, University of Oxford. This confirmed selenomethionine incorporation had been 100 % efficient, see **Figure 6.11**.

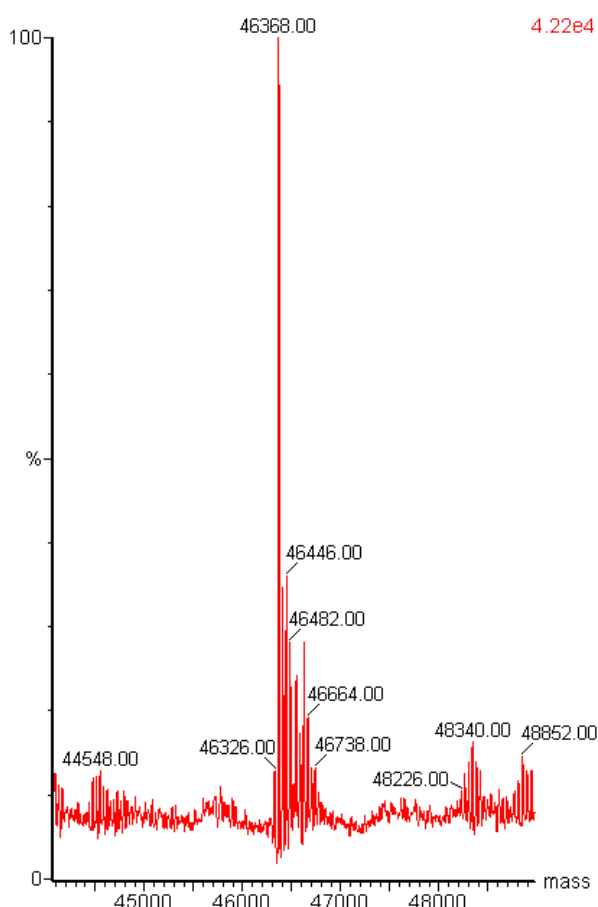


Figure 6.11 Mass spectrometry results of the selenomethionine derivatised ZapE from *N. meningitidis*

Assuming loss of the start methionine, there are 11 methionine residues in the protein structure. The difference in mass between a native methionine and a selenomethionine residue is 46.89 Da, giving a total expected mass difference of 515.79 Da. The expected mass for the native protein without the START methionine is 45847.2 Da, whilst the measured mass is 45852 Da. If there was 100 % selenomethionine incorporation, the calculated mass would be 46363 Da, but the expected mass given the previous observed mass would be 46368 Da. This is the mass observed in this experiment, indicating that selenomethionine incorporation is 100 %.

To investigate the oligomerisation state of this protein sample further, MALLS was conducted by Dr Steve Johnson, University of Oxford, with this sample at three different concentrations. Overall it was shown that the protein was monomeric with a calculated mass of approximately 46-48 kDa, see **Figure 6.12**. At all three concentrations the mass across the peak is increasing slightly, the

elution volume decreases at higher concentration and the peak shape is asymmetric, indicating a tendency towards aggregation and some heterogeneity but overall the protein is a stable monomer.

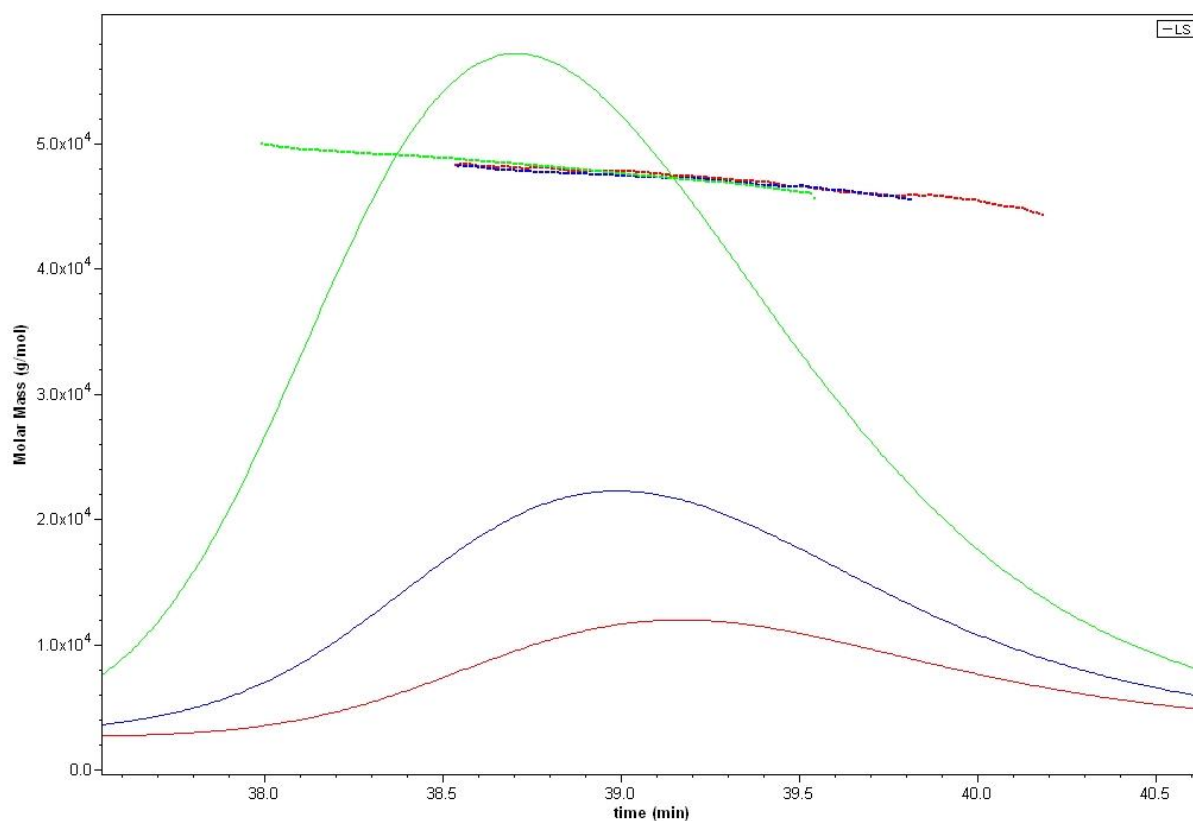


Figure 6.12 MALS analysis of *N. meningitidis* ZapE

Protein samples were run at 5mg/ml (green), 2mg/ml (blue) and 1mg/ml (red). Overall the mass is consistent at all three concentrations and the tendency towards a higher mass at the beginning of the elution peak is independent of sample concentration.

6.2.2 Crystallisation Attempts of Selenomethionine Derivatised *N. meningitidis* ZapE

The selenomethionine derivatised meningococcal ZapE was a highly homogenous sample which could be produced to yield a relatively large amount of protein sample which was stable; a sample amenable to crystallisation. Given the lack of any structural information available for ZapE, initial attempts to crystallise the protein were conducted with the full-length protein.

High throughput crystallisation trials were completed which used pre-made 96-well crystallisation screens (Molecular Dimensions). Crystallisation trials were set up using a Hydra 96 (Alpha Biotech Ltd) which transferred 75 µl mother liquor from the 96-well deep-well block of the crystallisation screen to the reservoirs of a MRC 2-drop crystallisation plate (Molecular Dimensions). The different screens trialled during these experiments are summarised in **Table 2.2**. 1:1 and 3:1 ratios of protein:mother liquor were pipetted to yield 600 nl drops using either an OryxNano Crystallisation Robot (Douglas Instruments) or a Mosquito Crystal (TTP Labtech). Plates were then sealed using a StarSeal Advanced Polyolefin Film (STARLAB) and incubated at 21 °C.

Crystallisation trials were set up at various concentrations of ZapE protein with the different screens described in **Table 2.2**. These trials yielded no diffracting protein crystals. As such further sample modifications were made to try and aid crystallisation. As ZapE is a reported ATPase, ATP analogue additives were also screened in combination with the various concentrations of protein to see if this would stabilise the protein further to promote crystallisation. The ATP additives chosen, were mimics of ATP transition states or non-hydrolysable ATP analogues as detailed in **Table 6.4**.

A comprehensive summary of the screens and ATP analogues trialled can be seen in **Table 6.5**.

Despite the extensive trials conducted, no diffracting protein crystals were found in any of the trays to date. This could be due to a number of problems:

- Intrinsic flexibility of the full-length protein molecule
- Extended unstructured regions at the hexa-his tag, N- or C- termini
- ATP analogues not bound by protein or they do not promote a more stable conformation

As such, further investigation into the predicted structure and domain arrangement of the ZapE protein were required.

Analogue Name	Structure	Feature	Protocol for Use
AMP-PNP		Non-hydrolysable ATP analogue. O-substituted with N in γ -phosphate.	Purchased in powder format. Stock solution 200 mM made with ddH ₂ O. Final concentration of 2 mM with 2 mM MgCl ₂ .
AMP-PCP		Non-hydrolysable ATP analogue. O-substituted with C in γ -phosphate.	Purchased in powder format. Stock solution 200 mM made with ddH ₂ O. Final concentration of 2 mM with 2 mM MgCl ₂ .
ADP·BeF ₃		ADP with BeF ₃ as a γ -phosphate mimic. Stable ATP analogue in “pre-transition state” conformation.	Final concentrations 2 mM ADP, 8 mM NaF, 2 mM MgCl ₂ , 2 mM BeCl ₂ to form the final derivative.
ADP·AlF ₄		ADP with AlF ₄ as a γ -phosphate mimic. Stable ATP transition state analogue.	Final concentrations 2 mM ADP, 8 mM NaF, 2 mM MgCl ₂ , 2 mM AlCl ₃ to form the final derivative.

Table 6.4 The ATP analogues used, their relative features and a protocol for use

In all cases R corresponds to adenosine. Final concentrations describe those in the crystallisation drop.

Protein sample used to set up trays	Purification buffer	96-well crystallisation screens trialled	ATP analogue
<i>N. meningitidis</i> ZapE 14 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , The Stura Footprint Combination	None
<i>N. meningitidis</i> ZapE 18 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , The Stura Footprint Combination	None
<i>N. meningitidis</i> ZapE 20 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	None
<i>N. meningitidis</i> ZapE 22 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, The PGA Screen™	None
<i>N. meningitidis</i> ZapE 24 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	None
<i>N. meningitidis</i> ZapE 21 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	ADP-AlF ₄
<i>N. meningitidis</i> ZapE 23 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	ADP-AlF ₄
<i>N. meningitidis</i> ZapE 25 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®	ADP-AlF ₄
<i>N. meningitidis</i> ZapE 21 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, Morpheus®, The PGA Screen™	ADP-BeF ₃
<i>N. meningitidis</i> ZapE 23 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, The Stura Footprint Combination	ADP-BeF ₃
<i>N. meningitidis</i> ZapE 25 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	ADP-BeF ₃
<i>N. meningitidis</i> ZapE 20 mg/ml	50 mM Tris-HCl pH7.5, 150 mM NaCl	Structure screen I + II, PACT <i>premier</i> , JCSG- <i>plus</i> , Proplex™, The Stura Footprint Combination, Morpheus®, The PGA Screen™	AMP-PCP

Table 6.5 All crystallisation trials conducted

All ATP analogues used in the crystallisation trials were present in the drop at the concentration and conditions described in **Table 6.4**

6.2.3 FFAS analysis of the Protein Sequence

As mentioned in **6.2**, there were no strong structural homologues of ZapE found from analysis of either the *E. coli* or *N. meningitidis* ZapE amino acid sequence using the FFAS online server.

However, the alignment determined by the server does indicate the presence of two regions of sequence which share some homology with previously determined structures. These regions are highlighted in **Figure 6.13**. Further analysis of the structures to which the ZapE sequence aligns in these regions indicates that they may represent two separate domains within the overall structure. This is in agreement with the published SAXS analysis on ZapE which shows the protein to be a two domain protein, apparently similar to FtsH (Marteyn et al., 2014). If this is indeed the case, a strategy forward in the crystallisation of ZapE could be to N-terminally truncate the sequence not included in the putative domain sequence or to attempt to crystallise individual domains.

All of the proteins described in the FFAS alignment have Walker box A and Walker B-like motifs within their sequence, like ZapE. Similarly to ZapE they are predominantly from the AFG-1 ATPase family. Despite low sequence identity between the different proteins highlighted, the overall fold within the first domain highlighted is very similar consisting of a compact $\alpha\beta\alpha$ nucleotide binding domain of five to six parallel β -strands sandwiched by α -helices.

<i>N. meningitidis</i> ZapE	MSNRDQLFKAPPFENHSPLTWYQAASQLPNFIRDDAQAAAIEHLDRWLWTE	50
<i>E. coli</i> ZapE	-----MQSVTPPTSQYLKALNEGSHQPDVQKEAVSRLEIYYQE	38
	: : . : * : * : . . * . * * : : : * :	
<i>N. meningitidis</i> ZapE	LMMFK-----RKRNRFLGRSLRSPQVP-KGLYFYGGVGRGKSFLM	89
<i>E. coli</i> ZapE	LINSTPPAPRTSGLMARVGLWVKREDTKHTPVRGLYMWGGVGRGKTWLM	88
	* : . : : : * : : : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	DAFFGCLPYRRKRRVHFHAFMAEIHQRLKTLKSESNP LKSVAAEIAKETR	139
<i>E. coli</i> ZapE	DLFYQSLPGERKQRLHFHFRFMLRVHEELTALQGQTDPLEIIADR FKAETD	138
	* * : . * . * : * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	VLCFDEFHVSDIADAMILGRLLLENLLNEGVLVATSNYAPSELYPQGQNR	189
<i>E. coli</i> ZapE	VLCFDEFFVSDITDAMLLGGLMKALFARGITLVATSNIPPELYRNGLQR	188
	* * * * * : * * : * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	SSFLPTIALIESSLTVLNVDGGEDYRLRLTLRP AEIFFTPANEENEAKLAK	239
<i>E. coli</i> ZapE	ARFLPAIDAIKQHCDVMNVDAGVDYRLRLTLTQAHLWLSPLHDETRAQMDK	238
	: * * : * : * : . : * : * : * : * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	LFKEMTGITDLNPGISTIHGREIPHKAESGRAIWFD FRALCFGPRSQSDY	289
<i>E. coli</i> ZapE	LWLALAGGKRENSPTLEINHRPLATMGVENQTLAVSFTTLCVDARSQHDY	288
	* : : : * . * . : * : * : . . : : . . * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	LYLAEHYEMVFISGLEQLSPQEKAEARRLTWLIDVLYDFRVKLCATGAVD	339
<i>E. coli</i> ZapE	IALSRLFHTVMLFDVPMTRLMSEARRFIALVDEFYERHVKLVS AEVP	338
	: * : . : . * : : : : : : : * * : * : * : * : * : * : * :	
<i>N. meningitidis</i> ZapE	VNHIYTEGDFAEETFRTASRMVEMQSEVYLEQPHLTLSPKASGG	383
<i>E. coli</i> ZapE	LYEIIYQGDR LKFEFQRCLSRLQEMQSE EYLKREHLAG-----	375
	: . * * : . : * * * * * : * * : * * : * * :	

Figure 6.13 Sequence alignment of the ZapE sequences from *N. meningitidis* and *E. coli* with FFAS regions of similarity highlighted

Of the 10 structures identified by FFAS, all aligned with the sequence highlighted in green, whilst just one (PDB ID: 2Z4R) aligned with the sequence highlighted in yellow as well.

6.2.4 Limited Proteolysis of *N. meningitidis* ZapE

Initial attempts to produce a ZapE protein product more amenable to crystallisation began with cleavage of the short N-terminal tag. Following successful expression and purification of ZapE, thrombin (GE Healthcare Life Sciences) cleavage of the N-terminal hexa-his tag was trialled to leave the full-length ZapE protein with a shorter tag remnant of Gly-Ser-His. Trial reactions were run with 0, 1, 2, 5 or 10 units of thrombin per mg of ZapE and incubated at 4 °C for 1 hour, 2 hours, 4 hours or overnight. Reactions were terminated by addition of 3x SDS-PAGE loading dye and boiling at 95 °C for 10 minutes. Under all reaction conditions which included thrombin, a 34

kDa product was visible on SDS-PAGE indicating that non-specific cleavage by thrombin had taken place. Although no strict thrombin cleavage consensus sequence is present within the protein sequence, the production of the 34kDa fragment indicated that cleavage may have occurred at the Walker Box A motif Gly-Val-Gly-Arg|Gly-Lys-Ser which is close in sequence to less specific thrombin cleavage sites. As such, no thrombin cleavage was performed in future preparations of these proteins.

Protease Used	Rationale	Features
Thrombin	Specific cleavage of N-terminal his-tag	Serine protease with apparent high specificity for Leu-Val-Pro-Arg/Gly-Ser
Trypsin	Limited proteolysis	Serine protease which cleaves after an Arg or Lys but not before a Pro
Subtilisin	Limited proteolysis	Serine protease with broad specificity

Table 6.6 *Proteases used to characterise ZapE structure and generate protein products more amenable to crystallisation*

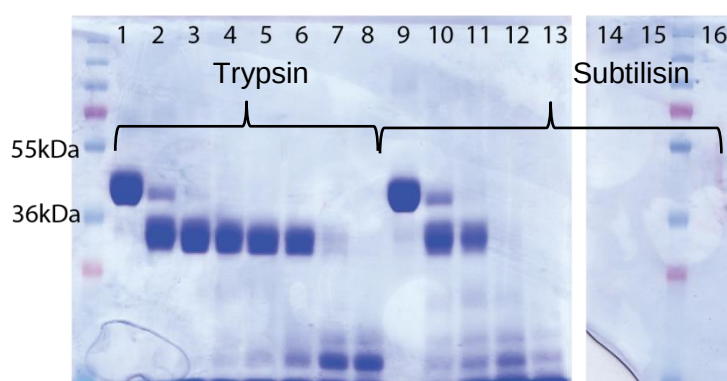


Figure 6.14 *Limited proteolysis of N. meningitidis ZapE*

Ratios of enzyme to protein (w:w).
 1 & 9 – 0, 2 & 10 - 1:10000, 3 & 11 - 1:5000, 4 & 12 - 1:1000, 5 & 13 - 1:500, 6 & 14 - 1:250, 7 & 15 - 1:50, 8 & 16 - 1:25.

Next, limited proteolysis was conducted with trypsin and subtilisin on the full-length *N. meningitidis* ZapE protein sample. In each case, reactions were set up with 20 µg protein and then a range of protease concentrations to provide a range of enzyme:protein ratios (w/w) from 1:10000 to 1:25. Control reactions were run with no protease present. All reactions were conducted in 50 mM Tris-HCl pH7.5, 150 mM NaCl and run at 4 °C overnight. Even at the lowest

concentration of protease, in the case of both trypsin and subtilisin, the majority of the full-length protein band disappears to show a 36 kDa product. A 1:50 ratio of trypsin:ZapE and a 1:1000 ratio of subtilisin:ZapE, is the point of disappearance of this 36 kDa product to form lower molecular weight species of approximately 18 kDa and less.

Given that trypsin was shown to be less aggressive in its proteolysis of ZapE, it was decided further experiments should be conducted with trypsin only. To discern the nature of the 36 kDa and 18 kDa fragments generated by the limited proteolysis, reactions were repeated at 1:1000 trypsin:ZapE and 1:50 trypsin:ZapE to generate these fragments respectively. The samples were then run on 4-20 % precast acrylamide gel (Biorad) by a standard SDS-PAGE protocol, though with 2 mM β -mercaptoethanol in the upper electrode buffer. The gel was then subject to semi-dry transfer to a PVDF membrane as per a standard Western blotting protocol. The membrane was then washed in ddH₂O and stained with coomassie before more washes in ddH₂O and air drying. The membrane was then sent for N-terminal sequencing at the PNAC Facility, Cambridge University, U.K. The band seen below the 18 kDa degradation product in **Figure 5.15** generated no signal through N-terminal sequencing. Within this study, it was investigated no further.

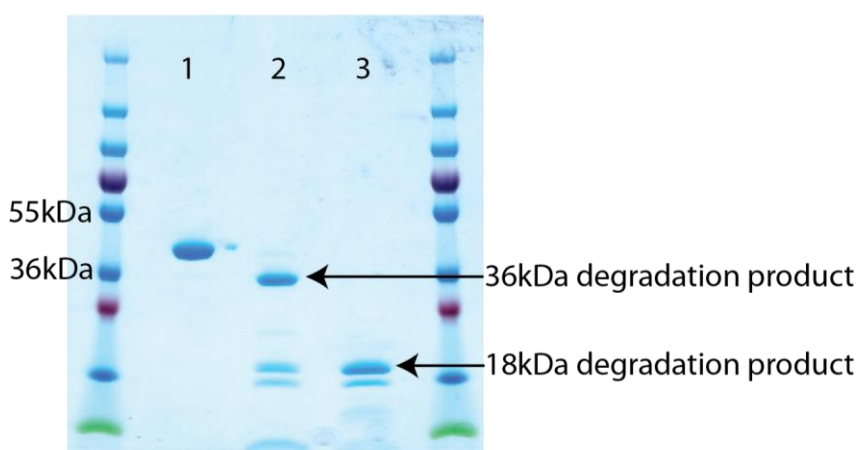


Figure 6.15 4-20% SDS-PAGE of limited proteolysis products for N-terminal sequence analysis
1 – undigested ZapE, 2 – 1:1000 trypsin:ZapE, 3 – 1:50 trypsin:ZapE. The bands subject to N-terminal sequence analysis are highlighted.

The results of the N-terminal sequencing analysis can be seen in **Table 6.7**. To determine the exact sequence of the 36 kDa and 18 kDa fragments, both samples were then subject to size-exclusion chromatography on S200 10/300 or S75 10/300 columns in 50 mM Tris-HCl pH7.5, 150 mM NaCl. The pooled peak corresponding to the proteolysed fragment was then analysed by mass spectrometry, conducted by Dr David Staunton, University of Oxford.

Fragment	Abundance within the gel band (pmoles)	N-terminal 5 amino acids
36 kDa	12	SLRSP
36 kDa	5	SPQVP
18 kDa	32	TLRPA
18 kDa	2	RLRTL

Table 6.7 N-terminal sequencing results of the *N. meningitidis* ZapE proteolysed fragments

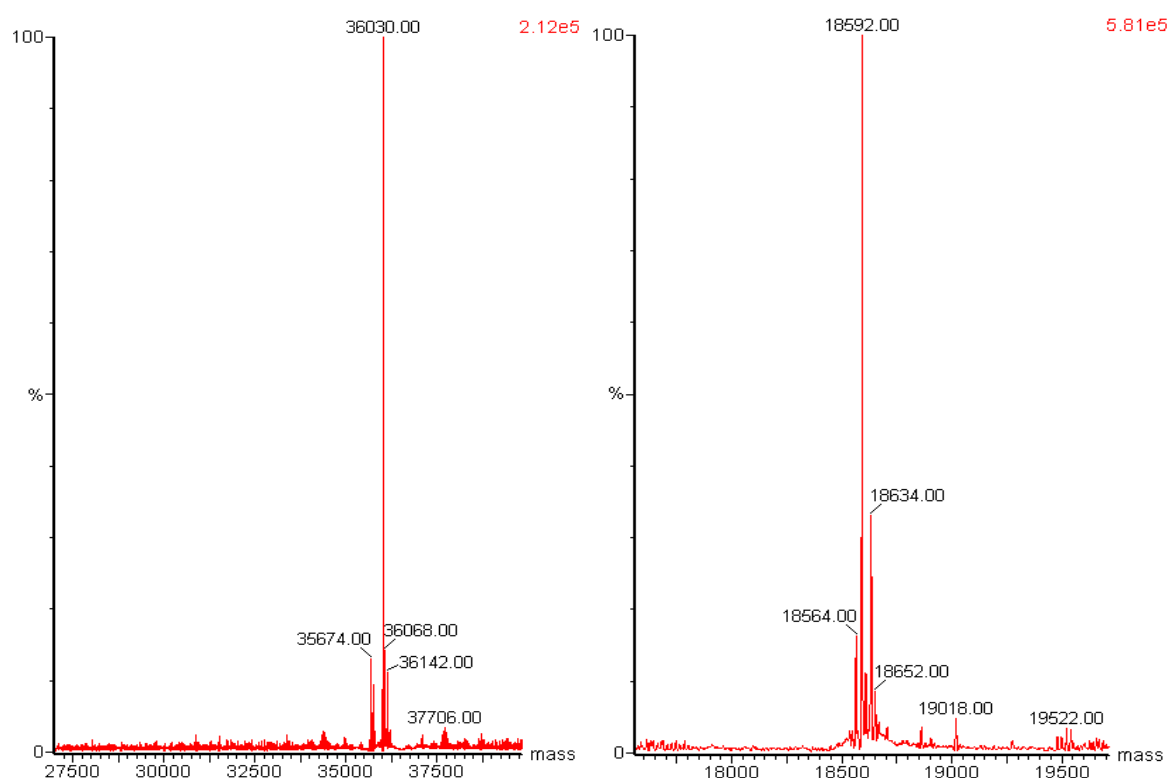


Figure 6.16 Mass spectrometry analysis of the proteolysed fragments

Using the N-terminal sequencing data in combination with the mass spectrometry results, it is possible to determine the complete sequence of the fragments. Only the fragments of higher

abundance were further analysed in each case. Both of the masses are as expected and correlate well to the less accurately predicted SDS-PAGE results.

Construct Name	Amino Acid Sequence	Molecular Weight (Da)	Theoretical pI	Difference in mass to MS results (Da)
<i>N. meningitidis</i> ZapE 36 kDa fragment	SLRSPQVPKGLYFYGGVGRGKSFLMD AFFGCLPYRRKRRVHFHAFMAEIHQR LKTLKSESNPLKSVA AEIAKETRVLCFD EFHVSDIADAMILGRLLNLLNEGVVL VATSNYAPSELYPQGQNRSSFLPTIALI ESSLTVLNVDGGEDYRLRTLPAEIFFT PANEENEAKLAKLFKEMTGITDLNPGI STIHGREIPHK AESGRAIWDFRALCF GPRSQSDYLYLA EHYEMVFISGLEQLS PQEKA EARRLTWLIDVLYDFRVKLCAT GAVDVNHIYTEGDFAEEFTRTASRMV EMQSEVYLEQPHLTSPKASGG	36025	6.03	5
<i>N. meningitidis</i> ZapE 18.5 kDa fragment	TLRPAEIFFTPANEENEAKLAKLFKEMT GITDLNPGISTIHGREIPHK AESGRAIW DFRALCFGPRSQSDYLYLA EHYEMVF ISGLEQLSPQEKA EARRLTWLIDVLYDF RVKLCATGAVDVNHIYTEGDFAEEFTR TASRMVEMQSEVYLEQPHLTSPK	18591	5.13	1

Table 6.8 Sequences of the dominant proteolytic fragments as determined by mass spectrometry and N-terminal sequencing

In the case of both fragments, the more abundant fragment sequence is shown. Both ExPASy predicted masses are in good agreement with the experimentally determined masses. The 36 kDa fragment shows proteolytic cleavage has only removed the N-terminal 64 amino acids, whilst the 18.5 kDa fragment has lost 217 N-terminal residues and the 4 C-terminal residues.

Both mass spectrometry and SDS-PAGE are denaturing techniques so to verify that the ZapE fragments observed are relevant to the physiological protein structure, it was important to determine whether these fragments represented true domains. Analysis of the FFAS structural hits indicate that, assuming correct alignment of the sequences, the N-terminal cleavage that produces the 36 kDa fragment would arise from proteolysis of an exposed loop, linking key components of the overall globular domain, see **Figure 6.17**. Assuming structural homology, this indicates that the 1:1000 trypsin:ZapE digest could be acting at this loop, but the two components would still hold together in solution to form the overall protein structure. Comparative analysis of

the size-exclusion chromatography profiles of the digested and undigested protein should allow discrimination of whether or not this is correct.

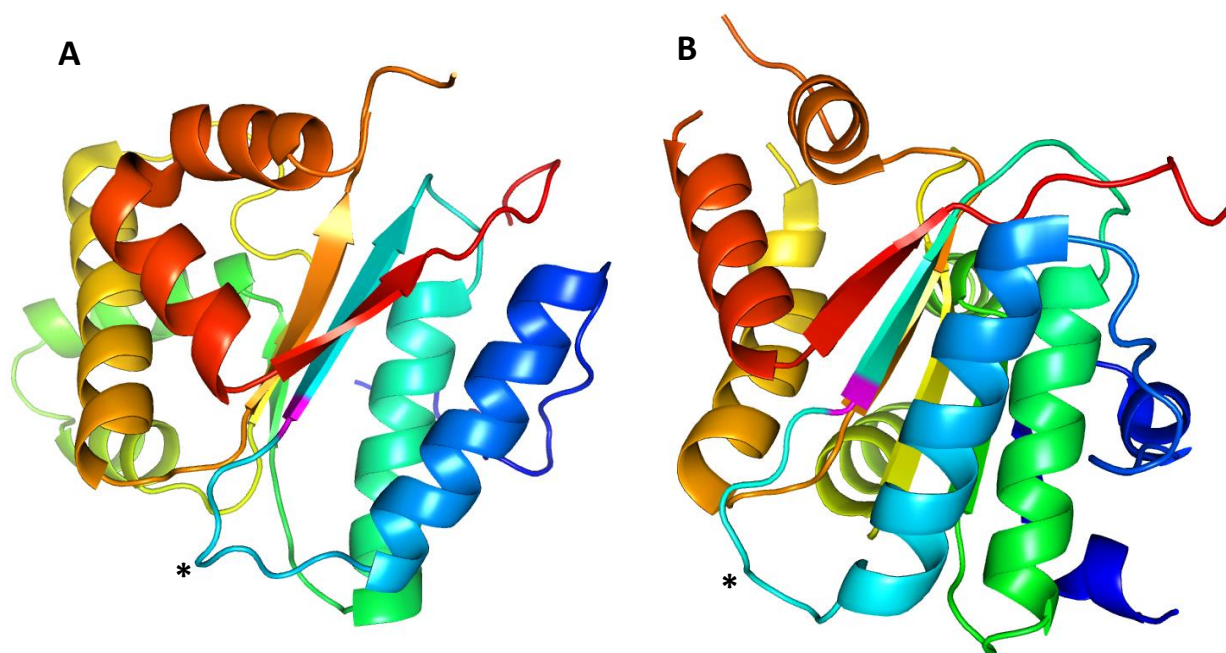


Figure 6.17 Comparable proteolysis site on FFAS top hits

The FFAS analysis indicates that the cleavage site to produce the 36kDa product seen under denaturing conditions would be at the residue in the exposed loop marked with a *. The residue marked on the following β -strand in magenta, is a highly conserved glycine residue at the start of the FFAS alignment. Panel A shows DnaC (PDB ID: 3ECC) with an FFAS score of -14.2, whilst panel B shows DnaI (PDB ID: 2W58) with an FFAS score of -14.5. Both of these protein sequences align to the green region of the alignment in **Figure 6.13**.

Reactions containing 2 mg of ZapE in 50 mM Tris-HCl pH7.5, 150 mM NaCl were incubated over night at 4 °C with either no trypsin, 1:50 trypsin:ZapE or 1:1000 trypsin:ZapE. Samples were subsequently analysed by analytical size-exclusion chromatography in 50 mM Tris-HCl pH 7.5, 150 mM NaCl using either an S200 10/300 GL in the case of the 1:1000 digestion or an S75 10/300 GL in the case of the 1:50 digestion.

For the 1:1000 digestion, the elution volume of the peaks of digested and undigested protein are identical. An additional shoulder can be seen on the peak of the digested sample, indicative of degradation, expected upon protease incubation. This supports the theory that the 36 kDa

fragment is irrelevant physiologically. However, the 18 kDa fragment elutes approximately 0.1 CV later than the undigested protein, indicating that this is indeed a true domain.

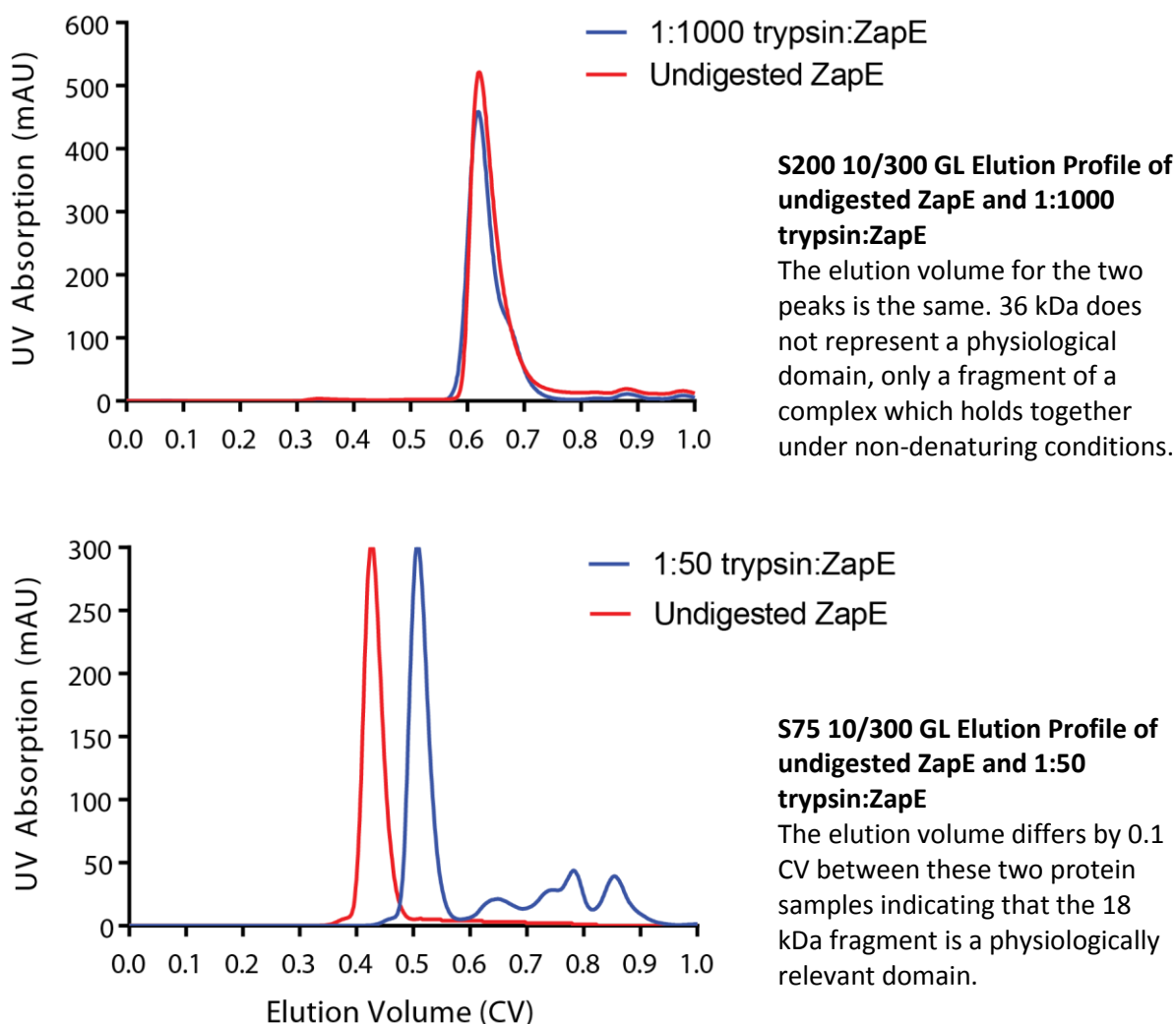


Figure 6.18 Size-exclusion profiles of digested and undigested ZapE

As determining molecular weight by size-exclusion chromatography is a low resolution technique, MALS analysis of the undigested and 1:1000 digested was conducted. This confirmed the finding that the 36 kDa fragment observed by mass spectrometry and SDS-PAGE is still bound to the 10 kDa fragment despite clipping of an exposed loop as the mass for both samples was determined to be 46-48 kDa. It is interesting that the first domain, highlighted in green in **Figure 6.19** and predicted to be 27 kDa, is not seen through this domain analysis, possibly due to the predicted

cleavage of an exposed loop, proteolysing it into two smaller fragments predicted to be 10 kDa and 17 kDa each. Perhaps the 17 kDa fragment is that described earlier from **Figure 6.15** which generated no signal on when attempts were made to N-terminally sequence this band.

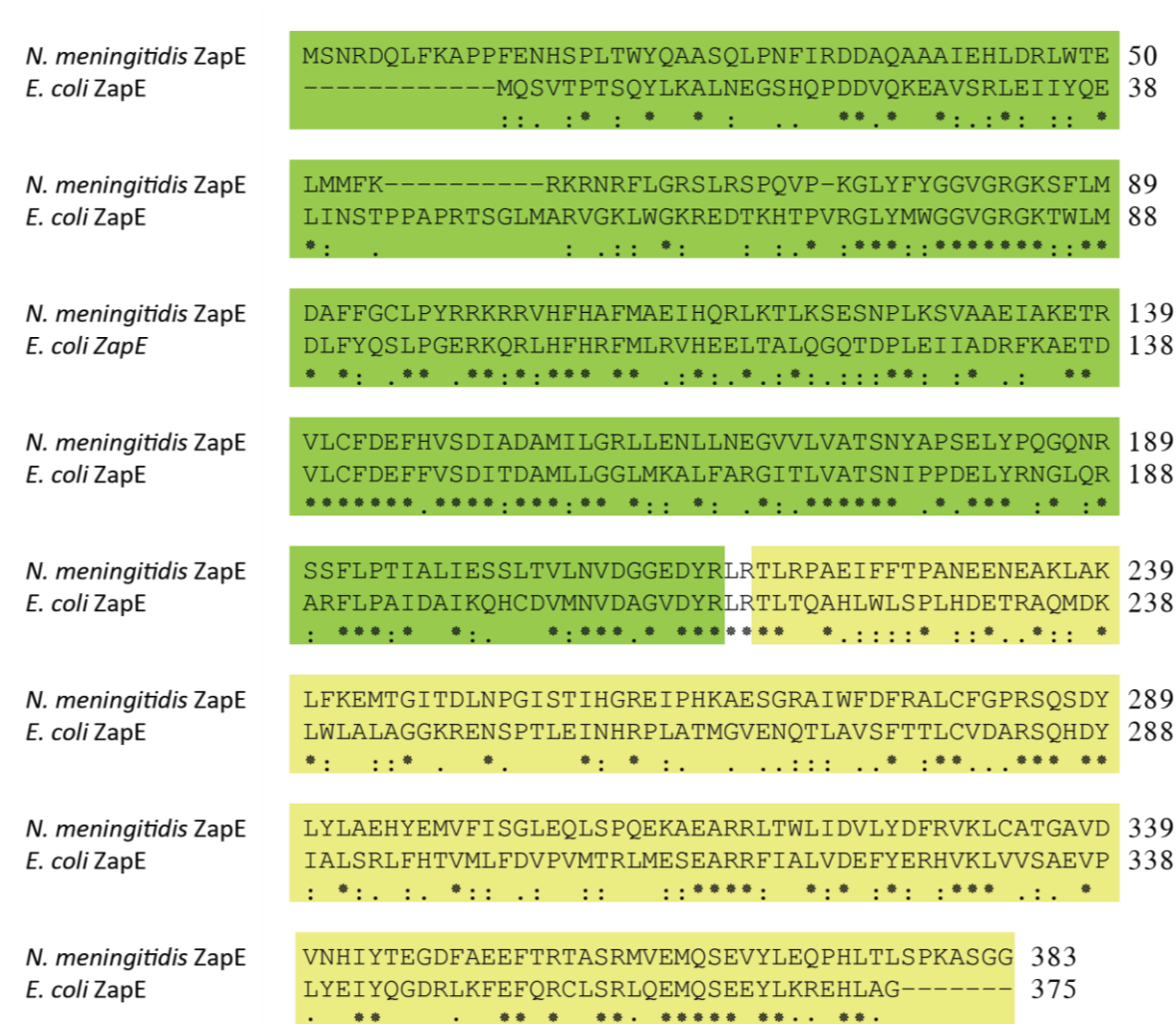


Figure 6.19 Predicted domains of ZapE

The putative domains of ZapE are highlighted in yellow and green.

Overall the domain structure of ZapE can now be interpreted as seen in **Figure 6.19** and this domain assignment correlates well with the original FFAS prediction. Future approaches to structure determination of ZapE should focus on producing the N and C-terminal domains as these smaller proteins are likely to be more amenable to crystallisation.

6.3 Functional Analysis of ZapE

ZapE is reported to be an ATPase which can bind FtsZ as well as modulate its functions in the cell (Marteyn et al., 2014). Many questions remain unanswered from this study regarding precise mechanistic details of ZapE function:

- What are the kinetic parameters of ZapE ATPase activity?
- What is the affinity of the interaction between ZapE and FtsZ?
- Does ZapE bind monomeric or polymerised FtsZ or both?

As such, experiments were designed and conducted to elucidate the role and function of ZapE in molecular detail.

6.3.1 Analysis of the ATPase Activity of ZapE

To measure the specific ATPase activity of ZapE, a protocol using the Biomol Green reagent (Enzo Life Sciences) was devised. Like malachite green based assays, this reagent allows spectrophotometric detection of free phosphate release which is directly proportional to ATP hydrolysis. Unlike such assays, Biomol green is safer to work with, has increased sensitivity to free phosphate and acts as a one-step reagent when performing such assays. In brief, reactions of 90 μ l were set up in triplicate in microtitre plates containing 2 μ M ZapE protein from *E. coli* or *N. meningitidis*, either wildtype or Walker box mutant, K84A or K85A respectively. The reactions were completed in a reaction buffer of 100 mM Tris-HCl pH 8. 5 mM MgCl₂. 10 μ l of an ATP stock ten times the final desired concentration was then added to each well to give a total reaction volume of 100 μ l. Final ATP concentrations of 15.6 μ M to 1 mM were trialled. Control reactions with no protein were run in parallel, also in triplicate. These reactions were then incubated for ten minutes at room temperature before addition of 100 μ l Biomol Green reagent. The reactions were then incubated for a further 5 minutes before A₆₅₀ was measured using a SpectraMax M5 (Molecular Devices).

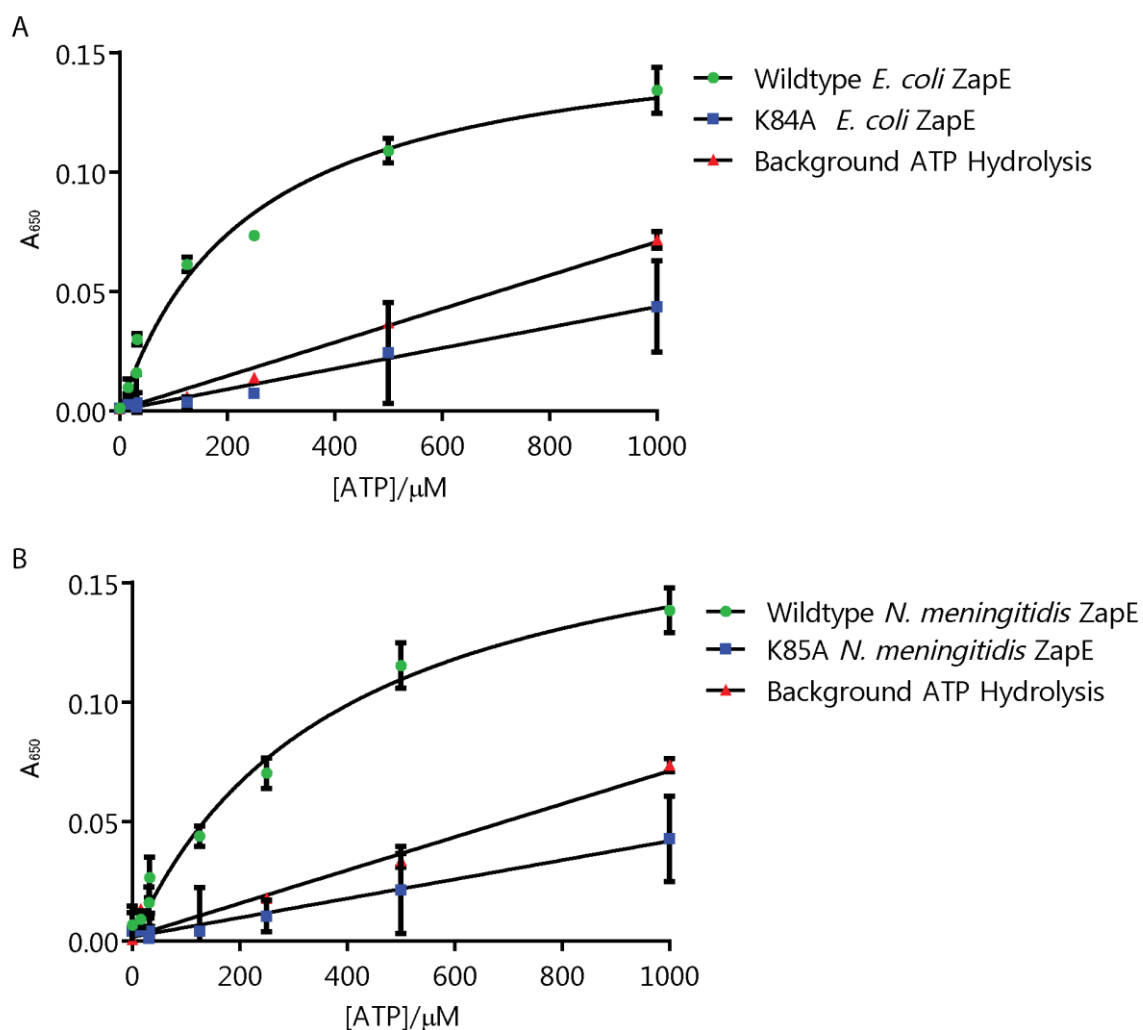


Figure 6.20 A_{650} at a range of ATP concentrations with wild type and mutant *E. coli* and *N. meningitidis* ZapE, compared to background ATP hydrolysis

Example raw data of a Biomol Green ATPase assay from analysis of *E. coli* ZapE, panel A and *N. meningitidis* ZapE, panel B. For both the wild type proteins, the A_{650} reading is significantly above that of the back ground ATP hydrolysis indicating specific ATPase activity.

Example raw data from such an assay can be seen in **Figure 6.20**. Specific ATPase activity is evident for both wild type proteins. Interestingly, the A_{650} reading for the mutant protein is lower than for the control. One explanation of this observation is that the ZapE mutant protein is able to bind ATP but not catalyse the hydrolysis reaction. As such it may stabilise the tri-phosphate form of the nucleotide, therefore giving a lower A_{650} reading than the control reaction monitoring spontaneous ATP hydrolysis.

Each microtitre plate assay had a series of standard phosphate concentrations with no protein added. Therefore, the corresponding A_{650} can be found for a certain concentration of free phosphate, see **Figure 6.21**. This allows calculation of the $V_{\max}$ in terms of nmoles phosphate produced per minute per mg ZapE.

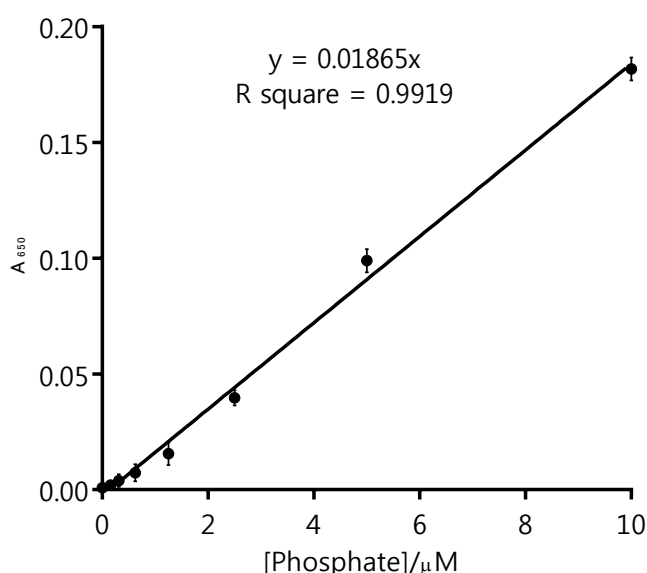


Figure 6.21 Standard curve for free phosphate and A_{650}

For each plate, the background ATP hydrolysis reaction A_{650} readings were then subtracted from the readings for the corresponding reaction containing the wild type ZapE protein. All A_{650} values for phosphate release from ATP specific to hydrolysis by ZapE were then averaged and plotted against the concentration of ATP. The standard error is shown as the error bar. Fitting this data with Michaelis-Menten kinetics allowed calculation of the K_m as a measure of ATP concentration and $V_{\max}$ in terms of A_{650} for each plate. $V_{\max}$ may be converted from A_{650} to units of nmol phosphate produced per minute per mg of protein.

Each plate was repeated in triplicate. The Michaelis-Menten fit for the data collected from each plate can be seen below in **Figure 6.22**.

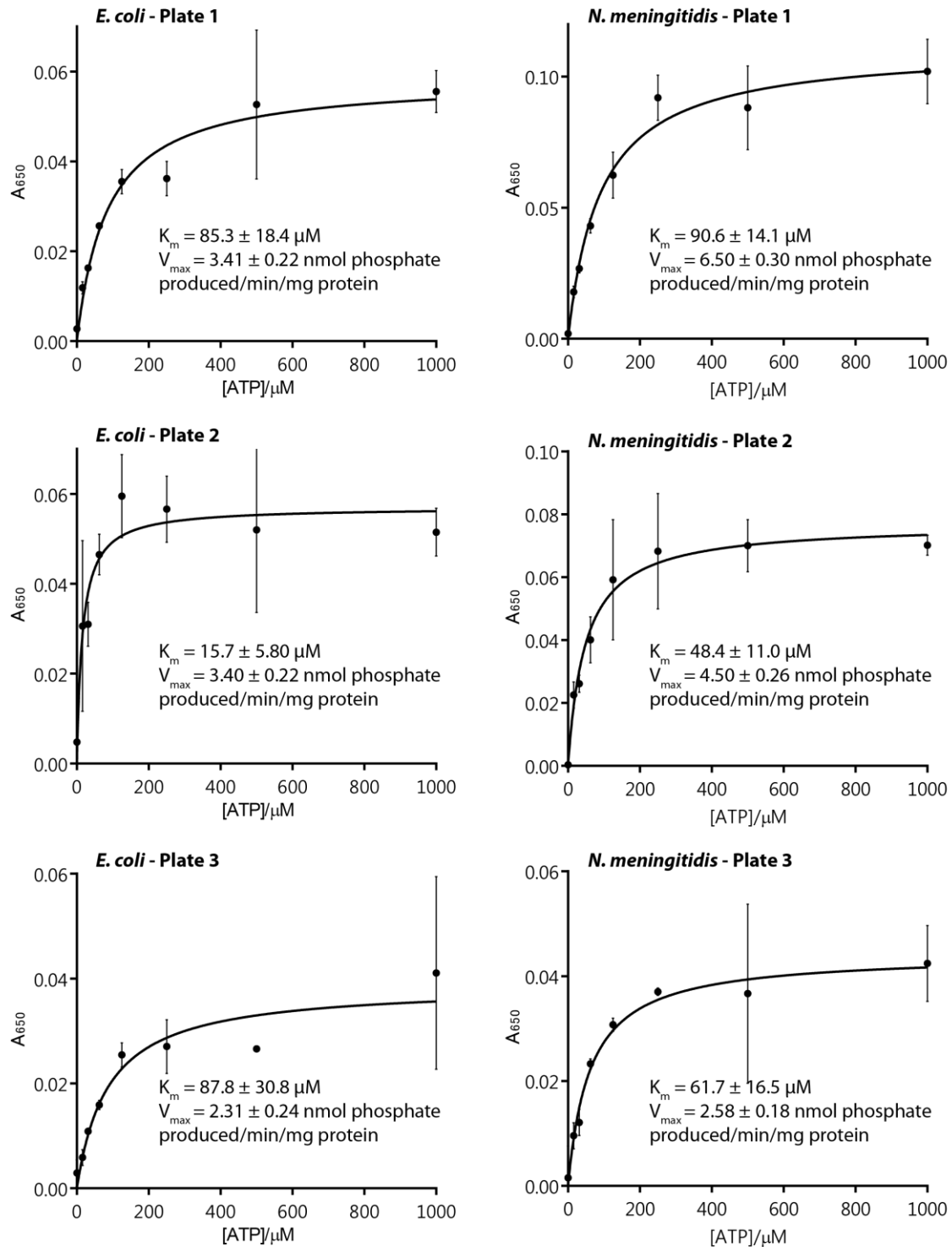


Figure 6.22 Michaelis-Menten fits of the specific change in A_{650} against ATP concentration

In general, the fit for each plate is good with all R^2 values > 0.72 . The error bars are generally larger for reactions with the highest concentrations of ATP. This is most likely due to the increased

error caused by the spontaneous hydrolysis of ATP in the reaction. Plate-to-plate variation with this assay is very large, most likely due to the number of variables which can modulate the A_{650} reading including pipetting errors, errors of timing during the colour development after Biomol Green addition and problems with consistent mixing. The averaged kinetic parameters calculated and their associated errors can be seen in **Table 6.9**.

Kinetic Parameter	<i>E. coli</i> ZapE	<i>N. meningitidis</i> ZapE
K_m (μM ATP)	62.9	66.9
Standard Error K_m (μM ATP)	40.92	21.5
$V_{\max}$ (nmoles phosphate produced/min/mg ZapE)	2.93	3.83
Standard Error $V_{\max}$ (nmoles phosphate produced/min/mg ZapE)	0.65	1.95

Table 6.9 Kinetic parameters of ZapE ATPase activity

Given the large plate-to-plate variation observed, the calculated kinetic parameters are most likely quite inaccurate due to this systematic error, a problem intrinsic to any assay using volatile or unstable substrates such as ATP. Despite the error in the calculated K_m and $V_{\max}$ values, they represent a good approximation which is in the same order of magnitude as other known bacterial cytoplasmic ATPases. The exact values of K_m and $V_{\max}$ themselves are not extremely important as they are calculated in conditions which are not representative of the physiological environment. Overall, the assay shows that both the *E. coli* and *N. meningitidis* versions of this protein have ATPase activity *in vitro*.

6.3.2 Investigations into FtsZ Binding by ZapE

In order to investigate the interaction between FtsZ and ZapE, homogeneous FtsZ was required. The full-length genes of *E. coli* and *N. meningitidis* FtsZ were ordered from GeneArt (Invitrogen) with flanking NcoI and XhoI restriction sites to allow “cut and paste” cloning by the same method described in 6.1.1 into pETM14, a vector allowing expression of the FtsZ protein with a N-terminal hexa-his tag. The expressed sequences are detailed in **Table 6.10**. Using the Ca^{2+} /GTP precipitation method of FtsZ purification (Rivas et al., 2000), *E. coli* FtsZ was successfully purified. Unfortunately this method failed for the expression and purification of *N. meningitidis* FtsZ due to low expression levels. Briefly, pETM14 FtsZ constructs were transformed into BL21 (DE3) *E. coli*. 2 x 10 ml overnight cultures were used to inoculate 2 x 1 l LB and were grown at 37 °C until $\text{OD}_{600} \sim 0.6$. After induction with 0.4 mM IPTG, cells were grown at 37 °C for 3 hours. After cells were harvested, the cell pellet was resuspended in 30 ml 50 mM PIPES pH 6.5, 5 mM MgCl_2 , 1 mM EDTA supplemented with 1 mg/ml lysozyme (Sigma), 400 U/ μl DNase I (Sigma) and EDTA-free protease inhibitor tablets (Pierce) and sonicated. The lysate was clarified by centrifugation at 26,000 $\times g$ for 20 minutes and then 200,000 $\times g$ for 45 minutes. After addition of 1 mM GTP and 20 mM CaCl_2 to precipitate the FtsZ, the samples were incubated at 30°C for 15 minutes and then centrifuged at 10,000 $\times g$ for 15 minutes. The FtsZ pellet was resuspended in 30 ml 50 mM PIPES pH 6.5, 5 mM MgCl_2 , 1 mM EDTA and centrifuged again at 10,000 $\times g$ for 15 minutes. The Ca^{2+} /GTP precipitation-resuspension was repeated once more with the final pellet resuspended in 50 ml 50 mM Tris-HCl pH 8.0, 250 mM KCl, 10 % glycerol. Yields were extremely high as over 500 mg of *E. coli* FtsZ could be produced per litre of culture in this manner.

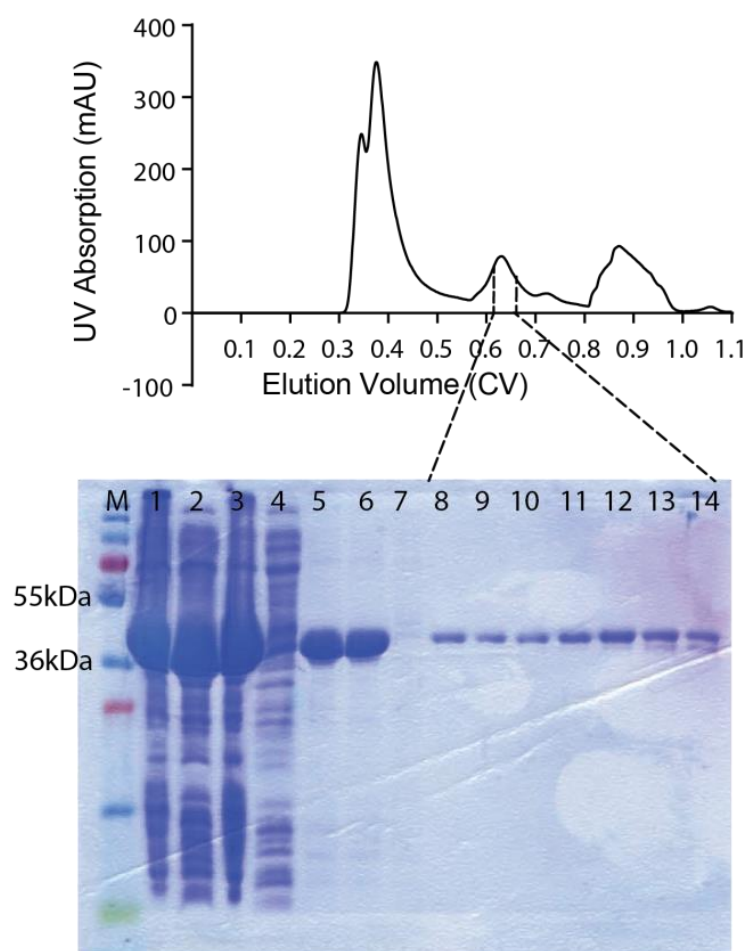
Construct Name	Amino Acid Sequence of Open Reading Frame	Molecular Weight (Da)	Theoretical pI
<i>E. coli</i> FtsZ pETM14	MKHHHHHSAGLEVLFGQ PMGFPEMELTNDAAVIK VIGVGGGGGNAVEHMRERIEGVEFFAVNTDAQAL RKTAVGQTIQIGSGITKGLGAGANPEVGRNAAEDR DALRAALEGADMVFIAAGMGGGTGTGAAPVVAEV AKDLGILTVAVVTKPFNFEGKKRMAFAEQGITELSKH VDSLITIPNDKLLKVLGRGISLLDAFGAANDVLKGA VQGLAELITRPLMNVDFAVTVMSMGYAMMGSG GVASGEDRAEEAAEMAISPLLEDIDLSGARGVLVNI TAGFDLRLDEFETVGNIRAFASDNATVVIGTSLDPD MNDELRTVVATGIGMDKRPEITLVNKKVQVQPV MDRYQQHGMALPTQEQKPVAKVVNDNAPQTAKE PDYLDIPAFLRKQAD-	42562	4.89
<i>N. meningitidis</i> FtsZ pETM14	MKHHHHHSAGLEVLFGQ PMEFVYDVAESAVSPA VIKVIKLGGGGCNAINNMVANNVRGVEFISANTDA QSLAKNHAAKRIQLGTNLTRGLGAGANPDIGRAAA QEDREAIEEAIRGANMLFITTMGGGTGTGSAPVV AEIAKSLGILTVAVVTRPFAYEGKRHVHAQAGLEQLK EHVDSLIIIPNDKLM TALGEDVTMREAFRAADNVLR DAVAGISEVVTCPEIINLDFADVKTVMNSNRGIAMM GSGYAQQIDRRMATDQAISSPLLDDVTLDGARGVL VNITTAPGCKMSELSEVMKIVNQSAHPDLECKFGA AEDETMSDAIRITIIATGLKEKGAVDVFPAREVEAV APSKQEQSHNVEGMIRTNRGIRTMNLTAADFNDQS VLDDFEIPAILRRQHNSDK-	43668	5.28

Table 6.10 Expressed sequences of FtsZ constructs

This table details the names of each of the constructs, the sequence expressed, the molecular weight of that protein product as well as its pI as calculated by ExPASy Protparam. The additional sequence of the N-terminal tag, derived from the pETM14 vector, is highlighted in blue. This contains the hexa-histidine affinity tag as well as the 3C cleavage sequence: LEVLFGQ/GP.

The sequence identity of the two FtsZ constructs is 46 %. Despite a fairly similar pI, such differences in sequence would easily account for the fact that preparations of the meningococcal version of the protein failed. As such, initial interaction studies were only conducted with *E. coli* ZapE and FtsZ. Investigations began looking at whether ZapE interacted with monomeric FtsZ. To purify the monomer, 5 ml of the resuspended FtsZ solution was further treated using Ni-affinity chromatography, as described in **6.1.3**, followed by size-exclusion chromatography using an S200 26/60 column run in 20 mM HEPES pH7.4, 250 mM KCl, 10 % glycerol run at 3 ml/min, see **Figure 6.23**.

Figure 6.23 Size-exclusion profile and SDS-PAGE analysis of the expression and purification of FtsZ from *E. coli*



Fractions taken throughout FtsZ purification: 1 - homogenate, 2 - homogenate supernatant, 3 - ultracentrifuge supernatant, 4 - supernatant after Ca^{2+} /GTP precipitation, 5 - resuspension from Ca^{2+} /GTP precipitation, 6 - supernatant from clarifying spin, 7 - supernatant after Ca^{2+} /GTP precipitation round 2, 8-14 - fractions from across SEC peak after filtered resuspension loaded.

FtsZ has a very low absorption at 280 nm. The Abs 0.1 % (=1 g/l) is approximately 0.1 so the yield is ten times higher than calculated by analysing the trace alone. The peak at 0.65 CV elution, corresponds to monomeric FtsZ. Most of the protein produced is in a higher oligomeric form. However, adding more glycerol to the buffer, whilst dissociating the higher order structures, would render the buffer not suitable for later studies. Given the high yield, it was not a problem to discard such a large percentage of the prepared protein at this stage. This method has previously been validated to produce folded and functional FtsZ so no further characterisation was completed of the FtsZ produced.

The interaction between FtsZ and ZapE was measured by surface plasmon resonance (SPR) using a Biacore 3000 (GE Healthcare). As FtsZ has a tendency to polymerise into higher order structures at

concentrations exceeding 50 μM , this protein was amine coupled to a CM5 chip (GE Healthcare) to minimise the potential for this to occur. After the CM5 chip was docked and primed in running buffer of 10 mM HEPES pH7.5, 150 mM NaCl, 1 mM EDTA 0.005 % Tween-20, 80 μl 50 mM NaOH was injected over all 4 flow channels followed by 120 μl 1:1 mix of NHS:EDC (N-Hydroxysuccinimide: ethyl(dimethylaminopropyl) carbodiimide). 33 μM stock of monomeric *E.coli* FtsZ was diluted 1:10 with 10 mM sodium acetate pH5. Flow channel 1 was left as a reference channel. Flow channels 2, 3 and 4 had approximately 1000, 2000 and 3000 response units (RU) of FtsZ coupled to their surface respectively. Next 120 μl 500 mM ethanolamine pH8.5 was injected over all four channels to quench any remaining activated carboxyl groups. Following equilibration of the chip in running buffer, ZapE injected into the system using the Kinject protocol. Each injection was 120 μl at a flow rate of 40 $\mu\text{l}/\text{minute}$. *E. coli* ZapE samples in running buffer ranged in concentration from 150 μM to 4.6 μM . After each injection, a dissociation time of 600 seconds was allowed and the chip was regenerated with three 20 μl injections of glycine pH3.0. An example trace of the ZapE injection can be seen in **Figure 6.24**.

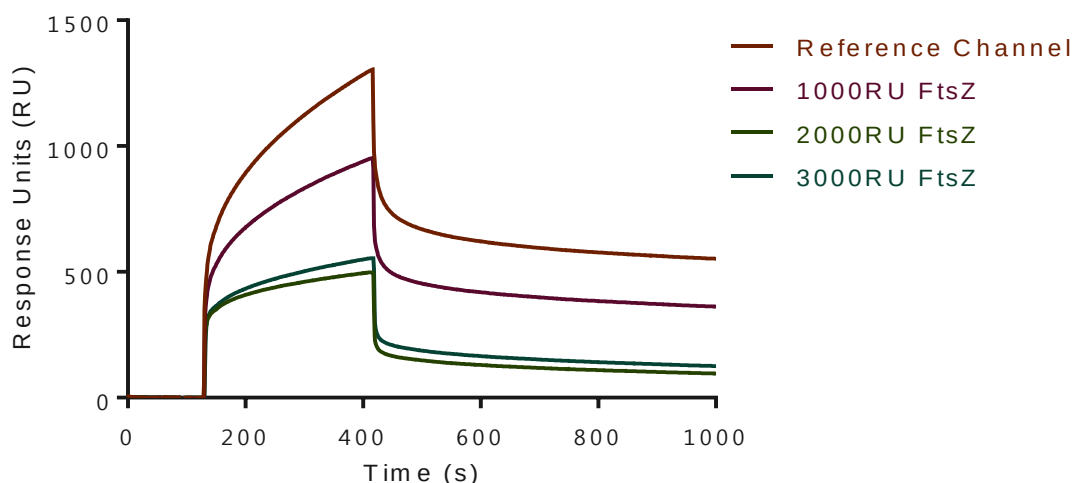


Figure 6.24 SPR trace of FtsZ “binding” ZapE

The trace shows a 120 μl injection of 18 μM *E. coli* ZapE at 40 $\mu\text{l}/\text{min}$ over the chip. The response change is largest for the reference channel and decreases in general with increasing levels of FtsZ on the chip surface. This indicates that ZapE appears to have a stronger interaction for the CM5 chip surface than the FtsZ coupled to the chip.

The trace shown in **Figure 6.24** indicates that ZapE actually interacts more strongly with the CM5 chip matrix than the amine coupled FtsZ on the surface. This is in disagreement with the findings of Marteyn *et al*, the paper discussed throughout this chapter. The CM5 chip surface is covered with a carboxymethylated dextran matrix, rendering the chip surface very hydrophilic and negatively charged. Traces, such as that seen in **Figure 6.24**, are common to very positively charged proteins, in particular DNA-binding proteins. The sequence of *E. coli* ZapE has 46 positively charged residues (arginines and lysines) so such a charged based interaction with the matrix is not unreasonable to postulate. However, the SPR assay is limited as it was only conducted once and with the components of the assay one way round i.e. bound to the chip or in solution.

Attempts to use micro-scale thermophoresis to investigate the interaction between ZapE and FtsZ were unsuccessful due to the lack of a stable fluorescent signal throughout the assay. FtsZ, which was fluorescently labelled, is known to undergo polymerisation with subtle shifts in environmental conditions such as changes to salt concentration or pH which is the most likely reason a stable fluorescent signal was not maintained.

Given the lack of evidence for an interaction between ZapE and FtsZ in these studies and the technical difficulties encountered in conducting such experiments, investigations continued to focus on other functional aspects of ZapE.

6.3.3 Investigations into DNA Binding by ZapE

The top five FFAS hits for both the *E. coli* and *N. meningitidis* ZapE protein sequence are predominantly DNA binding proteins, more specifically with a role in the primosome. Taken together with the SPR results, this suggests that ZapE could function as a DNA binding protein. Initial investigations to look at the potential DNA binding capabilities of ZapE used a basic electrophoretic mobility shift assay (EMSA) protocol. Briefly, each reaction containing 300 ng Miniprep (Qiagen) purified plasmid (pET28b *E. coli* ZapE construct) was incubated with or without

protein and/or nucleotide in a reaction buffer of 50 mM Tris-HCl pH 8.5, 5 mM MgCl₂ in a total volume of 10 µl, for 10 minutes at room temperature. 2 µl 6 x blue/orange loading dye (Promega) was then added to each reaction and then loaded onto a 1 % agarose-TAE gel. The gel was run at 100 V until complete. The gel was then incubated with 1:5000 10 mg/ml ethidium bromide: 1X TAE for 15 minutes or until bands could be visualised by UV using a Gel Doc system (Syngene).

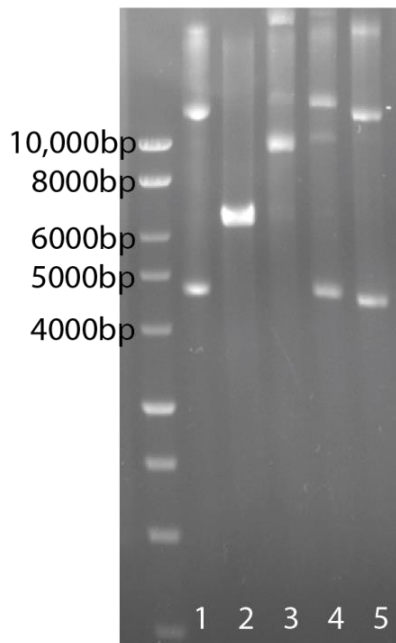


Figure 6.25 Initial EMSA of *E. coli* ZapE with pET28b *E. coli* ZapE plasmid

1 – supercoiled plasmid, 2 – linearised plasmid, 3 – supercoiled plasmid + WT *E. coli* ZapE, 4 – supercoiled plasmid + K84A *E. coli* ZapE, 5 – supercoiled plasmid + MxiGN.

100 pmol of each protein is present in reactions 3, 4 and 5. The molecular weight of the plasmid used in this assay is 1.7 MDa. 300 ng is equivalent to ~180 fmol plasmid, as such the protein is in vast molar excess, by approximately three orders of magnitude. Lane 3 shows a large increase in molecular weight of the band corresponding to supercoiled plasmid, from approximately 4800 bp (3.1 MDa) to 10000 bp (6.5 MDa). A change in molecular weight is also seen for the relaxed plasmid in lane 3 to a higher molecular weight species though this can not be measured accurately. This change is not seen for the ZapE mutant K84A though there does seem to be a smaller shift between the DNA incubated with the mutant protein compared to the control protein. MxiGN is the cytoplasmic N-terminal domain of *Shigella flexneri* protein MxiG, a structural component of the type III secretion system which is solubly expressed in *E. coli*. As this

protein most certainly has no function in DNA binding, it was used as a control protein in these assays.

The increase in molecular weight of the plasmid bands seen by comparing lanes 1 and 3 of **Figure 6.25** is very dramatic, 3.1 MDa to 6.5 MDa approximately. This change in apparent molecular weight may represent the bulk binding of the ZapE protein molecules to the plasmid or a DNA restructuring event such as unwinding or nicking of the plasmid. The event does not correspond to plasmid linearisation as the bands are too large to correspond to this. Interestingly the same change is not observed for the Walker box mutant though there does seem to be a smaller difference in band molecular weight compared to that of the control protein reaction in lane 5.

To try and further understand these results, further EMSAs were conducted to investigate binding to linearised plasmid compared to supercoiled plasmid, the presence of various nucleotides, binding to different plasmids as well as different incubation conditions. The results of all of these assays can be seen in **Figure 6.26**. All of these assays used the wild type *E. coli* ZapE protein and pET28b *E. coli* ZapE plasmid unless otherwise stated.

One of the key features of these analyses is the relative distribution between relaxed and supercoiled plasmid which alters between different EMSAs. Lane 1 in **Figure 6.25** is fundamentally the same reaction in terms of components, the associated concentrations and plasmid used, as **Figure 6.26A** lane 4, **Figure 6.26B** lane 11 and **Figure 6.26C** lane 1 pET. DNA supercoiling and gel migration is very much dependent on precise salt concentration so perhaps the variation between different plasmid preparations was sufficient to incur this change.

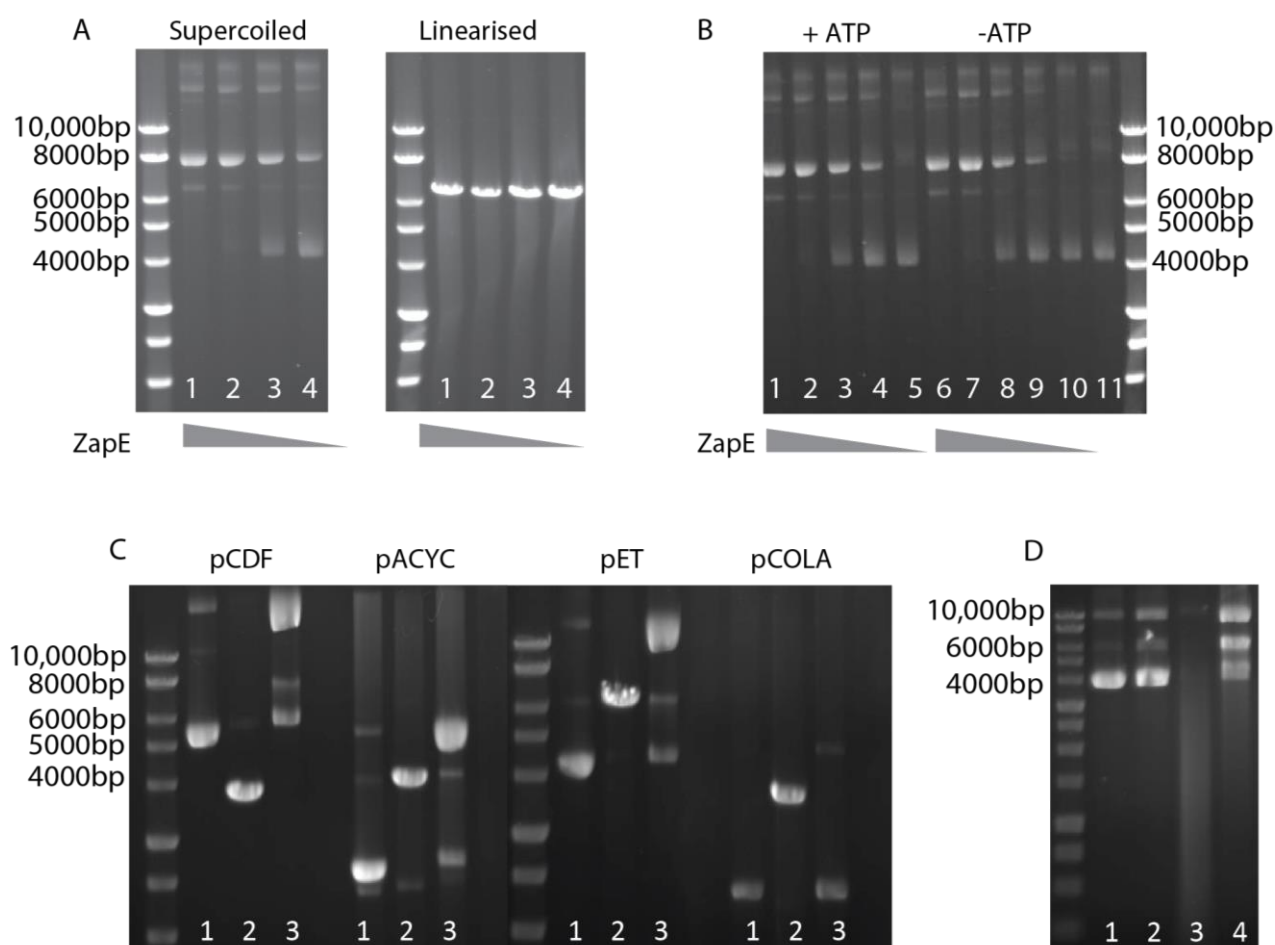


Figure 6.26 EMSAs conducted to investigate the DNA binding properties of *E. coli* ZapE

Panel A: For both gels depicted, 300 ng plasmid was incubated with 1 – 100 pmol, 2- 50 pmol or 3 – 10 pmol ZapE protein. A clear and discrete band shift can be seen for all reactions conducted with supercoiled plasmid but not with linearised plasmid.

Panel B: For all reactions, 300 ng supercoiled plasmid was incubated with 100 pmol, 50 pmol, 10 pmol, 5 pmol or 1 pmol ZapE protein (corresponding to lanes 1 to 5 and 6 to 10). Lane 11 shows plasmid incubated with no protein. Lanes 1 to 5 were incubated with 1mM ATP. As lanes 1-5 and 6-10 show the same pattern of DNA bands, it can be concluded that addition of ATP has no effect on DNA binding by ZapE.

Panel C: For each different plasmid, 1 – supercoiled plasmid, 2 – linearised plasmid and 3 – supercoiled plasmid with 100 pmol ZapE protein. In the case of all four vectors, a band shift is evident when comparing lanes 1 and 3 though this is less obvious with pCOLA.

Panel D: All reactions were incubated at 37 °C for 20 minutes. 1 – supercoiled plasmid, 2 – supercoiled plasmid + MxiGN, 3 – supercoiled plasmid + *E. coli* WT ZapE, 4 – supercoiled plasmid + *E. coli* K84A ZapE. Reactions 1 and 2 show equivalent DNA band patterns whilst 3 shows a smear of DNA and 4 shows a shift in the bands compared to the controls. The smear seen in 3 is indicative of nuclease activity in the wild type ZapE protein solution added to this reaction. As seen in **Figure 6.25**, the wild type and mutant versions of ZapE are not acting in the same way on the DNA but both are having a marked effect compared to the control reactions.

The results of the EMSA shown in **Figure 6.26A** indicate that the band shift is only observed for supercoiled DNA. However, a single binding event (increase in mass of 46 kDa) would not be easily resolved on such a gel. As such, if the wildtype protein is binding the supercoiled plasmid and then performing some form of DNA manipulation function, for example acting as a gyrase in unwinding the supercoiling of the plasmid, then one could postulate the initial binding event could still occur for linearised DNA but any subsequent activity would not. As such, it can not be ruled out that ZapE does not bind linearised DNA.

Given the different band shift pattern observed in the EMSA depicted in **Figure 6.25** between the wild type and K84A mutant of *E. coli* ZapE, it was decided that the next assay should investigate the effect of ATP on the reaction. **Figure 6.26B** shows no significant difference between reactions incubated with or without ATP. This implies a greater complexity in the role of the Walker box in the DNA binding mechanism of ZapE.

Next, given the discrete nature of the plasmid band shift when incubated with wild type *E. coli* ZapE, it was postulated that ZapE was acting at a discrete sequence on the plasmid. It was thought a possible candidate for this sequence was the plasmid origin, especially as the top FFAS hits are all primosome components, involved in DNA replication in bacteria. As such, ZapE was incubated with four plasmids from the Duet series (Novagen), each with a different origin. **Figure 6.26C** shows that in each case, DNA binding was observed with a similar phenotype to that seen with the pET vector originally investigated. This does not necessarily imply that the ZapE protein is not acting at a specific sequence as the four different origins of the different vectors may have a common short sequence motif, see **Figure 6.27**.

Lastly, supercoiled plasmid was incubated with the control MxiGN protein, wild type ZapE and K84A ZapE at 37 °C for 20 minutes. This represents more physiological conditions than incubation at room temperature. The results of this can be seen in **Figure 6.26D**. No change is observed for the control protein reaction. Degradation of the plasmid DNA is evident for the wild type ZapE

plasmid incubation only and a more marked band shift phenotype is observed for the K84A mutant. Either, this nuclease activity is from a contaminating protein specific to the wild type *E. coli* ZapE expression and purification, or this is activity specific to the ZapE itself. Designing an experiment in which it is shown that the nuclease activity is specific to ZapE would be extremely difficult. However, the mutant ZapE reaction shows formation of the higher molecular weight species seen in other wild types assays (compare lanes 1 and 3 (pCDF) or panel C with lanes 1 and 4 of **Figure 6.26D**). This could indicate that the mutant protein is still functional and has a similar DNA binding capacity compared with the wild type ZapE, but the rate at which it performs this function is greatly reduced and is only seen at elevated temperatures compared to the other assays carried out.

The next EMSA performed investigated the DNA binding activity of the *N. meningitidis* form of the ZapE protein. The results from this assay can be seen in **Figure 6.28**. For all reactions, irrespective of whether the protein added is wild type or mutant, the plasmid supercoiled or linearised or the presence of any nucleotide tested, extremely large protein-DNA complexes are visible. The bands seen represent molecular weights far greater than 6.5 MDa (10,000 bp marker). As such, any conclusions drawn from the mechanism of action of the *E. coli* ZapE, do not apply to the *N. meningitidis* version of the protein. The results from the EMSA shown in **Figure 6.28** could be interpreted as non-specific binding of the plasmid by all available molecules of ZapE in the reaction. The ratio of plasmid:protein is 1:555 in terms of numbers of molecules. If there was an equal number of protein molecules bound to every copy of plasmid, the expected mass increase would be 25 MDa, which is in line with the band shift observed. Some complexes are better resolved than others on this gel and some do appear to be of different molecular weights. However, a 1 % agarose-TAE gel is a poor method to resolve such large complexes and gel migration of such a large species may even be affected by the well shape or minor gel imperfections. Such non-specific binding could be observed as no meningococcal origin is present on the plasmid so a non-specific charge based interaction dominates as the observed phenotype.

```

pACYC      GCGCTAGCGGAGTGATATACTGGCTTACTATGTTGGCACTGATGAGGGTGTCACTGAAGTG 60
pCOLA      -----
pET         -----
pCDF        -----

pACYC      CTTCATGTGGCAGGAGAAAAAGGCTGCACCGTCCGTGACGAGAATATGTATACAGGA 120
pCOLA      -----
pET         -----
pCDF        -----

pACYC      TATATCCGCTTCCTCGCTCACTGACTCGCTACGCTCGGTGCTGCTGCGCGAGCGG 180
pCOLA      -----
pET         -----
pCDF        -----GCGCTG 6

pACYC      AAATGGCTTACGAACGGCGGAGATTCTTGAAGATGCCAGGAAGATACTTAACAGG- 239
pCOLA      -----AAACG-----TCTAGAAGATGCCAGGAGGATACTTAGCAGA- 37
pET         -----
pCDF        CGGACACATACAAAGTTACCCACAGATTCCGTGGATAAGCAGGGGACTAACATGTGAGGC 66

pACYC      GAAGTGAGAGGGCGCGGCAAGCCGTTTTTCCATAGGCTCCGCCCTTGCACAGCATC 299
pCOLA      GAGACAATAGGCGGAGCGAAGCCGTTTTTCCATAGGCTCCGCCCTTGCACAGCATC 97
pET         -----CGCGTGTGCGGTTTTTCCATAGGCTCCGCCCTTGCACAGCATC 47
pCDF        AAAACAGAGGGCGCGCGGTTTTTCCATAGGCTCCGCCCTTGCACAGCATC 126
          ** * * * * *

pACYC      ACGAAATCTGACGCTC-----AAATCA-GTGGT----- 326
pCOLA      ACGAAATCTGACGCTC-----AAATCA-GTGGT----- 124
pET         ACAAATCTGACGCTC-----AAGTCA-GAGGT----- 74
pCDF        ACATAACAGAGCGCTTTCCGGTGCATCTGTGGAGCCGTCAACCATGAATCTG 186
          ** * * * * *

pACYC      -----GGCGAAACCGACAGGACTATAAGATACCAGGCGTTTCCCTTGGCGGCTC 377
pCOLA      -----GGCGAAACCGACAGGACTATAAGATACCAGGCGTTTCCCTTGTATGGCTC 176
pET         -----GGCGAAACCGACAGGACTATAAGATACCAGGCGTTTCCCTTGGAAAGCTC 125
pCDF        ACAGTACGGGCGAAACCGACAGGACT-TAAAGATCCCAACCGTTTCCGGCGGCTGCTC 245
          * * * * *

pACYC      CCTCGTGCCTCTCCTGTTCTGCTTTCCGGTTTACCGGTGTC--ATTCCGCTGTAT-- 433
pCOLA      CCTCTTGGCTCTCCTGTTTCCCGTCTGCGG-----CG-----TCCG-TGTTGT-- 219
pET         CCTCGTGCCTCTCCTGTTCCGACCTGCCGCTTACCGGATACCTGT-CCGCTTCTTCC 184
pCDF        CCTCTTGGCTCTCCTGTTCCGACCTGCCGTTTACCGGATACCTGTCTCGGCTTCTCC 305
          * * * * *

pACYC      -----GGCGCGGT---TTGTCTCAT--TCCACGC--CTGACA-CTCAGTTCGGGTAG 478
pCOLA      -----GGTGGAG-----GCTTAC--CCAAATC--ACCACG-TCCCGTTCGTGTAG 261
pET         CTT-CGGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGGTAG 243
pCDF        CTTACGGGAAGTGTGGCGCTTTCTCATAGCTCACAC-ACTGGTATCTCGGCTCGGTGTAG 364
          * * * * *

pACYC      GCAGTTCGCTCCAAGCTGGACTGTATGCACGAACCCCGCTTCAGTCCGACCGGTGCGCC 538
pCOLA      ACAGTTCGCTCCAAGCTGGGCTGTGTGAAGAACCCCGTTCAAGCCGACTGCTGCGCC 321
pET         GTCGTTCGCTCCAAGCTGGGCTGTGTGACGAACCCCGTTCAAGCCGACTGCTGCGCC 303
pCDF        GTCGTTCGCTCCAAGCTGGGCTGTGAAGAAGTCCCGTTCAAGCCGACTGCTGCGCC 424
          * * * * *

pACYC      TTATCCGGTAACATATCGTCTTGAGTCCAACCCGGAAGACATGCAAAAGCACCCTGGCA 598
pCOLA      TTATCCGGTAACATATCATCTTGAGTCCAACCCGGAAGACACGCAAAAGCACCCTGGCA 381
pET         TTATCCGGTAACATATCGTCTTGAGTCCAACCCGGAAGACACGACTTATCGCACTGGCA 363
pCDF        TTATCCGGTAACATGTTCACTTGAGTCCAACCCGGAAGACAGCGTAAACGCACTGGCA 484
          * * * * *

pACYC      GCAGCCACTGGTAATTGA--TTTAGAGGAG-----TTAG-----TC 632
pCOLA      GCAGCCATTGGTAACATGAGAAATAGTGGAT-----TTAGATA-----TCGAGAGTC 427
pET         GCAGCCACTGGTAACAGGA--TTAGCAGAGCAGGTATGAGGCGGTGTACAGA-GTTC 420
pCDF        GCAGCCATTGGTAACATGGGAGTTCGACAGAG-GATTGTTAGCTAAACACGCGGTGCTC 543
          * * * * *

pACYC      TTGAAGTCATGCGCCGTTAAGGCTAAAGTAAAGTAAAGGACAAGTTTGGTGACTGCGCTCCT 692
pCOLA      TTGAAGTGGTGGCTAACAGAGGCTACACTGAAAGGACAGTATTGGTATCTGCGCTCCA 487
pET         TTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTG 480
pCDF        TTGAAGTGTGCGCCAAAGTCCGGCTACACTGGAAGGACAG-ATTGGTGTGTGCTCTCTG 602
          * * * * *

pACYC      CCAA-GCCAGTTACCTCGGTTCAAGAGTTGGT-AGCTCAGAGAACCTT-CGAA-AAACC 748
pCOLA      CTAAGCCAGTTACCT--AGGTTAAGCAGTTCCCAACTGACTTAACCTT-CGATCAAAAC 544
pET         CTGAAGCCAGTTACCTTCGAAAAGAGTTGGT-AGCT--CTTGATCCGCAAAACAAACC 537
pCDF        CGAAAGCCAGTTACCA-CGGTTAAGCAGTTCCCAACTGACTTAACCTT-CGATCAAAAC 660
          * * * * *

pACYC      GCGCTGC-AAG--GCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGAC-CAAAAC 804
pCOLA      GCCTCCC-CAG--GCGGTTTTTTCGTTTTACAGAGCAGGAGATTACG-ACGATCGTAAAG 600
pET         ACCGCTGGTAGCGGTGGTTTTTTTGTGTCAG-CAGCAGATTACGCGCAGAA-AAAAAG 595
pCDF        ACCTCCC-CAG--GTGGTTTTTTCGTTTTACAGGCAAAAGATTACGCGCAGAA-AAAAAG 716
          * * * * *

pACYC      GATCTCAAGAAGATCATCTTATTAATCAGATAAAATATTCTAGATTCAGTGCATTTA 864
pCOLA      GATCTCAAGAAGATCCT-----TTACGGATT----- 626
pET         GATCTCAAGAAGATCCT-----TT--GATCT----- 619
pCDF        GATCTCAAGAAGATCCT-----TT--GATC----- 739
          * * * * *

pACYC      TCTCTCAAATGTAGCACCTGAAGTCAGCCCCATACGATATAAGTTGTA 913
pCOLA      -----CCCGACACCA----- 636
pET         -----
pCDF        -----

```

Figure 6.27 Alignment of the vector origins of pCDF, pCOLA, pACYC and pET

Whilst the origin sequences of the various vectors are quite different, there are commonly shared motifs. These could be common binding sites for ZapE.

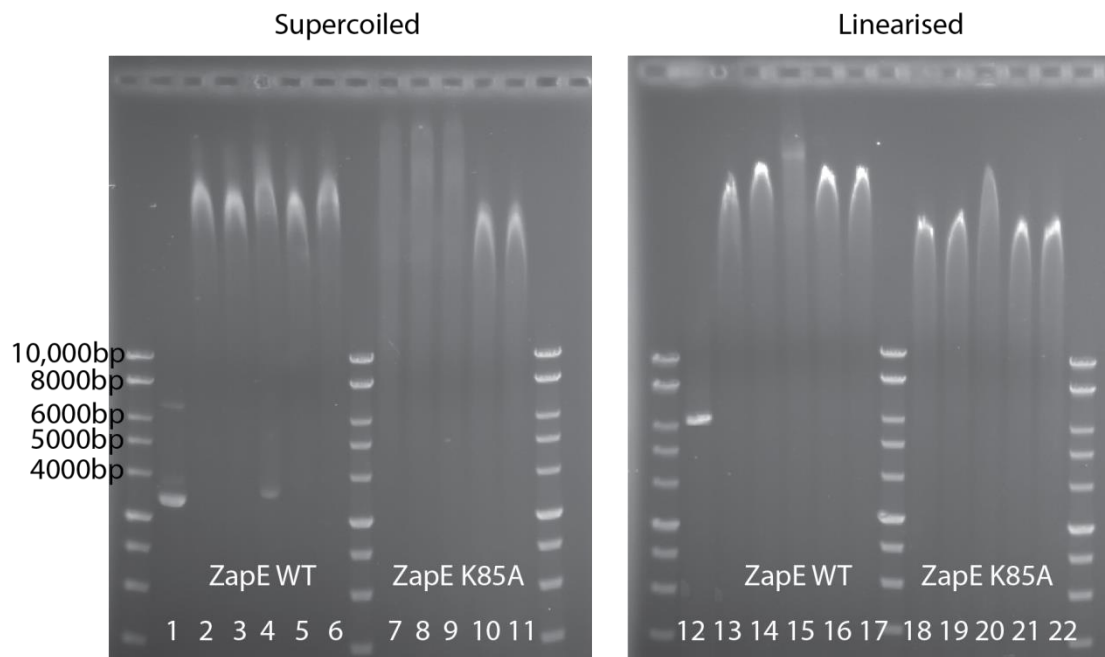


Figure 6.28 EMSA conducted to investigate the DNA binding properties of *N. meningitidis* ZapE
 Lanes 1-11 show reactions with supercoiled plasmid, lanes 12-22 show reactions with linearised plasmid. 1 and 12 have no protein, whilst all other reactions have 100pmol of the *N. meningitidis* ZapE protein indicated. Reactions 3, 8, 14 and 19 have 1 mM ADP; reactions 4, 9, 15 and 20 have 1 mM ATP; reactions 5, 10, 16 and 21 have 1 mM AMP-PCP; reactions 6, 11, 17 and 22 have 1 mM GTP.

Rationalising any of these results with those originally reported is very difficult (Marteyn et al., 2014). A thorough discussion of the lack of compatibility of the data is found in 6.4.

6.4 Perspectives

ZapE of both *E. coli* and *N. meningitidis*, both the wildtype and Walker box A mutants, were successfully expressed and purified in their full-length forms. The two domains of this protein were mapped for the *N. meningitidis* version of this protein by FFAS, limited proteolysis and N-terminal sequencing analysis. Attempts were made to crystallise the wild type *N. meningitidis* form of this protein though this has yet to prove successful. The domain analysis completed should help guide future attempts of crystallisation by focussing on the smaller N and C terminal modules.

The ATPase activity of the protein from both species was characterised, allowing calculation of both the K_m and V_{max} for the wild-type proteins. The Walker box mutant showed no ATPase activity and in fact appeared to stabilise ATP by preventing hydrolysis, reducing free phosphate release compared to background. As such, mutant versions of the ZapE proteins bound to nucleotide, could prove better candidates for crystallography trials.

The reported FtsZ binding function of ZapE was not reproduced in SPR experiments, in fact showing ZapE had a greater affinity for the negatively charged carboxymethylated dextran matrix of the chip surface than monomeric FtsZ. EMSA analysis shows that wild-type *E. coli* ZapE is able to bind supercoiled plasmid DNA. The discrete band shift observed in EMSA analysis is indicative of a specific DNA modulating event such as unwinding of supercoiled elements or nicking. A small shift is seen for the Walker box mutant compared to reaction run with control protein, a very different phenotype compared to the wild type. Despite different phenotypes being observed for the wild-type protein compared to the mutant, the addition of ATP to the reaction did not alter the observed band shift for the wild type protein.

One theory which would rationalise these results is that a fraction of the *E. coli* wildtype ZapE is purified bound to ATP in an “activated” form. It can then bind the plasmid and function as a gyrase in an ATP-dependent manner giving the observed band shift. Given that all EMSA reactions

are conducted with ZapE in vast molar excess (approximately 550 protein molecules for every plasmid molecule), if only one protein molecule is necessary to relax the plasmid in this manner, only 1 in 550 ZapE (0.2 %) molecules would need to be purified in this activated form to observe the phenotype seen in the EMSA results. The mutant version of the ZapE could also be purified with a fraction activated and bound to ATP. As the mutant is unable to hydrolyse ATP, it could bind the plasmid but not go on to perform the ATP-hydrolysis dependent gyrase activity. Assuming this hypothesis is correct and given that the difference in mass of the bands in the reaction run with *E. coli* K84A ZapE and MxiGN can be resolved on a 1 % agarose TAE gel, this indicates that the mass difference must be fairly large. As such it may not represent a 1:1 binding event of plasmid:ZapE. Higher resolution EMSA analysis is required to measure the mass difference between the bands of these two reactions.

The band shift is not apparently specific for the origin though may bind a specific sequence motif shared by the four different origins of the different vectors tested. More confusingly, EMSA analysis where reactions were incubated at 37 °C shows that the wild type protein has putative nuclease activity though it is not clear if this activity is intrinsic to the ZapE or a contaminant from the protein preparation. To add further complication to understanding ZapE function in the cell, the *N. meningitidis* form of the protein is able to non-discriminately bind supercoiled and linearised DNA in both its wild type and K85A mutant form.

Results from these investigations support the published findings of Marteyn et al only in the observed ATPase activity of ZapE (Marteyn et al., 2014). DNA binding and modulation is a novel function of ZapE. It is unlikely that FtsZ binding and DNA binding are compatible functions of ZapE. At the stage of Z-ring formation, DNA replication has terminated though nucleoid occlusion is ongoing (Lutkenhaus et al., 2012). However, postulating a role which rationalises the two functions in this context is difficult.

Given the large number of charged residues in the ZapE sequence from both species, 95 in both sequences, ZapE is likely to non-specifically bind proteins, as it non-specifically bound the SPR chip matrix. Pull-down analysis and co-localisation fluorescence microscopy are highly error-prone methods for determining a specific binding event. However, it must be considered that ZapE may only bind FtsZ in the GTP bound state or the protofilament state, neither of which were investigated in this work. Given all top FFAS hits are primosomal components, it is perhaps more convincing that ZapE is an ATPase with DNA binding functions. Given its putative role in bacterial cell division, perhaps it acts at the point of DNA replication, for example initiating the next round of DNA replication in bacteria. In *E. coli*, an attractive hypothetical role of ZapE would be to act on the *E. coli* virulence plasmid at the point of replication as this fits the timings observed for ZapE expression as well as playing a role in bacterial pathogenesis during host colonisation. However, *N. meningitidis* has no virulence plasmid or other DNA elements aside from the chromosomal so this would not be a conserved function. To validate any of these hypotheses, thorough further investigation is required.

This work has contributed greatly to the production of well characterised reagents to continue examining the function of ZapE. From a structural perspective, future experiments should focus on crystallisation of the N and C terminal domains of ZapE. The N-terminal domain contains the Walker motifs and as such should be sufficient for the observed ATPase activity. However, it would be of great interest to see which domains contribute to the DNA binding and/or modulating activity of ZapE from both species investigated. Elucidating the exact function of ZapE will be a challenge but attempts to find true interaction partners should provide insight into its mechanism of action.

7. Thesis Summary

This thesis describes an investigation into *N. meningitidis* pathogenesis using biophysical methods. Three distinct areas were investigated: the role of factor H sequestration for immune evasion and the host specific adaptations of *N. meningitidis* to do so; T4P and their function in meningococcal adhesion as well as an investigation into pilin cytotoxicity and immunogenicity; the colonisation factor ZapE which had a putative role in enhancing anaerobic bacterial growth.

N. meningitidis has long been described as a human specific pathogen (Granoff et al., 2009a). Solution of the structure of m-fH₆₇ allows for comparison with h-fH₆₇ and as such, an appreciation for meningococcal adaptations of the surface exposed protein fHbp to specifically bind the human form of this protein. This is one of many specific adaptations of *N. meningitidis* to evade the human immune system, though the only one characterised at this level of atomic detail in published literature (Johnson et al., 2012).

T4P have long been implicated as crucial virulence factors for *N. meningitidis*. Now, the first structure of a meningococcal major pilin has been solved revealing subtle differences to the *N. gonorrhoeae* PilE structure. More accurate modelling of meningococcal pili is now possible which reveals how slight changes in sequence in the hypervariable D-region of the protein, has a large impact on the surface chemistry of the pilus. For the first time, direct interaction studies were conducted with the putative pilus receptor CD46 and purified pilin using SPR and NMR, neither of which could detect an interaction. With the development of a protocol to produce large amounts of class I and class II pilins recombinantly, such studies may now be repeated with the recently described candidate pilus receptor CD147 (Bernard et al., 2014). Extensive investigations into pilin cytotoxicity proved that monomeric pilin is not cytotoxic in of itself, in contradiction to other published studies (Dunn et al., 1995), but most likely plays a role in the T4P assembly by increasing the proximity of meningococci to host epithelia allowing the action of other bacterial surface cytotoxic factors such as LPS. Pilin proteins were shown to be immunogenic, with the

immunodominant epitope most likely mapping to the D-region. However the sera produced from mouse immunisations were shown to be non-functional, perhaps suggesting why class II pilins do not face pressure from the human immune system to rapidly modulate their sequence through gene conversion events like class I pilins.

N. meningitidis PilC C-terminal domain was recombinantly produced, which has not been achieved in any published literature. However, the quality of the prepared protein and a lack of validation of its identity by mass spectrometry prevented any further studies.

ZapE was shown to be a novel monomeric ATPase which contrary to published literature did not bind FtsZ by the methods used (Marteyn et al., 2014). Instead a DNA modulation activity was shown although the precise nature or relevance of this function in bacterial pathogenesis and bacterial cell division remains unclear.

Overall, this work describes numerous aspects of *N. meningitidis* pathogenesis, all of which are relevant for a greater understanding of disease progression by this bacterium. Many aspects of the meningococcal literature have been addressed, the conclusions of some of which have been overturned through the investigations detailed herein. Future studies should focus on elucidating further the precised mechanisms of T4P-mediated adhesion to host surfaces as well as a thorough analysis of the DNA modulation functions of ZapE and the aspect of cell division biology in which this protein may function.

8. Appendix

8.1 Standard Laboratory Procedures

8.1.1 Standard Method of PCR

PCR set-up for 50µl reaction		
Reagent	Final concentration	Final volume used in successful PCR/µl
10 X Kod polymerase buffer	1 X	5
Forward primer (5 µM)	0.8 µM	2
Reverse primer (5 µM)	0.8 µM	2
Kod dNTP mix (2 mM)	0.2 mM	5
MgCl ₂ (25 mM)	1.5 mM	3
Autoclaved water (aH ₂ O)	-	19.1 or 28.1 depending on amount of template used
Kod polymerase (Merck) (2.5 U/µl)	0.02 U/µl i.e. 1 U per reaction	0.4
Template DNA (1 ng/µl)	0.02 ng/µl or 0.2 ng/µl	1 or 10

Thermocycler conditions	
Temperature and purpose	Time/s
Step 1: 98 °C – Initial denaturation	-
Step 2: 98 °C - Denaturation	15
Step 3: 60 °C – Annealing	2
Step 4: 72 °C – Extension	5
Cycle	Repeat steps 2 to 4, 30 times
Step 5: 72 °C – Final extension	-

Table 8.1 General PCR protocol

8.1.2 Site-Directed Mutagenesis Protocol

Reaction set up		Thermocycler conditions	
Reaction component	Volume /µl	Temperature /°C	Step length /minutes
10 X Pfu Turbo reaction buffer	5	Step 1: 95 °C – Initial denaturation	5
DNA template (10 ng/µl stock)	1	Step 2: 95 °C – Denaturation	0.5
Forward primer (10 µM)	1.25	Step 3: 55 °C – Primer annealing	1
Reverse primer (10 µM)	1.25	Step 4: 68 °C – Primer extension	9
dNTP mix (2 mM)	1	Step 5: 68 °C – Final extension	5
aH ₂ O	36.9	Cycle:	Repeat steps 2 to 4, 15 times
Pfu Turbo (2.5 U/µl)	1		
MgCl ₂ (25 mM)	3		
Total volume	50		

Table 8.2 Site-directed mutagenesis protocol

8.1.3 Transformation Protocol

Step number	Procedure
1	Add 5 µl ligation mix to a 20 µl thawed aliquot of Novablue cells in a 0.6 ml eppendorf, pre-chilled on ice
2	Tap the mixture gently to mix and leave on ice for 30 minutes
3	Heat-shock the cells at 42 °C for 45 s and leave on ice to recover for 5 minutes
4	Add 250 µl SOC media, transfer entire mixture to 15 ml falcon, then incubate at 37 °C with the falcon shaking gently on its side for 1 hour
5	Spin down the cells by centrifuging at 4000 rpm for 10 minutes, then remove the media and resuspend the cell pellet in 100 µl PBS
6	Plate out this mixture onto selective agar (kanamycin 50 µg/ml) and incubate the plates upside-down overnight at 37 °C

Table 8.3 Transformation protocol

8.1.4 TAE Agarose Gel Buffer System

A 50 X TAE buffer stock was made and diluted 1:50 with ddH₂O. The final concentrations of the components were, 40 mM Tris, 20 mM acetic acid, 1 mM EDTA. The pH of the 50X TAE buffer stock was 8.4. DNA samples were combined with 6 x Blue/Orange loading dye (Promega) and loaded onto a 1 % (w/v) agarose gel made with 1X TAE buffer, supplemented with 1 µg/ml ethidium bromide. The gel was run at 70 V in a Mini-Sub Cell GT Cell (Biorad) in 1X TAE buffer supplemented with 0.5 µg/ml ethidium bromide. Bands were visualised in a UV light box (Gel Doc 2000, Biorad). All gels run for EMSA analysis, were run in the absence of ethidium bromide and stained after the gel had finished running in 1X TAE supplemented with 2 µg/ml ethidium bromide.

8.1.5 SDS-PAGE Buffer System

SDS-PAGE gels were made according to the recipe in **Table 8.4**. Samples run on the gel were prepared by mixing with loading dye to a final concentration of 1X. Samples were incubated for 5 minutes at 95 °C and then cooled before being loaded. All gels were run at 200 V in a Mini-Protean Tetra Electrophoresis System (Biorad) in 25 mM Tris, 190 mM glycine and 1 % (w/v) SDS buffer. The gel was stained with Coomassie stain solution (0.12 % (w/v) Coomassie Brilliant Blue

R-250, 40 % (v/v) methanol and 10 % (v/v) glacial acetic acid and then destained in a solution of 30 % (v/v) methanol and 10 % (v/v) glacial acetic acid.

Solution	Component	Final Concentration
1 X Loading Dye	Tris-HCl (pH 6.8) Glycerol SDS Bromophenol Blue DTT	50 mM 10 % (v/v) 2 % (v/v) 0.1 % (w/v) 10 mM
SDS-PAGE 12 or 15 %	<u>Resolving Gel:</u> Acrylamide/methylene bisacrylamide solution 30 % (w/v) Tris-HCl pH 8.8 SDS Ammonium persulphate Tetramethylethylenediamine <u>Stacking Gel:</u> Acrylamide/methylene bisacrylamide solution 30 % (w/v) Tris-HCl pH 6.8 SDS Ammonium persulphate Tetramethylethylenediamine	12 or 15 % (w/v) 40 mM 0.1 % (w/v) 0.1 % (w/v) 0.1 % (w/v) 5.1 % (w/v) 130 mM 0.1 % (w/v) 0.1 % (w/v) 0.1 % (w/v)

Table 8.4 SDS-PAGE loading dye and gel components

9. References

- AAS, F. E., EGGE-JACOBSEN, W., WINTHER-LARSEN, H. C., LOVOLD, C., HITCHEN, P. G., DELL, A. & KOOMEY, M. 2006a. *Neisseria gonorrhoeae* type IV pili undergo multisite, hierarchical modifications with phosphoethanolamine and phosphocholine requiring an enzyme structurally related to lipopolysaccharide phosphoethanolamine transferases. *Journal of Biological Chemistry*, 281, 27712-27723.
- AGENCY FOR MENINGITIS RESEARCH, P. H. E. H. P. 2013. Meningococcal disease statistics.
- AHO, E. L., DEMPSEY, J. A., HOBBS, M. M., KLAPPER, D. G. & CANNON, J. G. 1991. Characterization of the opa (class 5) gene family of *Neisseria meningitidis*. *Molecular Microbiology*, 5, 1429-37.
- ANJUM, M. F., STEVANIN, T. M., READ, R. C. & MOIR, J. W. 2002. Nitric oxide metabolism in *Neisseria meningitidis*. *Journal of Bacteriology*, 184, 2987-93.
- ARCHIBALD, F. S. & DUONG, M. N. 1986. Superoxide dismutase and oxygen toxicity defenses in the genus *Neisseria*. *Infection and Immunity*, 51, 631-41.
- ASHKENAZY, H., EREZ, E., MARTZ, E., PUPKO, T. & BEN-TAL, N. 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research*, 38, W529-33.
- AUDETTE, G. F., VAN SCHAIK, E. J., HAZES, B. & IRVIN, R. T. 2005. DNA-binding protein nanotubes: Learning from nature's nanotech examples (vol 4 pg 1899, 2004). *Nano Letters*, 5, 201-201.
- BARDY, S. L., NG, S. Y. & JARRELL, K. F. 2003. Prokaryotic motility structures. *Microbiology*, 149, 295-304.
- BARLOW, P. N., BARON, M., NORMAN, D. G., DAY, A. J., WILLIS, A. C., SIM, R. B. & CAMPBELL, I. D. 1991. Secondary structure of a complement control protein module by two-dimensional ¹H NMR. *Biochemistry*, 30, 997-1004.
- BARLOW, P. N., STEINKASSERER, A., NORMAN, D. G., KIEFFER, B., WILES, A. P., SIM, R. B. & CAMPBELL, I. D. 1993. Solution structure of a pair of complement modules by nuclear-magnetic-resonance. *Journal of Molecular Biology*, 232, 268-284.
- BEEMINK, P. T. & GRANOFF, D. M. 2009. The modular architecture of meningococcal factor H-binding protein. *Microbiology*, 155, 2873-2883.
- BENTLEY, S. D., VERNIKOS, G. S., SNYDER, L. A., CHURCHER, C., ARROWSMITH, C., CHILLINGWORTH, T., CRONIN, A., DAVIS, P. H., HOLROYD, N. E., JAGELS, K., MADDISON, M., MOULE, S., RABBINOWITSCH, E., SHARP, S., UNWIN, L., WHITEHEAD, S., QUAIL, M. A., ACHTMAN, M., BARRELL, B., SAUNDERS, N. J. & PARKHILL, J. 2007. Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genetics*, 3, e23. [10.1371/journal.pgen.0030023](https://doi.org/10.1371/journal.pgen.0030023)
- BERNARD, S. C., SIMPSON, N., JOIN-LAMBERT, O., FEDERICI, C., LARAN-CHICH, M. P., MAISSA, N., BOUZINBA-SEGARD, H., MORAND, P. C., CHRETIEN, F., TAOUI, S., CHEVET, E., JANEL, S., LAFONT, F., COUREUIL, M., SEGURA, A., NIEDERGANG, F., MARULLO, S., COURAUD, P. O., NASSIF, X. & BOURDOULOUS, S. 2014. Pathogenic *Neisseria meningitidis* utilizes CD147 for vascular colonization. *Nature Medicine*, 20, 725-31.
- BERRY, J. L., CEHOVIN, A., MCDOWELL, M. A., LEA, S. M. & PELICIC, V. 2013. Functional analysis of the interdependence between DNA uptake sequence and its cognate ComP receptor during natural transformation in *Neisseria* species. *PLoS Genetics*, 9, e1004014. [10.1371/journal.pgen.1004014](https://doi.org/10.1371/journal.pgen.1004014)
- BI, E. F. & LUTKENHAUS, J. 1991. FtsZ ring structure associated with division in *Escherichia coli*. *Nature*, 354, 161-4.
- BLANC, E., ROVERSI, P., VONRHEIN, C., FLENSBURG, C., LEA, S. M. & BRICOGNE, G. 2004. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallographica Section D-Biological Crystallography*, 60, 2210-2221.

- BORDOLI, L., KIEFER, F., ARNOLD, K., BENKERT, P., BATTEY, J. & SCHWEDE, T. 2009. Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*, 4, 1-13.
- BREHONY, C., JOLLEY, K. A. & MAIDEN, M. C. 2007. Multilocus sequence typing for global surveillance of meningococcal disease. *FEMS Microbiology Reviews*, 31, 15-26.
- BREHONY, C., WILSON, D. J. & MAIDEN, M. C. J. 2009. Variation of the factor H-binding protein of *Neisseria meningitidis*. *Microbiology*, 155, 4155-4169.
- BRISSAC, T., MIKATY, G., DUMENIL, G., COUREUIL, M. & NASSIF, X. 2012. The meningococcal minor pilin PilX is responsible for type IV pilus conformational changes associated with signaling to endothelial cells. *Infection and Immunity*, 80, 3297-306.
- BROOKS, G. F. & JAWETZ, E. 2013. *Jawetz, Melnick, & Adelberg's medical microbiology*, New York ; London, McGrawHill Medical.
- BURROWS, L. L. 2005. Weapons of mass retraction. *Molecular Microbiology*, 57, 878-88.
- CAPECCHI, B., ADU-BOBIE, J., DI MARCELLO, F., CIUCCHI, L., MASIGNANI, V., TADDEI, A., RAPPUOLI, R., PIZZA, M. & ARICO, B. 2005. *Neisseria meningitidis* NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. *Molecular Microbiology*, 55, 687-98.
- CARTWRIGHT, K. 1995. *Meningococcal disease*, Chichester, Wiley.
- CEHOVIN, A., KROLL, J. S. & PELICIC, V. 2011. Testing the vaccine potential of PilV, PilX and ComP, minor subunits of *Neisseria meningitidis* type IV pili. *Vaccine*, 29, 6858-65.
- CEHOVIN, A., SIMPSON, P. J., MCDOWELL, M. A., BROWN, D. R., NOSCHESSE, R., PALLETT, M., BRADY, J., BALDWIN, G. S., LEA, S. M., MATTHEWS, S. J. & PELICIC, V. 2013. Specific DNA recognition mediated by a type IV pilin. *Proceedings of the National Academy of Sciences U S A*, 110, 3065-70.
- CEHOVIN, A., WINTERBOTHAM, M., LUCIDARME, J., BORROW, R., TANG, C. M., EXLEY, R. M. & PELICIC, V. 2010. Sequence conservation of pilus subunits in *Neisseria meningitidis*. *Vaccine*, 28, 4817-26.
- CHEN, B., TAI, P. C., HARRISON, R. & PAN, Y. 2006. Novel clustering algorithm combined with DSSP post processing for protein sequence motif discovering. *2006 IEEE International Conference on Granular Computing*, 449-452.
- CHEN, V. B., ARENDALL, W. B., HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D-Biological Crystallography*, 66, 12-21.
- CHENG, Y., JOHNSON, M. D., BURILLO-KIRCH, C., MOCNY, J. C., ANDERSON, J. E., GARRETT, C. K., REDINBO, M. R. & THOMAS, C. E. 2013. Mutation of the conserved calcium-binding motif in *Neisseria gonorrhoeae* PilC1 impacts adhesion but not piliation. *Infection and Immunity*, 81, 4280-9.
- CRAIG, L., PIQUE, M. E. & TAINER, J. A. 2004. Type IV pilus structure and bacterial pathogenicity. *Nature Reviews of Microbiology*, 2, 363-78.
- CRAIG, L., TAYLOR, R. K., PIQUE, M. E., ADAIR, B. D., ARVAI, A. S., SINGH, M., LLOYD, S. J., SHIN, D. S., GETZOFF, E. D., YEAGER, M., FOREST, K. T. & TAINER, J. A. 2003b. Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Molecular Cell*, 11, 1139-1150.
- CRAIG, L., VOLKMANN, N., ARVAI, A. S., PIQUE, M. E., YEAGER, M., EGELMAN, E. H. & TAINER, J. A. 2006. Type IV pilus structure by cryo-electron microscopy and crystallography: Implications for pilus assembly and functions. *Molecular Cell*, 23, 651-662.
- DAVIDSEN, T., TUVEN, H. K., BJORAS, M., RODLAND, E. A. & TONJUM, T. 2007. Genetic interactions of DNA repair pathways in the pathogen *Neisseria meningitidis*. *Journal of Bacteriology*, 189, 5728-37.
- DAVIES, J. K., HARRISON, P. F., LIN, Y. H., BARTLEY, S., KHOO, C. A., SEEMANN, T., RYAN, C. S., KAHLER, C. M. & HILL, S. A. 2014. The use of high-throughput DNA sequencing in the

- investigation of antigenic variation: application to *Neisseria* species. *PLoS One*, 9, e86704. [10.1371/journal.pone.0086704](https://doi.org/10.1371/journal.pone.0086704)
- DE BOER, P. A. 2010. Advances in understanding *E. coli* cell fission. *Current Opinion in Microbiology*, 13, 730-7.
- DE CORDOBA, S. R., ESPARZA-GORDILLO, J., DE JORGE, E. G., LOPEZ-TRASCASA, M. & SANCHEZ-CORRAL, P. 2004. The human complement factor H: functional roles, genetic variations and disease associations. *Molecular Immunology*, 41, 355-367.
- DE JONGE, M. I., HAMSTRA, H. J., VAN ALPHEN, L., DANKERT, J. & VAN DER LEY, P. 2003. Mapping the binding domains on meningococcal Opa proteins for CEACAM1 and CEA receptors. *Molecular Microbiology*, 50, 1005-15.
- DE VRIES, F. P., COLE, R., DANKERT, J., FROSCH, M. & VAN PUTTEN, J. P. 1998. *Neisseria meningitidis* producing the Opc adhesin binds epithelial cell proteoglycan receptors. *Molecular Microbiology*, 27, 1203-12.
- DELANO, W. L. 2004. Use of PYMOL as a communications tool for molecular science. *Abstracts of Papers of the American Chemical Society*, 228, U313-U314.
- DUNN, K. L., VIRJI, M. & MOXON, E. R. 1995. Investigations into the molecular basis of meningococcal toxicity for human endothelial and epithelial cells: the synergistic effect of LPS and pili. *Microbial Pathogenesis*, 18, 81-96.
- DURAND-HEREDIA, J., RIVKIN, E., FAN, G., MORALES, J. & JANAKIRAMAN, A. 2012. Identification of ZapD as a cell division factor that promotes the assembly of FtsZ in *Escherichia coli*. *Journal of Bacteriology*, 194, 3189-98.
- EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography*, 60, 2126-2132.
- ESPOSITO, S. & PRINCIPI, N. 2014. Vaccine profile of 4CMenB: a four-component *Neisseria meningitidis* serogroup B vaccine. *Expert Review of Vaccines*, 13, 193-202.
- EVANS, P. 2006. Scaling and assessment of data quality. *Acta Crystallographica Section D-Biological Crystallography*, 62, 72-82.
- FANG, F. C. 2004. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nature Reviews of Microbiology*, 2, 820-32.
- FERNANDO, A. N., FURTADO, P. B., CLARK, S. J., GILBERT, H. E., DAY, A. J., SIM, R. B. & PERKINS, S. J. 2007. Associative and structural properties of the region of complement factor H encompassing the Tyr402His disease-related polymorphism and its interactions with heparin. *Journal of Molecular Biology*, 368, 564-81.
- FIGUEROA, J., ANDREONI, J. & DENSEN, P. 1993. Complement deficiency states and meningococcal disease. *Journal of Immunology Research*, 12, 295-311.
- FINNE, J., FINNE, U., DEAGOSTINI-BAZIN, H. & GORIDIS, C. 1983. Occurrence of alpha 2-8 linked polysialosyl units in a neural cell adhesion molecule. *Biochemical and Biophysical Research Communications*, 112, 482-7.
- FOREST, K. T., BERNSTEIN, S. L., GETZOFF, E. D., SO, M., TRIBBICK, G., GEYSEN, H. M., DEAL, C. D. & TAINER, J. A. 1996. Assembly and antigenicity of the *Neisseria gonorrhoeae* pilus mapped with antibodies. *Infection and Immunity*, 64, 644-52.
- FOREST, K. T., DUNHAM, S. A., KOOMEY, M. & TAINER, J. A. 1999. Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Molecular Microbiology*, 31, 743-52.
- FU, H. 2004. *Protein-protein interactions : methods and applications*, Totowa, N.J., Humana Press.
- GALLI, E. & GERDES, K. 2010. Spatial resolution of two bacterial cell division proteins: ZapA recruits ZapB to the inner face of the Z-ring. *Molecular Microbiology*, 76, 1514-26.
- GARMAN, E. F. & MCSWEENEY, S. M. 2007. Progress in research into radiation damage in cryo-cooled macromolecular crystals. *Journal of Synchrotron Radiation*, 14, 1-3.
- GAULT, J., MALOSSE, C., MACHATA, S., MILLIEN, C., PODGLAJEN, I., PLOY, M. C., COSTELLO, C. E., DUMENIL, G. & CHAMOT-ROOKE, J. 2014. Complete posttranslational modification

- mapping of pathogenic *Neisseria meningitidis* pilins requires top-down mass spectrometry. *Proteomics*, 14, 1141-51.
- GEOFFROY, M. C., FLOQUET, S., METAIS, A., NASSIF, X. & PELICIC, V. 2003. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. *Genome Research*, 13, 391-8.
- GEORGIADOU, M., CASTAGNINI, M., KARIMOVA, G., LADANT, D. & PELICIC, V. 2012. Large-scale study of the interactions between proteins involved in type IV pilus biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus assembly. *Molecular Microbiology*, 84, 857-873.
- GILL, D. B. & ATKINSON, J. P. 2004. CD46 in *Neisseria* pathogenesis. *Trends in Molecular Medicine*, 10, 459-65.
- GILL, D. B., KOOMEY, M., CANNON, J. G. & ATKINSON, J. P. 2003. Down-regulation of CD46 by pilated *Neisseria gonorrhoeae*. *Journal of Experimental Medicine*, 198, 1313-22.
- GILTNER, C. L., HABASH, M. & BURROWS, L. L. 2010. *Pseudomonas aeruginosa* minor pilins are incorporated into type IV pili. *Journal of Molecular Biology*, 398, 444-61.
- GIULIANI, M. M., ADU-BOBIE, J., COMANDUCCI, M., ARICO, B., SAVINO, S., SANTINI, L., BRUNELLI, B., BAMBINI, S., BIOLCHI, A., CAPECCHI, B., CARTOCCI, E., CIUCCHI, L., DI MARCELLO, F., FERLICCA, F., GALLI, B., LUZZI, E., MASIGNANI, V., SERRUTO, D., VEGGI, D., CONTORNI, M., MORANDI, M., BARTALESI, A., CINOTTI, V., MANNUCCI, D., TITTA, F., OVIDI, E., WELSCH, J. A., GRANOFF, D., RAPPUOLI, R. & PIZZA, M. 2006. A universal vaccine for serogroup B meningococcus. *Proceedings of the National Academy of Sciences U S A*, 103, 10834-9.
- GRANOFF, D. M., WELSCH, J. A. & RAM, S. 2009a. Binding of complement factor H (fH) to *Neisseria meningitidis* is specific for human fH and inhibits complement activation by rat and rabbit sera. *Infection and Immunity*, 77, 764-9.
- GRANOFF, D. M., WELSCH, J. A. & RAM, S. 2009b. Binding of complement factor H (fH) to *Neisseria meningitidis* is specific for human fH and inhibits complement activation by rat and rabbit sera. *Infection and Immunity*, 77, 764-769.
- GREENWOOD, D. 2012. *Medical microbiology : a guide to microbial infections : pathogenesis, immunity, laboratory diagnosis and control*, Edinburgh, Churchill Livingstone.
- HAKOBYAN, S., HARRIS, C. L., TORTAJADA, A., DE JORGE, E. G., GARCIA-LAYANA, A., FERNANDEZ-ROBredo, P., DE CORDOBA, S. R. & MORGAN, B. P. 2008. Measurement of factor H variants in plasma using variant-specific monoclonal antibodies: Application to assessing risk of age-related macular degeneration. *Investigative Ophthalmology & Visual Science*, 49, 1983-1990.
- HALE, C. A., SHIOMI, D., LIU, B., BERNHARDT, T. G., MARGOLIN, W., NIKI, H. & DE BOER, P. A. 2011. Identification of *Escherichia coli* ZapC (YcbW) as a component of the division apparatus that binds and bundles FtsZ polymers. *Journal of Bacteriology*, 193, 1393-404.
- HAMMARSTROM, S. 1999. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Seminars in Cancer Biology*, 9, 67-81.
- HAMMERSCHMIDT, S., BIRKHOLZ, C., ZHRINGER, U., ROBERTSON, B. D., VAN PUTTEN, J., EBELING, O. & FROSC, M. 1994. Contribution of genes from the capsule gene complex (cps) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. *Molecular Microbiology*, 11, 885-96.
- HANSEN, J. K., DEMICK, K. P., MANSFIELD, J. M. & FOREST, K. T. 2007. Conserved regions from *Neisseria gonorrhoeae* pilin are immunosilent and not immunosuppressive. *Infection and Immunity*, 75, 4138-47.
- HARRISON, L. H. 2010. Epidemiological profile of meningococcal disease in the United States. *Clinical Infectious Diseases*, 50 Suppl 2, S37-44.
- HARRISON, O. B., CLAUS, H., JIANG, Y., BENNETT, J. S., BRATCHER, H. B., JOLLEY, K. A., CORTON, C., CARE, R., POOLMAN, J. T., ZOLLINGER, W. D., FRASCH, C. E., STEPHENS, D. S., FEAVERS, I.,

- FROSCH, M., PARKHILL, J., VOGEL, U., QUAIL, M. A., BENTLEY, S. D. & MAIDEN, M. C. J. 2013. Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerging Infectious Diseases*, 19, 566-573.
- HAWKEY, P. M. & LEWIS, D. A. 2004. *Medical bacteriology : a practical approach*, Oxford, Oxford University Press.
- HELAINÉ, S., CARBONNELLE, E., PROUVENSIER, L., BERETTI, J. L., NASSIF, X. & PELICIC, V. 2005. PilX, a pilus-associated protein essential for bacterial aggregation, is a key to pilus-facilitated attachment of *Neisseria meningitidis* to human cells. *Molecular Microbiology*, 55, 65-77.
- HELAINÉ, S., DYER, D. H., NASSIF, X., PELICIC, V. & FOREST, K. T. 2007. 3D structure/function analysis of PilX reveals how minor pilins can modulate the virulence properties of type IV pili. *Proceedings of the National Academy of Sciences U S A*, 104, 15888-93.
- HELM, R. A. & SEIFERT, H. S. 2010. Frequency and rate of pilin antigenic variation of *Neisseria meningitidis*. *Journal of Bacteriology*, 192, 3822-3.
- HERBERT, A. P., DEAKIN, J. A., SCHMIDT, C. Q., BLAUM, B. S., EGAN, C., FERREIRA, V. P., PANGBURN, M. K., LYON, M., UHRIN, D. & BARLOW, P. N. 2007. Structure shows that a glycosaminoglycan and protein recognition site in factor H is perturbed by age-related macular degeneration-linked single nucleotide polymorphism. *Journal of Biological Chemistry*, 282, 18960-8.
- HERBERT, A. P., UHRIN, D., LYON, M., PANGBURN, M. K. & BARLOW, P. N. 2006. Disease-associated sequence variations congregate in a polyanion recognition patch on human factor H revealed in three-dimensional structure. *Journal of Biological Chemistry*, 281, 16512-16520.
- HUBERT, K., PAWLIK, M. C., CLAUS, H., JARVA, H., MERI, S. & VOGEL, U. 2012. Opc expression, LPS immunotype switch and pilin conversion contribute to serum resistance of unencapsulated meningococci. *PLoS One*, 7, e45132. [10.1371/journal.pone.0045132](https://doi.org/10.1371/journal.pone.0045132)
- IMHAUS, A. F. & DUMENIL, G. 2014. The number of *Neisseria meningitidis* type IV pili determines host cell interaction. *EMBO Journal*.
- IMREY, P. B., JACKSON, L. A., LUDWINSKI, P. H., ENGLAND, A. C., 3RD, FELLA, G. A., FOX, B. C., ISDALE, L. B., REEVES, M. W. & WENGER, J. D. 1995. Meningococcal carriage, alcohol consumption, and campus bar patronage in a serogroup C meningococcal disease outbreak. *Journal Clinical Microbiology*, 33, 3133-7.
- JACK, D. L., JARVIS, G. A., BOOTH, C. L., TURNER, M. W. & KLEIN, N. J. 2001. Mannose-binding lectin accelerates complement activation and increases serum killing of *Neisseria meningitidis* serogroup C. *Journal Infectious Disease*, 184, 836-45.
- JARVIS, G. A. 1995. Recognition and control of Neisserial infection by antibody and complement. *Trends in Microbiology*, 3, 198-201.
- JARVIS, G. A. & VEDROS, N. A. 1987. Sialic-Acid of Group-B *Neisseria meningitidis* Regulates Alternative Complement Pathway Activation. *Infection and Immunity*, 55, 174-180.
- JOHANSSON, L., RYTKONEN, A., BERGMAN, P., ALBINGER, B., KALLSTROM, H., HOKFELT, T., AGERBERTH, B., CATTANEO, R. & JONSSON, A. B. 2003. CD46 in meningococcal disease. *Science*, 301, 373-375.
- JOHNSON, M. D., GARRETT, C. K., BOND, J. E., COGGAN, K. A., WOLFGANG, M. C. & REDINBO, M. R. 2011. *Pseudomonas aeruginosa* PilY1 binds integrin in an RGD- and calcium-dependent manner. *PLoS One*, 6, e29629. [10.1111/j.1365-2958.2008.06559.x](https://doi.org/10.1111/j.1365-2958.2008.06559.x)
- JOHNSON, S., TAN, L., VAN DER VEEN, S., CAESAR, J., GOICOECHEA DE JORGE, E., HARDING, R. J., BAI, X., EXLEY, R. M., WARD, P. N., RUIVO, N., TRIVEDI, K., CUMBER, E., JONES, R., NEWHAM, L., STAUNTON, D., UFRET-VINCENTY, R., BORROW, R., PICKERING, M. C., LEA, S. M. & TANG, C. M. 2012. Design and evaluation of meningococcal vaccines through structure-based modification of host and pathogen molecules. *Plos Pathogens*, 8, e1002981. [10.1371/journal.ppat.1002981](https://doi.org/10.1371/journal.ppat.1002981)

- JOKIRANTA, T. S., JAAKOLA, V. P., LEHTINEN, M. J., PAREPALO, M., MERI, S. & GOLDMAN, A. 2006. Structure of complement factor H carboxyl-terminus reveals molecular basis of atypical haemolytic uremic syndrome. *EMBO Journal*, 25, 1784-1794.
- KABSCH, W. 2010. Xds. *Acta Crystallographica Section D-Biological Crystallography*, 66, 125-132.
- KAHLER, C. M., MARTIN, L. E., SHIH, G. C., RAHMAN, M. M., CARLSON, R. W. & STEPHENS, D. S. 1998. The ($\alpha_2 \rightarrow 8$)-linked polysialic acid capsule and lipooligosaccharide structure both contribute to the ability of serogroup B *Neisseria meningitidis* to resist the bactericidal activity of normal human serum. *Infection and Immunity*, 66, 5939-47.
- KALLSTROM, H., BLACKMER GILL, D., ALBINGER, B., LISZEWSKI, M. K., ATKINSON, J. P. & JONSSON, A. B. 2001. Attachment of *Neisseria gonorrhoeae* to the cellular pilus receptor CD46: identification of domains important for bacterial adherence. *Cellular Microbiology*, 3, 133-43.
- KALLSTROM, H., ISLAM, M. S., BERGGREN, P. O. & JONSSON, A. B. 1998. Cell signaling by the type IV pili of pathogenic *Neisseria*. *Journal of Biological Chemistry*, 273, 21777-82.
- KALLSTROM, H., LISZEWSKI, M. K., ATKINSON, J. P. & JONSSON, A. B. 1997. Membrane cofactor protein (MCP or CD46) is a cellular pilus receptor for pathogenic *Neisseria*. *Molecular Microbiology*, 25, 639-647.
- KANTARDJIEFF, K. A. & RUPP, B. 2003. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science*, 12, 1865-1871.
- KANTARI, C., PEDERZOLI-RIBEIL, M. & WITKO-SARSAT, V. 2008. The role of neutrophils and monocytes in innate immunity. *Contributions to Microbiology*, 15, 118-46.
- KARUPPIAH, V., COLLINS, R. F., THISTLETHWAITE, A., GAO, Y. & DERRICK, J. P. 2013. Structure and assembly of an inner membrane platform for initiation of type IV pilus biogenesis. *Proceedings of the National Academy of Sciences U S A*, 110, E4638-E4647.
- KATZ, L. S., HUMPHREY, J. C., CONLEY, A. B., NELAKUDITI, V., KISLYUK, A. O., AGRAWAL, S., JAYARAMAN, P., HARCOURT, B. H., OLSEN-RASMUSSEN, M. A., FRACE, M., SHARMA, N. V., MAYER, L. W. & JORDAN, I. K. 2011. *Neisseria* Base: a comparative genomics database for *Neisseria meningitidis*. *Database (Oxford)*, 2011, bar035.
- KELLEY, L. A. & STERNBERG, M. J. E. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, 4, 363-371.
- KOLLS, J. K., MCCRAY, P. B., JR. & CHAN, Y. R. 2008. Cytokine-mediated regulation of antimicrobial proteins. *Nature Reviews in Immunology*, 8, 829-35.
- KORANSKY, J. R., SCALES, R. W. & KRAUS, S. J. 1975a. Bacterial hemagglutination by *Neisseria gonorrhoeae*. *Infection and Immunity*, 12, 495-498.
- KOROTKOV, K. V. & HOL, W. G. 2008. Structure of the GspK-GspI-GspJ complex from the enterotoxigenic *Escherichia coli* type 2 secretion system. *Nature Structural and Molecular Biology*, 15, 462-8.
- LEMAIRE, D., BARBOSA, T. & RIHET, P. 2012. Coping with genetic diversity: the contribution of pathogen and human genomics to modern vaccinology. *Brazilian Journal of Medical and Biological Research*, 45, 376-85.
- LI, J., EGELMAN, E. H. & CRAIG, L. 2012. Structure of the *Vibrio cholerae* Type IVb Pilus and stability comparison with the *Neisseria gonorrhoeae* type IVa pilus. *Journal of Molecular Biology*, 418, 47-64.
- LO, H., TANG, C. M. & EXLEY, R. M. 2009. Mechanisms of avoidance of host immunity by *Neisseria meningitidis* and its effect on vaccine development. *Lancet Infectious Disease*, 9, 418-27.
- LUTKENHAUS, J., PICHOFF, S. & DU, S. 2012. Bacterial cytokinesis: From Z ring to divisome. *Cytoskeleton (Hoboken)*, 69, 778-90.
- MADICO, G., WELSCH, J. A., LEWIS, L. A., MCNAUGHTON, A., PERLMAN, D. H., COSTELLO, C. E., NGAMPASUTADOL, J., VOGEL, U., GRANOFF, D. M. & RAM, S. 2006. The meningococcal

- vaccine candidate GNA1870 binds the complement regulatory protein factor H and enhances serum resistance. *Journal of Immunology*, 177, 501-510.
- MAIDEN, M. C., BYGRAVES, J. A., FEIL, E., MORELLI, G., RUSSELL, J. E., URWIN, R., ZHANG, Q., ZHOU, J., ZURTH, K., CAUGANT, D. A., FEAVERS, I. M., ACHTMAN, M. & SPRATT, B. G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences U S A*, 95, 3140-5.
- MAIER, B., POTTER, L., SO, M., LONG, C. D., SEIFERT, H. S. & SHEETZ, M. P. 2002. Single pilus motor forces exceed 100 pN. *Proceedings of the National Academy of Sciences U S A*, 99, 16012-7.
- MAITI, R., VAN DOMSELAAR, G. H., ZHANG, H. & WISHART, D. S. 2004. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Research*, 32, W590-W594.
- MAKOU, E., MERTENS, H. D. T., MACIEJEWSKI, M., SOARES, D. C., MATIS, I., SCHMIDT, C. Q., HERBERT, A. P., SVERGUN, D. I. & BARLOW, P. N. 2012. Solution structure of CCP modules 10-12 illuminates functional architecture of the complement regulator, Factor H. *Journal of Molecular Biology*, 424, 295-312.
- MALORNY, B., MORELLI, G., KUSECEK, B., KOLBERG, J. & ACHTMAN, M. 1998. Sequence diversity, predicted two-dimensional protein structure, and epitope mapping of Neisserial Opa proteins. *Journal of Bacteriology*, 180, 1323-30.
- MARTEYN, B., SCORZA, F. B., SANSONETTI, P. J. & TANG, C. 2011. Breathing life into pathogens: the influence of oxygen on bacterial virulence and host responses in the gastrointestinal tract. *Cellular Microbiology*, 13, 171-6.
- MARTEYN, B., WEST, N. P., BROWNING, D. F., COLE, J. A., SHAW, J. G., PALM, F., MOUNIER, J., PREVOST, M. C., SANSONETTI, P. & TANG, C. M. 2010. Modulation of Shigella virulence in response to available oxygen in vivo. *Nature*, 465, 355-8.
- MARTEYN, B. S., KARIMOVA, G., FENTON, A. K., GAZI, A. D., WEST, N., TOUQUI, L., PREVOST, M. C., BETTON, J. M., POYRAZ, O., LADANT, D., GERDES, K., SANSONETTI, P. J. & TANG, C. M. 2014. ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *Mbio*, 5, e00022-14. [10.1128/mBio.00022-14](https://doi.org/10.1128/mBio.00022-14)
- MATHEW, S. & OVERTURF, G. D. 2006. Complement and properdin deficiencies in meningococcal disease. *Pediatric Infectious Disease Journal*, 25, 255-6.
- MATTHEWS, B. W. 1968. Solvent content of protein crystals. *Journal of Molecular Biology*, 33, 491-8.
- MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, 40, 658-674.
- MCFERRIN, M. B. & SNELL, E. H. 2002. The development and application of a method to quantify the quality of cryoprotectant solutions using standard area-detector X-ray images. *Journal of Applied Crystallography*, 35, 538-545.
- MCNEARNEY, T., BALLARD, L., SEYA, T. & ATKINSON, J. P. 1989. Membrane cofactor protein of complement is present on human fibroblast, epithelial, and endothelial-cells. *Journal of Clinical Investigation*, 84, 538-545.
- MELICAN, K., MICHEA VELOSO, P., MARTIN, T., BRUNEVAL, P. & DUMENIL, G. 2013. Adhesion of *Neisseria meningitidis* to dermal vessels leads to local vascular damage and purpura in a humanized mouse model. *Plos Pathogens*, 9, e1003139. [10.1371/journal.ppat.1003139](https://doi.org/10.1371/journal.ppat.1003139)
- MILLER, F., PHAN, G., BRISSAC, T., BOUCHIAT, C., LIOUX, G., NASSIF, X. & COUREUIL, M. 2014. The hypervariable region of meningococcal major pilin PilE controls the host cell response via antigenic variation. *Mbio*, 5. [10.1128/mBio.01024-13](https://doi.org/10.1128/mBio.01024-13)
- MORAND, P. C., TATTEVIN, P., EUGENE, E., BERETTI, J. L. & NASSIF, X. 2001. The adhesive property of the type IV pilus-associated component PilC1 of pathogenic *Neisseria* is supported by the conformational structure of the N-terminal part of the molecule. *Molecular Microbiology*, 40, 846-56.

- MORGAN, H. P., MERTENS, H. D. T., GUARIENTO, M., SCHMIDT, C. Q., SOARES, D. C., SVERGUN, D. I., HERBERT, A. P., BARLOW, P. N. & HANNAN, J. P. 2012. Structural analysis of the C-terminal region (modules 18-20) of complement regulator factor H (fH). *Plos One*, 7. [10.1371/journal.pone.0032187](https://doi.org/10.1371/journal.pone.0032187)
- MOTHIBELI, K. M., DU PLESSIS, M., VON GOTTEBERG, A., MURPHY, E., HOISETH, S. K., ZLOTNICK, G. & KLUGMAN, K. P. Distribution of factor H binding protein beyond serogroup B: Variation among five serogroups of invasive *Neisseria meningitidis* in South Africa. *Vaccine*.
- NASSIF, X., BERETTI, J. L., LOWY, J., STENBERG, P., O'GAORA, P., PFEIFER, J., NORMARK, S. & SO, M. 1994. Roles of pilin and PilC in adhesion of *Neisseria meningitidis* to human epithelial and endothelial cells. *Proceedings of the National Academy of Sciences U S A*, 91, 3769-73.
- NASSIF, X., MARCEAU, M., PUJOL, C., PRON, B., BERETTI, J. L. & TAHA, M. K. 1997. Type-4 pili and meningococcal adhesiveness. *Gene*, 192, 149-53.
- NOINAJ, N., BUCHANAN, S. K. & CORNELISSEN, C. N. 2012. The transferrin-iron import system from pathogenic *Neisseria* species. *Molecular Microbiology*, 86, 246-257.
- OKEMEFUNA, A. I., GILBERT, H. E., GRIGGS, K. M., ORMSBY, R. J., GORDON, D. L. & PERKINS, S. J. 2008. The regulatory SCR-1/5 and cell surface-binding SCR-16/20 fragments of factor H reveal partially folded-back solution structures and different self-associative properties. *Journal of Molecular Biology*, 375, 80-101.
- ORANS, J., JOHNSON, M. D., COGGAN, K. A., SPERLAZZA, J. R., HEINIGER, R. W., WOLFGANG, M. C. & REDINBO, M. R. 2010. Crystal structure analysis reveals *Pseudomonas* PilY1 as an essential calcium-dependent regulator of bacterial surface motility. *Proceedings of the National Academy of Sciences U S A*, 107, 1065-70.
- PANGBURN, M. K., PANGBURN, K. L. W., KOISTINEN, V., MERI, S. & SHARMA, A. K. 2000. Molecular mechanisms of target recognition in an innate immune system: Interactions among factor H, C3b, and target in the alternative pathway of human complement. *Journal of Immunology*, 164, 4742-4751.
- PARGE, H. E., FOREST, K. T., HICKEY, M. J., CHRISTENSEN, D. A., GETZOFF, E. D. & TAINER, J. A. 1995. Structure of the fiber-forming protein pilin at 2.6-angstrom resolution. *Nature*, 378, 32-38.
- PEREZ-IRATXETA, C. & ANDRADE-NAVARRO, M. A. 2008. K2D2: estimation of protein secondary structure from circular dichroism spectra. *BMC Structural Biology*, 8, 25.
- PERKINS-BALDING, D., RATLIFF-GRIFFIN, M. & STOJILJKOVIC, I. 2004. Iron transport systems in *Neisseria meningitidis*. *Microbiology and Molecular Biology Reviews*, 68, 154-71.
- PERRAKIS, A., HARKIOLAKI, M., WILSON, K. S. & LAMZIN, V. S. 2001. ARP/wARP and molecular replacement. *Acta Crystallographica Section D-Biological Crystallography*, 57, 1445-1450.
- PETROV, A., LOMBARDO, S. & AUDETTE, G. F. 2013. Fibril-mediated oligomerization of pilin-derived protein nanotubes. *Journal of Nanobiotechnology*, 11, 24.
- POMPOSIELLO, P. J. & DEMPLE, B. 2002. Global adjustment of microbial physiology during free radical stress. *Advances in Microbial Physiology*, 46, 319-41.
- PORSCH, E. A., JOHNSON, M. D., BROADNAX, A. D., GARRETT, C. K., REDINBO, M. R. & ST GEME, J. W., 3RD 2013. Calcium binding properties of the *Kingella kingae* PilC1 and PilC2 proteins have differential effects on type IV pilus-mediated adherence and twitching motility. *Journal of Bacteriology*, 195, 886-95.
- PRINCE, S. M., ACHTMAN, M. & DERRICK, J. P. 2002. Crystal structure of the OpcA integral membrane adhesin from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences U S A*, 99, 3417-21.
- PROSSER, B. E., JOHNSON, S., ROVERSI, P., HERBERT, A. P., BLAUM, B. S., TYRRELL, J., JOWITT, T. A., CLARK, S. J., TARELLI, E., UHRIN, D., BARLOW, P. N., SIM, R. B., DAY, A. J. & LEA, S. M. 2007a. Structural basis for complement factor H-linked age-related macular degeneration. *Journal of Experimental Medicine*, 204, 2277-2283.

- PROSSER, B. E., JOHNSON, S., ROVERSI, P., HERBERT, A. P., BLAUM, B. S., TYRRELL, J., JOWITT, T. A., CLARK, S. J., TARELLI, E., UHRIN, D., BARLOW, P. N., SIM, R. B., DAY, A. J. & LEA, S. M. 2007b. Structural basis for complement factor H linked age-related macular degeneration. *Journal of Experimental Medicine*, 204, 2277-83.
- READ, R. J., MCCOY, A. J. & STORONI, L. C. 2007. Likelihood-based molecular replacement in Phaser. *Evolving Methods for Macromolecular Crystallography*, 245, 91-100.
- REDDICK, L. E. & ALTO, N. M. 2014. Bacteria fighting back: how pathogens target and subvert the host innate immune system. *Molecular Cell*, 54, 321-328.
- REYES, P. A., VARGAS, M. F., GARCIA, K. P., RUBILAR, P. S., NAVARRETE, P. A., FUENTES, P. M., RIOS, M. A., ORIHUELA, P. A., VARGAS, R. H., RUBIO, V. H., HECKELS, J., CHRISTODOULIDES, M., CARDENAS, H. & VELASQUEZ, L. A. 2007. Apoptosis related genes expressed in cultured fallopian tube epithelial cells infected in vitro with *Neisseria gonorrhoeae*. *Biological Research*, 40, 319-27.
- RIVAS, G., LOPEZ, A., MINGORANCE, J., FERRANDIZ, M. J., ZORRILLA, S., MINTON, A. P., VICENTE, M. & ANDREU, J. M. 2000. Magnesium-induced linear self-association of the FtsZ bacterial cell division protein monomer. The primary steps for FtsZ assembly. *Journal of Biological Chemistry*, 275, 11740-9.
- ROCK, J. D., MAHNANE, M. R., ANJUM, M. F., SHAW, J. G., READ, R. C. & MOIR, J. W. 2005. The pathogen *Neisseria meningitidis* requires oxygen, but supplements growth by denitrification. Nitrite, nitric oxide and oxygen control respiratory flux at genetic and metabolic levels. *Molecular Microbiology*, 58, 800-9.
- RUDEL, T., BOXBERGER, H. J. & MEYER, T. F. 1995a. Pilus biogenesis and epithelial cell adherence of *Neisseria gonorrhoeae* pilC double knock-out mutants. *Molecular Microbiology*, 17, 1057-71.
- RUDEL, T., FACIUS, D., BARTEN, R., SCHEURPFLUG, I., NONNENMACHER, E. & MEYER, T. F. 1995b. Role of pili and the phase-variable PilC protein in natural competence for transformation of *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences U S A*, 92, 7986-90.
- RUDEL, T., SCHEURPFLUG, I. & MEYER, T. F. 1995c. *Neisseria* PilC protein identified as type-4 pilus tip-located adhesin. *Nature*, 373, 357-9.
- RUDEL, T., VAN PUTTEN, J. P., GIBBS, C. P., HAAS, R. & MEYER, T. F. 1992. Interaction of two variable proteins (PilE and PilC) required for pilus-mediated adherence of *Neisseria gonorrhoeae* to human epithelial cells. *Molecular Microbiology*, 6, 3439-50.
- RUPP, B. & KANTARDJIEFF, K. 2010. *Biomolecular crystallography : principles, practice, and application to structural biology*, New York, Garland Science.
- RYLL, R. R., RUDEL, T., SCHEURPFLUG, I., BARTEN, R. & MEYER, T. F. 1997. PilC of *Neisseria meningitidis* is involved in class II pilus formation and restores pilus assembly, natural transformation competence and adherence to epithelial cells in PilC-deficient gonococci. *Molecular Microbiology*, 23, 879-92.
- RYTKONEN, A., JOHANSSON, L., ASP, V., ALBIGER, B. & JONSSON, A. B. 2001. Soluble pilin of *Neisseria gonorrhoeae* interacts with human target cells and tissue. *Infection and Immunity*, 69, 6419-6426.
- SCHEURPFLUG, I., RUDEL, T., RYLL, R., PANDIT, J. & MEYER, T. F. 1999. Roles of PilC and PilE proteins in pilus-mediated adherence of *Neisseria gonorrhoeae* and *Neisseria meningitidis* to human erythrocytes and endothelial and epithelial cells. *Infection and Immunity*, 67, 834-43.
- SCHMIDT, C. Q., HERBERT, A. P., MERTENS, H. D., GUARIENTO, M., SOARES, D. C., UHRIN, D., ROWE, A. J., SVERGUN, D. I. & BARLOW, P. N. 2010. The central portion of factor H (modules 10-15) is compact and contains a structurally deviant CCP module. *Journal of Molecular Biology*, 395, 105-22.

- SCHNEIDER, M. C., EXLEY, R. M., CHAN, H., FEAVERS, I., KANG, Y. H., SIM, R. B. & TANG, C. M. 2006. Functional significance of factor H binding to *Neisseria meningitidis*. *Journal of Immunology*, 176, 7566-7575.
- SCHNEIDER, M. C., PROSSER, B. E., CAESAR, J. J. E., KUGELBERG, E., LI, S., ZHANG, Q., QUORAISHI, S., LOVETT, J. E., DEANE, J. E., SIM, R. B., ROVERSI, P., JOHNSON, S., TANG, C. M. & LEA, S. M. 2009. *Neisseria meningitidis* recruits factor H using protein mimicry of host carbohydrates. *Nature*, 458, 890-U9.
- SCHOEN, C., BLOM, J., CLAUS, H., SCHRAMM-GLUCK, A., BRANDT, P., MULLER, T., GOESMANN, A., JOSEPH, B., KONIETZNY, S., KURZAI, O., SCHMITT, C., FRIEDRICH, T., LINKE, B., VOGEL, U. & FROSCH, M. 2008. Whole-genome comparison of disease and carriage strains provides insights into virulence evolution in *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences U S A*, 105, 3473-8.
- SCHOOLNIK, G. K., FERNANDEZ, R., TAI, J. Y., ROTHBARD, J. & GOTSCHLICH, E. C. 1984. Gonococcal pili. Primary structure and receptor binding domain. *Journal of Experimental Medicine*, 159, 1351-70.
- SCHRYVERS, A. B. & MORRIS, L. J. 1988. Identification and characterization of the human lactoferrin-binding protein from *Neisseria meningitidis*. *Infection and Immunity*, 56, 1144-9.
- SEIB, K. L., SERRUTO, D., ORIENTE, F., DELANY, I., ADU-BOBIE, J., VEGGI, D., ARICO, B., RAPPUOLI, R. & PIZZA, M. 2009. Factor H-binding protein is important for meningococcal survival in human whole blood and serum and in the presence of the antimicrobial peptide LL-37. *Infection and Immunity*, 77, 292-299.
- SEIB, K. L., TSENG, H. J., MCEWAN, A. G., APICELLA, M. A. & JENNINGS, M. P. 2004. Defenses against oxidative stress in *Neisseria gonorrhoeae* and *Neisseria meningitidis*: distinctive systems for different lifestyles. *Journal of Infectious Disease*, 190, 136-47.
- SERRUTO, D., ADU-BOBIE, J., SCARSELLI, M., VEGGI, D., PIZZA, M., RAPPUOLI, R. & ARICO, B. 2003. *Neisseria meningitidis* App, a new adhesin with autocatalytic serine protease activity. *Molecular Microbiology*, 48, 323-34.
- SIVONEN, A., RENKONEN, O. V., WECKSTROM, P., KOSKENVUO, K., RAUNIO, V. & MAKELA, P. H. 1978. The effect of chemoprophylactic use of rifampin and minocycline on rates of carriage of *Neisseria meningitidis* in army recruits in Finland. *Journal of Infectious Disease*, 137, 238-44.
- SJOLINDER, H., ERIKSSON, J., MAUDSDOTTER, L., ARO, H. & JONSSON, A. B. 2008. Meningococcal outer membrane protein NhhA is essential for colonization and disease by preventing phagocytosis and complement attack. *Infection and Immunity*, 76, 5412-20.
- SPINOSA, M. R., PROGIDA, C., TALA, A., COGLI, L., ALIFANO, P. & BUCCI, C. 2007. The *Neisseria meningitidis* capsule is important for intracellular survival in human cells. *Infection and Immunity*, 75, 3594-603.
- STABLER, R. A., MARSDEN, G. L., WITNEY, A. A., LI, Y., BENTLEY, S. D., TANG, C. M. & HINDS, J. 2005. Identification of pathogen-specific genes through microarray analysis of pathogenic and commensal *Neisseria* species. *Microbiology*, 151, 2907-22.
- STEIN, N. 2008. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography*, 41, 641-643.
- STEPHENS, D. S., GREENWOOD, B. & BRANDTZAEG, P. 2007. Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*. *Lancet*, 369, 2196-210.
- STERN, A. & MEYER, T. F. 1987. Common mechanism controlling phase and antigenic variation in pathogenic neisseriae. *Molecular Microbiology*, 1, 5-12.
- STIMSON, E., VIRJI, M., MAKEPEACE, K., DELL, A., MORRIS, H. R., PAYNE, G., SAUNDERS, J. R., JENNINGS, M. P., BARKER, S., PANICO, M. & ET AL. 1995. Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Molecular Microbiology*, 17, 1201-14.

- STIVALA, A., WYBROW, M., WIRTH, A., WHISSTOCK, J. C. & STUCKEY, P. J. 2011. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics*, 27, 3315-3316.
- STROM, M. S., NUNN, D. N. & LORY, S. 1993. A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proceedings of the National Academy of Sciences U S A*, 90, 2404-8.
- SZABO, Z., SANI, M., GROENEVELD, M., ZOLGHADR, B., SCHELERT, J., ALBERS, S. V., BLUM, P., BOEKEMA, E. J. & DRIESSEN, A. J. 2007. Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *Journal of Bacteriology*, 189, 4305-9.
- TOBIASON, D. M. & SEIFERT, H. S. 2001. Inverse relationship between pilus-mediated gonococcal adherence and surface expression of the pilus receptor, CD46. *Microbiology*, 147, 2333-2340.
- TURNER, D. P., MARIETOU, A. G., JOHNSTON, L., HO, K. K., ROGERS, A. J., WOOLDRIDGE, K. G. & ALA'ALDEEN, D. A. 2006. Characterization of MspA, an immunogenic autotransporter protein that mediates adhesion to epithelial and endothelial cells in *Neisseria meningitidis*. *Infection and Immunity*, 74, 2957-64.
- TZENG, Y. L., AMBROSE, K. D., ZUGHAIER, S., ZHOU, X., MILLER, Y. K., SHAFER, W. M. & STEPHENS, D. S. 2005. Cationic antimicrobial peptide resistance in *Neisseria meningitidis*. *Journal of Bacteriology*, 187, 5387-96.
- UNKMEIR, A., KAMMERER, U., STADE, A., HUBNER, C., HALLER, S., KOLB-MAURER, A., FROSCH, M. & DIETRICH, G. 2002a. Lipooligosaccharide and polysaccharide capsule: virulence factors of *Neisseria meningitidis* that determine meningococcal interaction with human dendritic cells. *Infection and Immunity*, 70, 2454-62.
- UNKMEIR, A., LATSCH, K., DIETRICH, G., WINTERMEYER, E., SCHINKE, B., SCHWENDER, S., KIM, K. S., EIGENTHALER, M. & FROSCH, M. 2002b. Fibronectin mediates Opc-dependent internalization of *Neisseria meningitidis* in human brain microvascular endothelial cells. *Molecular Microbiology*, 46, 933-46.
- UNNI, S., HUANG, Y., HANSON, R. M., TOBIAS, M., KRISHNAN, S., LI, W. W., NIELSEN, J. E. & BAKER, N. A. 2011. Web servers and services for electrostatics calculations with APBS and PDB2PQR. *Journal of Computational Chemistry*, 32, 1488-1491.
- VAN HOLDE, K. E., JOHNSON, W. C. & HO, P. S. 2006. *Principles of physical biochemistry*, Upper Saddle River, N.J., Prentice Hall.
- VIRJI, M. & HECKELS, J. E. 1983. Antigenic cross-reactivity of *Neisseria pili* - investigations with type-specific and species-specific monoclonal-antibodies. *Journal of General Microbiology*, 129, 2761-2768.
- VIRJI, M., MAKEPEACE, K., FERGUSON, D. J. & WATT, S. M. 1996a. Carcinoembryonic antigens (CD66) on epithelial cells and neutrophils are receptors for Opa proteins of pathogenic neisseriae. *Molecular Microbiology*, 22, 941-50.
- VIRJI, M., STIMSON, E., MAKEPEACE, K., DELL, A., MORRIS, H. R., PAYNE, G., SAUNDERS, J. R. & MOXON, E. R. 1996b. Posttranslational modifications of meningococcal pili. Identification of a common trisaccharide substitution on variant pilins of strain C311. *Annals of the New York Academy of Sciences*, 797, 53-64.
- WALPORT, M. J. 2001a. Advances in immunology: Complement (First of two parts). *New England Journal of Medicine*, 344, 1058-1066.
- WALPORT, M. J. 2001b. Advances in immunology: Complement (Second of two parts). *New England Journal of Medicine*, 344, 1140-1144.
- WELSCH, J. A., RAM, S., KOEBERLING, O. & GRANOFF, D. M. 2008. Complement-dependent synergistic bactericidal activity of antibodies against factor H-binding protein, a sparsely distributed meningococcal vaccine antigen. *Journal of Infectious Disease*, 197, 1053-61.
- WEYAND, N. J., WERTHEIMER, A. M., HOBBS, T. R., SISCO, J. L., TAKU, N. A., GREGSTON, L. D., CLARY, S., HIGASHI, D. L., BIAIS, N., BROWN, L. M., PLANER, S. L., LEGASSE, A. W.,

- AXTHELM, M. K., WONG, S. W. & SO, M. 2013. *Neisseria* infection of rhesus macaques as a model to study colonization, transmission, persistence, and horizontal gene transfer. *Proceedings of the National Academy of Sciences U S A*, 110, 3059-3064.
- WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISINEL, E. B., LESLIE, A. G. W., MCCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography*, 67, 235-242.
- WINN, M. D., MURSHUDOV, G. N. & PAPIZ, M. Z. 2003. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Macromolecular Crystallography, Pt D*, 374, 300-321.
- WINTER, G. 2010. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, 43, 186-190.
- WINTHER-LARSEN, H. C., HEGGE, F. T., WOLFGANG, M., HAYES, S. F., VAN PUTTEN, J. P. & KOOMEY, M. 2001. *Neisseria gonorrhoeae* PilV, a type IV pilus-associated protein essential to human epithelial cell adherence. *Proceedings of the National Academy of Sciences U S A*, 98, 15276-81.
- WORMANN, M. E., HORIEN, C. L., BENNETT, J. S., JOLLEY, K. A., MAIDEN, M. C., TANG, C. M., AHO, E. L. & EXLEY, R. M. 2014. Sequence, distribution and chromosomal context of class I and class II pilin genes of *Neisseria meningitidis* identified in whole genome sequences. *BMC Genomics*, 15, 253.
- YAZDANKHAH, S. P. & CAUGANT, D. A. 2004. *Neisseria meningitidis*: an overview of the carriage state. *Journal of Medical Microbiology*, 53, 821-32.
- YUSUF, D., DAVIS, A. M., KLEYWEGT, G. J. & SCHMITT, S. 2008. An alternative method for the evaluation of docking performance: RSR vs RMSD. *Journal of Chemical Information and Modeling*, 48, 1411-22.
- ZAHLANIE, Y. C., HAMMADI, M. M., GHANEM, S. T. & DBAIBO, G. S. 2014. Review of meningococcal vaccines with updates on immunization in adults. *Human Vaccines & Immunotherapeutics*, 10, 995-1007.
- ZETH, K., KOZJAK-PAVLOVIC, V., FAULSTICH, M., FRAUNHOLZ, M., HURWITZ, R., KEPP, O. & RUDEL, T. 2013. Structure and function of the PorB porin from disseminating *Neisseria gonorrhoeae*. *Biochemical Journal*, 449, 631-42.