Structural and functional studies of bacterial Tripartite α-Pore forming toxins

Alicia Churchill-Angus

Supervisor: Dr Patrick Baker

Department of Molecular Biology and Biotechnology

University of Sheffield

This thesis is submitted for the degree of

Doctor of Philosophy

September 2020

Declaration

I, Alicia M Churchill-Angus, declare that this thesis titled, "Structural and function studies of bacterial Tripartite α-Pore forming toxins" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Alicia Churchill-Angus

September 2020

Acknowledgments

During my PhD I have been helped and supported by many people. I would like to thank my supervisor Dr Patrick Baker, for giving me the opportunity to come and work in the crystallography group studying pore-forming toxins, he has truly inspired my love of Structural Biology. His patience (particularly with daily video calls since Covid began!) guidance, many interesting conversations, and climbing skills have been vital to the success of this project. Thank you also to my Advisors Prof. David Rice and Dr Rosie Staniforth for guidance. I'd also like to thank Prof. David Rice, Dr John Rafferty, Dr Svetlana Sedelnikova, Dr Claudine Bisson, Fiona Rodgers, Dr Sam Dix, Dr Jason Wilson, and Dr Adli Adziz from the crystallography group for giving your help and knowledge throughout my work, and for the many entertaining coffee conversations. I'd also like to thank those who have helped in small ways that have made all the difference to my work. Next, I would like to thank the friends I have made during my PhD, within the crystallography group, and within the department, most of all Dr Sam Dix, Dr Adli Aziz, Henry Wood and Dr Angus Robertson with whom I've shared many sleepless Diamond trips, and for many great times together outside of the lab. Furthermore, I would like to thank the project students I have had over the years, in particular Harry Schofield, and Thomas Marlow. To my brilliant housemates through my PhD, Maria Callaby, Bethany Ross, Lili Bidari and Gemma Walker thank you for always being there, for keeping me sane and the wonderfully entertaining conversations after long days in the Lab. Thank you also to all my other friends, and particularly Hannah Leflay for the many lunches and coffees through our 7 years at Sheffield. Finally, a big thank you to my family, especially my parents and sister who have supported me over the years and were unexpectedly trapped with me throughout the writing process, I could not have done it without them!

Abstract

Pore-forming toxins (PFTs) are vital in both Prokaryotes and Eukaryotes and make up 30% of all bacterial toxins identified to date. Members of the ClyA family of α-PFTs have been found with one (ClyA), two (e.g YaxAB, XaxAB, PaxAB) or three (Hbl, Nhe, and Ahl) protein components that make the active pore. Pore structures are known for the bipartite family members, but only soluble structures (NheA, Hbl-B and AhlC) of the tripartite toxins, (the most complex system within the family), have been solved. As such, there is a pressing need for more structural and biochemical data to fully understand tripartite pore activity and assembly.

The structure determination of *A. hydrophilia* AhlB in both soluble and pore forms as well as soluble AhlC head mutants have shown that AhlB is able to form pores (the first pore structure from a tripartite α-PFT), and also allowed roles to be proposed for both the B and C components of the Ahl toxin. TEM studies of Ahl pores, together with detailed biochemical studies of the activity of the AhlA, AhlB and AhlC components, has shown that AhlB, AhlBC and AhlABC pores can form, and identified three features common to all ClyA family α-PFTs.

Characterisation of a second ClyA family tripartite α-PFT, Smh from *Serratia marcescens*, has shown that structures of SmhA and SmhB are well conserved with AhlB and NheA allowing identification of a latch which must be broken for the soluble to pore conformational change to occur. Having determined structures for all three components of Gram negative tripartite α-PFTs, a structure-based model of a complete pore has been constructed. Finally, although the Smh proteins share high structural similarities with both the Ahl and Nhe systems, biochemical assays show assembly mechanisms have diverged within the family.

Table of contents

List of figures **xiii**

List of tables **xix**

Nomenclature **xxi**

1 Introduction **1**

1.1 Thesis introduction 1

1.2 Biological membranes 2

1.2.1 Properties of phospholipids and their assembly in an aqueous environment 2

1.2.2 Cell membranes as selectively permeable barriers 4

1.3 Bacterial pore-forming toxins 6

1.3.1 Role in Bacterial infections 9

1.3.2 General pore assembly mechanism 9

1.3.3 Excretion of the PFT from the bacteria cell 11

1.3.4 Membrane receptors and PFT binding 12

1.3.5 Pore assembly 16

1.3.6 Cell response to PFT attack 20

1.4 The ClyA family of PFTs 20

1.4.1 Cytolysin A (ClyA) 21

1.4.2 Bipartite ClyA family of α-PFTs 27
1.4.3 Tripartite ClyA family α-PFTs, NheABC and HblL1L2B 32
1.4.4 The Ahl toxin from *A. hydrophila* 38
1.5 Targeting PFTs for Antimicrobial therapy 40
1.6 Biotechnological applications of PFTs 40
1.7 Aims of this thesis 42

2 Materials and Methods 45

2.1 General methods and recipes 45
2.1.1 Antibiotics 45
2.1.2 Lysogeny Broth (LB) medium 45
2.1.3 LB Agar 45
2.1.4 SOC media 46
2.1.5 Strains and plasmid vector information 46
2.1.6 Gene cloning 46
2.1.7 Transformation of *E. coli* with plasmid vector 46
2.1.8 Glycerol stocks 48
2.1.9 SDS-PAGE 48
2.1.10 Preparation of samples for SDS-PAGE 49
2.1.11 Measuring protein concentration 49
2.1.12 Lipid stock preparation 50
2.1.13 Detergent screens 51
2.1.14 Liposome preparation 51
2.1.15 Proteoliposome preparation 52
2.2 Protein overexpression 52
2.2.1 Trial expression protocol 52
2.2.2 Large scale overexpression 52

2.2.3 Minimal media selenomethionine overexpression 53
2.3 Protein purification 54
2.3.1 Preparation of cell-free extract 54
2.3.2 Nickel affinity chromatography 55
2.3.3 Ammonium sulphate cut 55
2.3.4 Ion exchange chromatography 56
2.3.5 Gel filtration 56
2.4 Biochemical analysis 57
2.4.1 Ultra-centrifugation of detergent solubilised protein 57
2.4.2 Haemolytic assay 57
2.4.3 Protein cross-linkage 58
2.4.4 Liposome float assay 58
2.5 Crystallographic methods 59
2.5.1 Preparation of protein for crystallisation 59
2.5.2 Initial robot screening 59
2.5.3 Crystallisation optimisation 59
2.5.4 Mounting crystals 60
2.5.5 Crystal screening and data collection 61
2.5.6 Testing 61
2.5.7 Data collection 61
2.5.8 Data processing 62
2.5.9 Experimental phasing using Single wavelength anomalous dispersion (SAD) 63
2.5.10 Molecular replacement 63
2.5.11 Model building and refinement 64
2.5.12 Validation and analysis 64

2.5.13 Homology model generation 65
2.6 Electron Microscopy Methods 65
2.6.1 Glow discharging and loading carbon grids 65
2.6.2 Data collection on a CM110 transmission electron microscope. . . . 66
2.6.3 Data collection on Technai Arctica, FEI Cryo-electron microscope . 67
2.6.4 Image processing and analysis 67

3 Paper 1 **71**
3.1 Summary 71
3.2 Author contributions 72
3.3 Paper 1 72

Appendix A Extended methodology - Ahl toxin **109**
A.1 AhlA, AhlB, and AhlC construct design 109
A.2 Overexpression and purification of AhlA, AhlB and AhlC constructs 110
A.2.1 Purification of AhlA 110
A.3 Crystallisation, data collection and structure determination of AhlB 113
A.3.1 X-ray data collection of AhlB pore 113
A.3.2 Structure determination, model building and validation of AhlB pore 119
A.3.3 Interface analysis of AhlB pore 123
A.4 Model building and validation of soluble AhlB 126
A.5 Crystallisation, Data collection and structure determination of AhlC HM . . 130
A.5.1 Crystallisation trials of AhlC HM 130
A.5.2 X-ray Data collection of AhlC HM 131
A.5.3 Structure determination, model building and validation of AhlC HM 131
A.5.4 Interface analysis of head domains in the AhlC HM 134
A.6 Detergent screens and electron microscopy of AhlA, AhlB and AhlC 138

A.6.1 HPLC detergent screening against AhlB . . . 138
A.6.2 Detergent 18 trials with AhlB and AhlC . . . 139
A.6.3 Detergent 18 trials with AhlA, AhlB, and AhlC . . . 141
A.6.4 Optimisation of AhlB pores with detergent. . . . 143

4 Paper 2 **149**

4.1 Summary . . . 149
4.2 Author contributions . . . 149
4.3 Paper 2 . . . 150

5 Paper 3 **157**

5.1 Summary . . . 157
5.2 Author contributions . . . 158
5.3 Paper 3 . . . 158

Appendix B Extended methodology - Smh toxin **209**

B.1 SmhA, SmhB, and SmhC construct design . . . 209
B.2 Overexpression and purification of SmhA, SmhB and SmhC constructs . . . 210
B.2.1 Overexpression trials for SmhA, SmhB and SmhC . . . 210
B.2.2 Purification of SmhA . . . 210
B.2.3 Purification of SmhB . . . 212
B.2.4 Purification of SmhC . . . 212
B.3 Haemolytic sssays for the Smh toxin . . . 218
B.4 Electron microscopy of soluble complexes . . . 221
B.5 Crystallisation, data collection and structure determination of SmhB . . . 226
B.5.1 Crystallisation trials of SmhB . . . 226
B.5.2 X-ray data collection of SmhB . . . 226

B.5.3 Structure determination, model building and validation of SmhB crystal Form 1 and crystal Form 2 232
B.5.4 Structure determination, model building and validation of SmhB pore form. 233
B.6 Crystallisation, data collection and structure determination of SmhA 237
B.6.1 Crystallisation trials of SmhA 237
B.6.2 X-ray data collection of SmhA 237
B.6.3 Structure determination, model building and validation of SmhA . . 241
B.7 Bioinformatics analysis of SmhC 254
B.8 Homology searches and sequence alignments with SmhA and SmhB 257
B.9 Electron microscopy of Smh proteoliposomes 260

6 Conclusions and future work 265
6.1 Discussion 265
6.1.1 Structures of the A and B components show a conserved mechanism for conformational change 268
6.1.2 Modelling the tripartite pore 270
6.1.3 Mechanism of pore assembly 274
6.2 Future work 280
6.2.1 Structure of a tripartite pore 281
6.2.2 Structures of SmhC and NheC and Soluble complexes 281
6.2.3 Cell lysis assays 281
6.2.4 Chimeric Haemolytic assays 282
6.2.5 Protein engineering 282

References 285

List of figures

1.1 Characteristics of biological membrane lipids. 3

1.2 Characteristics and organisation of biological lipid bilayers. 5

1.3 The two classes of PFT. 8

1.4 Mechanisms of pore formation. 10

1.5 Export and membrane recognition of PFTs. 13

1.6 PFTs recognise and bind to the membrane via proteins, sugars and lipids. . 15

1.7 PFTs assemble to form a diverse range of pore types. 18

1.8 Conformational changes during soluble to pore transformation. 19

1.9 Soluble ClyA hides it hydrophobic residues in the α- tongue and N-terminal helix. 25

1.10 ClyA undergoes a large conformational change to expose its hydrophobic residues. 26

1.11 The YaxAB pore. 30

1.12 Structures of the bipartite toxins show different mechanisms for concealing the hydrophobic heads and pore assembly to ClyA. 31

1.13 Structure of Hbl-B and activity assays for the tripartite Hbl toxin. 35

1.14 Structure of NheA and proposed mechanism for pore assembly. 37

1.15 The Ahl toxin from Aeromonas hydrophila. 39

2.1 Details on the pET-21a vector taken from Novagen 47

A.1 Purification of AhlA. 112
A.2 Crystals of AhlB Semet. 114
A.3 AhlB SeMet SAD data collection. 116
A.4 Self-rotation plots of AhlB Semet in space group C2. 117
A.5 AhlB Semet (space group C2) Matthews calculation. 118
A.6 Visual analysis of density maps from SHELX (Sheldrick, 2008) for AhlB Semet. 121
A.7 Initial and final models and electron density for AhlB after rebuilding and refinement. 122
A.8 ePisa (Krissinel and Henrick, 2007) analysis of Interface 1 within the AhlB pore. 124
A.9 ePisa (Krissinel and Henrick, 2007) analysis of Interface 2 within the AhlB pore. 125
A.10 Model building and refinement of soluble AhlB. 128
A.11 Buccaneer and final validated models and electron density for each chain of soluble AhlB. 129
A.12 Crystals of AhlC HM (left) and mounted crystal in litholoop (right). 132
A.13 AhlC HM (PACT E12) data collection on beamline i04. 133
A.14 Model building and refinement of soluble AhlC HM. 135
A.15 Density around Thr 156 (top), Thr 160 (centre), Thr 161 (bottom). 136
A.16 Head domain Interface analysis for AhlC HM. 137
A.17 Chromatograms from best 3 detergents at 10 x CMC from initial screens with AhlB. 140
A.18 Chromatograms of AhlB and AhlC incubated with detergent 18 at 3 concentrations. 142
A.19 AhlABC forms pores when incubated with 1 x CMC of detergent 18. 144

A.20 Optimisation of AhlB pores for cryo-EM. 146
A.21 Optimisation of AhlB pores. 147

B.1 SDS-PAGE gels showing 5hr overexpression trials for SmhA, SmhB, and SmhC. 211
B.2 Purification of His tagged SmhA. UV trace chromatogram (left) and SDS-PAGE (right). 213
B.3 Purification of His tagged SmhB. UV trace chromatogram (left) and SDS-PAGE (right). 214
B.4 Purification of His tagged SmhC. UV trace chromatogram (left) and SDS-PAGE (right). 216
B.5 Purification of SmhC by AS cut and anion exchange chromatography. . . . 217
B.6 Lytic and pore-forming activity of Smh toxin. 220
B.7 TEM negative stain EM images of gel filtration fractions. 223
B.8 TEM negative stain EM images of SmhBC soluble complexes incubated with detergents. 224
B.9 TEM negative stain EM images of SmhABC soluble complexes incubated with detergents. 225
B.10 Crystals of SmhB (left) and mounted crystal in litholoop (right). 227
B.11 A representative 0.1 ^{o} oscillation image from SmhB crystal Form 1 229
B.12 A representative 0.1 ^{o} oscillation image from SmhB crystal Form 2 230
B.13 A representative 0.1 ^{o} oscillation image from SmhB pore Form 231
B.14 Molecular replacement, model building and refinement of SmhB crystal Form 1. 234
B.15 The final model of SmhB crystal Form 1 shows differences in $\alpha 7$ between chain A and B caused by crystal packing interactions. 235

B.16 SmhB crystal Form 2 shows further differences in α7 between chain A and B caused by crystal packing interactions. 236
B.17 Crystals of SmhA 239
B.18 SmhA SeMet SAD data collection. 242
B.19 SmhA crystal Form 1 Matthews calculation. 243
B.20 SmhA crystal Form 1 data collection. 244
B.21 SmhA crystal Form 2 Matthews calculation. 245
B.22 Experimental phasing and chain tracing by CRANK2.0 247
B.23 Final model of SmhA crystal Form 1. 248
B.24 Comparison of density around the head domain of chain D (yellow, left three) and chain A (green, right three). 249
B.25 Molecular replacement and model building for SmhA crystal Form 2. . . . 251
B.26 Difference in electron density for each chain of SmhA crystal Form 2. . . . 253
B.27 Sequence analysis of SmhC homologues. 255
B.28 Location of interfaces and conserved residues in the C tetramer. 256
B.29 Genomic analysis of regions upstream and downstream of SmhA and SmhB homologues identified by BlastP (Altschul et al., 1990). 259
B.30 TEM negative stain EM images of SmhBC and SmhABC showing pores when incubated with liposomes. 261
B.31 TEM negative stain EM images of SmhBC and SmhABC incubated with liposomes for 1hour at 37 ^{o}C. 263
B.32 3D classification of SmhBC filament. 264

6.1 All ClyA family members must have three features for a lytic pore. 267
6.2 Models of possible pore assemblies of the tripartite PFT from Gram negative bacteria. 273
6.3 A proposed assembly schematic of AhlB, AhlBC and AhlABC pores. . . . 276

6.4 Schematic showing pore assembly for the Smh toxin. 277

6.5 Schematic showing pore assembly for the Nhe and Hbl toxin. 278

List of tables

1.1 Members of the ClyA α-PFT family. 22

2.1 SDS-PAGE recipe. 49

A.1 Initial crystals hits for AhlB Semet grown in 98 well sitting drop plates . . . 115

A.2 Analysis of the two interfaces between subunits in the AhlB pore. 126

A.3 Initial crystals hits for AhlC HM grown in 98 well sitting drop plates 131

B.1 Data collection parameters for SmhB crystal forms. 228

B.2 Crystallisation conditions and data collection statistics for different crystal forms of SmhA. 238

B.3 Crystallisation conditions and data collection statistics for different crystal forms of SmhA. 240

Nomenclature

Acronyms / Abbreviations

AFM	Atomic force microscopy
Ahl	*Aeromonas* haemolytic toxin
APS	Ammonium persulphate
AU	Asymmetric unit
CA	Cardiolipin
CCD	charge-coupled device
CDC	cholesterol dependent cytolysin
CFE	Cell-free extract
ClyA	cytolisin A
Cryo-EM	Cryo electron microscopy
CV	Column volume
DDM	n-Dodecyl β-D-maltoside
DEAE	Diethylethanolamine

DLS	Diamond Light Source
DNA	Deoxyribonucleic acid
DOC	Sodium Deoxycholate
DTT	Dithiothreitol
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
FraC	fragaceatoxin C
fT3SS	flagella Type III secretion system
GPI	Glycosylphosphatidylinositol
GUI	Graphical user interface
Hbl	Haemolysin BL
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
Hla	α-haemolysin
HPLC	High-performance liquid chromatography
IEC	Ion Exchange Chromatography
IMAC	Immobilised Metal-ion Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny Broth
LDS	Lithium dodecyl-sulphate

LLG	Log-likelihood gain
MakA	motility associated killing factor A
MCS	Multiple Cloning Site
MES	2-(N-morpholino)ethanesulfonic acid
MPD	2-Methyl-2,4-pentanediol
MR	Molecular Replacement
MWCO	Molecular weight cut-off
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
Nhe	Non-haemolytic enterotoxin
NTA	nitrilotriacetic acid
OD	Optical Density
OMV	Outer membrane vesicles
PAGE	Poly-acrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Poly-ethylene glycol
PE	phosphatidylethanolamine

PFT	Pore forming toxin
PG	phosphatidylglycerol
PI	Propidium iodide
PISA	Proteins, Interfaces, Structures and Assemblies
R&D	Research and development
rmsD	root-mean-square deviation
SAD	Single Anomalous Dispersion
SDS	Sodium dodecyl sulphate
SEC	Size Exclusion Chromatography
SM	sphingomyelin
SOC	Super optimal broth
SRF	Self rotation function
T3SS	Type III secretion system
TAE	Tris-Acetate-EDTA
TEMED	Tetramethylethylenediamine
TEM	Transmission electron microscopy
TFZ	Translation function
TOF	time-of-flight
TRIS	2-Amino-2-(hydroxymethyl)propane-1,3-diol

VCC	*Vibrio cholerae* cytolysin
WHO	World Health Organisation
Xax	Xenorhabdus α-xenorhabdolysin
Yax	Yersinia α-xenorhabdolysin

Chapter 1

Introduction

1.1 Thesis introduction

Pore-forming toxins (PFTs) are used by both eukaryotic and prokaryotic organisms to disrupt biological membranes, acting as both offensive and defensive mechanisms. Their ability to undergo large scale conformational changes from soluble to membrane proteins as well as the variation in pore assembly adopted by different organisms makes them interesting candidates in the study of how proteins can interact with, and disrupt, membranes. The fact that PFTs are virulence factors in many antibiotic-resistant pathogenic bacteria also makes them important antimicrobial targets. In addition, an increasing understanding of these bacterial toxins has led to a range of biotechnological applications. This chapter explores the current understanding of Bacterial PFTs, beginning with a broad introduction to the family, before focusing specifically on the ClyA family of α-PFTs, which are the focus of the research undertaken.

1.2 Biological membranes

1.2.1 Properties of phospholipids and their assembly in an aqueous environment

Biological membranes can contain three types of lipids, these include phospholipids, glycolipids and sterols (Watson, 2015). All of these groups are amphipathic and consist of a hydrophilic head and hydrophobic tail. Both phospholipids and glycolipids use saturated and unsaturated fatty acids as a hydrophobic tail, with the chemical composition of the fatty acids playing an important role in membrane fluidity (Watson, 2015) (Figure 1.1A). Sterols are found mostly in eukaryotes and employ steroid rings to provide their hydrophobic domain (Watson, 2015)(Figure 1.1A). The hydrophilic head group varies between lipids, in the case of glycolipids, the sugar group is highly diverse and can act as a marker, such as an antigen, or provide a recognition site for membrane binding proteins, for example, the PFT aerolysin uses glycosylphosphatidylinositols as receptors for assisting pore formation (Watson, 2015; Wu and Guo, 2010). Due to their amphipathic nature, biological lipids arrange themselves in one of three structures, micelles, lipid bilayer sheets or lipid bilayer spheres (liposomes), when exposed to an aqueous environment (Watson, 2015) (Figure 1.1B). Due to the two fatty acid chains of a phospholipid or glycolipid being too large to pack easily into a micelle, they more readily form lipid bilayers making them perfect for biological membranes (Berg et al., 2012).

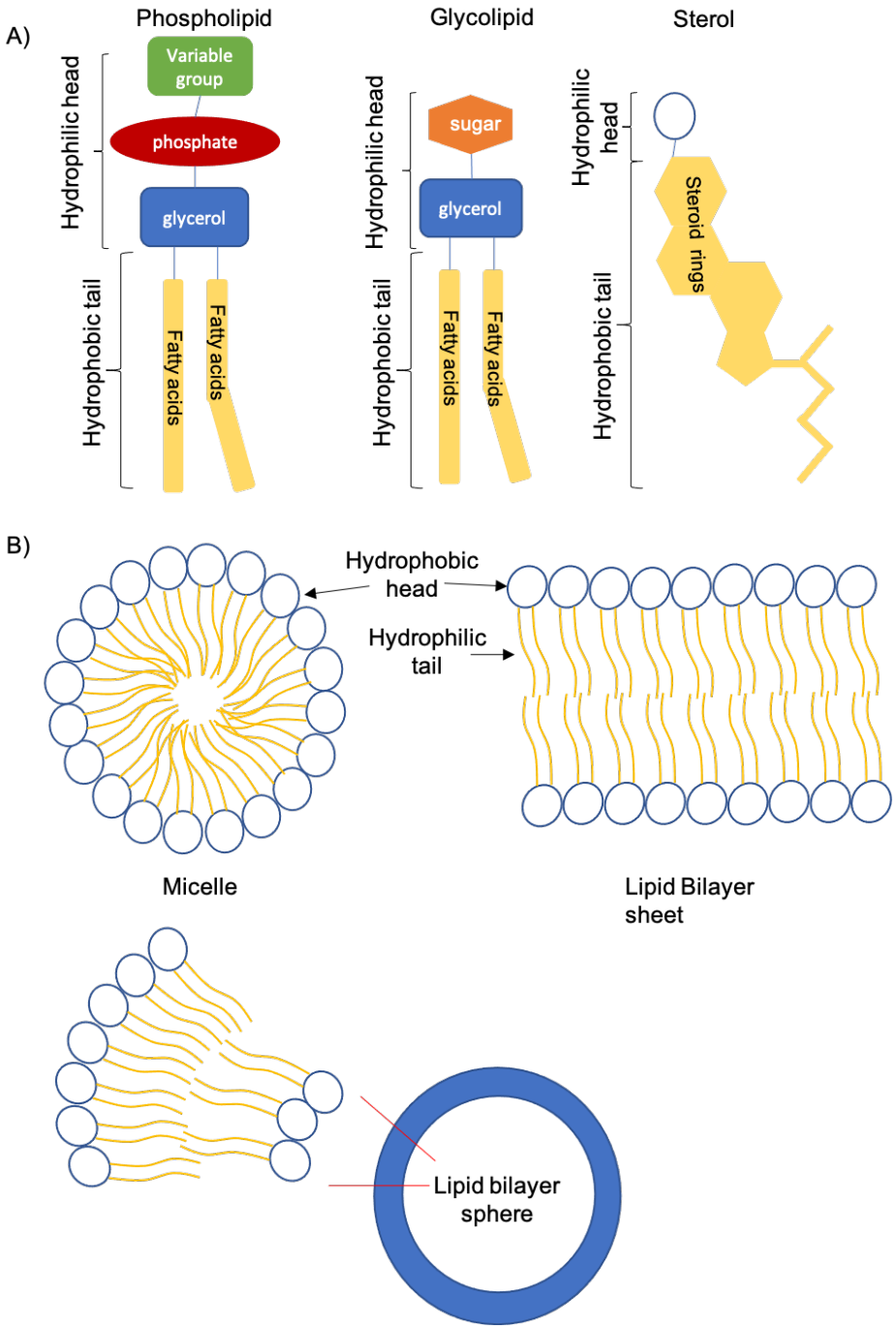


Figure 1.1 **Characteristics of biological membrane lipids.** A) Schematics of the structures of the three main types of membrane lipids. B) Membrane lipids can arrange themselves into three structures in an aqueous environment, micelles (top left), bilayer sheets (top right), and bilayer spheres (liposomes, bottom). Hydrophilic heads are shown as blue circles, hydrophobic tails as yellow lines.

1.2.2 Cell membranes as selectively permeable barriers

A cell membrane acts as a barrier between the interior and exterior of a cell. A lipid bilayer has low permeability to ions and most polar molecules, with the exception of water (Figure 1.2A) (Watson, 2015). The cell membrane controls the movement of ions and small molecules in and out of the cell, maintaining a favourable environment inside the cell for chemical processes to take place and allowing chemical signalling between cells, as well as generation of electrochemical gradients for energy generation (Nicholls and Ferguson, 2013). To do this a cell membrane contains many integral and peripheral membrane proteins,(Figure 1.2B) (Munro, 2003; Watson, 2015). Chain length, head group, and ratios of sterols:phospholipid: glycolipids, all affect curvature, fluidity and permeability of the membrane and can generate microenvironments in the membrane which concentrate certain proteins in certain regions (Munro, 2003). Integral membrane proteins interact with the hydrocarbon chains of the bilayer using either hydrophobic or amphipathic α-helices or β-barrels, (Figure 1.2C). Disruption of the cell membrane by these integral membrane proteins can allow the movement of ions and other impermeable molecules across the membrane by the formation of channels or pores. Channels allow movement of these molecules in a controlled manner such as sodium/potassium pumps, while pores allow movement in an uncontrolled manner and can be both pathogenic, by destroying the osmotic balance in the cell, and non-pathogenic, which allow removal of toxins from the cell without the disruption of osmotic balance (Castillo et al., 2015; Munro, 2003; Watson, 2015).

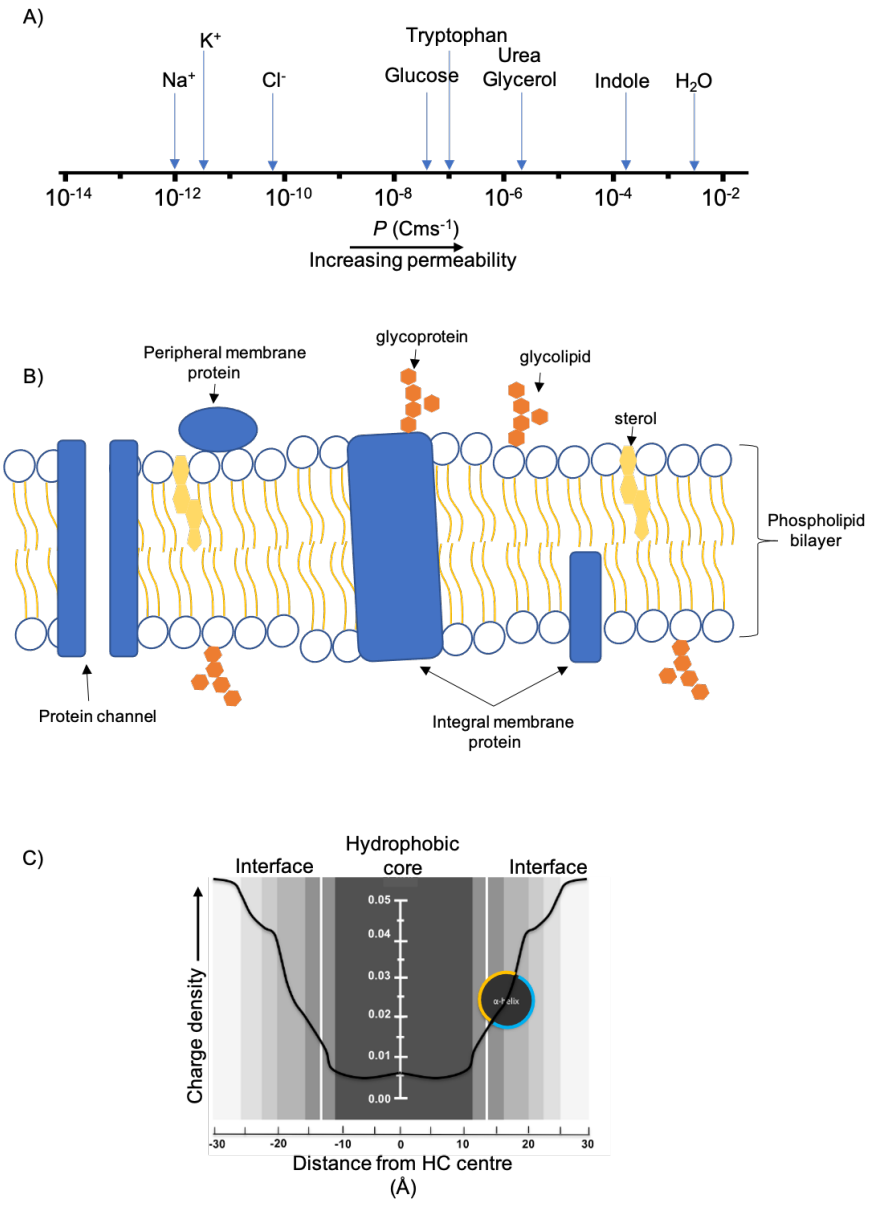


Figure 1.2 **Characteristics and organisation of biological lipid bilayers.** A) The permeability coefficient of different small molecules and ions across a bilayer. B) Schematic of a cellular membrane showing common molecules found within the membrane. Membranes concentrate a range of proteins, both integral and peripheral, including glycoproteins for signalling and cell adhesion and channels for import and export of important molecules from the cell. Cholesterol phospholipids and glycolipids all regulate membrane permeability and fluidity. C) Diagram showing the charge density over a lipid bilayer. Area between white lines at +14 Å and -14 Å show the region of highest hydrophobicity in the bilayer with the black circle illustrating how an amphipathic α-helix could insert into a lipid bilayer, the blue line represents the hydrophilic surface, and yellow represents the hydrophobic surface of the helix, figure was adapted from (Peraro and van der Goot, 2015)

1.3 Bacterial pore-forming toxins

Pore-forming toxins are membrane disrupting peptides found in many pathogenic bacteria including many of the antibiotic-resistant Enterobacteriaceae, and are often one of the major contributors to their virulence, making up 30 % of all bacterial toxins (Los et al., 2013). Enterobacteriaceae have recently been classified as critical in the WHO priority pathogens list for R&D of new antibiotics (WHO, 2017). Understanding how these PFTs work is therefore of great importance in the development of new antimicrobial therapies.

All PFTs promote infection by disrupting the cell membrane, puncturing a hole in the cell membrane, altering the permeability of the membrane, to disrupt osmotic balance, depolarise the membrane, or insert a secondary intracellular toxin, often leading to the death of the target cell, allowing colonisation of the host (Los et al., 2013; Peraro and van der Goot, 2015).

A remarkable feature of PFTs is their transformation from soluble proteins to transmembrane proteins. PFTs are secreted by the bacteria as water-soluble generally monomeric proteins, before undergoing a conformational change and oligomerisation at the host cell membrane to create a membrane-spanning pore (Cosentino et al., 2016). The trigger for these conformational changes can be caused by a range of events including receptor binding (protein or lipid), proteolysis, and in some cases lipid composition. The trigger however for many PFTs is still unknown (Rojko and Anderluh, 2015; Tanaka et al., 2015; Wu and Guo, 2010; Young and Collier, 2007).

There are two main classes of PFTs, categorised based on the secondary structure motif used to span the membrane, α-PFTs use α-helices while β-PFTs use β-barrels (Figure 1.3)(Cosentino et al., 2016; Peraro and van der Goot, 2015). Within these two classes, there are many types of PFTs, which vary by family, stoichiometry and receptor binding type. Many more structures of both the soluble and pore forms of β-PFTs have been determined than α-PFTs, this is due mostly to the greater stability of their β-barrel membrane domain.

As α-PFTs form much less stable pores, indirect biochemical methods have historically been employed to study their mechanisms of assembly, however recent developments in cryo-EM have resulted in the determination of new α-PFT structures (Bräuning et al., 2018; Cosentino et al., 2016; Schubert et al., 2018; Tamm et al., 2004).

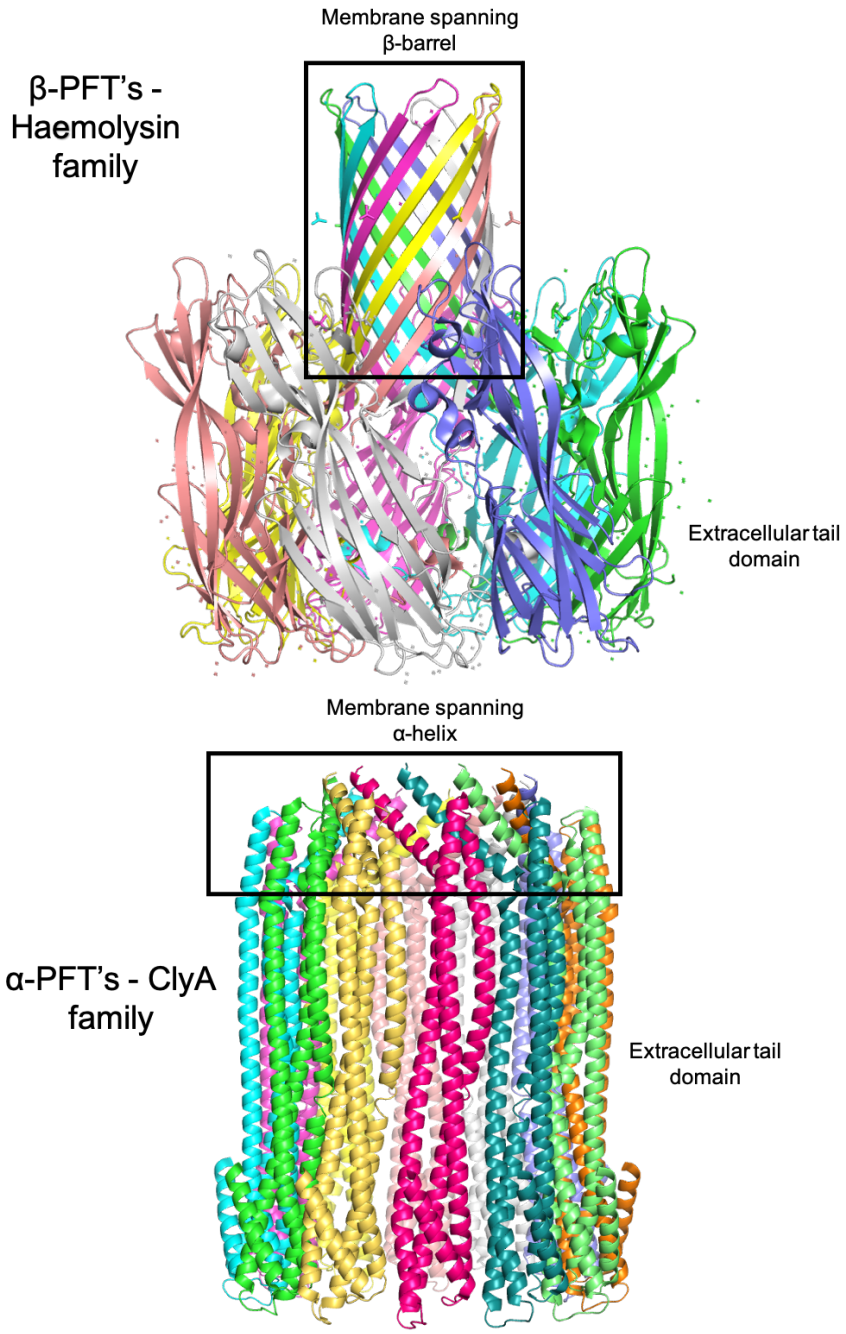


Figure 1.3 **The two classes of PFT.** β-PFTs (top) use β-barrels to span the membrane, while α-PFTs use α-helices (below). Membrane spanning regions highlighted by black boxes, with protomers shown in different colours.

1.3.1 Role in Bacterial infections

In vivo studies with PFTs show these toxins can aid in bacterial infection in one of two ways. The first way is by disrupting the host immune system. Nhe and Hbl in *Bacillus cereus* induce a strong inflammatory response causing serious tissue damage, LLO in Listeria allows internalisation in the host cell by breaking down phagosomes, while YaxAB from *Yersinia enterolitica* directly targets and kills immune cells (e.g. macrophages) (Fox et al., 2020; Gonzalez et al., 2008; Mathur et al., 2019; Wagner et al., 2013). The second method is to damage the host epithelial and endothelial barriers and thus allow the bacteria to enter different organs and spread through the host during early infection. In *Streptococcus pneumoniae* the PLY toxin has been shown to disrupt alveolar-capillary boundary allowing *S. pneumoniae* to enter the blood stream (Gonzalez et al., 2008).

1.3.2 General pore assembly mechanism

As all PFTs share the same requirement of overcoming the membrane barrier, all families have a common mode of action; the soluble protein will bind the membrane, oligomerise and undergo a conformational change to penetrate the membrane (Peraro and van der Goot, 2015). There is great diversity, however, in the mechanisms used in pore formation by different PFTs even within the same family and as such there are a variety of ways by which the above mentioned steps are achieved (Peraro and van der Goot, 2015). These include three general mechanisms (outlined in Figure 1.4);

- Soluble protein binds the membrane followed by pre-pore formation then membrane penetration, forming the active pore.
- Soluble protein binding followed by membrane penetration and oligomerisation.
- Pore assembly in an outer membrane vesicle (OMV) followed by fusion of this OMV with the host membrane.

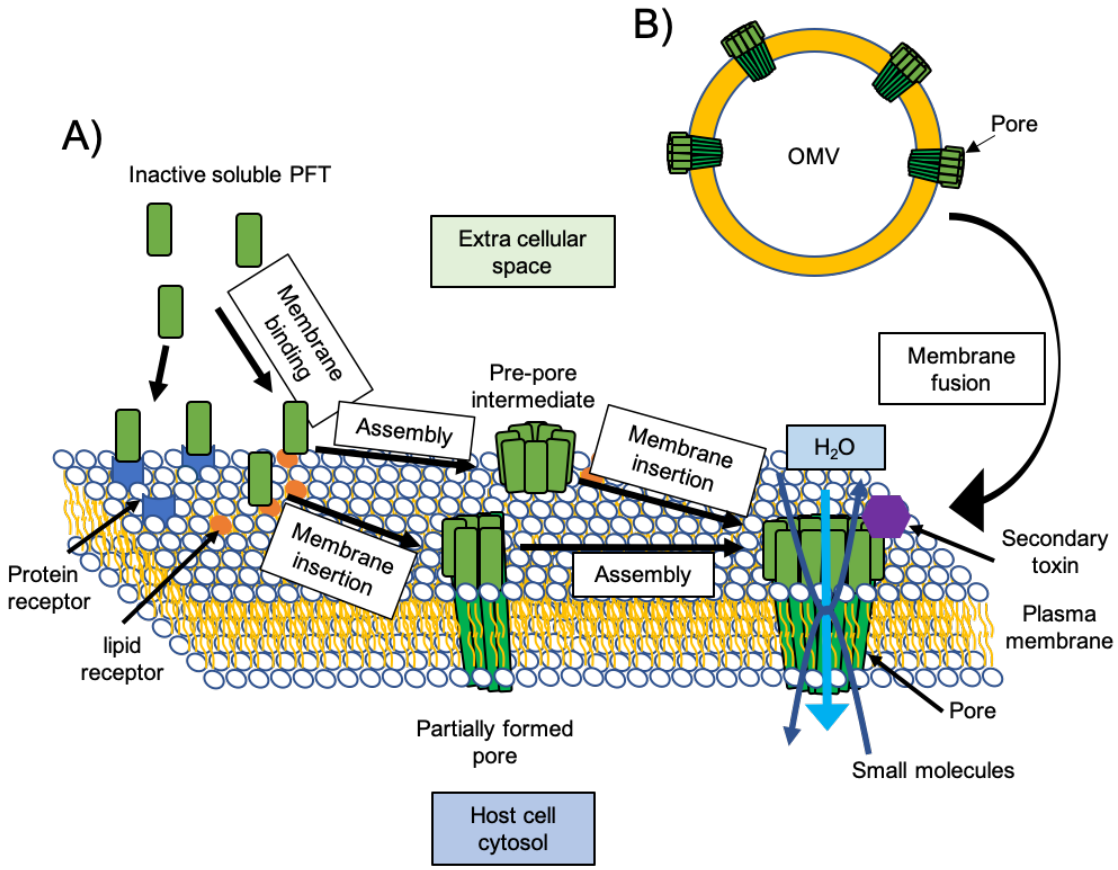


Figure 1.4 **Mechanisms of pore formation.** A schematic showing the different stages of pore formation. A) inactive soluble protein binds to the host cell membrane by interacting with either protein receptors or glycolipid receptors. The inactive protein is activated and will either insert into the membrane then assemble into a pore by oligomerisation, or oligomerise on the membrane surface forming a pre-pore before membrane insertion. The complete pore allows the flow of small molecules, water or a secondary toxin across the membrane. B) Alternatively, the pore assembly mechanisms described in A can occur at the bacteria cell membrane, where the active pores are enclosed in an outer membrane vesicle which fuses with the host cell membrane delivering the active pore.

1.3.3 Excretion of the PFT from the bacteria cell

To date, four methods have been identified by which PFTs are excreted by the bacterial cell. One is by outer membrane vesicles (OMVs), while the remaining three are by bacterial secretion systems (Dongre et al., 2018; Fagerlund et al., 2010; Wai et al., 2003).

Pores have been observed to be excreted by OMVs, such as in the excretion of cytolysin A (ClyA) from *Escherichia coli*, and *Vibrio cholerae* cytolysin (VCC) (Elluri et al., 2014; Wai et al., 2003). ClyA is present in the periplasm as a soluble monomer, these monomers then assemble into inactive pores in the membrane of OMVs, which then associate to the target membrane by an as yet unknown mechanism. In the case of ClyA, it is believed that the OMV provides a favourable environment for disulphide bond reduction. In soluble ClyA found in the bacterial periplasm, the disulphide bond between Cys 87 and Cys 285 prevents the formation of an active PFT. Once in the OMV, the disulphide bond is reduced and ClyA undergoes a conformational transition and oligomerisation to the active pore, which is delivered to the host cell membrane (Wai et al., 2003; Wyborn et al., 2016). These OMVs containing ClyA pores have been visualised using AFM and electron microscopy (Wai et al., 2003)(Figure 1.5A).

Aerolysin, like ClyA, is found as a soluble monomeric protein in the periplasm. This soluble monomer is then transported across the outer membrane to the extracellular space by the type II secretion system, where it then migrates to the host cell membrane and assembles into the active pore (Alouf et al., 2015).

Another method by which PFTs are excreted is the Sec pathway. The unfolded protein polypeptide containing a signal sequence is transported across the bacterial cell membrane by the ATPase SecA and the SecYEG channel (Green and Mecsas, 2016), (Figure 1.5B). In the case of the ClyA family tripartite toxins Nhe and Hbl from *Bacillus cereus*, all three proteins contain a Sec like N-terminal signal sequence. Mutation of this signal sequence prevented secretion, as did the addition of SecA inhibitor azide (Fagerlund et al., 2010). It is

as yet unknown whether Nhe and Hbl then assemble into inactive pores in MVs like ClyA or migrate to the host cell as soluble monomers before assembling into pores like Aerolysin.

More recently, a close homologue of the ClyA family toxins, motility associated killing factor A (MakA) from the motility-associated killing factor operon of *V. cholerae*, was observed to be excreted by the flagella T3SS (fT3SS), with the energy for translocation provided by the proton motive force. MakA was visualised within the flagellar export channel, mutations in the flagellar gene *flha* (encoding the regulator of early flagella genes) reduced secretion of MakA, suggesting it is recognised as an early secretion substrate by the flagella T3SS (Figure 1.5B). This represents a previously unidentified secretion system of bacterial toxins (Dongre et al., 2018).

1.3.4 Membrane receptors and PFT binding

For PFTs that are excreted as soluble proteins from the bacterial cell, these soluble proteins must have a means of attaching to the target cell membrane while the hydrophobic region of the protein is still hidden in the core of the soluble protein. To be able to target a specific cell membrane, and recognise self from non-self, PFTs interact with a range of membrane receptors including sugar, lipid and protein receptors (Peraro and van der Goot, 2015). Figure 1.6 shows examples of PFTs and their interaction with the cell membrane.

Some PFTs interact with specific protein receptors on the host cell membrane. For example, α-haemolysin (Hla), a β-PFT from *Staphylococcus aureus*, has been shown to use ADAM10 as a cellular receptor for binding to the target membrane. As well as acting as a cellular receptor ADAM10 has also been shown to be required for the heptamerisation of Hla. Adam10 is only found in certain membranes (for example it is found on the surface of rabbit erythrocytes but not human), and the requirement for this receptor in toxin assembly explains the species-specific target preference shown by Hla (DuMont and Torres, 2014; Geny and Popoff, 2006). Other examples of interactions with different classes of protein

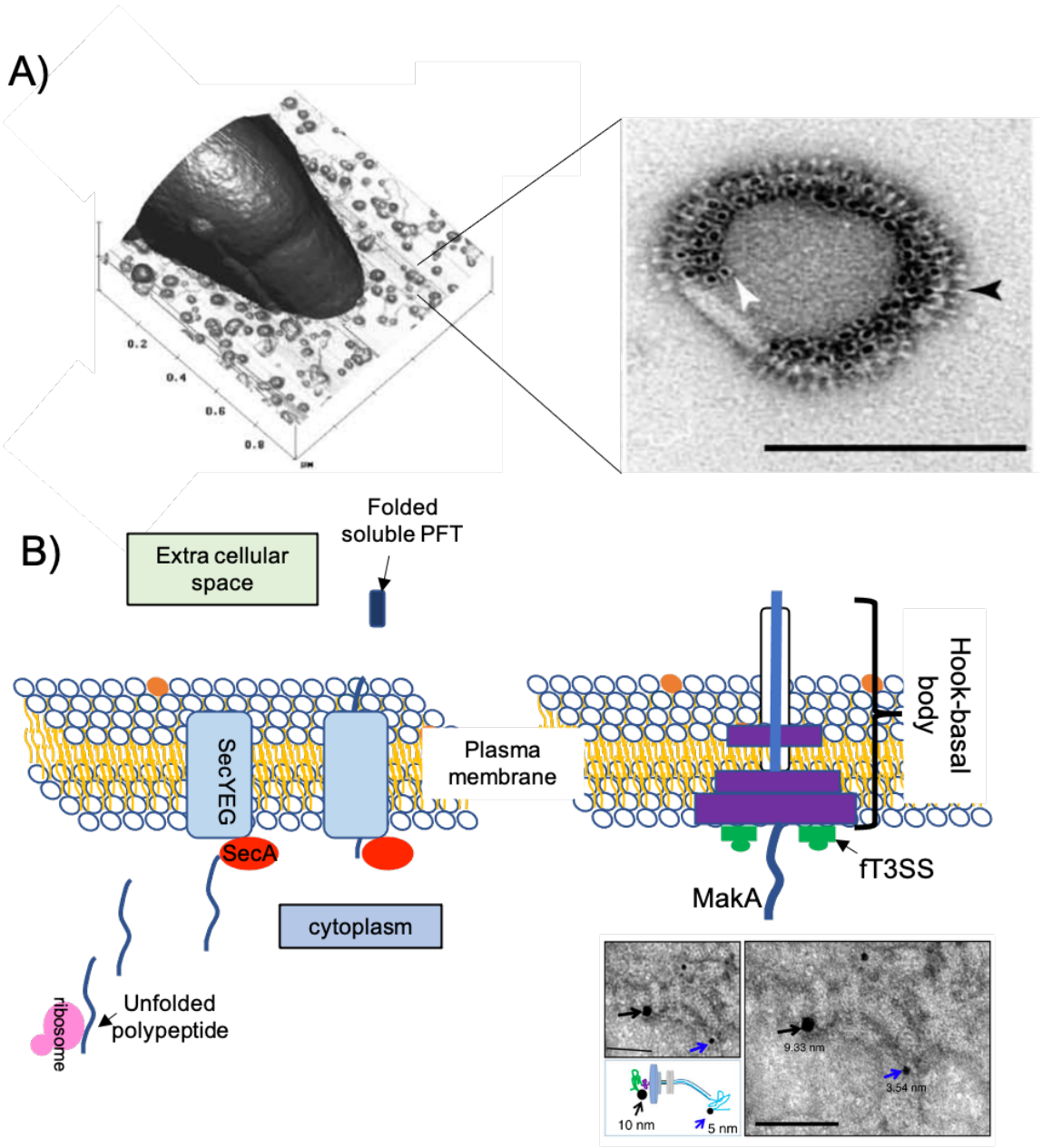


Figure 1.5 **Export and membrane recognition of PFTs.** A) ClyA is exported from the bacterial cell as assembled pores in OMVs. Atomic force microscopy of a bacterial cell surrounded by OMVs (left) and negative electron microscopy of ClyA pores in OMV (right), (image adapted from (Wai et al., 2003)). B) Left, B. cereus tripartite toxins are exported via the Sec system as an unfolded polypeptides, these then fold into soluble proteins on reaching the extracellular space. Right, a homologue of the tripartite Nhe toxin, MakA is exported through the flagella T3SS, electron microscopy showing localisation of MakA in the flagella, black dot shows immunogold labelled MakA (image adapted from (Dongre et al., 2018).

include aerolysin (in all *Aeromonas* species), which interacts with the GPI anchor of GPI anchored proteins as well as to the N-linked sugar domain, leading to very high-affinity interactions (Peraro and van der Goot, 2015).

Other PFTs have been shown to favour certain lipids or lipid environments. Both EqtII (from *Actinia equina*) and Lysenin produced by earthworms (*Eisenia fetida*) are β-PFTs which are sphingomyelin (SM) specific toxins. It has been found that although both interact with SM, lysenin prefers clustered SM while EqtII prefers dispersed SM and as such bind different regions of the cell membrane (Bhat et al., 2013; Makino et al., 2015). Actinoporins (from Sea anemones) also bind SM, however, the addition of cholesterol improves binding, this is likely due to cholesterol promoting phase separation making SM more accessible (Rojko et al., 2016). These three proteins show that although they all need SM there is also a requirement for these lipids to be in a specific lipid environment in the membrane.

Cholesterol is also directly involved in membrane interaction in several PFTs, including Aerolysin, CDCs and ClyA, which interact with the 3'-OH and steroid rings of cholesterol. In cholesterol-dependent dependent cytolysins (CDCs), removal of 90 % of erythrocyte membrane cholesterol led to <0.1 % of the CDC, perfringolysinO's (from *Clostridium perfringens*), original lytic activity (Giddings et al., 2003). Cholesterol has also been shown to play an important role in stabilisation, and activation of the α-PFT ClyA. (Sathyanarayana et al., 2018). Cholesterol has been proposed to be the mechanism by which ClyA can recognise self from non-self as bacterial membranes do not contain cholesterol while mammalian cells do. The requirement for cholesterol to stabilise the N-terminal helix of ClyA and activate the pore acts as a control mechanism to prevent lysis of the bacterial cell when assembling in the outer membrane for transport in OMVs. On reaching the host membrane the presence of cholesterol stabilise the N-terminal helix, activates the pore and lysis occurs. (Sathyanarayana et al., 2018).

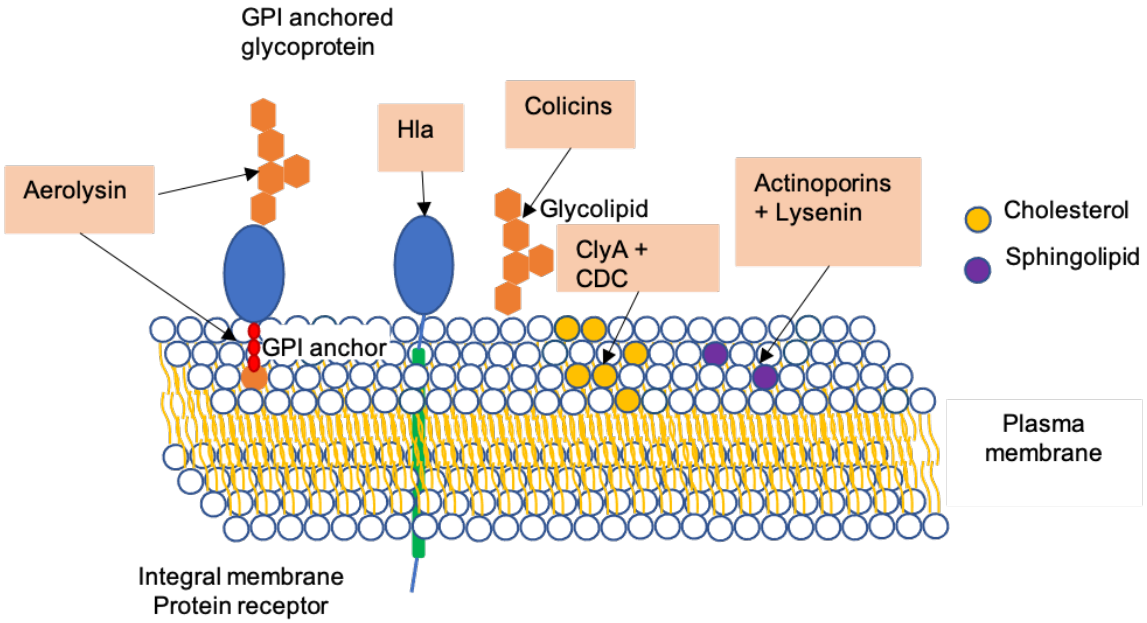


Figure 1.6 **PFTs recognise and bind to the membrane via proteins, sugars and lipids.** Interaction sites of different PFTs are highlighted.

1.3.5 Pore assembly

Once PFTs have bound to the target membrane they must next undergo a conformational change to expose their hydrophobic domains, insert into the membrane, and oligomerise. As explained above, the order in which this conformational change and oligomerisation occurs varies between toxins. Generally the β-PFTs oligomerise at the membrane surface before inserting their hydrophobic domains into the membrane, while the α-PFTs tend to insert into the membrane before oligomerisation (Cosentino et al., 2016).

PFTs can assemble into:

1. Barrel-stave pores (e.g ClyA, aerolysin), with hydrophilic residues on either side of the membrane and in the pore lumen, and hydrophobic residues partitioning the lipid from the solvent and forming an enclosed pore.

2. Protein lipid pores (e.g fragaceatoxin C (FraC) from *Actinia fragacea*), made up of alternating lipid and protein (Cosentino et al., 2016) (Figure 1.7).

Aerolysin is an example of a β-PFT barrel-stave pore. Like most β-PFTs a pre-stem-loop (Figure 1.7) is extracted and anchors the protein to the membrane, this must then join with other pre-stem-loops of other protomers to form a pre-pore structure on the membrane surface, before undergoing a final large conformational change to insert its β-barrel into the membrane,(Figure 1.7) (Iacovache et al., 2016). ClyA, like aerolysin, is a barrel-stave pore, which undergoes a large conformational change from soluble to pore form (involving 60 % of the residues). However, ClyA is an α-PFT and inserts into the membrane before oligomerisation. The conformational change results in a hydrophobic section becoming exposed and inserting into the membrane, acting in a similar way to the stem-loop of aerolysin. At the same time, the N-terminal helix swings through 180 degrees to span both leaflets of the membrane. The oligomerisation of a number of protomers then creates the final active pore (Figure 1.7)(Roderer and Glockshuber, 2017). A detailed outline of these conformational

changes is shown in Figure 1.8. The final stoichiometry of the ClyA-family α-PFTs has been shown to vary, with ClyA forming 10-13mer pore, equally some β-PFTs such as the anthrax toxin have also been shown to vary in size.The diameter of each pore also plays an important role in what can pass through these PFT. CDCs are examples of large β-PFTs (diameter 200-300 Å) which allow movement of proteins as well as ions across the target cell membrane (Reboul et al., 2016). The ClyA family of α-PFTs, however, is comprised of pores ranging from 12 Å (Hbl) – 50 Å (Nhe) in diameter, much smaller than CDCs, and have only been observed to allow the flow of ions such as K+ through the membrane (Eifler et al., 2006; Mathur et al., 2019; Wagner et al., 2013; Young and Collier, 2007).

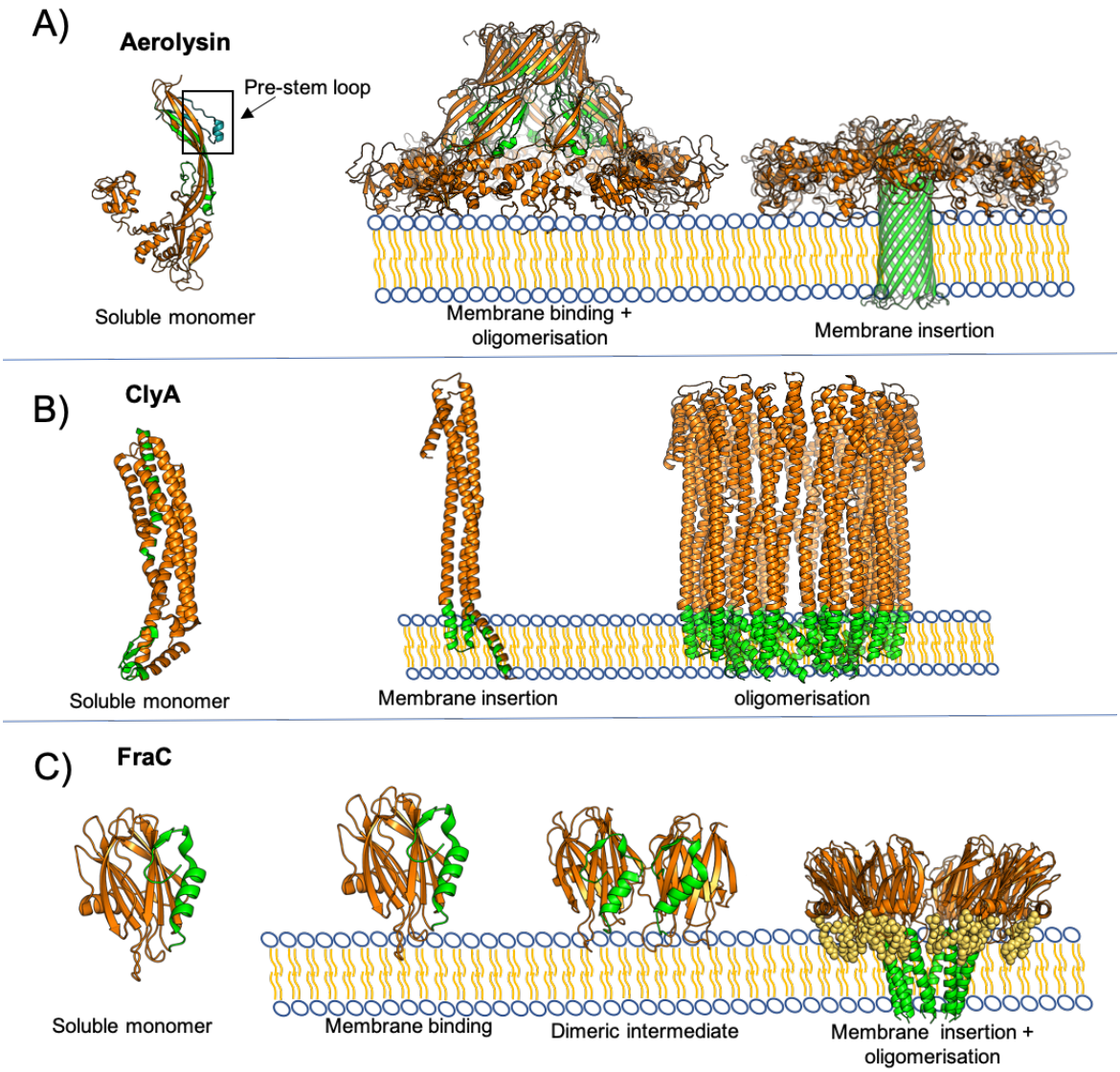


Figure 1.7 **PFTs assemble to form a diverse range of pore types.** Orange represents the extracellular domain, while membrane spanning domains are shown in green. A) β-PFT, aerolysin forms a pre-pore on the membrane before insertion of the β-barrel membrane-spanning domain. B) α-PFT, ClyA, inserts into the membrane before oligomerisation. C) The protein-lipid pore, FraC dimerises on the membrane surface before forming a nanopore with alternating protein and lipid molecules. Yellow spheres represent detergent

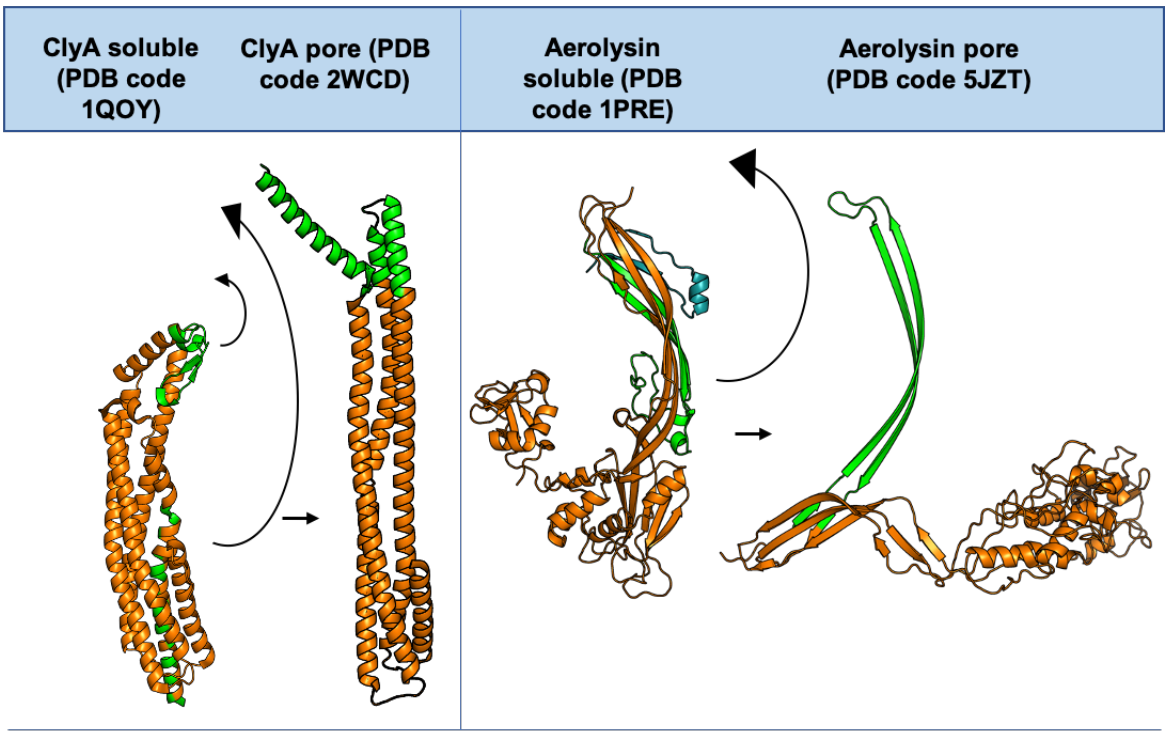


Figure 1.8 **Conformational changes during soluble to pore transformation.** Like most PFTs, a large conformational change occurs in the transition from soluble to pore for both ClyA (left) and aerolysin (right), membrane regions shown in green, and exposed in the pore conformation.

1.3.6 Cell response to PFT attack

Puncturing of a mammalian membrane by PFTs triggers a counter-attack by the host in the form of the inflammatory response, as well as membrane repair mechanisms. NOD-like receptors in the cytoplasm of mammalian cells can recognise PFT inserted in the membrane and activate the pro-inflammatory response, including the production of IL8 which recruits neutrophils to clear the bacterial infection (Gonzalez et al., 2008). K^+ efflux caused by the ClyA family PFTs also triggers the inflammasome. Over-activation of the inflammatory response can, however, be detrimental to the host, in the case of *B.cereus*, efflux of K^+ by Nhe and HBL toxins result in over-stimulation of the inflammatory response and devastating tissue damage (Fox et al., 2020; Mathur et al., 2019). Cells targeted by PFTs are also able to repair their membranes, three mechanisms have been proposed by which membrane repair is triggered by PFTs, these include the activation of the p38 mitogenic kinase pathway, calcium influx and a less-understood mechanism whereby the pore is closed by constriction of the transmembrane region of the pore (Gonzalez et al., 2008).

1.4 The ClyA family of PFTs

The ClyA family of α-PFTs is made up of single (ClyA), two (YaxAB, XaxAB and PaxAB) and three component (NheABC (Nhe), HblL1L2B (Hbl), and AhlABC (Ahl)) members, (Structures, PDB codes and bacterial species are listed in Table 1.1 (Beecher et al., 1995; Bräuning et al., 2018; Lindbäck et al., 2004; Oscarsson et al., 1999; Schubert et al., 2018). Although these members share low sequence identity (18-40 %), their shared structural similarities (rmsd Cα 2.4-3.8 Å) allowed for them to be classified as a new α-PFT family (Bräuning et al., 2018; Fagerlund et al., 2008).

All members of the ClyA family confer lytic activity following pore assembly on the target cell membrane. Structures of both the soluble protein and assembled pore have been

solved for the single component ClyA toxin as well as the two component YaxAB and XaxAB toxins (Bräuning et al., 2018; Eifler et al., 2006; Schubert et al., 2018; Wallace et al., 2000). Together with biochemical assays these structures have provided a good understanding of the activity and assembly of both the single and two component members of the family. To date, for the tripartite members of the ClyA family, only the soluble structures of the A component of the Nhe toxin (NheA) and the C component from the Hbl toxin (HblB) have been solved (Ganash et al., 2013; Madegowda et al., 2008), and thus a large amount of the current knowledge of these tripartite systems has been gained from in-depth biochemical analyses.

The method of export, however, of these ClyA family proteins from the bacterial cell is not well understood. As discussed in Section 1.3.4, ClyA is found in the periplasmic space as a soluble monomeric protein which then assembles into an inactive pore in OMVs which deliver it to the target cell membrane. The proteins of both Nhe and Hbl are transported across the inner membrane by the Sec secretory system, however, the method by which they are delivered to the target membrane is as yet unknown.

1.4.1 Cytolysin A (ClyA)

ClyA is found in many pathogenic species of Salmonella, *E. coli*, and *Shigella*. As ClyA shared no homology to other haemolysins it was classified as a new family of cytotoxin (Libby et al., 1994). Studies in *Salmonella typhi* showed that mutation of the *clyA* gene resulted in a reduced invasion of human epithelium cells and growth in macrophages, It was therefore proposed that ClyA was an important virulence factor which aided colonisation of the host (Fuentes et al., 2008; Libby et al., 1994). Further to this ClyA was shown to be expressed in anaerobic conditions in pathogenic strains of *E. coli* and was cytotoxic to macrophages and lysed mammalian cells (Oscarsson et al., 1999), further highlighting the importance of ClyA in pathogenicity.

Table 1.1 Members of the ClyA α-PFT family.

Structure name	Bacterial species	PDB code	Mono, bi or tripartite	Soluble or Pore
CytolysinA (ClyA)	*E.coli*	1QOY	Mono	Soluble
ClyA pore	*E.coli*	2WCD	Mono	12mer-Pore
Yersinia α - xenorhabdolysin (Yax) A	*Y. entrerocolitica*	6EK7	Bi	Soluble
YaxB	"	6EK8	Bi	Soluble
YaxAB	"	6EL1	Bi	10mer-Pore
Xenorhabdus α - xenorhabdolysin (Xax) A	*X. nematophila*	6GY8	Bi	Soluble
XaxB	"	6GY7	Bi	Soluble
XaxAB	"	6GY6	Bi	13mer-Pore
Non-haemolytic enterotoxin (Nhe) A	*B.cereus*	4K1P	Tri	Soluble
Haemolysin BL (Hbl) B	*B.cereus*	2NRJ	Tri	Soluble
Aeromonas haemolytic toxin (Ahl) C	*Aeromonas hydrophila*	2H2E	Tri	Soluble

Structure and assembly of ClyA

ClyA, like all PFTs, is first expressed as a soluble protein, where the hydrophobic regions used to span the membrane in its pore conformation are hidden within the core of the soluble protein. Soluble ClyA is composed of a 5 helical bundle tail domain, and a head domain made up of two short α-helices and two small β-strands (β-tongue) which is connected to the tail domain by two long α-helices that span the length of the protein (α3 and α4). The β-tongue is entirely hydrophobic and folds so that the hydrophobic residues pack against α1, α3, α4 and the short helices within the head (Figure 1.10), hiding them from the solvent. The N-terminal helix (α1) is amphipathic, again packing to hide its hydrophobic residues within the core of the protein against α2, α3, α4, and α5 (Figure 1.9) (Wallace et al., 2000).

In the pore structure of ClyA the head domain has undergone a conformational change with the β-tongue now providing a hydrophobic, 14 Å helix-turn-helix extension to α4 and α5, long enough to extend through a single leaflet of the membrane (Figure 1.10A). α1 is now able to swing down into a position previously occupied by the head domain, forming an extension to α2 with a kink at the membrane interface. The N-terminus packs its hydrophobic residues against the hydrophobic helix-turn-helix and the lipid tails of the membrane (Figure 1.10B). This results in an entirely helical final pore conformation, with exposed hydrophobic domains. This protomer oligomerises to form the pore, with the N-terminal amphipathic helix spanning both leaflets of the membrane and providing the hydrophilic lumen to the pore (Mueller et al., 2009).

These structures, together with studies of the assembly dynamics of ClyA (Roderer and Glockshuber, 2017), have allowed for a pore assembly mechanism to be proposed, whereby ClyA forms a molten globule intermediate state on its transition from soluble to pore form (Benke et al., 2015). As described above the pore conformation is stabilised by cholesterol in the membrane. Once in the pore conformation, protomers dimerize and associate with existing oligomers forming a sealed pore via extensive surface interactions

with neighbouring protomers (Benke et al., 2015). While it is clear that the N-terminal helix provides the transmembrane spanning region of the ClyA pore, it is less clear what the role of the β-tongue is. Three outward-facing hydrophobic residues in the β-tongue were mutated to polar residues, abolishing activity and preventing ClyA binding the membrane (Wallace et al., 2000), leading to a proposition that the β-tongue anchored the soluble ClyA monomer to the membrane. However, more recent studies using both mutations and deletions of the β-tongue have disputed this finding, showing ClyA can still bind the membrane, but with severely impaired activity, suggesting instead that the β-tongue is needed to maintain the open conformation of the pore (Fahie et al., 2018). It is possible that the β-tongue may enhance binding to the membrane as well as maintain the open pore.

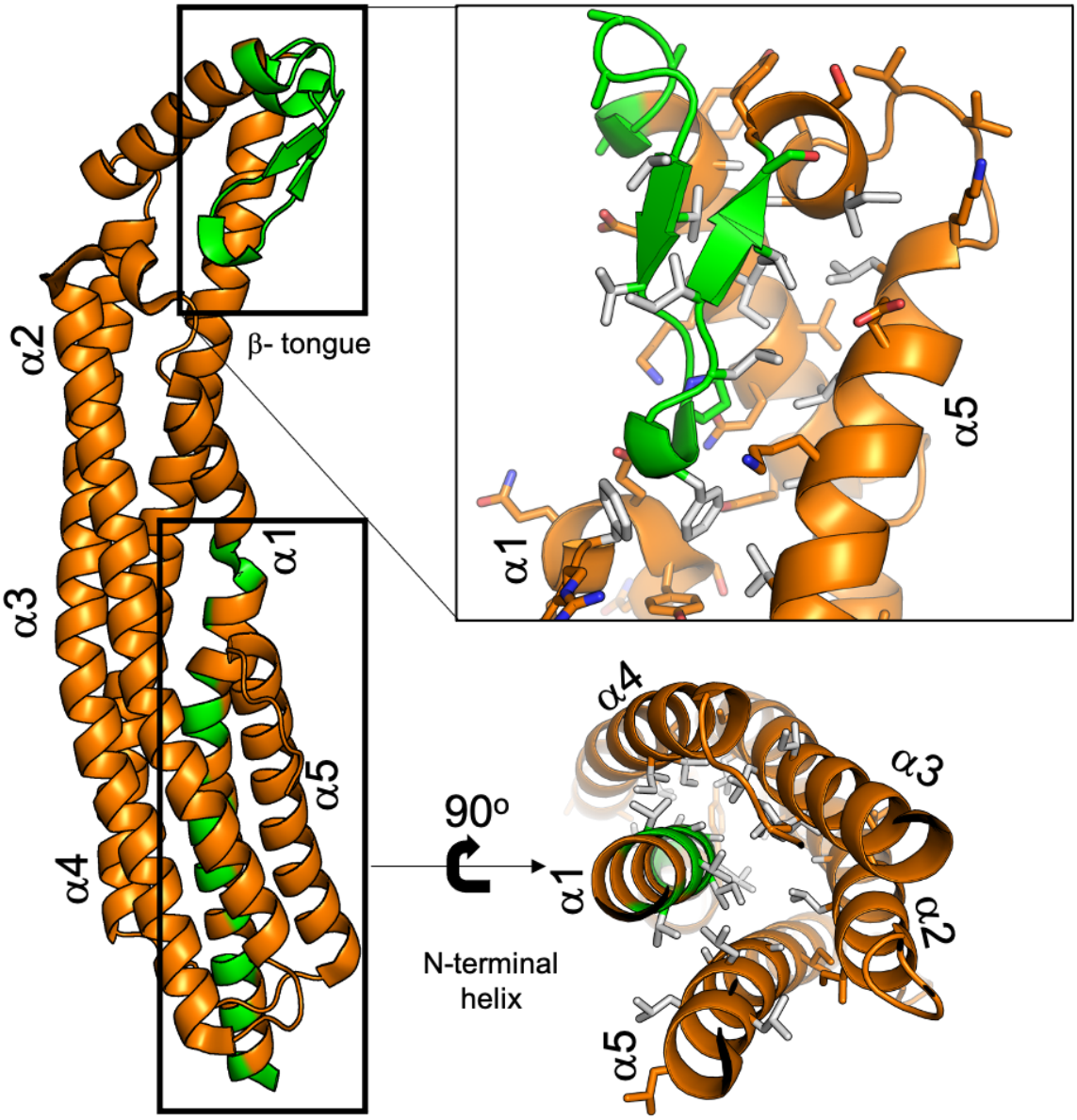


Figure 1.9 **Soluble ClyA hides it hydrophobic residues in the β- tongue and N-terminal helix.** Green shows the membrane-spanning domains, while hydrophobic residues are shown in white. Hydrophobic residues in the β- tongue and N-terminal helix are hidden within the core of the soluble protein.

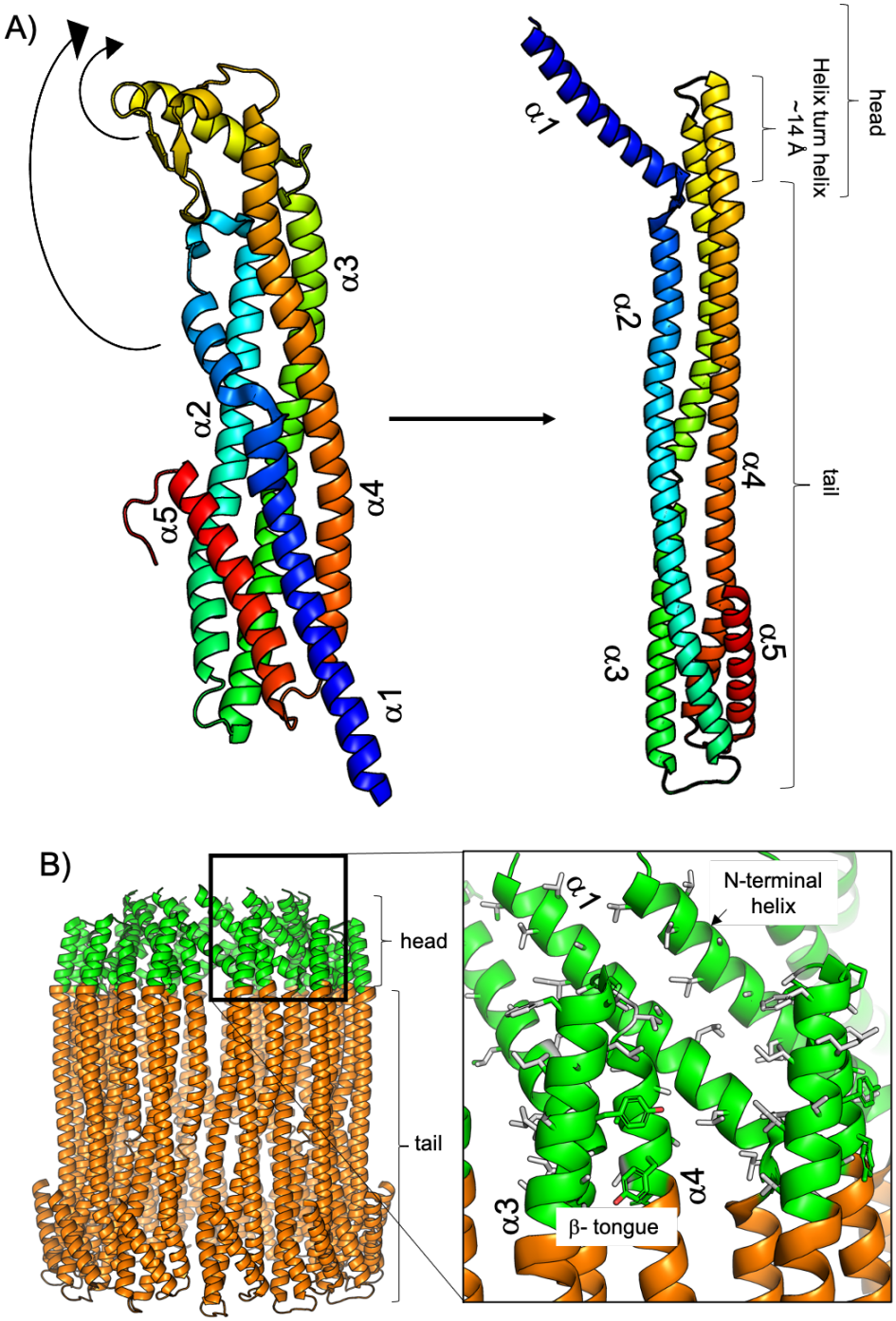


Figure 1.10 **ClyA undergoes a large conformational change to expose its hydrophobic residues.** A) A large conformational change results in the movement of the N-terminal α1 and the β- tongue (coloured blue, N-terminus to red, c-terminus). B) On pore formation the hydrophobic residues are now exposed and face the exterior of the pore to interact with the lipid tails of the membrane.

1.4.2 Bipartite ClyA family of α-PFTs

The bipartite ClyA family members were first identified in the insect pathogen *Xenorhabdus nematophila*, a potent killer of many insects. Knowledge of its toxins could therefore provide important insights for the development of new pesticides. The XaxAB cytotoxin was shown to be made up of two proteins (XaxA and XaxB), both of which were required for cytotoxic, and haemolytic activity. Furthermore, inactivation of the XaxAB toxin resulted in complete loss of cytotoxic activity in *X. nematophila* (Vigneux et al., 2007). Further bipartite toxins have been identified in many Enterobacteriaceae, including *Yersinia enterocolitica* (YaxAB) and *Photorhabdus luminescens* (PaxAB). As with XaxAB, both YaxAB and PaxAB conferred cytotoxic activity to cells in their host organism, but only if both proteins are expressed simultaneously (Wagner et al., 2013; Zheng et al., 2012).

Structure and assembly of the Bipartite α-PFTs

The pore structures of YaxAB and XaxAB form the same oligomeric structures containing 8 – 12 dimers of the A and B proteins (Bräuning et al., 2018; Schubert et al., 2018). These bipartite YaxAB and XaxAB pores share many similarities to the ClyA pore. The B component contains an amphipathic helix-turn-helix in its membrane-spanning head domain and fulfils the role of the N-terminal helix in ClyA, forming the hydrophilic lumen of the pore (Figure 1.11A)(Bräuning et al., 2018; Schubert et al., 2018). The A component provides the short hydrophobic helix-turn-helix, which would span a single leaflet, in the same way as the β-tongue of ClyA (Figure 1.11A)(Bräuning et al., 2018; Schubert et al., 2018). These features resulted in YaxAB and XaxAB being classified as members of the ClyA-family (Bräuning et al., 2018; Schubert et al., 2018). The dimer itself is held together by two interfaces between A and B, one in the tail by a close network of polar and hydrophobic residues and the second at the head between well conserved hydrophobic residues (Figure 1.11B). The interface between neighbouring dimers is found at a hydrophobic patch in A and B located at

the juncture between the neck and tail (Figure 1.11B)(Bräuning et al., 2018; Schubert et al., 2018).

The soluble structures of YaxA, YaxB, XaxA and XaxB, have also been determined. All A structures and all B structures share the same overall fold. The A components consist of a 5 helical bundle tail domain, with two extended helices (neck) connecting the hydrophobic head domain to the tail (Figure 1.12A). Unlike other PFTs, the A component is in an almost identical conformation in both the soluble and pore structures. In the crystal structures of YaxA and XaxA, the hydrophobic residues in the head domain are exposed to the solvent, rather than hidden within a β-tongue as seen in ClyA. Only a slight twist in the head domain helices are required to allow it to pack against the hydrophobic residues of the B component head domain in the pore (Figure 1.12B). There is no major conformational change to expose the hydrophobic residues, posing the question of whether this exposed head conformation is an intermediate, and the soluble conformation is similar to that of ClyA (Bräuning et al., 2018; Schubert et al., 2018). The B component of these bipartite toxins like the A component contains a five-helical bundle tail domain and two extended helices (neck) connecting the head and tail domains. In the B component, however, the head is comprised of 2 short α-helices (rather than a β-tongue) which pack perpendicular to the neck helices, allowing the hydrophobic residues of this head domain to be hidden from the solvent in a new fold not seen in ClyA (Figure 1.12A) (Bräuning et al., 2018; Schubert et al., 2018). In the B component only the head domain undergoes a conformational change with the two short helices in the head straightening, this results in the extrusion of the hydrophobic residues of the head for insertion through both leaflets of the membrane (Figure 1.12B)(Bräuning et al., 2018; Schubert et al., 2018).

Liposome float assays with YaxA and YaxB have shown that YaxA can bind membranes alone, while YaxB cannot bind unless YaxA is present. This together with mutagenic analysis of the hydrophobic head of YaxA, which showed that removal or addition of polar

residues to the head significantly reduced haemolytic activity, suggested that A plays a role in both recruitment of B to the membrane and stabilising the pore form of B, similar to the β-tongue of ClyA. Furthermore, premixing of YaxA and YaxB resulted in no binding of either component to the membrane, instead an off-pathway soluble complex was formed. A mechanism of pore assembly for bipartite α-PFTs has thus been proposed whereby the A component binds the membrane first, followed by the B component, these dimers then oligomerise to the final pore, with possible off-pathway soluble complexes also being formed (Bräuning et al., 2018).

Regulation of expression of the YaxAB toxin is controlled by the transcriptional regulator RovA (responsible for temperature-dependent expression) (Wagner et al., 2013).

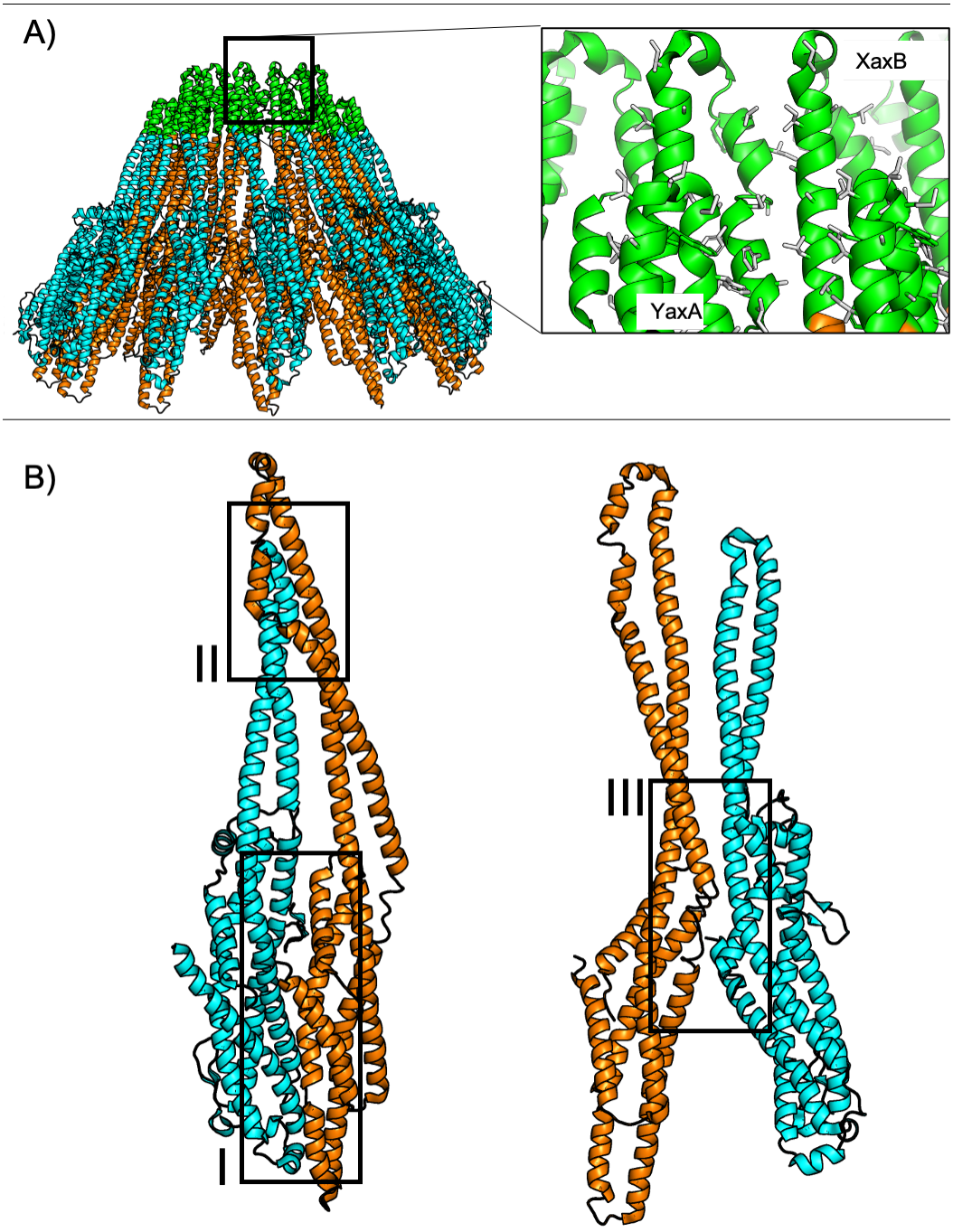


Figure 1.11 **The YaxAB pore.** A) The pore structure of YaxAB (blue, YaxA, orange, YaxB) shows packing of the hydrophobic residues of A against those of B, and facing the exterior of the pore. Green shows the membrane-spanning domains, while hydrophobic residues are shown in white. B) The YaxAB dimer has three interfaces, I - polar and hydrophobic residues, II- hydrophobic residues (left), and between dimers, III- hydrophobic patch (right). YaxA, blue, YaxB, orange.

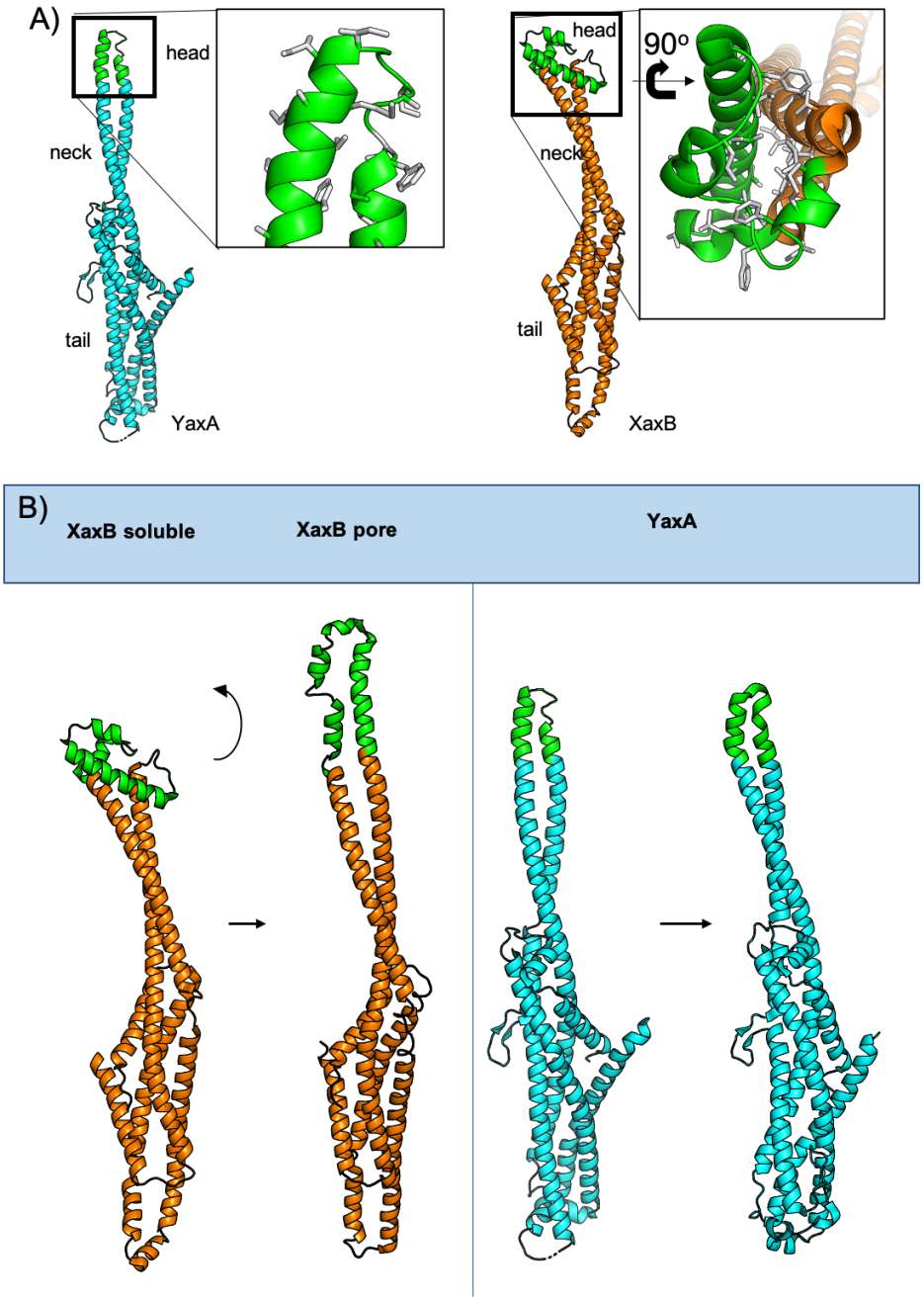


Figure 1.12 **Structures of The bipartite toxins show different mechanisms for concealing the hydrophobic heads and pore assembly to ClyA.** Green shows the membrane-spanning domains, while hydrophobic residues are shown in white. A) YaxA crystal structure shows the hydrophobic head is exposed to solvent in the soluble conformation (left). Soluble XaxB crystal structure shows a novel fold used for concealing the hydrophobic residues. B) The bipartite toxins undergo much smaller changes, with no major conformational change seen for the A component.

1.4.3 Tripartite ClyA family α-PFTs, NheABC and HblL1L2B

Of all the members of the ClyA family α-PFTs, the tripartite α-PFTs are the least well characterised due to the lack of structural information. This has led to detailed biochemical analysis of the family to try and elucidate the pore assembly mechanisms, and has shown that although both Nhe and Hbl are tripartite α-PFTs they have adopted different mechanisms of pore assembly.

The first tripartite ClyA family α-PFT to be identified was the Hbl toxin in diarrheal food poisoning strains of *B. cereus*, which was shown to possess haemolytic, cytotoxic and vascular permeability activity (Thompson et al., 1984). The Hbl toxin was identified as the virulence factor responsible for fluid accumulation in rabbit ileal loops and the key virulence factor for diarrhoea in *B. cereus* (Beecher et al., 1995). After a large food poisoning outbreak in Norway a second tripartite ClyA family α-PFT was also identified in *B. cereus*. The strain of *B. cereus* isolated from this outbreak, however, did not contain the Hbl toxin, instead only a non-haemolytic cytotoxin named Nhe (Lund and Granum, 1996). Both of these toxins are composed of three proteins, namely Hbl-L1, Hbl-L2, Hbl-B for the Hbl toxin and NheA, NheB, and NheC, for the Nhe toxin. In both cases, it was shown that all three proteins were required for cytotoxic activity, and EM has shown the formation of pores in liposome membranes (Fox et al., 2020; Lindbäck et al., 2004; Mathur et al., 2019; Tausch et al., 2017). Furthermore, the ratio at which these three proteins are present has a direct effect on the activity of the toxin, a feature of the tripartite toxins that will be described in more detail below (Lindbäck et al., 2004; Tausch et al., 2017). Recent studies of the Nhe and Hbl tripartite α-PFTs show that both can activate the NLRP3 inflammasome leading to septic shock in mice, and inhibition of this response prevents lethality, demonstrating the importance of this toxin family in bacterial pathogenicity (Fox et al., 2020; Mathur et al., 2019).

Transcription of the Nhe and Hbl operons has been shown to be regulated by a number of global regulators and 5' UTRs to fine-tune expression (Böhm et al., 2016). Among the regulators are PlcR (activated by the onset of stationary phase), and CodY (nutrient responsive repressor). The presence of CodY suggests a possible nutrient regulation mechanism for the Nhe system, indeed strong activation of Nhe and Hbl resulted from depletion of free amino acids (Böhm et al., 2016).

Structure of Hbl-B and role in pore assembly

The structure of the soluble Hbl-B protein is currently the only known structure for the Hbl toxin. Although sequence identity between ClyA and Hbl-B is low (32 %) both proteins share a common structure with a β-tongue containing head domain (Madegowda et al., 2008)(Figure 1.13A). Like the other ClyA family α-PFTs, Hbl-B has a five helical bundle tail domain and a head domain comprised of two short helices and a β-tongue. Like in ClyA the β-tongue domain contains a 17 residue hydrophobic segment, however the head domain of Hbl-B makes significant interactions with the tail domain unlike in ClyA, where these interactions are minimal (Figure 1.13A). Based on the structural similarities in the head domains of both ClyA and Hbl-B it was proposed that a similar conformational change occurs whereby the β-tongue refolds into an extended helix-turn-helix exposing the hydrophobic residues, allowing interaction with the membrane (Madegowda et al., 2008).

For pore formation, Hbl components must assemble sequentially with an obligatory binding order of HblB - HblL1 – HblL2 (Figure 1.13B). Hbl-B has been shown to be able to bind membrane alone, similarly to YaxA and XaxA (Bräuning et al., 2018; Schubert et al., 2018). Priming of the membrane with B is a crucial step in pore formation and it has been shown that high amounts of L1 or L2 do not advance pore formation if insufficient B is bound to the target membrane. Further to this, excess of L1 or depletion of B hindered pore formation, while L2 did not seem to be crucial. Finally excess of L1 or L2, or depletion

of B resulted in a reduction in toxic activity. These results led to a proposal that a ratio of 1:1:10 = L2: L1: B is required for fastest pore formation and 1:2:2 is required for maximum cytotoxicity (Jessberger et al., 2019). The inhibitory effects seen with excess L1 and L2 and the requirement for excess B could be explained by the formation of off-pathway complexes observed to form between L1 and B, and L2 and B, similar to the soluble off-pathway YaxAB complex. It is proposed that naturally, most Hbl-B is bound in soluble complexes with L1 or L2 and likely acts as a control to pore assembly (Jessberger et al., 2019; Tausch et al., 2017).

A mechanism for Hbl has recently been proposed where Hbl-B binds to the target membrane as the first step. This is the crucial and rate-limiting step in pore formation and occurs in a few minutes. Free Hbl-L1 then binds B followed by L2 forming the final pore, these two processes are rapid once Hbl-B has bound (Jessberger et al., 2019).

NheA structure and the role of each Nhe component in pore formation.

NheA is the only component of the Nhe system with a known structure. It shares very low sequence identity with ClyA as well as Hbl-B (20 % and 18 %, respectively). However, as observed with the other family members it shares a very similar structure, again with a five helix tail domain and a β-tongue containing head domain (Figure 1.14A). In NheA the β-tongue is significantly larger than both that of Hbl-B and ClyA, however it packs closely against the tail as in Hbl-B (Figure 1.14A). In addition, the β-tongue is amphipathic, not hydrophobic as previously described in Hbl-B and ClyA. It was therefore proposed that the β-tongue in NheA could undergo the same conformational change as in ClyA to form two significantly longer amphipathic α-helices that could traverse the membrane (Ganash et al., 2013).

Monoclonal antibody binding studies with NheA have revealed that NheB is the only binding partner for NheA, as NheA is only present on the cell surface when NheBC complexes are present, not when NheC is bound to the membrane alone (Didier et al., 2016).

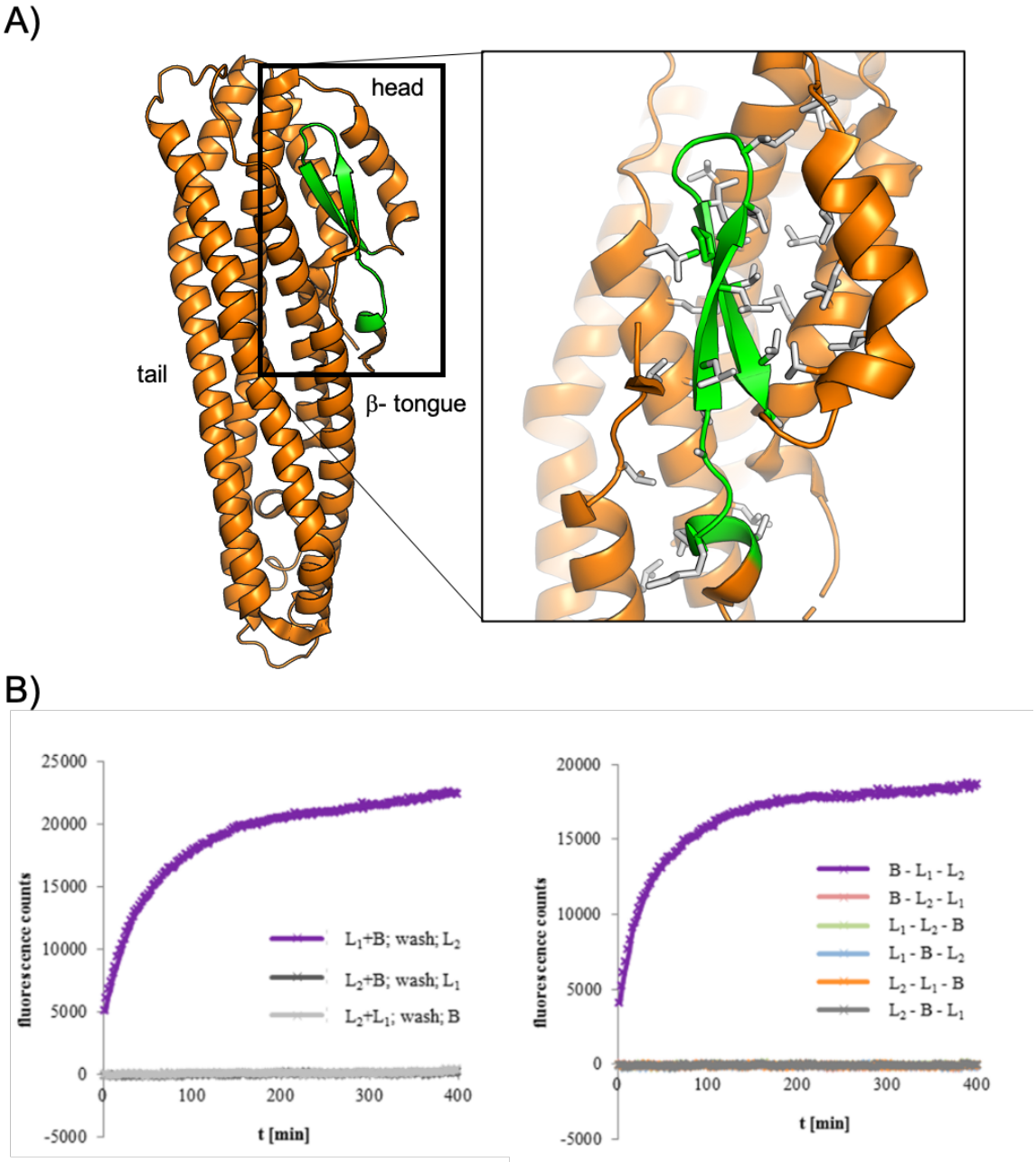


Figure 1.13 **Structure of Hbl-B and activity assays for the tripartite Hbl toxin.** A) Hbl-B hides its hydrophobic head within a β- tongue motif in the same way as ClyA. Green shows proposed membrane-spanning domains, while hydrophobic residues are shown in white. B) Activity assays for Hbl. When Vero cells are lysed propidium iodide (PI) is released, the level of PI is measured by fluorescence counts. Lytic activity is only seen when cells are preincubated with L1+B before the addition of L2 (Left) or when proteins are added sequentially in the order of B-L1-L2 (right). (Figure adapted from (Jessberger et al., 2019).

Furthermore, NheB binds to NheA via the N-terminus, and the conformational transition from soluble to pore in NheA occurs about the C-terminus (Didier et al., 2016). It has previously been shown that a ratio of NheB: NheC of between 50:1 and 5:1 is optimum for cytotoxicity, due to the formation of NheBC complexes in solution that prevent the binding of free NheB to the membrane to form an active pore. However, these NheBC complexes can bind to the membrane and are more stable than NheC binding alone, it is therefore vital that a balance between complexed NheB and NheC, and free NheB is maintained. This allows stable binding of NheC in complex with NheB to the membrane, followed by the subsequent binding of free NheB to this complex, forming intermediate small transmembrane pores (Heilkenbrinker et al., 2013; Zhu et al., 2016). The ratio of NheB: NheC is also important for ensuring the optimum binding of NheA to the membrane, as increasing the concentration of NheC reduces the presence of NheA at the membrane (Heilkenbrinker et al., 2013; Lindbäck et al., 2010). For the Nhe system, it is clear optimum pore formation occurs with reduced NheC. The 109 nucleotide non-coding region between *nheB* and *nheC* contains an inverted repeat which results in reduced translation of NheC, Granum et al. (1999). However similar regulatory elements are yet to be identified in other ClyA family tripartite PFT operons.

Following the monoclonal antibody studies of NheA, a detailed assembly mechanism was proposed for the Nhe toxin. NheA is present in solution as single molecules, while NheB and NheC form complexes with some free NheB being present. The BC complexes bind to the membrane and recruit further NheB forming small permeable pro-pores, finally, NheA binds to NheB in the pro-pore and undergoes a conformational change centring around the movement of the C-terminus leading to full pore formation and death of the cell (Figure 1.14B)(Didier et al., 2016).

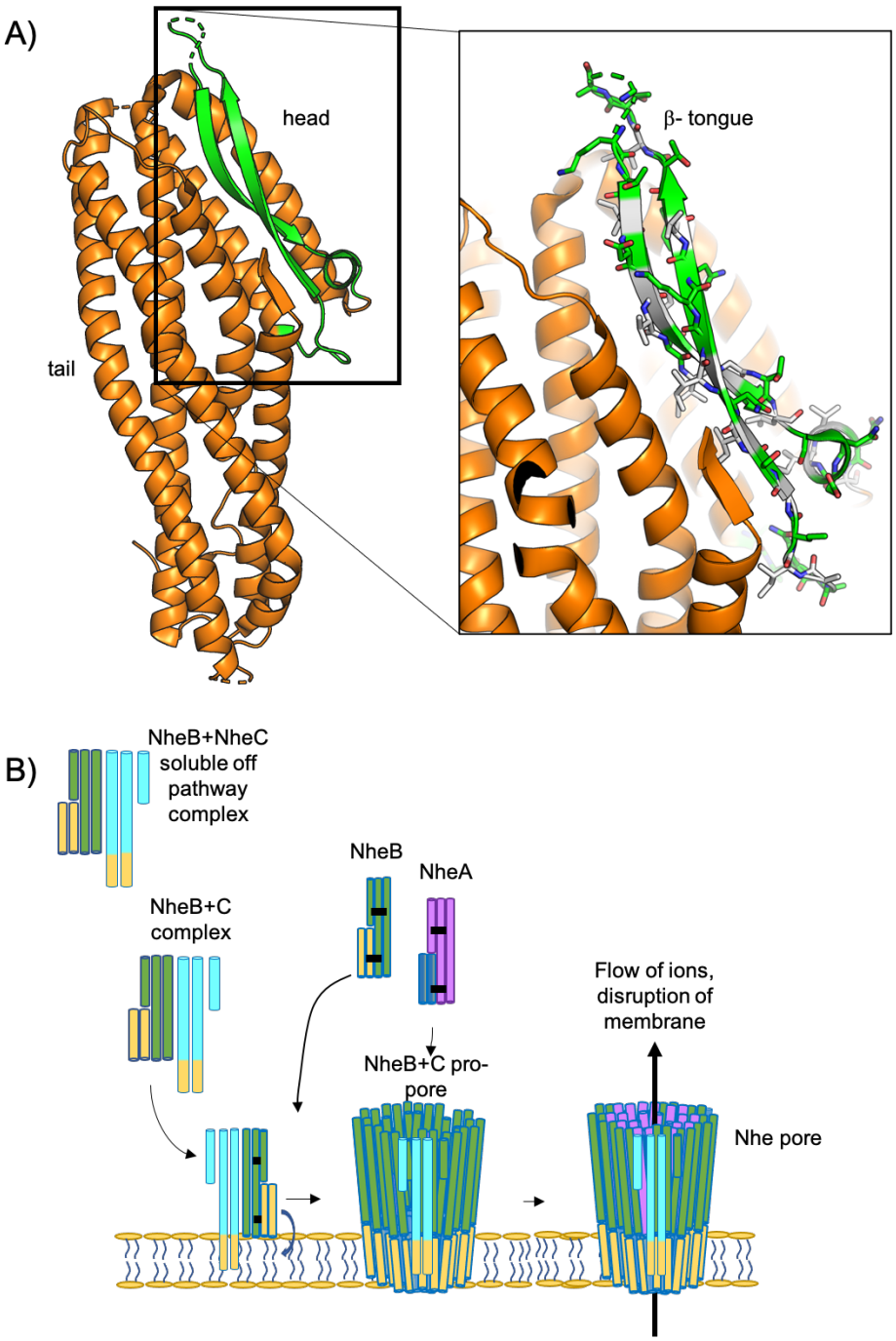


Figure 1.14 **Structure of NheA and proposed mechanism for pore assembly.** A) Crystal structure of NheA shows a large β- tongue motif, with two amphipathic β- strands. Green shows proposed membrane-spanning domains, while hydrophobic residues are shown in white. B) Proposed Nhe assembly mechanism. NheB and NheC form complexes in solution before oligomerisation and pro-pore assembly in the membrane, followed by binding of NheA to NheB and a conformational change leading to membrane disruption.

1.4.4 The Ahl toxin from *A. hydrophila*

More recently, work by Jason Wilson at the University of Sheffield has identified other species of bacteria that contain possible ClyA-family tripartite α-PFTs (Wilson, 2016). The Ahl toxin from *A. hydrophila*, a pathogen which causes fatal haemorrhagic septicaemia in farm fish (Hatha et al., 2005), contained three proteins AhlA, AhlB and AhlC, which although share very low sequence identity (<40 %) to Nhe shared almost identical hydrophobicity plots and were all found in a single operon (Figure 1.15A). The three Ahl proteins were purified and haemolytic assays showed all three were needed for lysis (Figure 1.15B) and further to this pores were observed by negative stain electron microscopy when protein was incubated with liposomes, confirming that the Ahl toxin was indeed a tripartite α-PFT(Wilson, 2016). The Structure of AhlC was solved, showing a tetramer of AhlC in the extended pore conformation observed in ClyA, YaxAB and XaxAB pores, with the hydrophobic heads hidden within the tetramer rather than in a β-tongue (Figure 1.15C)(Wilson, 2016). Analysis of the self-rotation function (SRF) of data collected from crystals of AhlB showed a 10-fold symmetry, suggesting that AhlB alone may be able to form a pore structure (Figure 1.15C) (Wilson, 2016). Both of these results show new features not previously observed in the Nhe or Hbl toxins.

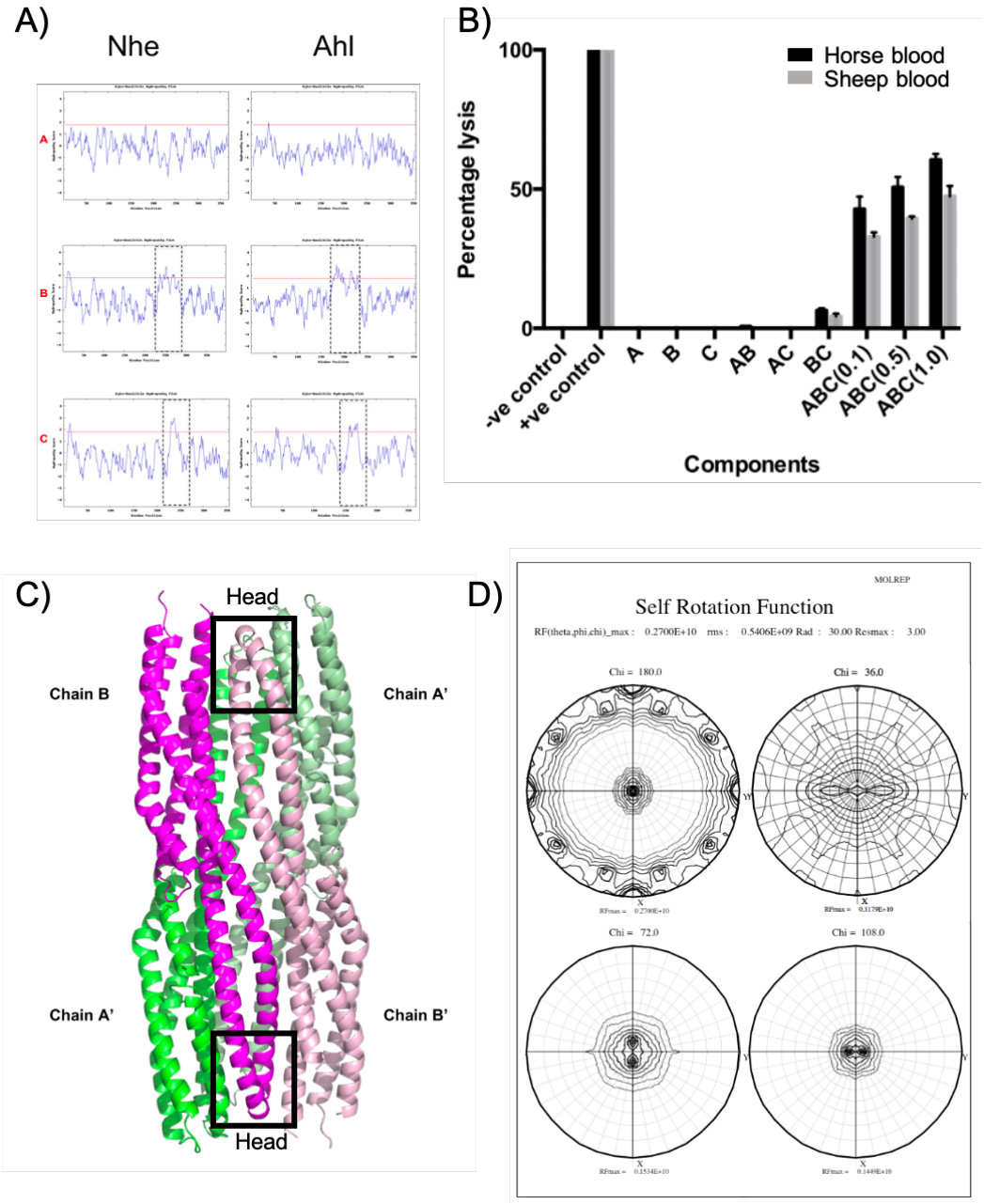


Figure 1.15 **The Ahl toxin from *Aeromonas hydrophila.*** A) Hydrophobicity plots of Nhe and Ahl A, AhlB and AhlC (top to bottom), dashed box shows hydrophobic regions. B) haemolytic assays showing % lysis for each Ahl protein incubated with erythrocytes alone and in combination. C) Crystal structure of AhlC, shows a tetrameric assembly with each AhlC chain in an extended conformation, with the hydrophobic heads hidden within the tetramer. D) Self rotation function for AhlB, showing sections for Chi= 180°, Chi = 36°, Chi = 72°, and Chi = 108°, generated with a radius of integration of 30 Å, and a resolution cut-off of 3 Å. In the Chi = 180° section, peaks are present at omega = 0 and phi intervals of 36 °. All figures were adapted with permission from (Wilson, 2016)

1.5 Targeting PFTs for Antimicrobial therapy

As shown in the previous sections PFTs play an important role in disrupting the immune response and early colonisation, during bacterial infection. Key examples being LLO which is needed for escape from the phagosome by *Listeria*, PLY, which disrupts the alveolar-capillary juncture allowing *Streptococcus pneumoniae* to spread ultimately into the bloodstream, and Nhe and Hbl, from *B.cereus*, which destroy epithelial cells and induce septic shock (Gonzalez et al., 2008; Mathur et al., 2019). Targeting these toxins for antibacterial therapy will therefore likely slow or even prevent early colonisation and spread giving the immune system chance to clear the infection (Escajadillo and Nizet, 2018).

1.6 Biotechnological applications of PFTs

Due to the abundance of PFTs in nature and their ability to create a channel across a membrane, as well as their variability in terms of different size, conductance and mechanisms of pore formation, they offer an exciting target for biotechnological applications. Almost all biotechnological applications for PFTs exploit the ability to apply a voltage across the pore and then measure the change in current as different molecules are passed through the pore allowing characterisation of a range of different molecules (Wang et al., 2018).

As discussed in section 1.3.5, the diameter of different PFTs varies significantly and also determines what can pass through them, whether this is proteins or ions. This difference in size is taken advantage of when selecting a PFT for a specific biotechnological application. When sequencing DNA or polypeptides the diameter of the transmembrane pore lumen must be small enough to tightly fit a single nucleotide or amino acid allowing the small change in current to be measured as each residue passes through. As described below α-haemolysin and FraC have the perfect diameter pores for DNA sequencing and Peptide sequencing respectively (Franceschini et al., 2013; Huang et al., 2019, 2017). For measuring

single-molecule kinetics (e.g. enzyme kinetics), the pore must have a large central cavity able to be easily mutated to allow for successful trapping of single proteins of interest, the large extracellular domain of the ClyA family proteins make them ideal for this role. The cylindrical shape of ClyA also makes it a good candidate for biotechnology applications as the current through a cylinder can be easily modelled (Willems et al., 2019). Finally, in all cases, the PFT chosen must readily assemble into a stable pore (Maglia et al., 2010).

Since it was established that ss-DNA could be translocated through a α-haemolysin pore (Franceschini et al., 2013), PFTs are now widely established as a method by which to rapidly sequence DNA on a small scale at low cost. Through engineering the lumen of the pore, as well as coupling of the pore to secondary DNA binding proteins (e.g. DNA polymerase), DNA can be threaded through the pore at 100-200bp per second, slow enough for current detection technologies to easily measure the distinct current fingerprint of each nucleotide (Wang et al., 2018).

More recently, focus has turned to exploiting PFTs for detection and analysis of proteins as well as peptide sequencing. It has been shown that by altering the charge of the lumen of ClyA, different proteins can be trapped within the pore and the length of time the protein is trapped can be increased. This will allow enzymatic reactions on a single molecule level to be observed, by measuring changes in current through the pore as conformational changes occur in the trapped enzyme (Huang et al., 2017). Furthermore, aerolysin pores have been used to measure single amino acid differences in short polypeptides, building a basis for the possibility of peptide sequencing using PFTs (Ouldali et al., 2020). FraC nanopores have also been successfully engineered to allow the discrimination of peptides and proteins from 25kDa down to 1.2kDa by measuring nanopore currents, and were able to measure differences in the presence or absence of a single tryptophan. If the speed of translocation of peptides can be controlled, as with DNA sequencing, it is feasible to believe that FraC

nanopores or aerolysin could be used for peptide sequencing in the near future (Piguet et al., 2018).

1.7 Aims of this thesis

Although a large amount of biochemical and kinetic data has been obtained on the two tripartite ClyA family α-PFTs identified in *B. cereus*, there is only limited structural knowledge of the components, and thus a full understanding of these toxins remains elusive. Both Nhe and Hbl are important virulence factors in *B. cereus*, thus better understanding of the mechanisms of this family will be important in the development of new antimicrobials, as well as possibly providing new candidates for the development of biotechnology applications. Previous work in the group by Jason Wilson, described above, has already identified a tripartite haemolytic α-PFT homologue of Nhe, Ahl, in the Gram negative bacteria *A. hydrophila*. Work described in this thesis will begin with building on the previous work of Jason Wilson to further characterise the Ahl toxin and go on to expand the tripartite family through the identification and study of other homologues, in the hope that they would be more amenable to structural characterisation. Gaining structures of these proteins in both their soluble form by X-ray crystallography, and as pores using a combination of X-ray crystallography and electron microscopy will help to rationalise the biochemical results so far obtained for both the Hbl and Nhe toxins, and allow the generation of a detailed mechanism by which these tripartite toxins undergo the conformational change from soluble to pore. Furthermore, the ClyA family of α-PFTs appear to adopt different assembly pathways and although the family members share similar structural features, the family has developed a diverse range of tools for pore assembly. Biochemical analysis of any potentially new tripartite family members will help better understand this diversity.

The aims of this project are:

- Expand the structural knowledge of the tripartite family through X-ray crystallography and electron microscopy.
- Suggest possible roles of each protein in the tripartite toxins.
- Characterise the mechanisms of pore assembly adopted by new members of the tripartite family.
- Compare structures and mechanisms to identify requirements for pore formation in the ClyA family of α-PFTs.

Chapter 2

Materials and Methods

2.1 General methods and recipes

2.1.1 Antibiotics

Ampicillin was used as the antibiotic in this study for the selection of constructs containing the ampicillin resistance marker present in pET21a vectors. Ampicillin was prepared by dissolving 0.1 g in 1 ml of MilliQ H_2O to make a 1000 x stock of 100 mg/ml

2.1.2 Lysogeny Broth (LB) medium

LB Media (Miller, 1972) was prepared by adding 10 g tryptone, 10 g NaCl, and 5 g yeast extract and to 1 L deionised water. LB-media was then immediately autoclaved at 121°C and stored at room temperature until needed.

2.1.3 LB Agar

LB agar was made by adding 15 g bacterial agar to 1 L LB media. The mixture was then immediately sterilised by autoclave at 121°C resulting in solubilisation of the agar. LB-agar was stored at room temperature until needed. To use, LB-agar was melted in the microwave

on medium power and 25 ml poured into petri dishes per single plate. The plates were supplemented with 100 μg/ml ampicillin, and, if needed, 1 % w/v glucose to suppress leaky expression. Plates were used immediately once set.

2.1.4 SOC media

Super optimal broth with catabolite repression (SOC) media (Hanahan, 1983) was purchased from New England Biotech (NEB) and contained 20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 0.186 g/L KCl, 0.952 g/L, MgCl2, 2.47 g/L MgSO4 and 3.6 g/L glucose.

2.1.5 Strains and plasmid vector information

The pET21a vector was used throughout this study, and allowed tightly controlled expression which could be induced by the addition of isopropyl -D-1-thiogalactopyranoside (IPTG). Details of the vector are shown below (Figure 2.1).

2.1.6 Gene cloning

Open reading frames for genes in this study were synthesised and cloned into expression vectors, and codon optimised for expression in *E.coli* by Geneseq so as to contain a C-terminal His_6 tag. Restriction enzymes NdeI and XhoI were used throughout.

2.1.7 Transformation of *E. coli* with plasmid vector

Plasmid vectors were transformed into either *E. coli* DH5 (NEB) for cloning or BL21 (DE3) (NEB) for overexpression. These strains were purchased as chemically competent cells and stored in glycerol at -80 ^{o}C. Cells were thawed on ice before the addition of 50-100 ng of plasmid. Cells and plasmid were mixed gently then incubated on ice for 30 minutes. Cells were heat-shocked at 42 ^{o}C for 1 minute then incubated on ice for 2 minutes. 700 μl of SOC

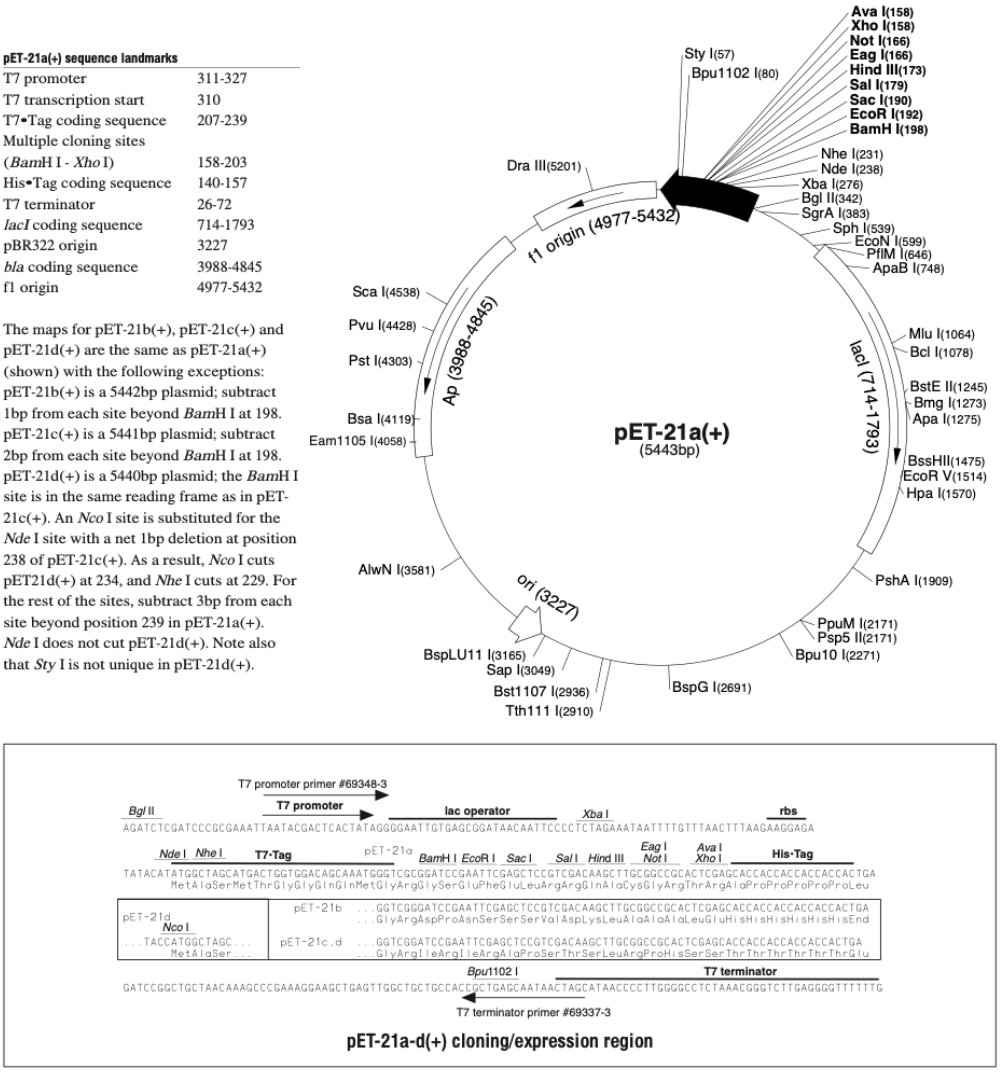


Figure 2.1 Details on the pET-21a vector taken from Novagen

was then added to the cells before incubating at 37 °C with shaking at 220 rpm for 1 hour. Finally, cells were plated onto LB-agar and 100 μg/ml ampicillin and incubated at 37 °C overnight.

Plasmid propagation, purification, and sequencing

For transformation experiments, plasmids needed to be at high concentrations and high purity. Plasmid propagation was carried out by transformation into *E. coli* DH5 (NEB) cells as described in Section 2 above and grown overnight at 37 °C in 5ml LB media supplemented with antibiotics. Cells were harvested and plasmid was purified using Monarch Plasmid Miniprep Kit (NEB) which uses alkaline lysis to extract the plasmid DNA (protocol is provided in the kit and available online). For sequencing, 5 μl of 60 ng μl^{-1} plasmid DNA was mixed with 5 μl of either forward or reverse T7 primer and sent for Sanger sequencing with GATC Biotech (Germany).

2.1.8 Glycerol stocks

For long term storage of transformed *E. coli* cells, glycerol stocks were made. *E. coli* cells in LB broth were mixed with glycerol at a 1:1 ratio in a total volume of 500 l, to give a 50 % v/v glycerol stock, before freezing at -80 °C.

2.1.9 SDS-PAGE

SDS-PAGE was used to determine the level of protein during protein overexpression and purity of protein during purification. SDS electrophoresis separates proteins based on molecular weight, with a molecular weight marker as reference. Gels were made by hand to a 1 mm thickness in a BioRad Mini-PROTEAN casting glass plate using stacking and resolving gels as described below. A comb was put in the top of the gel to produce 15 wells.

Table 2.1 SDS-PAGE recipe. *APS (Ammonium persulphate) – 10 % solution prepared by adding 0.1 g APS to 1 ml MilliQ H_2O

12 % resolving	6 % Stacking
2.5 ml 30 % acrylamide	0.75 ml 30 % acrylamide
2.35 ml Tris pH 8.8	0.47 ml Tris pH6.8
1.28 ml water	2.46 ml water
62.5 µl 10 % SDS	37.5 µl 10 % SDS
10 µl TEMED	5 µl TEMED
62.5 µl 10 % APS*	5 µl 10 % APS*

2.1.10 Preparation of samples for SDS-PAGE

15-20 μg of protein was mixed with SDS and reducing agent (Novex NuPage reducing agent) so as to contain 1 % of each. The solution was then either incubated at 95 ^{o}C for 10 minutes before loading or loaded directly onto the gel. The stacking stage of the gel was run at 60 V for 10 minutes, followed by the resolving stage, at 200 V for 45 minutes.

The resultant gel was stained using coomassie stain (0.1 % Coomassie blue, 1 volume methanol, 1 volume glacial acetic acid, 1 volume deionised water) for 20 minutes, followed by destaining solution (10 % (v/v) glacial acetic acid, 10 % (v/v) methanol) for 2 hours.

2.1.11 Measuring protein concentration

Protein concentration was measured either by Bradford assay or absorbance at 280 nm. For the Bradford assay 0.2 ml BioRad Bradford reagent (Bradford, 1976), was added to the protein sample (volume dependent on the protein being analysed) and the total volume made

up to 1ml with MilliQ H_2O in a plastic cuvette. The contents of the cuvette were mixed by inversion and the absorbance reading at a wavelength of 595 nm was taken (OD_{595}). Protein concentration was calculated using the following formula:

$$Concentration\ (mg\ ml^{-1}) = \frac{OD_{595}\ x\ 15}{volume\ of\ sample\ (\mu l)}$$

Protein concentration was also calculated using the extinction coefficient for each protein (where sufficient aromatic residues were present) and absorbance at 280 nm (OD_{280}). Extinction coefficients were calculated from the amino acid sequence using ProtParam software (Gasteiger et al., 2005). Absorbance measurements were carried out using an Implen Nanophotometer 7122 nanodrop, so minimal sample was used. A lid attachment was used to alter the path length by lid factor. Protein concentration was calculated using the equation given below:

$$Concentration\ (M) = \frac{OD_{280}\ x\ lid\ factor}{extinction\ coefficient\ (M^{-1}\ cm^{-1})}$$

2.1.12 Lipid stock preparation

Lipid stocks were prepared from *E. coli* total lipid extract (Avanti Polar Lipids). 100 mg lyophilised total lipid extract was resuspended in 1 ml 2:1 chloroform: methanol solution, to give 100 mg/ml lipid stock solution, then stored at -20 ^{o}C. The composition of *E. coli* total lipid extract is listed below.

- PE 57.5 %
- PG 15.1 %
- CA 9.8 %
- Unknown 17.6 %

2.1.13 Detergent screens

2 mg/ml of purified protein in 50 mM tris pH8, 0.2 M NaCl was incubated with 25 detergents (Hampton research detergent screen) at 5x CMC and 10x CMC for 4 hours at 37 ^{o}C. Each sample was automatically applied on to a Shodex protein KW-803 high-performance liquid chromatography (HPLC) column equilibrated with 50 mM Tris pH8 and 0.5 M NaCl, by a liquid handler. The column was then washed with 50 mM Tris, 0.2 M NaCl to elute the protein, which then passed through a UV detector and a chromatogram was produced for each sample. Two controls were also set up, one containing detergent incubated on its own and one with protein incubated on its own.

2.1.14 Liposome preparation

100 μl of *E. coli* total lipid extract stock solution was placed in a round bottom flask and solvent was evaporated off under a gentle nitrogen stream which was slowly increased as a film formed at the bottom of the flask. To ensure a thin film formed the flask was swirled while the nitrogen stream was applied, evaporation was complete when the flask no longer felt cold to the touch under the nitrogen stream.

The lipid film was flash-frozen by submerging the bottom of the flask in liquid nitrogen. The flask was then placed on a vacuum pump at 1.5 mBar for 2 hours to remove excess solvent.

1 ml of 10 mM phosphate buffer saline (PBS) pH 7.4 was added to the lipid film and then vortexed for 10 minutes to resuspend the lipids. Resuspended lipids were extruded through a mini extruder (Avanti polar lipids) containing a 0.1 um filter to produce liposomes. The lipids were syringed through the extruder between two Hamilton 1 ml syringes >10 times to form a uniform solution of unilamellar liposomes. The extruded liposomes were always taken from the second syringe as this should contain no pre-extruded contaminants.

Extruded liposomes were used immediately or stored at -20 ^{o}C until needed. Liposome quality was always checked by electron microscopy before use.

2.1.15 Proteoliposome preparation

Proteoliposomes were used for both electron microscopy and liposome float assays. The amount of protein described in each specific case was incubated at 37 ^{o}C for 1 hour with either 20 μg liposome for electron microscopy, or 200-300 μg of liposomes for lipid float assays. Proteoliposomes were then loaded immediately on to carbon grids or used in float assays.

2.2 Protein overexpression

2.2.1 Trial expression protocol

New constructs produced for this project required testing to see if they produced suitable quantities of protein for downstream processes.

5 ml LB with 100 μg/ml ampicillin was inoculated with a single colony of *E. coli* transformant and incubated overnight at 37 ^{o}C and 220 rpm shaking. 1 % v/v of the overnight suspension was used to inoculate 50 ml LB with 100 μg/ml ampicillin and incubated at 37 ^{o}C until an OD_{600} of 0.8 was reached then induced with 1 mM IPTG and incubated at test temperatures of 16, 25, and 37 ^{o}C for 5 hours. Protein expression was analysed by SDS-PAGE and the best temperature for soluble protein expression was chosen.

2.2.2 Large scale overexpression

Colonies were selected and grown in 5 ml LB with 100 μg/ml ampicillin for 8 hours. This suspension was used to inoculate 50 ml LB, which was then grown overnight at 37 ^{o}C and

220 rpm shaking. 4 x 500 ml flasks containing LB and 100 μg/ml ampicillin were inoculated with 10ml of the overnight culture. The 500ml flasks were incubated at 37 ^{o}C until an OD_{600} of 0.8 was reached, then induced with 1 mM IPTG and incubated at 37 ^{o}C for 5 hours, or left overnight at 25 ^{o}C or 16 ^{o}C.

Cells were harvested by centrifugation at 20,000 g for 15 minutes at 37 ^{o}C, supernatant was removed and the remaining pellet was frozen at - 80 ^{o}C until needed.

2.2.3 Minimal media selenomethionine overexpression

To produce proteins incorporating selenomethionine, the common method of using minimal media supplemented with amino acids to repress methionine biosynthesis was used in the overexpression step. Minimal media was made by dissolving the following components in deionised water:

Minimal media

10.5 g/L K_2HPO_4

1 g/L $(NH_4)SO_4$

4.5 g/L KH_2PO_4

0.5 g/L Tri-sodium citrate

5 g/L glycerol

0.5 g/L Adenine, Guanine, Thymine, Uracil

the resulting mixture was autoclaved at 121 ^{o}C. A mixture of amino acids was also made by dissolving the following components in deionised water:

Amino acid mixture

1 g/L $MgSO_4.7H_2O$

4 mg/L Thiamine

100 mg/ml L-lysine, L-phenylalanine, L-Threonine

50 mg/ml L-isoleucine, L-Leucine, L-Valine

The additive mixture was then stored at -20 ^{o}C until needed. For use, 10 ml of the amino acid mix was added to 500 ml of minimal media, followed by 40 mg/L of selenomethionine.

For selenomethionine overexpression, colonies of transformed BL21 cells were grown as described for LB expression until an OD_{600} 0.8 was reached. Cells were then spun down at 5000 g at 20 ^{o}C for 15 minutes. The resulting pellet was washed twice in minimal media before resuspending in 500ml of minimal media, supplemented with amino acid mixture and selenomethionine, and incubating at 37 ^{o}C for 1 hour, and the OD_{600} was used to monitor the cell recovery. So long as OD_{600} between 0.6 and 0.8 was achieved, the culture was then induced with 1 mM IPTG. Cells were harvested by centrifugation at 20,000 g for 15 minutes at 37 ^{o}C, supernatant was removed and the remaining pellet was frozen at -80 ^{o}C until needed.

2.3 Protein purification

Methods used to purify proteins followed one of two routes. Route one was used to exploit the His_6 tag using nickel affinity chromatography, followed by size exclusion gel filtration. Route two was used for proteins where a His_6 tag was not available and included ammonium sulphate precipitation, followed by anion exchange chromatography before gel filtration chromatography.

2.3.1 Preparation of cell-free extract

To recover the desired protein from the overexpression, cells were broken open using sonication, then cell debris was separated from the cell-free extract by centrifugation.

Cell paste was suspended in lysis buffer (50 mM Tris pH 8, 0.5 M NaCl) in a 1:8 ratio. The suspension was put into small containers on ice for sonication. Sonication was done in 2x20 s bursts at 15 kHz using the largest probe possible for the container. Samples were sonicated in rotation and kept on ice between bursts to allow for cooling. Lysate was centrifuged at 40000 g for 15 minutes at 4 ^{o}C and the cell-free extract collected for purification.

2.3.2 Nickel affinity chromatography

Immobilised metal ion affinity chromatography was used to purify His_6 tagged protein. This technique uses immobilising Ni^{2+} ions on a cross-linked agarose matrix for binding with poly His clusters of tagged protein. Binding is reversed by competitive replacement of the poly-His with increasing imidazole concentration, so allowing selective purification of tagged protein. Cell-free extract (CFE) in 50 mM Tris pH 8, 0.5 M NaCl was applied to a 5 ml Hi-trap Nickle column (GE Healthcare) mounted on an Akta Pure (GE Healthcare), which was then washed with buffer A (50 mM Tris pH 8, 0.5 M NaCl) until the UV trace returned to baseline. A 0-100 % gradient of buffer B (50 mM Tris pH 8, 0.5 M NaCl, and 0.5 M imidazole) over 50 ml was applied to elute protein. 2.5 ml fractions of elute were collected, fractions believed to contain protein were run on SDS-PAGE, those containing the correct protein were pooled and concentrated for crystallisation or gel filtration.

2.3.3 Ammonium sulphate cut

To determine the appropriate concentration of ammonium sulphate to use for precipitation of the desired protein, concentrations of 0.5-1.5 M ammonium sulphate were added to CFE for 10 minutes. Precipitated protein was pelleted by centrifugation at 70,000 g for 5 minutes at 4 ^{o}C. The supernatant was collected and the pellet was resolubilised in 50 mM Tris pH 8. Resolubilised pellet and supernatant were loaded on to SDS-PAGE to analyse the purity of

the desired protein. Resolubilised pellets that contained the desired protein were combined for further purification.

2.3.4 Ion exchange chromatography

All proteins have an isoelectric point (pI), at pH values above this point the protein has an overall negative charge and allows it to bind to a positively charged medium (anion exchange), below this value the protein would have a positive charge and bind a negatively charged medium (cation exchange). Based on the pI of the protein in this study a 5 ml Hi-trap DEAE sepharose Fast Flow column (GE Healthcare) equilibrated with 50 mM Tris pH 8, was used. Protein or CFE was loaded on to the column which was then washed with the same buffer as used in equilibration. The protein was then eluted from the column by changing the ionic strength using a gradient of 0 - 0.3 M NaCl. Fractions were collected and analysed on SDS-PAGE and those containing protein were pooled and depending on purity used for crystallisation and assays or applied to a gel filtration column for further purification and analysis.

2.3.5 Gel filtration

Gel filtration chromatography separates protein based on size and shape by using a porous matrix of spherical particles, usually crossed linked Dextran beads. Proteins smaller than the pores in the beads enter the beads as they move through the column, while large proteins do not and therefore travel through a smaller column volume and elute faster than smaller proteins.

The column was calibrated using Buffer A (0.5 M NaCl, 50 mM Tris pH 8) for purification or 10 mM PBS pH 7.4 for analytical purposes. Protein was loaded on to the column and eluted using buffer A. The elute was collected in fractions, with fractions believed to contain protein run on SDS-PAGE and then pooled for crystallisation or use in assays. Details of

columns and buffers used for specific proteins and analysis can be found in the relevant papers and appendices.

2.4 Biochemical analysis

2.4.1 Ultra-centrifugation of detergent solubilised protein

Ultracentrifugation was used to determine the composition of complexes formed by proteins in the presence of detergent. Large complexes > 400 kDa would pellet while smaller proteins would remain in solution.

10 μM of protein was incubated with 45 μM N-heptyl-thioglucopyranoside at a total volume of 100 μl for 1 hour at 37 ^{o}C. Samples were then spun for 30 minutes at 214,500 g in a Beckman Optima MAX ultracentrifuge at 4 ^{o}C. Supernatant was removed, and the pellet was resuspended in 50 mM Tris pH8 and 0.5 M NaCl. Both the pellet and supernatant were loaded onto an SDS-PAGE gel for analysis.

2.4.2 Haemolytic assay

Defibrinated horse blood was purchased from Thermo Scientific and stored at 4 ^{o}C. Blood was washed in 10 mM PBS pH 7.4 by diluting to 10 % w/v and centrifuged at 1500 g for 5 minutes. The supernatant was removed and the washing step repeated two more times. The final pellet of blood cells was resuspended to 0.25 – 0.5 % w/v in PBS. Each protein was buffer exchanged into 10 mM PBS pH 7.4 and concentration measured using absorbance at 280 nm.

Assays were set up in 0.2 ml Eppendorf PCR tubes with 200 μl 0.25 - 0.5 % w/v blood. Protein was aliquoted into the lid of the tube and mixed with blood by inversion once each component had been added. Reactions were then incubated at 37 ^{o}C on a blood wheel to mix throughout the experiment.

After incubation, the tubes were centrifuged at 1500 g for 5 minutes at 4 ^{o}C to pellet any remaining erythrocytes. Supernatant was removed onto an Implen Nanophotometer 7122 nanodrop and haemolysis was measured by the release of haemoglobin into the supernatant using absorbance at 542 nm. A positive control containing 0.2 ml 0.25 – 0.5 % w/v erythrocytes centrifuged at 1500 g for 5 minutes and resuspended in 1 ml MilliQ H_2O, followed by pipetting to mix and induce osmotic lysis. A negative control containing 0.2 ml erythrocytes 0.25-0.5 % w/v was also set up. Both were incubated at 37 ^{o}C.

2.4.3 Protein cross-linkage

Cross-linking was carried out in a two-step reaction. 25 % w/v Glutaraldehyde was added to 1.4 mg/ml of protein in PBS buffer then incubated at room temperature for 15 minutes. Protein solution was then buffer exchanged into 0.1 M MES pH 6 using a Zebaspin 7K desalting column (Sigma) before adding 40 % w/v 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and incubating for a further 30 minutes at room temperature. Finally, the protein was buffer exchanged into PBS using a Zebaspin 7K for assays and SDS-PAGE.

2.4.4 Liposome float assay

Freshly prepared proteoliposomes made by incubating 200-300 μg liposomes with 100 μg protein were used in liposome float assays. Negative controls were run in parallel and contained only protein. The total volume of proteoliposome was made up to 800 μl with 55 % w/v sucrose in PBS and transferred into Ultraclear ultracentrifuge tubes (Beckman Coulter). This was then overlaid with 3.8 ml 40 % w/v sucrose in PBS followed by 400 μl PBS. Samples were then centrifuged at 200,000 g for 4 hours in a Beckman Coulter SW 55 Ti rotor, 6 fractions of 100 μl were taken from the top of the tube 3.6 ml from the middle was discarded and 6 fractions were collected from the bottom. All fractions were run on SDS-PAGE for analysis.

2.5 Crystallographic methods

2.5.1 Preparation of protein for crystallisation

Protein was concentrated to between 6-15 mg/ml using a Vivaspin concentrator (Sartorius) with a molecular weight cut off smaller than that of the desired protein. The concentration was then altered depending on the result of crystallisation trials. Prior to crystallisation, proteins were buffer exchanged into crystallisation buffer, containing 10-200 mM NaCl and 50 mM Tris pH 8, using a Zebaspin desalting column (Sartorius).

2.5.2 Initial robot screening

Initial screens for crystallisation conditions were carried out using a TTP Labtech Mosquito LCP crystallisation robot in a 96 well plate. Each plate contained 96 individual experiments which consisted of a large well for the mother liquor and 2 small wells, one of which was used to contain a 1:1 mix of mother liquor to protein. The large well contained 50 μl of the crystallisation solution, while the small well contained 100 nl of crystallisation solution and 100 nl of protein. The plate was then sealed with crystal clear tape before being stored at 17 ^{o}C or 7 ^{o}C. During storage vapour-water equilibration between the large well and small well could occur, leading to the formation of protein crystals. Crystal growth was checked by observing the drops under a microscope. Initial crystal trials used the commercially available screens PACT, JCSG+, MPD and Proplex (Molecular Dimensions).

2.5.3 Crystallisation optimisation

Crystals from screening were further optimised to increase quality and size. Optimisation was carried out using 24 well plates and micro bridges, or sitting drop 96 well plates. The conditions in which the initial protein crystals grew were optimised by varying pH, salt concentration, and precipitate concentration. Some optimisations involved seeding the drops

with crushed crystals from previous trials. These were prepared by diluting crushed crystals 10,000 x and 1000 x in mother liquor. Horsehair was then dipped in the diluted seed solution and dragged through the optimisation drop in micro bridges.

Microbridge, 24 well plates

A 24 well plate was used to scale up the experiment and yield bigger crystals. 500 μl of crystallisation solution was made up in the well and a microbridge, which contained a small well, was placed in the large well. 1 μl of crystallisation solution and 1 μl of protein were placed in the small well and mixed by gently pipetting up and down before sealing with crystal clear tape. The plates were incubated at 17 ^{o}C and checked for crystals by observing drops under a microscope.

Sitting drop optimisation

96 well sitting drop plates were used to set up optimisation screens which allowed a much larger chemical space to be screened than with the 24 well plates, but with small volume drops. 50 μl of crystallisation solution was made up in large wells using a Formulatrix Formulator robot. 100 nl:100 nl mix of crystallisation solution to protein drops were made in the small wells using a TTP Labtech Mosquito LCP crystallisation robot. The plate was then sealed with crystal clear tape before storing at 17 ^{o}C, or 7 ^{o}C. Trays were checked for crystals by observing drops under a microscope.

2.5.4 Mounting crystals

To collect data, crystals were removed from the drop and mounted into loops that matched the size of the crystal. The drop is exposed under a microscope by removing the tape covering the well and the loop attached to a magnetic wand was used to loop a single crystal. The loop was then placed in liquid nitrogen to cool to 100 K. During cryo-cooling ice crystals can form

disrupting the protein crystal lattice, as well as this ice-rings can compromise the quality of data collected. To reduce the formation of ice crystals, protein crystals were either grown in conditions that act as cryoprotectants (e.g MPD, low molecular weight PEGs, ethylene glycol or glycerol) and so were placed immediately into liquid nitrogen, or cryoprotectants were made using 1 ml well solution with 25 % ethylene glycol or glycerol and crystals in loops were dipped in this solution before freezing. The well was then resealed with more crystal clear tape, and put back in the incubation room.

2.5.5 Crystal screening and data collection

Crystals were shipped to the Diamond Light Source, Oxfordshire, UK, for data collection using high-intensity X-rays. All testing and data collection was carried out at 100 K on the MX beamlines. MX beamlines, in general, were tunable wavelength and equipped with Pilatus or Eiger (DECTRIS) detectors, which allows for various collection strategies and rapid data collection, details of which are described below.

2.5.6 Testing

A standard beamline strategy was used to test protein crystals for diffraction as well as provide evidence of potential resolution data could be collected to. 5 images were collected with 0.2 s exposure, 0.2 ^{o} oscillation, 40 % transmission and a delta of 45 ^{o}. Imosflm (Battye et al., 2011) or EDNA (Leal et al., 2013) were used for the initial determination of space group and unit cell parameters.

2.5.7 Data collection

Once test images had shown that a protein crystal was of sufficient quality for data collection a data collection strategy was calculated. For a native dataset data was collected over a rotation range based on the estimated space group, to cover the reciprocal space required.

Typically an image would be collected over an oscillation of 0.1 ^{o} and an exposure of between 0.1-0.008 s, with a beam intensity of 50-100 % depending on the stability of the crystal. Typically data were collected at a wavelength of 0.09795 Å, this was dependent on the beamline the experiment was carried out on.

Experimental phasing used a SAD data collection strategy. Selenomethionine incorporated protein was used for experimental phasing. A fluorescence scan was carried out near the K edge of selenium and analysed using COOCH (Evans and Pettifer, 2001) to identify the peak wavelength as well as f' and f". The peak wavelength was used to collect high multiplicity data. To generate this data the beam was attenuated to between 5-30 % transmission to reduce radiation damage, and multiple isomorphous datasets would be collected usually from the same crystal to allow for images to be processed together. Typically a single SAD dataset would be collected over double the rotation of a native dataset. Details of each specific data collection can be found in the relevant appendices and manuscripts.

2.5.8 Data processing

Data used in this report was processed automatically at the Diamond Light Source using three different pipelines, Xia2 3dii, Xia2 Dials and FastDP (Winter, 2010; Winter and McAuley, 2011; Winter et al., 2018). Each pipeline is made up of programs that carry out spot detection, indexing of reflection, spacegroup determination, scaling, merging and data reduction. FastDP uses XDS (Kabsch, 2010) for spot detection and indexing, XDS and POINTLESS (Evans, 2006; Kabsch, 2010) for indexing and space group determination and finally XDS, POINTLESS and Scala (Winter and McAuley, 2011) for scaling, merging and data reduction. Xia2 DIALS uses DIALS (Winter et al., 2018) for spot finding and indexing while Xia2 3dii uses XDS (Kabsch, 2010), both utilise POINTLESS and Aimless (Evans and Murshudov, 2013) for downstream processing.

2.5.9 Experimental phasing using Single wavelength anomalous dispersion (SAD)

Selenomethionine incorporated crystals were used for experimental phasing throughout this thesis. The single-wavelength anomalous dispersion (SAD) method was used, which takes advantage of the anomalous scattering properties of selenium atoms at or near their absorption edge. Either the SHELX suite (Sheldrick, 2010) or CRANK2 (Skubák and Pannu, 2013) software package was used for phasing and production of an initial density map and model. In both cases, the process is guided by the number of heavy atoms expected based on the sequence (the number of selenium atoms is related to the number of methionines in the sequence), and the position of the heavy atom sites determined is used for calculation of initial phases. The resultant electron density map is used in model building. Details of software used can be found in the appendices.

2.5.10 Molecular replacement

Molecular replacement was used for phasing when a suitable homologue structure was available. PhaserMR (McCoy et al., 2007) was used for all cases. Rotation and translation of the Patterson map for the known structure onto the Patterson map of X-ray data from the unknown structure is used to place the search model (or multiple copies of the search model, depending on the predicted cell contents) in the asymmetric unit of the unknown structure. Phases generated from placing the search model are then used to generate an initial electron density map from target structure factors. A successful solution is determined by its TFZ (translation function) score, LLG (log-likelihood gain) and packing clashes, and analysis of resultant electron density.

2.5.11 Model building and refinement

After a starting map and model were produced using either experimental methods or molecular replacement, iterative model building and refinement was carried out using a combination of the following programs: COOT, Parrot, Buccaneer and Refmac5 (Adams et al., 2010; Emsley et al., 2010; Murshudov et al., 1997).

2.5.12 Validation and analysis

When all interpretable features in the electron density and difference Fourier map had been built, rebuilding was tentatively deemed complete and the model validated using the Molprobity server (Williams et al., 2018). This programme checks steric clashes, side-chain rotamers, Ramachandran angles, bond lengths and bond angles, and generates a validation report, and an overall score, based on other structures at a comparable resolution. Riding hydrogen atoms were added to the model and where there was clear evidence sidechains were flipped to satisfy hydrogen bonding. If the validation highlighted errors in the model, Coot and Refmac5 (Emsley et al., 2010; Murshudov et al., 1997) were used to further refine the model to improve the Molprobity score and the improved model was then compared with other comparable models in the PDB (Berman et al., 2002). All figures containing protein models were made using PyMOL (Schrödinger, LLC, 2015), while figures containing electron density were made using COOT (Adams et al., 2010; Emsley et al., 2010; Murshudov et al., 1997) or PyMOL and FFT (Read and Schierbeek, 1988; Schrödinger, LLC, 2015). Structure-based sequence alignments were carried out using Dali (Holm and Laakso, 2016), and calculated rmsD between models based on Cα alignment over equivalent residue regions. Finally, Docking predictions were done using the HADDOCK2.2 server (Van Zundert et al., 2016).

2.5.13 Homology model generation

All homology models presented in this thesis were generated using Phyre2 (Kelley et al., 2015). Details of Structures used as templates can be found in relevant appendices and papers.

2.6 Electron Microscopy Methods

2.6.1 Glow discharging and loading carbon grids

Glow discharging

Carbon grids were pre-prepared by Svetomir Tsokov by applying a thin carbon film to copper grids. Grids were then stored at room temperature for up to 4 weeks in a petri dish. The grids were glow discharged before the sample was applied to encourage absorption of the sample onto the grid. The required number of grids were placed on a parafilmed microscope slide inside a glow discharge chamber of a Cressington 208 Carbon Coater. A vacuum was then established inside the chamber, a small amount of air was allowed back in for plasma formation. Once the vacuum was stable a timer for glow discharge was set based on the age of the grid. Grids were used immediately after and kept out of direct sunlight.

Negative staining by uranyl formate

50 μl drops, 2 of distilled water and 1 drop of 1 %(w/v) uranyl formate were placed on parafilm. The grids were clamped between a pair of tweezers and 5 μl of the sample was applied to the carbon side of the grid. The sample was left to absorb for 1 minute before excess liquid was blotted off. The grid was then placed in the first drop of water, blotted and placed in the second drop of water. The grid was blotted again before finally placing in

uranyl formate for 20 seconds, then dried using an air vacuum until all liquid was removed. Dried carbon grids were put in a grid box and stored for up to 5 weeks.

Cryo sample loading

Quantifoil 300 mesh copper grids were glow discharged as described in Section 2.6.1 A Leica GP automatic plunge freezer was set up with fresh blotting paper and humidity was set to 80 %, liquid ethane was cooled in the ethane cup to 103 K. 5 μl sample was applied to either side of the grid and left to absorb for 60 seconds, the grid was then blotted for 1-10 seconds and then plunged immediately into liquid ethane. Grids were then stored in liquid nitrogen until used.

2.6.2 Data collection on a CM110 transmission electron microscope.

Data was collected on a Phillips CM100 with a tungsten filament set at HT 100 kV, and an attached Gatan MultiScan 794, 1 K x1 K CCD camera, using Gatan Digital Software. A reference image was taken at the beginning of data collection once CCD cooling had finished to determine and minimise background noise. Objective aperture was set to 1, and the spot size of the electron beam was set to 2.

Samples were observed first with a "search" setting, where images were collected with an exposure time of 0.1 s, image binning of 2 and full CCD. Once an area of interest had been found in search mode, a second focus was used with 0.4 s exposure, image binning of 1 and the central half of the CCD. Finally, images were captured in "record' mode with 1 s exposure, image binning of 1, and using the full CCD. If the sample quality was good, images were collected a 28500 x magnification and a range of defoci for downstream analysis.

2.6.3 Data collection on Technai Arctica, FEI Cryo-electron microscope

All imaging using Technai Arctica was carried out by Svet Tsokov. Grids stored in liquid nitrogen first had to be clipped with C-clips and C-clip rings for use in the Technai Arctica autoloader. Grids were loaded into the autoloader and then onto the microscope. Eucentric focus was set and the objective aperture was removed for screening, an atlas was collected at 165 x magnification to look at the quality of ice. For areas of good ice test exposures at 200 kV were collected at a range of magnifications and exposures.

2.6.4 Image processing and analysis

Image processing was done by Claudine Bisson and Jason Wilson. Micrographs were converted from tiff to MRC file format using the e2proc2d.py program in EMAN2 (Egelman et al., 2018). Negative stain micrographs were imported into Relion 3.1(He and Scheres, 2017; Scheres, 2012) or cisTEM (Egelman et al., 2018) for processing. Particles were either picked manually and extracted or picked using the "Ab-initio" particle picking mode in cisTEM (Egelman et al., 2018; Sigworth, 2004). 20 iterations of 2D or 25 iterations of 3D classification were carried out. Classes were visually inspected. Specific details of Image processing can be found in Appendices A and B.

Publications

The publications included in this thesis are listed below.

1. Identification and structural analysis of the tripartite α-pore forming toxin of *Aeromonas hydrophila* **Wilson et al. (2019)**

2. The A component (SmhA) of a tripartite pore forming toxin from *Serratia marcescens*: expression, purification, and crystallographic analysis . **Churchill-angus et al. (2020)**

3. Identification and characterisation of a tripartite α-pore forming toxin from *Serratia marcescens* **Submitted to *Scientific reports***

Chapter 3

Paper 1

3.1 Summary

The first paper presented in this thesis is an article describing the expansion of the ClyA family tripartite α-PFTs into Gram negative bacteria, focusing on the identification and characterisation of the Ahl toxin from *Aeromonas hydrophila*. Biochemical assays alongside the determination of structures of the soluble AhlB, AhlB pore, and soluble AhlC lead to a proposed mechanism of transition from soluble to pore form AhlB, and confirm that AhlB is the pore-forming component of the tripartite system. Further to this, the structure of AhlC shows it forms a tetramer to hide its hydrophobic domain, rather than a β-tongue, a mechanism previously unobserved in the family. This paper shows how, although Ahl shares many structural similarities with the ClyA-family of α-PFTs, it has a unique mechanism of action whereby off-pathway AhlB and AhlBC pores can form if AhlA is not present, also no inhibitory soluble complexes form, both features not seen in the other tripartite toxins Nhe or Hbl from *B. cereus* ((Jessberger et al., 2019; Zhu et al., 2016)). Finally, three requirements for activity by all pores in the ClyA-family are proposed and it is shown how each protein within the tripartite α-PFT fulfils one of these requirements. Details of work done that was

omitted, or was only briefly mentioned, from the results and discussion in this paper can be found in Appendix A (page 109)

3.2 Author contributions

P.J.B. conceived the project; P.J.B., A.M.C.-A., J.S.W. and C.B. designed the experiments; A.M.C.-A., J.S.W., S.E.S., S.B.T. and S.P.D. performed the experiments; P.J.B., A.M.C.-A., J.S.W., C.B., J.B.R. and P.A.B. interpreted the data and P.J.B., A.M.C.-A. and J.S.W. prepared the paper with input from all authors. Experimentally, I carried out the repeat protein production and purification of all three proteins, carried out haemolytic assays and liposome float assays, and did negative stain electron microscopy. In addition, I crystallised Selenomethionine AhlB pore, and AhlC HM, carried out X-ray data collection for both as well as subsequent data analysis and structure determination. Finally, I built and refined AhlB Se-met, AhlC HM and soluble AhlB, and carried out bioinformatics analysis and modelling of AhlA.

3.3 Paper 1

Paper 1 manuscript will start on the next page.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-10777-x OPEN

Identification and structural analysis of the tripartite α-pore forming toxin of *Aeromonas hydrophila*

Jason S. Wilson[1,4], Alicia M. Churchill-Angus[1,4], Simon P. Davies[1,2], Svetlana E. Sedelnikova[1], Svetomir B. Tzokov[1], John B. Rafferty[1], Per A. Bullough[1], Claudine Bisson[1,3] & Patrick J. Baker[1]

The alpha helical CytolysinA family of pore forming toxins (α-PFT) contains single, two, and three component members. Structures of the single component *Eschericia coli* ClyA and the two component *Yersinia enterolytica* YaxAB show both undergo conformational changes from soluble to pore forms, and oligomerization to produce the active pore. Here we identify tripartite α-PFTs in pathogenic Gram negative bacteria, including *Aeromonas hydrophila* (AhlABC). We show that the AhlABC toxin requires all three components for maximal cell lysis. We present structures of pore components which describe a bi-fold hinge mechanism for soluble to pore transition in AhlB and a contrasting tetrameric assembly employed by soluble AhlC to hide their hydrophobic membrane associated residues. We propose a model of pore assembly where the AhlC tetramer dissociates, binds a single membrane leaflet, recruits AhlB promoting soluble to pore transition, prior to AhlA binding to form the active hydrophilic lined pore.

[1] Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield, South Yorkshire S10 2TN, UK. [2] School of Biomedical Sciences, Faculty of Biological Sciences and Astbury Centre for Structural and Molecular Biology, University of Leeds, Leeds LS2 9JT, UK. [3] ISMB, Department of Biological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK. [4] These authors contributed equally: Jason Wilson, Alicia Churchill-Angus. Correspondence and requests for materials should be addressed to P.J.B. (email: p.baker@sheffield.ac.uk)

Pore forming toxins (PFTs) are critical components of the molecular offensive and defensive machinery of cells in virtually all kingdoms of life. In eukaryotes, PFTs are largely involved in the innate immune response[1], whilst in bacteria, PFTs form the major group of virulence factors in many pathogenic bacteria and constitute 30% of all toxins identified to date[2]. By puncturing holes in the membrane, bacterial PFTs facilitate the takeover of host resources, destroy the osmotic balance of the target cell[3] or insert a secondary intracellular toxin (such as with the anthrax binary toxin), through the pore formed in the membrane[4]. Many pathogenic bacteria, including highly antibiotic resistant strains, employ PFTs in their invasive arsenal. This makes them an attractive target for the development of virulence-targeted therapies that may have broad spectrum activity against human and other pathogens, a strategy that has potential to reduce the acquired resistance seen in conventional antimicrobial therapy[5].

All PFTs are produced as a soluble, generally monomeric form, which recognises the target cell by binding to specific receptors, thus concentrating the proteins to the membrane surface, before exposure of the transmembrane hydrophobic regions, oligomerization, and membrane insertion[6,7]. Depending on the secondary structure of the membrane component, PFTs can be classified as α-PFTs, using a ring of amphipathic helices to construct the pore or as β-PFTs, where a β-barrel is used to traverse the membrane[7].

Although a number of structures of α-PFTs have been determined by X-ray crystallography and electron microscopy[7], few are known in both pore and soluble form. One α-PFT where this transition is well understood is Cytolysin A (ClyA) from enteropathogenic *Escherichia coli* and *Salmonella* species[8,9]. In this protein a large conformational change occurs, involving 80% of residues, as the single soluble subunit rearranges to expose its membrane spanning hydrophobic residues (the β-tongue) and oligomerises to form the active pore of 6–13 subunits[9,10]. In this pore the hydrophobic β-tongue refolds to form the ends of two adjacent helices and these insert into a single leaflet of the target bilayer, with the amphipathic N-terminal helix providing the membrane spanning component and thus a hydrophilic lined pore. A number of α-PFTs have been identified which show structural similarity to ClyA but are composed of two protein components (bipartite PFTs). These include YaxAB from the human pathogenic bacteria *Yersinia enterolitica*[11,12], the PaxAB system of the insect pathogen *Photorhabdus luminescens*[11,13] and the XaxAB system from the insect pathogen *Xenorhabdus nematophila*[14,15]. The structures of the YaxAB and XaxAB pores and their component parts have been determined[11,15], and all show structural similarity to ClyA, but with low-sequence identity (approx 22% identity)[11]. In the active YaxAB pore the two proteins are arranged as ten heterodimers in a pore with C10 symmetry, while the XaxAB pore shows variability in pore protomer count, with 12–15 heterodimers observed. These bipartate pores have a pronounced funnel shape compared to the largely cylindrical ClyA pore due to a more extensive extracellular component[11,15], with the A component attaching to a single leaflet and the B component forming the membrane spanning region using two amphipathic helices. During pore formation the membrane spanning residues in YaxB and XaxB are exposed by a conformational change that is smaller in extent, but similar in effect, to that seen in ClyA[10,11,15].

Two α-PFT homologues to ClyA have also been identified in pathogenic strains of *Bacillus cereus*, but in these systems three proteins encoded on the same operon are required for maximal haemolytic activity (tripartite PFTs). One of these α-PFTs in *B. cereus* is formed from the Hbl-L_1, Hbl-L_2 and Hbl-B proteins[16–19], with the second comprised of proteins NheA, NheB and NheC[20,21]. In the latter system, NheC primes the host cell for the formation of ion permeable NheB/C pores[22], prior to the addition of NheA to construct the complete pore. Structures of the soluble NheA and HblB monomers are known, which both show similarity to ClyA, but again with less than 20% sequence identity[23,24]. As yet, the mode of assembly of this tripartite class of ClyA family α-PFTs is unknown.

Here, we identify tripartite ClyA family toxins in the genomes of medically and economically important pathogenic Gram-negative bacteria, including *Serratia marcescens*, a causative agent of nosocomial infections[25]; *Erwinia mallotivora*, responsible for papaya dieback disease[26], resulting in $58 million/annum damage to crops in Malaysia; and highly virulent strains of *Aeromonas hydrophila*, which cause fatal haemorrhagic septicaemia in a number of important farmed fish species with epidemics in US and Asian aquaculture causing multimillion dollar losses[27,28], as well as being a contributor to opportunistic human disease[29]. We show that the orthologue identified in *A. hydrophila* (AhlABC) is a haemolytic tripartite α-PFT, requiring all three components for maximal lysis, and we present the structures of the soluble forms of monomeric AhlB and tetrameric AhlC and the decameric pore form of AhlB. Our data presented here has allowed the construction of a model to show how a three component α-PFT system could assemble to create a lethal pore.

We show, guided by functional and mutagenic studies, that the AhlC tetramer first disassembles into monomers in order to form the initial membrane-binding event. We also show that AhlB undergoes a large conformational change similar to ClyA YaxB and XaxB[10,11,15], involving the beta tongue becoming an extended alpha helix and the tail domain helices sliding relative to one another as AhlB assembles into a hydrophobic pore. Secondary structure modelling of the predicted head domain of AhlA shows that similar conformational changes would expose an amphipathic helix, which on assembly of the AhlABC pore could provide a hydrophilic lining to the pore as seen in the other family members ClyA, YaxAB and XaxAB[10,11,15] .

Results

***A. hydrophila* has a tripartite α-PFT**. Although single and two component α-PFTs have been identified, the only known three component ClyA family members are the tripartite NheABC and $HblL_1L_2B$ systems of *B. cereus*. We thus instigated a bioinformatics search based on both the sequence and structure of the NheA, NheB and NheC proteins to discover further tripartite α-PFTs. The hypothetical protein translated from the AXH33180.1 gene of *A. hydrophila* species AL09-71, was identified with 24% sequence identity and 46% similarity, as defined by TCoffee[30], to NheB. The proteins coded by the upstream and downstream adjacent genes to AXH33180.1 in the *A. hydrophila* genome (AHX33179.1 and AHX33181.1, respectively) had only 6% and 9% sequence identity to NheA and NheC, yet their sequence similarity was 39% and 37%, respectively. Furthermore, hydropathy plots of these three *A. hydrophila* gene products showed a very similar pattern of hydrophobic regions as those seen in the NheABC tripartite toxin (Supplementary Fig. 1). The AHX33180.1 gene product contained a possible membrane-spanning region of 63 hydrophobic residues, similar in length and position to the 54 residue predicted membrane spanning region of NheB. Similarly, the AHX33181.1 gene product had a stretch of 23 hydrophobic residues in the same place in the sequence as seen for 22 hydrophobic residues in NheC. In addition, the sequence of the AHX33179.1 gene product was devoid of significant stretches of hydrophobic residues, a pattern also observed in NheA (Supplementary Fig. 1). These three *A. hydrophila* proteins were thus provisionally identified as a tripartite α-PFT and named AhlA, AhlB and AhlC. The AhlA and AhlB sequences

have since been annotated in NCBI as having a HBL type fold (PFAM05791).

To expand the family further, the three genes identified in *A. hydrophila* were used to identify orthologues in other Gram-negative bacteria. A total of 25 gene products with homology to the AhlA and 36 gene products with homology to AhlC were identified with E values < 0.01, including the known bipartite Gram-negative α-PFT components YaxA, PaxA and XaxA, whereas over fifty gene products were identified with homology to AhlB (Supplementary Data 1) suggesting that numerous α-PFT systems may exist in Gram-negative bacteria. Seven examples of full tripartite α-PFT systems were identified, including those in three human pathogens (*Serratia marcescens, Serratia liquefaciens* and *Spirosoma fluviale*), and two plant pathogens (*Erwinia mallotivora* and *Chromobacterium piscinae*) (Supplementary Table 1).

All three AHL components are required for maximum AHL lysis. Having identified that the AhlABC system of *A. hydrophila* could potentially be a tripartite PFT, each component was cloned, expressed and purified and used in haemolytic assays with horse erythrocytes (Fig. 1a, b). In isolation, each of the Ahl components exhibited no haemolytic activity. When tested in 1:1 binary combinations the AhlB/AhlC mixture showed 50% of the lysis observed when cells were lysed by osmotic shock with water, whereas the AhlA/AhlB and AhlA/AhlC mixtures exhibited no activity. Size exclusion chromatography of AhlB mixed with an excess of AhlC in aqueous solution did not show any AhlB/AhlC association (Supplementary Figure 2), indicating that any complex formation of these two components of *A. hydrophila* only occurs on the membrane. This is in contrast to both YaxAB and XaxAB where the A and B components are seen to form a complex before binding the membrane[11,15] and also for the *B. cereus* tripartite PFT where equimolar amounts of NheB and NheC produced an inhibitory complex in solution that prevented cell lysis[20,31]. A 1:1:1 mixture of all three Ahl components showed 95% of the lysis of the positive control, indicating that all three components of the AhlA/AhlB/AhlC toxin system are required for maximal lysis.

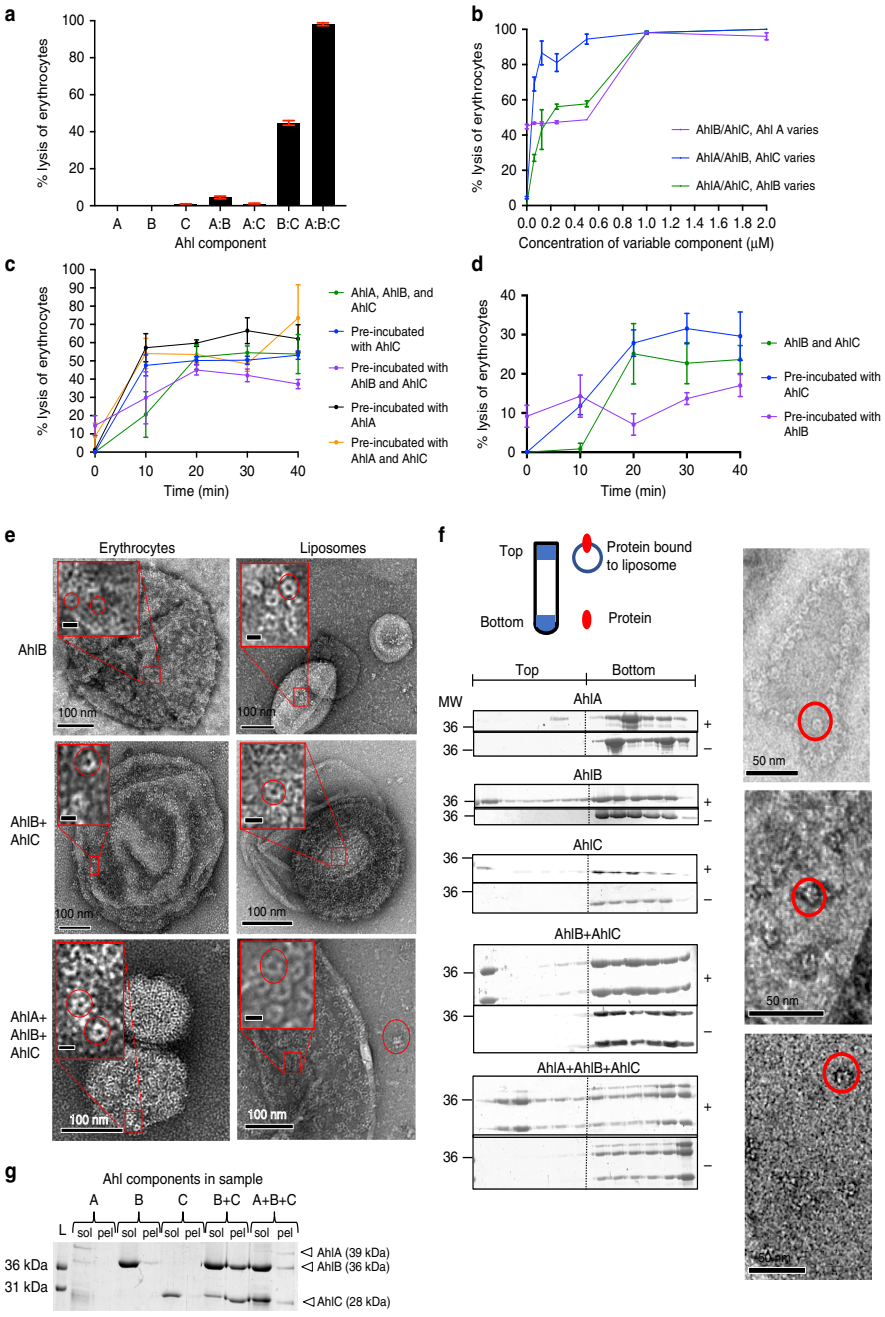
To elucidate the effect of the concentration of each protein on lytic activity, the ratio of individual proteins was varied within a mixture of AhlA, AhlB and AhlC by serial dilutions (Fig. 1b). When AhlA was diluted by one-half, lytic activity was reduced by 50%, to a level similar to that seen for AhlB and AhlC. However, when AhlC was diluted eightfold in the AhlA/AhlB/AhlC mixture 80% of the haemolytic activity remained. In contrast, an increasing reduction in lytic activity was seen across the dilution series when AhlB was diluted in the AhlA/AhlB/AhlC mixture. These results indicate that maximal lysis requires equal amounts of AhlA, AhlB and AhlC, and decreasing the concentration of any one of these components reduces lytic activity, with the most dramatic effects seen for AhlB and AhlA. Furthermore, when all three proteins are present at equal concentrations or increased above 1:1:1 no inhibitory effects are seen, in contrast to the *B. cereus* NheABC toxin where concentrations of NheC above a NheA:NheB:NheC ratio of 10:10:1 are inhibitory[20].

Pre-incubation of erythrocytes with AhlC promotes rapid lysis. To determine the binding order of each protein, time course assays were run using erythrocytes incubated with one or two of the Ahl components before addition of the other components (Fig. 1c, d). When AhlA/AhlB/AhlC, or AhlB/AhlC, were incubated together with erythrocytes there was a lag time of 20 min before lysis. However, if the cells were pre-incubated with AhlC before addition of AhlA and AhlB, the lag time was removed (or reduced if just AhlB was added) indicating that AhlC may prime the membrane for pore assembly. In contrast, cells pre-incubated with AhlB prior to the addition of AhlC showed reduced lysis compared to cells incubated with AhlB and AhlC together, suggesting AhlB is inhibitory in the presence of a membrane, similar to that observed in the Nhe PFT, where pre-incubation of NheB with vero cells prevented formation of a functional complex[31]. Cells pre-incubated with AhlB and AhlC, prior to the addition of AhlA, showed a lag time, with lysis not reaching that of cells incubated with AhlA/AhlB/AhlC together, whereas removal of the lag time and maximal lysis occurred for erythrocytes pre-incubated with AhlA, prior to addition of AhlB and AhlC, showing that addition of AhlA to a pre-formed AhlB/AhlC pore is slower than if the full tripartite pore can assemble with all three components present. The end points of these time course assays were viewed by negative stain electron microscopy. Pores could only be seen in erythrocyte membranes incubated with AhlB, AhlB/AhlC or AhlA/AhlB/AhlC (Fig. 1e). The diameter of the AhlB pores was estimated from 100 particles to be 10 nm (s.d. = 2 nm), with the AhlB/AhlC and AhlA/AhlB/AhlC pores appearing significantly larger with a diameter of 16 nm (n = 100, s.d. = 2 nm), a similar size to that of the widest head region of the YaxAB and XaxAB pores[11,15].

The interaction between the Ahl components with lipid bilayers was investigated using both liposome float assays and ultracentrifugation in the presence of detergent to separate large assemblies, (>400 kDa) from soluble protein components (Fig. 1f, g). AhlA, AhlB and AhlC together bound to the liposomes with negative stain EM images showing that pores had been formed. Incubation of all three Ahl components with detergent showed they were also present in the large assemblies isolated by ultracentrifugation, with pores identified by negative stain EM (Supplementary Fig. 3). Similarly, pores and large assemblies could be identified for both AhlB and AhlB/AhlC in both liposomes and by ultracentrifugation. No pores were seen for AhlA/AhlC together, or separately, despite both these components binding to liposomes and no protein was present in the respective ultracentrifugation pellets, which implies that the pores observed for the AhlA/AhlB/AhlC mixture contained all three PFT components, with equal amounts present despite AhlB and AhlC being in excess. As no pores could be identified in negative stain EM images of the float assay control fractions without liposomes (Supplementary Fig. 3), we concluded that large complexes of both AhlB/AhlC and AhlA/AhlB/AhlC can only be formed in the presence of detergent or lipid bilayers, with a preference for the AhlABC complexes and that large oligomers of AhlB can be formed under the same conditions.

Negative stain EM showed that when AhlC was incubated together with AhlB in liposome or erythrocyte preparations, there was a dramatic increase in the number of pore-like structures in individual liposomes, compared to the situation with AhlB alone, where only a few pores were seen (Fig. 1e), indicating that AhlC modulates AhlB association with lipid. As sufficient protein was present for pores to be present in all liposomes it is interesting to note that some liposomes were saturated with AhlBC pores and others remained free of pores (Supplementary Fig. 3), suggesting cooperativity in pore formation. Indeed, when AhlA, AhlB and AhlC were incubated together all liposomes were saturated, a situation also observed for ClyA[32].

Structure of the soluble form of AhlB. The structure of the soluble form of AhlB was determined to 2.3 Å resolution (Table 1, Supplementary Fig. 4a). AhlB folds into a compact five helical bundle structure (α1–α5), with an associated domain constructed from a mix of three alpha helices and three beta strands. The



overall structure of soluble AhlB is very similar to that seen in the soluble forms of ClyA (pdb:1QOY), MakA (pdb:6EZV), NheA (pdb: 4K1P) and HblB (pdb:2NRJ) (Fig. 2a). Within this compact soluble structure, residues G191–A222, identified from hydropathy plots as a hydrophobic domain that possibly inserts into the membrane (Supplementary Fig. 1), fold into two antiparallel β-strands (β1 and β2; the β-tongue) (Fig. 2b). These hydrophobic residues are thus shielded from solvent by packing against residues of helix α4 (E254–V270), α3a (residues V170–L186), α4a (residues G225–G239), the N-terminal helix (α1, residue G12–V34) and the C-terminus (residues T331–A359), which forms a short beta strand (β3) and helix (α5a). In this soluble

Fig. 1 Lytic activity of the AHL toxin against horse erythrocytes. **a** Percentage lysis when each protein was incubated with horse erythrocytes alone and in combination at 1:1 ratios for 1 h. **b** % Lysis against varying AhlA, AhlB or AhlC concentration, when all three proteins were incubated together for 1 h. The other two proteins were at a fixed concentration of 1 μM. **c** % lysis against time when erythrocytes were incubated with 500 nM AhlA, AhlB and AhlC (green), or pre-incubated with 500 nM AhlA (black), AhlC (blue), AhlB and AhlC (purple), or AhlA and AhlC (orange), for 1 h before addition of the remaining proteins at time 0. **d** % lysis against time when erythrocytes were incubated with 500 nM AhlB and AhlC (green), or pre-incubated with 500 nM AhlC (blue) or AhlB (purple) for 1 hour before addition of the second protein at time 0. **e** TEM negative stain images of AhlB, AhlBC and AhlABC pores in both erythrocytes and liposomes. Scale bars on magnified images represent 10 nm, and pores are highlighted in red circles. **f** Liposome floatation assay to determine lipid bilayer binding ability of each Ahl toxin component. Membrane binding ability of each AhlA, AhlB and AhlC alone and AhlB+AhlC, and AhlA+AhlB+AhlC (+) was assessed by SDS-PAGE analysis of the six top and six bottom fractions (left to right gel lanes with molecular weight indicated, kDa), together with a control without liposomes (−). A schematic of the ultracentrifuge tube shows the location of the top and bottom fractions and in which fractions liposomes and protein are expected. TEM negative stain EM images are shown of top fractions of AhlB (top), AhlBC (centre) and AhlABC (bottom), with pores circled in red. **g** SDS-PAGE of soluble (sol) and pelleted (pel) fractions from ultracentrifugation of AhlA, AhlB and AhlC with detergent, alone and in combination. Invitrogen Mark12 ladder (L) is labelled in the first lane. All assays were carried out in triplicate ($n = 3$) with error bars corresponding to the mean ± standard deviation. Source data are provided as a Source Data file

Table 1 X-ray data collection and refinement statistics for AhlB structures

	AhlB soluble PDB: 6GRK	AhlB pore SeMet PDB: 6GRJ	AhlB pore Form 2 PDB: 6H2F
Data collection			
Beamline	I03	I03	I02
Wavelength (Å)	0.9794	0.9794	0.9795
Space group	C2	C2	$C222_1$
Cell dimensions			
a, b, c (Å) =	133.5, 79.8, 111.0	363.6, 116.5, 217.4	117.7, 178.2, 485.6
α, β, γ(°) =	90, 90.2, 90	90, 118.0, 90	90, 90, 90
Molecules per asymmetric unit	3	10	10
Resolution (Å)[a]	2.33–58.35 (2.33–2.39)	2.94–107.58 (2.94–2.99)	2.55–39.55 (2.55–2.73)
Total reflections[a]	162245 (12843)	1150390 (52927)	1053972 (195594)
Unique reflections[a]	49042 (3660)	169188 (8317)	165659 (29762)
R_{merge}[a,b]	0.074 (0.50)	0.20 (1.8)	0.19 (0.63)
R_{pim}[a,c]	0.070 (0.46)	0.082 (0.76)	0.12 (0.40)
Mean I/σ(I)[a]	9.1 (2.0)	6.7 (1.0)	6.6 (2.7)
Completeness (%)[a]	98.1 (99.8)	99.1 (98.0)	100 (100)
Multiplicity[a]	3.3 (3.5)	6.8 (6.4)	6.4 (6.6)
Mid-slope		1.025	
dF/F		0.186	
Refinement			
No. of non-H atoms	7875	24386	24766
Rwork/Rfree	0.22/0.25	0.22/0.24	0.19/0.28
Average B factors (Å^2)	49	70	42
Bond length rmsd (Å)	0.0095	0.011	0.011
Bond angle rmsd (°)	1.41	1.57	1.57
Ramachandran favoured/allowed (%)	95.8/100	97.9/100	96.1/100

[a]Values in brackets are for data in the high-resolution shell
[b]$R_{merge} = \Sigma_{hkl} \Sigma_i | I_i - I_m| / \Sigma_{hkl} \Sigma_i I_i$
[c]$R_{pim} = \Sigma_{hkl} \sqrt{1/n-1} \Sigma_{i=1} |I_i - I_m| / \Sigma_{hkl} \Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively

structure of AhlB all non-Ala/Gly hydrophobic residues of the hydrophobic head are packed away in the core of the protein as the hydrophobic β-tongue. The use of such a β-tongue to shield membrane inserting hydrophobic residues is a characteristic of the soluble conformations of many ClyA family members, including Hbl-B, ClyA, MakA and NheA[8,23,24,33].

Structure of the pore form of AhlB. We determined the structure of an oligomeric pore conformation of AhlB, by growing crystals in the presence of MPD, a reagent which has been previously shown to induce pore formation in PFTs[34]. The structure of AhlB was determined to 2.9 Å in space group C2 (Form 1), using Se-Methionine SAD and subsequently the resolution was extended to 2.5 Å, using data from a crystal in space group $C222_1$ (Form 2, Table 1, Supplementary Fig. 4). The structures of both crystal forms are closely related (rmsd = 0.78 Å), with the asymmetric unit in each case containing ten copies of AhlB arranged as a ring of subunits in C5 symmetry, forming a funnel shape around a central pore, with an overall length of 143 Å and an external diameter of 115 Å at the large end of the funnel (the tail), reducing to a minimum diameter of 30 Å at the neck before finally expanding to a final diameter of 46 Å (the head), dimensions in agreement with those observed in negative stain EM for AhlB pores (Fig. 3a). The internal dimensions of the pore were calculated using HOLE[35], which showed a minimal internal diameter at the neck of 20 Å (Fig. 3b).

Each subunit of the pore is constructed from two extended α-helices (α3 and α4) that join the head of the pore to the five α-helices that fold into the pore tail. Major conformational changes have occurred when compared to the soluble structure, which are centred on rotations about two hinges, hinge 1 (K152-L160 and R248-A254) and hinge 2 (L186-G191 and A222-G225; Fig. 4). In this soluble to pore transition, changes in ϕ and ψ angles of up to 180° for the hinge residues and movements of up 94 Å result in both the β-tongue and α3a/α4a (D156-L186 and G225-Q246) unpacking and the secondary structure rearranging to form the two 140 Å extended helices (α3 and α4), constructing the head of the funnel and presenting this hydrophobic head to the membrane (Fig. 4 and Supplementary Movie 1). The stretch of adjacent hydrophobic head residues at the C-terminus of α3 (G176–A201) and the N-terminus of α4 (V212–L234), each of length 39 Å is sufficient for these two helical segments to insert fully through both leaflets of a lipid bilayer and are highly conserved in all Gram negative bacteria (Supplementary Fig. 5), with the extent of the insertion of the helices into the membrane defined by rings of tyrosine residues on α4 (Y245) and phenylalanine residues on α3 (F203; Fig. 3c), residues that are known to delineate transmembrane helices[36]. In the two crystal forms of the AhlB pore, the helices of the head domain adopt slightly different arrangements at their distal ends, due to differences in crystal packing between the two structures (Supplementary Fig. 6), indicating that there is some flexibility in the packing in the transmembrane part of the pore.

Within the oligomeric pore, AhlB adopts two subunit conformations, which vary at both the tail and head regions. In the Type 2 conformation residues from both hinges adopt helical

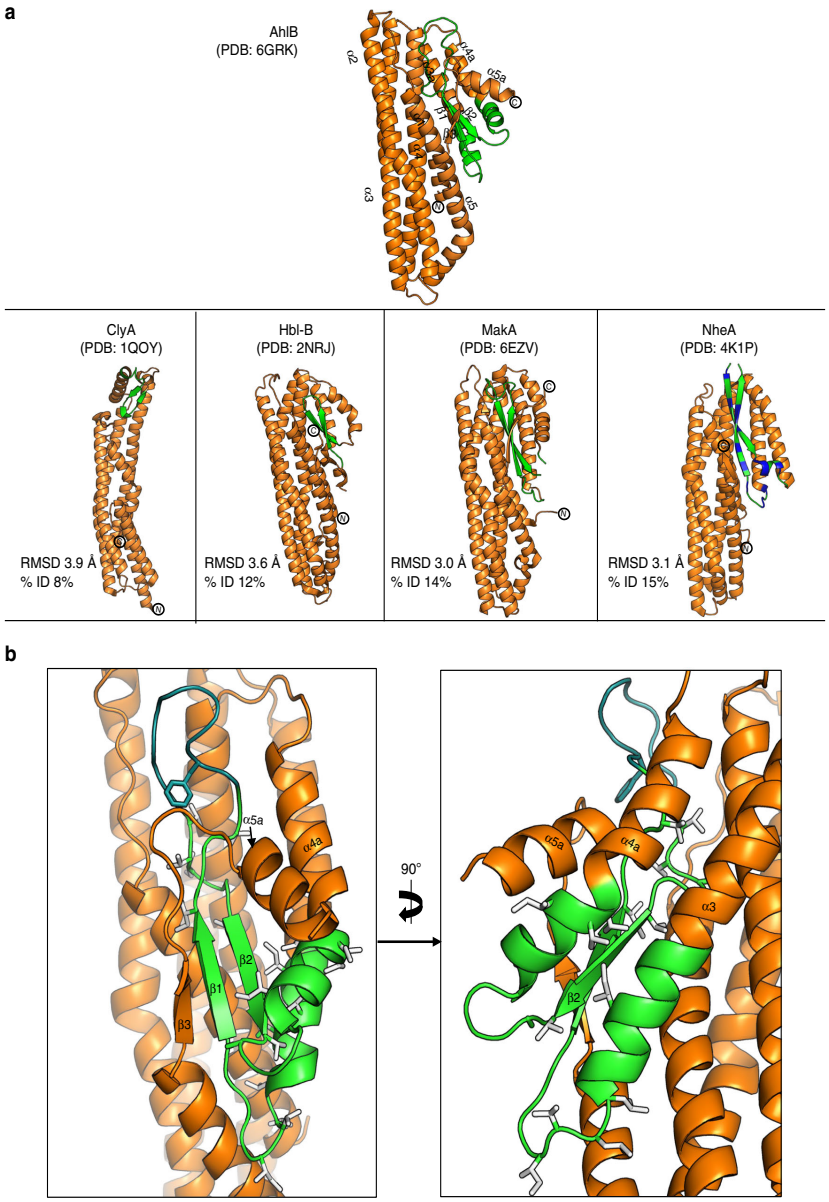


Fig. 2 β-tongue structure of soluble AhlB. **a** Structure of Soluble AhlB and comparison with other related structures. The region of the head that forms the two membrane spanning helices in the pre-pore form of AhlB are highlighted in green, below are soluble structures of ClyA, Hbl-B, MakA and NheA, the proposed hydrophobic regions coloured green with the exception of NheA where the amphipathic β-tongue is coloured green (hydrophobic) and blue (hydrophilic), RMSD and % sequence Identity to soluble AhlB are given. **b** Hydrophobic β-tongue of soluble AhlB, coloured as in (**a**). All non-Ala/Gly hydrophobic residues of the head are highlighted as sticks

conformations, forming sections of α3 and α4, whereas in the Type 1 conformation, the hinge 1 residues in α3 (K152–L160) remain in a loop, indicating that the Type 2 conformation is the fully extended form. Indeed, the transition to form AhlB Type 2 from AhlB Type I is an extension of the same hinge movement observed between the soluble and type I conformations (Supplementary Movie 1). Furthermore, in Type 1 the N-terminal helix (α1) is sandwiched between helices α4 and α5, whereas in the Type 2 conformation the first 15 residues of α1 are disordered, with the N-terminus of α1 packing against the C terminus of α5 in an end-to-end arrangement (Fig. 3c). Thus, the Type 1 conformation has a five-helix bundle at the tail, whereas

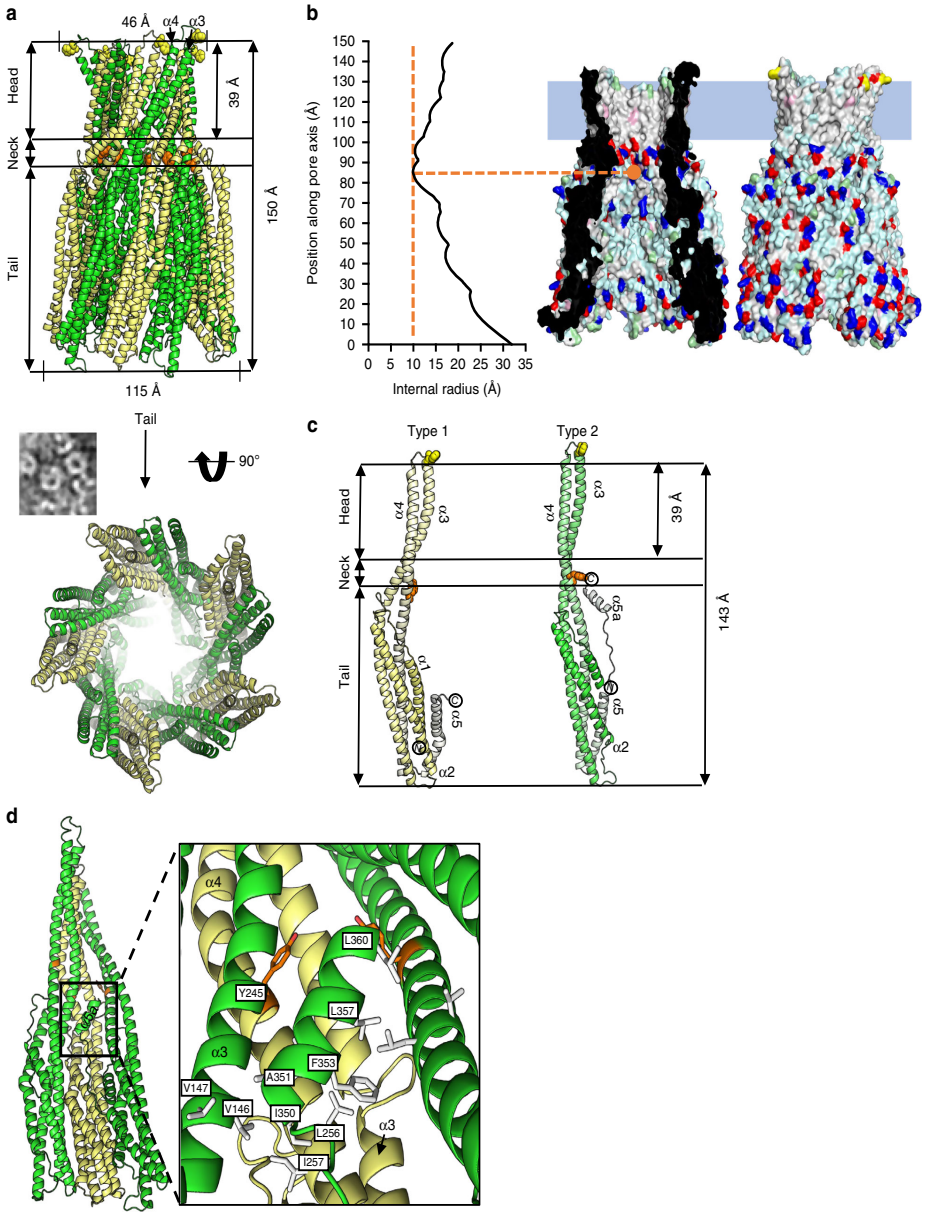


Fig. 3 Structure of the AhlB pore. **a** Cartoon representation of AhlB pore, Type1 conformation is pale yellow while Type 2 is coloured green. Y245 is shown as orange spheres and delineates the beginning of the hydrophobic head, with the ring of F203 highlighted as yellow spheres at the end of the head. Below is a negative stain EM image of AhlB pores in liposomes next to the crystal structure, orientated as seen in the EM, with the a view looking down the pore from the tail. **b** Surface rendering generated in PyMOL[61], shows hydrophobic (white), negative (red), positive (blue) and polar (cyan) surfaces. The membrane bound head domain is entirely hydrophobic, with a blue rectangle defining the proposed membrane region. The internal radius against position along the vertical pore axis is plotted with the narrowest region in the neck labelled, calculations were performed using HOLE[35]. **c** Side view of the AhlB Type 1 and Type 2 protomers. Protomers are coloured dark (N-terminus) to light (C-terminus). The Head, neck and tail are highlighted along with residues F203 and Y245 coloured as in (**a**). **d** Interactions between α5a with the neighbouring subunits in the Type 1/Type 2 dimer, with hydrophobic residues highlighted in white and Y245 coloured orange

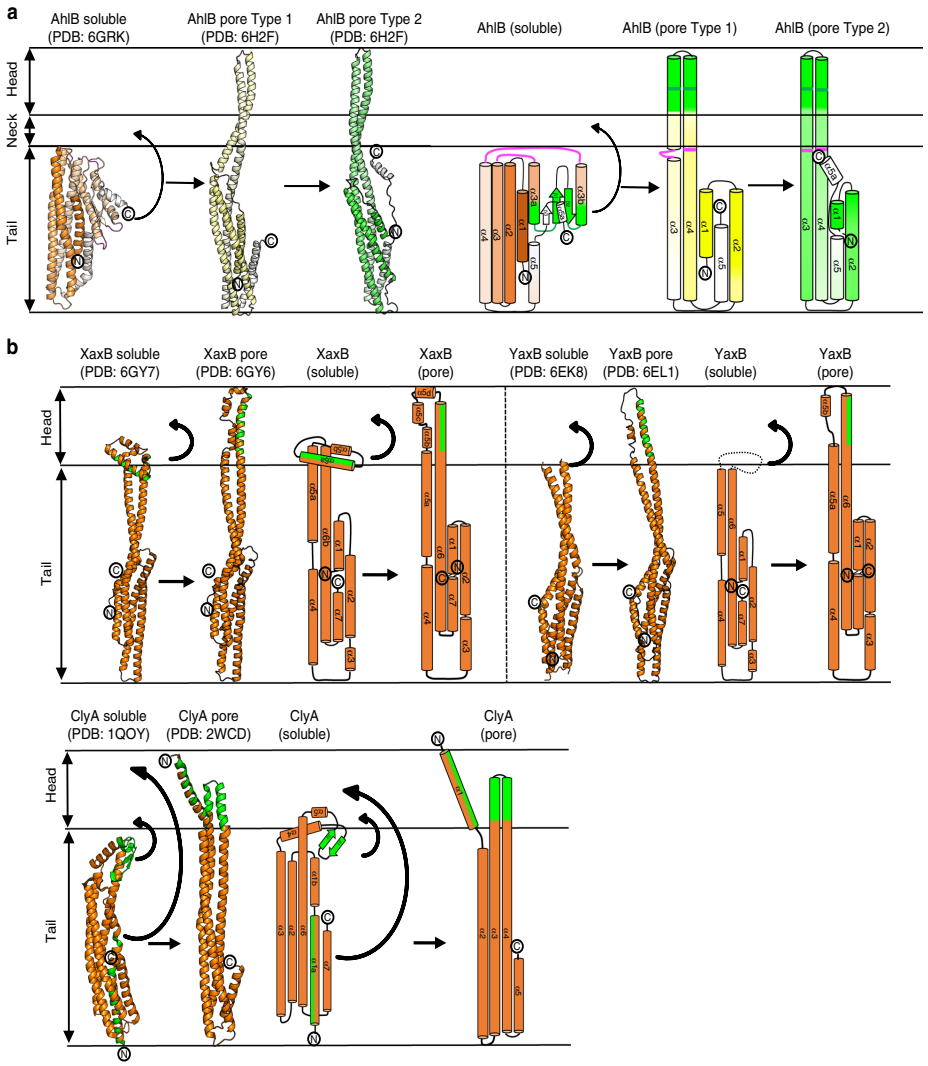


Fig. 4 Differences in conformation between soluble and pore form AhlB. The β-tongue of soluble AhlB undergoes two 180° rotations into the AhlB pore Type 1 structure to expose its hydrophobic head, followed by untwisting of the head and movement of α1 and α5 to the Type 2 structure. **a** Cartoon structures of soluble AhlB (orange) and pore form AhlB Type 1 (yellow) and AhlB Type 2 (green), coloured from dark (N-terminus) to light (C-terminus). Schematic diagrams of AhlB coloured according to (**a**), with the addition of hydrophobic regions coloured in green. The location of the two hinge regions in the soluble AhlB structure are marked on the pore structures as pink and green lines. **b** Cartoon (left) and schematic (right) diagrams of XaxB, YaxB and ClyA soluble and pore forms. The β-tongue of soluble ClyA undergoes two rotations similar to that of AhlB to form a short hyrophobic head, long enough to span a single leaflet of a lipid bilayer, while the N-terminus swings out to form an amphipathic membrane spanning helix. Unlike ClyA and AhlB, YaxB and XaxB have no β-tongue and a single hinge, which unfolds to expose a single amphipathic helix

the Type 2 conformation has a four-helix bundle at the tail. The packing arrangement of the N and C terminal helices in the tail of Type 1 is the same as that observed in soluble AhlB, as well as soluble ClyA, NheA, Hbl-B and MakA[8,23,24,33]. In contrast, the Type 2 conformation in this tail region packs into an arrangement similar to that seen in the pore structures of both YaxB and XaxB[11,15] (Supplementary Fig. 7), suggesting that the fully extended Type 2 conformation of AhlB is likely to be the conformation present in the active AhlABC pore.

In the AhlB pore one of each of the subunit conformations of AhlB pack together to form a dimer and five of these Type 1/Type 2 dimers assemble into the decameric oligomer (Fig. 3). The interface between dimers is constructed from α5a of the Type 2 AhlB monomer (residues I350, A251, F353, L357 and L360) packing against α4 of the Type 1 AhlB monomer (residues L256 and I257) and α3 in AhlB Type 2 (residues V146 and V147) of the neighbouring dimer. The first 15 N-terminal residues of the Type 2 subunits point towards the centre of the oligomer where

discontinuous electron density is observed (Supplementary Fig. 8), perhaps suggesting that these residues may form a plug in the entrance to the funnel resulting in an inactive pore, possibly explaining the lack of haemolytic activity of AhlB alone. We note, however, that the equivalent residues in the structures of the active YaxAB and XaxAB pores are also disordered[11,15].

Crystal structure of AhlC soluble tetramer. As we have shown that the binary complex of AhlB and AhlC is partially haemolytic, the structure of the soluble form of AhlC was determined to 2.8 Å resolution in space group $P6_522$ using Se-Methionine SAD, and refined to higher resolution (2.35 Å) using native data (Table 2, Supplementary Fig. 9). The asymmetric unit contained a dimer of AhlC (chains P and Q), which by rotation about the crystallographic twofold axis formed a 222 tetramer (subunits PP', QQ') (Fig. 5), consistent with gel filtration results where AhlC eluted as a tetramer (Supplementary Fig. 2).

Each subunit of AhlC had a very similar structure to that of the Type 2 conformation of the AhlB pore structure, (rmsd 2.4 Å, Fig. 6a), with a bundle of five α-helices forming the tail and helices α3 and α4 running the entire length of the protein and extending from the tail by 50 Å to form the neck and head domains (Fig. 5a). A stretch of 16 residues (L155-V171) at the end of α3 and the start of α4 in the head domain are hydrophobic, clearly indicating a possible method of membrane attachment. However, these sections of the two helices only extend over 15 Å, sufficient to span just one leaflet of the membrane, unlike AhlB where the hydrophobic parts of the equivalent structural elements can span the whole bilayer. The extent of this hydrophobic head domain is marked by a group of conserved lysine residues (K150, K151 and K152) and a tyrosine (Y154) on α3, which would delineate the extent of membrane insertion (Fig. 5b).

Within the tetrameric assembly of soluble AhlC residues L156, L158, L160 and L163 in the α3–α4 hydrophobic head of subunit P pack against the hydrophobic residues L246 and L249 in the tails of both subunit Q and the symmetry related subunit P', to form leucine zippers in the head to tail arrangement of the tetramer. Y154 of subunit P also forms π-stacking interactions with F250 in P', and K152 of subunit P forms an ionic interaction with E257 of chain Q (Fig. 5b). The head of subunit P' adopts the equivalent interactions in the tetramer. The other two hydrophobic heads from subunits Q and Q' are more exposed on the tetramer surface and in the crystal pack against neighbouring hydrophobic heads from symmetry related tetramers (Supplementary Fig. 10). We have also determined the structure of a second crystal form of soluble AhlC in space group $P2_1$ (Table 2, Supplementary Fig. 9), which contains a complete tetramer in the asymmetric unit. In this second crystal form the head domains of subunit Q and S do not form crystal contacts and the electron density for these areas is poor, indicating that these hydrophobic head domains are flexible within the tetramer (Supplementary Fig. 10). Thus, AhlC appears to use assembly of the quaternary structure to hide its hydrophobic head from the cytosol, rather than the conformational change mechanism seen in AhlB. Alignment of AhlC with YaxA or XaxA in their respective pores show that AhlC is closely related to both these structures, but that the extended conformation of AhlC subunit Q is more similar to YaxA or XaxA than

Table 2 X-ray data collection and refinement statistics for AhlC structures

	AhlC Form 1 (SeMet)	AhlC Form 1 PDB: 6H2E	AhlC Form 2 PDB: 6H2D	AhlC head mutant PDB: 6R1J
Data collection				
Beamline	I04	I04	I03	I04
Wavelength (Å)	0.9763	0.9763	0.9763	0.97951
Space group	$P6_522$	$P6_522$	$P2_1$	$P6_122$
Cell parameters				
a, *b*, *c* (Å) =	134.6, 134.6, 145.3	134.7, 134.7, 145.3	65.1, 61.7, 130.0	88.5, 88.5, 291.0
α, *β*, *γ* (°) =	90, 90, 120	90, 90, 120	90, 92, 90	90, 90, 120
Molecules per asymmetric unit	2	2	4	2
Resolution (Å)[a]	2.81–29.54 (2.81–2.88)	2.35–58.33 (2.35–2.41)	2.62–46.74 (2.62–2.69)	1.92–76.68 (1.92–1.95)
Total reflections[a]	751,732 (51,617)	1,293,084 (97,395)	105,013 (7858)	988,355 (22,848)
Unique reflections[a]	19,468 (1369)	32,942 (2392)	31,150 (2309)	52,765 (2550)
R_{merge}[a,b]	0.129 (1.013)	0.138 (3.967)	0.074 (0.710)	0.084 (2.073)
R_{pim}[a,c]	0.029 (0.230)	0.023 (0.639)	0.058 (0.509)	0.019 (0.710)
Mean $I/\sigma(I)$[a]	28.1 (4.8)	19.8 (1.4)	11.0 (1.8)	16.1 (0.9)
Completeness (%)[a]	99.8 (98.1)	99.8 (99.8)	99.7 (100)	100 (100)
Multiplicity[a]	38.6 (37.7)	39.3 (40.7)	3.4 (3.4)	18.7 (9.0)
Mid-slope	1.36			
dF/F	0.059			
Refinement				
No. of non-H atoms		3866	6152	4108
R_{work}/R_{free}		0.22/0.27	0.27/0.33	0.23/0.28
Average B factors (Å^2)		68	58	51
Bond length rmsd (Å)		0.012	0.011	0.0092
Bond angle rmsd (°)		1.50	1.48	1.61
Ramachandran favoured/allowed (%)		96.4/100	93.7/100	99.0/100

[a]Values in brackets are for data in the high-resolution shell
[b]$R_{merg} = \Sigma_{hkl} \Sigma_i | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$
[c]$R_{pim} = \Sigma_{hkl} \sqrt{1/n-1} \Sigma_{i=1} | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively

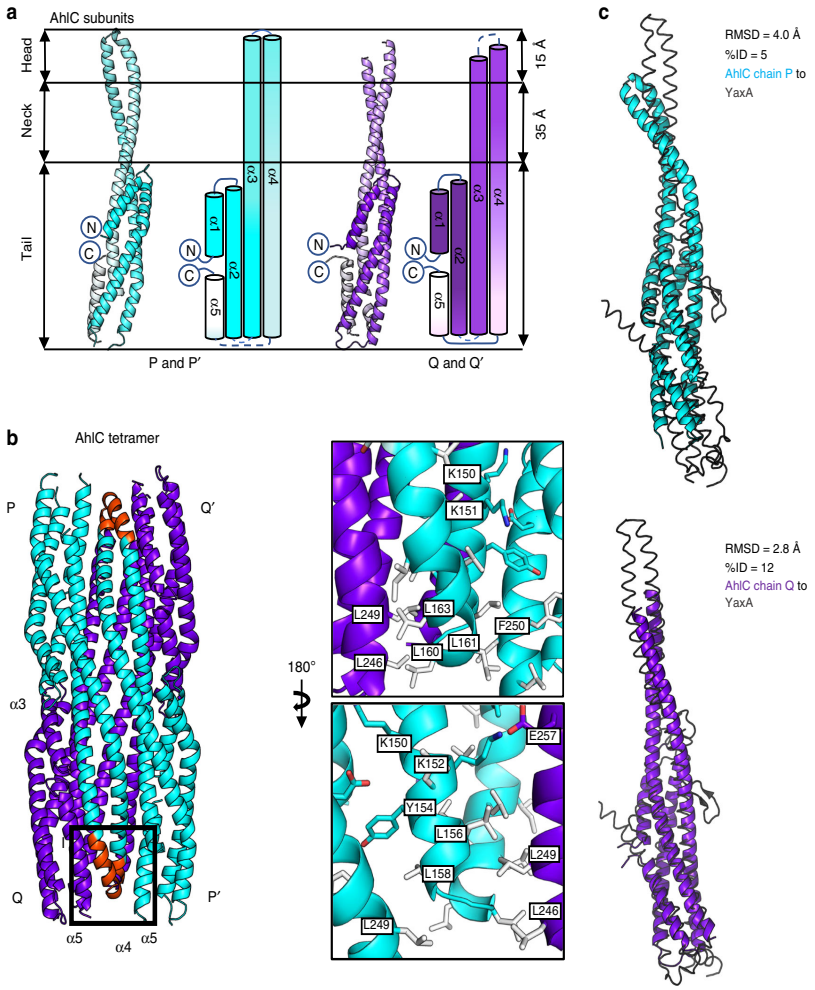


Fig. 5 The AhlC soluble tetramer buries its hydrophobic heads. **a** Cartoon and schematic structures of AhlC Form 1 subunits P and P′ (cyan) and Q and Q′ (purple), coloured dark (N-terminus) to light (C-terminus). **b** Tetrameric assembly of AhlC, hydrophobic heads of α3 and α4 are highlighted (red) in subunit P and form leucine zipper interactions with α5 in the tails of both subunit Q and subunit P′. Hydrophobic residues are coloured white. **c** Superposition of YaxA (grey) with the two independent subunits of subunit P and Q of the AhlC tetramer, as aligned by the DALI server [59], showing higher structural similarity to subunit Q

that of AhlC subunit P, where the head is buried within the tetramer (rmsd 2.8 and 2.7 Å; and 4.0 and 2.8 Å, for subunits Q and P against YaxA and XaxA, respectively, Fig. 5c).

To determine if disassembly of the tetramer is required for membrane binding, we cross-linked soluble AhlC with glutaraldehyde to produce equal proportions of AhlC monomer and AhlC dimer, whereas when AhlC was cross-linked with both glutaraldehyde and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, the dimeric species was formed exclusively (Supplementary Fig. 11a). Haemolytic assays using these cross-linked samples showed that lysis was abolished for the purely dimeric form and reduced by about 50% for the mixture of monomeric and dimeric AhlC. Based on the length of the hydrophobic head of AhlC, its delineating charged residues and the observation that it is not lytic in isolation, we suggest that AhlC inserts into a single leaflet of the bilayer. We thus constructed a triple leucine to threonine mutation (α3 L156T, and α4 L160T, L161T) of AhlC to reduce the hydrophobicity of the head. Haemolytic assays using the triple mutant of AhlC resulted in an almost 80% reduction in activity of the AhlABC complex (Supplementary Figure 11b). We also determined the structure of this triple mutant of AhlC (Table 2), which showed it adopts the same tetrameric structure as that observed for the wild-type AhlC (rmsd 0.6 Å) . Taken together, these results show that monomeric AhlC is required for lysis, and that disassembly of the tetramer of AhlC and the hydrophobic nature of the head of AhlC are both vital for pore formation.

Discussion

Within the greater ClyA family there are a number of requirements that need to be met in constructing a fully active pore

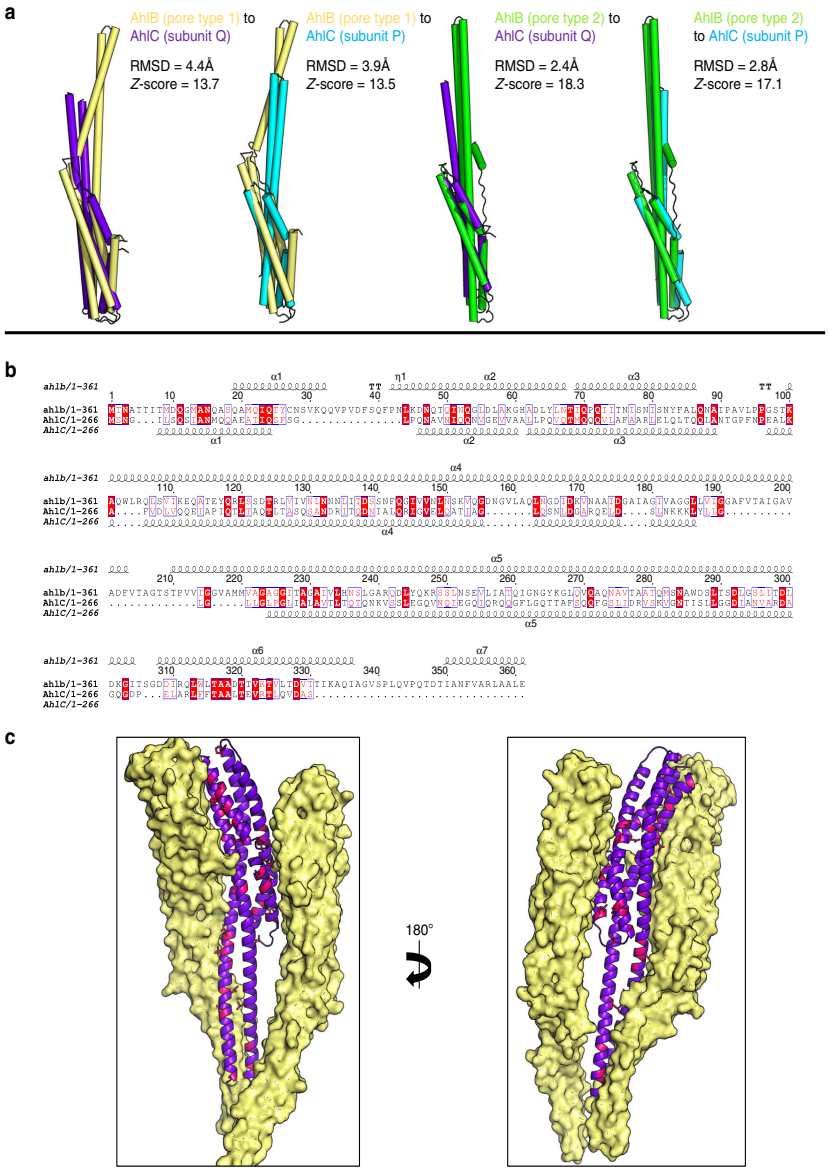


Fig. 6 Conserved residues between AhlB and AhlC. **a** Superposition of AhlB Type1 (yellow) and Type 2 (green) with subunits P and Q of the AhlC tetramer (blue and purple respectively), as aligned by the DALI server[59], showing higher structural similarity to subunit Q. **b** Sequence alignment of AhlB and AhlC. Identical residues are highlighted in red boxes, with similar residues in red lettering. Secondary structure is shown above (AhlB) and below (AhlC). **c** One AhlB Type 2 subunit is replaced by a single AhlC subunit Q (purple) in the AhlB pore. Conserved interface residues on AhlC are coloured pink, with AhlB Type 1 coloured pale yellow

forming system. These include masking the hydrophobic membrane attachment residues from solvent in the soluble form of the protein; the transition from soluble to pore form; the initial attachment to the membrane and the oligomerization to form the active pore. For the simplest member of the family, the *E. coli* protein ClyA itself, these four functions are carried on the same polypeptide. The ClyA protomer unpacks from a compact soluble form, refolding its β-tongue region into the hydrophobic α3–α4 head that inserts into one leaflet of the membrane, whilst further conformational changes expose the amphipathic N-terminal helix which penetrates both membrane leaflets and forms the lining of the pore upon oligomerization[9,10].

Comparison of the soluble and pore structures of AhlB shows that a major rearrangement of the protein occurs between them,

in a similar way to that seen in the ClyA pore formation. The hydrophobic β-tongue of AhlB is shielded from solvent by packing against α3a, α4a, the N-terminal helix and C-terminal helix of the tail domain in the soluble conformation. A bifold hinge mechanism operates to unpack the β-tongue from the tail domain to form the 140 Å extended α3 and α4 helices, refolding the hydrophobic β-tongue into a α-helical conformation to provide the membrane spanning α3–α4 head domain, in ClyA these conformational changes have been proposed to occur via a molten globule state[10]. Unlike in ClyA where the α3 and α4 head inserts into just one leaflet[9], the equivalent hydrophobic head region of AhlB is of sufficient length to cross both membrane leaflets, indicating that AhlB may be the pore forming component of the system. A similar role is played by both YaxB and XaxB in their respective bipartite pores, where the α5–α6 head (equivalent to α3–α4 in AhlB) inserts through both leaflets of the membrane, with a similar, but smaller conformational change to that seen for AhlB or ClyA[10,11,15] (Fig. 4). The α5–α6 head of both YaxB and XaxB protomers is formed from amphipathic helices making a hydrophilic pore lining, whereas both the head in AhlB, and thus the pore lining, are hydrophobic.

The attachment of a toxin component to a single leaflet of the target bilayer occurs in ClyA, and for the YaxA and XaxA components of the YaxAB and XaxAB bipartite PFTs[10,11,15] . The extent of hydrophobic residues in the α3–α4 head of AhlC is only sufficient to insert into one leaflet and thus AhlC most likely carries this role in the AhlABC tripartite PFT, supported by liposome float assays and associated negative stain EM which show that AhlC associates with a bilayer, but that no pores are formed from AhlC alone. In addition, the triple Leu–Thr head mutant of AhlC designed to decrease its hydrophobicity, reduces lytic activity by 80%, emphasising the head's critical role. We suggest the single leaflet insertion of AhlC is the initial binding event, as the lag time for lysis is abolished when erythrocytes are pre-incubated with AhlC, a situation also observed for the Nhe and Hbl tripartite PFTs of *B. cereus* where NheC and Hbl-B (each equivalent to AhlC) have been shown to attach to membrane lipids to start pore assembly[37,38].

For AhlC membrane attachment the hydrophobic head must become exposed. In ClyA this is achieved during the conformational change from the soluble to pore form[9,10]. However, the AhC subunits of the soluble tetramer adopt a pore type conformation with the hydrophobic heads buried within the tail domains of symmetry related subunits. Thus some disassembly of the AhlC tetramer must occur before membrane attachment, supported by the cross-linking of AhlC, which abolishes lytic activity. Flexibility of the head to disassociate from the tetramer when presented with a hydrophobic surface is also shown in the crystals of AhlC, where the α3–α4 hydrophobic heads of two subunits protrude from the tetramer to pack against equivalent regions of symmetry related molecules in one crystal form and make no contacts and are disordered in the other. Such disassembly of the AhlC tetramer on binding to membrane would produce a distinct concentrating effect of initiating subunits on the membrane surface, explaining the rapid lysis of erythrocytes pre-incubated with AhlC before addition of the other pore components. As the structure of AhlC is very similar to the AhlB Type 2 conformation, with 30 out of the 55 conserved residues between AhlB and AhlC present at the AhlB Type1/Type 2 interface (Fig. 6) a further role for AhlC may be to instigate the conformational change in AhlB by providing a similar structure and binding surface to that seen between protomers in the AhlB pore.

Multiple sequence alignment of the C components of the α-PFTs that we have identified from other Gram-negative bacteria show that residues involved in the packing interactions that conceal the hydrophobic head in the AhlC tetramer are conserved (Supplementary Fig. 12). Residues K152 and E257 which form a salt bridge between the head and tail domains in AhlC are identical in all the proteins with the exception of *Moritella* sp., *Janthinobacterium lividum* and *Serratia* sp. where conservative substitutions occur (K152Y and E257S, respectively). The hydrophobic heads of all the C components are 15 residues long and are rich in Leu and Ile with L158 in α3, L163 in α4 and L249 in α5 conserved and could thus form the same leucine zippers as seen in AhlC. This suggests that the tetrameric assembly of soluble AhlC is conserved across these tripartite α-PFTs of Gram-negative bacteria (Supplementary Fig. 12). It thus appears that in the tripartite AhlABC system the initial transition from soluble to membrane forms of AhlC and binding into a single leaflet occurs via a different mechanism to ClyA, based on the disassembly of a tetramer rather than the conformation change mechanism seen in ClyA[9,10]. In contrast, the hydrophobic residues of YaxA and XaxA (equivalent to AhlC) are not concealed in their respective monomeric soluble structures, although these residues do form hydrophobic interactions in their crystal lattices, perhaps suggesting that soluble YaxA and XaxA may fold in a manner similar to ClyA or AhlB to hide their hydrophobic residues.

The structures of the tripartite Ahl PFT proteins AhlC and the pore form of AhlB presented here are closely related to those of the bipartite PFTs YaxAB and XaxAB (Supplementary Figs. 7 and 13). The decameric AhlB pore is constructed from five dimers of the Type1 and Type 2 conformations, with the Type 1 conformation an intermediate on the soluble to pore transition (Supplementary Movie 1). As we have shown that AhlBC pores are partially active, and as an equal number of the single leaflet insertion heads surround the membrane spanning pore in both the ClyA and YaxAB/XaxAB pores, it would thus seem reasonable to propose that AhlC and AhlB could assemble into a pore containing 10 copies of AhlB and 10 of AhlC, with the ring of AhlB subunits surrounded by a ring of AhlC subunits (AhlC is equivalent to YaxA). However, as 80% lysis can also be achieved with a ratio of AhlC:AhlB:AhlA of 0.25:1:1 (Fig. 1b), pores with fewer copies of AhlC might also be lytic. Nevertheless, using the YaxAB pore as a guide, we modelled AhlB and AhlC into a decameric pore. Using alternating AhlB Type 1 and Type 2 conformations with AhlC produced a pore with severe steric clashes at the head. However, if a pore is made using the AhlB type 2 conformation alone with AhlC, no steric clashes occur (Supplementary Fig. 14), further indicating that the AhlB Type 2 conformation is likely to be that of the active pore structure.

This model of an AhlBC pore has a constriction diameter of approximately 30 Å, slightly larger than that seen in the AhlB pore structure (20 Å), but similar in size to both the YaxAB and XaxAB structures (31 Å)[11,15] and consistent with the observation that molecules larger than ~20 Å cannot pass through a similar NheBC pore from the tripartite *B. cereus* Nhe system[22]. However, this proposed AhlBC pore has a hydrophobic pore lining, as the α3–α4 head of AhlB is formed exclusively from hydrophobic residues. This is in direct contrast to ClyA, XaxAB and YaxAB, where the pore lining itself is hydrophilic. As maximal lysis only occurs when all three components of the AhlABC system are present, AhlA must play an important role in the pore construction. Indeed, the liposome float assays and complementary ultracentrifugation experiments with detergent show that all three components are present in pores formed from AhlA, AhlB and AhlC and EM images of liposome preparations show that efficient pore formation, with pores visible in every liposome, only occurs when AhlA, AhlB and AhlC are incubated together with the membranes, compared to the situation with AhlB/AhlC alone, when pores can be seen in just some of the liposomes (Fig. 1e and Supplementary Fig. 3e).

One possible function of AhlA is that it provides the hydrophilic pore lining. AhlA shares closest sequence similarity with *B. cereus* NheA (39%) enabling a structure based sequence alignment to be constructed between NheA (PDB code 4k1p) and AhlA (Supplementary Fig. 15). In this alignment, β1 of NheA (N196–T214) is equivalent to residues D189–T207 of AhlA, with residues K230-S248 in AhlA equivalent to β2 of NheA (T217–A235). There is an insertion in the AhlA sequence between β1 and β2, compared to NheA, which is also seen in the other A components identified in Gram-negative bacteria. Assuming a similar bifold soluble to pore transition for AlhA as seen in AhlB, then these β1–β2 residues would form the extended membrane spanning helices (α3–α4) and would be of sufficient length to pass through both leaflets of the membrane, with S208–A229 in AhlA forming a intracellular loop between them. The sequence of these helices shows that α3 of the head would be amphipathic with α4 hydrophilic (Supplementary Fig. 15). These helices could pack together via their hydrophilic surfaces to produce a membrane-spanning element with an overall amphipathic character. It is, therefore, plausible that within the fully active pore, AhlA produces the hydrophilic pore lining with the hydrophobic surface of AhlA α3 packing against the hydrophobic surface of AhlB, and the hydrophilic α4 lining the inside of the pore. As the kinetic assays show that equal amount of AhlA and AhlB are present in the most active pores, the pore could be formed with alternating AhlA and AhlB subunits at the tail and a ring of AhlA α3–α4 heads providing the hydrophilic lining of the membrane-spanning region. AhlA could be accommodated within the pore without reducing the internal diameter by movement of the AhlB heads, which have been shown to be flexible between the two crystal forms of the AhlB pore (Fig. 7). Such an AhlABC pore would necessitate an increased outside diameter for the transmembrane region, compared to the AhlB or AhlBC pores alone, as this pore model would have a splayed funnel shape, like YaxAB and XaxAB, the pore diameter at the funnel entrance would only have to be slightly bigger in the AhlABC pore than the AhlBC pore, in agreement with negative stain EM images that show a similar size for the AhlABC and AhlBC pores. Flexibility in assembling pores is shown in both ClyA and XaxAB PFTs, where pores of variable protomer number

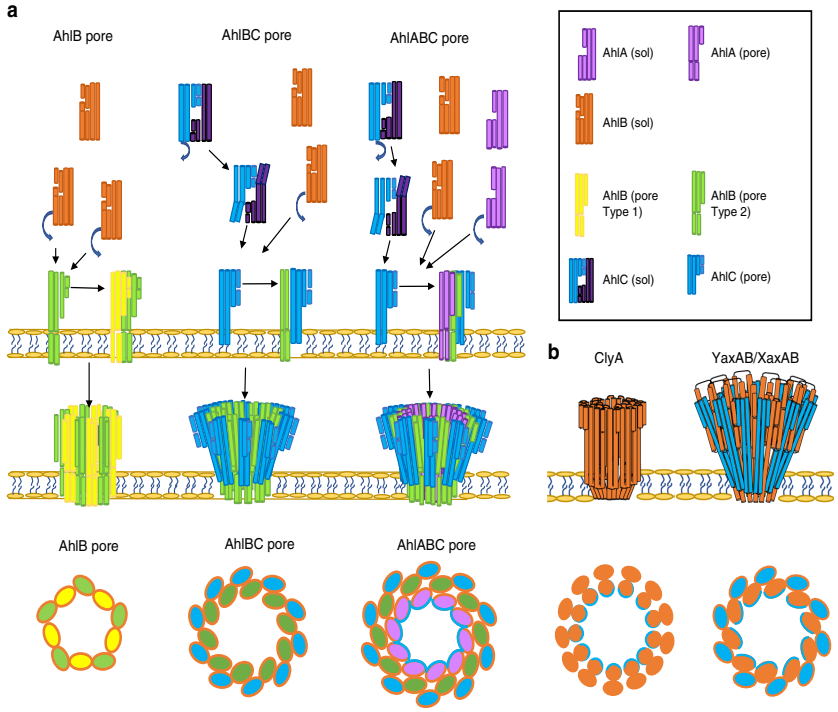


Fig. 7 A proposed assembly schematic of AhlB, AhlBC and AhlABC pores. **a** Assembly of AhlB pores; soluble AhlB (orange) reconfigures to the Type 1 pore conformation (yellow) on exposure to the lipid bilayer and recruits more AhlB monomers to form an inactive pore of mixed Type 1, Type 2 (green) conformation. Assembly of AhlBC pores; when AhlC (cyan) is present tetramers of AhlC disassemble at the membrane and monomers insert into one leaflet. Soluble AhlB (orange) is recruited to the lipid bilayer where it unpacks (green) to form a hetero dimer with AhlC. Further AhlB and AhlC are recruited until a complete pore is formed with a ring of AhlC on the outside and a hydrophobic ring of AhlB on the inside. Assembly of AhlABC pores; AhlC inserts in to the membrane as in AhlBC pore assembly, soluble AhlA and AhlB are then recruited to the membrane where they associate with AhlC and further AhlA, AhlB and AhlC are recruited until a complete pore is formed with a hydrophilic lining from AhlA on the inside of the membrane spanning region. Shown below is a cross-section views through the membrane bound region of each pore, with AhlB type 1 (yellow), AhlB type 2 (green), AhlC (blue) and AhlA (pink) with hydrophobic surfaces (orange) and hydrophilic (light blue) highlighted. Each oval represents the two α3, α4 head helices. Relative sizes of the proposed pores are not implied by these schematics. **b** Schematics of the ClyA and the YaxAB/XaxAB pores. Below is a cross-section view through the membrane bound region of each pore. ClyA and YaxB/XaxB are coloured orange, while YaxA/XaxaA is coloured blue. Hydrophilic (blue) and hydrophobic (orange) surfaces are highlighted. Both pores have a hydrophilic internal lining

have been observed[15,32], which may also be the case in AhlABC, (for example, with alternating protomers of AhlA and AhlB providing the membrane spanning region), but the overall architecture would be the same. The amphipathic α3 and hydrophilic α4 sequence pattern of AhlA is also conserved within the A component of the other tripartite Gram-negative bacteria α-PFTs we have identified, suggesting that a similar role for the A component may operate throughout the tripartite family (Supplementary Fig. 15c).

The large scale conformational change seen in both AhlB and ClyA and proposed for AhlA presumably also occurs for at least the tripartite PFT components *B. cereus* NheA and HblB, as structures of the soluble forms of these two proteins are in the same conformation as seen in the soluble forms of AhlB and ClyA (Fig. 2). In addition, the *Vibrio cholerae* MakA protein, of the motility associated cytotoxin operon, also shares the same general structure as soluble AhlB, but as yet it is unclear whether the *mak* operon encodes a ClyA family PFT[33]. A similar, but smaller scale conformational change occurs for the bipartite YaxB and XaxB proteins[11,15], indicating that this mechanism occurs across the whole ClyA family.

It is thus clear that initial insertion into a single leaflet by two short hydrophobic α helices is a common feature of all these α-PFT's, carried out by AhlC in the three component AhlABC system, by YaxA or XaxA in the two component YaxAB or XaxAB pores and by the β-tongue in the single component ClyA pore, albeit the method by which the hydrophobic residues are occluded in the soluble forms is different. The membrane spanning pore itself is then assembled using the N-terminal helix of ClyA or the components YaxB or XaxB in the bipartite YaxAB and XaxAB pores. However, it appears that in the tripartite AhlABC pore each component carries a separate role. AhlC provides the initial single leaflet insertion, assembly with AhlB then producing an oligomeric hydrophobic pore, and finally recruitment of AhlA produces a hydrophilic fully active AhlABC pore. Structural and sequence analysis suggests this method of pore assembly may well occur throughout the tripartite α-PFT family.

The structures of the tripartite AhlABC toxin components described here show that each are related with those from the greater ClyA family, although low sequence similarity is seen to the greater family members. The known importance of ClyA, NheABC, $HblL_1L_2B$, YaxAB and XaxAB for virulence in their respective organisms[11,17,20,21], suggests that the large number of similar α-PFT systems that we have identified in pathogenic Gram-negative bacteria may also be important for the virulence of these organisms. Disruption of the pore assembly may well thus provide a means to develop new virulence-targeted therapies. We also note that biotechnological applications of protein membrane pores to encapsulate enzymes for biosensors[39] and for nucleotide sequencing[40] are increasing and identification of this greater tripartite α-PFT family may well provide a wider variety of different assemblies to be exploited.

Methods

Bioinformatic search. The amino acid sequence of *B. cereus* NheB was submitted to BlastP[41], which identified the gene AXH33180.1 (*ahlB*) from *A. hydrophila*. Upstream and downstream analysis of the *A. hydrophila* genome identified genes AXH33179.1 (*ahlA*) and AHX33181.1 (*ahlC*). All three genes were aligned with the respective NheABC components using Tcoffee[30] and hydrophobicity plots were generated for each gene using Protscale[42]. The amino acid sequences of AHX33180.1, AXH33179.1 and AHX33181.1 were submitted to BlastP[41] excluding *Aeromonas* species to identify orthologues with *E* values < 0.01.

Protein cloning and purification. The ahl genes *ahlA*, *ahlB* and *ahlC* were amplified from genomic DNA of *A. hydrophila* strain AL09–71 using polymerase chain reaction with synthetic oligonucleotide primers (Eurofins) and a Q5 high-fidelity polymerase (NEB) (Supplementary Table 2). The amplified fragments were subjected to double restriction digest using NdeI/XhoI (NEB) for AhlB, AhlC and $AhlC^{HM}$, and NdeI/HindIII (NEB) for AhlA, before ligation by T4 ligase (NEB) into a pET21a expression vector (Novagen). The AhlC head mutant was generated from the cloned AhlC wild-type gene using a Q5 mutagenesis kit (NEB; for primer sequences see Supplementary Table 2). All constructs were designed to contain a C-terminal His_6 tag.

Each protein was expressed in an *E. coli* BL21 DE3 expression cell line (NEB). All cultures were grown in LB media at 37 °C, until an OD_{600} of 0.6 was reached, at which point protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). For AhlA, protein expression was carried out at 16 °C overnight, whilst for AhlB and AhlC protein expression was carried out at 25 °C overnight. Purification of AhlA, AhlB and AhlC was carried out using the same protocol. Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 8, 0.2 M NaCl), sonicated (3×20 s bursts at 16,000 nm λ) and insoluble material was removed by centrifugation (40,000*g*). Soluble protein was applied to a 5 ml Nickel Hi-trap column (GE Healthcare) in binding buffer (50 mM Tris-HCl pH 8 and 0.5 M NaCl) at a flow rate of 5 ml min^{-1}. Protein was eluted on a gradient of binding buffer to 50 mM Tris-HCl pH 8, 0.5 M NaCl and 1 M imidazole. Proteins were further purified by size exclusion chromatography using a Superdex 200 pg column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 and 0.5 M NaCl. AhlA purification buffers also contained cOmplete protease inhibitors (Roche), and AhlA was stored in buffer containing 50 mM L-Arg and L-Glu[43], to prevent degradation. AhlB ran as a partially oligomerized species of 90 kDa on sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) when boiled with 1 mM DTT, but this high molecular weight species disappeared and AhlB ran at its expected molecular weight (36 kDa) when heated to 60 °C. This phenomenon may occur due to partial protection of the extensive hydrophobic domain of AhlB by SDS[44], but it was not observed for either of the AhlA or AhlC proteins.

Selenomethionine derivatized proteins were expressed in the same way, but prior to induction the culture was centrifuged for 10 min at 20 °C (30,000*g*) and the pellet was washed in 50 ml of minimal media (10.5 g/l K_2HPO_4, 1 g/l $(NH_4)SO_4$, 4.5 g/l KH_2PO_4, 0.5 g/l tri-sodium citrate, 5 g/l glycerol, and 0.5 g/l of each nucleobase adenine, guanosine, thymine and uracil, 1g/l $MgSO_4{\cdot}7H_2O$, 4 mg/l Thiamine, 100 mg/ml of each of L-lysine, L-phenylalanine, L-threonine, and 50 mg/ml of each of L-isoleucine, L-leucine, and L-valine, and 40 mg/l of seleno-L-methionine). Cells were pelleted again, this time at 5000*g*, then resuspended in 500 ml of minimal media. The culture was then induced by addition of 1 mM IPTG, and grown for 2 days at 16 °C before harvesting. Selenomethione derivatized proteins were purified in the same way as nonderivatized proteins.

AhlB and AhlC complex formation analysis. Totally, 8 mg/ml of purified AhlB and 12 mg/ml of purified AhlC were mixed together in equal volumes and then applied to a Superdex 200 GE column which had been pre-equilibrated with 0.1 M NaCl and 50 mM HEPES pH 7.5. Fractions were collected and protein content was analysed by SDS-PAGE.

Haemolytic assays. The haemolytic activity of AhlA, AhlB and AhlC, was determined by measuring the release of haem from lysed cells photometrically at 542 nm, as described by Rowe and Welch[45]. A 0.25% (w/v) erythrocyte suspension was prepared from horse blood (Themo scientific) by washing cells repeatedly by centrifugation at 1500*g* for 5 min and resuspending cells in 10 mM phosphate buffer pH 7.4, 2.7 mM KCl and 137 mM NaCl. Varying concentrations of each AhlABC protein were incubated with 1 ml of the erythrocyte suspension and incubated on a blood wheel at 37 °C for 1 h. Erythrocytes were centrifuged at 1500*g* for 5 min, and the supernatant was removed for photometric analysis. A positive control of cells lysed in ddH_2O and a negative control with no protein added were used to normalise the data for 0 and 100% lysis. Haemolytic assays using AhlC head mutant were performed using the same method as used for the wild-type protein. For measurement of lysis against time, the rate of lysis was reduced in order to take measurements every 10 min, using a 0.5% (w/v) erythrocyte suspension with a protein concentration of 500 nM. All assays were carried out in triplicate.

Ultracentrifugation. Totally, 10 µM of purified AhlA, AhlB or 5 µM of purified AhlC were incubated with 45 µM N-heptyl-thioglucopyranoside at a total volume of 100 µl for 1 h at 25 °C. A third sample containing a mixture of 10 µM AhlB and 5 µM AhlC, or 10 µM AhlA, AhlB and 5 µM AhlC, was incubated with 45 µM N-hepty-thioglucopranoside at a total volume of 100 µl for 1 h at 25 °C. After incubation, samples were spun for 30 min at 214,500*g* in a Beckman Optima MAX ultracentrifuge at 4 °C. Supernatant was separated from the pellet and the pellet was resuspended in 100 µl 50 mM Tris-HCl pH 8 and 0.5 M NaCl. Finally, the soluble fraction and pelleted fraction from each sample was analysed by SDS-PAGE. In the ultracentrifugation assays AhlA partially degraded in the absence of other components, possibly due to disruption of the stabilising interactions between AhlA and the buffer components Glu and Arg, required to stabilise soluble AhlA.

Liposome floatation assay. Liposomes were prepared from *E. coli* total lipid extract (Avanti Polar lipids), using the extrusion method[46]. Totally 100 mg of solid

lipids were resuspended in 1 ml of 2:1 chloroform:methanol. A total of 100 µl of this solution was dried with nitrogen gas and then flash frozen in liquid nitrogen, before further desiccation under a vacuum for 2 h. The final lipid film was resuspended in 1 ml of 10 mM PBS pH 7.4 and vortexed until fully dissolved. The lipid solution was then extruded through a 100 µm filter 11 times to generate a uniform 100 µm liposome suspension, as gauged by negative stain electron microscopy. Totally, 300 µg of liposomes with 100 µg of protein for were incubated for 15 min at 37 °C before addition of 100 µg of protein 2 or the equivalent volume of PBS buffer. This mixture was then incubated for a further 45 min at 37 °C. After incubation the total volume was made up to 800 µl with 55% (w/v) sucrose in PBS. The liposome mixture was transferred to an ultraclear ultracentrifuge tube (Beckman Coulter), overlaid with 3.8 ml 40% (w/v) sucrose in PBS, followed by a layer of 400 µl PBS. Samples were centrifuged at 200,000×*g* for 4 h at 4 °C in a Beckman Coulter SW 55 Ti rotor. Six 100 µl fractions were taken from the top of the tube, 3.8 ml removed and a further six 100 µl fractions taken from the bottom, which were analysed by SDS-PAGE.

Cross-linking of AhlC. AhlC was cross-linked using both glutaraldehyde and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) in a two step reaction. First 10 µl of 25% w/v glutaraldehyde was added to 90 µl of 1.4 mg/ml AhlC and incubated at room temperature for 15 min, to give 50% cross-linked AhlC, before buffer exchanging into 0.1 M MES pH 6 buffer using a Zebaspin 7K desalting column (sigma). Next 160 µg EDC was added to the partially cross-linked AhlC and incubated at room temperature for 30 min, to give fully cross-linked AhlC. This cross-linked AhlC was finally buffer exchanged into PBS buffer using a Zebaspin 7K desalting column for use in assays.

AhlC crystallisation and structure determination. Purified AhlC was concentrated to 40 mg/ml by centrifugation in a Vivaspin 10 kDa MWCO concentrator (Sartorius), and then buffer exchanged into 50 mM Tris-HCl pH 8, 100 mM NaCl with a Zebaspin desalting column (Sartorius). Native AhlC was crystallised at 16 °C by sitting drop vapour diffusion with 0.2 M $MgCl_2$, 0.1 M Tris-HCl pH 7.0, and 10 % (w/v) PEG 10,000 (200:200 nl drop). Se-Met derivatized AhlC was crystallised using the same conditions, with crystals growing in a week. Prior to cooling in liquid nitrogen, crystals were cryoprotected in mother liquor containing an additional 20 % (v/v) ethylene glycol.

X-ray diffraction data were collected from a single crystal of Se-Met AhlC on beamline i04 of the Diamond Light Source a wavelength of 0.9790 Å. Images were integrated and scaled using FastDP[47]. The crystals diffracted to 2.8 Å resolution and belonged to the space group $P6_522$, (Form 1, Table 1). Heavy atom sites and an initial map were calculated by the FastEP[48] pipeline. The asymmetric unit contained two chains of AhlC and the map was of sufficient quality to autobuild a model using Buccaneer (CCP4)[49,50], which was optimised and completed in COOT[51], before refinement of the model against a higher resolution native dataset (2.35 Å), collected on the same beamline and integrated using Xia2 3dii[52]. (Table 1). Iterative rebuilding and refinement were carried out using COOT and REFMAC, respectively[51,53] to give a final model (PDB: 6H2E) with an R and Rfree of 0.22 and 0.27, respectively. Residues 231–241, and 268–274 of chain A and residues 155–160, 239–242, and 271–274 of chain B were omitted from the final model due to poor electron density.

The second crystal form of AhlC also crystallised in a week by sitting drop vapour diffusion in 0.2 M sodium chloride, 0.1 M tris pH 8, and 20% PEG6000 (200:200 nl drop). Crystal Form 2 was cryoprotected in mother liquor with an additional 20% ethylene glycol and crystals diffracted to 2.6 Å resolution and belonged to the space group $P2_1$. Data were collected on beamline i03 at the Diamond Light Source and processed by the Xia2 3dii pipeline[52,54], (Table 1). The structure was determined by molecular replacement using one subunit from the existing Se-Methionine AhlC structure as a search model. The asymmetric unit was comprised of four subunits of AhlC. Iterative rounds of manual model building and refinement were completed using COOT and REFMAC, respectively[51,53] to give a final model (PDB: 6H2D) with R and Rfree of 0.27 and 0.33, respectively. Residues 1–4 in all chains, and 71–89, 155–162, 229–248, 265–274 of chain P, 62–86, 147–170, 235–249, and 266–274 of chain Q, 68–89, 147–172, 230–274 of chain R, and 70–88, 158–164, 234–244, and 266–274 of chain S, were omitted from the final model due to poor electron density.

The triple L156T, L160T, L161T AhlC head mutant was crystallised using sitting drop vapour diffusion in 0.1 M imidazole pH8, and 10% Peg 8000 (100:100 nl drop). The crystal was cryoprotected in mother liquor with an additional 20 % ethylene glycol. The crystals diffracted to 1.92 Å resolution and belonged to the space group P 6_122. Data were collected on beamline i04 at the Diamond Light Source and processed by the Xia2 Dials[55] pipeline (Table 1). The structure was determined by molecular replacement using one subunit from the existing Se-Methionine AhlC structure as a search model. The asymmetric unit was comprised of two subunits of AhlC With the tetramer assembled by crystal symetry. Iterative rounds of manual model building and refinement were completed using COOT and REFMAC[51,53] respectively to give a final model (PDB: 6R1J) with R and Rfree of 0.23 and 0.28, respectively. Residues 1–2 and 239–242 of chain J, and 1–3, 74–76 and 232–243 from chain D were omitted from the final model due to poor electron density.

AhlB crystallisation and structure determination. Purified AhlB was concentrated to 10 mg/ml by centrifugation using a Vivaspin 30 kDa MWCO concentrator (Sartorius), and then buffer exchanged into 0.2 M NaCl, 50 mM Tris pH8 using a Zebaspin desalting column (Sartorius). Crystals were grown by sitting drop vapour diffusion on microbridges in 60% methyl-2,4-pentanediol (MPD), 0.2 M ammonium phosphate and 0.1 M Tris pH8.5 (16 °C) (1:1 µl drop) by streak seeding with crushed microcrystals that had been diluted 10^6-fold. Crystals grew in a month and were plunge-cooled in liquid nitrogen directly from the drop.

X-ray diffraction data were collected at a wavelength of 0.9790 Å on beamline i03 at the Diamond Light Source. Data were integrated to 2.94 Å resolution in space group C2 using the DIALS[55] pipeline and scaled and merged using Aimless[56]. Initial phases of AhlB Se-Methionine were obtained by SAD. Heavy atom sites and an initial electron density map were calculated using SHELXC,D,E[48], and an initial model was built using Phenix[57] followed by Buccaneer[50]. In this initial model, a ring of 10 AhlB subunits could clearly be seen. Rebuilding and refinement using COOT and REFMAC, respectively[51,53], resulted in a final model (PDB: 6GRJ) with R and RFree of 0.22 and 0.24, respectively. Residues 1–15, 337–360 and 202–208 of chains A, C, E, G and I, and residues 205–208 and 342–360 of chains B, D, F, H and J, were omitted from the final model due to poor electron density.

The second crystal form of the AhlB pore was grown by sitting drop vapour diffusion in a month, in 0.2 M ammonium phosphate, 0.1 M tris pH 8.5, and 50% MPD (200:200 nl drop). The crystals were plunge-cooled in liquid nitrogen directly from the drop and X-ray diffraction data were collected at beamline i02 of the Diamond Light Source and processed using Xia2 2da[52,54] in space group $C222_1$ to 2.55 Å (Table 1). The structure was determined by molecular replacement with PhaserMR[58] using the 10-mer Se-Methionine AhlB pore as a search model. Iterative rounds of refining and rebuilding with Coot[51] and REFMAC[53] were carried to give a final model (PDB: 6H2F) with R and Rfree of 0.19 and 0.28, respectively, with residues 202–208 in A and D, 200–208 in E and G, 205–207 in H and 342–367 in F, as well as residues 1–17 in B, D, F, H and J, and 339–367 in A, C, E, G and I omitted from the final model due to poor density.

The soluble form of AhlB was concentrated to 10 mg/ml using a Vivaspin 30 kDa MWCO concentrator (Sartorius), and buffer exchanged into 50 mM tris pH 8 and 0.2 M NaCl by Zebaspin column (Sartorius). Crystals were grown by sitting drop vapour diffusion in 0.2 M potassium thiocyanate, 0.1 M bis–tris propane pH 6.5, and 20% PEG3350 (200:200 nl drop). Crystals took a month to grow at 16 °C, and were cryoprotected in mother liquor containing an additional 20% ethylene glycol, and plunge-cooled in liquid nitrogen. Data were collected to 2.33 Å resolution at i03 of the Diamond Light Source, and processed in space group C2 using Xia2 3da[52], (Table 1). The structure was determined by molecular replacement in PhaserMR[58] using a single subunit of the AhlB pore with the hydrophobic head domain removed as search model to give a solution with three AhlB chains in the asymmetric unit. The molecular replacement solution model was further pruned to remove overlapping residues and an initial model was built using Buccaneer[50]. Refinement of the data and iterative rebuilding were carried out in Coot[51] and REFMAC[53], respectively, to give a final model (PDB: 6GRK) with R and Rfree of 0.22 and 0.25, respectively. Structural alignments described in this paper were undertaken using Dali[59].

Electron microscopy. Samples from lytic assays using erythrocytes, ultra-centrifugation pellets, liposome flotation assays and prepared proteoliposomes were all visualised using negative stain TEM. Proteoliposomes were generated by incubating 10 µg of Ahl proteins with 20 µg of fresh liposomes in 100 µl of 10 mM PBS buffer pH 7.4, at 37 °C for 1 h and kept on ice before use. Totally, 5 µl of each of the respective samples was pipetted onto a glow discharged (copper 300 mesh) carbon-coated grid, and then stained with 1% (w/v) uranyl formate. Carbon grids were then air dried before storing for up to 5 weeks at room temperature. Electron micrographs were collected using a Philips CM100 100 kV transmission electron microscope, equipped with a Gatan 1 K CCD camera. Micrographs were collected with a pixel to nm ratio of 0.72 pixels per nm, and this was used in the calculation of the pore sizes.

Modelling of AhlA. A homology model of AhlA was generated using Phyre2[60] and the structures of soluble MakA (pdb: 6EZV), Hbl-B (pdb: 2NRJ) and XaxA (pdb: 6GY8) as templates. The model of AhlA contained 312/354 residues (84%) at a Phyre2 accuracy of >90%.

Generation of figures. All protein cartoon diagrams and surface rendering used in figures and movies were made using PyMOL[61], with intermediates in the movie morph generated using LSQmann[62].

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the finding of this manuscript are available from the corresponding author upon reasonable request. Atomic coordinates for AhlC Form 1 (PDB code 6H2E), AhlC Form 2 (PDB code 6H2D), AhlC Head mutant (PDB code 6R1J), AhlB soluble

(PDB code 6GRK), AhlB pore SeMet (PDB code 6GRJ) and AhlB pore Form 2 (PDB code 6H2F) have been deposited in the RCSB Protein Data Bank. The source data underlying Fig. 1a–d, f, g and Supplementary Figs. 3a, b and 11 are provided as a Source Data file.

Received: 6 August 2018 Accepted: 29 May 2019
Published online: 01 July 2019

References

1. Dunkelberger, J. R. & Song, W.-C. Complement and its role in innate and adaptive immune responses. *Cell Res.* **20**, 34–50 (2010).
2. Los, F. C. O., Randis, T. M., Aroian, R. V. & Ratner, A. J. Role of pore-forming toxins in bacterial infectious diseases. *Microbiol. Mol. Biol. Rev.* **77**, 173–207 (2013).
3. Rojko, N., Dalla Serra, M., Maček, P. & Anderluh, G. Pore formation by actinoporins, cytolysins from sea anemones. *Biochim. Biophys. Acta* **1858**, 446–456 (2016).
4. Young, J. A. T. & Collier, R. J. Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu. Rev. Biochem.* **76**, 243–265 (2007).
5. Allen, R. C., Popat, R., Diggle, S. P. & Brown, S. P. Targeting virulence: can we make evolution-proof drugs? *Nat. Rev. Microbiol.* **12**, 300–308 (2014).
6. Cosentino, K., Ros, U. & García-Sáez, A. J. Assembling the puzzle: oligomerization of α-pore forming proteins in membranes. *Biochim. Biophys. Acta - Biomembr.* **1858**, 457–466 (2016).
7. Peraro, M. D. & van der Goot, F. G. Pore-forming toxins: ancient, but never really out of fashion. *Nat. Rev. Microbiol.* **14**, 77–92 (2015).
8. Wallace, A. J. et al. *E. coli* hemolysin E (HlyE, ClyA, SheA): X-ray crystal structure of the toxin and observation of membrane pores by electron microscopy. *Cell* **100**, 265–276 (2000).
9. Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R. & Ban, N. The structure of a cytolytic a -helical toxin pore reveals its assembly mechanism. *Nature* **459**, 726–730 (2009).
10. Benke, S. et al. The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Commun.* **6**, 6198 (2015).
11. Bräuning, B. et al. Structure and mechanism of the two-component α-helical pore-forming toxin YaxAB. *Nat. Commun.* **9**, 1806 (2018).
12. Wagner, N. J., Lin, C. P., Borst, L. B. & Millera, V. L. YaxAB, a *Yersinia enterocolitica* pore-forming toxin regulated by RovA. *Infect. Immun.* **81**, 4208–4219 (2013).
13. Zhang, X. et al. XaxAB-like binary toxin from *Photorhabdus luminescens* exhibits both insecticidal activity and cytotoxicity. *FEMS Microbiol. Lett.* **350**, 48–56 (2014).
14. Vigneux, F. et al. The xaxAB genes encoding a new apoptotic toxin from the insect pathogen *Xenorhabdus nematophila* are present in plant and human pathogens. *J. Biol. Chem.* **282**, 9571–9580 (2007).
15. Schubert, E., Vetter, I. R., Prumbaum, D., Penczek, P. A. & Raunser, S. Membrane insertion of α-xenorhabdolysin in near-atomic detail :https://doi.org/10.1101/312314 (2018).
16. Thompson, N. E., Ketterhagen, M. J., Bergdoll, M. S. & Schantz, E. J. Isolation and some properties of an enterotoxin produced by *Bacillus cereus*. *Infect. Immun.* **43**, 887–894 (1984).
17. Beecher, D. J., Schoeni, J. L. & Wong, A. C. Enterotoxic activity of hemolysin BL from *Bacillus cereus*. *Infect. Immun.* **63**, 4423–4428 (1995).
18. Ryan, P. A., Macmillan, J. D. & Zilinskas, B. A. Molecular cloning and characterization of the genes encoding the L1 and L2 components of hemolysin BL from *Bacillus cereus*. *J. Bacteriol.* **179**, 2551–2556 (1997).
19. Beecher, D. J. & Macmillan, J. D. Characterization of the components of hemolysin BL from *Bacillus cereus*. *Infect. Immun.* **59**, 1778–1784 (1991).
20. Lindbäck, T., Fagerlund, A., Rødland, M. S. & Granum, P. E. Characterization of the *Bacillus cereus* Nhe enterotoxin. *Microbiology* **150**, 3959–3967 (2004).
21. Lund, T. & Granum, E. Characterisation of a non-haemolytic enterotoxin complex from *Bacillus cereus* isolated after a foodborne outbreak. *Microbiol. Lett. FEMS Microbiol. Lett.* **141**, 151–156 (1996).
22. Zhu, K. et al. Formation of small transmembrane pores: an intermediate stage on the way to Bacillus cereus non-hemolytic enterotoxin (Nhe) full pores in the absence of NheA. *Biochem. Biophys. Res. Commun.* **469**, 613–618 (2016).
23. Ganash, M. et al. Structure of the NheA component of the Nhe toxin from *Bacillus cereus*: implications for function. *PLoS One* **8**, e74748 (2013).
24. Madegowda, M., Eswaramoorthy, S., Burley, S. K. & Swaminathan, S. X-ray crystal structure of the B component of Hemolysin BL from *Bacillus cereus*. *Proteins Struct. Funct. Genet.* **71**, 534–540 (2008).
25. Su, L. H. et al. Extended epidemic of nosocomial urinary tract infections caused by *Serratia marcescens*. *J. Clin. Microbiol.* **41**, 4726–4732 (2003).
26. Maktar, N. H., Kamis, S., Mohd Yusof, F. Z. & Hussain, N. H. Erwinia papayae causing papaya dieback in Malaysia. *Plant Pathol.* **57**, 774–774 (2008).
27. Hossain, M. J. et al. An asian origin of virulent aeromonas hydrophila responsible for disease epidemics in united states-farmed catfish. *MBio* **5**, e00848–14 (2014).
28. Zheng, W., Cao, H. & Yang, X. Aeromonas veronii infection in the cultured snakehead fish, *Ophiocephalus argus* (Cantor). *Afr. J. Microbiol. Res.* **6**, 7218–7223 (2012).
29. Rasmussen-Ivey, C. R., Figueras, M. J., McGarey, D. & Liles, M. R. Virulence factors of *Aeromonas hydrophila*: in the wake of reclassification. *Front. Microbiol.* **7**, 1337 (2016).
30. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
31. Heilkenbrinker, U. et al. Complex formation between NheB and NheC is necessary to induce cytotoxic activity by the three-component *Bacillus cereus* Nhe enterotoxin. *PLoS One* **8**, e63104 (2013).
32. Tzokov, S. B. et al. Structure of the hemolysin E (HlyE, ClyA, and SheA) channel in its membrane-bound form. *J. Biol. Chem.* **281**, 23042–23049 (2006).
33. Dongre, M. et al. Flagella-mediated secretion of a novel Vibrio cholerae cytotoxin affecting both vertebrate and invertebrate hosts. *Commun. Biol.* **1**, 59 (2018).
34. Tanaka, Y. et al. 2-Methyl-2,4-pentanediol induces spontaneous assembly of staphylococcal α-hemolysin into heptameric pore structure. *Protein Sci.* **20**, 448–456 (2011).
35. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360 (1996).
36. Ulmschneider, M. B. & Sansom, M. S. P. Amino acid distributions in integral membrane protein structures. *Biochim Biophys. Acta* **1512**, 1–14 (2001).
37. Lindbäck, T. et al. Cytotoxicity of the *Bacillus cereus* Nhe enterotoxin requires specific binding order of its three exoprotein components. *Infect. Immun.* **78**, 3813–3821 (2010).
38. Sastalla, I. et al. The *Bacillus cereus* Hbl and Nhe tripartite enterotoxin components assemble sequentially on the surface of target cells and are not interchangeable. *PLoS One* **8**, e76955 (2013).
39. Soskine, M., Biesemans, A. & Maglia, G. Single-molecule analyte recognition with ClyA nanopores equipped with internal protein adaptors. *J. Am. Chem. Soc.* **137**, 5793–5797 (2015).
40. Franceschini, L., Soskine, M., Biesemans, A. & Maglia, G. A nanopore machine promotes the vectorial transport of DNA across membranes. *Nat. Commun.* **4**, 2415 (2013).
41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
42. Gasteiger, E. et al. *The Proteomics Protocols Handbook* 571–607, https://doi.org/10.1385/1592598900 (2005).
43. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A. & Lian, L. Y. A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* **126**, 8933–8939 (2004).
44. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl Acad. Sci.* **106**, 1760–1765 (2009).
45. Rowe, G. E. & Welch, R. A. Assays of hemolytic toxins. *Methods Enzym.* **235**, 657–667 (1994).
46. Morano, J. K., Martin, F. J. & Martin, W. *Liposome Extrusion Method. United States Patent. Patent number* 4,927,637 (1990).
47. Winter, G. & McAuley, K. E. Automated data collection for macromolecular crystallography. *Methods* **55**, 81–93 (2011).
48. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 479–485 (2010).
49. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
50. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
52. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
53. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255 (1997).
54. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
55. Winter, G. et al. DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 85–97 (2018).
56. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).

57. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
58. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
59. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44**, 351–355 (2016).
60. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
61. Schrödinger, L. L. C. *The {PyMOL} Molecular Graphics System, Version~1.8.* (2015).
62. Kleywegt, G. J. Use of non-crystallographic symmetry in protein structure refinement. *Acta Cryst. D* **52**, 842–857 (1996).

Acknowledgements

We would like to thank Julia Pridgeon of USDA for the kind gift of *A. hydrophila* genomic DNA, the Diamond Light Source for access to beamlines I02, I03 and I04 that contributed to these results, and the beamline scientists for their assistance with data collection. Electron Microscopy and mass spectrometry were performed in the Faculty of Science Electron Microscopy Facility and the ChemMS Mass Spectrometry Centre, at the University of Sheffield. This work was supported by BBSRC Doctoral Training Grant BB/F016832/1 awarded to J.S.W., University of Sheffield scholarships awarded to J.S.W. and A.M.C.-A., BBSRC grant BB/D524975/1 and EPSRC grant EP/L026872/1. We dedicate this paper to Peter Artymiuk (decd.) who initiated this work in Sheffield.

Author contributions

P.J.B. conceived the project; P.J.B., A.M.C.-A., J.S.W. and C.B. designed the experiments; A.M.C.-A., J.S.W., S.E.S., S.B.T. and S.P.D. performed the experiments; P.J.B., A.M.C.-A., J.S.W., C.B., J.B.R. and P.A.B. interpreted the data and P.J.B., A.M.C.-A. and J.S.W. prepared the paper with input from all authors.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10777-x.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information: *Nature Communications* thanks Gregor Anderluh and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

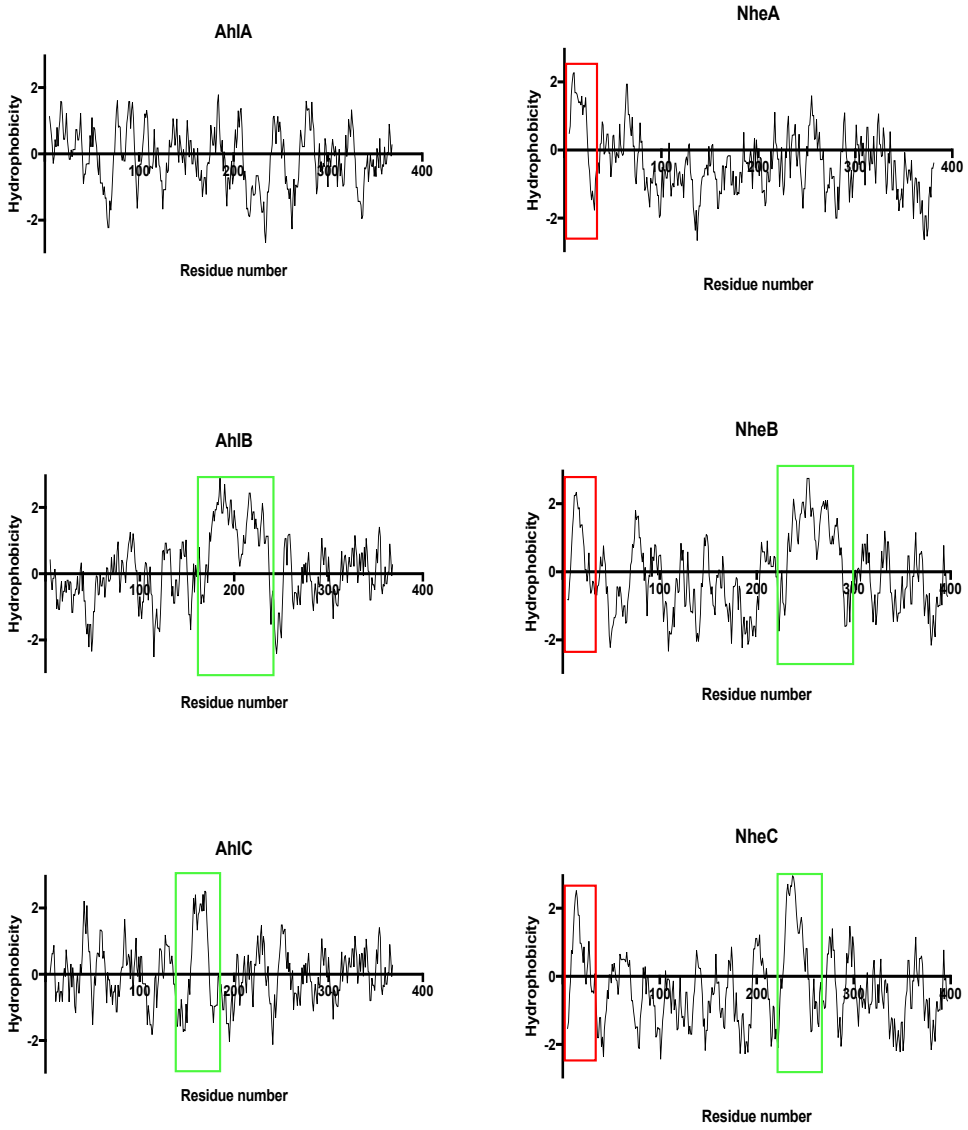
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© Crown 2019

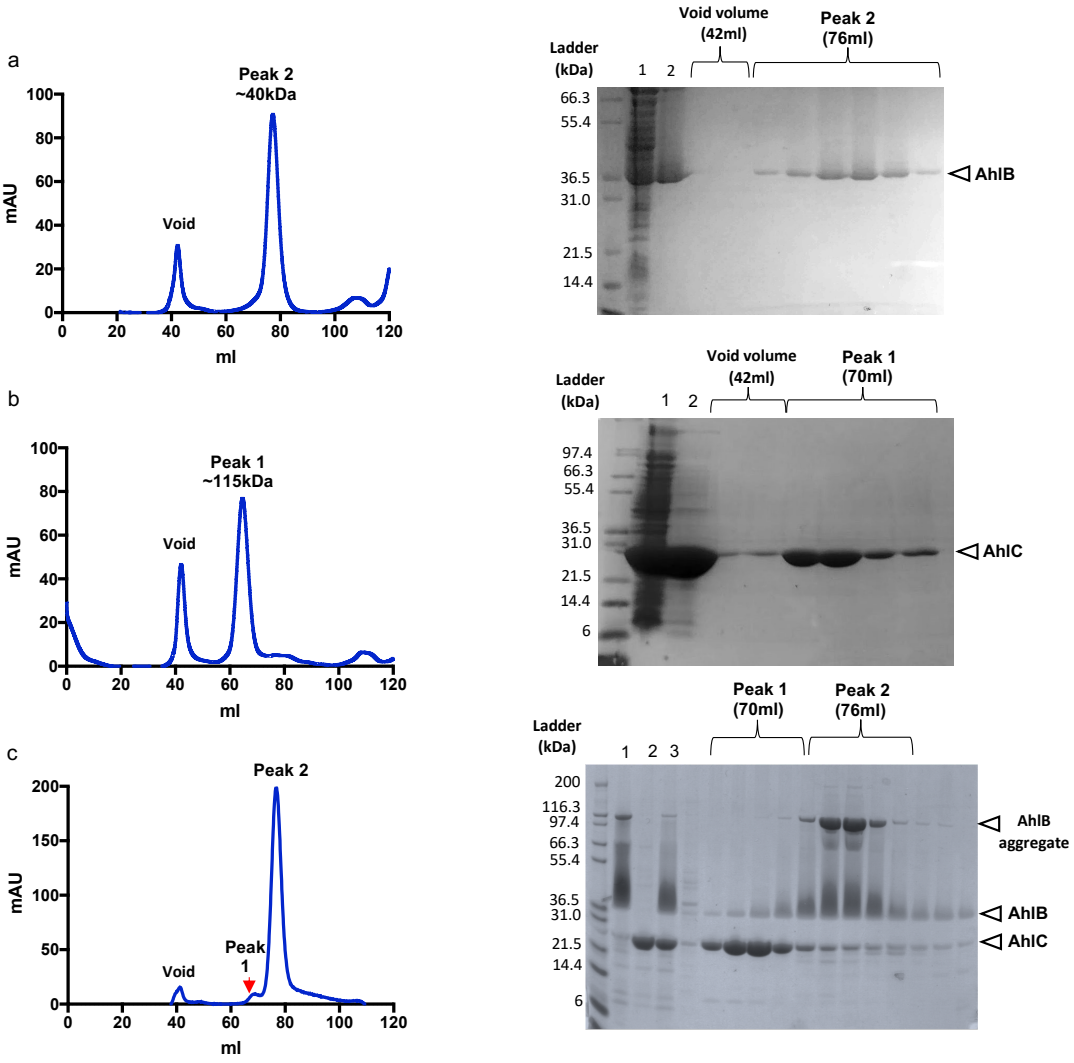
Identification and structural analysis of the tripartite α-pore forming toxin of *Aeromonas hydrophila.*

J. Wilson, A. Churchill-Angus, et al.

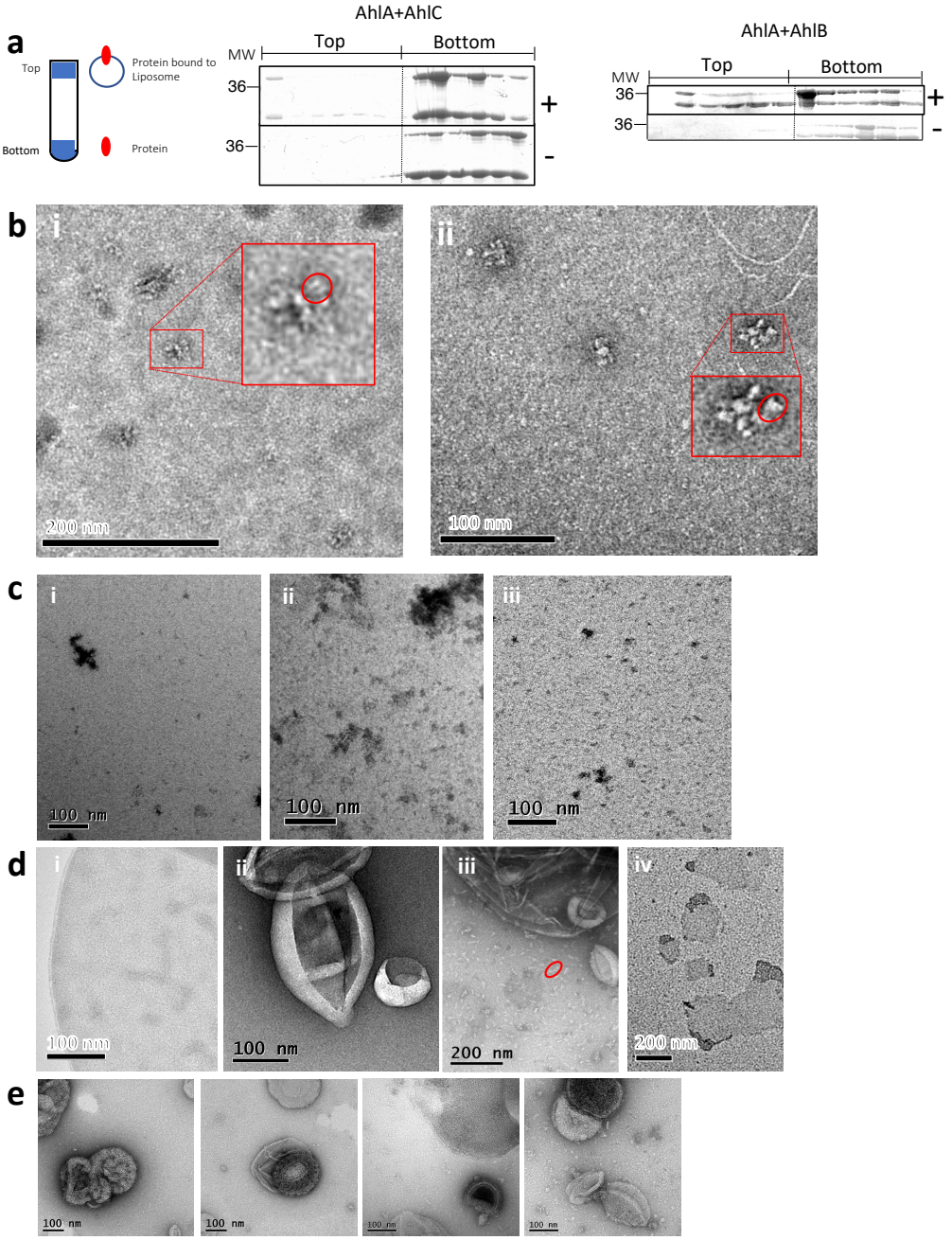
Supplementary Information



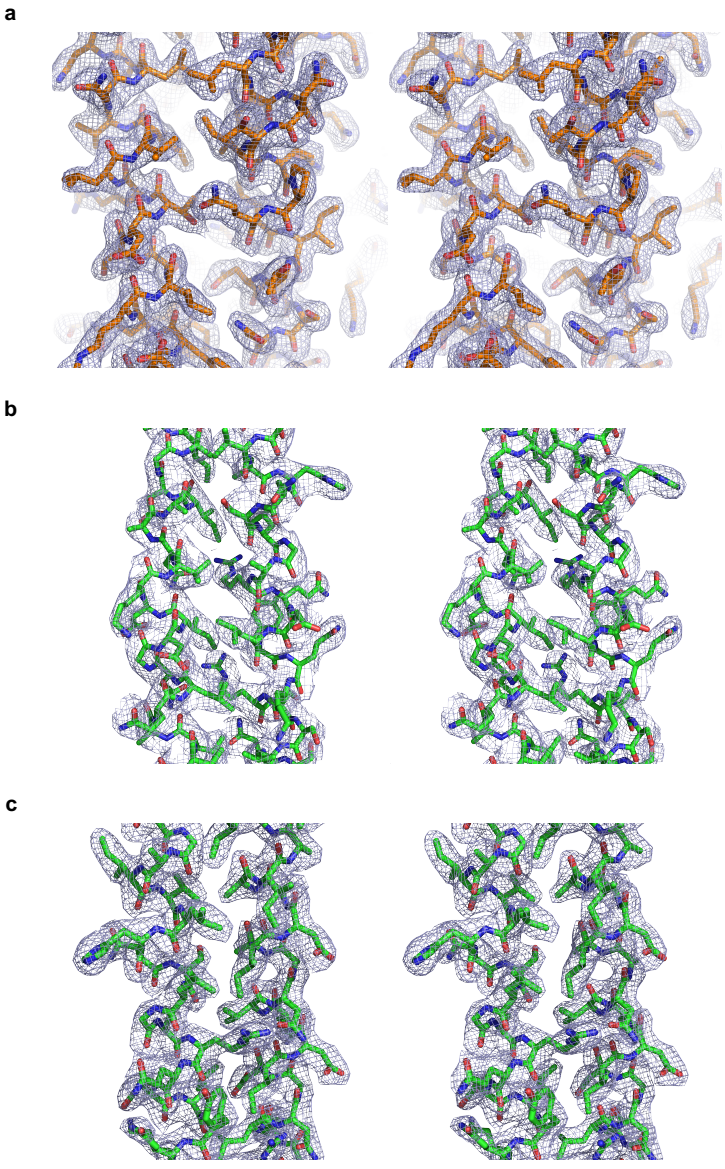
Supplementary Figure 1. Kyte and Doolittle hydrophathy plots. Kyte and Doolittle plots for *B. cereus* NheA, NheB, NheC and *A. hydrophila* AhlA, AhlB, AhlC generated using ProtScale[1] . AhlB, NheB, AhlC and NheC have a single predicted transmembrane region (green box) (>1.5), while AhlA and NheA have no predicted transmembrane region. Red boxes highlight the predicted signal sequences (SignalP [2]) in Gram positive NheA, NheB, NheC proteins



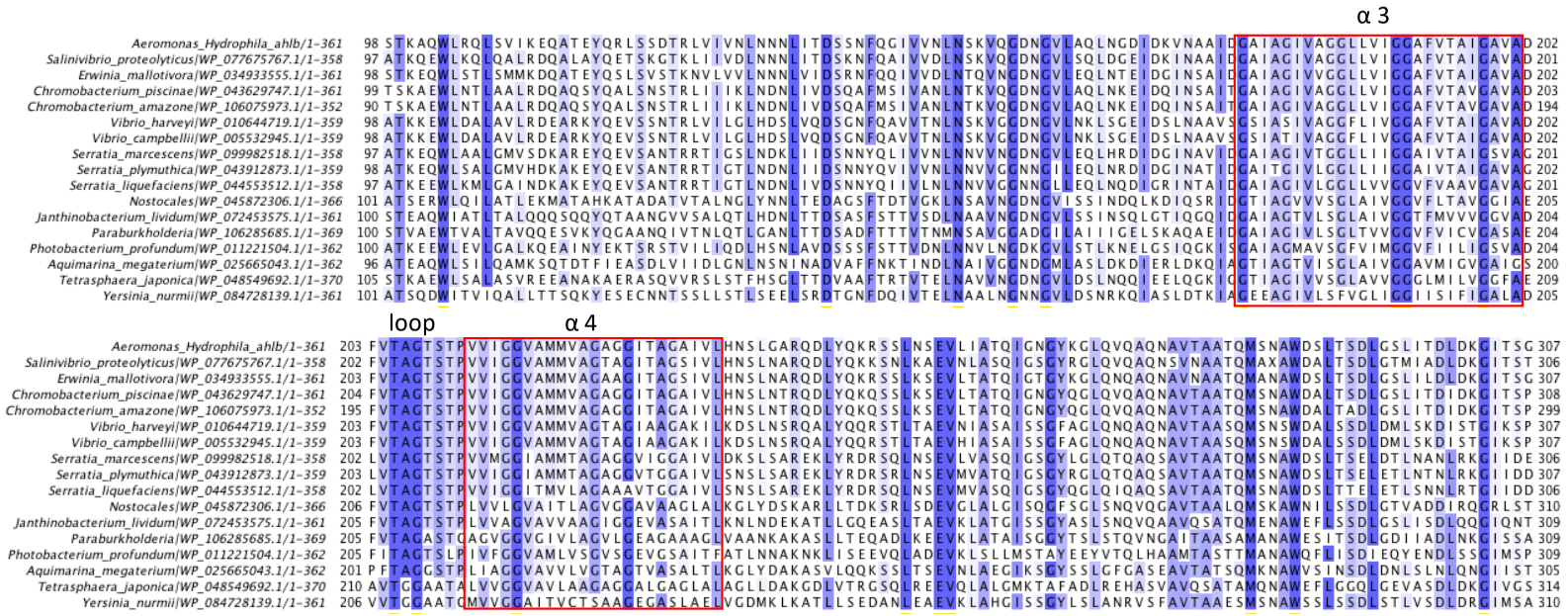
Supplementary Figure 2. Gel filtration purification for AhlB, AhlC, and AhlB+C in solution. Gel filtration UV trace chromatogram (left) and SDS-PAGE (right). (**a**) AhlB forms a monomer in solution. Lane 1 and 2 of the SDS-PAGE show cell free extract (CFE) and Gel filtration load respectively. (**b**) AhlC forms a tetramer in solution. Lane 1 and 2 of the SDS-PAGE show CFE and Gel filtration load (**c**) A 1:2(w/w) mixture of AhlB (peak 2) and AhlC (peak 1) elutes as two separate peaks, showing AhlB and AhlC do not interact in solution. The reduced absorbance of AhlC relative to AhlB because AhlC contains no Trp residues and the extinction coefficients are 1490 M^{-1} cm^{-1} and 25440 M^{-1} cm^{-1} for AhlC and AhlB respectively. Lane 1, 2 and 3 of the SDS-PAGE show AhlB, AhlC, and AhlB/AhlC Gel filtration load respectively. AhlB runs as a partially aggregated (90 kDa) species on SDS-PAGE when boiled[3]. Monomric AhlB is smeared as no stack was used on the gel.



Supplementary Figure 3 EM micrographs of erythrocytes and liposomes. (a) SDS-gels of top and bottom fractions from liposome float assays where AhlA and AhlC, and AhlA and AhlB were incubated with liposomes (+) and without liposomes (-). A schematic of the ultracentrifuge tube shows the location of the top and bottom fractions and in which fractions liposomes and protein are expected. Source data are provided as Source Data file. **(b)** Negative stain EM of ultracentrifugation pellet from AhlA+AhlB+AhlC (i), and AhlB +AhlC (ii). Side views of groups of pores in micelles can be seen in both cases. Enlarged images are highlighted in red with individual pores in red circles. (**c)** Fractions from negative controls of liposome float assays for AhlB (i), AhlBC (ii) and AhlABC (iii). No ordered pore like complexes are seen when these proteins are incubated together in the absence of liposomes. (**d**) Erythrocyte membrane (i) and liposomes (ii) in the absence of AhlABC components. Liposomes incubated with AhlA (iii), and AhlC (iv). AhlA forms extended fibrous aggregates on the surface of the EM grid (highlighted by red oval), but does not form any ordered assemblies on the surfaces of the liposomes. Liposomes are destroyed by the presence of AhlC and form flat sheets on the carbon grid. (**e**) AhlB and AhlC together saturate some liposomes with pores while others are left empty.

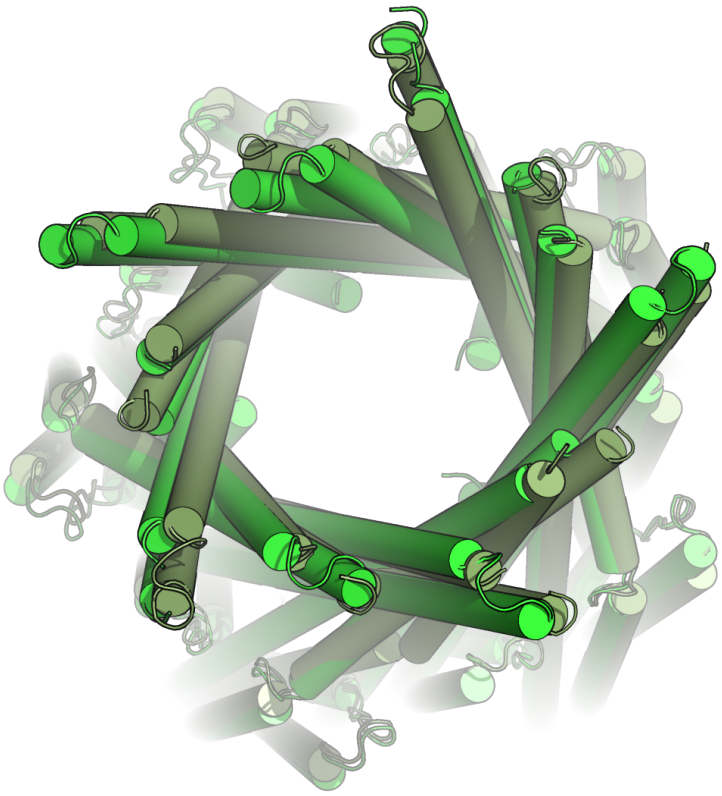


Supplementary Figure 4. Stereo images of areas of the 2Fo-Fc electron density maps of AhlB. An area of the density map (blue mesh), contoured at 1.0 σ, **(a)** AhlB soluble(D63-N79 alpha 2 and Y117-N132 alpha 3) **(b)** SeMet AhlB pore, (V159-I178 alpha 3 and V233-L251 alpha 4) **(c)** AhlB pore Form 2 (Q162-G180 alpha 3 and A229-Q246 alpha 4).

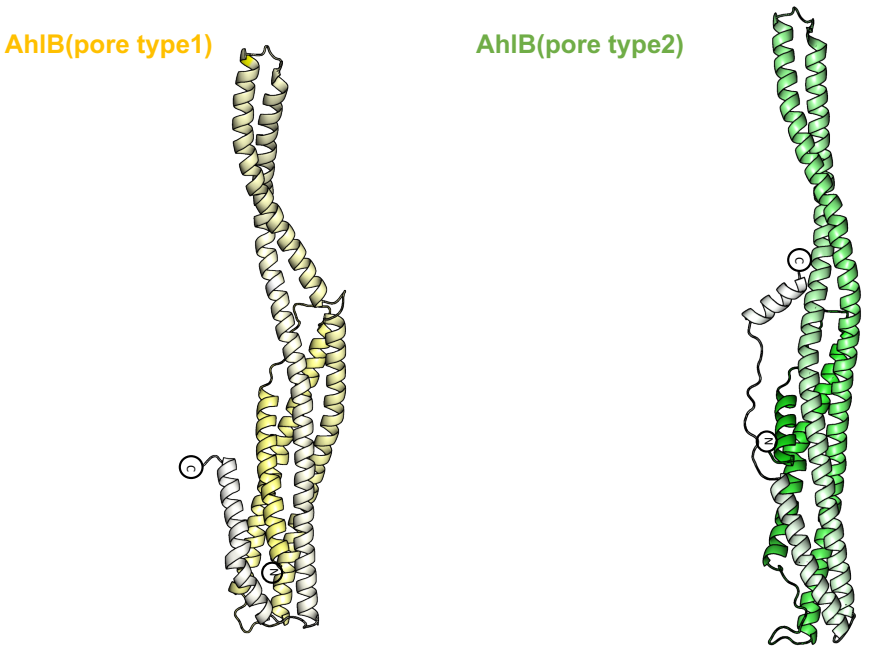


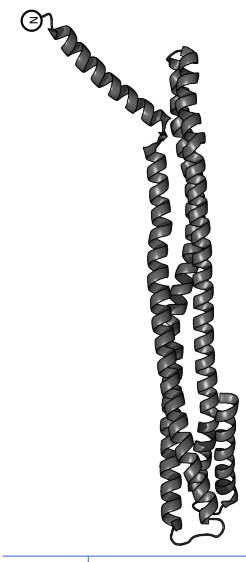
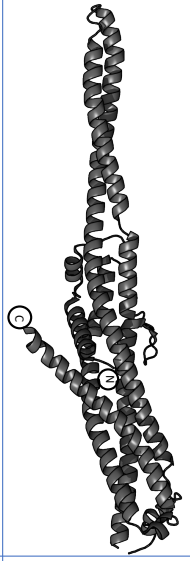
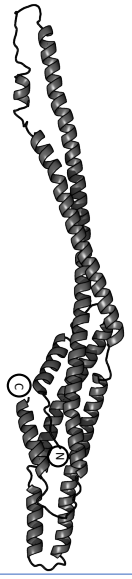
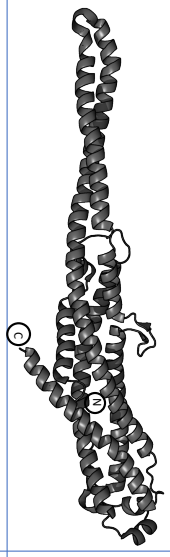
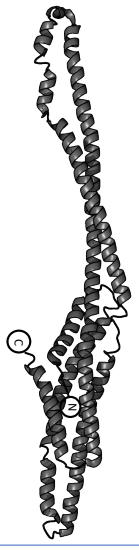
Supplementary Figure 5. Sequence analysis of AhlB. Sequence alignment, using Tcoffee [4], of AhlB against other Gram negative bacteria identified using BlastP [5]. Region G176-A201 (α 3), loop, and V212-L234 (α 4) are highlighted. The two hydrophobic helices are conserved in all bacteria as well as the highly conserved loop region.

a

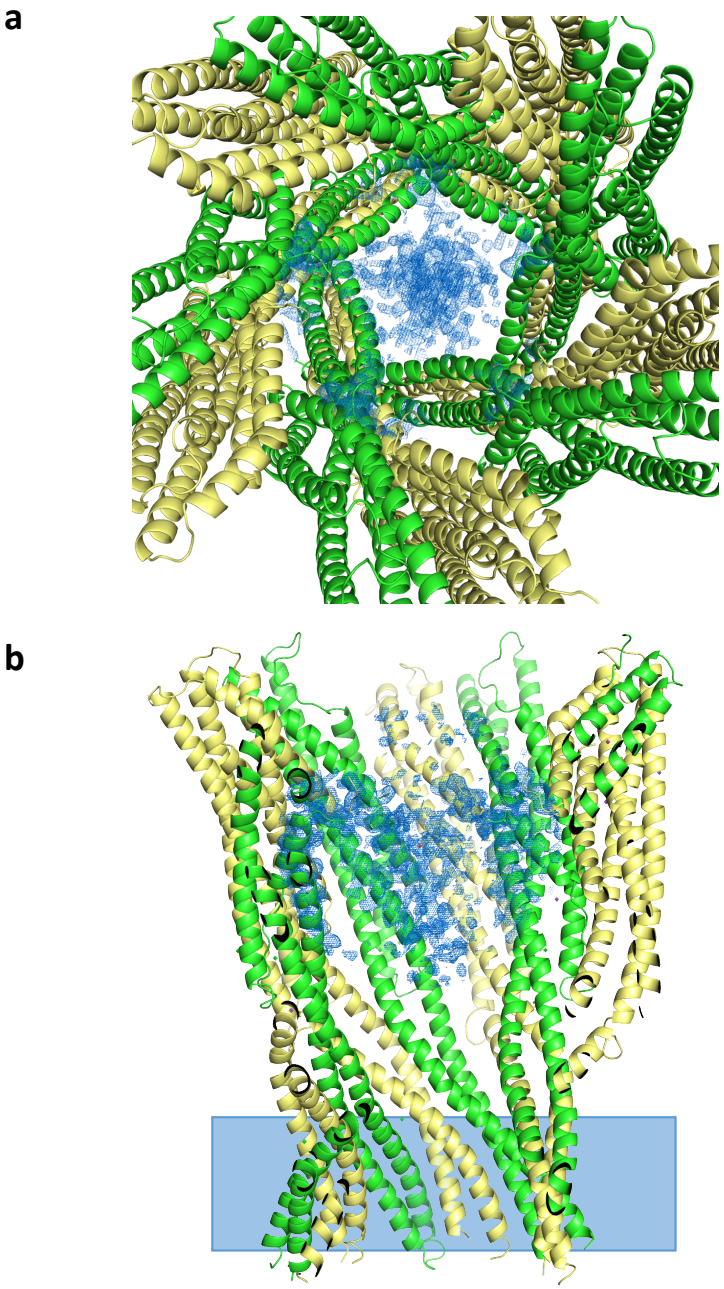


Supplementary Figure 6. Different packing of head domains in each crystal form of AhlB pore shows flexibility in the head domain. (**a**) Superposition of AhlB Form 1 (dark green) and AhlB Form2 (light green) based on residues from the tail domains, viewed from the head showing differences in the positioning of the heads in each form, suggesting flexibility of the AhlB head domain.

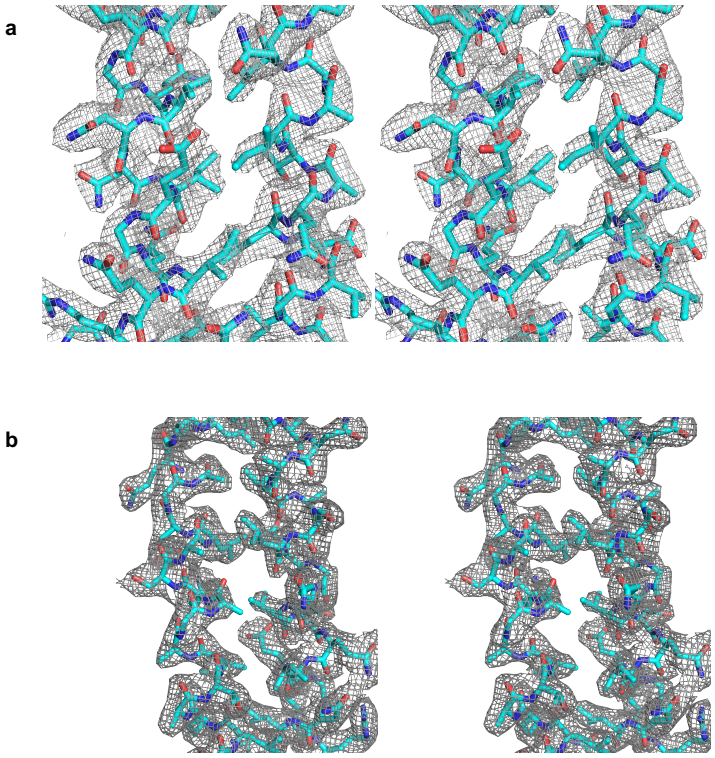


	ClyA(pore) (PDB code 2WCD)	YaxA(pore) (PDB code 6EL1)	YaxB(pore) (PDB code 6EL1)	XaxA(pore) (PDB code 6GY6)	XaxB(pore) (PDB code 6GY6)
					
AhlB Type1	Z-score 10.9 RMSD 5.2 Å % ID 7%	Z-score 15.5 RMSD 7.5 Å % ID 7%	Z-score 12.7 RMSD 4.3 Å % ID 7 %	Z-score 13.5 RMSD 4.5Å % ID 8%	Z-score 5.6 RMSD 3.8Å % ID 7%
AhlB Type2	Z-score 13.7 RMSD 4.1 Å % ID 10%	Z-score 18.1 RMSD 3.7 Å % ID 9%	Z-score 11.1 RMSD 3.5 Å % ID 7%	Z-score 20 RMSD 5.6Å % ID 9%	Z-score 12.4 RMSD 3.5Å % ID 7%

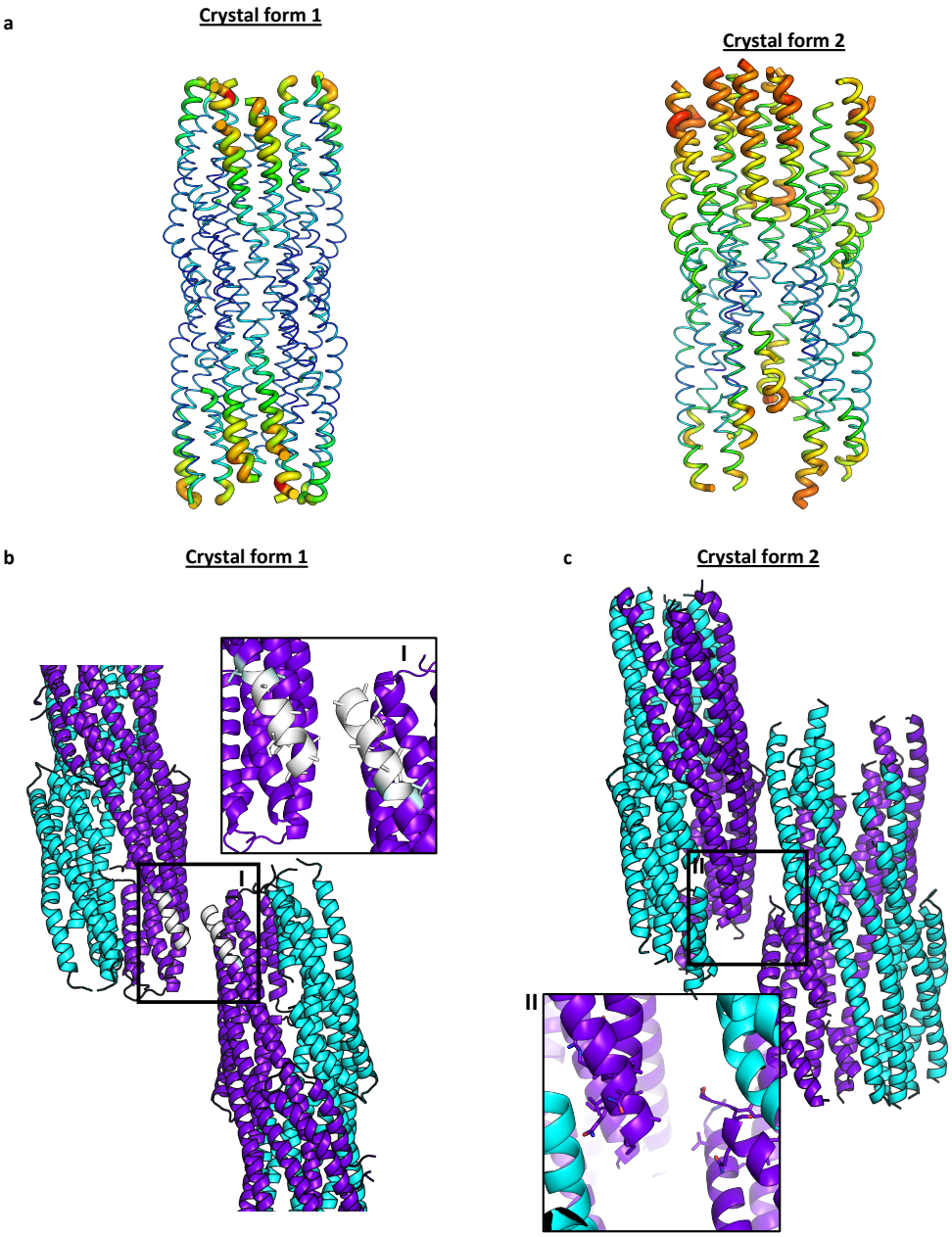
Supplementary Figure 7. Structural superposition of AhlB (pore) with ClyA α-PFT family toxins. Superposition of AhlB Type1 (yellow) and AhlB Type2 (green) with ClyA family α-PFT's (grey) (aligned by the DALI server [6]), with Z-scores, RMSD(Cα) and percentage sequence identity (%ID).



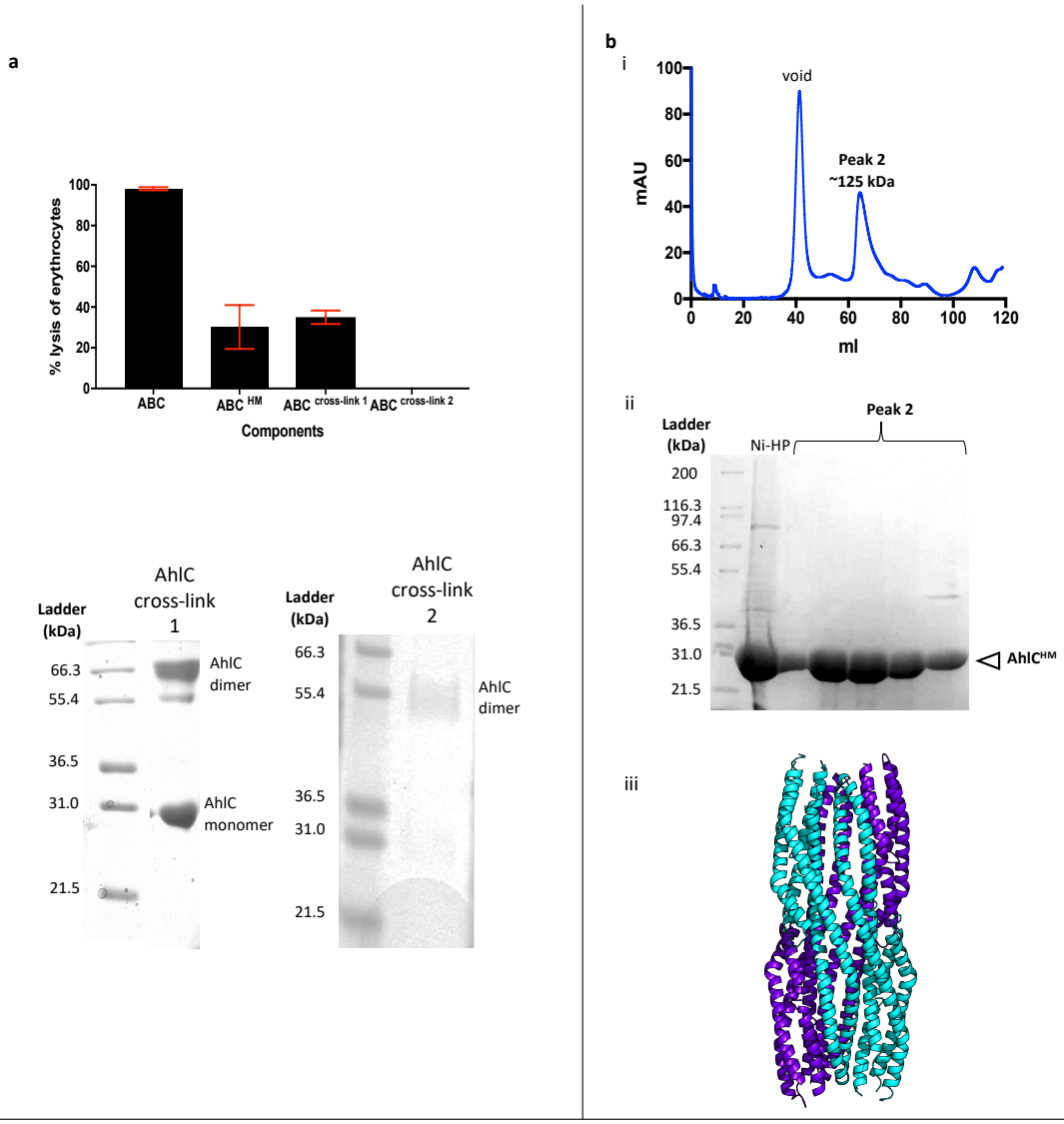
Supplementary Figure 8. Undefined density at the center of the AhlB pore. a) View looking down from the tail of the AhlB pore (green and yellow cartoon helices). 2Fo-Fc Density map, contorted at 0.9 σ (blue mesh), shows a large region of undefined density blocking the center of the AhlB pre-pore. b) side view of the AhlB pore slabbed to show the undefined density at the center. 2Fo-Fc Density map, contorted at 0.9 σ (blue mesh). The membrane is shown as a blue box.



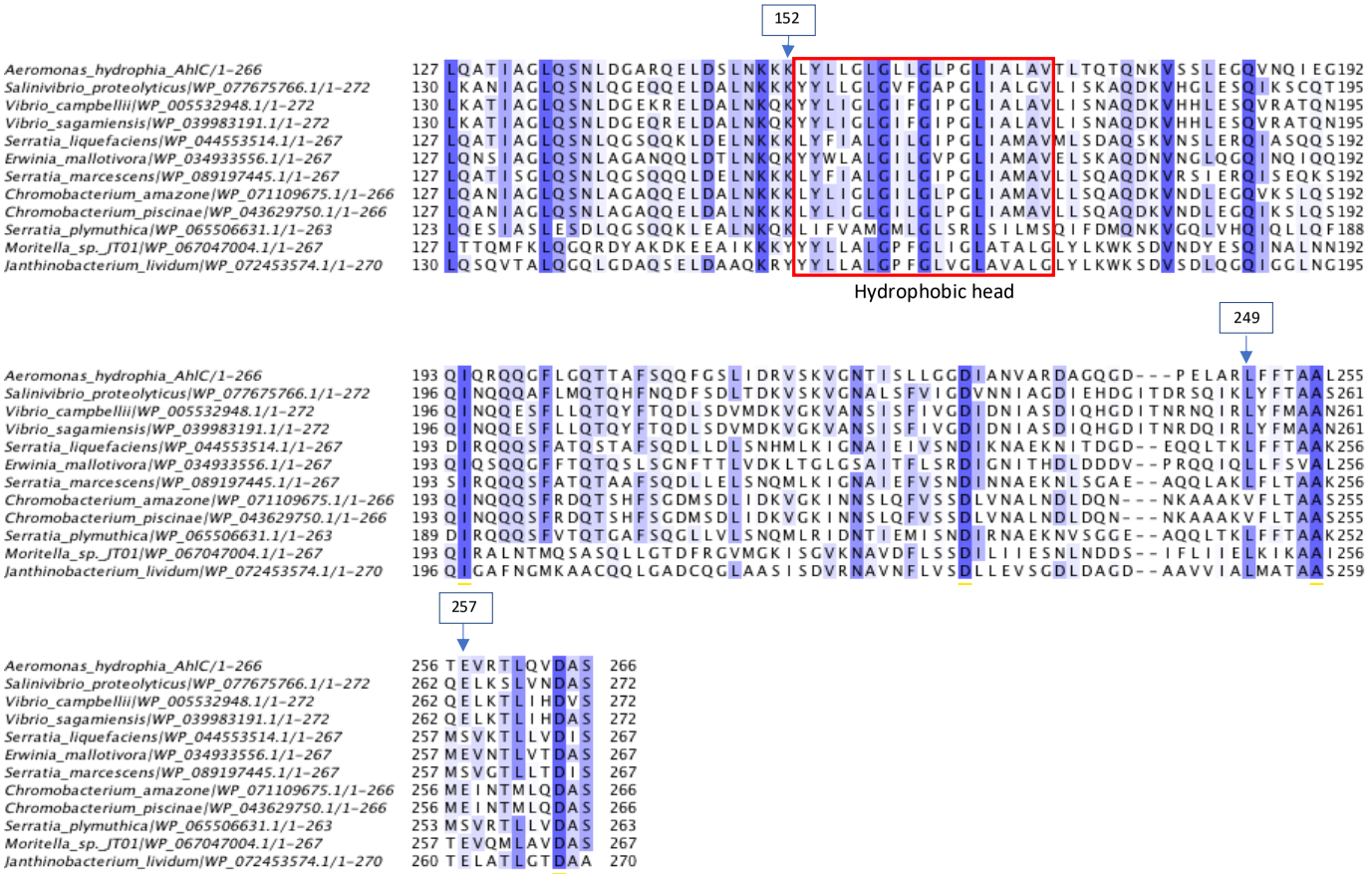
Supplementary Figure 9. Stereo images of an area of the 2Fo-Fc electron density maps of AhlC. An area of the density map (blue mesh), contoured at 1.0 σ, around alpha helices 3 and 4 of chain P from **(a)** AhlC Form 1, (G184-N197 alpha3 and Q127-Q135 alpha4) **(b)** AhlC Form 2 (L127-L145 alpha 3 and G177-R196 alpha 4)



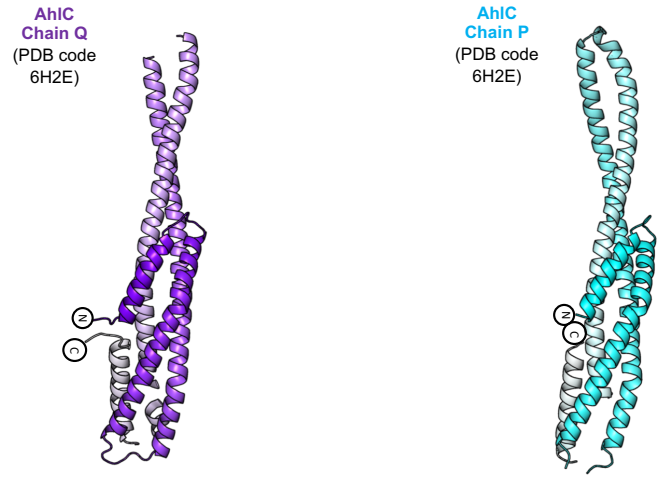
Supplementary Figure 10. Crystal packing of crystal Form 1 and crystal form 2 of AhlC shows flexibility in the head of AhlC. (a) B-factor putty diagram of crystal Form 1 and 2, low B-factor (blue) to high B-factor (red). Crystal form 1 has lower B-factors in the head domain and a more complete structure. Crystal form 2 has higher B-factors in the head domain and is incomplete showing high flexibility. Cartoon diagrams of the crystal packing of AhlC Crystal form 1 (**b**) and Crystal form 2 (**c**). The hydrophobic head (white) of Q and Q' (purple) in Crystal form 1 pack against neighboring hydrophobic heads in symmetry related molecules (**I**). In Crystal form 2 packing prevents hydrophobic interactions and results in disordered heads in Q and S (**II**).

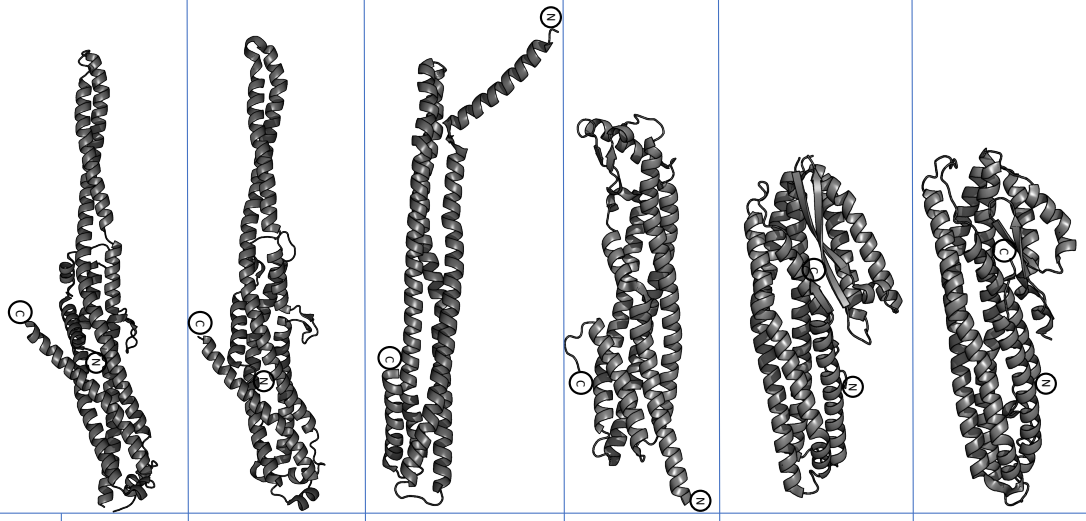


Supplementary Figure 11. AhlC head mutant L156T, L160T, L161T (AhlC^{HM}) purification and haemolytic assays. (**a**) 1μM AhlA and AhlB with AhlC, AhlC^{HM} , AhlC $^{cross\text{-}link\ 1}$ or AhlC $^{cross\text{-}link\ 2}$, were added to 0.25% (w/v) horse erythrocytes and incubated for 1hr at 37°C. Percentage haemolysis of 0.25% (w/v) horse erythrocytes by the AHL α-PFT was calculated by measuring absorbance of the supernatant at 542nm, all assays were done in triplicate. Below are SDS-Page gels of cross linked AhlC product in cross-link1 (glutaraldehyde) and cross-link2 (glutaraldehyde + 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride). All assays were carried out in triplicate (n=3) with values corresponding to the mean ± standard deviation. (**b**) Gel filtration UV trace chromatogram (i) and SDS-PAGE (ii),AhlC^{HM} forms a tetramer in solution and in the crystal structure (iii). Source data are provided as Source Data file.

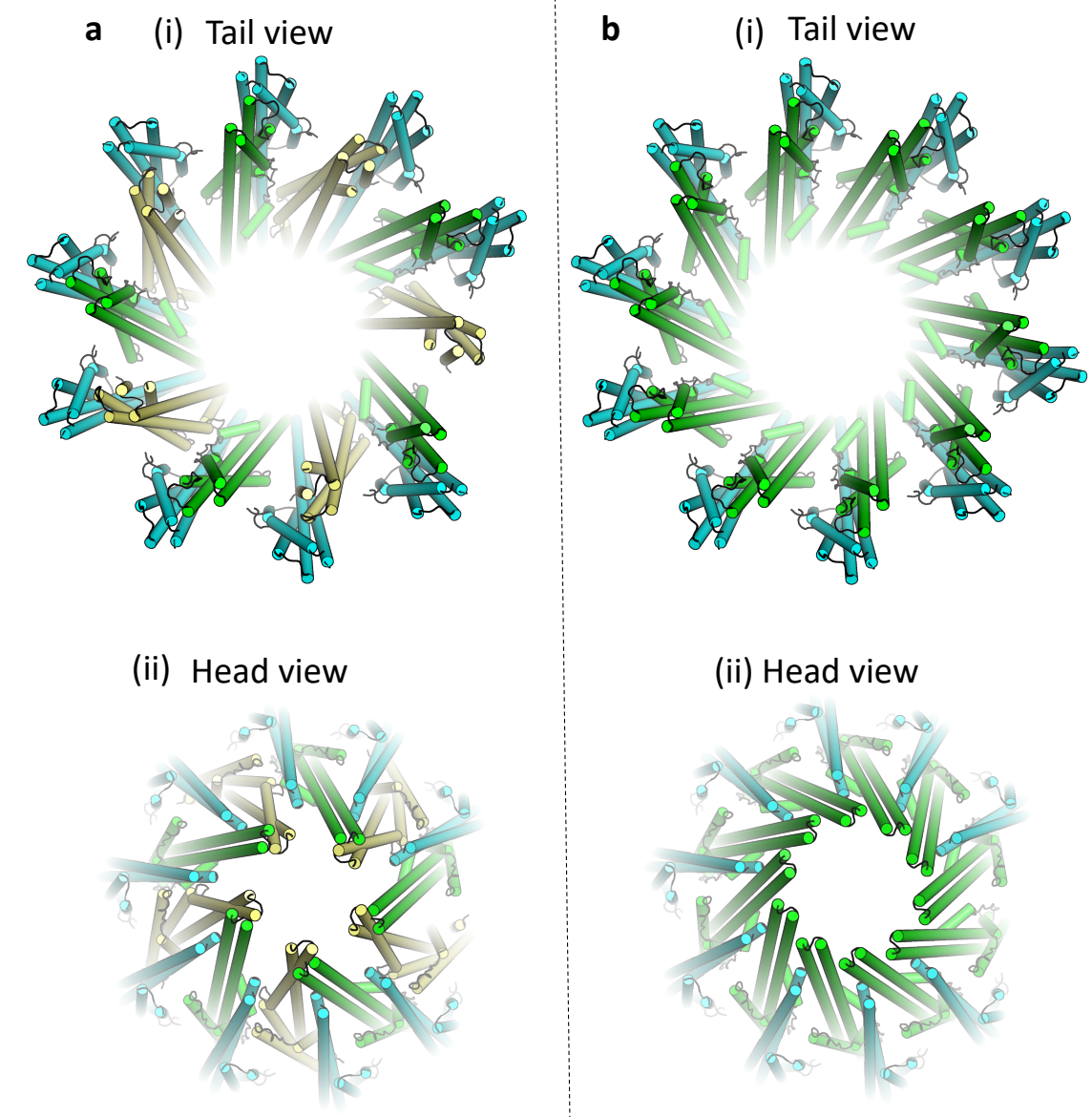


Supplementary Figure 12. Sequence analysis of AhlC. Sequence alignment, using Tcoffee [4], of AhlC against orthologous proteins from other Gram negative bacteria, identified using BlastP [5]. The hydrophobic head and residues K152, L249 and E257 in *A. hydrophila* AhlC are highlighted.

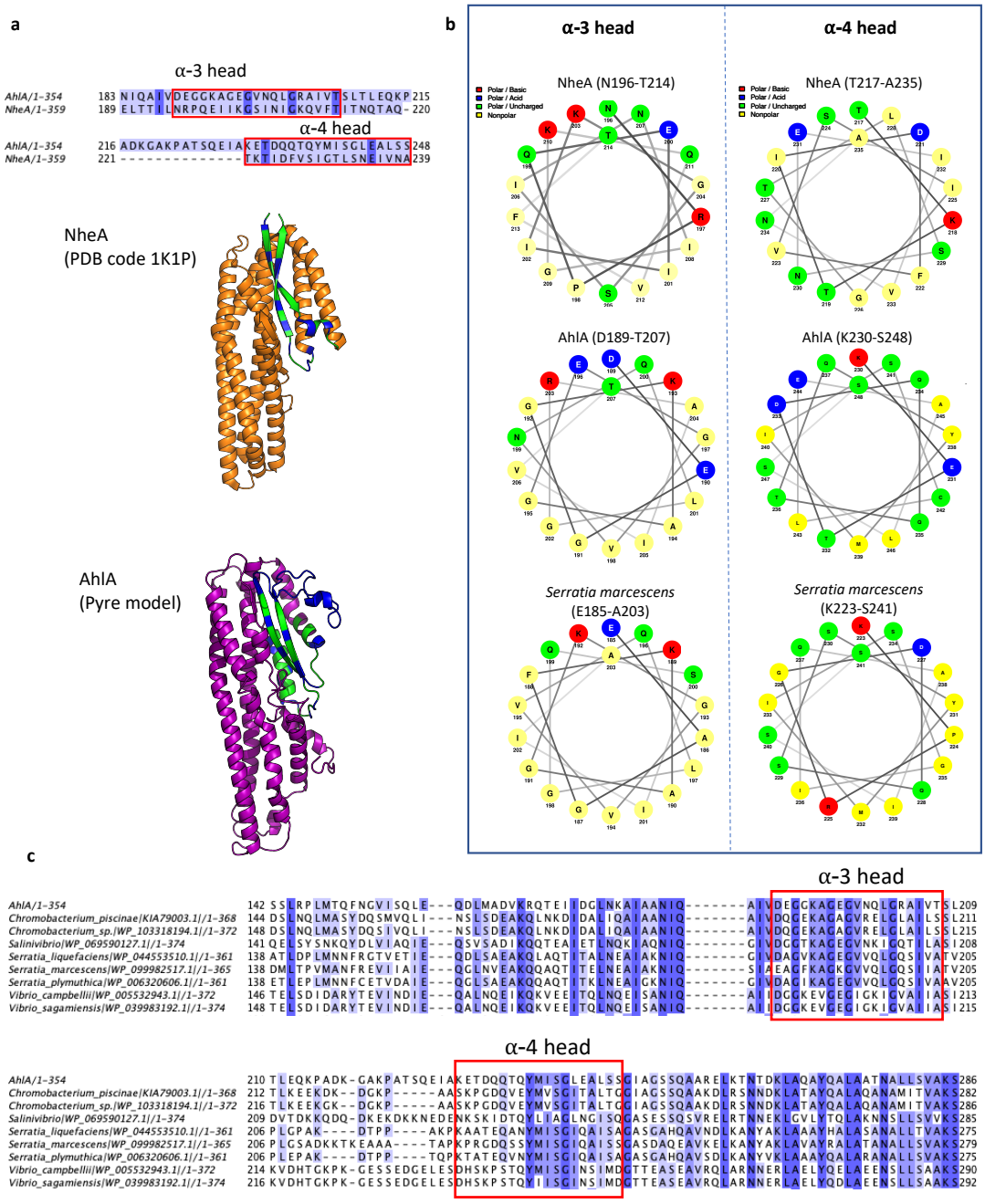


	YaxA (sol) (PDB code 6EK7)	XaxA (sol) (PDB code 6GY8)	ClyA(pore) (PDB code 2WCD)	ClyA (sol) (PDB code 1QOY)	NheA (sol) (PDB code 4K1P)	Hbl-B (sol) (PDB code 2NRJ)
						
AhlC Chain Q	Z-score 17.7 RMSD 2.8 Å % ID 12 %	Z-score 18.3 RMSD 2.7 Å % ID 11 %	Z-score 13 RMSD 3.6 Å % ID 8%	Z-score 9 RMSD 7.5 Å % ID 8 %	Z-score 14.2 RMSD 3.0 Å % ID 13 %	Z-score 14.3 RMSD 3.8 Å % ID 12%
AhlC Chain P	Z-score 13.4 RMSD 4.0 Å % ID 5 %	Z-score 17 RMSD 2.8 Å % ID 11 %	Z-score 13.2 RMSD 4.9 Å % ID 7%	Z-score 9 RMSD 5 Å % ID 10 %	Z-score 13.2 RMSD 2.9 Å % ID 13 %	Z-score 13.4 RMSD 3.5 Å % ID 13%

Supplementary Figure 13. Structural superposition of AhlC with ClyA α-PFT family toxins. Superposition of AhlC chain Q (purple) and AhlC chain P (blue) with ClyA family α-PFT's (grey) (aligned by the DALI server[6]), with Z-scores, RMSD(Cα) and percentage sequence identity (%ID).



Supplementary Figure 14. Modelling of AhlBC pore into YaxAB
(**a**) Alternating AhlB type1 (yellow) and type2 (green) were aligned with YaxB in the YaxAB pore, while AhlC (cyan) was aligned with YaxA. Clashes can be seen between AhlB type1 and AhlB type2 in both the tail (i) and the head domains (ii). (**b**) AhlB type2 aligned with YaxB to produce a homo-oligomeric AhlB pore, with AhlC aligned with YaxA. In this pore both the tails (i) and heads (ii) of AhlB pack without clashes.



Supplementary Figure 15. Sequence analysis of AhlA. (**a**) Alignment of NheA and AhlA β-tongues which are predicted to construct the α-3 and α-4 heads in the pore forms (upper), with Phyre2 [7] model of AhlA (pink, lower) compared to crystal structure of NheA (orange), both contain β-tongues (green hydrophobic, blue hydrophilic), which could form 2 extended amphipathic helices assuming a similar mechanism of unfolding to that seen in AhlB. (**b**) Helical wheels of predicted transmembrane region of the α-3 and α-4 heads in pore forms of NheA (PDB code 1K1P), AhlA, and *S. marcesecens* generated using NetWheels (Mól, A.R., Castro, M.S. and Fontes, W. NetWheels Tool. http://lbqp.unb.br/NetWheels/). In all three species the α-3 head has an amphipathic helix whith α-4 heads hydrophilic. (**c**) Sequence alignment, using Tcoffee[4], of AhlA against other Gram negative bacteria identified using BlastP [5]. Region D189-T207 (α-3 head) and K230-S248 (α-4 head) are highlighted.

Supplementary Table 1| Gram negative Bacteria identified from BlastP[5] searches as containing full tripartite PFT systems

Organism	AhlA homologue accession no. and sequence identity (sequence aligned)	AhlB homologue accession no. and sequence identity (sequence aligned)	AhlC homologue accession no. and sequence identity (sequence aligned)
Salinovibrio sp.	WP_069590127.1 43% (over 99%)	WP_069590129.1 77% (over 98%)	WP_069590131.1 51% (over 99%)
Erwinia mallotivora	WP_034933552.1 46% (over 99%)	WP_034933555.1 79% (over 99%)	WP_034933556.1 52% (over 100%)
Chromobacterium piscinea	WP_052247043.1 50% (over 98%)	WP_043629747.1 78% (over 97%)	WP_043629750.1 47% (over 100%)
Vibrio campbelli	WP_005532943.1 35% (over 98%)	WP_005532945.1 62% (over 99%)	WP_00532948.1 48% (over 99%)
Serratia marcescens	WP_099982517.1 44% (over 99%)	WP_09982518.1 62% (over 99%)	WP_09982519.1 43% (over 98%)
Serratia plymuthica	WP_006320606.1 43% (over 99%)	WP_043912873.1 62% (over 99%)	WP_06550663.1 38% (over 98%)
Serratia liquefaciens	WP_044553510.1 45% (over 99%)	WP_044553512.1 62% (over 99%)	WP_044553514.1 46% (over 100%)

Supplementary Table 2|Primers used for cloning (restriction sites underlined)	
Primer	**Primer sequence**
AhlA his forward	GGC GCT AGG TAC TAC ATA TGA CCA CGA TCG CCA CCC TGG
AhlA his reverse	ATC TAA AGC TTT AG TGG TGG TGG TGG TGG TGA GCG TCT GGC AGG ATG CC
AhlA link his forward	GGC GGA CAC CAC CAC CAC CAC CAC
AhlA link his reverse	TCC GCC AGC GTC TGG CAG GAT GCC
AhlB his forward	CAG CCC ATA TGA CCA ACG CAA CAA CCA TCA CCA TGG ACC AG
AhlB his reverse	ATC TCG AGG GCG GCC AGG CGC G
AhlC his forward	GGC GCT AGG TAC TAC ATA TGA GCA ACG GCA TTC TTT CC
AhlC his reverse	TAA TCC TCG AGT AG TGG TGG TGG TGG TGG GAA GCG TCC ACC TGC
AhlC HM SDM forward	GGAACTACTGGGCTGCCGGGCCTCATC
AhlC HM SDM reverse	CAGCCCAGTCAGGTAGAGTTTTTTCTTGTTGAGGGAGTC G

Description of Additional Supplementary Files

File name: Supplementary Data 1

Description: BlastP results for AhlA, AhlB, and AhlC with E value<0.01, Aeromonas Taxid excluded. ClyA was not identified by any of the BLAST searches due to very low sequence similarity to any of the Ahl components (E values of 3.7, 0.8 and 2.2 against AhlA, AhlB, AhlC, respectively).

To view file scan QR code or use link below:
https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-019-10777-x/MediaObjects/41467_2019_10777_MOESM4_ESM.xlsx

File name: Supplementary Movie 1

Description: A Bi-fold hinge action is required for the conformational change from soluble AhlB to Pore form AhlB. Soluble AhlB (starting image) undergoes a large-scale rearrangement about two hinges to form the Type 1 AhlB pore form. To illustrate this, the soluble AhlB structure (start of movie) was subjected to two rotations of 135º about axes defined through the alpha carbons of 155-250, to form intermediate 1, and 189-224 to form intermediate 2. This structure was then morphed into the AhlB type 1 crystal structure to form intermediate 3 (pause 1), and then morphed again into the AhlB type 2 structure (end of movie), during which α-3 and α4 untwist slightly, and the N and C terminal helices rearrange. The hydrophobic head of AhlB is coloured green, with the loop that protrudes from the distal side of the membrane coloured blue/green. Intermediate coordinate files were produced using Pymol [8] and LSQman [9].

To view file scan QR code or use link below:
https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-019-10777-x/MediaObjects/41467_2019_10777_MOESM5_ESM.mov

Appendix A

Extended methodology - Ahl toxin

This appendix is an overview of the details of work done that was omitted or was only briefly mentioned, from the results and discussion in Paper 1 (Section 3.3). This includes details of the expression and purification of AhlA, AhlB, AhlC and AhlC head mutant (AhlC HM), crystallisation and subsequent structure determination of the pore form and soluble AhlB, as well as the AhlC HM. It also outlines details of detergent trials used to select the best detergent for use in subsequent assays and EM studies of the AhlABC pore. Finally, details of sample optimisation for cryo-EM of the final pore are presented.

A.1 AhlA, AhlB, and AhlC construct design

Constructs for AhlA, AhlB, AhlC, and AhlC HM were provided by Jason Wilson, Department of Molecular Biology and Biotechnology, Sheffield. All constructs were full length, contained a C-terminal His_6 tag and generated from genes AHX33179.1 (*ahlA*), AHX33180.1 (*ahlB*) and AHX33181.1 (*ahlC*) from *A. hydrophila* strain AL09–71. Full details of cloning are found in Paper1 (Section 3.3).

A.2 Overexpression and purification of AhlA, AhlB and AhlC constructs

Protocols for overexpression of each construct had previously been established by Jason Wilson, also purification protocols for AhlB, AhlC and AhlC HM had been established and are described in Paper1 (Section 3.3). However previous attempts to purify AhlA resulted in a partially digested product and therefore a method was developed to produce stable full-length AhlA for assays.

A.2.1 Purification of AhlA

To try and improve the stability of AhlA a twofold approach was used. First EDTA free protease inhibitor was added to lysis buffer to reduce degradation by proteases, and secondly, 50 mM L-Arg and L-Glu were added to purified samples. Addition of L-Arg and L-Glu has been shown to reduce aggregation and increase the solubility of proteins by masking charged surface residues and also hydrophobic patches thus reducing surface interactions that drive aggregation. Addition of L-Arg and L-Glu have also been shown to reduce proteolytic cleavage and increase protein stability although the method by which this is achieved is as yet unclear (Golovanov et al., 2004).

2 g of cell paste was resuspended in 25 ml lysis buffer (50 mM Tris-HCl pH 8 and EDTA free cOmplete protease inhibitor (Roche) before lysis by sonication (3×20 s bursts at 16,000 nm λ). Insoluble material was removed by centrifugation at 40,000 g and soluble material was applied to a 5 ml Nickel Hi-trap column (GE Healthcare) in binding buffer (50 mM Tris-HCl pH 8 and 0.5 M NaCl). Protein was eluted on a gradient of binding buffer to 50 mM Tris-HCl pH 8, 0.5 M NaCl and 0.5 M imidazole. AhlA eluted at 0.19 M Imidazole in fractions 13-18 (Figure A.1A). 50 mM L-Glu and 50 mM L-Arg were added to all fractions containing AhlA before concentrating down to 2 ml using a Vivaspin 30

kDa MWCO concentrator (Sartorius). This sample was loaded on to a Superdex 200 pg column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 and 0.5 M NaCl. AhlA eluted at a volume of 45 ml (fractions 3-7) as a large molecular weight complex, (501 kDa), again 50 mM L-Glu and 50 mM L-Arg were added to all fractions containing AhlA. To assess the stability of AhlA after purification by size exclusion chromatography, samples of each fraction were run on SDS-PAGE. In all fractions, a single band of molecular weight 40 kDa was observed (corresponding to the molecular weight of AhlA) (Figure A.1B), in contrast to previous purifications of AhlA which gave two distinct peaks of 40 kDa and 24 kDa (Figure A.1C). All fractions containing AhlA were pooled and concentrated down to a final concentration of 4 mg/ml in a total volume of 200 μl for assays. Although the yield of soluble AhlA was still low this new purification method produced a stable undigested AhlA, which could be used for biochemical assays.

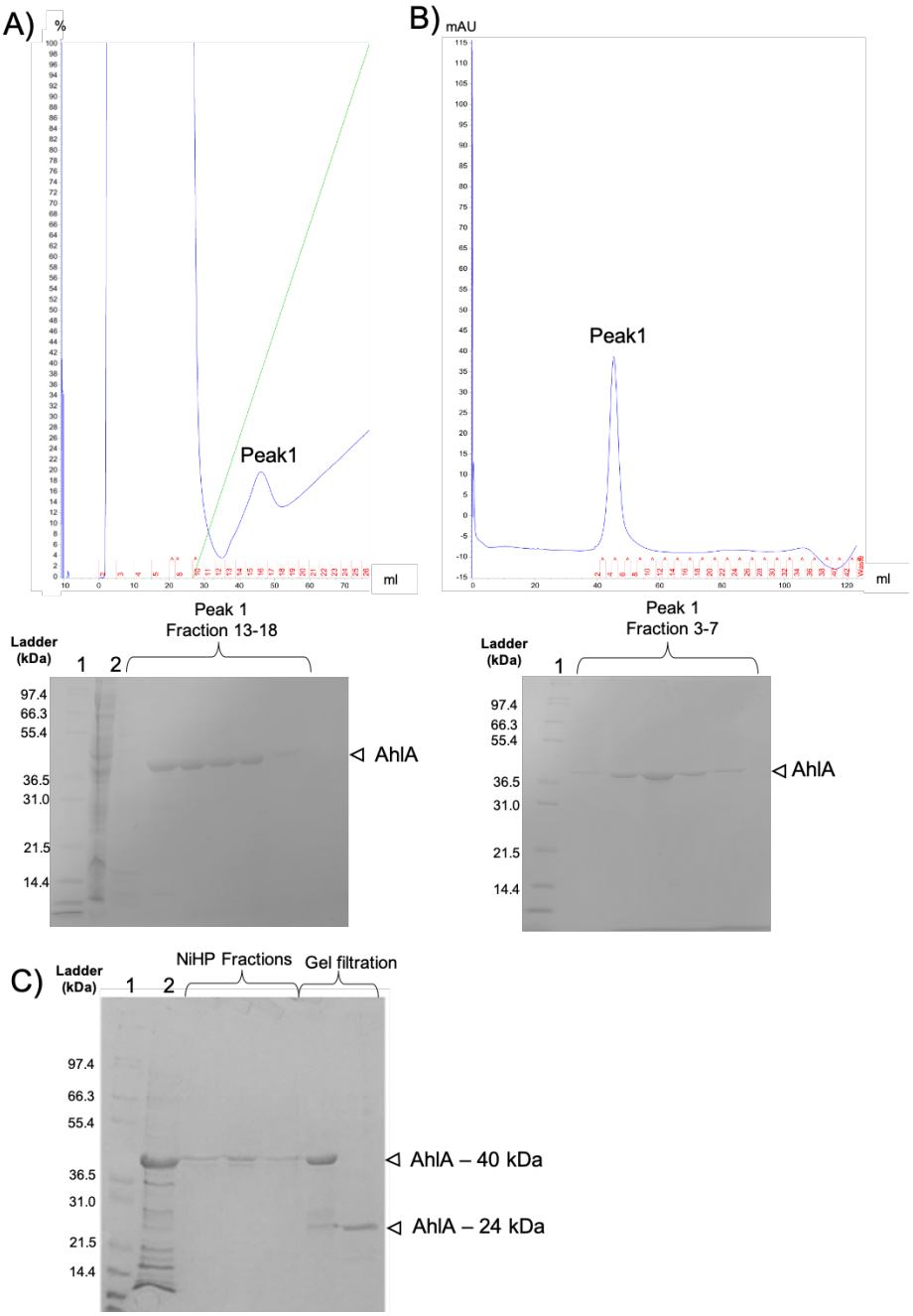


Figure A.1 **Purification of AhlA.** UV trace chromatogram (blue line) (top) and SDS-PAGE (below). A) Purification by NiHP. AhlA eluted at 0.19 M imidazole (Green line, imidazole gradient). Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and NiHP load respectively. B) Gel Filtration for AhlA with 50 mM L-Glu + L-Arg. AhlA eluted as a single peak at a volume of 45 ml. Lane 1 of the SDS-PAGE shows Mark 12 ladder. C) SDS-PAGE analysis of samples from NiHP and Gel-filtration of AhlA without the addition of L-Glu and L-Arg shows after Gel filtration, AhlA has been cleaved into 2 fragments of MWs 40 kDa and 24 kDa. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and NiHP load respectively.

A.3 Crystallisation, data collection and structure determination of AhlB

Crystallisation trials of AhlB pore

Selenomethionine derived AhlB was purified and concentrated before being subject to crystallisation trials as described in Section 2.5.1 and Paper1 (Section 3.3). Trials were carried out in a 96 well sitting drop plate. Plates were then incubated at 17 ^{o}C and checked regularly.

Crystals grew in five conditions shown in Table A.1. All conditions contained MPD and either contained buffer or no buffer. All crystals were of sufficient size for subsequent data collection (Figure A.2), and as such were mounted on litholoops (Molecular Dimensions) and preserved in liquid nitrogen as described in Section 2.5.4.

The largest crystals from JCSG+ A11 were harvested and used for seeding in optimisations of the A11 crystallisation condition as described in Section 2.5.3 and Paper1 (Section 3.3). These produced larger single crystals in 60 % MPD, 0.2 M ammonium phosphate and 0.1 M Tris pH 8.5, which were mounted on litholoops (Molecular Dimensions) and preserved in liquid nitrogen (described in Section 2.5.4), as shown in Figure A.2F.

A.3.1 X-ray data collection of AhlB pore

As discussed in Paper1 (Section 3.3) and Section 2.5.9, selenomethionine incorporated AhlB protein was used in crystallisation of AhlB to provide heavy atoms for initial phasing.

Data was collected from a single crystal at Diamond Light Source beamline i03 as described in Section 2.5.7, with full data processing statistics given in section 3.3 (Paper1, Table1). First, a selenium edge scan was done to determine if selenium had been incorporated into the protein and to find the peak and inflection energy for data collection (Figure A.3A). To determine the % transmission to use, a test collection was done at peak energy (12659 eV).

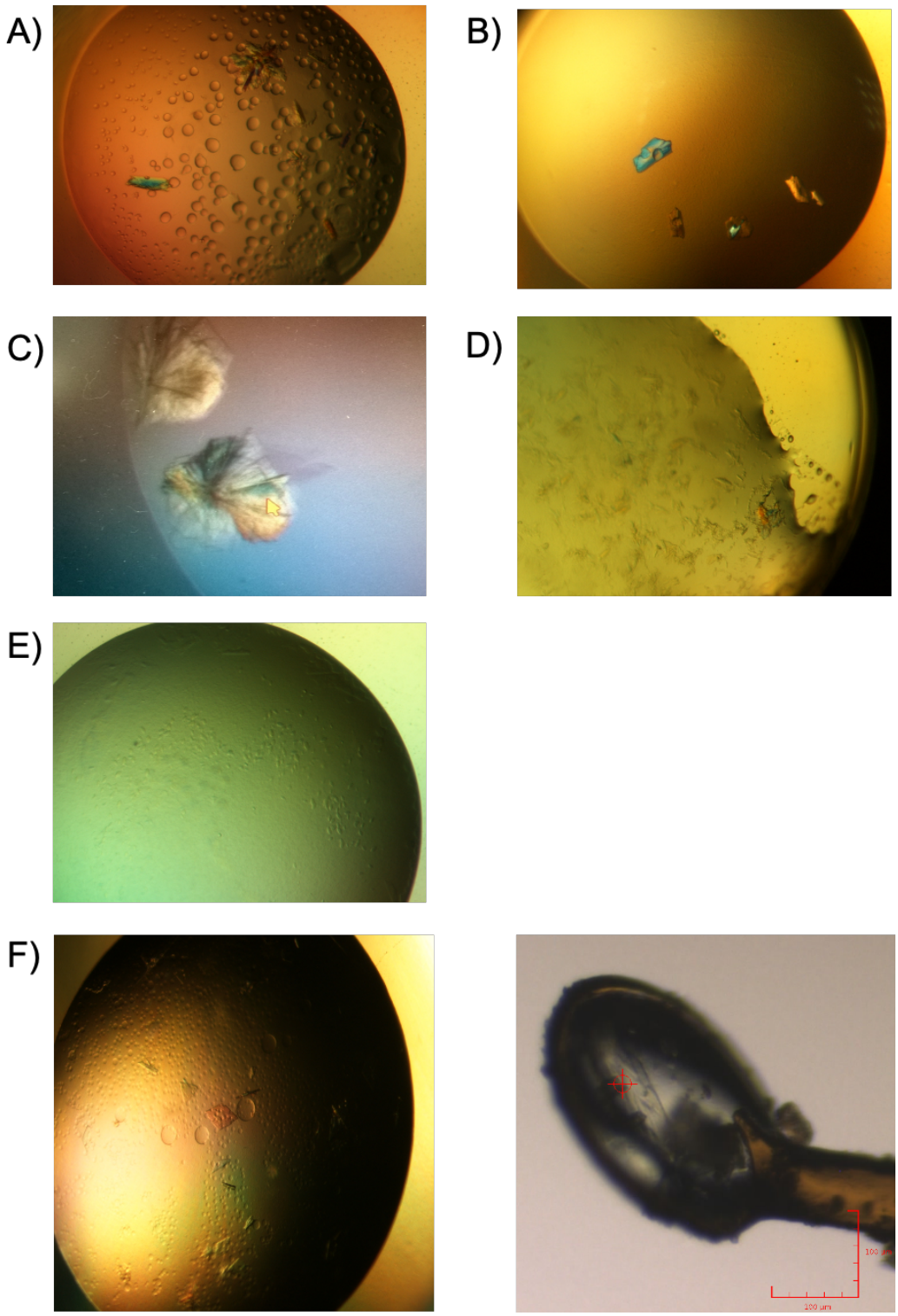


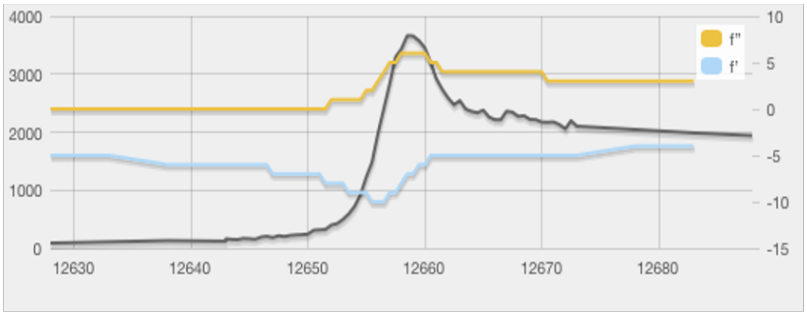
Figure A.2 **Crystals of AhlB Semet.** A) MPD H2, B) MPD H3, C) JCSG A11, D) MPD D6, E) MPD D11 F) Optimised AhlB crystals grown in micro bridges from seeding (left), mounted crystal in litholoop (right).

Table A.1 Initial crystals hits for AhlB Semet grown in 98 well sitting drop plates, (crystallisation conditions, and resolution is given for each crystal.)

Crystallisation condition	Salt	Buffer	Precipitant	Diffraction
MPD H2	0.5 M Ammonium sulfate	0.1 M HEPES pH7.5	30 % MPD	6.8 Å
MPD H3	0.2 M Na citrate	0.1 M HEPES pH8.5	30 % MPD	6.5 Å
JCSG A11	0.2 M Ammonium phosphate	0.1 M Tris pH8.5	50 % MPD	7 Å
MPD D6	0.2 M K phosphate	N/A	40 % MPD	-
MPD D11	0.2 M K citrate	N/A	40 % MPD	-

Five images were collected with oscillation 0.5 ^{o}, exposure time of 0.5 s and transmission of 20 % and 30 %. Finally, a dataset of 3600 images was collected with an oscillation per image of 0.1 ^{o} (total 360 ^{o}), and transmission of 30 %. An example of a diffraction image is provided in (Figure A.3B.), showing diffraction out to 3.2 Å. Analysis of the self rotation function in space group C2 at Chi=180 ^{o} showed peaks at omega=0 ^{o} and phi intervals of 36 ^{o}, peaks were also present at Chi=72 ^{o} omega=0 ^{o} and Chi=36 ^{o} omega=0 ^{o}, consistent with 10-fold non-crystallographic symmetry (Figure A.4). The unit cell parameters of a=363.6 Å, b=116.5 Å, c=217.4 Å, α=90 ^{o}, β=118 ^{o}, γ=90 ^{o} suggested a solvent content of 70 % would give an asymmetric unit content of 10 molecules, as determined by the Matthews calculation (Matthews, 1968) (Figure A.5). As AhlB contains 7 methionine residues in its sequence these values indicate that 70 Selenomethionines were present in the asymmetric unit, and therefore 70 Se atoms were searched for in the initial substructure determination.

A)



	Wavelength	f'	f''
Peak	0.9794 Å	-7.25e	6.47
Inflection	0.9796 Å	-10.25e	3.65

B)

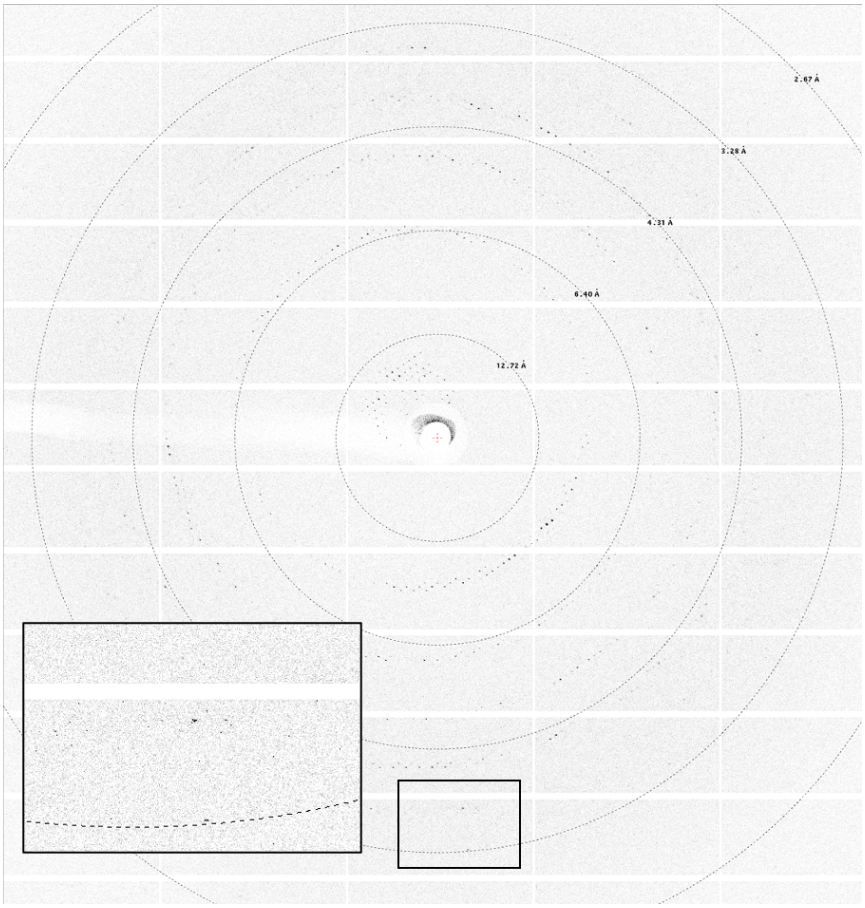


Figure A.3 **AhlB SeMet SAD data collection.** A) Fluorescence scan of AhlB SeMet crystal at Se K edge carried out with 1s exposure and 3.2 % transmission. The black line shows fluorescence at each wavelength, the blue line plots f' and yellow f''. Calculated using CHOOCH (Evans and Pettifer, 2001). B) A representative 0.1 ^{o} oscillation image from AhlB Semet. An enlarged view of the region highlighted by the square shows diffraction extends to around 3.2 Å.

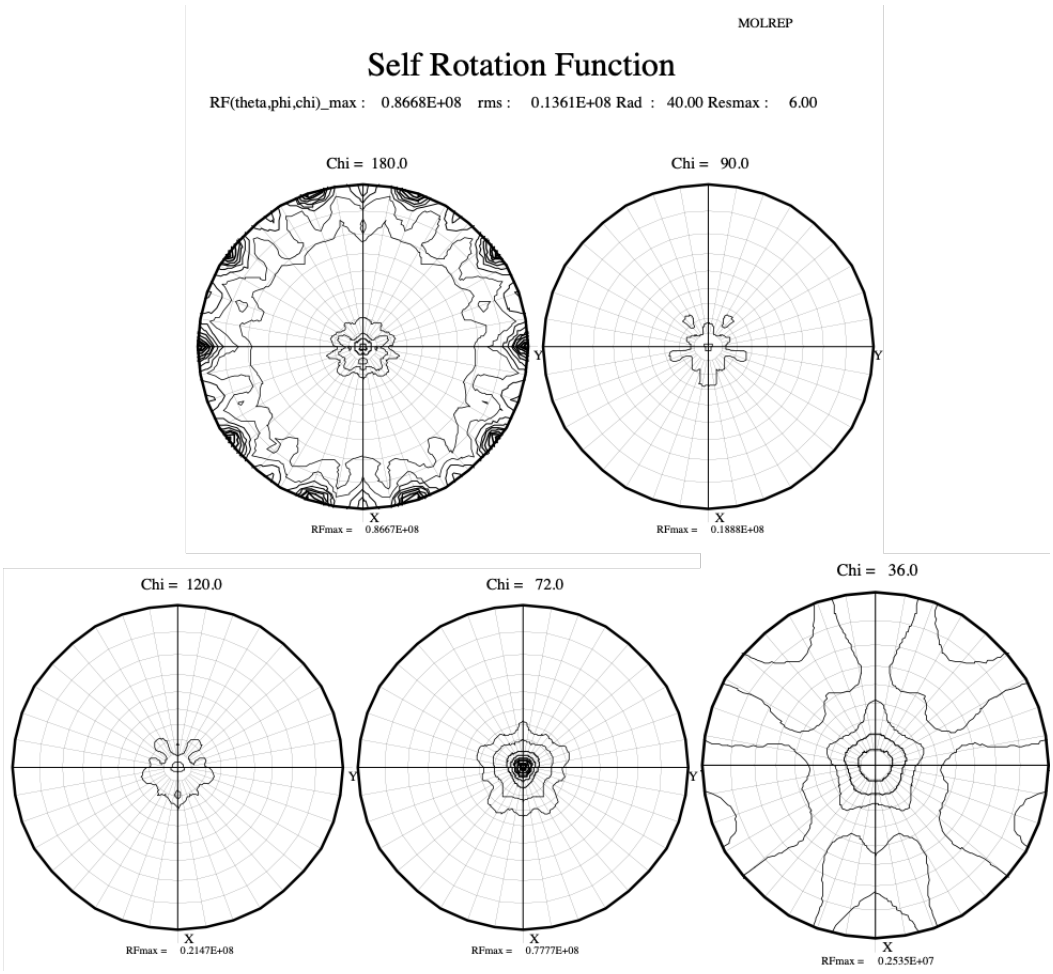


Figure A.4 **Self-rotation plots of AhlB Semet in space group C2.** Self-rotation plots for Chi=180 º, Chi=90 º, Chi=120 º, Chi=72 º and Chi= 36 º generated with a high res cut off =6.0 Å. At Chi=180 º peaks can be seen at omega=0 º and phi intervals of 36 º.

A)

N(mol)	Prob(N) for resolution	Prob(N) overall	Vm A**3/Da	Vs % solvent	Mw Da
1	0.0000	0.0000	52.54	97.66	38689.00
2	0.0000	0.0000	26.27	95.32	77378.00
3	0.0000	0.0000	17.51	92.98	116067.00
4	0.0000	0.0000	13.14	90.64	154756.00
5	0.0000	0.0001	10.51	88.29	193445.00
6	0.0003	0.0007	8.76	85.95	232134.00
7	0.0008	0.0014	7.51	83.61	270823.00
8	0.0016	0.0025	6.57	81.27	309512.00
9	0.0028	0.0040	5.84	78.93	348201.00
10	0.0045	0.0058	5.25	76.59	386890.00
11	0.0079	0.0096	4.78	74.25	425579.00
12	0.0118	0.0141	4.38	71.91	464268.00
13	0.0172	0.0197	4.04	69.57	502957.00
14	0.0221	0.0248	3.75	67.23	541646.00
15	0.0278	0.0301	3.50	64.88	580335.00
16	0.0393	0.0411	3.28	62.54	619024.00
17	0.0514	0.0527	3.09	60.20	657713.00
18	0.0641	0.0649	2.92	57.86	696402.00
19	0.0793	0.0794	2.77	55.52	735091.00
20	0.0918	0.0906	2.63	53.18	773780.00
21	0.0996	0.0974	2.50	50.84	812469.00
22	0.1032	0.1001	2.39	48.50	851158.00
23	0.0998	0.0964	2.28	46.16	889847.00
24	0.0892	0.0859	2.19	43.82	928536.00
25	0.0720	0.0692	2.10	41.47	967225.00
26	0.0493	0.0475	2.02	39.13	1005914.00
27	0.0289	0.0278	1.95	36.79	1044603.00
28	0.0168	0.0162	1.88	34.45	1083292.00
29	0.0093	0.0090	1.81	32.11	1121981.00
30	0.0045	0.0044	1.75	29.77	1160670.00
31	0.0025	0.0024	1.69	27.43	1199359.00
32	0.0011	0.0010	1.64	25.09	1238048.00
33	0.0007	0.0007	1.59	22.75	1276737.00
34	0.0004	0.0004	1.55	20.40	1315426.00
35	0.0000	0.0000	1.50	18.06	1354115.00
36	0.0000	0.0000	1.46	15.72	1392804.00
37	0.0000	0.0000	1.42	13.38	1431493.00
38	0.0000	0.0000	1.38	11.04	1470182.00

B)

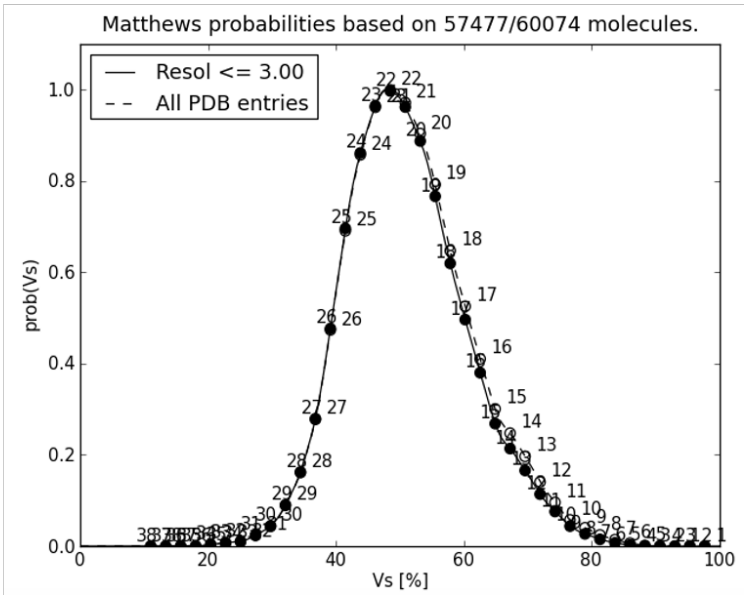


Figure A.5 **AhlB Semet (space group C2) Matthews calculation** A)Table of the possible number of molecules in the AU, with respective probability and percentage solvent content, suggesting 22 molecules is the AU. B) A plot of the probability against the solvent content. Graph and table generated using online software at http://www.ruppweb.org/mattprob/ (Matthews, 1968).

A.3.2 Structure determination, model building and validation of AhlB pore

The AhlB Selenomethionine data set had an anomalous mid-slope of 1.025 as shown in Section 3.3(Paper1 Table1). Initial phases were calculated using SHELX (Sheldrick, 2010). SHELXC showed that anomalous signal was present (>0.8 d"/sig) out to 3.36 Å. SHELXD found 98 Selenium atom sites, within the range expected for 10 molecules of AhlB. Phase calculation, density modification and heavy atom refinement using SHELXE reduced this to 54 selenium atom sites (Figure A.6A). Inspection of the sites showed strong density for 50 of the 54 sites, 20 less than the expected value of 70. The missing sites were all located in the first 13 N-terminal residues. These heavy atom sites and a solvent content of 76 %, were used to calculate phases and run density modifications in SHELXE and the best density map was produced for the inverted hand. Manual examination of the map for each hand against heavy atom coordinates confirmed that the inverted hand was correct (Figure A.6B), and initial phases were enough to see continuous density of helices within the protein (Figure A.6B).

Phenix "Find Helices and Strands" (Adams et al., 2010), was used in initial model building to find and place alpha helices at a resolution of 5 Å. Density modification using this model and a solvent content of 76 % was carried out using the software Parrot (Cowtan, 2010). After two rounds of building using Phenix, and three rounds of refinement by Parrot, a FOM=0.69 was achieved and the distinctive shape of a pore could be observed (Figure A.7A).

After initial placement of helices and density modification using Phenix and Parrot, five further rounds of automatic building and refinement were carried out using Buccaneer (Cowtan, 2010). The final cycle of building in Buccaneer had placed 4363 residues in 210 fragments, with the longest fragment being 145 residues. The final cycle of refinement gave an R-factor of 0.41 and Free R of 0.47. This produced a model with almost complete sequence for each of the 10 AhlB molecules in the structure and allowed the NCS in the AU to be easily determined by eye. It was immediately apparent that each pore was constructed

from 5 dimers of AhlB, with each dimer being made up of two conformations of AhlB. This gives the pore a distinctive 5-fold NCS. This discovery made it possible to use NCS in model building. One dimer was chosen for manual model building using Coot and then superimposed onto the other dimers using the LSQ function of Coot (Emsley et al., 2010). Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.22 and Free R of 0.24 (Figure A.7B), full validation details can be found Paper1 Table 1. Analysis of the crystal packing clearly showed the two-fold symmetry down the B cell axis and the 10 and 5-fold non-crystallographic symmetry of the pore down the C^* cell axis, as well as large solvent channels as predicted from a solvent content of 70 %. Hydrophobic heads of each subunit pack against heads of neighbouring pores, providing large hydrophobic interfaces to stabilise the crystal packing, likewise large hydrogen bonding interfaces between neighbouring tails further stabilize packing (Figure A.7C).

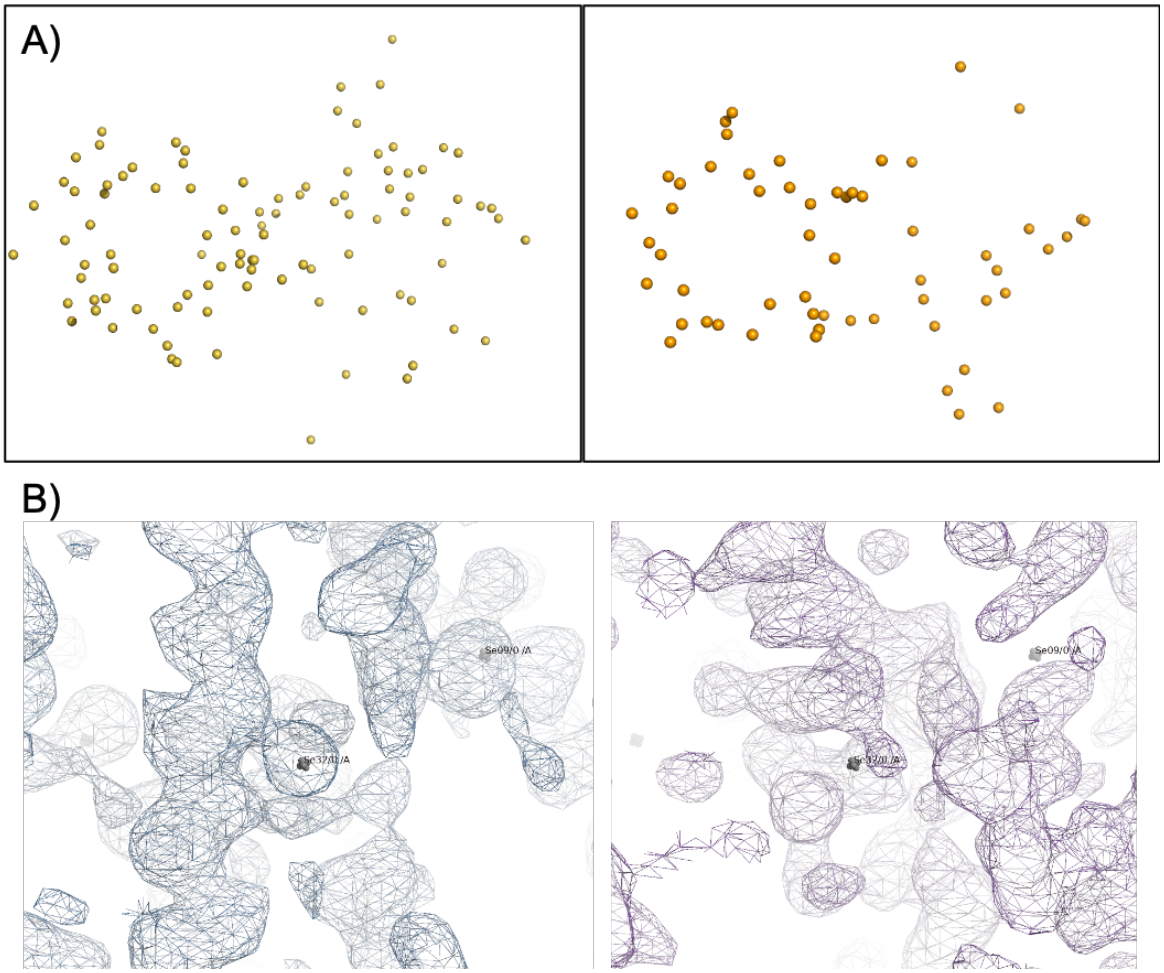


Figure A.6 **Visual analysis of density maps from SHELX (Sheldrick, 2008) for AhlB Semet.** A) Selenium atom co-ordinates from ShelxD (left) and after density modification by ShelxE (right). ShelxD placed 98 Se atoms. After density modification by ShelxE 44 atoms were removed. B) 2Fo-Fc maps for both hands generated from the initial phase determination of AhlB by SHELXE. Se atoms fit density for Inverted hand (left) and continuous density for helices was visible, as such the inverted hand was chosen.

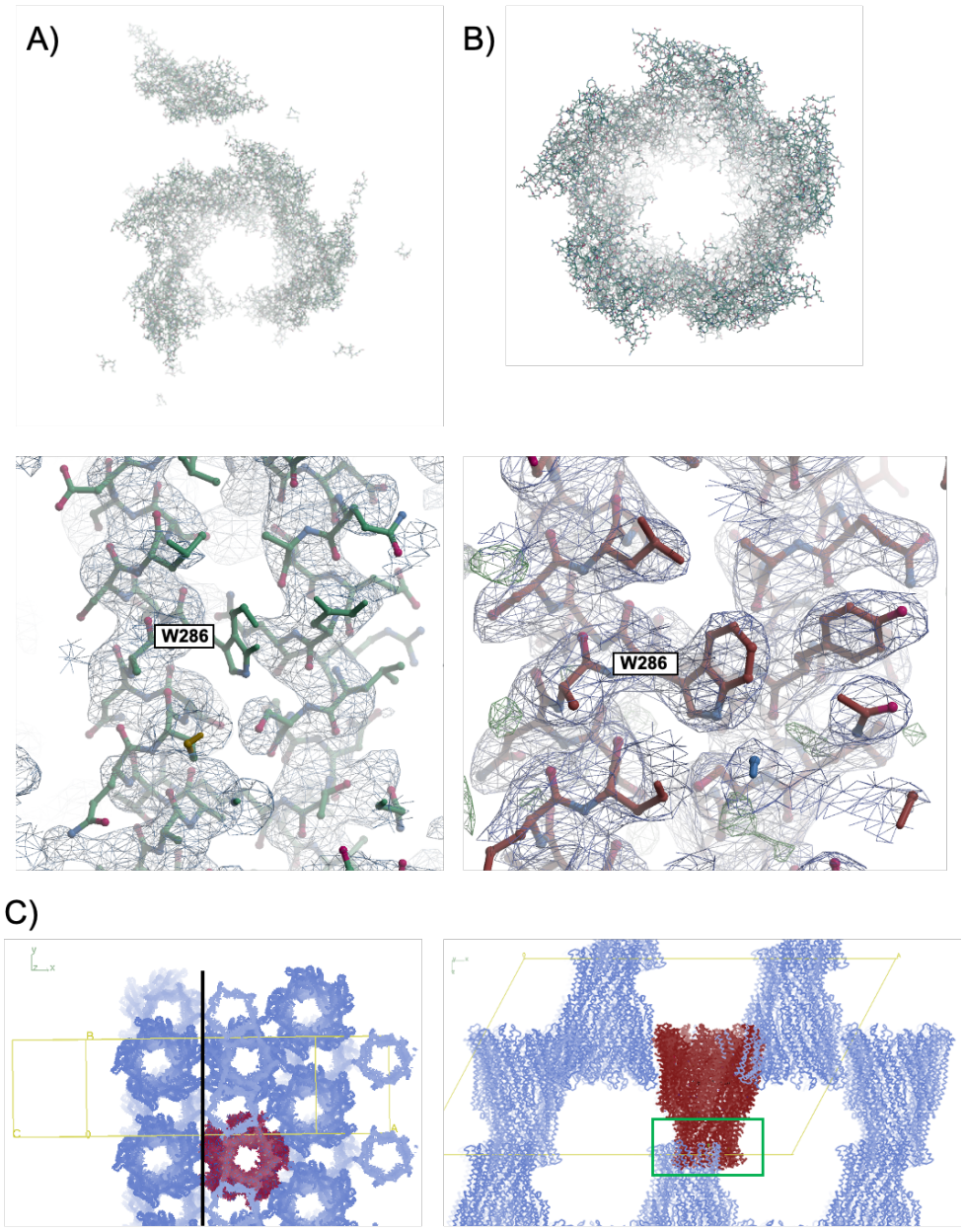


Figure A.7 **Initial and final models and electron density for AhlB after rebuilding and refinement.** 2Fo-Fc map contoured at 1.0 σ (blue mesh) and positive difference map contoured at 2.99 σ (green, positive, and red, negative, mesh). A) Model (Top) and density map around Trp 286 (Below) after building and density modification with Phenix find helices and Parrot. B) Final model after validation (top) and density around Trp 286, shows great improvement with clear density for Trp286 and surrounding residues. C) Symmetry (Blue) and unit cell (yellow lines) shown with final refined structure (red). Left, view looking down C^*, clearly shows two-fold symmetry down the B axis (black line), and the 10 and 5-fold non-crystallographic symmetry of the pore down the C^* axis. Right, Hydrophobic heads of pores pack against neighbouring heads within the crystal (green box). Large solvent channels are seen explaining high solvent content when estimating the number of molecules in the AU.

A.3.3 Interface analysis of AhlB pore

The structure of the AhlB pore was analysed using the ePISA server (Krissinel and Henrick (2007)) to identify the interfaces between monomeric AhlB subunits within the pore complex. Two unique interfaces were identified between subunits (Interface 1, A<>B, C<>D, E<>F, G<>J, H<>I) and (Interface 2, F<>A, B<>H, J<>C, D<>E, I<>G). The first of these interfaces represents the interface between monomers within the five dimers, while the second represents the interface between monomers of neighbouring dimers, as discussed in Section 3.3 (Paper 1). Both interfaces can be divided into three main regions, one in the tail, one in the neck and one in the head. The tail and neck regions contain the interface hydrogen bonds and salt bridges, while the head contains a large hydrophobic surface interface (Figure A.8 and Figure A.9). The first interface has the greatest interface area (2800 $Å^2$, 66 % of the total buried surface area), and interface residues (27 %) showing a stronger interface between the two dimers than between neighbouring dimers (TableA2), however, both interfaces involve a large Interface area, have a negative delta G and binding energy and both are involved in forming a stable biological assembly (Table A.2).

A)

Hydrogen bonds

	Monomer 1	Length	Monomer 2
1	B:GLN 19 [NE2]	3.1	C:ALA 337 [O]
2	B:GLN 19 [NE2]	2.9	C:GLY 338 [O]
3	B:ARG 125 [NH2]	3.3	C:ASN 67 [OD1]
4	B:ARG 125 [NH2]	3.1	C:LEU 66 [O]
5	B:SER 139 [OG]	2.7	C:GLN 32 [OE1]
6	B:SER 140 [OG]	3.2	C:GLN 33 [O]
7	B:GLN 143 [NE2]	3.5	C:GLN 32 [O]
8	B:ASN 284 [ND2]	2.8	C:MET 336 [O]
9	B:ASN 284 [ND2]	3.1	C:GLN 24 [OE1]
10	B:SER 291 [OG]	2.9	C:THR 75 [OG1]
11	B:ASP 292 [N]	3.5	C:SER 78 [OG]
12	B:GLN 335 [NE2]	3.0	C:ALA 337 [O]
13	B:ASN 284 [OD1]	2.6	C:GLN 24 [NE2]
14	B:THR 280 [OG1]	2.6	C:ASN 28 [ND2]
15	B:ASN 133 [OD1]	2.6	C:LYS 31 [NZ]
16	B:SER 288 [O]	3.8	C:SER 78 [OG]
17	B:ASP 292 [OD1]	2.6	C:SER 78 [OG]
18	B:ASP 299 [OD1]	2.8	C:ASN 82 [ND2]
19	B:ASP 309 [O]	3.0	C:GLN 106 [NE2]
20	B:ASP 309 [OD2]	2.6	C:GLN 106 [NE2]
21	B:ASP 156 [OD1]	3.2	C:GLN 242 [NE2]
22	B:GLU 254 [OE2]	3.7	C:GLN 242 [NE2]
23	B:ASP 156 [OD1]	2.6	C:GLN 246 [NE2]
24	B:LYS 152 [O]	3.3	C:GLN 246 [NE2]
25	B:SER 151 [O]	2.6	C:SER 249 [OG]

Salt bridges

	Monomer 1	Length	Monomer 2
1	B:LYS 169 [NZ]	3.4	C:ASP 175 [OD1]
2	B:LYS 169 [NZ]	3.8	C:ASP 175 [OD2]
3	B:ASP 309 [OD2]	3.7	C:ARG 105 [NH1]
4	B:ASP 309 [OD1]	3.0	C:ARG 105 [NH1]
5	B:ASP 166 [OD2]	2.9	C:ARG 241 [NH1]
6	B:ASP 166 [OD2]	3.2	C:ARG 241 [NH2]
7	B:ASP 166 [OD1]	3.8	C:ARG 241 [NH2]
8	B:ASP 292 [OD2]	2.8	C:LYS 333 [NZ]
9	B:ASP 328 [OD1]	3.0	C:LYS 333 [NZ]

B)

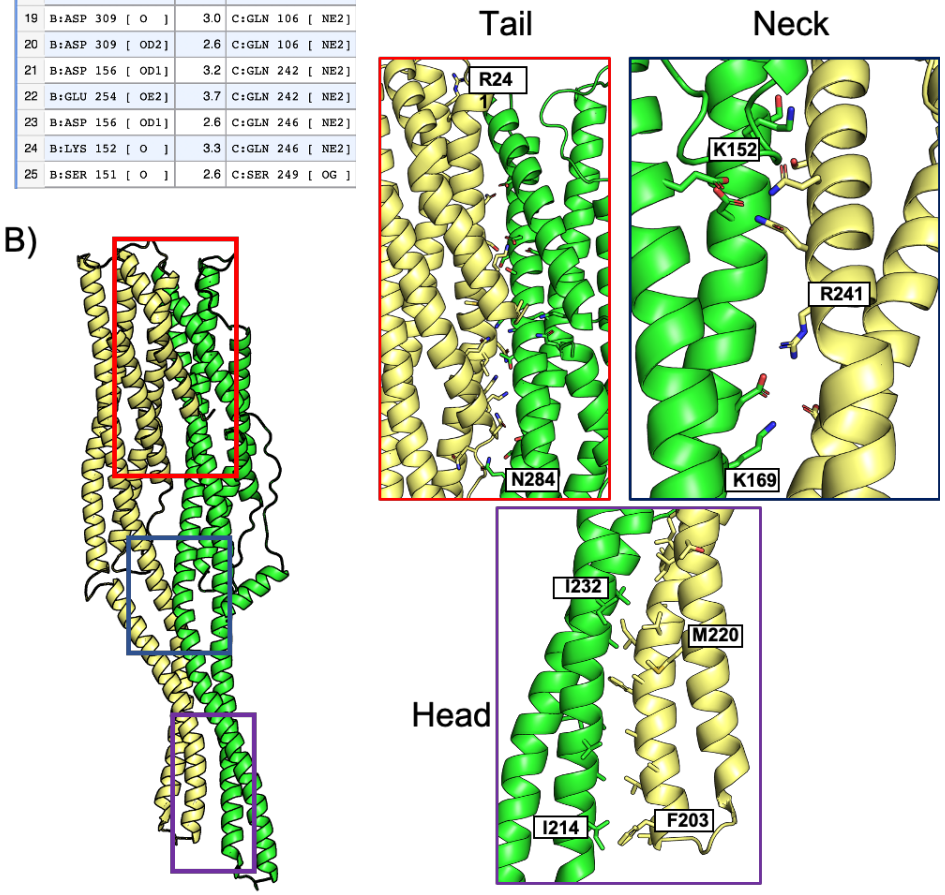


Figure A.8 **ePisa (Krissinel and Henrick, 2007) analysis of interface 1 within the AhlB pore.** A) Table of hydrogen bonds and salt bridges between Chain A (AhlB type 2 conformation) and B (AhlB type 1 conformation). Each table contains residue number, interacting atoms and bond length. B) Dimer of AhlB chain A (green) and chain B (yellow). Red, blue and purple boxes highlight interface residues in the tail, neck and head regions respectively.

A)

Hydrogen bonds

	Monomer 1	Length	Monomer 2
1	H:LYS 31 [NZ]	2.4	C:GLN 19 [OE1]
2	H:GLN 40 [NE2]	3.3	C:GLN 154 [OE1]
3	H:GLN 70 [NE2]	2.7	C:GLN 11 [OE1]
4	H:GLN 242 [NE2]	3.0	C:ALA 161 [O]
5	H:TYR 245 [OH]	3.0	C:ASP 168 [OD1]
6	H:TYR 245 [OH]	3.7	C:ASN 164 [O]
7	H:TYR 245 [OH]	3.1	C:ASP 168 [OD2]
8	H:ASN 28 [OD1]	3.4	C:GLN 19 [NE2]
9	H:LYS 31 [O]	3.1	C:ASN 274 [ND2]
10	H:GLN 32 [OE1]	3.2	C:GLN 271 [NE2]
11	H:LEU 66 [O]	3.5	C:GLN 281 [NE2]
12	H:ASN 79 [OD1]	2.7	C:ASN 3 [ND2]
13	H:ASN 82 [OD1]	3.2	C:ASN 3 [ND2]
14	H:VAL 339 [O]	3.0	C:GLN 335 [NE2]

Salt bridges

	Monomer 1	Length	Monomer 2
1	H:ARG 241 [NH2]	3.2	C:ASP 168 [OD1]

B)

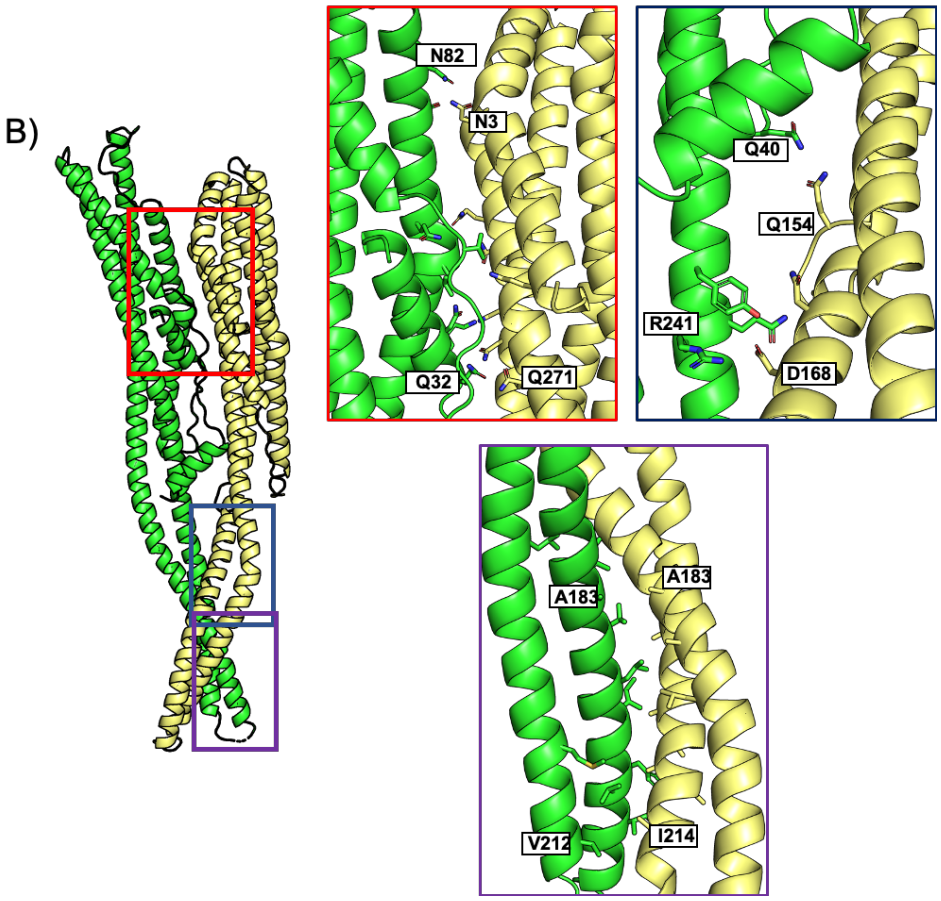


Figure A.9 **ePisa (Krissinel and Henrick, 2007) analysis of Interface 2 within the AhlB pore.** A) Table of hydrogen bonds and salt bridges between Chain B (AhlB type 1 conformation) and H (AhlB type 2 conformation). Each table contains residue number, interacting atoms and bond length. B) Dimer of AhlB chain H (green) and chain B (yellow). Red, blue and purple boxes highlight interface residues in the tail, neck and head regions respectively.

Table A.2 Analysis of the two interfaces between subunits in the AhlB pore. ePISA (Krissinel and Henrick, 2007) results for AhlB pore dimer Interface 1 and Interface 2.

ePISA results	Interface 1	Interface 2
Interface surface area (Å^2)	2800	1500
Interface residues (%)	27	17
Delta G	-37.9	-14.9
Binding energy	-51.4	-20.6

A.4 Model building and validation of soluble AhlB

Having solved the pore structure of AhlB, previous data sets collected from other crystal forms of AhlB by Jason Wilson were revisited (Wilson, 2016). Using data from AhlB crystals he had grown, that were suspected to be in the soluble conformation, Jason used a truncated model of a single chain of the type 2 conformation of the pore form of AhlB as MR model (Figure A.10A). The MR model had residues 150-259 of α3 and α4 as well as large loop regions, removed. Details of crystallisation conditions and data collection are described in Paper1. Initial phases were calculated using Phaser MR (McCoy et al., 2007) as described in Section 2.5.10 and Paper1 (Section 3.3). It could be seen that this crystal form was in the soluble conformation and thus I built and refined this structure. Refinement of the model placed by Phaser resulted in clear density for side chains and positive density for missing helices (Figure A.10B). Subsequent automated model building using Buccaneer (Cowtan, 2006) placed 1071 residues in 7 fragments with three chains, 95.8 % of the residues were built. The final cycle of refinement gave an R-factor of 0.30 and Free R of 0.39 (Figure A.10C). The missing head domain was successfully built by buccaneer in two of the three chains (Figure A.10C). Subsequent rounds of manual building and refinement using Coot (Emsley

et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.22 and Free R of 0.25 (Figure A.11), full validation details can be found in Section3.3 (Paper1, Table 1). A full comparison between the soluble and pore forms of AhlB can be found in Section3.3 (Paper1, Table 1).

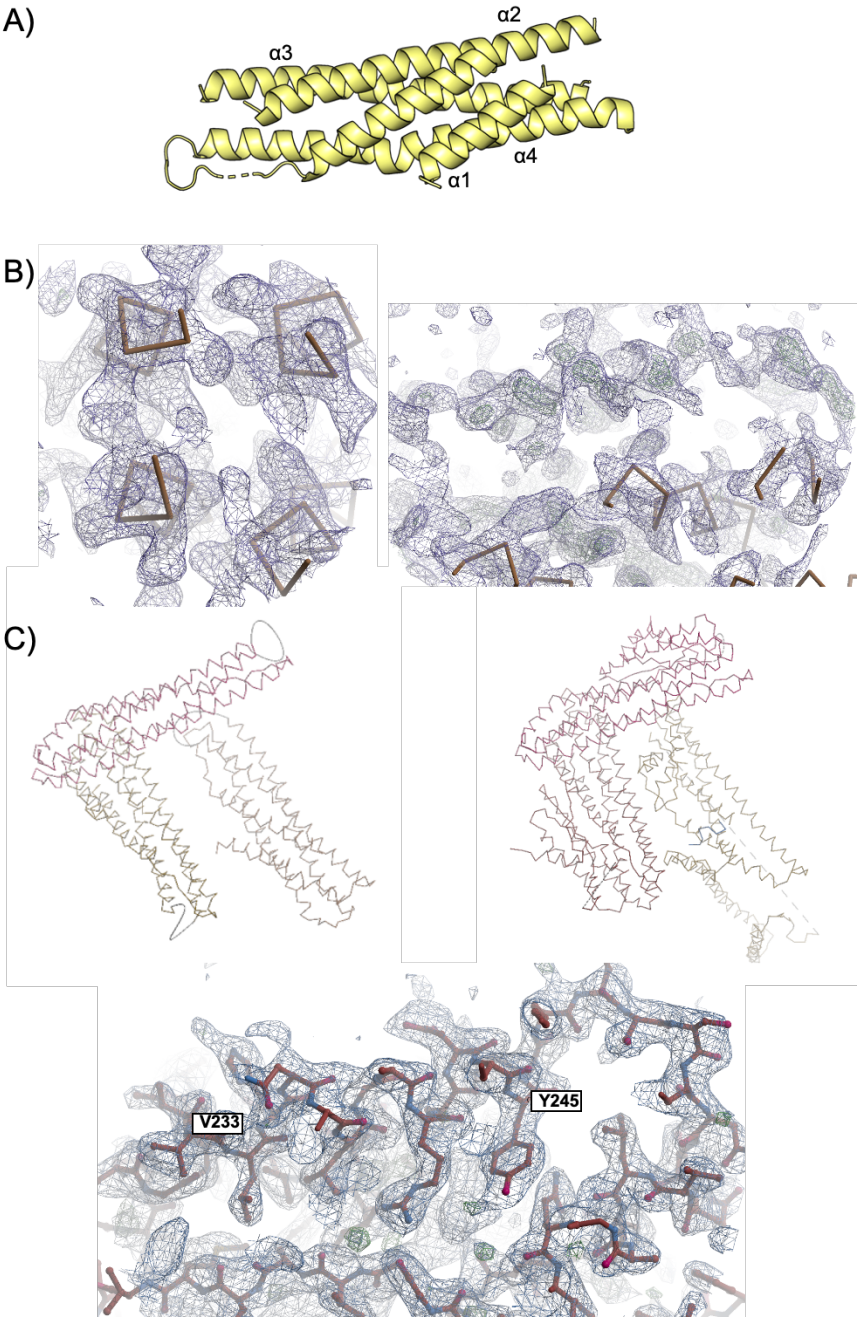


Figure A.10 **Model building and refinement of soluble AhlB.** A) Truncated AhlB type 2 model used as the search model for molecular replacement. B) 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh), after refinement of PhaserMR (McCoy et al., 2007) model in Refmac (Murshudov et al., 1997). Clear side-chain density can be seen around helices (left), while positive density is visible for missing helices (right). C) left, model after PhaserMR placed three copies of the search model, Right, Model after automated model building and refinement with Buccaneer (Cowtan, 2006). Bottom, Buccaneer successfully built residues 150-259 into positive density in A. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh).

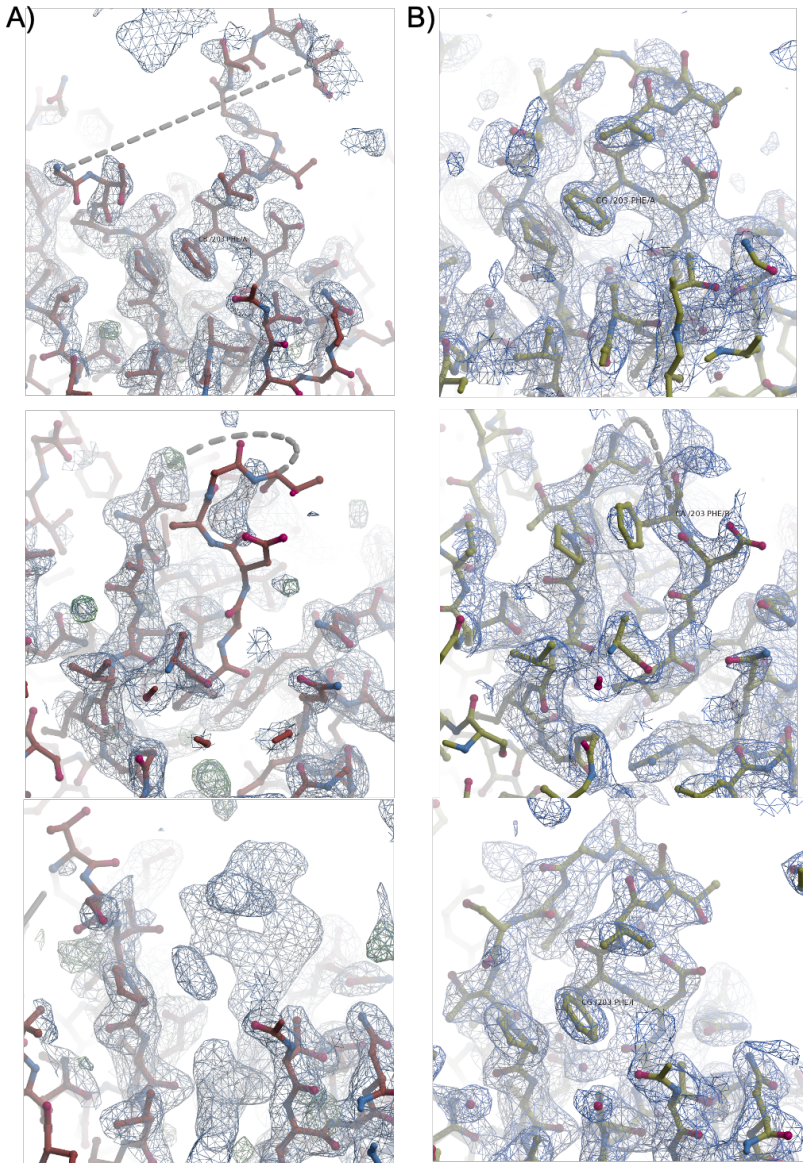


Figure A.11 **Buccaneer (Cowtan, 2006) and final validated models and electron density for each chain of soluble AhlB.** A) Model (Top) and density map around Phe 203 (Below) in chain A (top), B (centre) and I (Bottom), after automated model building and refinement with Buccaneer. B) Final model after validation and density around Phe 203, shows great improvement in all chains. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and positive difference map contoured at 2.99 σ (green, positive, and red, negative, mesh).

A.5 Crystallisation, Data collection and structure determination of AhlC HM

A.5.1 Crystallisation trials of AhlC HM

To investigate whether the hydrophobic head of AhlC was important in binding to the membrane, three leucine residues in the proposed binding segment were mutated to threonine (L156T, L160T, and L161T), and the mutant (AhlC HM) expressed, purified and used in assays by Jason Wilson. To confirm that the mutations had not disrupted the folding of AhlC, I determined its structure.

AhlC HM was purified and concentrated before being subject to crystallisation trials as described in Section 2.5.5 and Paper1 (Section 3.3). Trials were carried out in a 96 well sitting drop plate. Plates were then incubated at 17 ^{o}C and checked regularly. Crystals grew in three conditions shown in Table A.3. All crystals were of sufficient size for subsequent data collection (Figure A.12) and were mounted on litholoops (Molecular Dimensions) and preserved in liquid nitrogen as described in Section 2.5.4.

Table A.3 Initial crystals hits for AhlC HM grown in 98 well sitting drop plates, (Crystallisation conditions, Space group and resolution is given for each crystal.)

Crystallisation condition	Salt	Buffer	Precipitant	Diffraction	Space group
JCSG E12	-	0.1 M imidazole pH8.0	10 % PEG 8000	1.92 Å	$P6_122$
PACT C2	-	0.1 M PCTP pH 5.0	25 % PEG 1500	2.59 Å	$P2_12_12_1$
PACT H10	0.2 M Sodium/potassium phosphate	0.1 M Bis-Tris propane pH8.5	20 % PEG 3350	2.23 Å	$P6_122$

A.5.2 X-ray Data collection of AhlC HM

Data was collected from a single crystal (PACT E12, Table A.3) at Diamond Light source beamline i04 as described in Section 2.5.7 and Paper1 (Section 3.3) with full data processing statistics given in Section 3.3 (Paper1, Table2). A dataset of 2000 images was collected with an oscillation per image of 0.1 ^{o} (total 200 ^{o}), exposure of 0.04 s, and transmission of 60 %. An example of a diffraction image is provided in (Figure A.13A), showing diffraction out to 2 Å. Cell parameters (Section 3.3 (Paper1, Table2)) suggested 2 molecules in the asymmetric unit with a solvent content of 45 %, as determined by the Matthews calculation (Matthews, 1968) (Figure A.13B).

A.5.3 Structure determination, model building and validation of AhlC HM

Initial phases were calculated using a single chain of AhlC (PDB code 2H2E) as a search model and Phaser MR (McCoy et al., 2007) as described in Section2 and gave 11 possible solutions. The best solution had a refined TFZ score of 20.7 and LLG of 6418 and placed

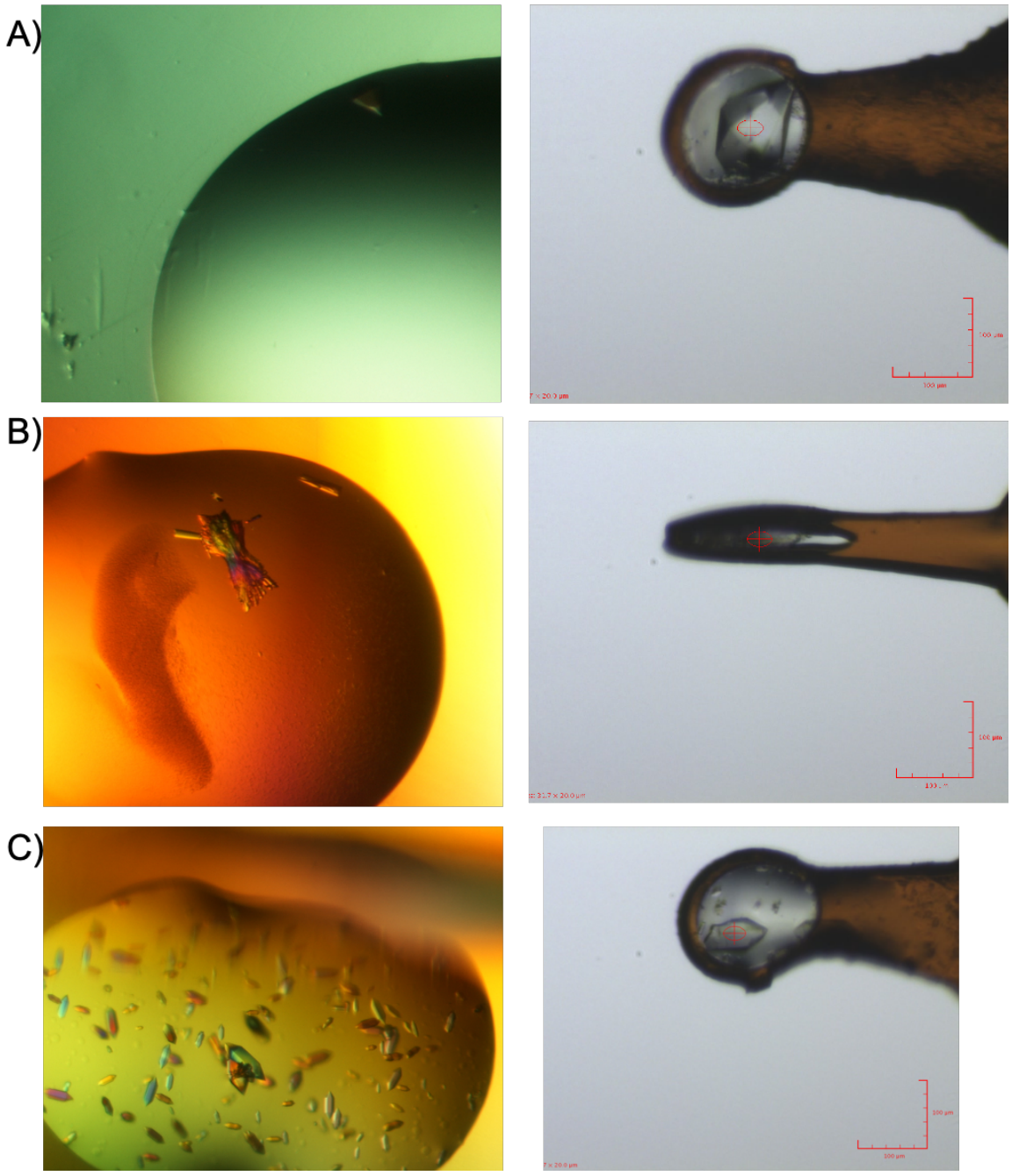
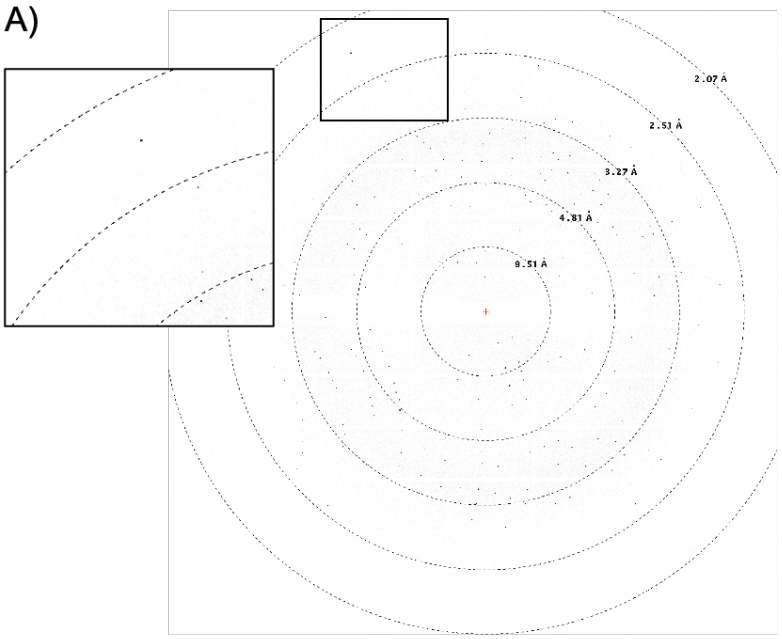


Figure A.12 **Crystals of AhlC HM (left) and mounted crystal in litholoop (right).** A) JCSG E12, B) PACT C2, C) PACT H10.



B)

N(mol)	Prob(N) for resolution	Prob(N) overall	Vm A**3/Da	Vs % solvent	Mw Da
1	0.0041	0.0403	5.88	79.07	28000.00
2	0.4595	0.6464	2.94	58.14	56000.00
3	0.5364	0.3134	1.96	37.21	84000.00
4	0.0000	0.0000	1.47	16.28	112000.00

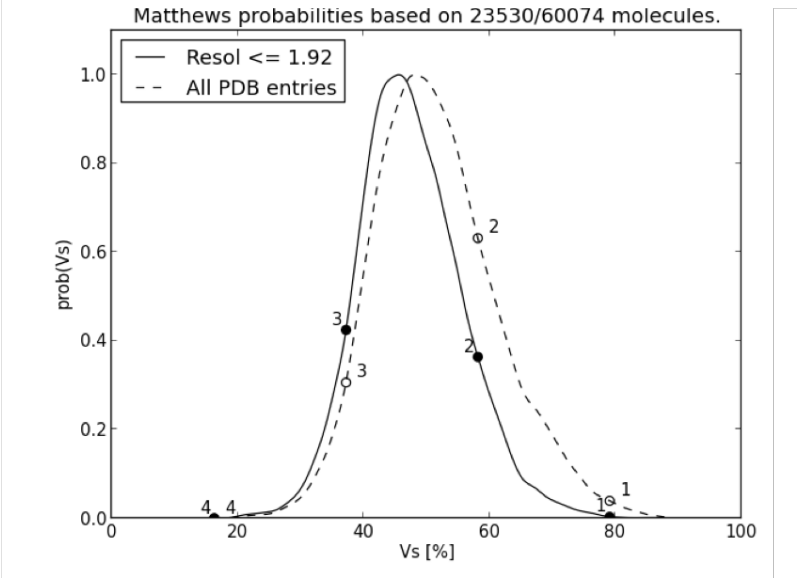


Figure A.13 **AhlC HM (PACT E12) data collection on beamline i04.** A) A representative 0.1 ^{o} oscillation image from AhlC HM. An enlarged view of the region highlighted by the square shows diffraction extends to around 2 Å B) Soluble AhlB Matthews calculation. Top, table of the possible number of molecules in the AU, with respective probability and percentage solvent content, suggesting 2 molecules is the AU. Below, A plot of the probability against the solvent content. Graph and table generated using online software at http://www.ruppweb.org/mattprob/ (Matthews, 1968).

two molecules in the AU. Refinement of this model using REFMAC5 (Murshudov et al., 1997) gave an R-factor of 0.35 and Free R of 0.38. Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.23 and Free R of 0.28 (Figure A.14), full validation details can be found in (Section 3.3) Paper1, Table2 (Section 3.3. In both chains, clear electron density could be seen for all residues in the head domains (Figure A.14B), in contrast to previous models of AhlC. In the final model, all three Leu to Thr (L156T, L160T, and L161T) mutations fit density well (Figure A.15). The asymmetric unit contained a dimer of AhlC (chains D and J), which by rotation about the crystallographic twofold axis formed a 222 tetramer (Chain JJ' and DD') (Figure A.16A).

A.5.4 Interface analysis of head domains in the AhlC HM

Analysis of the crystal packing of symmetry mates in the unit cell shows that the head domain of chain J packs against the tail domain of chain J' within the tetramer as well as Chain D in the neighbouring symmetry-related tetramer. The head domain of chain D, however, packs against Chain J' and Chain D' tail domains within the tetramer (Figure A.16A). Closer analysis of these regions showed that in Chain D Tyr 154 packs tightly into the tail domain of neighbouring chain D forming a hydrogen bond interaction with Asp 231 resulting in disorder in the tail helix after Asp 231 as shown by the positive difference density where the residues could not be successfully built (Figure A.16B). This is not the case in Chain J where Tyr 154 does not pack within the tetramer (Figure A.16B), instead the hydrophobic residues Ile 167 and Val 171 form hydrophobic interactions with Ile 123 and Leu 127 in Chain J of the neighbouring tetramer (Figure A.16C). In the head domain of Chain D Ile 123 and Leu 127 are exposed to the solvent (Figure A.16C). The Tyr-Asp interaction likely plays an important role in tetramer assembly and stability, while the exposed hydrophobic residues

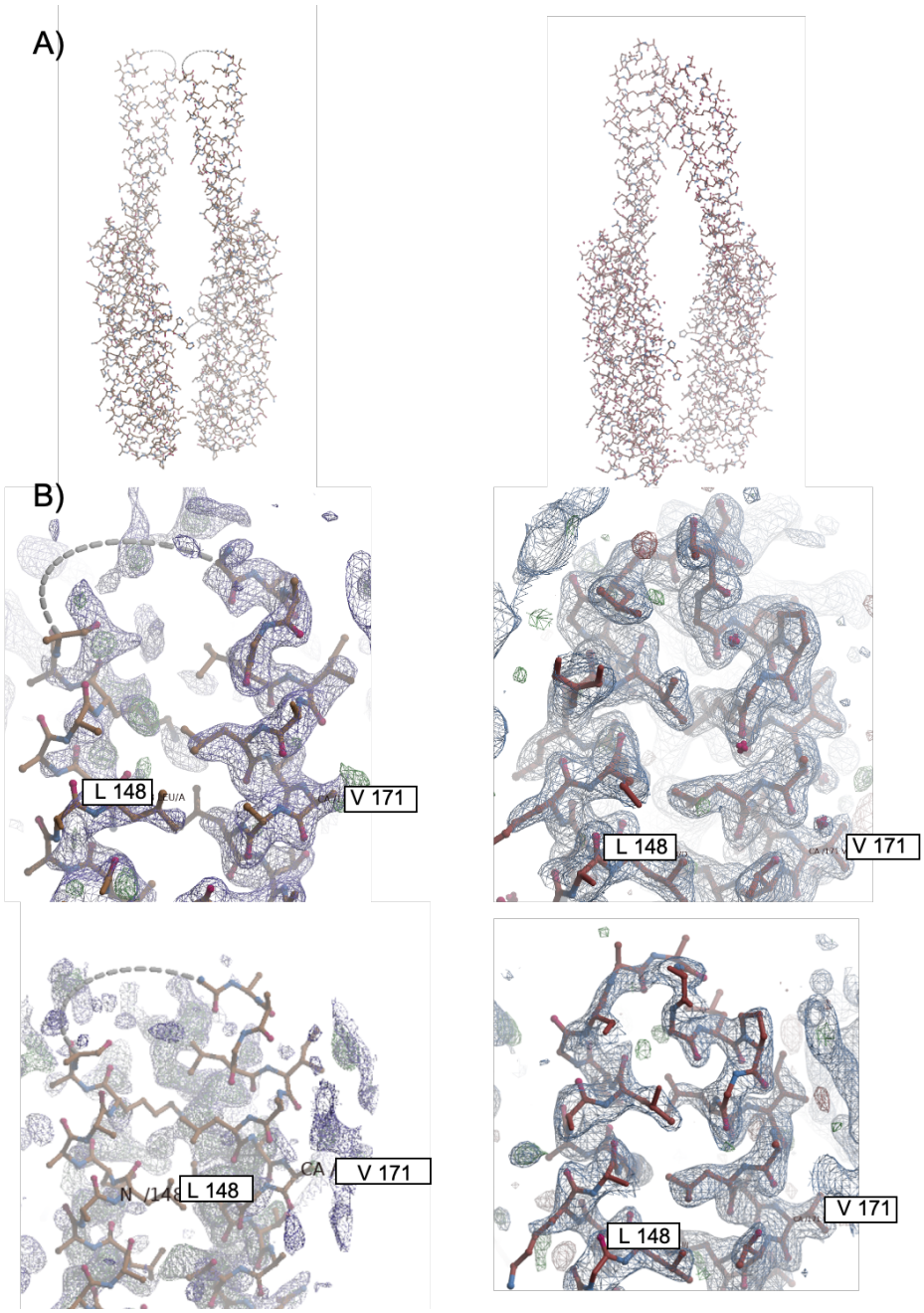


Figure A.14 **Model building and refinement of soluble AhlC HM.** A) Left, model after PhaserMR (McCoy et al., 2007) placed two copies of the search model. Right, model after model building and Validation, the head domains of each chain is now complete. B) Model and density map of the head domain of Chain A (top) and B (bottom) between Lue 148 – Val 171, after refinement of PhaserMR output (left) and the final model after validation (Right). Both Chains are complete with clear density for most side chains. Significant movement of the head domain in chain B was required in order to fit to density. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh).

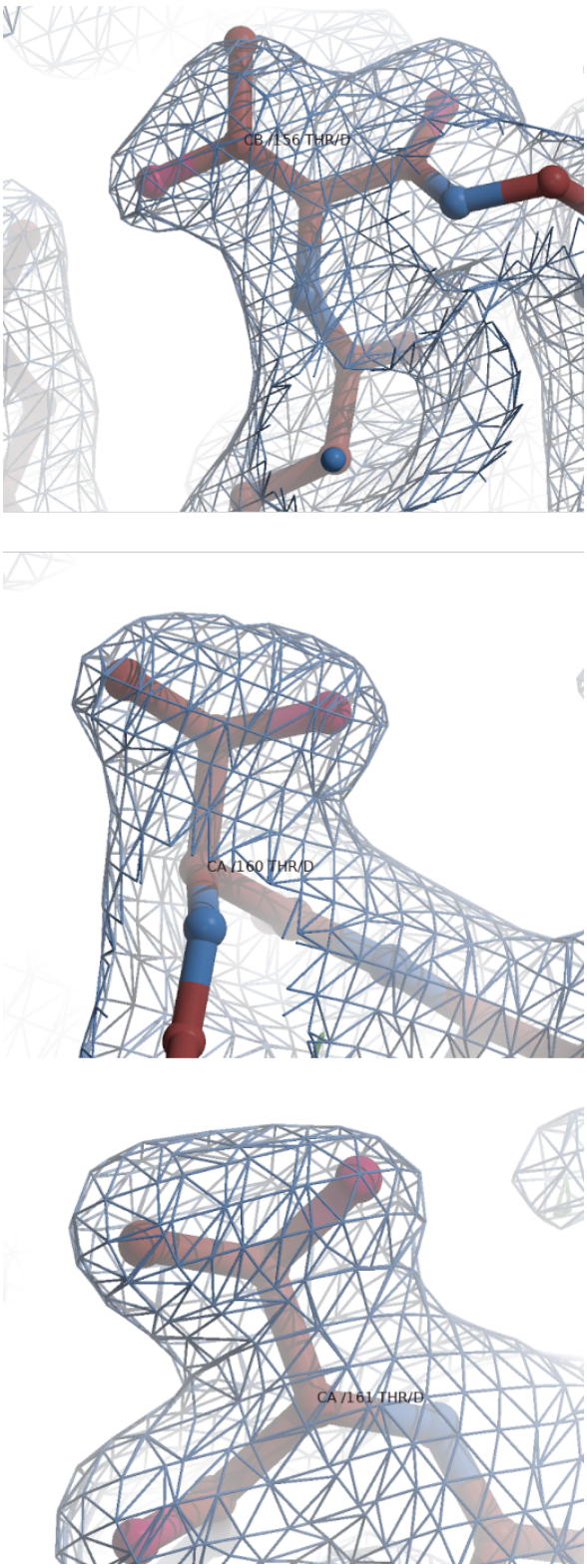


Figure A.15 **Density around Thr 156 (top), Thr 160 (centre), Thr 161 (bottom).** 2Fo-Fc map contoured at 1.0 σ (blue mesh).

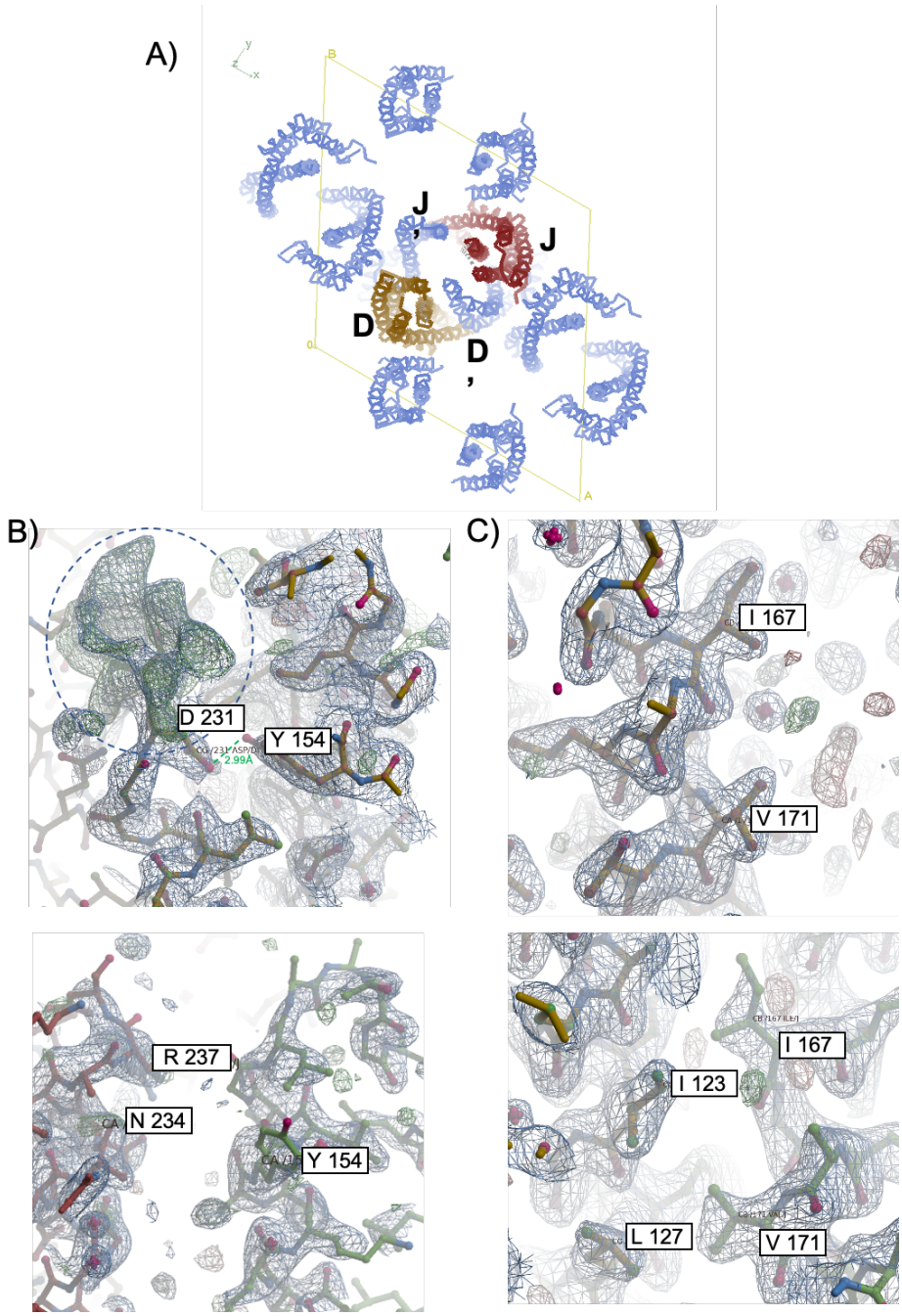


Figure A.16 **Head domain Interface analysis for AhlC HM.** A) Crystal packing for AhlC HM, Symmetry generated molecules are shown in blue while the unit cell as yellow lines. B) Chain D head domain packs tightly with neighbouring Chain D in the tetramer, Tyr 154 is 2.99 Å from Asp 231 forming a hydrogen bond. Positive difference density where residues could not be successfully built is highlighted in the tail domain by the blue circle. Below, Chain J head domain does not pack and no hydrogen bond forms between Try 154 and Asp 231. C) Top, Hydrophobic residues (Ile 167 and Val 171) are exposed to solvent in Chain D, however, in chain J they form hydrophobic interactions with Chain J tail in neighbouring tetramer through crystal packing. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh).

at the ends of α 3 - α 4 are important for membrane recognition and initiating disassembly of the tetramer for the insertion of the hydrophobic head as described in Paper1 (Section 3.3).

A.6 Detergent screens and electron microscopy of AhlA, AhlB and AhlC

To study the pore-forming properties of AhlABC a suitable detergent had to be identified that was able to effectively drive pore formation. To do this a new method was developed using HPLC that allowed rapid screening of a large number of detergents. The successful identification of a detergent allowed for both electron microscopy studies of single pores and ultra-centrifuge assays as described in Paper1 (Section 3.3), Section 2.6 and Section 2.4.1.

For AhlABC to form a pore, a conformational change must be induced by biological membranes and/or detergents, with the pore purified in conditions that maintain this conformation. Previous studies by Jason Wilson with the most common detergents (DDM, DOC), failed to purify enough complex for structural studies. As a result, a more thorough screen of detergents needed to be carried out to identify the best detergent for induction of pore formation.

As screening many detergents is very time-consuming using a standard gel filtration method, a method for screening detergents was developed whereby up to 100 detergents could be screened using HPLC, using only 70 μl protein in each trial. All detergents screened in this study were sourced from the Hampton Research detergent screen 1.

A.6.1 HPLC detergent screening against AhlB

AhlB was screened against 25 detergents, each at 10x CMC. To reduce the manual demand in screening, a single buffer (50 mM Tris-HCl pH8 and 0.5 M NaCl) was used for all 25 detergents. 70 μl of 2 mg/ml AhlB was incubated with each of the 25 detergents for 4 hours

at 37 ^{o}C. Two controls were also set up, one containing detergent incubated on its own and one with protein incubated on its own.

A good detergent would be classified as one that produced a single peak on the chromatogram at an elution volume of 7 ml, without any significant peaks for the controls (the detergent or the protein incubated on its own) and no degradation of the protein. Figure A.17A shows chromatograms for three of the 25 detergents alongside the two controls. Each of these three detergents produced a single peak, with no peaks for the controls. Detergent 10 (Lauryldimethylamine oxide) and 13 (Nonyl-β-D-glucopyranoside) produced a small peak at 6.5 ml, while detergent 18 (N-Heptyl-β-D-thioglucopyranoside) had a large peak at 7 ml. These peaks were distinct from the detergent peaks (4.7 ml and 5.7 ml) and the protein without detergent (10.9 ml) and showed no significant peaks at high volumes, which would represent degraded protein. As a result, these samples were collected for negative stain EM (Figure A.17). The most pores were observed in sample incubated with detergent 18 and as such this detergent was chosen for further optimisation and trials with AhlB and C.

A.6.2 Detergent 18 trials with AhlB and AhlC

As detergent 18 showed the best results in the initial trials with AhlB, it was taken forward for studies to purify a homogenous population of the AhlB+C pore for EM. Due to the cost of this detergent in the first test, only low concentrations of detergent were used. AhlB and AhlC at a 1:1 ratio (giving a total protein concentration of 2 mg/ml) were incubated with detergent 18, at concentrations of 0.1, 1 and 2 x CMC. As with studies with AhlB alone, controls were run containing AhlB+C without detergent and detergent without protein.

These samples were then run on the HPLC as described in Section 2.1.13. Figure A.18A shows the chromatograms for each concentration and the controls. AhlB+C without detergent shows a large peak at the expected elution volume for AhlB, however there is no discernible peak for AhlC alone. There is however a small peak at the expected elution volume for the

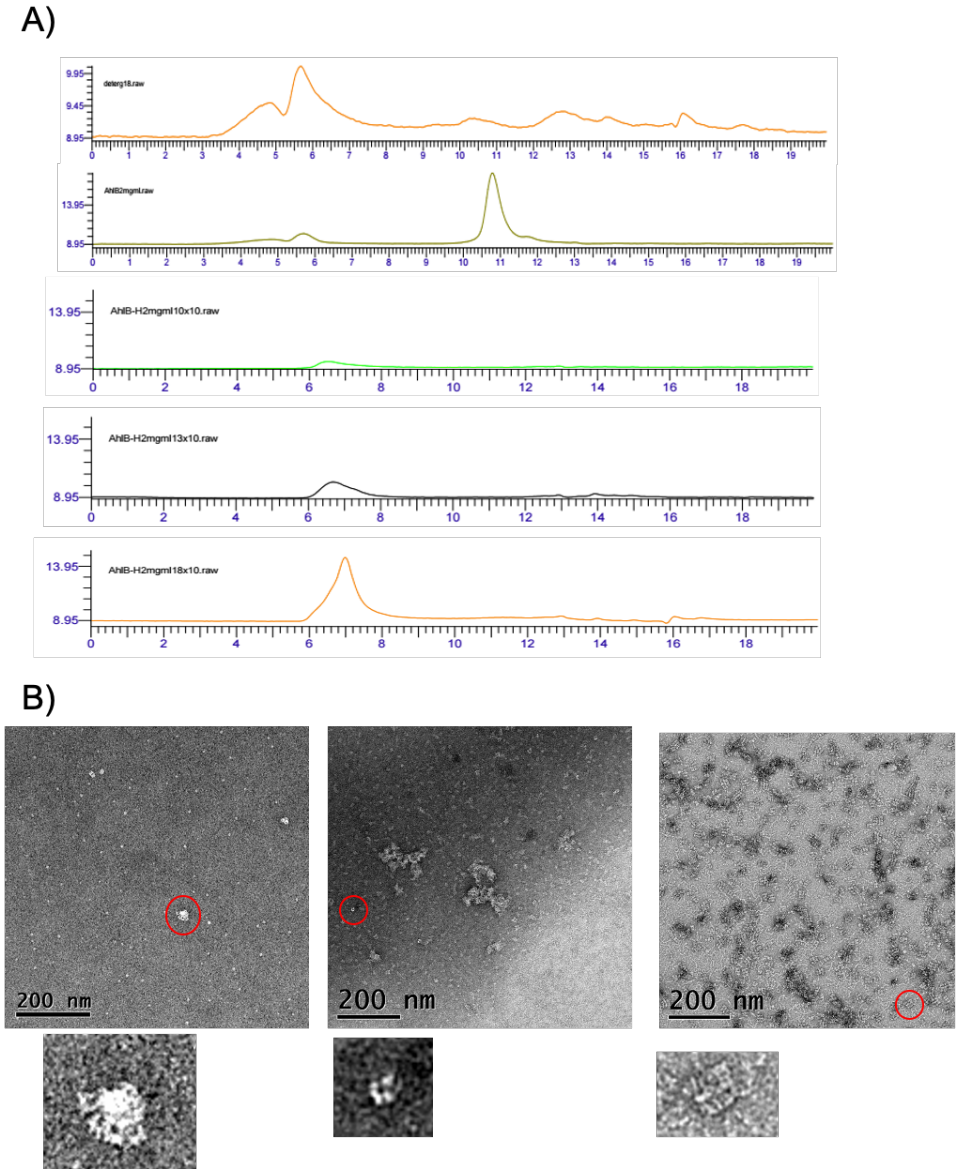


Figure A.17 **Chromatograms from best 3 detergents at 10 x CMC from initial screens with AhlB.** The first 2 chromatograms show detergent on its own and protein on its own. The bottom 3 chromatograms show each detergent incubated with AhlB, the best detergent being 18 at the bottom and the worst being 10 at the top. B) TEM negative stain images of samples taken from the initial detergent screen, magnified images of regions highlighted in red are shown below. Left, image taken of AhlB incubated with detergent 10 for 4 hours at 37 ^{o}C. Small groups of pores can be seen highlighted in red. Centre, AhlB incubated with detergent 13 for 4 hours at 37 ^{o}C. Individual pores are visible, highlighted in red, also large aggregates can be seen. Right, AhlB incubated with detergent 18 for 4 hours at 37 ^{o}C. Side on views of Multiple pores in micelles can be seen.

pore, suggesting AhlB+C may be forming oligomers without the help of detergent. There are also small peaks at a high elution volume, this might be degraded protein or contamination. As the detergent concentration increases, the peak at 7 ml increases to a maximum at 1 x CMC. At 1 x CMC no monomeric protein peak is present but the peaks at high elution volume have increased. By 2 x CMC the peak at 7 ml has fallen away and a large broad peak has appeared around 11 ml. At this concentration, the detergent seems to be hindering pre-pore formation (Figure A.18A).

Samples of potential pore complexes in detergent concentrations 0.1 x CMC and 1 x CMC were collected and diluted down to 0.02 mg/ml and 0.2 mg/ml protein concentration for negative stain EM studies (Figure A.18B). However, samples at 1 x CMC formed precipitate and could not be used. Side views of pores in detergent micelles could be seen for the 0.1 x CMC sample on the carbon grids, showing successful induction of pore formation.

A.6.3 Detergent 18 trials with AhlA, AhlB, and AhlC

Due to low yields of AhlA from purification, not enough sample was available to carry out HPLC screens of detergent with AhlA, AhlB and AhlC together, Instead, all three proteins were incubated with detergent 18 at 1x CMC as described in Section 2.1.13. The resulting sample was loaded onto a carbon coated grid and observed negative stain EM to see if pore formation had been induced. Single pores were visible in this sample (Figure A.19A) showing that the detergent N-Heptyl-β-D-thioglucopyranoside was able to induce pore formation in AhlB, AhlBC and AhlABC and could be taken forward for ultracentrifugation studies to determine the composition of the pores (Section 2.4.1 and Paper1 (Section 3.3)).

Two negative stain micrographs of AhlABC incubated with detergent 18 were used for 2D classification. Negative stain micrographs (7.18 Å/pix) were imported to cisTEM (Egelman et al., 2018) for processing. The CTF (contrast transfer function) of each image was estimated to a maximum resolution of 18 Å. Particles were picked with a maximum

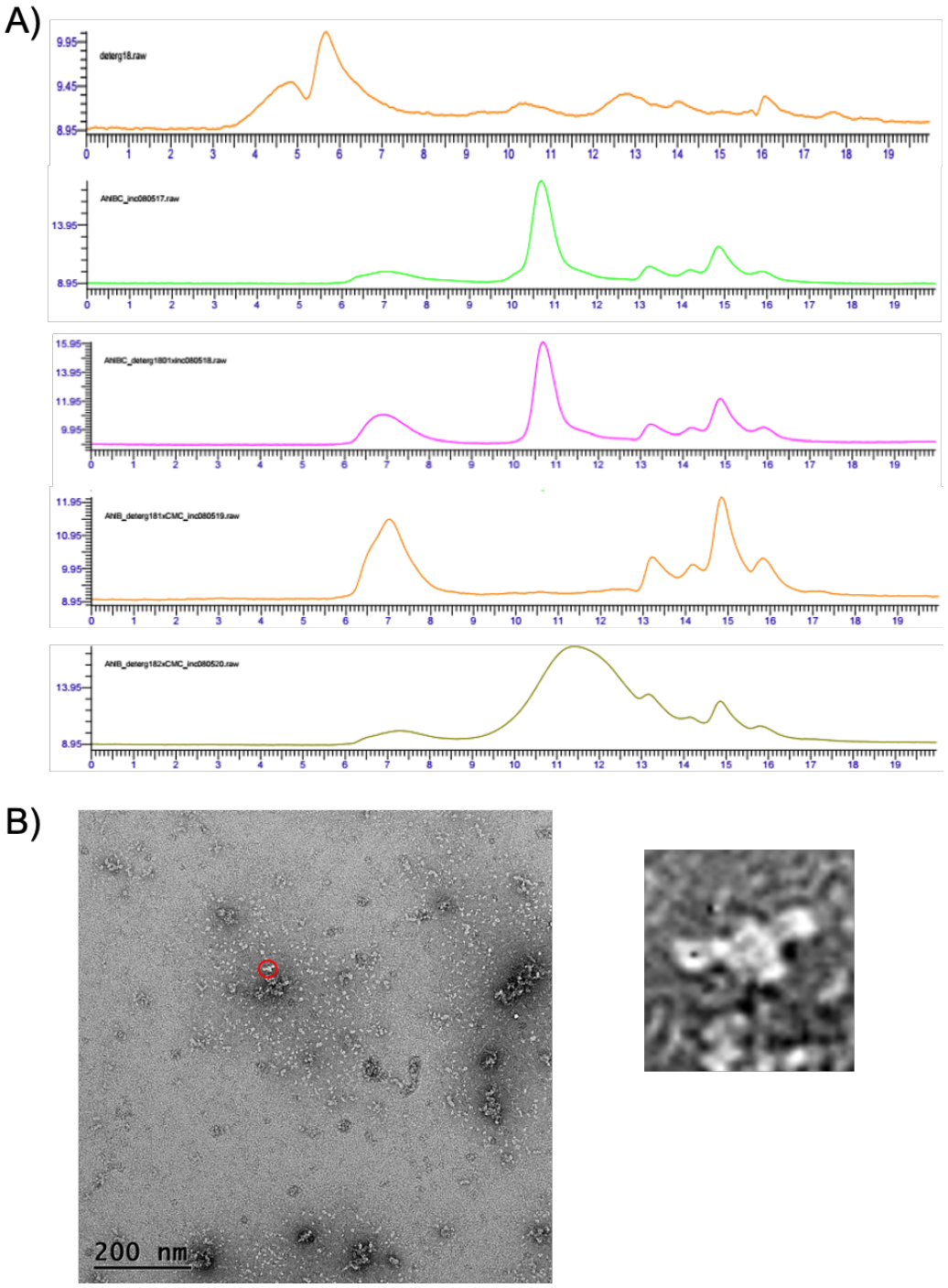


Figure A.18 **Chromatograms of AhlB and AhlC incubated with detergent 18 at 3 concentrations.** The first 2 chromatograms show detergent on its own and protein on its own. The bottom 3 chromatograms show AhlBC incubated with Detergent 18, detergent concentration increases from 0.1-2 x CMC top to bottom. B) TEM negative stain images of AhlB and AhlC incubated with 0.1 x CMC detergent 18. Side on views of pores are visible in detergent micelles.

particle size of 180 Å and a characteristic particle radius of 80 Å (based on expected AhlABC pore size from Paper1 (Section 3.3)). A few extra 'pores' were added manually resulting in 966 particles. Images in Figure A.19B show some particles, but many of these look like aggregates as well as more obvious pores. These may be oblique views of the pore, however more particles would be required to fill these views to generate good 2D classes. Particles were extracted with a box size of 48 pixels (450 Å), and used in 2D classification, with 10 classes, 20 iterations, a max resolution of 18 Å, a mask radius of 90 Å, and an angular search step of 5 ^{o}. Signal was lost by iteration 20, so iteration 7 was selected (Figure A.19C). Classes are quite ambiguous as expected from the low image number. Class 4 shows a possible pore with a diameter of 140 Å, this could represent an AhlB pore (Figure A.19C). The other classes are more ambiguous and vary in size, these could represent other pore complexes or detergent micelles. Further optimization of the sample to remove micelles and ensure a single pore species is present is needed, as well as the collection of more images to generate better classes of other pore views.

A.6.4 Optimisation of AhlB pores with detergent.

The crystal structure of the AhlB pore suggests a plug is formed in the centre of the oligomer by the 15 missing N-terminal residues of the type 2 AhlB monomers. To try and determine if this is the case, samples of AhlB pores were generated and optimised for single-particle cryo-electron microscopy. 2 mg/ml AhlB was incubated with the detergent N-heptyl- β-D-thioglucopyranoside for 1hr at 37 ^{o}C. The sample was loaded onto a Superose 6 column pre-equilibrated with 0.5 M NaCl, 50 mM Tris pH8 and 1x CMC N-heptyl-β-D-thioglucopyranoside. AhlB eluted from the column at a volume of 9.5-13.5 ml (Figure A.20A). Fractions 12, 14, 16, and 18 were loaded onto carbon coated grid as described in Section 2.6.1, for analysis by negative stain EM. Pores could be observed in all fractions,

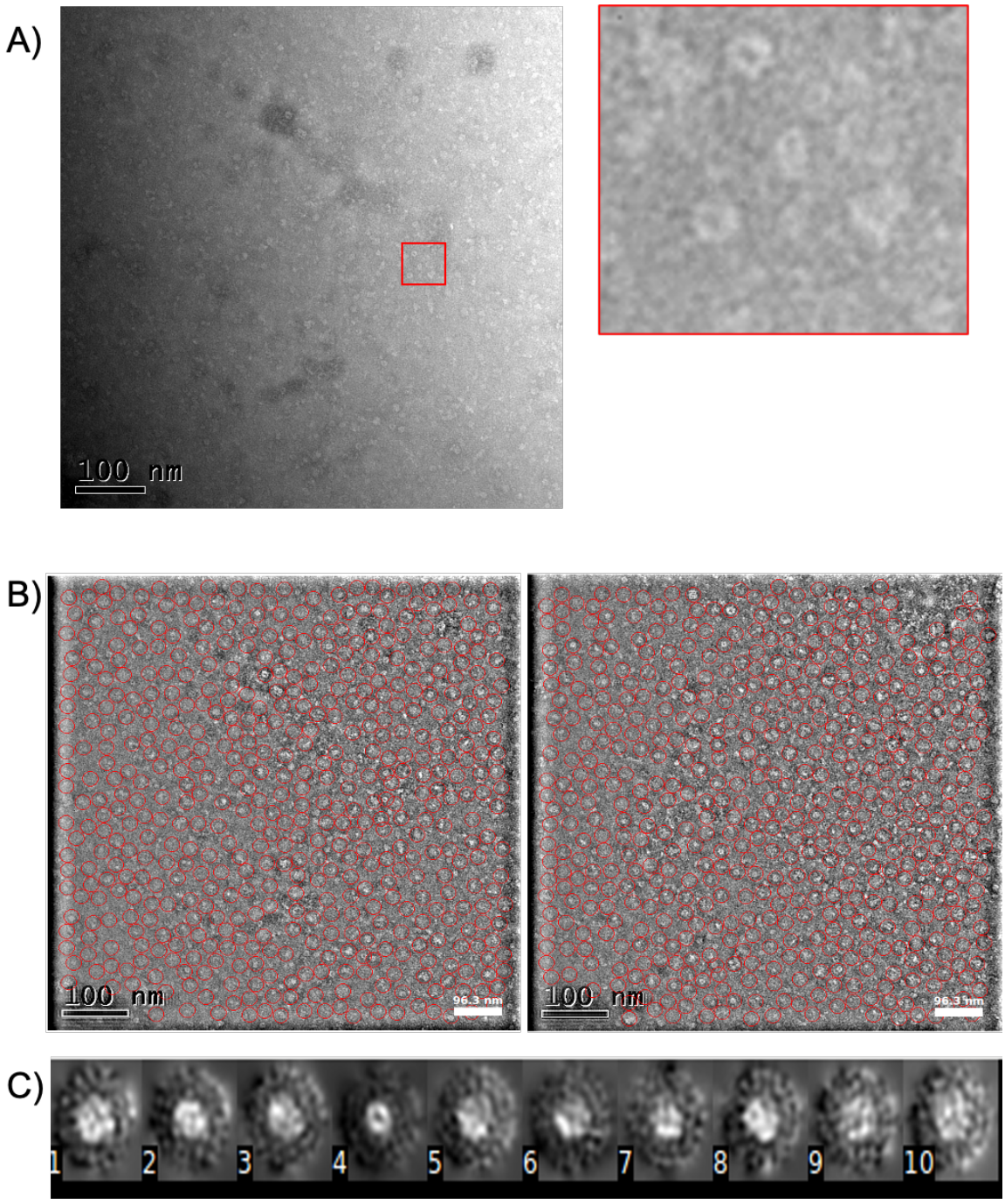


Figure A.19 **AhlABC forms pores when incubated with 1 x CMC of detergent 18.** A)TEM negative stain images of AhlA+AhlB+AhlC incubated with detergent shows top and side views of single pores of diameter 130Å +/- 20 Å. Right - a magnification of the area of the original image highlighted in red showing two top views of pores and a side view. B) Particle picking based on a maximum particle radius of 180 Å. C) 2D classes from iteration 7. Class 4 is the only class to contain a pore-like object.

with the best pores in sample 14 (Figure A.20B), however the pores in these fractions were mostly grouped around small detergent micelles with only a few single particles.

To try and increase the number of single pore particles, samples from fractions 14, 15, 16 and 17 were buffer exchanged into 0.5 M NaCl, 50 mM Tris pH8 with the aim of breaking up the large micelles while maintaining pore assembly. Figure A.21A shows buffer exchanged sample from fraction 17. Dispersed single particles of approximately 10 nm diameter could be seen, as a result, this sample was loaded onto Quantifoil holey carbon grids and plunge frozen in liquid ethane, as described in Section 2.6.1. The sample was loaded into a Techni Arctica microscope under cryo conditions as outlined in Section 2.6.3. As with the negative stain samples single particles could be seen on the carbon film however no particles were observed in the holes. As images are collected at a low dose for high-resolution Cryo-EM, particles must be concentrated in vitreous ice-filled holes to ensure the background is low and contrast is high. Unfortunately, this meant that no high-resolution data could be collected. (Figure A.21). Optimisation by varying the pH, glow discharge times and grid type (graphene oxide vs carbon) failed to improve particle distribution. Further optimisation of the grid and sample preparation will be needed before any data can be collected.

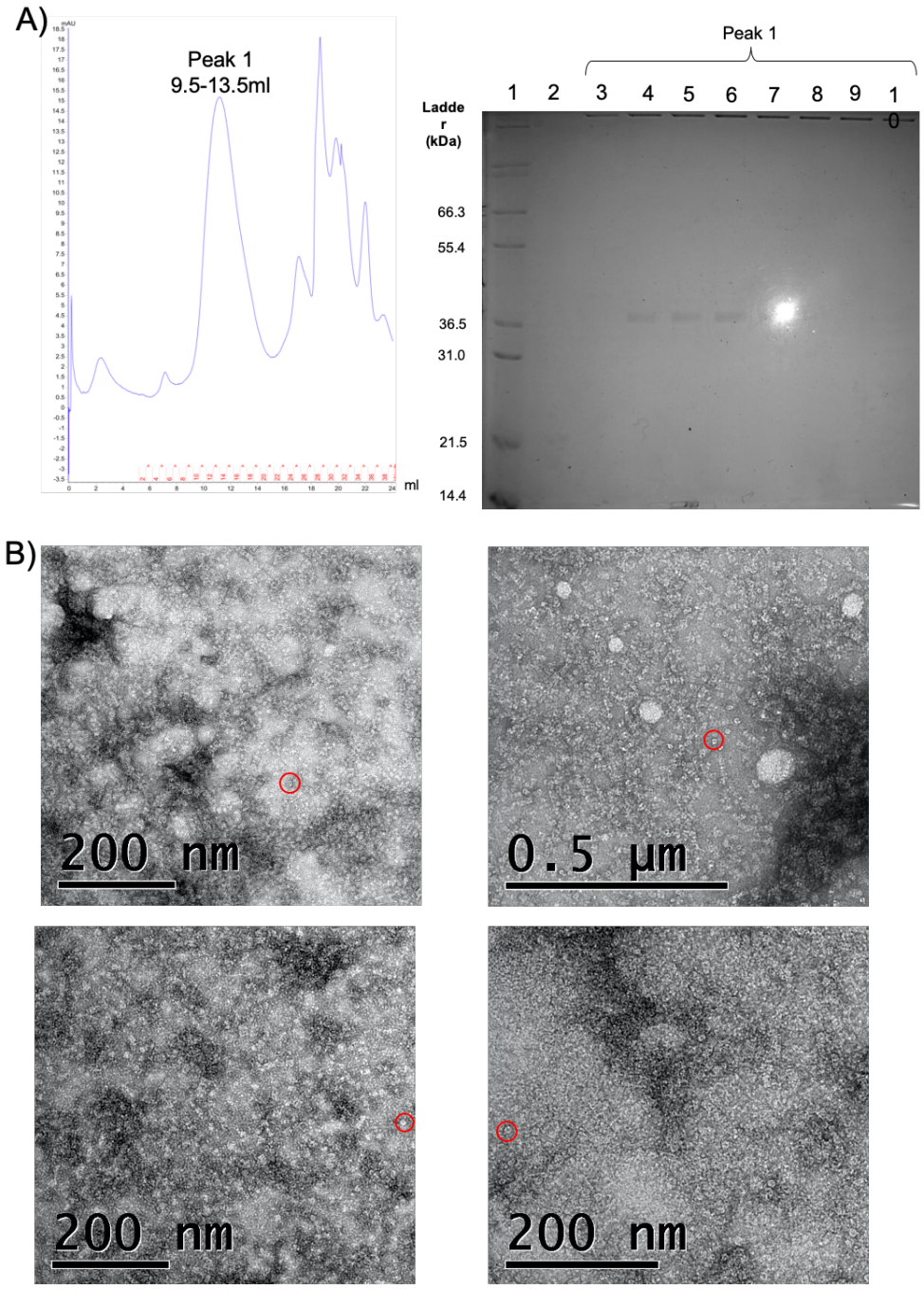


Figure A.20 **Optimisation of AhlB pores for Cryo-EM.** A) Gel filtration UV trace chromatogram (left) and SDS-PAGE (right) of AhlB incubated with 1 x CMC N-heptyl-β-D-thioglucopyranoside (detergent 18). SDS-PAGE show Lane 1, Mark 12 ladder; Lane 2, Fraction 6; Lane 3 – 10, Fraction 11-18. AhlB elutes mostly in fractions 12-15. B) TEM negative stain images of Fractions 12 (top left), 14 (top right), 16 (bottom left), 18 (bottom right). Pores are visible in all fractions highlighted in red.

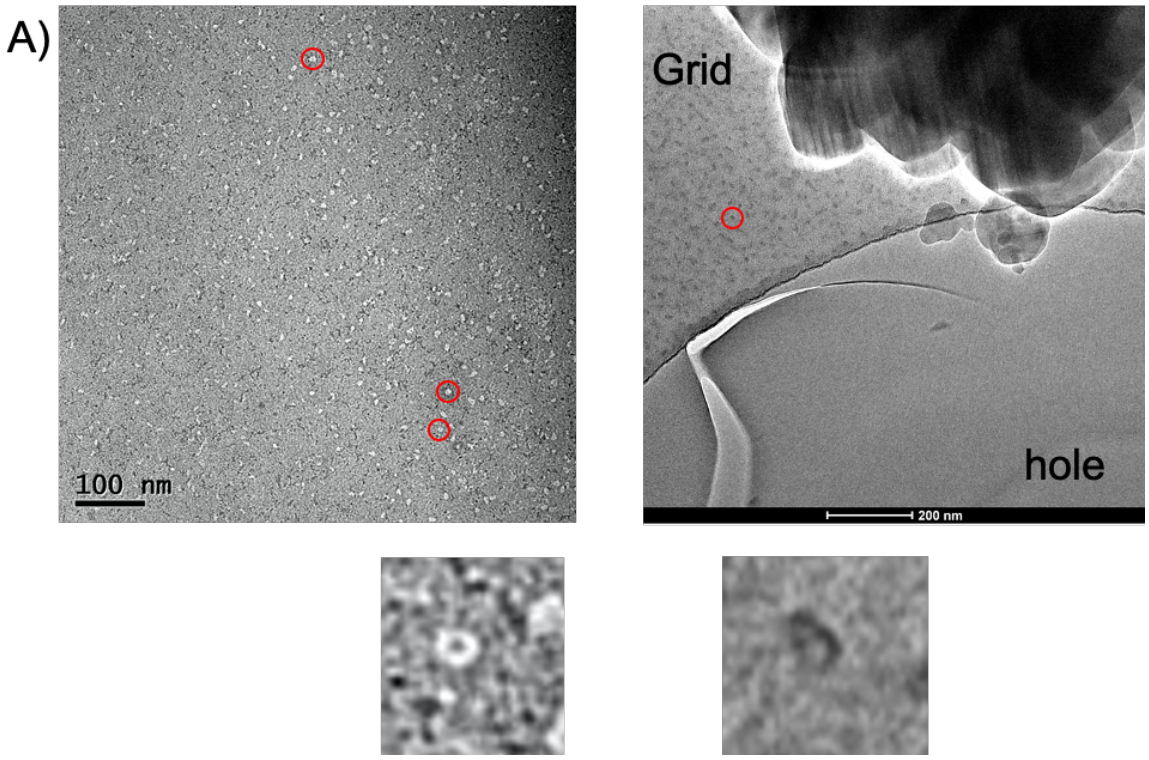


Figure A.21 **Optimisation of AhlB pores.** A) Removal of detergent from purified samples by buffer exchange into 0.5 M NaCl, 50 mM Tris pH 8 gave more disperse single pores. B) Cryo-EM images of sample from A. All pores are concentrated on the carbon film (red circle) with no particles in the holes.

Chapter 4

Paper 2

4.1 Summary

The second paper in this thesis is a crystallisation article discussing the generation, expression, and purification of SmhA protein construct generated from the ORF for *Serratia marcescens* MSU-97 (NCBI accession number OKB64935.1). The paper also details suitable conditions to generate diffraction quality crystals as well as subsequent SAD data collection and phasing, using Selenomethionine derivative SmhA. Finally, preliminary data analysis is discussed for SmhA, the first structural data collected from an A component of a Gram negative tripartite α-PFT. Details of work done that was omitted, or was only briefly mentioned, from the results and discussion in this paper can be found in Appendix B (page 209)

4.2 Author contributions

P.J.B. conceived the project; P.J.B., A.M.C.-A., T.H.S. designed the experiments; A.M.C.-A. T.H.S and S.E.S performed the experiments; P.J.B., A.M.C.-A. interpreted the data and P.J.B. and A.M.C.-A. prepared the paper with input from all authors. Experimentally, I carried out

the construct design, protein production purification, crystallisation and data collection of SmhA.

4.3 Paper 2

Paper 2 manuscript will start on the next page.

research communications

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

The A component (SmhA) of a tripartite pore-forming toxin from *Serratia marcescens*: expression, purification and crystallographic analysis

Alicia M. Churchill-Angus,[a] Svetlana E. Sedelnikova,[a] Thomas H. B. Schofield[a,b] and Patrick J. Baker[a]*

[a]Department of Molecular Biology and Biotechnology, The University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom, and [b]Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom. *Correspondence e-mail: p.baker@sheffield.ac.uk

Received 27 August 2020
Accepted 17 October 2020

Edited by M. W. Bowler, European Molecular Biology Laboratory, France

Keywords: pore-forming toxin; crystallization; ClyA family; *Serratia*.

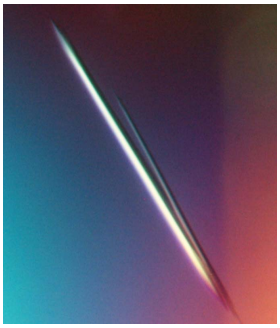
Supporting information: this article has supporting information at journals.iucr.org/f

Tripartite α-pore-forming toxins are constructed of three proteins (A, B and C) and are found in many bacterial pathogens. While structures of the B and C components from Gram-negative bacteria have been described, the structure of the A component of a Gram-negative α-pore-forming toxin has so far proved elusive. SmhA, the A component from the opportunistic human pathogen *Serratia marcescens*, has been cloned, overexpressed and purified. Crystals were grown of selenomethionine-derivatized protein and anomalous data were collected. Phases were calculated and an initial electron-density map was produced.

1. Introduction

Tripartite α-pore-forming toxins (α-PFTs) are members of the ClyA α-pore-forming toxin family (Fagerlund *et al.*, 2008; Wilson *et al.*, 2019); however, unlike ClyA, where the pore is formed from an oligomer of a single protomer, three proteins (A, B and C) are involved in active pore formation (Sastalla *et al.*, 2013; Lindbäck *et al.*, 2004; Wilson *et al.*, 2019; Beecher & Macmillan, 1991). In the ClyA family the active pore is formed when the soluble protein(s) undergo a large-scale conformational change to expose the membrane-binding regions, with protomers assembling into a hydrophilic-lined pore (Benke *et al.*, 2015; Roderer & Glockshuber, 2017). It is proposed that each protein of the tripartite α-PFT (A, B and C) fulfils a role in the active pore that is provided by different regions of the ClyA protomer (Wilson *et al.*, 2019). The C component makes the first attachment to the target cell, binding to a single leaflet of the membrane, and is equivalent in function to the β-tongue region of soluble ClyA. In ClyA the pore is completed by the N-terminal amphipathic helix of each protomer assembling to construct the membrane-spanning, hydrophilic-lined pore of the oligomer (Roderer & Glockshuber, 2017; Wallace *et al.*, 2000; Benke *et al.*, 2015). In the tripartite α-PFTs the A and B components are functionally equivalent to this region of ClyA. The B component acts as the pore-forming unit, using two hydrophobic helices to span the membrane, with the A component proposed to provide amphipathic helices that produce the hydrophilic interior lining of the oligomeric pore (Wilson *et al.*, 2019; Mueller *et al.*, 2009; Benke *et al.*, 2015).

Tripartite α-PFTs were first identified in the pathogenic Gram-positive bacterium *Bacillus cereus*, when the Hbl



OPEN ACCESS

research communications

system, and later the NheABC system, were identified as vital toxins in its pathogenicity and the cause of a major food-poisoning outbreak in Norway (Thompson *et al.*, 1984; Lund & Granum, 1996; Beecher *et al.*, 1995). The tripartite α-PFT family has recently been expanded into a large number of clinically and economically important Gram-negative bacteria, including the fish and opportunistic human pathogen *Aeromonas hydrophila* (Wilson *et al.*, 2019). The α-PFT toxin AhlABC from *A. hydrophila*, like NheABC, has been shown to be lytic to mammalian cells and forms pores in membranes (Wilson *et al.*, 2019; Lindbäck *et al.*, 2004).

Structures of soluble AhlB and AhlC, and also a pore structure of AhlB, have been solved by X-ray crystallography; however, a structure of the A component from a Gram-negative α-PFT has yet to be determined (Wilson *et al.*, 2019) and thus the structural role of this protein in the active pore is as yet unknown. Within the Gram-positive *B. cereus* α-PFTs, only the structures of NheA (PDB entry 4k1p; Ganash *et al.*, 2013) and HblB (PDB entry 2nrj; Madegowda *et al.*, 2008) have been determined. HblB is functionally equivalent to AhlC, yet these two proteins share less than 10% sequence identity and their structures are significantly different (Wilson *et al.*, 2019). Similarly, NheA and AhlA share only 6% sequence identity, and thus structures of the A component from the Gram-negative bacterial α-PFT systems may also vary substantially from that of NheA.

Serratia marcescens is a nosocomial human-pathogenic Gram-negative bacteria (Su *et al.*, 2003; Kurz *et al.*, 2003; Iguchi *et al.*, 2014). Genomic analysis has shown that it possesses an α-PFT with three proteins (SmhABC) homologous to the AhlABC proteins (Wilson *et al.*, 2019).

In this paper, we present the overexpression, purification and crystallization of SmhA and show the first electron-density map for an A component of a Gram-negative tripartite α-PFT.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Cloning and overexpression. The open reading frame for SmhA from *S. marcescens* MSU97 (NCBI accession No. OKB64935.1) was synthesized and cloned into the pET-21a expression vector by GenScript to contain a C-terminal His_6 tag.

The plasmid was transformed into an *Escherichia coli* BL21 (DE3) expression cell line (NEB). One colony was used to inoculate a 250 ml flask containing 50 ml Luria–Bertani (LB) broth supplemented with 100 µg ml^{-1} ampicillin and was grown overnight at 37°C. 10 ml of this overnight culture was then used to inoculate 500 ml LB broth supplemented and incubated as described above until an OD_{600} of 0.6 was reached, at which point protein expression was induced by the addition of 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG). Protein expression was carried out overnight at 16°C.

To prepare selenomethionine-incorporated SmhA, 2 × 500 ml of cells were grown as described above and harvested prior to induction. The cells were washed and resuspended in selenomethionine minimal medium [10.5 g l^{-1} K_2HPO_4, 1.0 g l^{-1} $(NH_4)_2SO_4$, 4.5 g l^{-1} KH_2PO_4, 0.5 g l^{-1} trisodium citrate·$2H_2O$, 5.0 g l^{-1} glycerol and 0.5 g l^{-1} each of adenine, guanosine, thymine and uracil; medium *A*] and added to 2 × 500 ml of medium *A* supplemented with 1.0 g l^{-1} $MgSO_4$·$7H_2O$, 4.0 mg l^{-1} thiamine; 100 mg l^{-1} each of L-lysine, L-phenylalanine and L-threonine; 50 mg l^{-1} each of L-isoleucine, L-leucine and L-valine; and 40 mg l^{-1} seleno-L-methionine. Growth was continued until an OD_{600} of 0.6 was reached before induction with 1 m*M* IPTG. The protein was expressed overnight at 16°C. The cells were harvested and pelleted before storage at −25°C.

2.1.2. Purification. Harvested cells of either native or selenomethionine-derivatized (SeMet) SmhA were defrosted, resuspended in lysis buffer (50 m*M* Tris pH 8.0) and lysed by sonication (3 × 20 s bursts at 16 µm amplitude). Insoluble material was removed by centrifugation at 40 000*g* for 15 min. The supernatant was applied onto a 5 ml nickel HiTrap column (GE Healthcare) in binding buffer (50 m*M* Tris pH 8.0, 0.5 *M* NaCl). The protein was eluted with a linear gradient of 0–1 *M* imidazole in binding buffer and fractions containing protein were pooled, concentrated and buffer-exchanged into 50 m*M* Tris pH 8.0, 10 m*M* NaCl for crystallization; the purification was analysed by SDS–PAGE. Macromolecule-production information is summarized in Table 1.

Table 1
Macromolecule-production information.

Source organism	*S. marcescens* MSU97
Restriction sites	NdeI/XhoI
Cloning vector	pET-21a(+)
Expression vector	pET-21a(+)
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced	MNNLTSIDLSPQTLMAMHISISSQALLNQS YSNLLLSQQLLTSQSMDPGLTVKIKAYQ NQLRQQAQVFKQNTVAELIGLYTKASNF AALVNAVNALYSTEDPQVSQKGAEMVAA LSDVAQHYQAAAQAVHTQLQAKREMLEP LMGNFLNVIDAIEQGLNAEAKQQAQTIA ELNEAIAKNIQSIADAGFKAGEGVVQLG QSIVAAVPLGPTDKKPKEAPTAPPKPLS DQASYMISGIQAISAGASGAQQAVNELK ANYAKLAVAYRALATANALLSVAKSVQA QAQLFVDTYVLTEQRMALLPTEWGKVAE AYLTAAPIINQAGSAAEIKQAKQIISLN AEKWQLFSKSIDNAKANYAGNNILPEVL EHHHHHH

2.2. Crystallization

Purified SmhA was concentrated to 7 mg ml^{-1} for crystallization using a Vivaspin 30 kDa molecular-weight cutoff concentrator (Sartorius). The concentrated protein was used to set up 96-well sitting-drop crystallization trials using a TTP LabTech Mosquito LCP robot, with both 200 nl:200 nl and 200 nl:100 nl well solution:protein solution drops, and stored at 7°C. Crystallization-condition suites used for preliminary screens included JCSG+, PACT *premier*, MPD, Morpheus, ProPlex and AmSO4 (Qiagen and Molecular Dimensions).

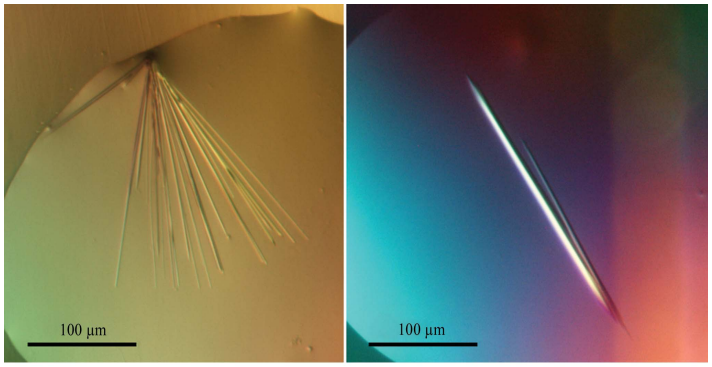
Initial crystals of both native and SeMet SmhA grew in PACT *premier* condition B11 (0.2 *M* MES pH 6, 0.2 *M* $CaCl_2$, 20% PEG 6000). Optimization (using a Formulatrix Formulator robot) of the SeMet SmhA crystals around PACT *premier* condition B11 in a 96-well sitting-drop plate with 200 nl:200 nl drops gave larger more defined crystals from 0.1 *M* MES pH 6.1, 0.14 *M* $CaCl_2$, 21% PEG 6000 (Fig. 1). Crystallization information is summarized in Table 2.

Table 2
Crystallization.

Method	Sitting-drop
Plate type	96-well sitting drop
Temperature (K)	280
Protein concentration (mg ml^{-1})	7
Buffer composition of protein solution	50 m*M* Tris pH 8.0, 10 m*M* NaCl
Composition of reservoir solution	0.2 *M* MES pH 6, 0.2 *M* $CaCl_2$, 20% PEG 6000
Volume and ratio of drop	200 nl:200 nl
Volume of reservoir (µl)	50

2.3. Data collection and processing

A single SeMet SmhA crystal was flash-cooled in liquid nitrogen using a cryoprotectant consisting of 20% ethylene glycol, 0.2 *M* MES pH 6, 0.2 *M* $CaCl_2$, 20% PEG 6000 and data were collected at the selenium absorption edge (0.9792 Å) on beamline I03 at Diamond Light Source (DLS; Fig. 2). Data



(*a*) (*b*)

Figure 1
(*a*) Native SmhA crystals grown in PACT *premier* condition B11 (0.2 *M* MES pH 6, 0.2 *M* $CaCl_2$, 20% PEG 6000). (*b*) SeMet SmhA crystals grown in optimized conditions based on PACT *premier* condition B11 (0.1 *M* MES pH 6.1, 0.14 *M* $CaCl_2$, 21% PEG 6000).

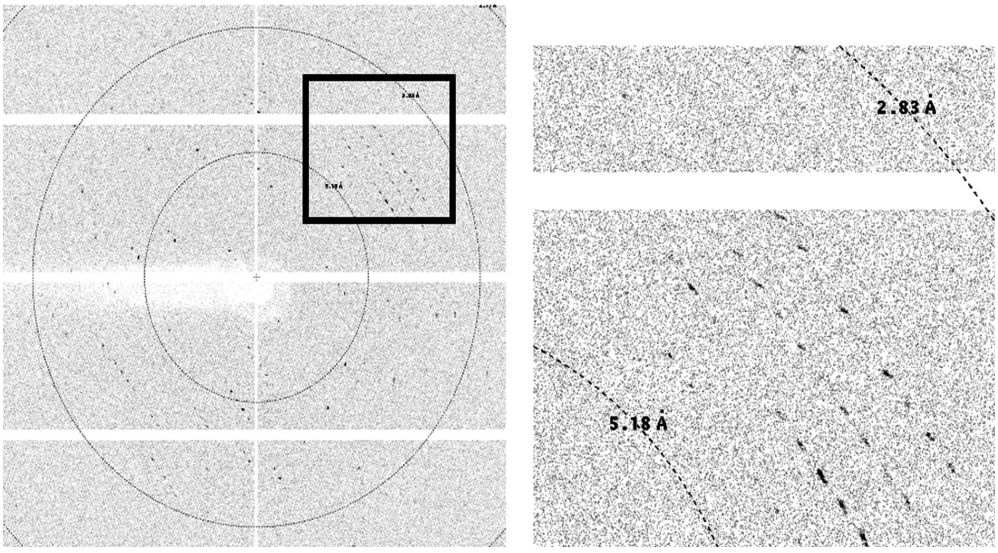


Figure 2
A representative 0.1° oscillation image from an SeMet SmhA crystal collected using an EIGER2 XE 16M detector on beamline I03 at Diamond Light Source. An enlarged view of the region highlighted by the square shows that diffraction extends to around 3.3 Å resolution.

were processed to 3.3 Å resolution using the *xia2*/*DIALS* pipeline (Winter *et al.*, 2018) and showed that the crystal belonged to space group $P4_2$, with unit-cell parameters $a = b = 151.8$, $c = 134.0$ Å (Table 3).

3. Results and discussion

3.1. Construct design

Structural studies of the tripartite toxins have been hampered in part by difficulties in producing large quantities of stable protein. The expression of *A. hydrophila* AhlA using constructs generated from genomic DNA in *E. coli* BL21 cells (Wilson *et al.*, 2019) produced protein for assays, but the yield was low and insufficient for crystallization. Pairwise sequence alignment of AhlA with *S. marcescens* SmhA shows 43% identity and 53% similarity (Supplementary Fig. S1), identifying SmhA as a good candidate for structural studies of the A component from a Gram-negative bacterium. As with AlhA, initial attempts to overexpress SmhA using constructs from genomic DNA also proved unsuccessful. To try to improve expression in *E. coli* BL21, the SmhA gene (NCBI accession

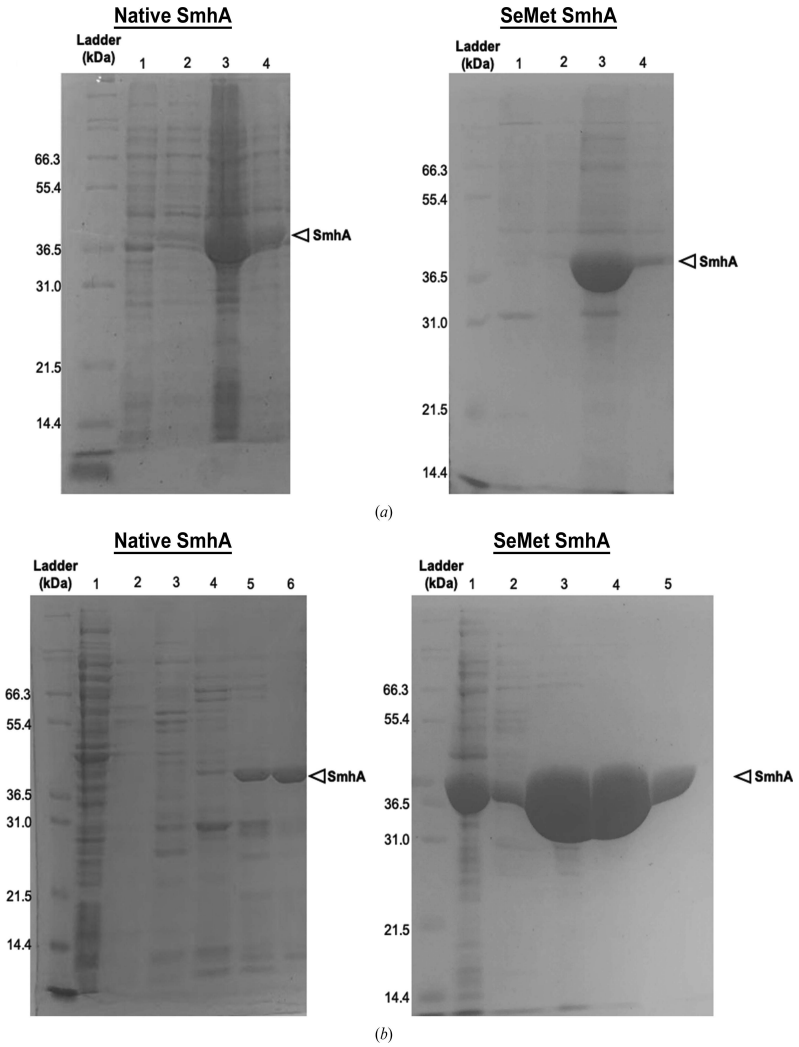


Figure 3
(*a*) SDS–PAGE gels showing overexpression of native and SeMet SmhA. Lane 1, pre-induction insoluble fraction; lane 2, pre-induction soluble fraction; lane 3, post-induction insoluble fraction; lane 4, post-induction soluble fraction. (*b*) SDS–PAGE gels showing nickel HiTrap column purification of native and SeMet SmhA. Lane 1, cell-free extract; lanes 2–5/6, elution fractions from the nickel HiTrap column. Fraction 6 and fractions 4 and 5 (native and SeMet, respectively) were >90% pure and were used for crystallization.

No. OKB64935.1) was synthesized and optimized in both GC content and codon usage for expression in *E. coli* (GenScript). This resulted in 38% of the codons being altered (Supplementary Fig. S2). This new construct was successfully used to express SmhA in *E. coli* with a C-terminal 6×His tag and, following nickel column purification, resulted in protein with >90% purity and a good yield (9.5 mg l^{-1}; Fig. 3).

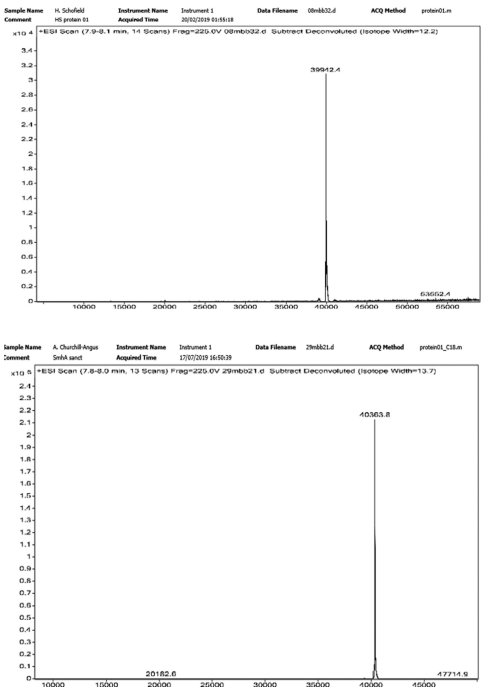


Figure 4
Mass spectrum for native SmhA (top) and SeMet SmhA (bottom) as used for crystallization. A molecular weight of 40 363.8 Da for SeMet SmhA (the molecular weight of native SmhA with a His_6 tag is 39 942.4 Da) shows the incorporation of selenium at all nine methionine sites.

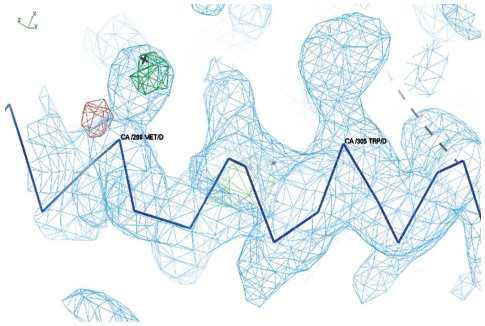


Figure 5
Initial electron-density map contoured at 1.0σ (blue) and an anomalous difference map (positive, green) showing a helical section of SmhA. Density can be seen for the side chains of Trp305 and Met298, with positive difference for the Se atom (shown as a cross) in Met298.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	I03, DLS
Wavelength (Å)	0.9792
Temperature (K)	100
Detector	EIGER2 XE 16M
Rotation range per image (°)	0.1
Total rotation range (°)	360
Exposure time per image (s)	0.008
Space group	$P4_2$
a, b, c (Å)	151.8, 151.8, 134.0
α, β, γ (°)	90, 90, 90
Mosaic spread (°)	0.19
Resolution range (Å)	67.9–3.34 (3.40–3.34)
Total No. of reflections	612546 (30358)
No. of unique reflections	44270 (2175)
Completeness (%)	100 (98.5)
Multiplicity	13.8 (14.0)
$\langle I/\sigma(I)\rangle$†	4.7 (1.0)
$R_{r.i.m.}$	0.118 (0.759)
Overall B factor from Wilson plot (Å^2)	56.626

† The high-resolution cutoff for the data was automatically determined in the *DIALS* pipeline, which uses $CC_{1/2}$ = 0.5 as the limit of usable data. For the outer shell the mean $I/\sigma(I)$ is 2.0 at 3.7 Å resolution.

3.2. Data analysis of SmhA

Analysis of the Matthews coefficient for SeMet SmhA showed that the asymmetric unit most likely contained between six and ten molecules with a solvent content between 63% and 38%, respectively, with eight molecules being the most probable, with a V_M value of 2.39 Å^3 Da^{-1} and a solvent content of 48% (Matthews, 1968; Kantardjieff & Rupp, 2003). Mass spectrometry showed that the molecular weight of SeMet SmhA was 40 363.8 Da, which is 421.4 Da more than the native sample, indicating full incorporation of Se atoms into the nine methionine residues of the protein and also indicating that the protein was of high purity (Fig. 4). In order to maximize the quality of the single-wavelength selenium anomalous signal, data-collection parameters were chosen to minimize radiation damage, whilst still providing good multiplicity, albeit at the expense of resolution. A beam size of 80 × 20 µm was selected to match the dimensions of the crystal, with a beam transmission of 20%, giving a flux of 7.44 × 10^{11} photons s^{-1}. 3600 images of 0.1° and an exposure of 0.008 s gave a data set with a half-set correlation coefficient of 1.0, an anomalous multiplicity of 7.0 and an anomalous correlation coefficient of 0.2 with no obvious signs of radiation damage. Selenium positions were calculated from these Se SAD data, and an initial density map and model were generated using the *CRANK*2 pipeline (Skubák & Pannu, 2013; Fig. 5). A preliminary initial model of SmhA, placing 2809 residues assigned to 28 fragments, was automatically built into the electron density. Visual inspection of the map and model confirmed that eight molecules were present in the asymmetric unit, with the side-chain positions of the methionine residues clearly aligned with the positive density of the

anomalous difference map. A self-rotation function calculated using the data between 50 and 6 Å resolution showed the presence of a noncrystallographic twofold axis perpendicular to the crystallographic fourfold (peak of 88% of the origin at polar coordinates 90.0, 111.2, 180°). Inspection of the initial model showed that two subunits were related by this rotation axis, but were separated by 20 Å. In addition, a self-Patterson indicated the presence of noncrystallographic translational symmetry with a peak of 43% of the origin at 0, 0, 0.467 and the model showed that six of the eight molecules were related by this translational symmetry. However, despite these noncrystallographic symmetry relationships, no high-order oligomeric arrangement could be observed for the eight subunits, indicating that the structure of SmhA was of the monomeric soluble form of the protein, rather than an oligomeric structure assembled around a central rotation axis as required for the proposed pore form.

Further work to extend the resolution of the data and refine the SmhA structure is ongoing.

Acknowledgements

We would like to thank Diamond Light Source for access to beamline I03 under proposal MX17773, which contributed to these results, and the beamline scientists for their assistance with data collection. Mass spectrometry was performed at the ChemMS Mass Spectrometry Centre at the University of Sheffield.

Funding information

Funding for this research was provided by: University of Sheffield (scholarship to Alicia M. Churchill-Angus; studentship to Thomas H. B. Schofield); EPSRC (grant No. EP/L026872/1 to Professor M. D. Ward).

References

Beecher, D. J. & Macmillan, J. D. (1991). *Infect. Immun.* **59**, 1778–1784.
Beecher, D. J., Schoeni, J. L. & Wong, A. C. (1995). *Infect. Immun.* **63**, 4423–4428.
Benke, S., Roderer, D., Wunderlich, B., Nettels, D., Glockshuber, R. & Schuler, B. (2015). *Nat. Commun.* **6**, 6198.
Fagerlund, A., Lindbäck, T., Storset, A. K., Granum, P. E. & Hardy, S. P. (2008). *Microbiology*, **154**, 693–704.
Ganash, M., Phung, D., Sedelnikova, S. E., Lindbäck, T., Granum, P. E. & Artymiuk, P. J. (2013). *PLoS One*, **8**, e74748.
Iguchi, A., Nagaya, Y., Pradel, E., Ooka, T., Ogura, Y., Katsura, K., Kurokawa, K., Oshima, K., Hattori, M., Parkhill, J., Sebaihia, M., Coulthurst, S. J., Gotoh, N., Thomson, N. R., Ewbank, J. J. & Hayashi, T. (2014). *Genome Biol.* **6**, 2096–2110.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kurz, C. L., Chauvet, S., Andrès, E., Aurouze, M., Vallet, I., Michel, G. P. F., Uh, M., Celli, J., Filloux, A., De Bentzmann, S., Steinmetz, I., Hoffmann, J. A., Finlay, B. B., Gorvel, J. P., Ferrandon, D. & Ewbank, J. J. (2003). *EMBO J.* **22**, 1451–1460.
Lindbäck, T., Fagerlund, A., Rødland, M. S. & Granum, P. E. (2004). *Microbiology*, **150**, 3959–3967.
Lund, T. & Granum, P. E. (1996). *FEMS Microbiol. Lett.* **141**, 151–156.
Madegowda, M., Eswaramoorthy, S., Burley, S. K. & Swaminathan, S. (2008). *Proteins*, **71**, 534–540.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R. & Ban, N. (2009). *Nature*, **459**, 726–730.
Roderer, D. & Glockshuber, R. (2017). *Philos. Trans. R. Soc. B*, **372**, 20160211.
Sastalla, I., Fattah, R., Coppage, N., Nandy, P., Crown, D., Pomerantsev, A. P. & Leppla, S. H. (2013). *PLoS One*, **8**, e76955.
Skubák, P. & Pannu, N. S. (2013). *Nat. Commun.* **4**, 2777.
Su, L.-H., Ou, J. T., Leu, H.-S., Chiang, P.-C., Chiu, Y.-P., Chia, J.-H., Kuo, A.-J., Chiu, C.-H., Chu, C., Wu, T.-L., Sun, C.-F., Riley, T. V., Chang, B.-J. & The Infection Control Group (2003). *J. Clin. Microbiol.* **41**, 4726–4732.
Thompson, N. E., Ketterhagen, M. J., Bergdoll, M. S. & Schantz, E. J. (1984). *Infect. Immun.* **43**, 887–894.
Wallace, A. J., Stillman, T. J., Atkins, A., Jamieson, S. J., Bullough, P. A., Green, J. & Artymiuk, P. J. (2000). *Cell*, **100**, 265–276.
Wilson, J. S., Churchill-Angus, A. M., Davies, S. P., Sedelnikova, S. E., Tzokov, S. B., Rafferty, J. B., Bullough, P. A., Bisson, C. & Baker, P. J. (2019). *Nat. Commun.* **10**, 2900.
Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Cryst.* D**74**, 85–97.

Chapter 5

Paper 3

5.1 Summary

The final paper in this thesis is an article describing the identification and characterisation of the ClyA family tripartite α-PFT Smh from *Serratia marcescens*. The work has shown that Smh is a tripartite toxin and that although it shares very low sequence identity to the Gram positive tripartite toxins Nhe and Hbl from *B.cereus*, its components share many functional similarities. These include their ability to form BC pro-pores before activation by addition of A, and their ability to form large inhibitory soluble complexes, characteristics not observed in *A. hydrophila* Ahl, the most closely sequence-related toxin to Smh. Further to this, structures of SmhA, soluble SmhB, and pore SmhB allowed detailed structural analysis of the family, illustrating how the β-tongue head and helical tail are well conserved in the family and contain conserved latches which must be broken in the transition from soluble to pore form. A structure-based model of a tripartite pore was generated using docking prediction, monoclonal antibody studies of Nhe (Didier et al., 2016; Dietrich et al., 2005), and structures from the Smh and Ahl toxins, and shows how the A component would be placed inside the pore providing a hydrophilic lumen as proposed in Paper 1 (Section 3.3). Finally, the paper describes a model of assembly for Smh and discusses how, although the

members of the tripartite α-PFT family are structurally similar, and will form very similar final pore arrangements, the path of assembly by which this is achieved varies greatly and likely aids each bacteria to colonise unique niches in the environment. Details of work done that was omitted, or was only briefly mentioned, from the results and discussion in this paper can be found in Appendix B, page 209.

5.2 Author contributions

PJB conceived the project; PJB, AMC.-A, THS, TRM designed the experiments; AMC-A, THS, TRM and SES performed the experiments; PJB, AMC-A, JSW, JBR. interpreted the data and PJB and AMC-A prepared the paper with input from all authors. Experimentally, I carried out the construct design, protein production and purification for all constructs. I carried out haemolytic assays, as well as all negative stain electron microscopy. I also carried out crystallisation, and X-ray data collection, as well as subsequent data analysis and phasing for all structures described. Finally, I built and refined SmhA, soluble SmhB, and pore SmhB structures and carried out all bioinformatics analysis and modelling.

5.3 Paper 3

Paper 3 manuscript will start on the next page.

Identification and characterisation of a tripartite α-pore forming toxin from *Serratia marcescens*

Alicia M. Churchill-Angus[1], Thomas H B Schofield[1,2], Thomas R. Marlow[1], Svetlana E. Sedelnikova[1], Jason S. Wilson[1], John B. Rafferty[1], and Patrick J. Baker[1*]

[1] Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield, South Yorkshire, S10 2TN, United Kingdom. [2] Current address: Astbury Centre for Structural and Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom. Correspondence and requests for material should be addressed to P.J.B. (email: p.baker@sheffield.ac.uk)

Abstract

Tripartite members of the ClyA family of α-PFTs have recently been identified in a number of pathogenic Gram-negative bacteria, including the human pathogen *Serratia marcescens*. Structures of a Gram-negative A component and a tripartite α-PFT complete pore are unknown and a mechanism for pore formation is still uncertain. Here we characterise the tripartite SmhABC toxin from *S. marcescens* and propose a mechanism of pore assembly. We present the structure of soluble SmhA, as well as the soluble and pore forms of SmhB. We show that the β-tongue soluble structure is well conserved in the family and propose two conserved latches between the head and tail domains that are broken on the soluble to pore conformational change. Using the structures of individual components, sequence analysis and docking predictions we illustrate how the A, B and C protomers would assemble on the membrane to produce a complete tripartite α-PFT pore.

Introduction

The Gram negative bacteria *Serratia marcescens* is widespread throughout the environment, with strains involved in marine, plant and animal infections [1–3]. Multiple strains have been identified as important antibiotic resistant nosocomial human pathogens [2], and as the causal agent in a number of infections including respiratory, urinary tract, septicaemia and meningitis [1,4]. Conversely, other plant-associated strains have been reported to promote plant growth and protect plants from infection by the production of antifungal and antibacterial compounds [5].

Recently a number of strains of *S. marcescens,* including plant-associated strain MSU97[5] and human pathogenic strain UMH7 [6], have been identified as containing homologues of the tripartite pore-forming toxin (PFT) AhlABC from *Aeromonas hydrophila* [7]. Tripartite α-PFTs are members of the ClyA α-PFT family [8]. The ClyA α-PFT family contains pores composed of one (*E.coli*, ClyA) [9–11], two (*Yersinia enterocolitica*, YaxAB and *Xenorhabdus nematophila*, XaxAB) [12,13] or three (*A. hydrophila*, AhlABC and *Bacillus cereus* NheABC and HblL_1L_2B) [7,8,14,15] protein components. Proteins in this family undergo a large scale conformational change from a compact soluble structure, where the hydrophobic residues are hidden often within a β-tongue motif, to an extended pore structure with the hydrophobic residues exposed in two extended α-helices which insert into the cell membrane [7,9,12–14]. To form a complete oligomeric pore all ClyA α-PFT family members share three common features, which are carried on different components in the bipartite and tripartite pores. These are: a short single leaflet spanning hydrophobic helix-turn-helix motif that provides the initial membrane binding event, carried on the ClyA β-tongue motif; or by YaxA and XaxA, or AhlC, NheC and Hbl-B, in the single, bipartite and tripartite toxins, respectively; a longer

membrane-spanning hydrophobic or amphipathic helical region (ClyA N-terminal helix; YaxB, XaxB; AhlB, NheB, $HblL_2$); and finally a hydrophilic lining to the interior of the pore (ClyA N-terminal helix; YaxB, XaxB; AhlA, NheA, $HblL_1$) [7]. In ClyA, and the bipartite pores YaxAB and XaxAB, this results in an oligomeric hydrophilic pore assembled from amphipathic helices, surrounded by the hydrophobic single leaflet anchoring components [10,12,13]. In addition to their haemolytic activity, recent studies of the Nhe and Hbl tripartite PFTs found in food poisoning strains of *B. cereus* show that both are able to activate the NLRP3 inflammasome leading to septic shock in mice, and inhibition of this response prevents lethality, demonstrating the importance of this toxin family in bacterial pathogenicity [16,17]. Understanding how these tripartite α-PFTs function is thus an important step in combatting infection and designing virulence targeted therapies.

Although significant progress has been made on structural studies of tripartite α-PFTs driven by the discovery of tripartite α-PFTs in Gram negative bacteria [7], structures of the A component from a Gram negative species or the complete tripartite pore remain elusive. Structures of either of these would allow for a better understanding of how the A, B and C components fulfil the three features observed in other ClyA α-PFT family members. The high sequence similarity between the proteins in Gram negative tripartite PFTs means that the structure and function of the component proteins must be closely related across the different species [7]. As *S. marcescens* has implications in human infection and antibiotic resistance, together with its antimicrobial properties and prevalence in the environment, it is an ideal choice for further studies of tripartite PFTs.

Here we show that *S. marcescens* has a tripartite haemolytic α-PFT (SmhABC), and propose a mechanism of pore assembly. We present the structures of the soluble A component (SmhA), as well as the soluble and pore conformations of the B component (SmhB). We show how these structures share high structural similarity with other members of the ClyA α-PFT family, especially those of NheA and AhlB, and show that a chimeric Smh/Ahl pore retains full lytic activity. Using these structures along with those of the closely related *A. hydrophila* AhlABC toxin, we propose a mechanism of soluble to pore transformation, model membrane-associated complexes of the three proteins of the Gram negative tripartite α-PFTs and present a structure based model of the location of the A component in a tripartite PFT pore.

Results

SmhABC lyses erythrocytes and forms pores in erythrocytes and liposome membranes

Homology searches with *A. hydrophila* AhlABC have previously identified *S. marcescens* SmhABC as a possible tripartite α-PFT [18]. To confirm this assignment, SmhA, SmhB and SmhC were expressed, purified and used in haemolytic assays with erythrocytes. Individual components showed no lytic activity alone, neither did mixtures of SmhA+SmhB and SmhA+SmhC (Fig. 1A). However, incubating equimolar concentrations of SmhB and SmhC together with erythrocytes showed 5% lysis after 2 hours, while SmhA, SmhB and SmhC together showed 40% lysis in 1 hour and 80% in 2 hours (Fig. 1A). Negative stain EM images of reaction mixtures for SmhA+SmhB+SmhC, and SmhB+SmhC showed that for

95 SmhA+SmhB+SmhC, individual pores and pores bound to the membrane were present,
96 while for SmhB+SmhC pores bound to membranes were visible, indicating that SmhABC is
97 indeed an α-PFT (Fig. 1B).

98 To determine an assembly order for the SmhABC pore, erythrocytes were pre-incubated
99 with different components for 1 hour before addition of the remaining components. Pre-
100 incubation with either SmhB or SmhC resulted in 82 and 84% lysis after 1 hour, respectively,
101 double that of the mixture of SmhA+SmhB+SmhC without pre-incubation (p=0.005, and
102 p=0.03) (Fig. 1C). When erythrocytes were pre-incubated with mixtures of SmhA+SmhC or
103 SmhB+SmhC, 100% lysis was achieved in 1 hour, significantly greater than with no pre-
104 incubation (SmhA+SmhB+SmhC, p=0.0009) and to the single components alone (p=0.01 and
105 p=0.009) (Fig. 1C). Pre-incubation with SmhA or SmhA+SmhB followed by addition of the
106 other components resulted in 55% and 40% lysis respectively, with no significant difference
107 to the SmhA+SmhB+SmhC control with no pre-incubation (p=0.08 and p =0.90) (Fig. 1C). It
108 thus appears that pre-binding of either SmhB or SmhC alone, or together, to the membrane
109 increases pore efficiency, whereas SmhA only increases efficiency if SmhC is also present.

110 For many ClyA family α-PFTs, efficacy of the pore is controlled by regulating the formation
111 of soluble off pathway complexes [14,19,20]. Lytic assays with varying concentrations of SmhA,
112 SmhB or SmhC were used to study any effect of ratio on potency of the SmhABC system.
113 When SmhA concentration was varied maximal lysis was achieved up to a 1:1:1 ratio (A:B:C),
114 lysis then decreased steadily to a minimum of 34% at a ratio of 4:1:1 and remained at this
115 value up to a ratio of 8:1:1 (Fig. 1D). When varying both SmhB and SmhC, lysis increased up
116 to a maximum at a ratio of 1:1:1, however, while no reduction of activity was observed for
117 SmhC at higher concentrations, lytic activity dropped to 30% at a ratio of 1:2:1 and

remained at 30% lysis at higher concentrations of SmhB (Fig 1D). This showed that both SmhA and SmhB have an inhibitory effect on lysis at high concentrations. Size exclusion analysis of soluble Smh proteins at 37 °C showed that mixtures of SmhB and SmhC form a large soluble complex (~890 kDa) while mixtures of SmhA and SmhC precipitate at 37°C. In addition, SmhA, SmhB and SmhC together in solution formed a 1 MDa complex containing all three proteins (Supplementary fig. 1). Incubation of this soluble SmhABC complex with erythrocytes showed no lysis after 1 hour. Similar assays with the soluble SmhBC complex also showed no lysis, and addition of SmhA also resulted in no lysis after a further 1 hour incubation. As both the soluble SmhABC and SmhBC complexes are both inhibitory to lysis, this suggests that the Smh pore efficacy is regulated by the formation of off pathway soluble complexes, as seen in other family systems from *Bacillus, Yersinia* and *Xenorhabdus* [7,13,19,21]. This control, however, does not involve any potential SmhA/SmhB complex, as none could be observed on gel filtration and pre-incubating mixtures of SmhA and SmhB with erythrocytes followed by addition of SmhC gave no difference in lysis efficiency ($p = 0.9411$, Fig. 1C).

Structure of SmhB

The proteins of the Smh PFT system of *S. marcescens* share high levels of sequence identity with those of the Ahl system from *A. hydrophilia* (46, 62, 43% identity, respectively), however, activity of the Ahl PFT is not regulated by off pathway soluble complexes [7]. To further characterise the Smh PFT we determined the structure of SmhB to compare to AhlB. The soluble form of SmhB was solved at a resolution of 1.84 Å in space group $P2_12_12_1$ by molecular replacement using the soluble form of AhlB [7] (PDB code 6GRK) as a search model (Supplementary fig. 2A) (Table 1). As expected, soluble SmhB folds into the family observed

β-tongue structure, with a five-helix bundle tail (α1, α2, α3, α6, α7), and a head domain containing three α-helices (α4, α5, α8) alongside a four-strand β-sheet (β1, β2, β3), which includes the β-tongue motif (β1 and β2). The head domain of SmhB is largely hydrophobic. This is particularly evident in the β-tongue and helices α4 and α5, which together constitute the predicted hydrophobic transmembrane region (G175 - L233), (Fig. 2A). The crystal asymmetric unit contained two closely related monomeric AhlB molecules (RMSD Cα 0.56 Å) with the only significant differences in the tail domain, where α7 from chain A folds as a single helix with a kink at S320, while in chain B α7 is divided into α7a and α7b by a short loop region (T319-V321). Chain A will be described in this section unless otherwise stated.

A structure of the SmhB pore was also determined at a resolution of 6.98 Å by molecular replacement using the AhlB pore (PDB code: 6H2F) as a search model (Supplementary fig. 2B) (Table 1). Although the resolution of this structure was low, the electron density was clear and continuous for the protein chain, enabling a model to be built with residues truncated at Cβ. In addition, clear positive difference density could be observed in omit maps calculated with sections of the model deleted (residues 157-251 in chain A and B), indicating that the structure solution was correct (Supplementary figure 2C). Like the AhlB pore, the SmhB pore contained ten monomers of SmhB in two conformations (type 1 and type 2) which assemble into a ring with pseudo 10-fold symmetry. As seen in AhlB, the head domain of SmhB undergoes a large scale conformational change transforming from the soluble to pore forms, with the β-tongue substructure refolding to form extended hydrophobic helices to α3 and α4 (Fig. 2B).

The *S. marcescens* Smh and the *A. hydrophila* Ahl α-PFTs have greatly varying rates of lysis of horse erythrocytes, with the AhlABC α-PFT requiring 100 times the concentration of each component to give equivalent rates of lysis as the SmhABC system [7]. As the overall structures of the soluble and pore forms of SmhB are almost identical to those of AhlB (RMSD Cα 0.64 Å and 0.46 Å, respectively), (Fig. 2C), not surprising given the high level of sequence identity (62%), we investigated if these structural similarities also conveyed functional similarities. Haemolytic assays, replacing the SmhB component of the SmhABC pore with AhlB in a 1:1:1 ratio showed that this chimeric pore had the same lytic activity as the SmhABC pore (Fig. 2C). The differences in lytic activity between the *Serratia* and *Aeromonas* tripartite α-PFTs are thus not due to the B component alone.

Structure of soluble SmhA

To further investigate the differences in lytic activity between SmhABC and AhlABC the crystal structure of soluble SmhA was determined to 2.78Å in space group $P2_12_12_1$. This crystal form contained four closely related molecules in the asymmetric unit (rmsD Cα between molecules of 0.4-0.6 Å), with no oligomeric structures observed (Table 1)(Supplementary fig. 2D).

SmhA folds into two distinct domains, a compact five helical bundle (α1, α2, α3, α6 and α7), and an associated head domain containing two helices (α4 and α5), and the β-tongue domain (two long anti-parallel β-strands, β1 and β2) (figure 3A). In all monomers, there is missing density for approximately 20 residues between β1 and β2 (G208 – A229). Residues E192-V205 (β1 and loop to α3) and Y231-N250 (β2 and loop to α4) of the β-tongue domain are largely hydrophobic containing only polar uncarged and nonpolar residues, and are

188 buried in the core of the protein, packing against α1, α3, α4, α5 (residues A272-Q283), and
189 residues N351-E367 at the C-terminus (Fig. 3A).
190
191 Structure based sequence alignments showed that SmhA shares strong similarities with
192 components from other tripartite systems that have a β-tongue substructure, such as SmhB
193 (PDB code 6ZZ5, rmsD 2.9 Å), NheA (PDB code 4K1P, rmsD 3.2 Å), AhlB (PDB code 6GRJ,
194 rmsD 3.3 Å), Hbl-B (PDB code 2NRJ, rmsD 2.8 Å), and also the *Vibrio cholera*e cytotoxin
195 MakA (PDB code 6DFP, rmsD 3.5 Å), that is, at present, not identified as a ClyA tripartite
196 family member (figure 3B). All of these structures share the same compact helical bundle
197 tail domain and β-tongue domain. Within the tail domain, greatest variation is seen in the
198 length of the N-terminal region. All structures have a N-terminal helix of between 22 and 28
199 residues, which packs against the β-tongue domain. However, in SmhA, NheA and Hbl-B,
200 this region has an additional extended loop and short helix that extends the length of the
201 tail domain, a feature not observed in SmhB or AhlB. In all of these structures, the C-
202 terminus is used to shield the hydrophobic residues of the β-tongue, but the method by
203 which the C-terminus hides the hydrophobic tail is different. MakA, NheA, and Hbl-B all use
204 β-strands packing against the hydrophobic region of β1. AhlB and SmhB use both a short β-
205 strand and a short helix connected by a loop, while SmhA uses only a short loop
206 perpendicular to β1 (Fig. 3B).
207
208 It has been shown that the soluble β-tongue conformation of the ClyA family proteins
209 rearranges into an extended pore conformation in both ClyA and AhlB [7,9], and we have
210 shown the same occurs in SmhB**.** As soluble SmhA, SmhB and AhlB all share very similar
211 structures, we propose that the head domain of SmhA would also unfold into its pore form

about 2 hinges in the same way as seen in SmhB and AhlB, with hinge 1 between α3 and α4, and α5 and α6, and hinge 2 between α4 and β1, and α5 and β2. This would result in the pore form of SmhA having a similar extended helical structure to that of AhlB and SmhB.

A model of this proposed pore form of SmhA was generated using a Dali [22] structure based sequence alignment to map the residues of SmhA onto the structure of AhlB type 2 head (supplementary fig. 3). In this model the mixed hydrophobic/hydrophilic residues of the β-tongue head of SmhA form two extended amphipathic helices 36 Å in length at the ends of α3 and α4, (Fig. 3C), sufficient to cross the membrane, and thus with the potential to form a hydrophilic lining to the pore. In the pore conformation of SmhA, residues G208 – A229 lie between the C-terminal and N-terminal ends of the amphipathic helices α3 and α4, respectively, and would thus occur on the intracellular side of the target membrane. These residues, omitted in the soluble structure due to very weak electron density, are presumably quite flexible. This intracellular loop between α3 and α4 is present in the A components of all the Gram negative tripartite toxins, but is not found in either the Gram positive Nhe or Hbl tripartite toxins, the single component ClyA [10,23], or bipartite YaxAB [12] and XaxAB [13] toxins. Residues on this loop, which do not show any pattern of conservation within the Gram-negative A component sequences, could thus interact with different intracellular components of the target cell, depending on the bacterial species the toxin is found in.

Soluble to pore conformational change

The transformation from a soluble β-tongue containing structure to an extended helical pore structure seen in SmhB and modelled in SmhA, occurs in a similar way for *A. hydrophila*

AhlB [7]. In order for this conformational change to occur, the C-terminal loop and helix α7 must move to free the β-tongue to form the two extended trans-membrane helices in the pore conformation (α3-α4), with the N-terminal helix (α1) also moving from parallel to α7 to an end to end packing, resulting in a narrower more compact tail. Sequence analysis of possible family member homologues and mapping of conserved residues onto the soluble structures of SmhA, SmhB and NheA shows that many of the conserved residues are located in the head domain and are focused on two regions. First, a leucine zipper forms between conserved Leu and Ile residues in α4 and α5 (Fig. 4), packing the hydrophobic membrane-spanning residues in the core of the domain. The second region is between the N-terminus and β-tongue, which includes a family conserved glutamine (Q38 SmhA, Q32 SmhB, Q45 NheA) on the N-terminal helix α1 that forms hydrogen bonds with the β-sheet backbone of the β-tongue (Fig. 5A). In SmhA Q38 and in NheA Q45 also interact with Q263 and Q282, respectively on α6, a residue conserved in all A component proteins (Supplementary fig. 4). In addition, both SmhA and SmhB contain a family conserved (conserved in all but 1 sequence) lysine residue at the C-terminus (K333 in both SmhA and SmhB) that forms a hydrogen bond with a conserved hydroxyl (T81 SmhA; T78 SmhB) on helix α2 (Fig. 5B, Supplementary fig. 4). These three motifs also occur in the structures of AhlB [7] and NheA [24], but in NheA two glutamines (Q351, Q83) hydrogen bond at the position of the Lys to Thr interaction in SmhA and SmhB (Fig. 5B). As the interactions between these conserved residues are present in the known soluble structures of all the tripartite A and B components and are disrupted during the soluble to pore transition on pore formation, it is likely that these residues play an important role in the interaction between the head and tail domains, possibly acting as latches to precipitate the conformational change of all of this β-tongue containing tripartite α-PFT proteins.

260

261 **Discussion**

262 Studies of the ClyA family tripartite PFTs have shown a number of mechanisms by which
263 different family members assemble their active pores. Nhe is proposed to first form BC pro-
264 pores by binding of BC soluble complexes to the membrane before recruitment of further B
265 to form the pro-pore, followed by the rapid addition of A and lysis [21,25,26]. Hbl first binds Hbl-
266 B (equivalent to the SmhC component) before rapid recruitment of Hbl-L_1 (B-component)
267 and finally Hbl-L_2 (A-component) [20]. Ahl primes the membrane with AhlC allowing the
268 recruitment of AhlB and AhlA[7]. All of these three tripartite toxins also vary in their required
269 ratio for optimum activity, with both Nhe and Hbl forming inhibitory complexes when either
270 the C component is in excess (NheC) or the A or B component are in excess (Hbl-L_1, Hbl-L_2),
271 while no inhibitory effects are seen for any of the Ahl components, with proposed ratios of
272 10:10:1, 1:1:10 and 1:1:1 (A:B:C) for Nhe, Hbl and Ahl respectively [14,18–20,25].

273 Here we have shown how *S. marcescens* contains a ClyA family tripartite PFT, Smh, which
274 like the other tripartite family members requires all three proteins, SmhA, SmhB and SmhC
275 to form a lytic pore. Our assays suggest that the Smh toxin employs yet another variation on
276 the tripartite pore assembly mechanism. We have shown that the presence of both SmhB
277 and SmhC is a minimum requirement for pore formation, with maximum lysis achieved
278 when erythrocytes are preincubated with SmhC together with either SmhB or SmhA.
279 Preincubation with SmhB or SmhC alone enhances lysis but not to the same extent, while
280 pre-incubation with SmhA has no effect. This suggests that although SmhC and SmhB are
281 both able to independently enhance lysis, in combination they are most efficient and can
282 form pro-pores before the addition of SmhA and full lysis (Fig. 6). A similar pore assembly
283 mechanism is also seen in the Nhe toxin [25], but is in contrast to Ahl, where formation of

AhlBC pores is an off-pathway process and inhibitory to lysis [7]. Our assays have also shown that SmhA and SmhB act to inhibit lysis when present at concentrations higher than that of SmhC, and form large inactive complexes with SmhC, a feature also shared by Hbl, but in contrast to Nhe, where NheC is inhibitory even at low concentrations [19–21]. The *A. hydrophila* AhlABC system, which is the most closely related in sequence to SmhABC, does not exhibit this inhibitory soluble complex formation, instead, the soluble AhlC component forms soluble tetramers that appear to preclude the formation of any soluble complexes with AhlA or AhlB, by occluding any binding surfaces of AhlC within the tetramer [7]. In contrast, SmhC is monomeric in solution, enabling soluble inhibitory complexes to form in the Smh toxin. Off pathway soluble complex formation is also observed in the bipartite ClyA family α-PFTs YaxAB and XaxAB [12,13]. In these toxins, both the A (equivalent to SmhC) and B components are monomeric in their soluble forms, as is also the case Hbl-B and so accessible for soluble complex formation like SmhC[12,13,27]. We propose the off pathway soluble SmhBC and SmhABC complexes act to control the efficacy of the Smh toxin. Although there is still some ambiguity in the exact order of membrane association for the Smh toxin, our TEM and haemolytic assays show that SmhBC pro-pores are definitely able to assemble, with SmhA interacting with the SmhBC pro-pore to destroy the cell. However, our data are also consistent with the Smh components binding sequentially at the membrane to form a fully lytic pore, as seen with Hbl [20] (Fig. 6).

We have identified two potential latches in the soluble tripartite β-tongue containing components, and the residues involved are located in surface patches of conserved residues in both SmhA and SmhB (Supplementary fig. 5). Monoclonal antibody binding studies that interfere with interactions between the three components of the *B. cereus* Nhe α-PFT [25,28,29]

have identified the binding sites between NheA (Mab site 2G11) to NheB (Mab site 1E11) and NheB (Mab site 2B11) to NheC (Supplementary fig. 6). Mapping of these binding sites onto the SmhA and SmhB structures, (Supplementary fig. 6), show that SmhC would interact with SmhB at the patches of conserved residues on SmhB containing both latch 1 and latch 2. Likewise, the interaction of SmhB with SmhA would also involve the latch motifs on SmhA. It thus seems likely that the β-tongue containing components of the tripartite ClyA family toxins share a conserved mechanism of transformation from soluble to pore form. The B and C components interact at a binding interface located around residues 95-123 of SmhB, this could break latches 1 and 2 on the B component, releasing the β-tongue, C-terminus, and N-terminus to initiate the transformation into the B component pore conformation (Fig. 7). The A component then binds to B at an interface around residues 289-308 in SmhB and 57-71 in SmhA, similarly breaking the two latches on the A component, and allowing its conformational change from soluble to pore form (Fig. 7). The full conformational change of the β-tongue to the pore forms of both A and B components requires a membrane, otherwise off pathway soluble complexes occur [12,19,21]. The release of the C-terminus from the β-tongue agrees with the observation that the start of the conformational transition in NheA occurs at the C-terminus [25]. This is further evidenced in the structure of a second crystal form of SmhB (Table 1), where comparison of the B-factors for the two chains in the asymmetric unit revealed high flexibility at residues 85-124 and residues 280-320 (Supplementary Fig. 7), corresponding to the proposed binding sites to SmhC and SmhA, indicating that movement in these binding regions can occur to facilitate complex formation.

These predicted binding sites between the A, B and C components, the structures of SmhA, SmhB and AhlC (as a surrogate for SmhC), and structure based sequence alignments were used to construct models of AhlC bound to soluble SmhB; AhlC bound to SmhB type 2 pore conformation and the AhlC: SmhB complex bound to soluble SmhA with the HADDOCK server [30] (Fig. 7). Ten of these SmhA, SmhB, SmhC units were subsequently assembled into the complete pore with C10 symmetry by rotation of increments of 36° around the central axis of the SmhB pore (Fig. 8). This model of the SmhABC pore is consistent with that proposed for AhlABC [7] with an inner ring of alternating membrane-spanning hydrophobic B and amphipathic A components forming the hydrophilic pore, surrounded by the single leaflet anchoring C components. This general pore architecture is also seen in the greater ClyA family, but with the individual roles of the three tripartite PFT proteins of a single leaflet anchoring function, a membrane-spanning function and a hydrophilic pore lumen carried on two proteins for the bipartite YaxAB/XaxAB PFTs [12,13] and on a single protein for the prototypical ClyA toxin [9,10]. Within the tripartite α-PFT family members, both the structures of the A and B components and the mechanism for transformation from soluble to pore forms of these components are well conserved. Nevertheless, the involvement of three proteins in pore formation has enabled differences in both the assembly and in the regulation of each tripartite PFT system, helping these bacterial species to thrive in their different environmental niches.

Materials and Methods

Protein cloning and purification

354 The ORFs for SmhA, SmhB, and SmhC from *S. marcescens* MSU-97 were synthesised and
355 cloned into pET21a expression vectors by GenScript Biotech Corporation, so as to contain a
356 C-terminal 6-His tag. Each protein was expressed in *E.coli* BL21 DE3 expression cell line
357 (NEB). Cultures were grown in LB media at 37 ^{o}C until an OD_{600} 0.6 was reached, protein
358 expression was then induced using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). For
359 all three proteins expression was carried out at 16 ^{o}C overnight. SmhA and SmhB were
360 purified using the same protocol. The cell pellet was resuspended in lysis buffer (50 mM Tris
361 pH8) and sonicated (3x20 s burst at 16000 nm λ), insoluble material was removed by
362 centrifugation at 40000 *g* for 15 minutes. Soluble protein in binding buffer (50 mM Tris
363 pH8, 0.5 M NaCl) was applied to a 5ml Nickel Hi-trap column (GE Healthcare) and protein
364 was eluted using a gradient of 0-1 M imidazole in binding buffer. Protein was further
365 purified by size exclusion chromatography on a Superdex 200pg column (GE Healthcare)
366 pre-equilibrated with 50 mM Tris pH8, 0.5 M NaCl. The SmhC cell pellet was sonicated and
367 centrifuged as described for SmhA and SmhB. As the protein did not bind to the Ni column,
368 an ammonium sulphate cut was carried out on soluble SmhC to a final ammonium sulphate
369 concentration of 0.5-1.5M. After 10 minutes precipitated protein was pelleted by
370 centrifugation at 70000 *g* for 5 minutes at 4 ^{o}C, the pellet was resuspended in 50mM Tris
371 pH8. Soluble protein was applied to a 10 ml (2 x 5 ml) HiTrap DEAE Sepharose Fast Flow
372 column (GE Healthcare) pre-equilibrated in 50 mM Tris pH 8 and eluted on a gradient of 0-
373 0.3 M NaCl in 50 mM Tris pH8, 1 M NaCl. Selenomethionine SmhA protein was expressed in
374 the same way as the S-met SmhA, as detailed previously [31].
375
376 **Generation of homologue sequence alignments**

Homologues were identified using the BlastP server [32] and all results with an E <0.01 were assessed to determine if they were part of a tripartite toxin operon. All homologues identified as part of a tripartite PFT were then used in subsequent sequence alignments. Sequence alignments of homologues of SmhA and SmhB were generated using Tcoffee [33].

Haemolytic assays

The lytic activity against horse erythrocytes of SmhA, SmhB, and SmhC was determined by measuring the release of haem from lysed cells as described by Rowe and Welch [34]. A 0.5 % v/v suspension of horse erythrocytes (Thermo scientific) was prepared by repeated washing of cells and resuspension in 10 mM PBS pH 7.4. Varying concentrations of each SmhABC protein were incubated with 0.1 ml erythrocyte suspension on a blood wheel at 37 ^{o}C for 1 hour, or cells were preincubated with various components for one hour followed by addition of the remaining components followed by a further 1 hour incubation. Erythrocyte mixtures were centrifuged at 1500 *g* for 5 minutes, and supernatant removed for photometric analysis at 542 nm. A positive control of erythrocytes lysed with ddH_2O and a negative control with no protein were used to normalised the data for 0 and 100% lysis. All assays were carried out in triplicate.

Analysis of soluble complexes of SmhABC

To determine whether individual components of Smh could bind to each other in the soluble conformation, 40 μM of each protein in 10 mM PBS pH 7.4 were mixed in various combinations and incubated for 45 minutes at 37 ^{o}C and then applied to a Superdex 200 increase gel-filtration (GE Healthcare) column pre-equilibrated with PBS. Fractions were collected and analysed by SDS-PAGE.

401

402 **SmhA crystallisation and structure determination**

403 Purified SmhA was concentrated to 7 mg/ml in a Vivaspin 30 KDa MWCO concentrator
404 (Sartorius), and then buffer exchanged into 50 mM Tris pH8, 10 mM NaCl. Crystals were
405 grown by sitting drop vapour diffusion in 96-well plates. Se-methionine crystals grew in 0.1
406 M MES pH 6.5, 0.16 M $CaCl_2$ and 20 % PEG 6000 (7 ^oC, 200 nl:200 nl drop), while S-
407 methionine crystals grew in 0.2 M potassium nitrate, 20% PEG 3350 (7 ^oC, 200 nl:200 nl
408 drop). Crystals were cryo-protected in mother-liquor containing an additional 20% (v/v)
409 ethylene glycol before freezing in liquid nitrogen.

410

411 X-ray diffraction data from a single Se-methionine SmhA crystal were collected on beamline
412 i03 of the Diamond Light Source (DLS) at wavelength 0.9792 Å. Images were integrated and
413 scaled using the Xia2 Dials pipeline[35,36] into space group $P4_2$ at a resolution of 2.98 Å (Table
414 1). Initial phases were obtained by SAD, with the heavy atom sites and initial electron
415 density map and model calculated using the CRANK2 pipeline[37]. The model was optimised
416 and completed in Coot[38]. This crystal form 1 (PDB code 7A26) contained eight independent
417 molecules in the asymmetric unit and the structure of a single chain was used as a
418 molecular replacement model to determine the structure for a second higher resolution S-
419 methionine data set (2.57 Å), in space group $P2_12_12_1$. This second data set was collected on
420 beamline i04-1 of the DLS at wavelength 0.9159 Å and images were processed using Xia2
421 3dii (Table 1). Rebuilding and refinement were carried out using COOT and REFMAC[39], to
422 give a final model with R and Rfree of 0.23 and 0.25, respectively (PDB code 7A27). Residues
423 210-228 in chain A, 209-228 in chain B, 209-225 in chain C and 210-225 in chain D were
424 omitted from the final model due to poor density.

425

426 **SmhB crystallisation and structure determination**

427 Purified SmhB was concentrated to 14 mg/ml in a Vivaspin 30 KDa MWCO concentrator
428 (Sartorius) and then buffer exchanged into 50 mM Tris pH8, 10 mM NaCl. All crystals were
429 grown by sitting drop vapour diffusion in 96-well plates. Crystals grew in 0.17 M ammonium
430 sulphate, 25.5 % PEG4000 (16 ^{o}C, 100 nl:100 nl drop) and were cryo-protected in mother-
431 liquor containing an additional 20% (v/v) ethylene glycol before freezing in liquid nitrogen.
432 X-ray diffraction data from a single SmhB crystal was collected on beamline i04-1 of the DLS
433 at wavelength 0.9159 Å. Images were integrated and scaled using the Xia2 3dii[36] pipeline
434 into space group $P2_12_12_1$ at a resolution of 1.84 Å (Table 1). The structure was determined
435 by molecular replacement with PhaserMR[40] using the soluble conformation of AhlB (PDB:
436 6GRK) as a search model. Rounds of model building and refinement were carried out using
437 COOT[38] and REFMAC[39] to give a final model with an R and Rfree of 0.19 and 0.22,
438 respectively (PDB code 6ZZ5). Residues 206-207 in chain A and 206-208 in chain B were
439 omitted from the final model due to poor density.

440

441 A second crystal form of SmhB was grown in 0.22 M Magnesium chloride hexahydrate
442 0.1 M Na acetate pH 5.15, 26% (w/v) PEG 6000 (16 ^{o}C, 100 nl:100 nl). Data was collected on
443 beamline i04-1 of the DLS. Images were integrated and scaled using AutoPROC+STARANISO
444 [41–45] into space group $P2_1$ at a resolution of 1.86 Å. The structure was determined by
445 molecular replacement with PhaserMR [40] using a single chain of the Form 1 SmhB structure
446 as a search model. Rounds of model building and refinement were carried out using COOT [38]
447 and REFMAC (39) to give a final model with an R and Rfree of 0.21 and 0.25 respectively

448 (PDB code 6ZZH). Residues 205-206 in chain A and B were omitted from the final model due
449 to poor density.
450
451 The pore structure of SmhB was crystallised in 0.2 M Calcium chloride, 40% MPD (16 °C, 100
452 nl:100 nl). Data were collected from on beamline i24 of the DLS at wavelength 0.9795 Å.
453 Images were integrated and scaled using Dials[35] into space group C2 at a resolution of 6.98
454 Å. The structure was determined by molecular replacement with PhaserMR (40) using the
455 AhlB pore conformation (PDB: 6GRJ) as a search model. Rounds of model building and
456 refinement were carried out using COOT (38) and REFMAC [39] to give a final model with an R
457 and Rfree of 0.33 and 0.33 respectively (PDB code 7AOG). The final model was constructed
458 from poly-alanine as density for side chains could not be resolved at this resolution.
459 Structure alignments in this paper were done using Dali[22].
460
461 **Modelling NheB and C, and docking of A, B, C complexes**
462 Homology models of NheB and NheC were generated using Phyre2 [46] and the structure of
463 Hbl-B (PDB: 2NRJ) as the template. Models of potential complexes between the components
464 in the tripartite a-PFT family were generated using structures of individual components from
465 the Smh and Ahl systems, predicted binding regions determined from antibody binding
466 studies on the Nhe system and the HADDOCK2.2 [30] server. Multiple runs of Haddock were
467 completed and only consistent solutions were used in subsequent modelling. In this way
468 models of complex structures between AhlC (PDB:6H2E) and SmhB (PDB:6ZZ5) (predicted
469 docking regions residue 97-125 in SmhB, 79-118 in AhlC chain Q); AhlC (PDB:6H2E) and AhlB
470 T2 (PDB: 6H2F) (residues 97-125, 176-184 and 227-234 AhlB T2 , 79-118 and 153-171 in

AhlC chain Q); and AhlBT2 (PDB: 6H2F) and SmhA (PDB: 7A27) (residues 289-319 in AhlBT2, 57-71 in SmhA) were produced.

Electron microscopy

Lytic assay samples were visualised using negative stain TEM. 5µl of each respective sample was pipetted onto a glow discharged carbon-coated grid (copper 300 mesh), then stained with 1% (w/v) uranyl formate and air dried. Electron micrographs were collected with a Philips CM100 100 kV transmission electron microscope, equipped with a Gatan 1 K CCD camera. Micrographs were collected with a pixel to nm ratio of 0.72 pixels per nm.

Generation of figures

All protein cartoon and surface rendering used in the figures were generated using Pymol [47].

Data availability

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. Atomic coordinates for SmhA crystal form1 (PDB code 7A26), SmhA Form 2 (PDB code 7A27), soluble SmhB form1 (PDB code 6ZZ5), soluble SmhB form2 (PDB code 6ZZH), SmhB pore (PDB code 7AOG), have been deposited in the RCSB Protein Data Bank.

References

1. Yu, V. Serratia marcescens. *N. Engl. J. Med.* **46**, 903–912 (1979).
2. Iguchi, A. *et al.* Genome evolution and plasticity of serratia marcescens, an important multidrug-resistant nosocomial pathogen. *Genome Biol. Evol.* **6**, 2096–2110 (2014).

3. Lacey, L. A. *et al.* Insect pathogens as biological control agents: Back to the future. *J. Invertebr. Pathol.* **132**, 1–41 (2015).
4. Su, L. H. *et al.* Extended epidemic of nosocomial urinary tract infections caused by Serratia marcescens. *J. Clin. Microbiol.* **41**, 4726–4732 (2003).
5. Matilla, M. A., Udaondo, Z., Krell, T. & Salmond, G. P. C. Genome Sequence of Serratia marcescens MSU97, a Plant-Associated Bacterium That Makes Multiple Antibiotics. *Genome Announc.* **5**, (2017).
6. Anderson, M. T., Mitchell, L. A., Zhao, L. & Mobleya, H. L. T. Capsule production and glucose metabolism dictate fitness during Serratia marcescens bacteremia. *MBio* **8**, (2017).
7. Wilson, J. S. *et al.* Identification and structural analysis of the tripartite α-pore forming toxin of Aeromonas hydrophila. *Nat. Commun.* **10**, 2900 (2019).
8. Fagerlund, A., Lindbäck, T., Storset, A. K., Granum, P. E. & Hardy, S. P. Bacillus cereus Nhe is a pore-forming toxin with structural and functional properties similar to the ClyA (HlyE, SheA) family of haemolysins, able to induce osmotic lysis in epithelia. *Microbiology* **154**, 693–704 (2008).
9. Roderer, D. & Glockshuber, R. Assembly mechanism of the α-pore-forming toxin cytolysin A from Escherichia coli. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, 20160211 (2017).
10. Eifler, N. *et al.* Cytotoxin ClyA from Escherichia coli assembles to a 13-meric pore independent of its redox-state. **25**, 2652–2661 (2006).
11. Benke, S. *et al.* The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Commun.* **6**, 6198 (2015).
12. Bräuning, B. *et al.* Structure and mechanism of the two-component α-helical pore-

519 forming toxin YaxAB. *Nat. Commun.* **9**, 1806 (2018).
520 13. Schubert, E., Vetter, I. R., Prumbaum, D., Penczek, P. A. & Raunser, S. Membrane
521 insertion of α-xenorhabdolysin in near-atomic detail. *Elife* **7**, 1–25 (2018).
522 14. Sastalla, I. *et al.* The Bacillus cereus Hbl and Nhe Tripartite Enterotoxin Components
523 Assemble Sequentially on the Surface of Target Cells and Are Not Interchangeable.
524 *PLoS One* **8**, e76955 (2013).
525 15. Beecher, D. J., Schoeni, J. L., Lee Wong, A. C. & Wong, A. C. Enterotoxic activity of
526 hemolysin BL from Bacillus cereus. *Infect. Immun.* **63**, 4423–4428 (1995).
527 16. Fox, D. *et al.* Bacillus cereus non-haemolytic enterotoxin activates the NLRP3
528 inflammasome. *Nat. Commun.* **11**, (2020).
529 17. Mathur, A. *et al.* A multicomponent toxin from Bacillus cereus incites inflammation
530 and shapes host outcome via the NLRP3 inflammasome. *Nat. Microbiol.* **4**, 362–374
531 (2019).
532 18. Wilson, J. Structural and Functional Studies of the AHL Tripartite Pore Forming Toxin
533 Proteins from Aeromonas hydrophila. (University of Sheffield, 2016).
534 19. Tausch, F. *et al.* Evidence for complex formation of the Bacillus cereus haemolysin BL
535 components in solution. *Toxins (Basel).* **9**, 288 (2017).
536 20. Jessberger, N. *et al.* Binding to the target cell surface is the crucial step in pore
537 formation of hemolysin BL from Bacillus cereus. *Toxins (Basel).* **11**, (2019).
538 21. Zhu, K. *et al.* Formation of small transmembrane pores: An intermediate stage on the
539 way to Bacillus cereus non-hemolytic enterotoxin (Nhe) full pores in the absence of
540 NheA. *Biochem. Biophys. Res. Commun.* **469**, 613–8 (2016).
541 22. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44**, 351–355 (2016).
542 23. Wallace, A. J. *et al.* E. coli hemolysin E (HlyE, ClyA, SheA): X-ray crystal structure of the

543 toxin and observation of membrane pores by electron microscopy. *Cell* **100**, 265–76
544 (2000).
545 24. Ganash, M. *et al.* Structure of the NheA Component of the Nhe Toxin from Bacillus
546 cereus: Implications for Function. *PLoS One* **8**, e74748 (2013).
547 25. Didier, A., Dietrich, R. & Märtlbauer, E. Antibody Binding Studies Reveal
548 Conformational Flexibility of the Bacillus cereus Non-Hemolytic Enterotoxin (Nhe) A-
549 Component. *PLoS One* **11**, e0165135 (2016).
550 26. Heilkenbrinker, U. *et al.* Complex Formation between NheB and NheC Is Necessary to
551 Induce Cytotoxic Activity by the Three-Component Bacillus cereus Nhe Enterotoxin.
552 *PLoS One* **8**, e63104 (2013).
553 27. Madegowda, M., Eswaramoorthy, S., Burley, S. K. & Swaminathan, S. X-ray crystal
554 structure of the B component of Hemolysin BL from Bacillus cereus. *Proteins Struct.*
555 *Funct. Genet.* **71**, 534–540 (2008).
556 28. Dietrich, R., Moravek, M., Bürk, C., Granum, P. E. & Märtlbauer, E. Production and
557 characterization of antibodies against each of the three subunits of the Bacillus
558 cereus nonhemolytic enterotoxin complex. *Appl. Environ. Microbiol.* **71**, 8214–8220
559 (2005).
560 29. Didier, A. *et al.* Monoclonal antibodies neutralize Bacillus cereus Nhe enterotoxin by
561 inhibiting ordered binding of its three exoprotein components. *Infect. Immun.* **80**,
562 832–8 (2012).
563 30. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative
564 Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
565 31. Churchill-angus, A. M., Sedelnikova, S. E., Schofield, T. H. B. & Patrick, J. The A
566 component (SmhA) of a tripartite pore forming toxin from Serratia marcescens :

expression , purification , and crystallographic analysis. submitted to *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* (2020).

32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

33. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).

34. Rowe, G. E. & Welch, R. A. Assays of hemolytic toxins. *Methods Enzymol.* **235**, 657–667 (1994).

35. Winter, G. *et al.* DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 85–97 (2018).

36. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).

37. Skubák, P. & Pannu, N. S. Automatic protein structure solution from weak X-ray data. *Nat. Commun.* **4**, 2777 (2013).

38. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K., Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D* **66**, 486–501 (2010).

39. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography* **53**, 240–255 (1997).

40. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

41. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 293–302 (2011).

42. Kabsch, W. *et al.* XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
43. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
44. Winn, M. D. *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
45. Tickle, I. . *et al.* STARANSIO. (2019).
46. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
47. Schrödinger, LLC. *The {PyMOL} Molecular Graphics System, Version~1.8*. (2015).

Acknowledgements

We would like to thank the Diamond Light Source for access to beamlines I03, I24 and I04-1 that contributed to these results, and the beamline scientists for their assistance with data collection. Electron Microscopy was performed in the Faculty of Science Electron Microscopy Facility at the University of Sheffield. Mass spectrometry was performed in the ChemMS Mass Spectrometry Centre at the University of Sheffield, EPSRC (Grant No. EP/L026872/1 to Prof. M D Ward). This work was supported by BBSRC Doctoral Training Grant BB/F016832/1 awarded to JSW and University of Sheffield scholarships awarded to JSW and AMC-A.

Author Contributions

PJB conceived the project; PJB, AMC.-A, THS, TRM designed the experiments; AMC-A, THS, TRM and SES performed the experiments; PJB, AMC-A, JSW, JBR. interpreted the data and

615 PJB and AMC-A prepared the paper with input from all authors.

616

617 **Additional Information**

618 Supplementary information accompanies the paper.

619

620 **Competing Interests**

621 The authors declare no competing interests.

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

Table 1: X-ray data collection and refinement statistics for Smh structures

	SmhB soluble (PDB:6ZZ5)	SmhB soluble Form2 (PDB:6ZZH)	SmhB Pore (PDB:7AOG)	SmhA Semet (PDB:7A26)	SmhA Form2 (PDB:7A27)
Data collection					
Beamline	I04-1	I04-1	I24	I03	I04-1
Wavelength (Å)	0.9159	0.9159	0.9795	0.9792	0.9159
Space group	$P2_12_12_1$	$P2_1$	C2	$P4_2$	$P2_12_12_1$
Cell parameters					
a (Å) =	50.94	95.93	224.48	151.36	81.24
b (Å) =	113.95	49.73	118.06	151.36	92.58
c (Å) =	130.37	99.49	209.53	133.46	221.1
α (°) =	90	90	90	90	90
β (°) =	90	118.43	109.44	90	90
γ (°) =	90	90	90	90	90
Molecules per asymmetric unit	2	2	10	8	4
Resolution (Å)	1.84-65.18 (1.84-1.87)	1.86 – 87.51 (1.86-1.88)	6.98 – 66.77 (6.98 – 7.10)	2.98 – 67.69 (2.98 – 3.03)	2.57-110.55 (2.57-2.61)
Total reflections[a]	445968 (15821)	286930 (8241)	28759 (1072)	1727990 (88850)	703016 (35733)
Unique reflections[a]	66582 (3207)	50372 (2519)	8305 (350)	61518 (3065)	53914 (2644)
R_{merge}[a,b]	0.139 (1.214)	0.1 (0.899)	2.295 (3.384)	0.406 (3.048)	0.190 (2.129)
R_{pim}[a,c]	0.078 (0.842)	0.044 (0.525)	1.403 (2.356)	0.078 (0.575)	0.055 (0.595)
Mean I/**σ**(I)[a]	9.5 (1.0)	10.4 (1.4)	1.3 (0.5)	7.2 (1.3)	8.5 (1.1)
Completeness (%)[a]	99.8 (98.1)	90.2 (54.6)	98.8 (88.2)	100 (99.9)	99.9 (99.8)
Multiplicity[a]	6.7 (4.9)	5.7 (3.3)	3.5 (3.1)	28.1 (29.0)	13.0 (13.5)
Mid-slope				1.126	
dF/F				0.140	
Refinement					
No. of non-H atoms	11426	11266	28485	21017	21019
R_{work}/R_{free}	0.19/0.22	0.21/0.24	0.33/0.34	0.24/0.25	0.23/0.25
Average B factors (Å^2)	27	31	0.0	78	78
Bond length rmsd (Å)	0.0130	0.0107	0.0053	0.0087	0.0073
Bond angle rmsd (°)	1.759	1.65	1.49	1.63	1.54
Ramachandran favoured/allowed (%)	99.01/100	98/100	93.32/100	95.60/100	96/100

[a]Values in brackets are for data in the high-resolution shell

[b]$R_{merg} = \Sigma_{hkl} \Sigma_i | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$

[c]$R_{pim} = \Sigma_{hkl} \sqrt{1/n-1} \Sigma_{i=1} | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively

641 **Figure legends**

642 **Figure 1 Lytic and pore forming activity of Smh toxin.** (A) Percentage lysis of erythrocytes
643 with each protein alone and in combination at a ratio of 1:1:1. (B) TEM negative stain
644 images of erythrocytes incubated with SmhA+SmhB+SmhC (top), showing side and top
645 views of pores in membrane fragments, and SmhB+SmhC (below) showing top views of
646 pores which line the membranes of intact erythrocytes. (C) Percentage lysis after pre-
647 incubation of erythrocytes for 1 hour at 37°C with either SmhA, SmhB or SmhC alone or
648 together, before addition of the remaining components with a further 1 hour incubation.
649 Red stars above bars represent P value, * P ≤0.05; ** P ≤0.01; *** P ≤0.001. (D) Percentage
650 lysis with varying concentrations of SmhA, SmhB or SmhC and fixed concentrations (10 nM)
651 of the remaining 2 components.
652
653 **Figure 2 Structure of soluble and pore form SmhB** (A) Structure of soluble SmhB
654 (PDB:6ZZH), the hydrophobic β-tongue (yellow) is buried from the solvent. (B) Pore
655 Structure of SmhB (PDB:7AOG). 10 monomers of SmhB, in two conformations (type 1 –
656 green, and type 2 – orange), assemble to form a helical pore with a hydrophobic head
657 (yellow) made up of extended hydrophobic helices (top), with the individual type1-type2
658 dimers shown (bottom). (C) The structures of both soluble and pore SmhB are almost
659 identical to those of AhlB (top), haemolytic assay results (below) for SmhABC and chimeric
660 toxin, SmhA + AhlB + SmhC, showing both toxin combinations reach 100% lysis after 1 hr
661 incubation with erythrocytes.
662
663 **Figure 3 Structure of SmhA.** (A) Crystal Structure of SmhA (purple) (PDB:7A27), with
664 residues of the predicted head domain coloured as in helical wheels (C, top right). (B) rmsD

values for structural alignments of soluble SmhA to soluble SmhB, AhlB, NheA, MakA, and Hbl-B. Hydrophobic residues are coloured yellow, and C-terminus is labelled. (C) Predicted pore structure of SmhA showing the two extended head domain helices as seen in AhlB[7] coloured as in the plotted helical wheels (right). Surface representation of pore SmhA model (bottom right) shows how the sequence of the two extended helices create hydrophobic and hydrophilic surfaces on opposite faces of the monomer.

Figure 4 Conserved mechanisms for head domain folding. SmhA (left, purple), SmhB (right, green) and NheA (below, pale blue) adopt a leucine zipper between α4 and α5, formed from family conserved residues (light brown) to hide hydrophobic residues and stabilise the head domain.

Figure 5 Family conserved hinge residues show conserved latch residues for head domain folding. A) Conserved Lys and Thr residues (light brown) interact between α2 and the C-terminal helix in soluble SmhA (top), and SmhB (centre), an equivalent interaction (Q83-E351) is seen in NheA (bottom). (B) Hydrogen bonding between conserved glutamine residues to the main chain of the β-tongue in SmhA (top), SmhB (centre) and NheA (bottom).

Figure 6 Schematic showing pore assembly for the Smh toxin. SmhC (cyan) binds to a single leaflet of the membrane using its hydrophobic head (yellow). Soluble SmhB is recruited (green) and undergoes its conformational change to pore form, with its hydrophobic head traversing the membrane. At this stage either oligomerisation occurs (top, mechanism 1) to form SmhBC pro-pores, followed by recruitment of the SmhA component (purple) to form a

689 lytic pore. Alternatively (bottom, mechanism 2), SmhC, SmhB and SmhA assemble
690 sequentially on the membrane prior to oligomerisation to form the lytic pore. Schematic top
691 views of the pores are shown, with each oval representing two helices with hydrophobic
692 (yellow) and hydrophilic (blue) surfaces indicated. Soluble inhibitory complexes can form
693 between SmhB+SmhC and SmhA+SmhB+SmhC.
694
695 **Figure 7 Model of tripartite α-PFT pore assembly.** A) The C component (cyan) inserts its
696 hydrophobic head (yellow) through a single leaflet of the membrane. B) Soluble B (green)
697 binds to C with the N-terminal residues of α3 (G97- R125) in B interacting with the C-
698 terminal residues of α3 on C (regions of both coloured pink). C) The B component then
699 undergoes a large conformational change, inserting two hydrophobic helices through the
700 membrane. D) Finally, the A component (purple) binds to the B component with the region
701 Q58-T72 on A (blue) interacting with E294-L313 on B (blue). The head region of A (circled in
702 dashed lines) is predicted to undergo a conformational change to extended helices, which
703 insert through the membrane. All complexes modelled using the HADDOCK2.2 web server
704 [30].
705
706 **Figure 8 Model of the complete tripartite α-PFT ABC pore**. (A) Based on docking predictions
707 the A component (purple) fits between neighbouring B components (green) on the inside
708 surface of the pore, surrounded by C components (cyan), highlighting the membrane (blue
709 bar) and hydrophobic surfaces (yellow). (B) top view of pore. (C) Surface rendered full and
710 cutaway side views of the pore model generated in Pymol [47], showing the hydrophilic lining
711 to the pore with residues coloured hydrophobic (white), polar (pale blue), positive (blue)
712 and negative (red).

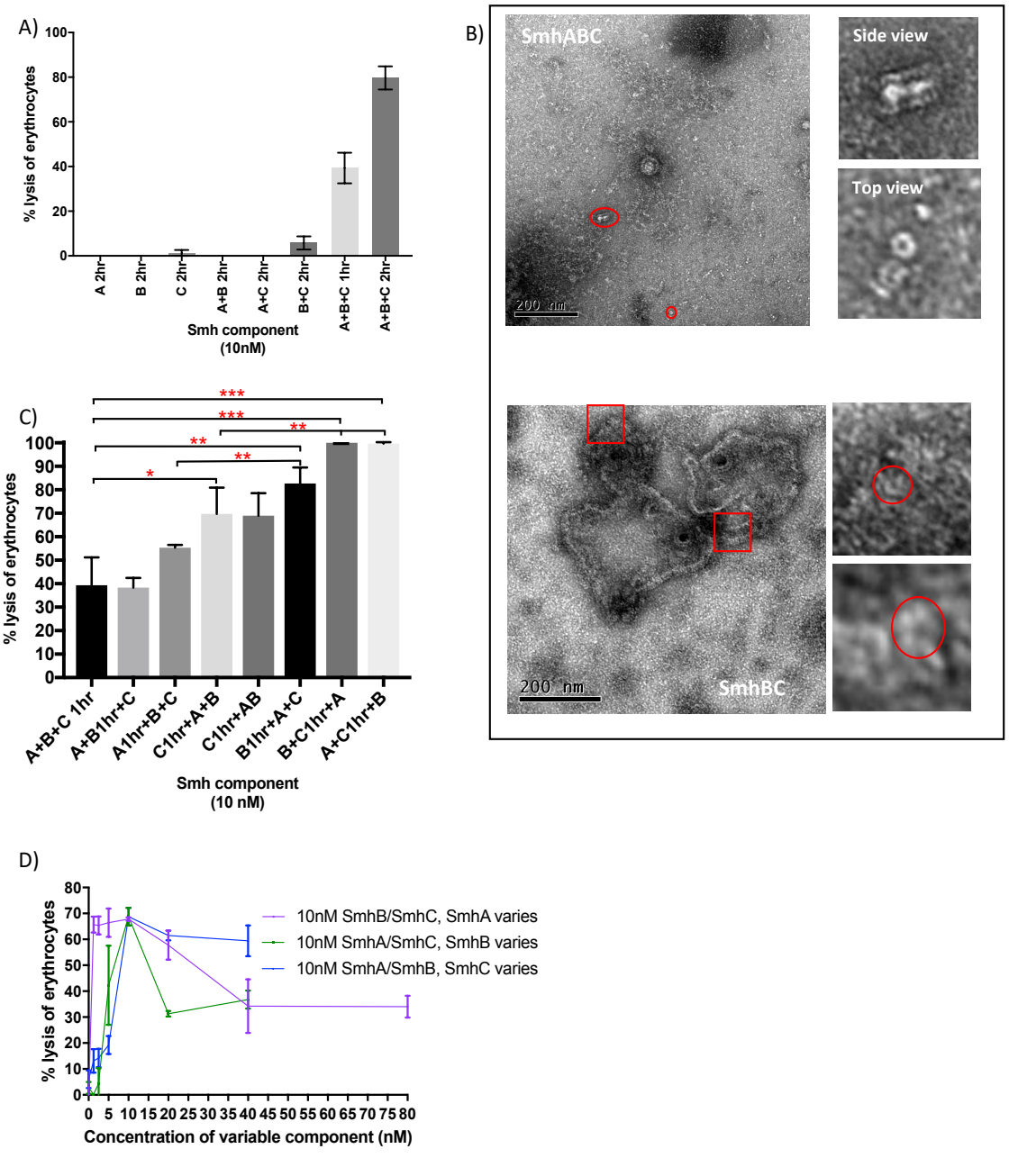

A)
% lysis of erythrocytes
A 2hr
B 2hr
C 2hr
A+B 2hr
A+C 2hr
B+C 2hr
A+B+C 1hr
A+B+C 2hr
Smh component (10nM)
B)
SmhABC
Side view
Top view
200 nm
SmhBC
C)
% lysis of erythrocytes
A+B+C 1hr
A+B1hr+C
A1hr+B+C
C1hr+A+B
C1hr+AB
B1hr+A+C
B+C1hr+A
A+C1hr+B
Smh component (10 nM)
D)
% lysis of erythrocytes
10nM SmhB/SmhC, SmhA varies
10nM SmhA/SmhC, SmhB varies
10nM SmhA/SmhB, SmhC varies
Concentration of variable component (nM)

Figure 1

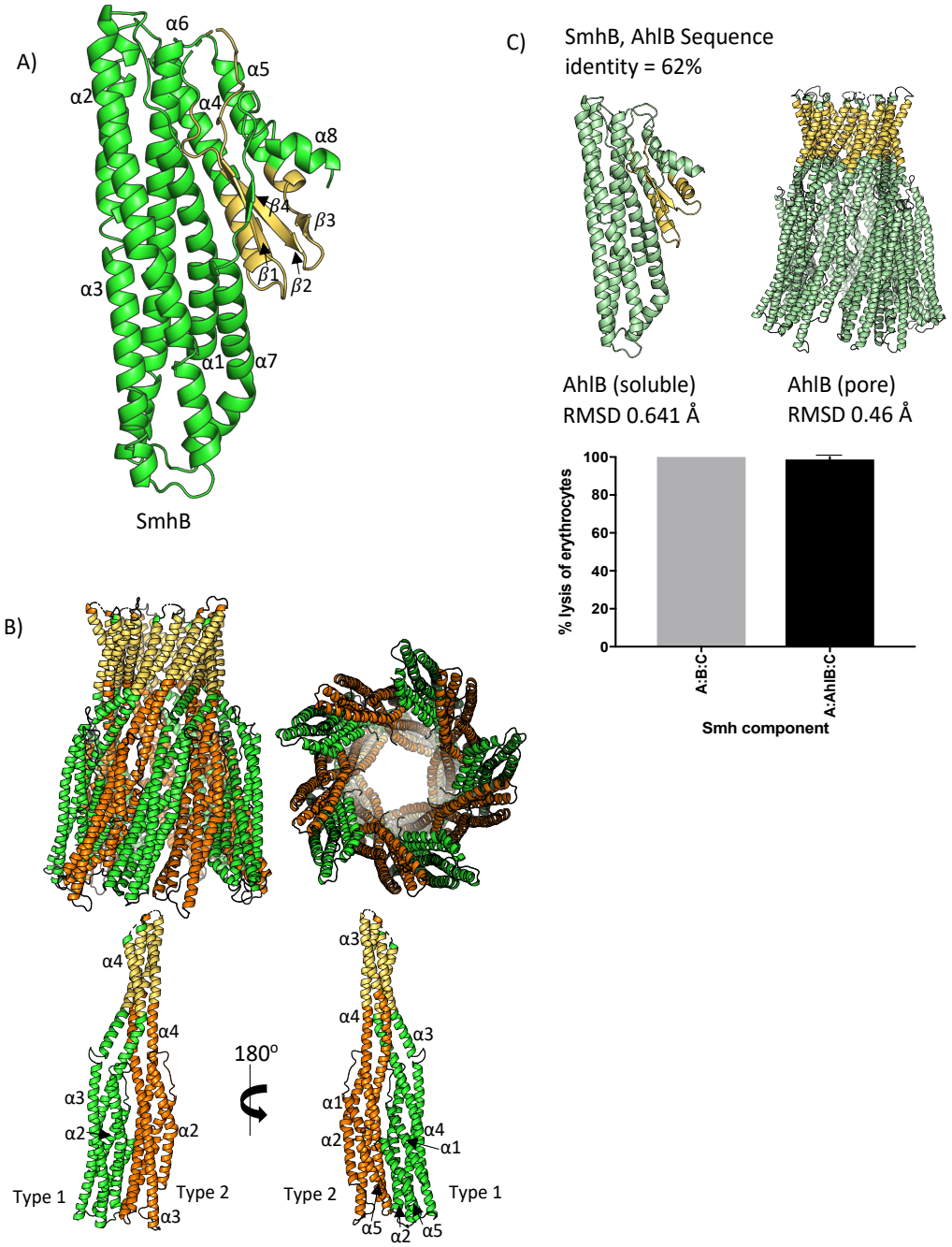

A)
α6
α5
α2
α4
α8
β4
β3
β1
β2
α3
α1
α7
SmhB
B)
α4
α4
α3
α2
α2
Type 1
Type 2
α3
180°
α3
α4
α3
α1
α2
α4
α1
Type 2
Type 1
α5
α2
α5
C)
SmhB, AhlB Sequence identity = 62%
AhlB (soluble)
RMSD 0.641 Å
AhlB (pore)
RMSD 0.46 Å
% lysis of erythrocytes
100
80
60
40
20
0
A:B:C
A:AhlB:C
Smh component

Figure 2

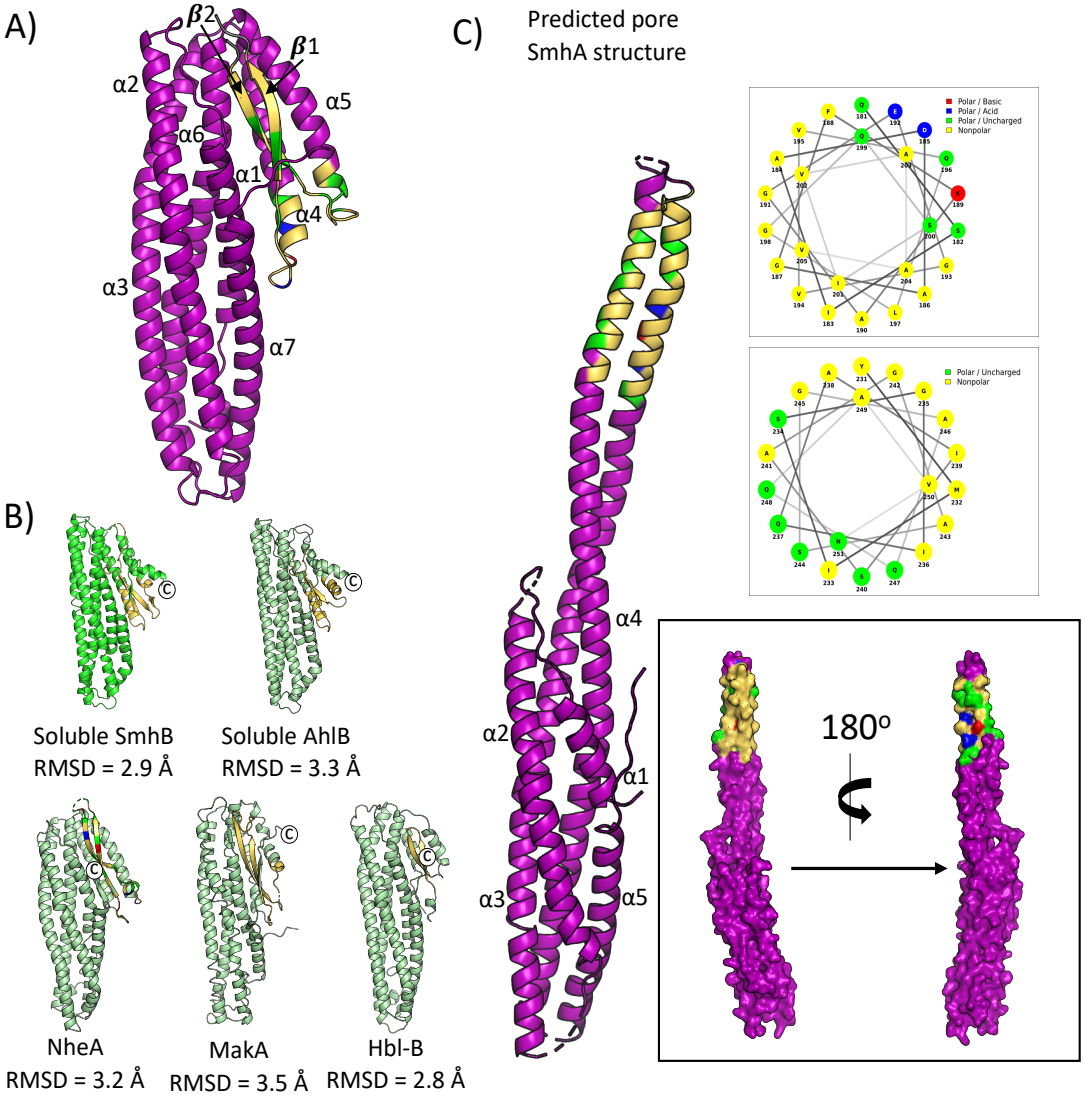

A)
β2
β1
α2
α5
α6
α1
α4
α3
α7
B)
Soluble SmhB
RMSD = 2.9 Å
Soluble AhlB
RMSD = 3.3 Å
NheA
RMSD = 3.2 Å
MakA
RMSD = 3.5 Å
Hbl-B
RMSD = 2.8 Å
C)
Predicted pore
SmhA structure
Polar / Basic
Polar / Acid
Polar / Uncharged
Nonpolar
Polar / Uncharged
Nonpolar
α4
α2
α1
α3
α5
180°

Figure 3

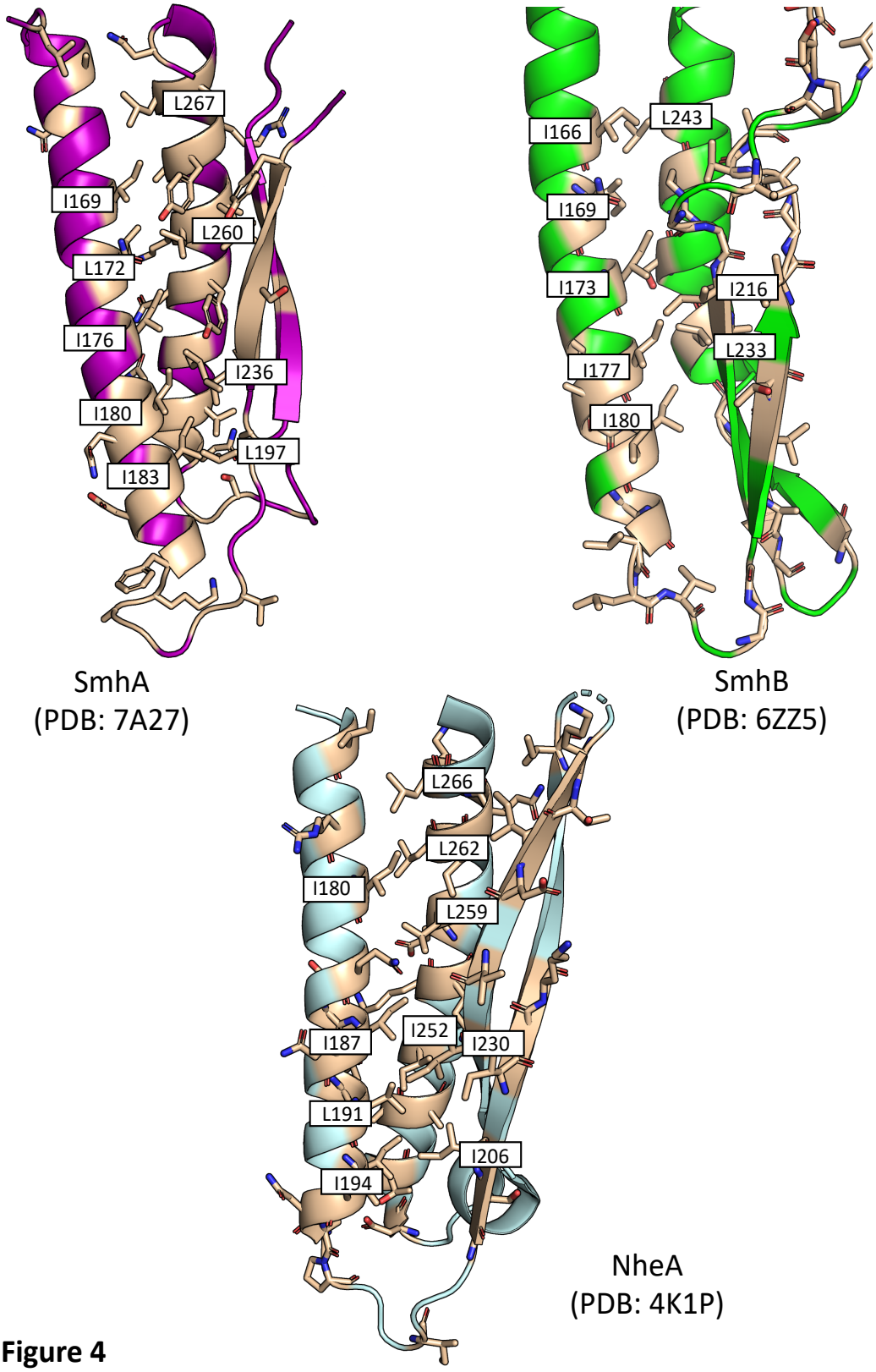

L267
I169
L260
L172
I176
I236
I180
L197
I183
SmhA
(PDB: 7A27)
I166
L243
I169
I173
I216
L233
I177
I180
SmhB
(PDB: 6ZZ5)
L266
L262
I180
L259
I252
I187
I230
L191
I206
I194
NheA
(PDB: 4K1P)

Figure 4

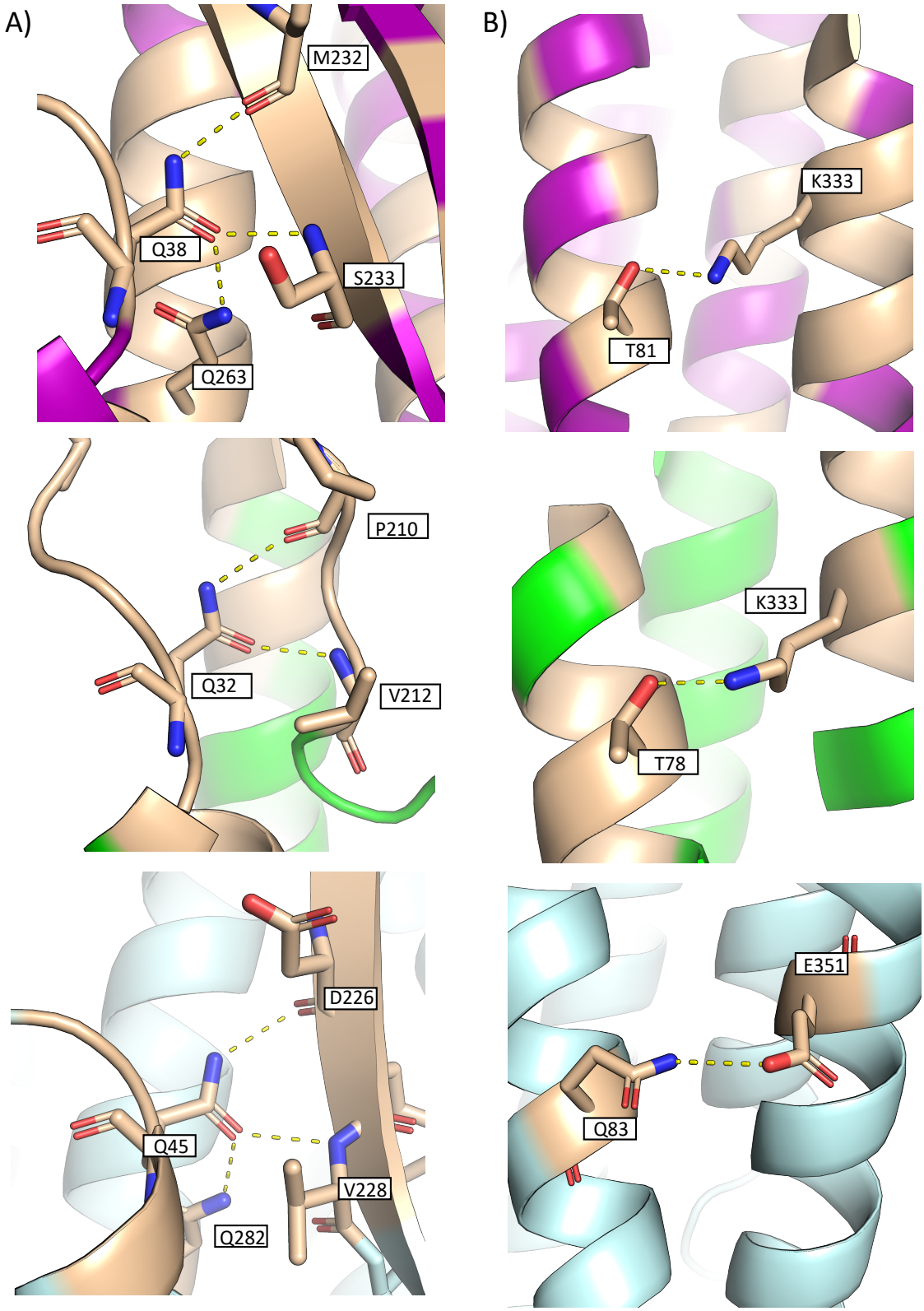

A)
M232
Q38
S233
Q263
B)
K333
T81
P210
Q32
V212
K333
T78
D226
Q45
V228
Q282
E351
Q83

Figure 5

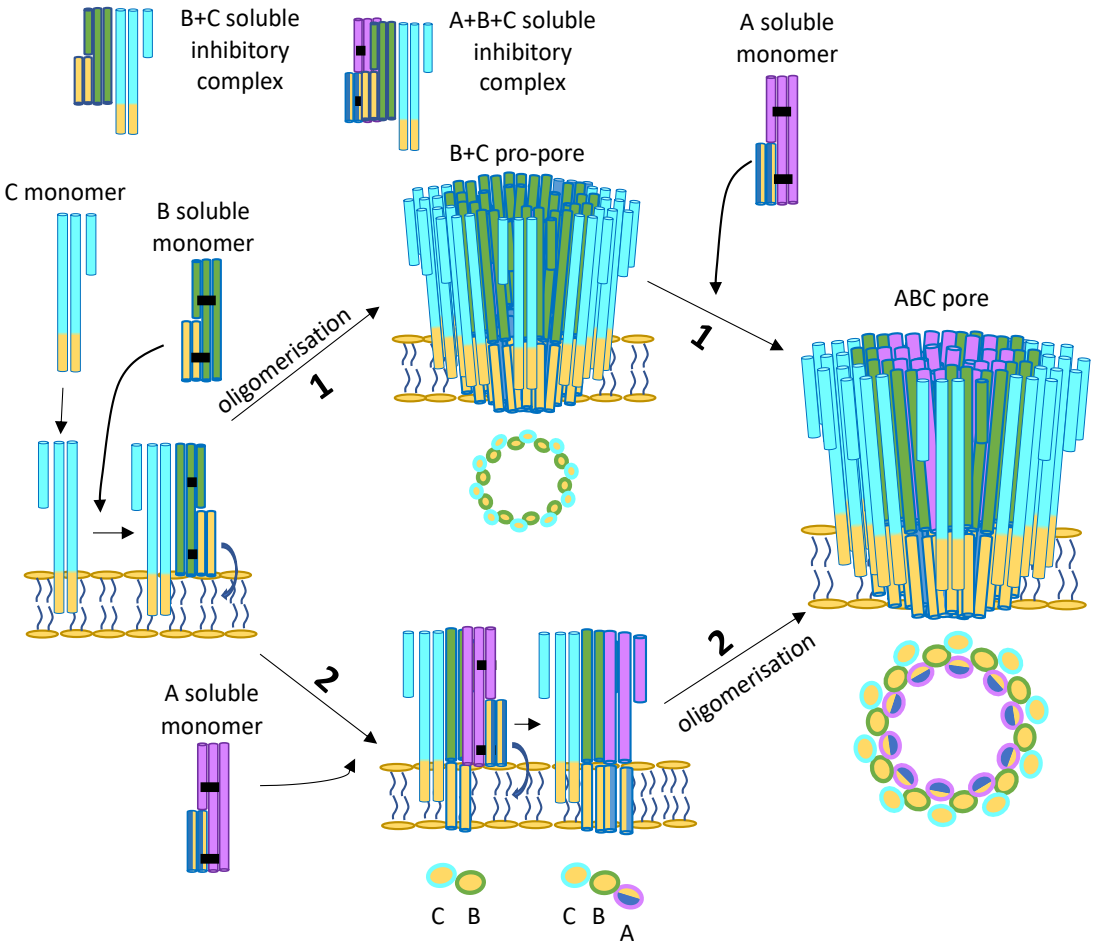

B+C soluble inhibitory complex
A+B+C soluble inhibitory complex
A soluble monomer
B+C pro-pore
C monomer
B soluble monomer
oligomerisation
1
ABC pore
1
2
A soluble monomer
2
oligomerisation
C B
C B A

Figure 6

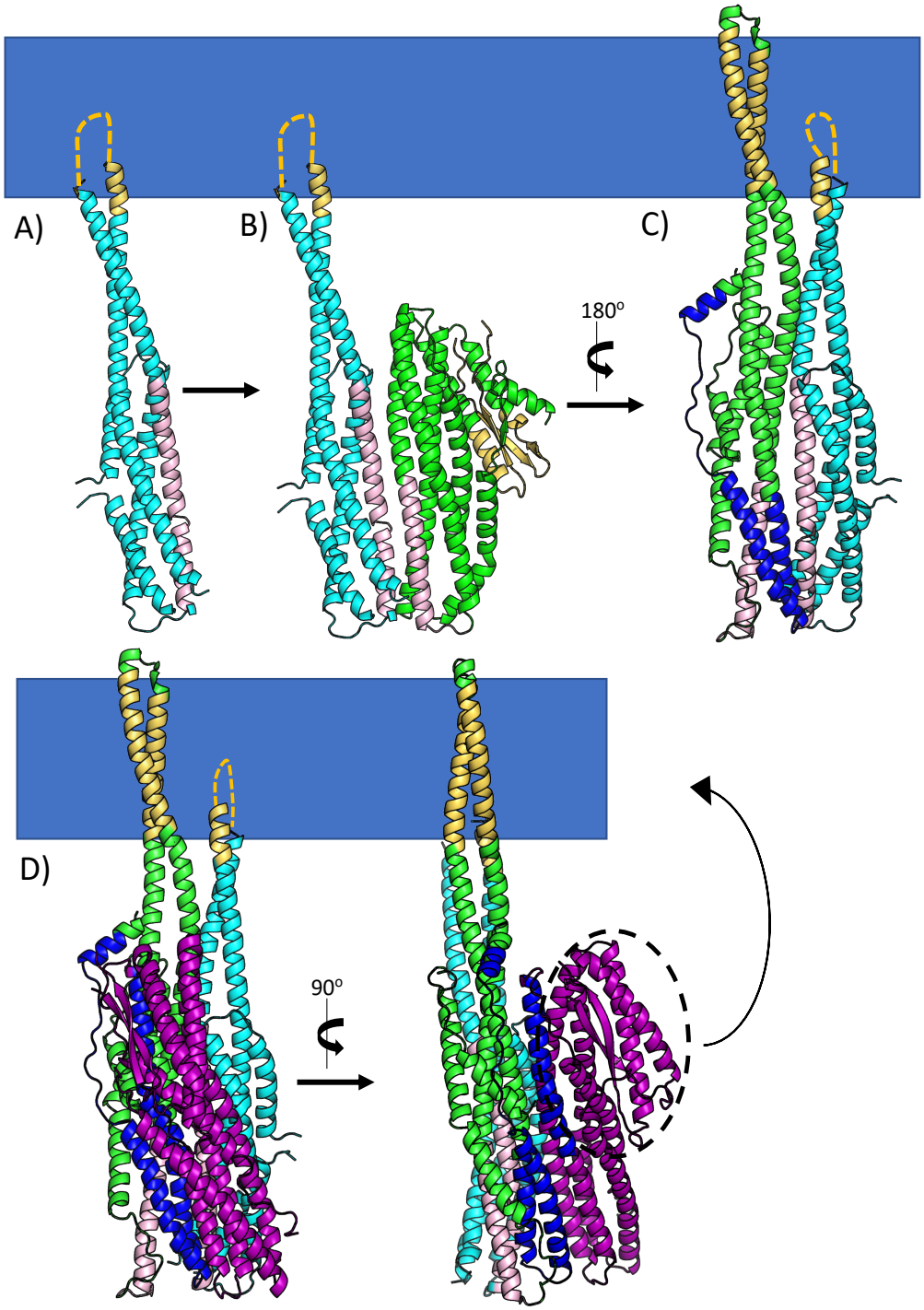


Figure 7

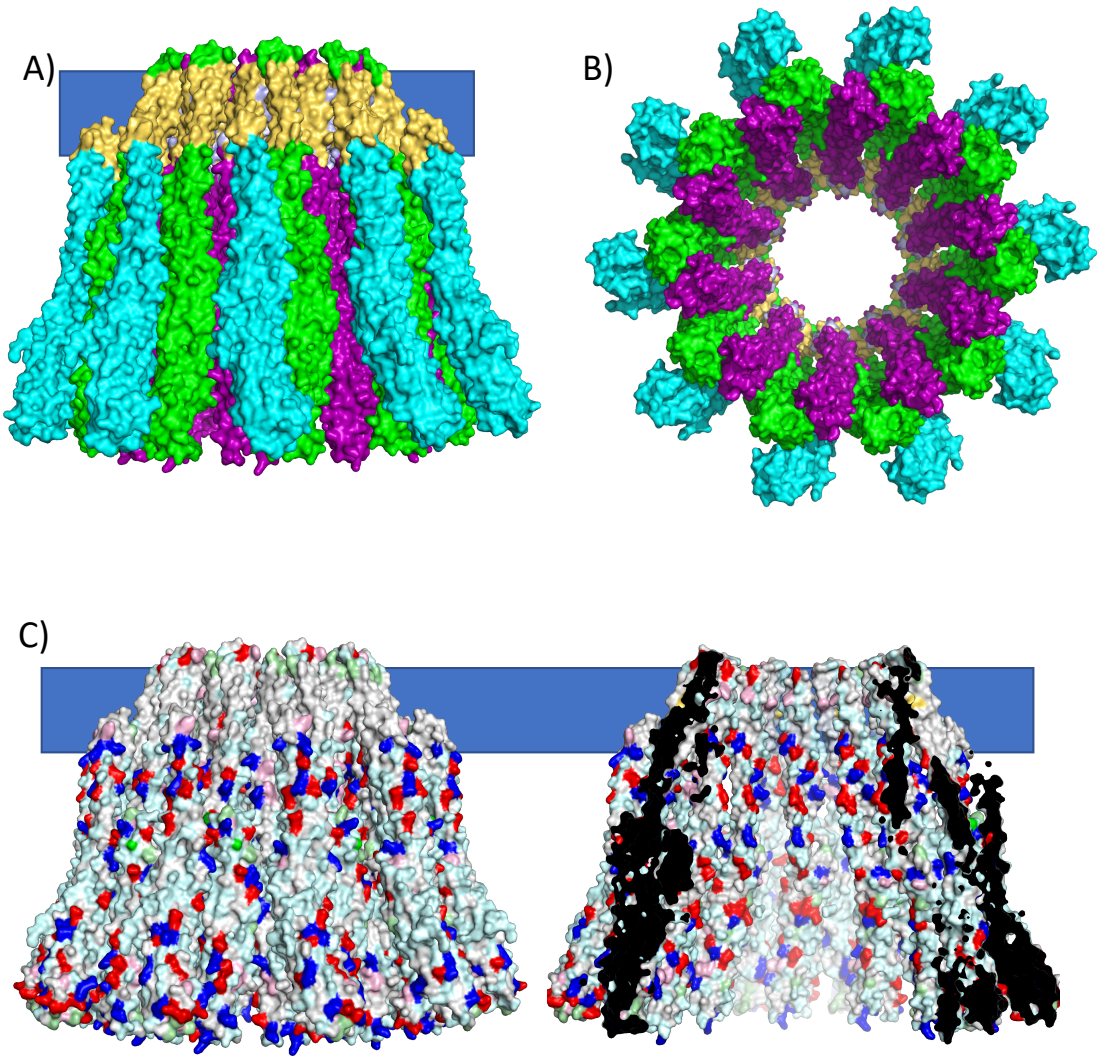
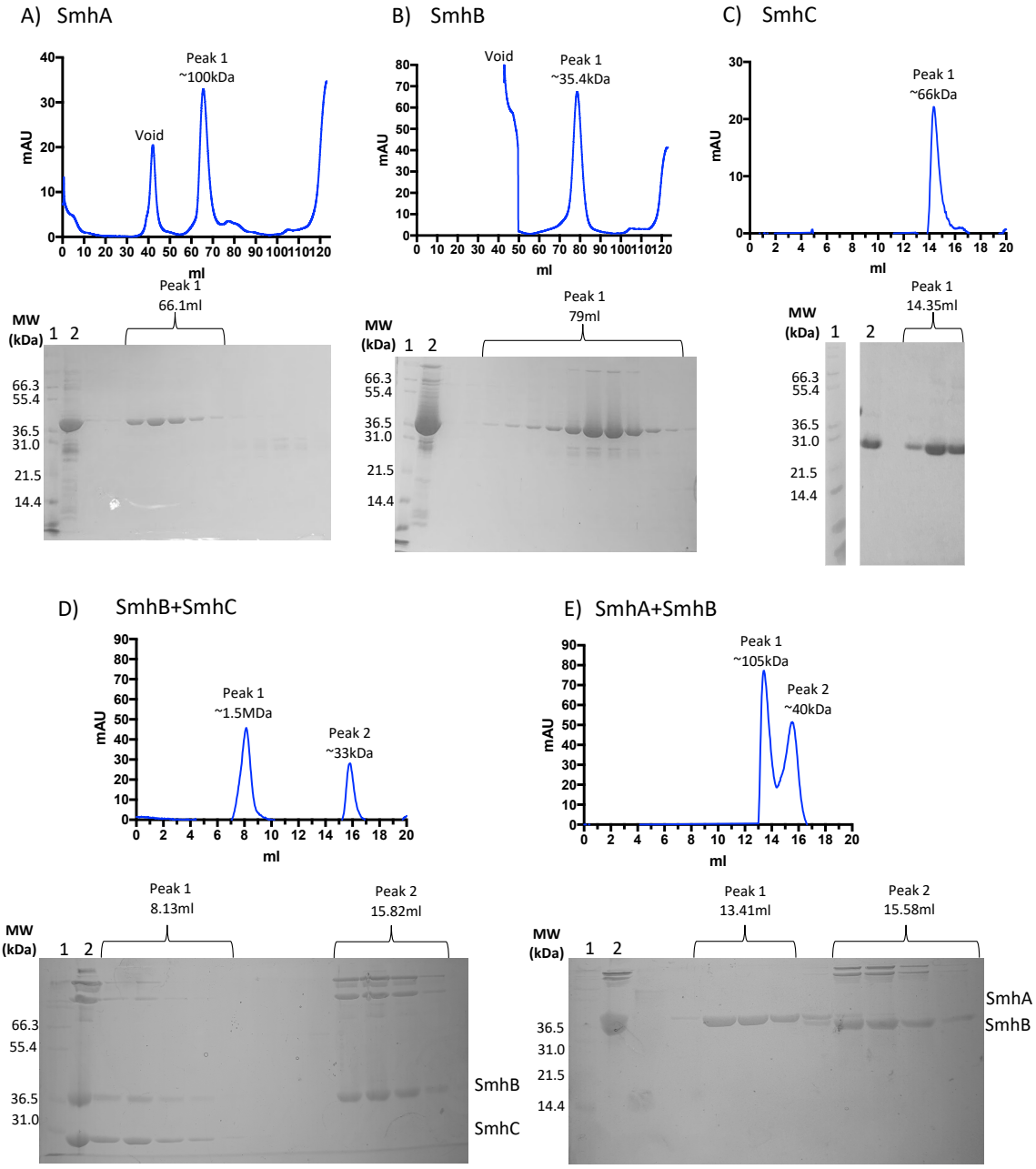


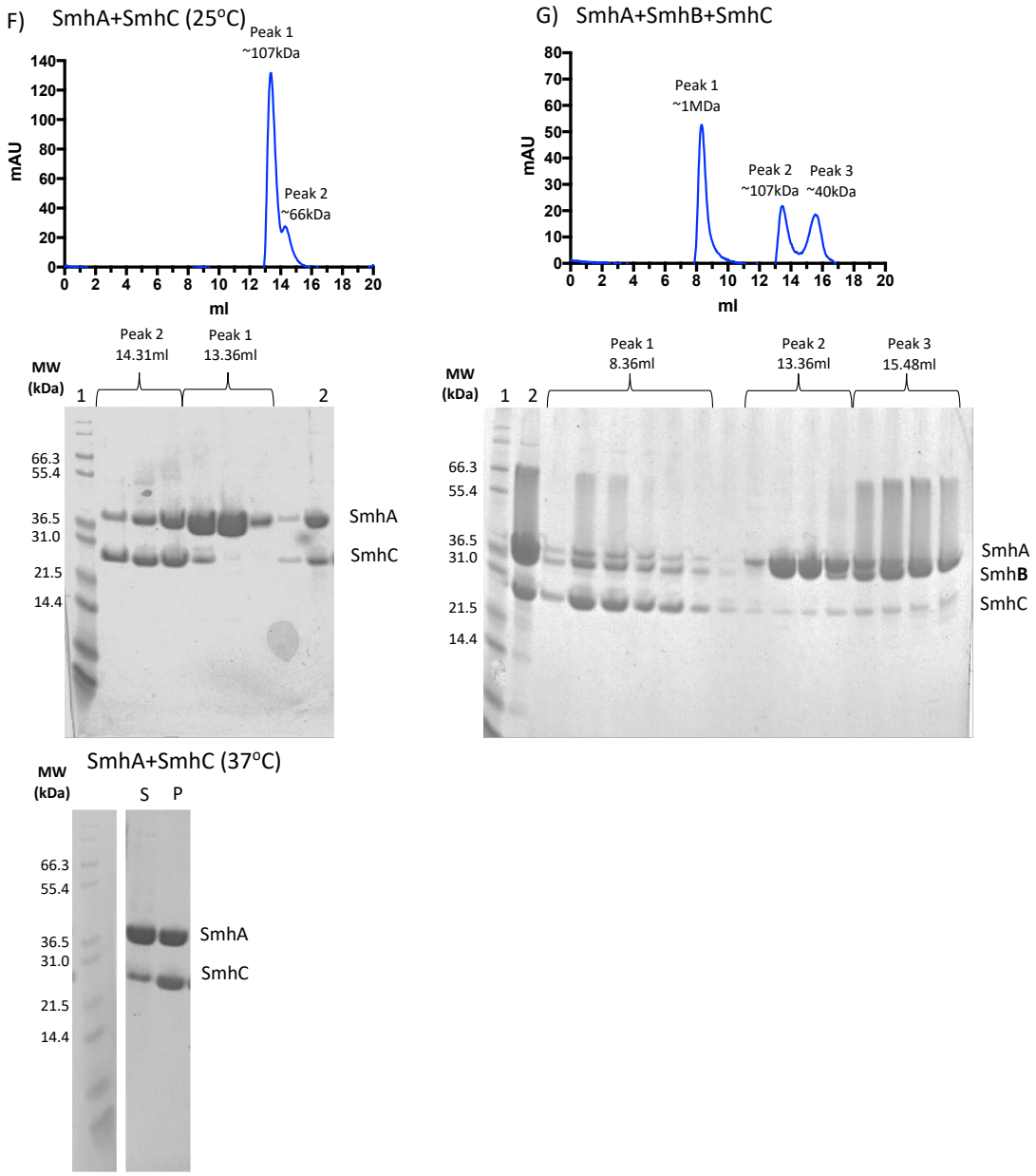
Figure 8

Identification and characterisation of a tripartite α-pore forming toxin from *Serratia marcescens*

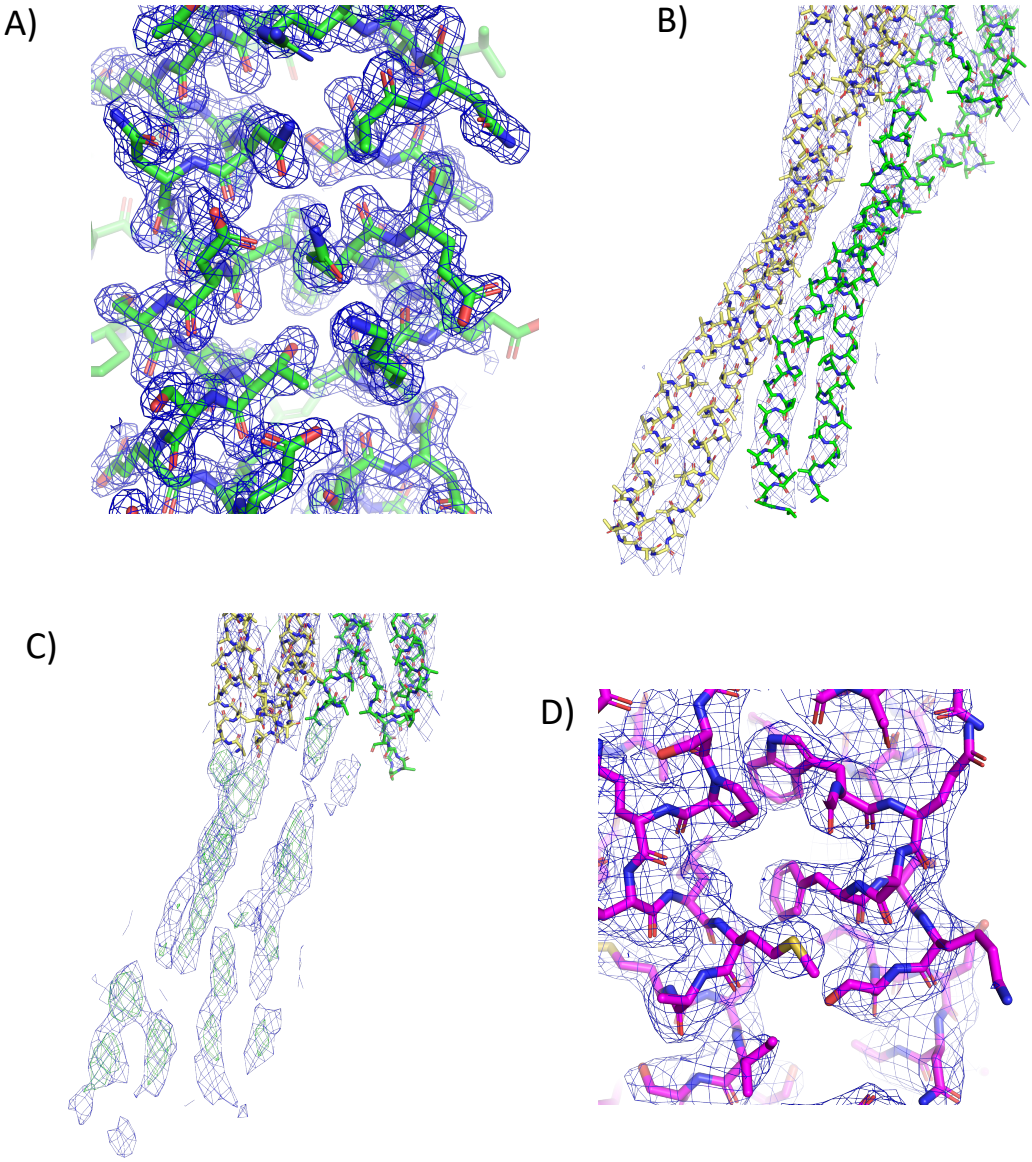
A.Churchill-Angus, et al.

Supplementary Information

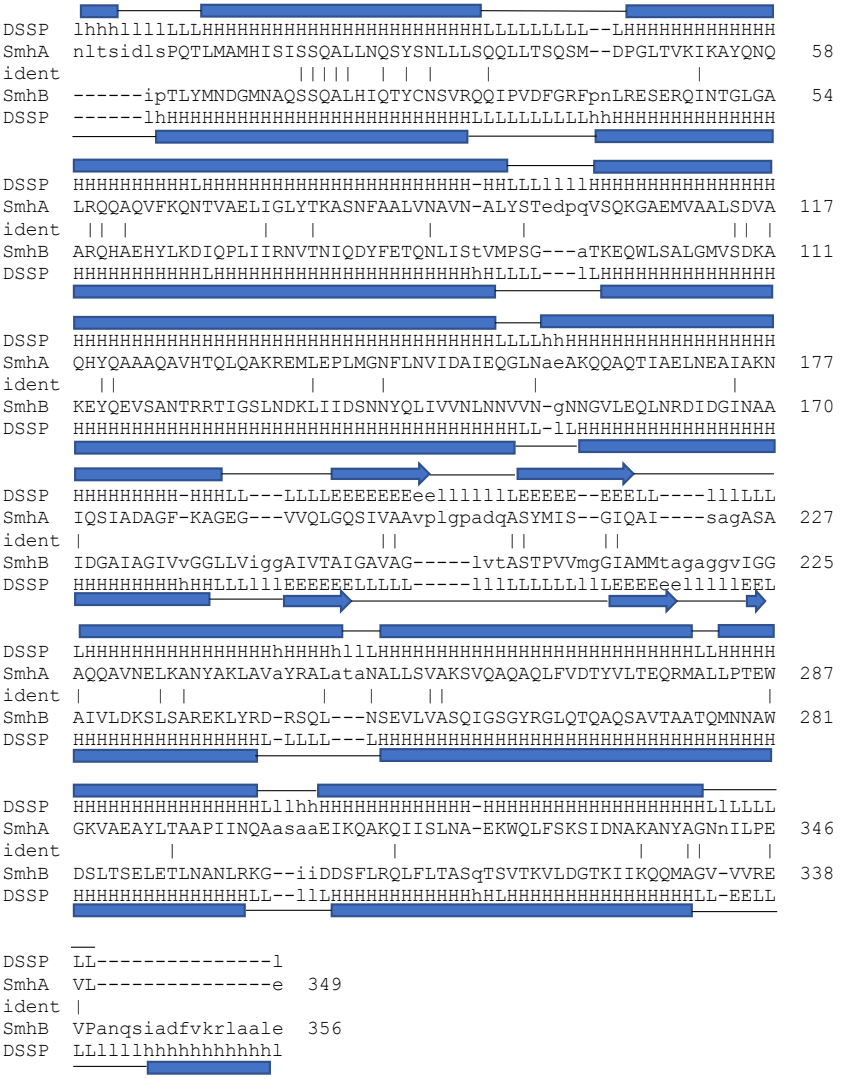

A) SmhA
Peak 1
~100kDa
Void
mAU
ml
Peak 1
66.1ml
MW
(kDa)
1 2
66.3
55.4
36.5
31.0
21.5
14.4
B) SmhB
Void
Peak 1
~35.4kDa
mAU
ml
Peak 1
79ml
MW
(kDa)
1 2
66.3
55.4
36.5
31.0
21.5
14.4
C) SmhC
Peak 1
~66kDa
mAU
ml
Peak 1
14.35ml
MW
(kDa)
1 2
66.3
55.4
36.5
31.0
21.5
14.4
D) SmhB+SmhC
Peak 1
~1.5MDa
Peak 2
~33kDa
mAU
ml
Peak 1
8.13ml
Peak 2
15.82ml
MW
(kDa)
1 2
66.3
55.4
36.5
31.0
SmhB
SmhC
E) SmhA+SmhB
Peak 1
~105kDa
Peak 2
~40kDa
mAU
ml
Peak 1
13.41ml
Peak 2
15.58ml
MW
(kDa)
1 2
36.5
31.0
21.5
14.4
SmhA
SmhB



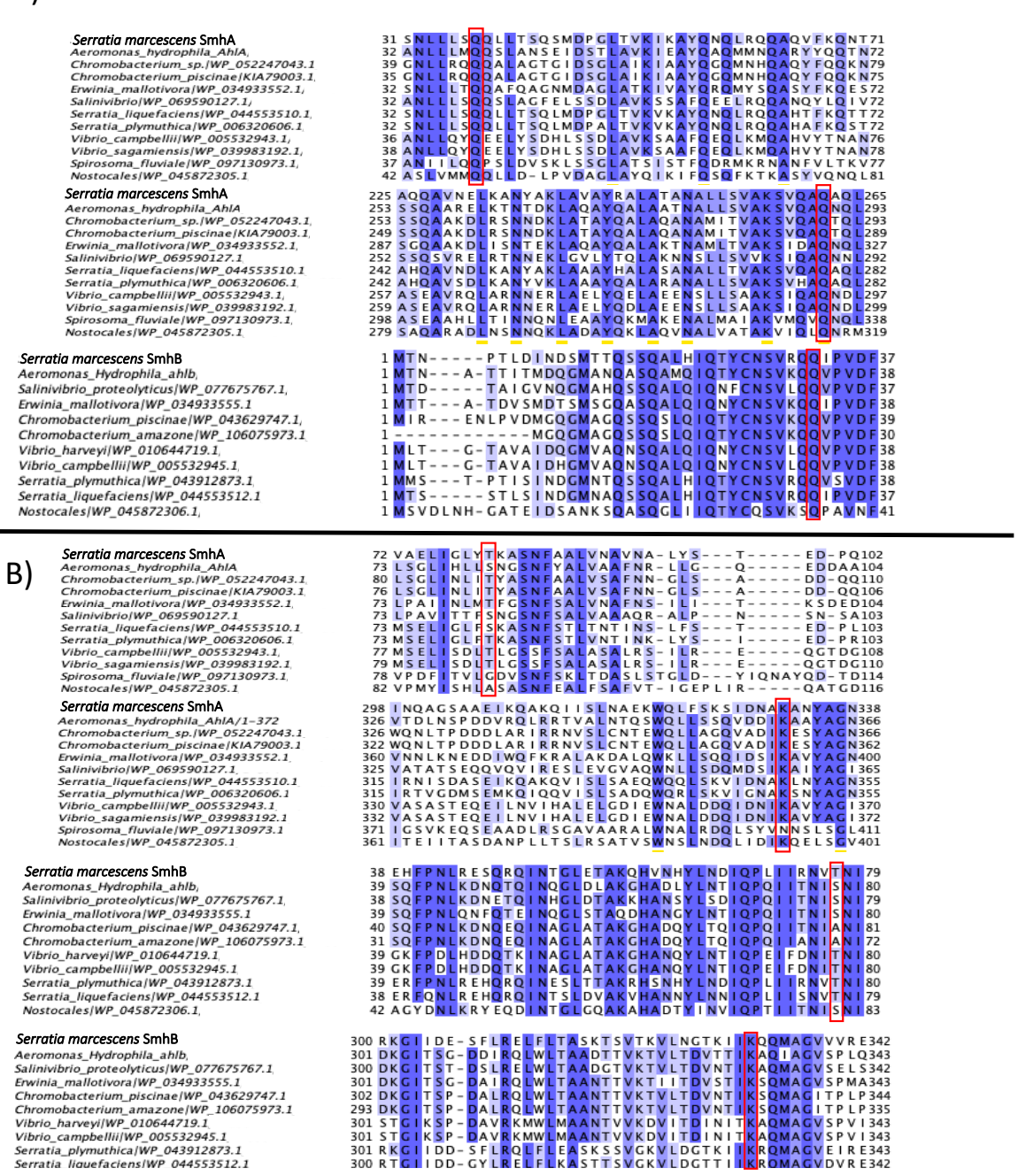
Supplementary Figure 1. Analysis of SmhA, SmhB and SmhC by size exclusion chromatography. Gel filtration chromatograms with SDS-PAGE gels below showing the content of the labelled peaks, Lane 1 contains the MW ladder, Lane 2 contains the gel filtration load. A) SmhA purification. Superdex 200pg B) SmhB purification. C) SmhC purification. Lane 1, 2 and peak1 are cropped images from the same gel. D) A 1:1 ratio of SmhB + SmhC elutes as 2 peaks, one high molecular weight peak containing both SmhB and SmhC and a second peak low mw peak containing only SmhB. E) A 1:1 ratio of SmhA + SmhB elutes as 2 separate peaks one containing SmhA and the other SmhB. F) A 1:1 ratio of SmhA + SmhC , peak 1 containing SmhA and peak 2 containing SmhC. SDS-PAGE of resulting supernatant (S) and pellet (P) after incubation of SmhA and SmhB at 37°C for 1 hr (below), showing that most of SmhC precipitates with SmhA. Ladder, lane S and P are cropped images from the same gel. G) A 1:1:1 ratio SmhA + SmhB +SmhC which elutes as 3 separate peaks, peak 1 – high molecular weight species containing SmhA, SmhB and SmhC. Peak 2 and 3 contain the soluble SmhA, SmhB and SmhC components. A) and B) purified using Superdex 200pg, C)-G) purified using Superdex 200 increase.



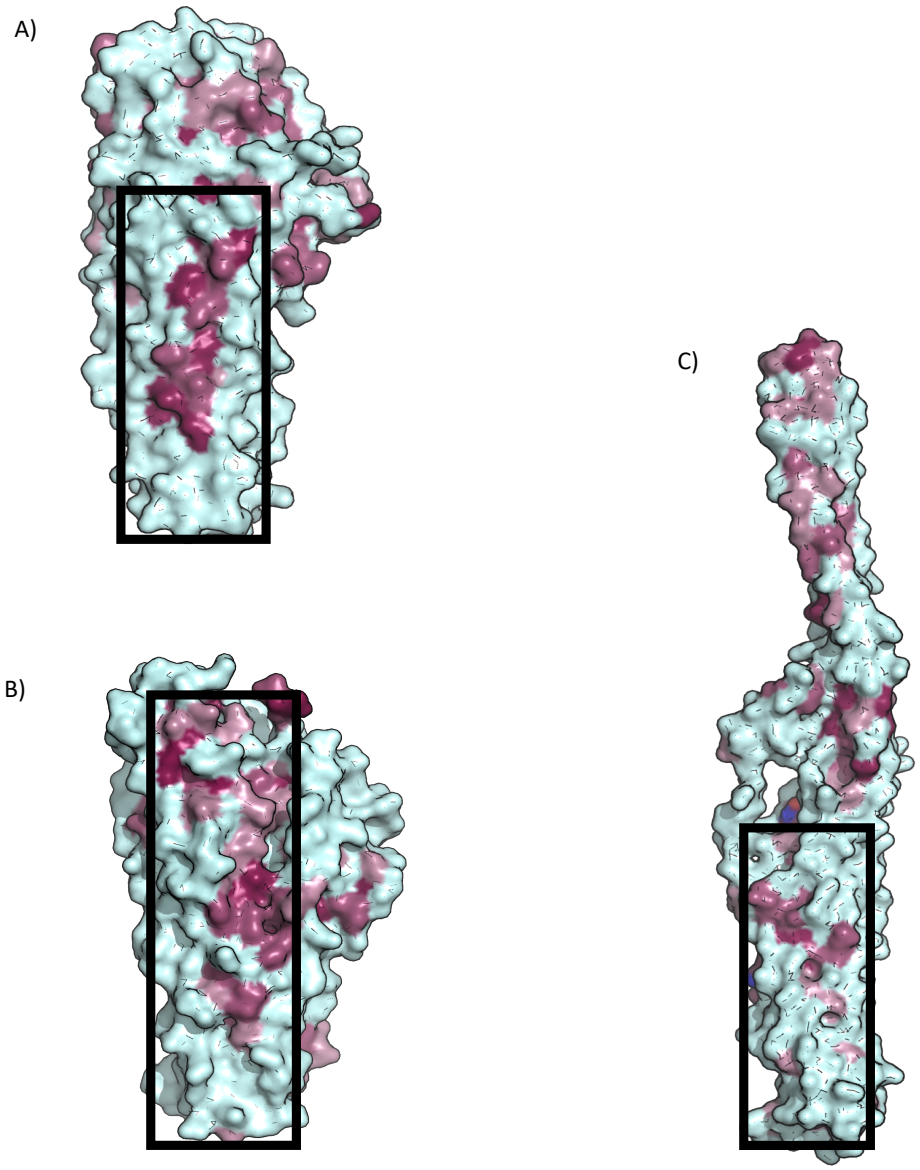
Supplementary Figure 2. Electron density maps. A) 1.8 Å resolution 2Fo-Fc map contoured at 1.0σ for soluble SmhB residues N122-D111 and E293-N282 B) 6.98 Å resolution 2Fo-Fc map contoured at 1.0σ (blue mesh) for SmhB pore conformation chain A (yellow) and B (green) . C) 2Fo-Fc map contoured at 1.0σ (blue mesh) and positive difference map contoured at 2.99σ (green mesh) from an omit map after refinement deleting the head region of chains A and B from the SmhB pore. Positive density returns for both head domains. D) 2.6 Å resolution 2Fo-Fc map contoured at 1.0σ for SmhA residues W342-N351 and F345 and S10-S20.



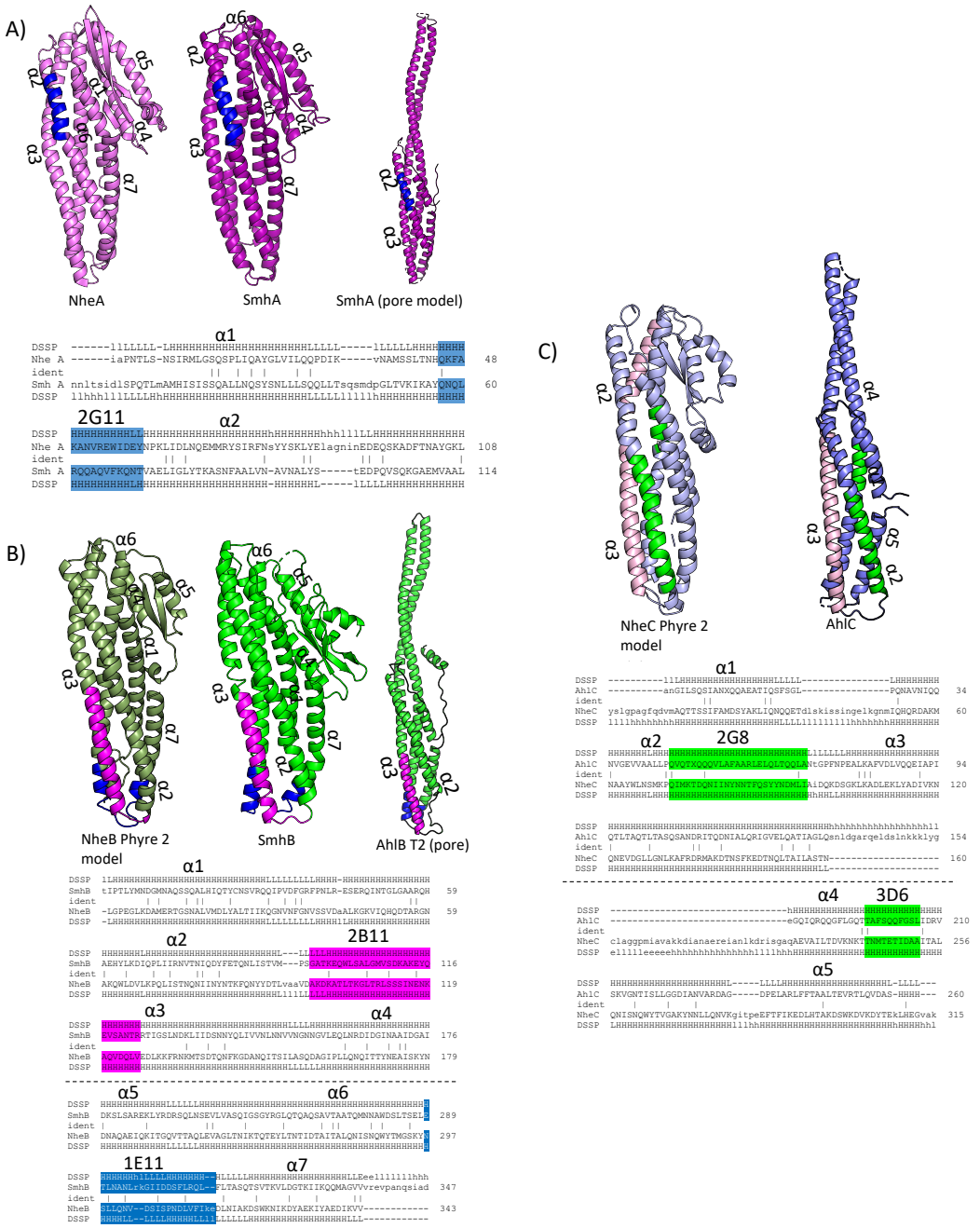
Supplementary Figure 3. Structure based sequence alignment of SmhA and SmhlB, helices shown as rectangles, strands as arrows (H-helix, L- loop, E-strand), prepared using Dali [1] .



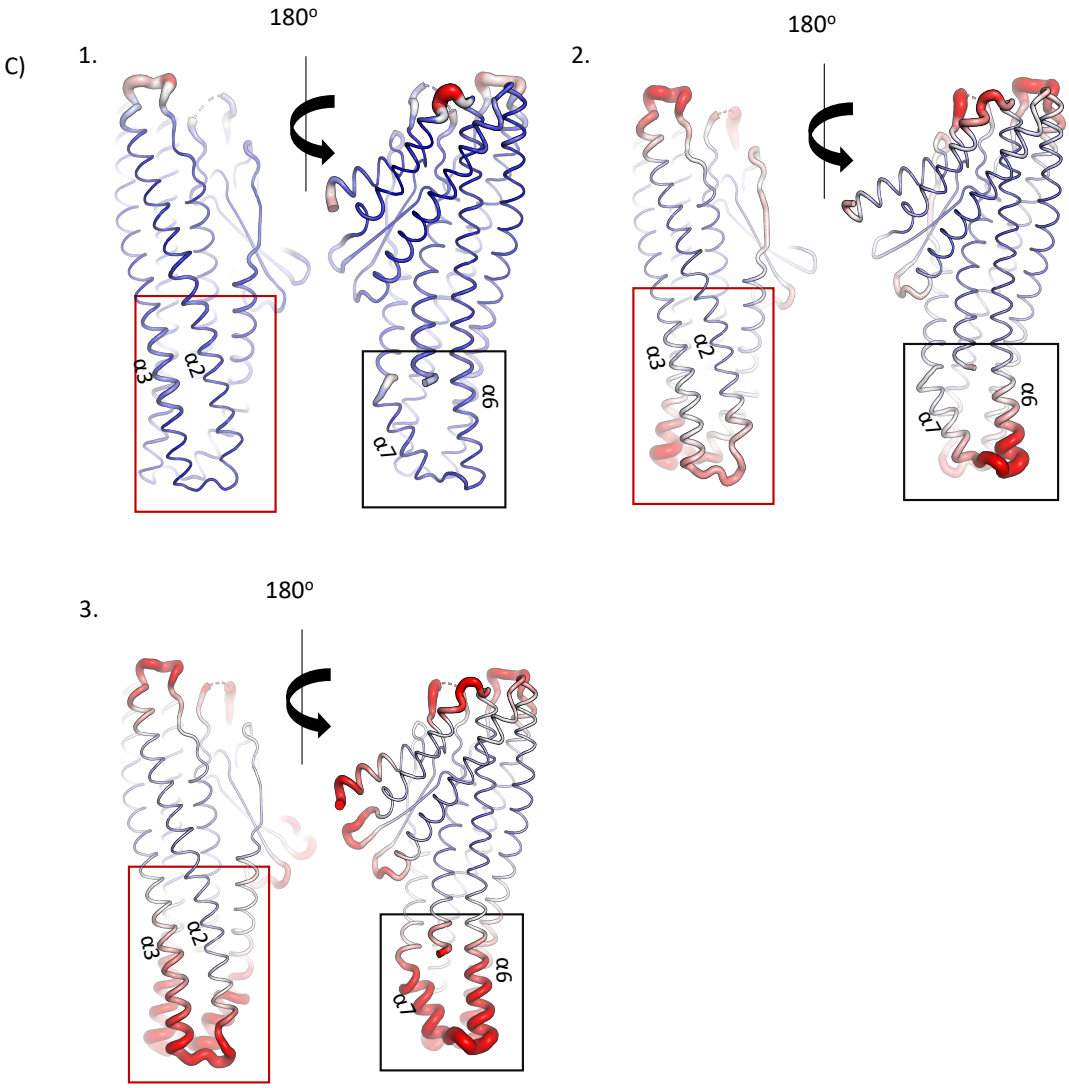
Supplementary Figure 4. Sequence alignments of SmhA and SmhB with A and B components from other species, highlighting (red boxes) residues involved in hydrogen bonding networks that are broken on the soluble to pore transition and with family conserved residues in blue. A) Q38 and Q263 in SmhA and Q32 in SmhB. B) T81 and K333 Lys in SmhA and T78 and K333 SmhB.



Supplementary Figure 5. Surface representation of (A) SmhA, (B) SmhB, and (C) AhlBT2. Family conserved residues are shown in purple. Black boxes outline predicted binding surfaces between the Smh components for (A) SmhA to SmhB, (B) SmhB to SmhC, (C) SmhB to SmhA.



Supplementary Figure 6 Mapping of Nhe antibody binding sites onto Smh and Ahl proteins (A) Structures of NheA (left) showing the 2G11 (blue) Mab binding site (identified by Didier et al [2] and SmhA (right) with NheA 2G11 Mab site mapped onto structure based on Dali [1] structural alignment (below). (B) Phyre 2 [3] model of NheB (left) showing 2B11 (pink) and 1E11 (blue) Mab binding sites and structures of SmhB soluble and pore forms, with the NheB Mab sites mapped onto these structures based on Dali structural alignments (below). (C) Phyre 2 model of NheC (left), showing 2G8 and 3D6 Mab binding sites (green), which do not disrupt complex formation with predicted binding site for the B component identified by Heilkenbrinker [4] (pink) and the structure of AhlC (Right) with NheC Mab sites mapped onto the structure based on Dali structural alignment (below).



Supplementary Figure 7. Flexible regions of SmhB. Variations in B-factor (blue – low B-factor, to red - High B-factor) between monomers of different crystal structures of the soluble form of SmhB**.** An increasing B-factor can be seen at residue 85-124 (red box) and 280-320 (black box) from panels A-C. (A) Crystal form 1, chain A (PDB code 6ZZ5). Crystal form 2 (PDB code 6ZZH) chain A (B) and chain B (C) .

Appendix B

Extended methodology - Smh toxin

This appendix is an overview of the details of work done that was omitted or was only briefly mentioned, from the results and discussion in Paper 2 (Section 4.3) and Paper 3 (Section 5.3). This includes details on the expression and purification SmhA, SmhB, and SmhC, crystallisation and subsequent structure determination of SmhB pore form and soluble SmhB, as well as soluble SmhA. It also outlines details of time course haemolytic assays and EM studies of the Smh toxin.

B.1 SmhA, SmhB, and SmhC construct design

Constructs for SmhA, SmhB, SmhC, were synthesised by Genescript Biotech Corporation so as to incorporate a His_6 tag, as described in Section 2.1.6. All constructs were full length and generated from genes OKB64935.1 (*smhA*), OKB64936.1 (*smhB*) and OKB64937.1 (*smhC*) from *S. marcescens* strain MSU97.

B.2 Overexpression and purification of SmhA, SmhB and SmhC constructs

As all three constructs were previously unstudied, expression and purification protocols had not been established, initial expression trials were required. All constructs were transformed into *E. coli* BL21 (DE3) cells for overexpression as described in Section 2.1.7.

B.2.1 Overexpression trials for SmhA, SmhB and SmhC

To test if the Smh constructs would produce sufficient quantities of soluble protein when expressed in *E. coli* BL21 (DE3) cells, small scale expression trials in 50 ml LB broth were carried out as described in Section 2.2.1, before continuing to large scale overexpressions. Overexpression trials were carried out at 16 ^{o}C, 25 ^{o}C and 37 ^{o}C. For all three proteins, the greatest soluble expression was achieved at 16 ^{o}C, SDS-PAGE gels for these trials can be seen in Figure B.1 B1. As a result, large scale overexpressions for SmhA, SmhB, and SmhC were carried out at 16 ^{o}C overnight as described in Section 2.2.2. Selenomethionine expressions were carried out using the same conditions as sulphur methionine expression, but using the growth protocol detailed in Section 2.2.3, and methods for both are described in Section 2.2.

B.2.2 Purification of SmhA

2 g of cell paste was resuspended in 25 ml lysis buffer (50 mM Tris-HCL pH 8) before lysis by sonication (3×20 s bursts at 16,000 nm λ). Insoluble material was removed by centrifugation at 40,000 g and soluble material was applied to a 5ml Nickel Hi-trap column (GE Healthcare) in binding buffer (50 mM Tris-HCl pH 8 and 0.5 M NaCl). Protein was eluted on a gradient of binding buffer to 50 mM Tris-HCl pH 8, 0.5 M NaCl and 0.5 M imidazole. SmhA eluted at 0.125 M Imidazole (Figure B.2A). Fractions containing SmhA were concentrated down to

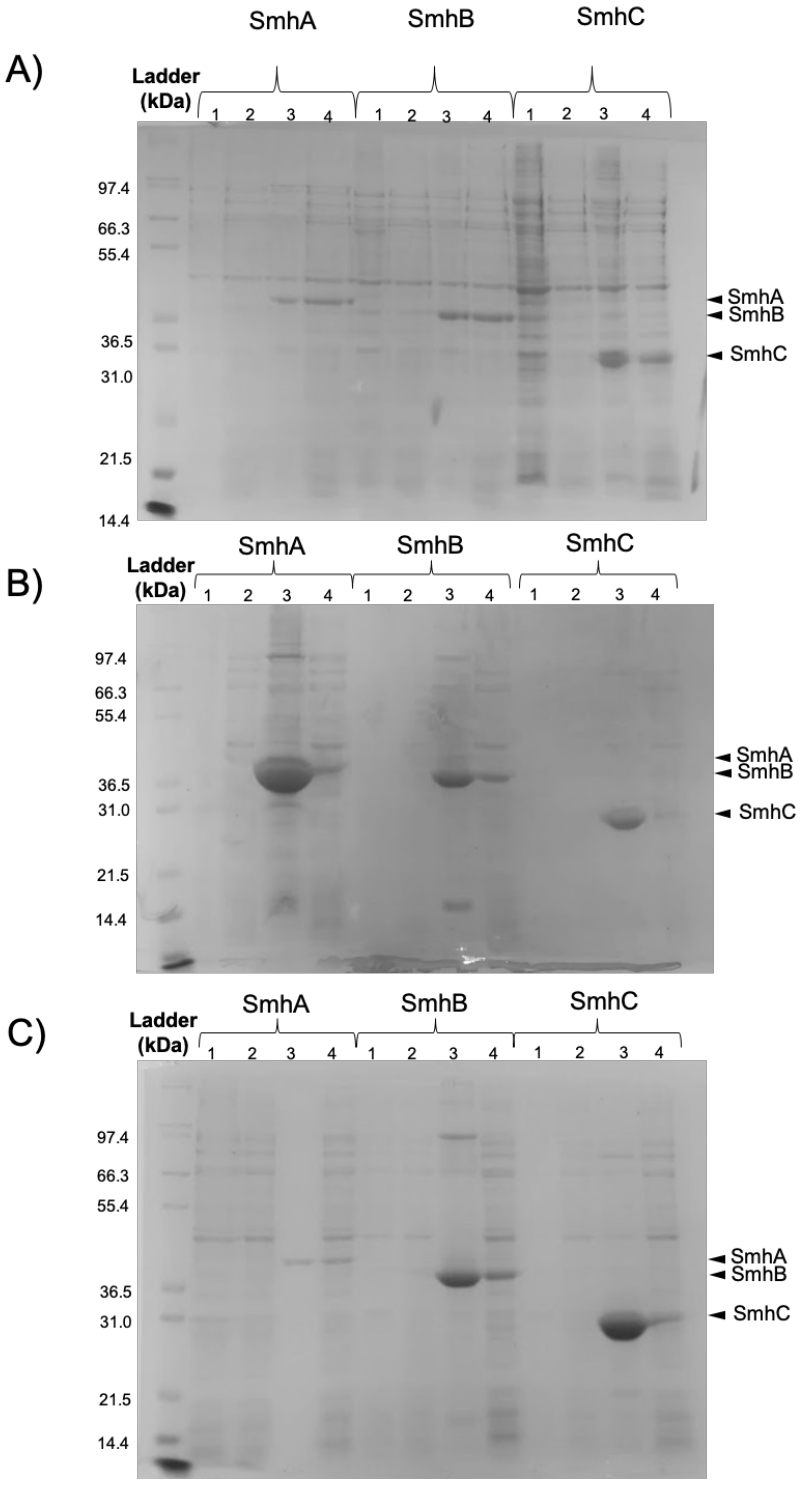


Figure B.1 **SDS-PAGE gels showing 5hr overexpression trials for SmhA, SmhB, and SmhC.** Mark 12 ladder with molecular weights of each band given. For each SmhA, SmhB, and SmhC, lanes 1 and 2 contain pre-induction insoluble and soluble samples respectively, while lanes 3 and 4 contain post-induction insoluble and soluble samples respectively. A) Trials at 16 ^{o}C. B) Trials at 25 ^{o}C. C) Trials at 37 ^{o}C.

2 ml using a Vivaspin 30 kDa MWCO concentrator (Sartorius). This sample was then loaded on to a Superdex 200 pg column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 and 0.5 M NaCl. SmhA eluted at a volume of 65 ml (100 kDa), suggesting a dimer of SmhA in solution (Figure B.2B). All fractions containing SmhA were pooled and concentrated down to a final concentration of 7 mg/ml in a total volume of 0.5 ml for crystallisation and assays. Selenomethionine purification was carried out using the same methods as sulphur methionine expression and resulted in similar yields of soluble protein.

B.2.3 Purification of SmhB

2 g of cell paste was resuspended in 25 ml lysis buffer (50 mM Tris-HCl pH 8) before lysis by sonication (3×20 s bursts at 16,000 nm λ). Insoluble material was removed by centrifugation at 40,000 g and soluble material was applied to a 5 ml Nickel Hi-trap column (GE Healthcare) in binding buffer (50 mM Tris-HCl pH 8 and 0.5 M NaCl). Protein was eluted on a gradient of binding buffer to 50 mM Tris-HCl pH 8, 0.5 M NaCl and 0.5 M imidazole. SmhB eluted at 0.175 M Imidazole (Figure B.3A). Fractions containing SmhA were concentrated down to 2ml using a Vivaspin 30 kDa MWCO concentrator (Sartorius). This sample was then loaded on to a Superdex 200 pg column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 and 0.5 M NaCl. SmhB eluted at a volume of 79 ml (35.4 kDa), suggesting a monomer of SmhB in solution (Figure B.3B). All fractions containing SmhB were pooled and concentrated down to a final concentration of 14 mg/ml in a total volume of 0.5 ml for crystallisation and assays.

B.2.4 Purification of SmhC

2 g of cell paste was resuspended in 25 ml lysis buffer (50 mM Tris-HCL pH 8) before lysis by sonication (3×20 s bursts at 16,000 nm λ). Insoluble material was removed by centrifugation at 40,000 g and soluble material was applied to a 5 ml Nickel Hi-trap column

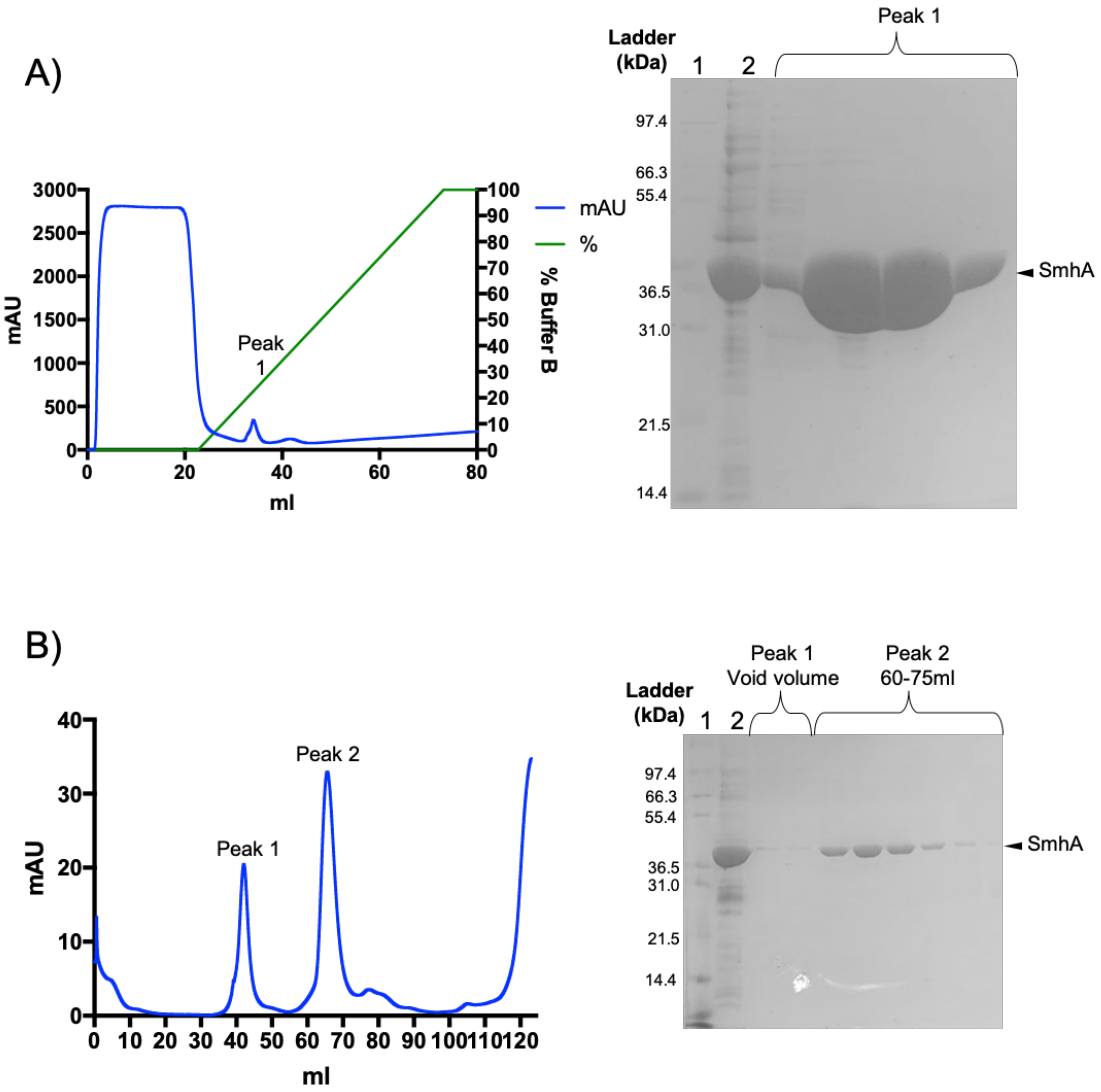


Figure B.2 **Purification of His tagged SmhA. UV trace chromatogram (left) and SDS-PAGE (right).**A) Purification by NiHP. SmhA eluted at 25 % Buffer B. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and NiHP load respectively, with fractions from peak 1 in lanes 3-6, as indicated. B) Gel filtration of SmhA NiHP eluent. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and gel filtration load respectively, with factions from peak 1 and peak 2 from the gel filtration shown in lanes 3-10, as indicated.

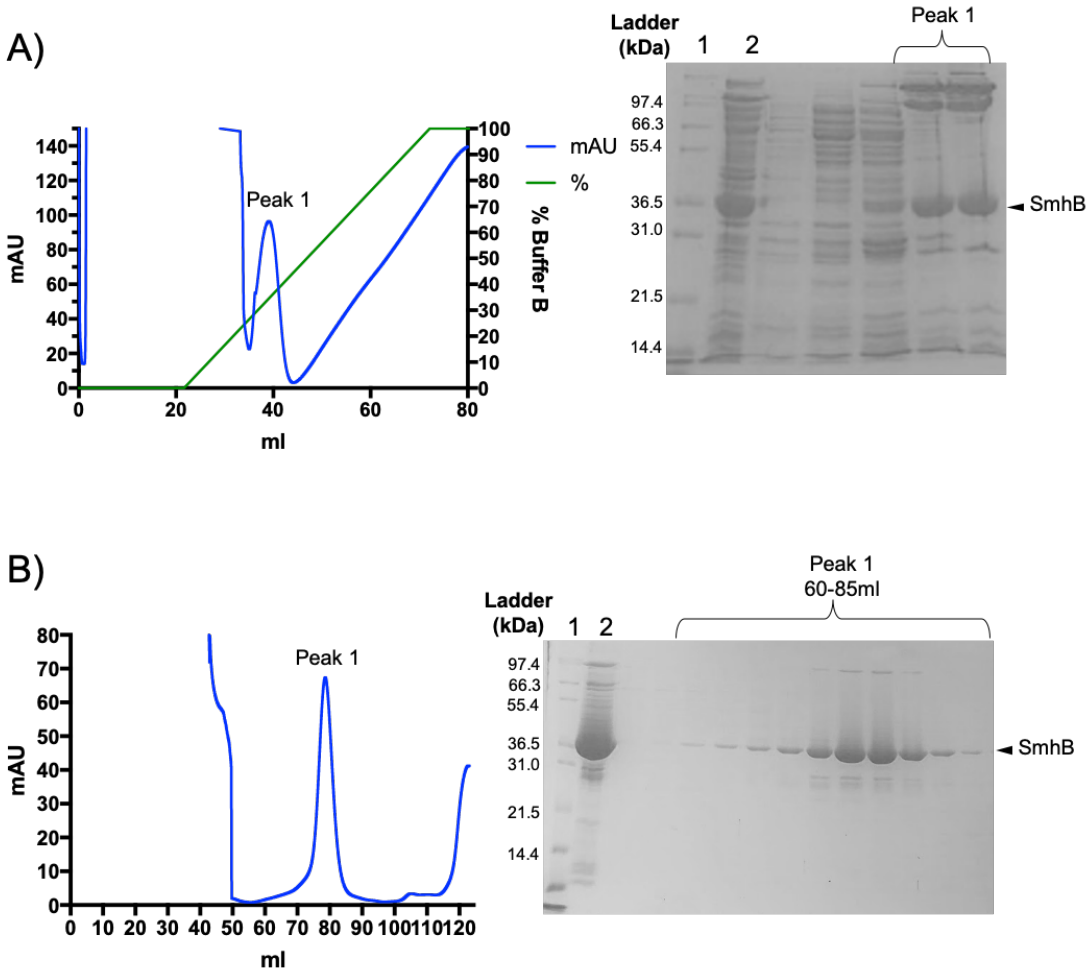


Figure B.3 **Purification of His tagged SmhB. UV trace chromatogram (left) and SDS-PAGE (right).**A) Purification by NiHP. SmhB eluted at 35 % Buffer B. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and NiHP load respectively, with fractions from peak 1 in lanes 6-7, as indicated. B) Gel Filtration of SmhB NiHP eluent. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and Gel filtration load respectively, with factions from peak 1 from the gel filtration shown in lanes 5-14, as indicated.

(GE Healthcare) in binding buffer (50 mM Tris-HCl pH 8 and 0.5 M NaCl). Protein was eluted on a gradient of binding buffer to 50 mM Tris-HCl pH 8, 0.5 M NaCl and 0.5 M imidazole. Although the SmhC construct contained a His_6 tag very little SmhC eluted from the column, at 0.125 M imidazole with only faint bands visible at a molecular weight of 28 kDa on SDS-PAGE. However, total protein was measured as 3.12 mg in 20 ml from absorbance using the nanodrop, which did not correspond with the levels seen in SDS-PAGE analysis suggesting protein concentration measurements were inaccurate (Figure B.4A). Fractions containing SmhC were concentrated down to 2 ml using a Vivaspin 30 kDa MWCO concentrator (Sartorius). This sample was then loaded on to a Superdex 200 pg column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 and 0.5 M NaCl. SmhC eluted as a broad peak between 60-85 ml, and produced two faint bands on SDS-PAGE, with an approximate 50 % purity for SmhC (Figure B.4B). The low levels of protein recovery perhaps suggest that the His_6 tag was occluded in the structure or perhaps cleaved from the construct.

Due to the low yield from the nickel column a second purification method was used which involved a two-step process of ammonium sulphate (AS) cut followed by anion-exchange chromatography. An analytical AS cut was carried out on 0.5 ml SmhC CFE and AS concentrations of 0.5 M, 0.66 M, 0.75 M and 1.5 M. The best ratio of precipitated SmhC relative to impurities was found at 0.75 M AS (Figure B.5A). Pelleted SmhC was resolubilised in 30 ml 50 mM Tris pH 8. Solubilised protein was then applied to a 5 ml DEAE -Sepharose column equilibrated with 50 mM Tris pH 8, and eluted over a gradient of 0 – 0.3 M NaCl using 50 mM Tris pH8, 1 M NaCl. SmhC eluted at 0.05-0.2 M NaCl (Figure B.5B). The first 5 fractions were pooled and concentrated down to a final concentration of 7.5 mg/ml in a total volume of 0.5 ml for crystallisation and assays.

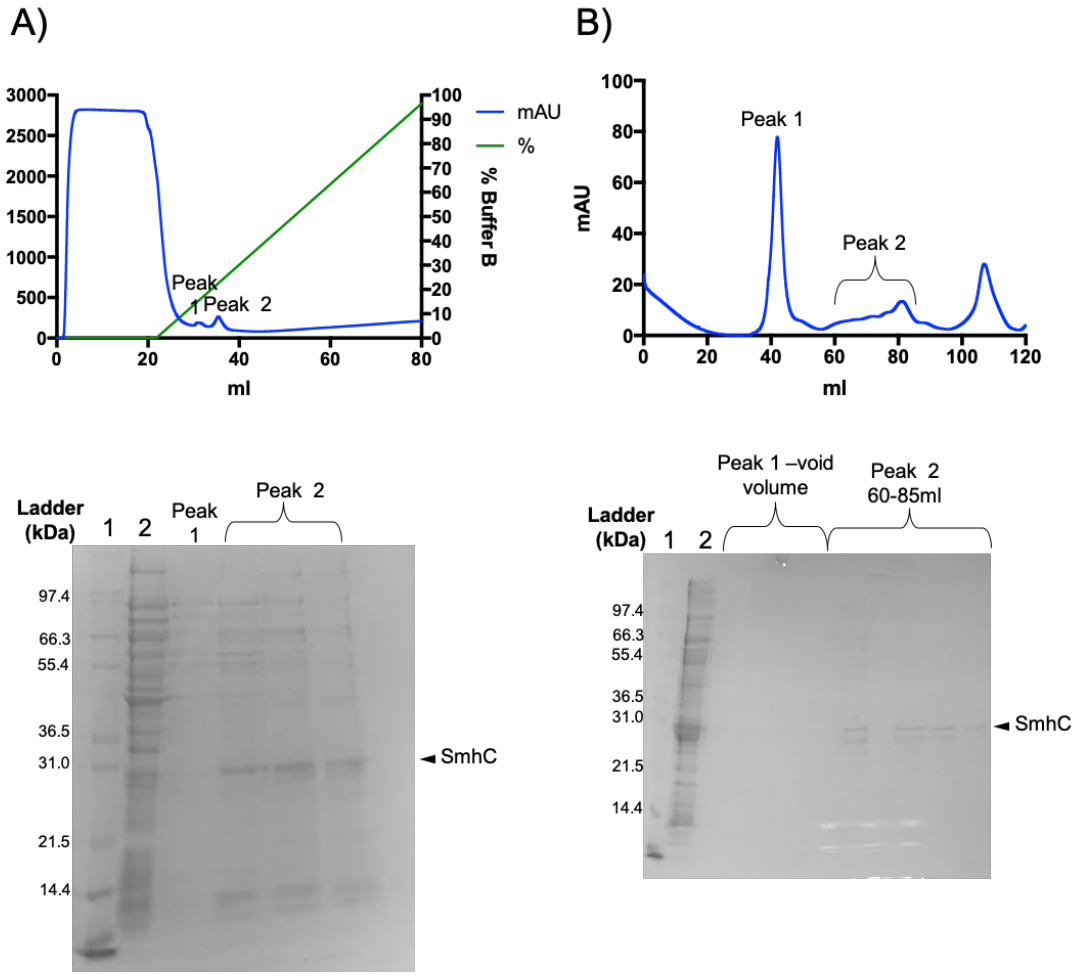


Figure B.4 **Purification of His tagged SmhC. UV trace chromatogram (left) and SDS-PAGE (right).**A) Purification by NiHP. SmhB eluted at 25 % Buffer B. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and NiHP load respectively, with fractions from peak 1 and peak 2 in lane 3-6, as indicated. B) Gel Filtration of SmhB NiHP eluent. Lane 1 and 2 of the SDS-PAGE show Mark 12 ladder and Gel filtration load respectively, with factions from peak 1 and peak 2 from the gel filtration shown in lanes 3-10, as indicated.

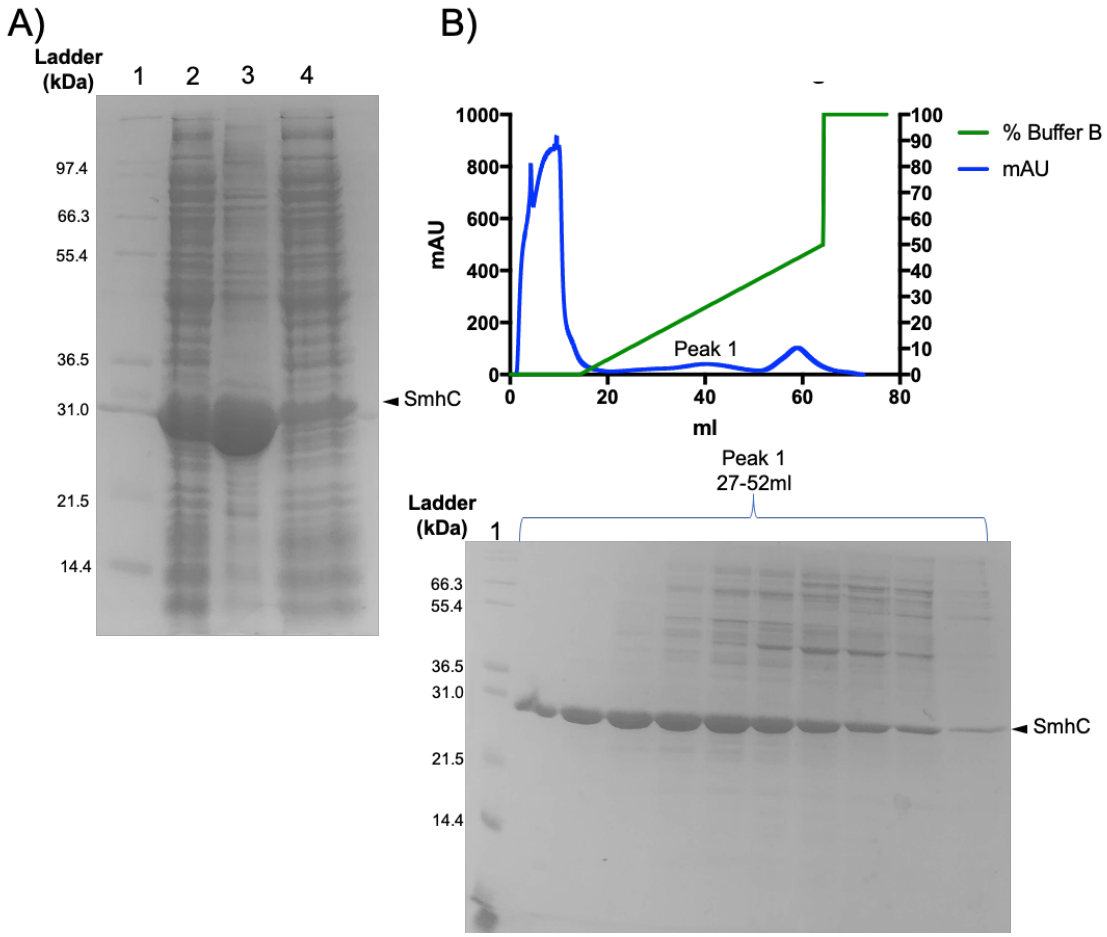


Figure B.5 **Purification of SmhC by AS cut and anion exchange chromatography.**A) SDS-PAGE of ammonium sulphate cut with 0.75M ammonium sulphate. Lane 1, Mark12 ladder; Lane 2, CFE Lane 3, pellet; Lane 4, supernatant. B) Anion exchange UV trace chromatogram of purification of solubilised pellet from AS cut. SmhC eluted between 10 and 40 % NaCl. Below, Right, SDS-PAGE of Anion exchange. Lane 1, Mark 12 ladder. with fractions from peak 1 in lanes 2-11, as indicated.

B.3 Haemolytic sssays for the Smh toxin

As discussed in Paper 3 (Section 5.3), haemolytic assays have shown how the different components of the Smh toxin can act to either enhance or retard lysis when preincubated with erythrocytes. Both SmhB and SmhC slightly enhance lysis if preincubated with erythrocytes before the addition of the remaining two components, however lysis is further enhanced if SmhC is preincubated in combination with SmhA or SmhB before the addition of the final component. In contrast excess SmhA or SmhB retard lysis and form inhibitory soluble complexes when SmhB and SmhC or SmhA, SmhB and SmhC are mixed in solution.

Further time course haemolytic assays were carried out to determine any effect of concentration on the rate of lysis. At equal ratios, SmhA, SmhB, SmhC took 20 minutes to reach 100 %. Reducing the concentration of SmhA to a ratio of 0.5:1:1 resulted in 100 % lysis in 10 minutes while increasing the concentration above 1:1:1 resulted in a reduction of maximum lysis, and at a ratio of 4:1:1 a reduction of rate to 40 minutes for maximum lysis (Figure B.6A). This effect is likely due to off-pathway soluble complex formation between SmhA, SmhB and SmhC, as more SmhA is now available in solution to bind SmhB and SmhC before these two components are able to bind to the membrane. In contrast at low concentrations of SmhA, SmhB and SmhC are free to bind the membrane before interacting with SmhA. Increasing SmhC relative to SmhA and SmhB resulted in an increased rate of lysis and total lysis up to a maximum at a ratio of 1:1:4 (Figure B.6C). Both Increasing or decreasing the concentration of SmhB relative to SmhA and SmhC resulted in decreasing lysis, however while the rate of lysis remained the same as the concentration was increased, decreasing the concentration resulted in a significant reduction in rate (Figure B.6B). At a ratio of 1:0.5:1 lysis remains low until 40 minutes when a sudden increase of 40 % is observed, this could indicate a critical pore concentration is reached on the membrane resulting in complete cell lysis (Figure B.6B). TEM negative stain micrographs of liposomes show that some membranes are saturated while others remain empty, a phenomenon also

observed with Ahl (Paper 1 (Section 3.3))(Wilson et al., 2019) suggesting a preference of pores to concentrate together on a membrane rather than randomly insert (Figure B.6D). Together, these results indicate pore formation by SmhABC is fastest with excess SmhC. Excess SmhB or SmhA hinders pore formation, while reducing SmhA below 0.5:1:1 has little effect. The effects of each Smh protein on lysis are also shared by their equivalent proteins in the Hbl toxin, where excess Hbl-B (equivalent to SmhC) increases rate and lysis, while excess Hbl-L1 (SmhB) or Hbl-L2 (SmhA) both reduce toxic activity (Jessberger et al., 2019).

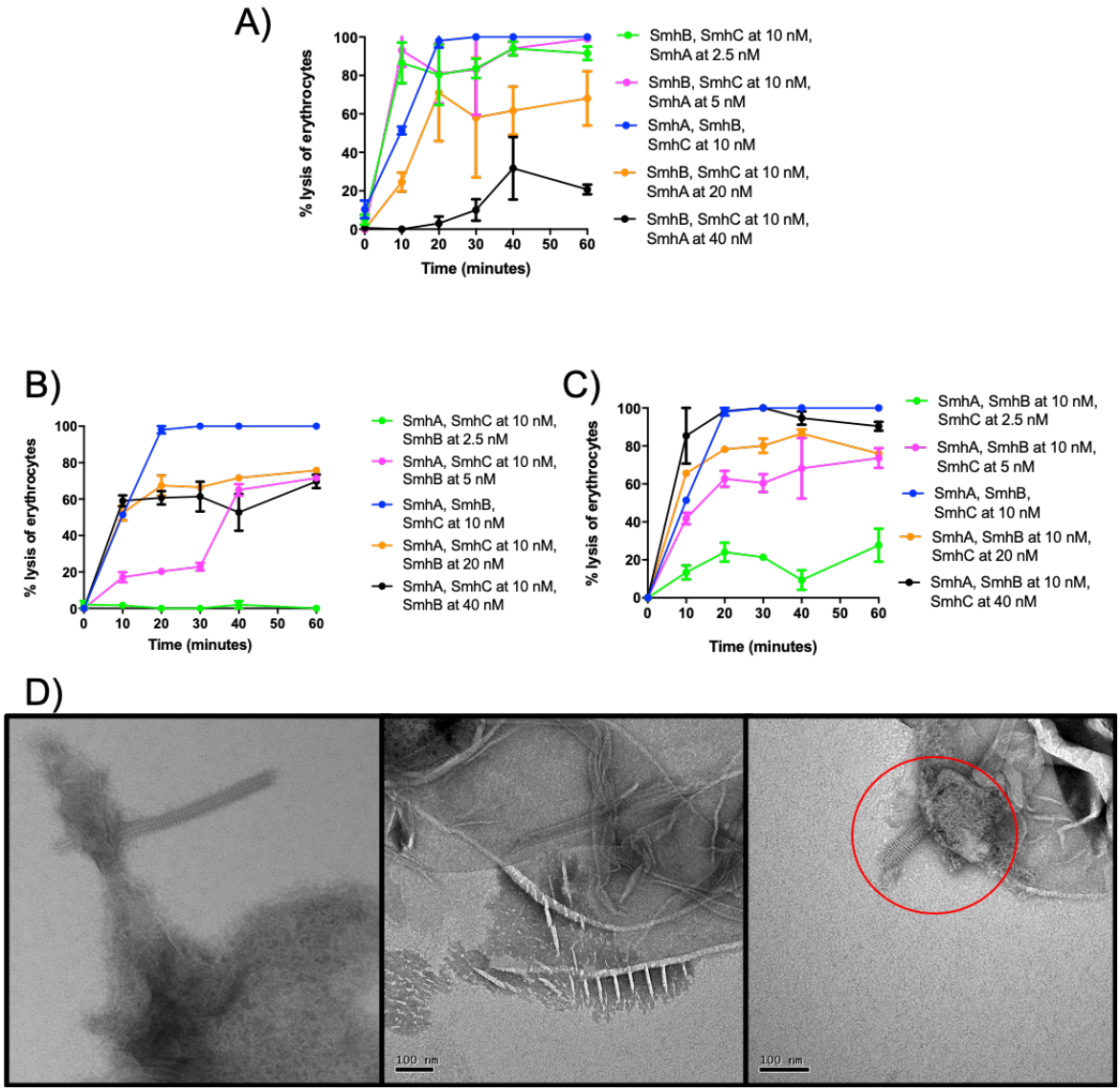


Figure B.6 **Lytic and pore-forming activity of Smh toxin.**A) % lysis against time with varying concentrations of SmhA and fixed concentrations (10 nM) of SmhB and SmhC. B) % lysis against time with varying concentrations of SmhB and fixed concentrations (10 nM) of SmhA and SmhC. C) % lysis against time with varying concentrations of SmhC and fixed concentrations (10 nM) of SmhA and SmhB. D) TEM negative stain EM images of liposomes after incubation with SmhA+SmhB+SmhC. Some liposome membranes are broken down and saturated with pores (left) while others remain empty (centre) or contain small concentrated regions of pores (red circle, right).

B.4 Electron microscopy of soluble complexes

As shown in Paper 3 (Section 5.3) soluble SmhB incubated with soluble SmhC for 1 hour at 37 ^{o}C produced a large molecular weight complex (1.5 MDa), similarly SmhA, SmhB and SmhC incubated together also produced a large molecular weight complex (1 MDa). Both complexes were soluble and the gel filtration peak for SmhABC complex suggested a homogeneous sample, while the SmhBC peak was broader suggesting a less homogeneous sample. Each complex was analysed by negative stain TEM as described in Section 2.6. SmhBC particles were elongated (Figure B.7A) with an approximate diameter estimated from 40 particles of 64 Å $\pm$ 20 Å. These particles resemble two pores side by side. SmhABC particles were much larger (diameter 170 Å $\pm$ 20 Å, estimated from 40 particles) and looked like crowns. Similar crown-like structures were also observed for YaxAB complexes (YaxA and YaxB are equivalent to SmhC and SmhB, respectively) in the absence of membranes or detergent (Figure B.7B) (Bräuning et al., 2018). When YaxAB complexes were mixed with detergent (Cymal-6) single pore particles were produced, which were taken forward for Cryo-electron microscopy and structure solution. To test if this was also the case for the Smh complexes, each complex was mixed with 2 detergents, Cymal-6 (used for YaxAB (Bräuning et al., 2018)) and N-heptyl-thioglucopyranoside (used for Ahl (Wilson et al., 2019)). 0.15 mg/ml of SmhABC or SmhBC was mixed with 0.56 mM Cymal-6 or 30 mM N-heptyl-thioglucopyranoside. The resulting samples were then analysed by TEM negative stain microscopy. In both cases, N-heptyl-thioglucopyranoside produced the cleanest sample with the most single particles (Figure B.8 and Figure B.9). These particles had an average diameter estimated from 40 particles of 81Å $\pm$ 20 Å (SmhABC) and 67 Å $\pm$ 20 Å (SmhBC), comparable to the diameter of AhlB pores and SmhB pores. The diameter of SmhABC particles with detergent is approximately half that observed without detergent, suggesting particles without detergent may be composed of two pores side by side, or that SmhA is stripped off the soluble inhibitory complex by detergent, resulting in a smaller SmhB or

SmhBC pore. Some larger complexes are still visible in both samples and some particles may be detergent micelles rather than pores, further optimisation is therefore needed to produce a sample for cryo-EM and structure solution.

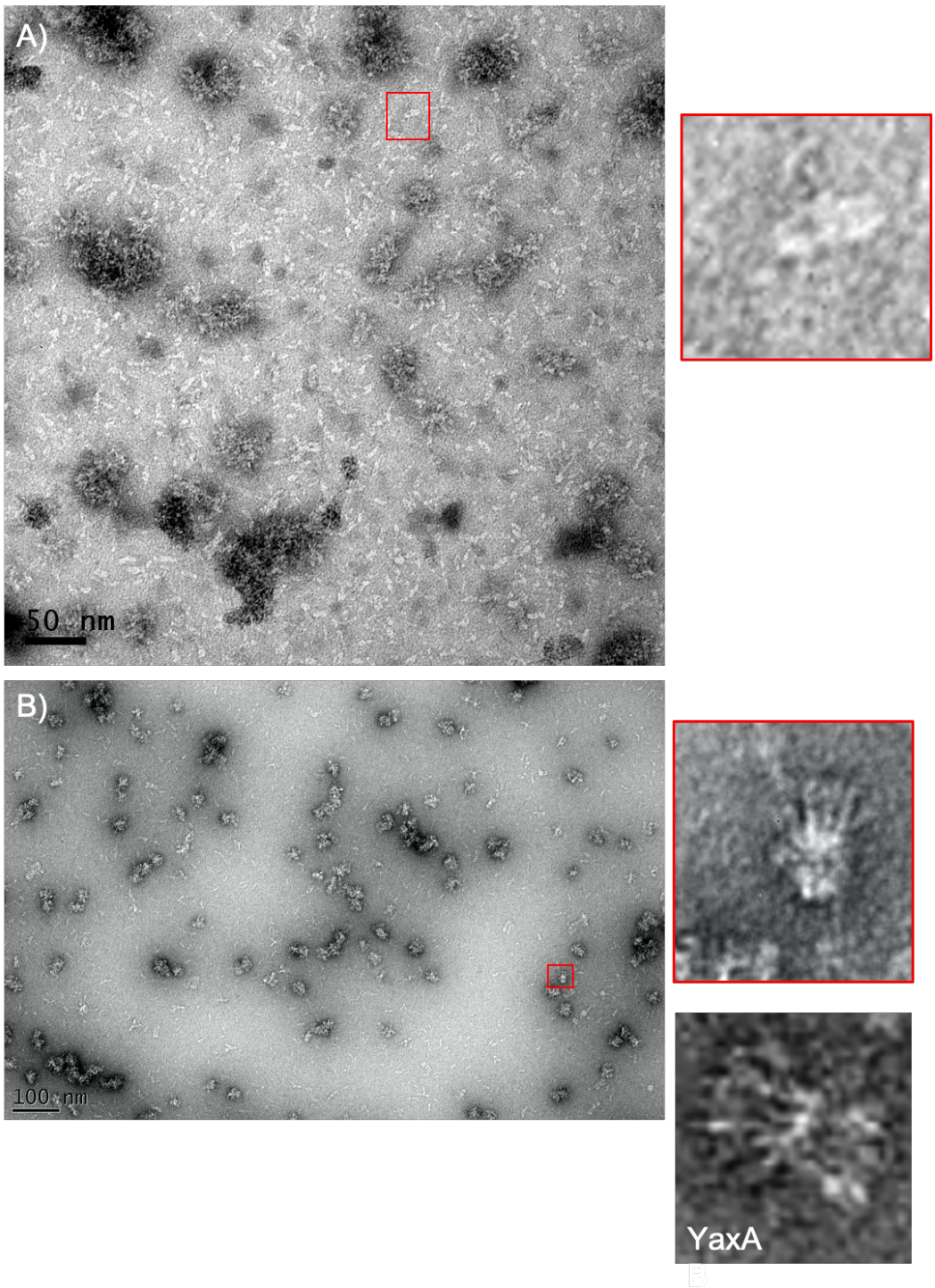


Figure B.7 **TEM negative stain EM images of gel filtration fractions.** A) SmhBC soluble complexes, long structures are visible (right) which could represent two pores side by side. B) SmhABC soluble complexes. Large splayed crown-like structures are seen (right), similar to those observed for soluble YaxAB complexes shown below (image adapted from Bräuning et al. (2018)).

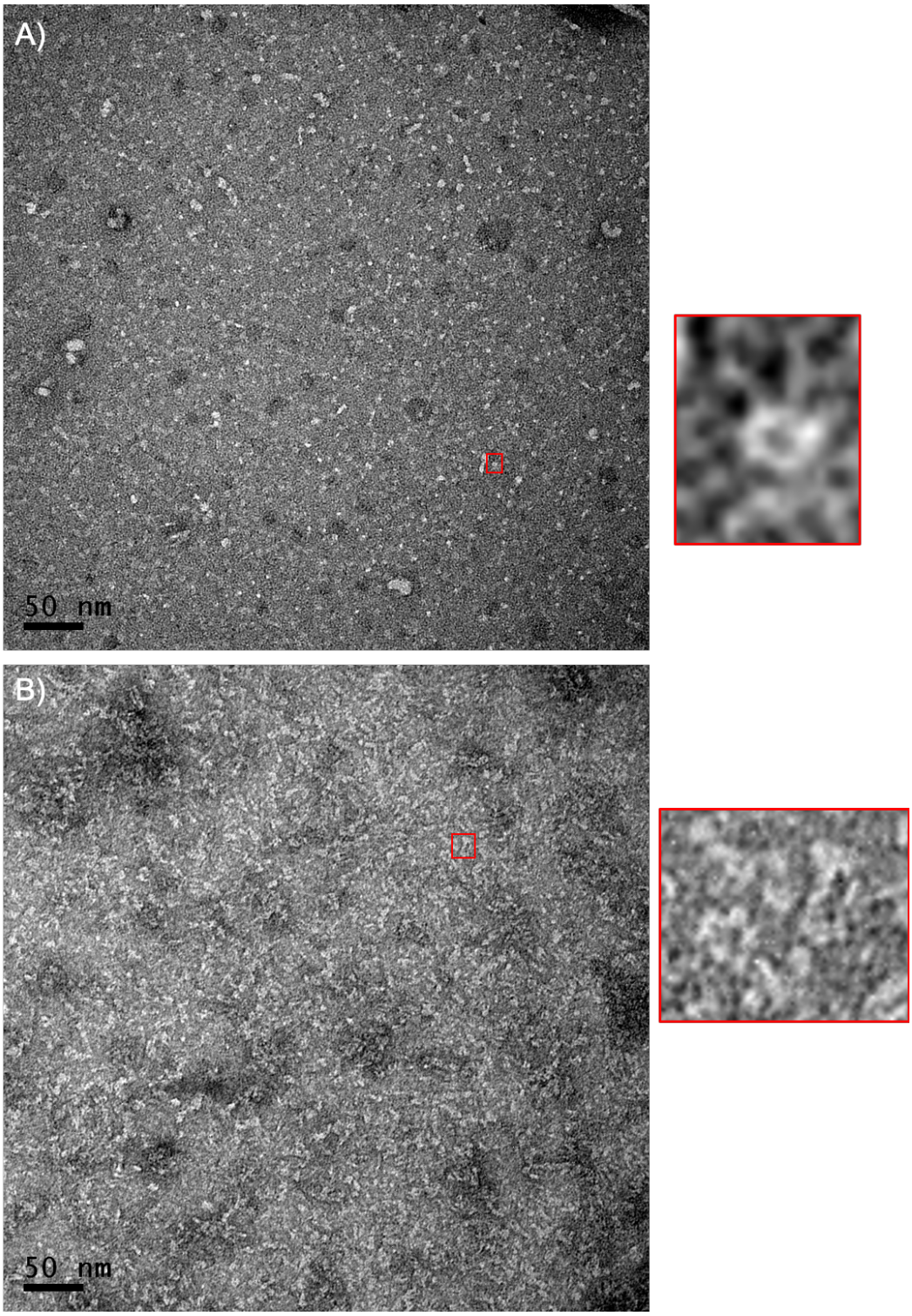


Figure B.8 **TEM negative stain EM images of SmhBC soluble complexes incubated with detergents.** A) SmhBC complex incubated with detergent N-heptyl-thioglucopyranoside. SmhBC separates into more single particles, top views of pores are also visible (right). B) SmhBC complex incubated with detergent Cymal-6, some single particles and many aggregates are visible (right).

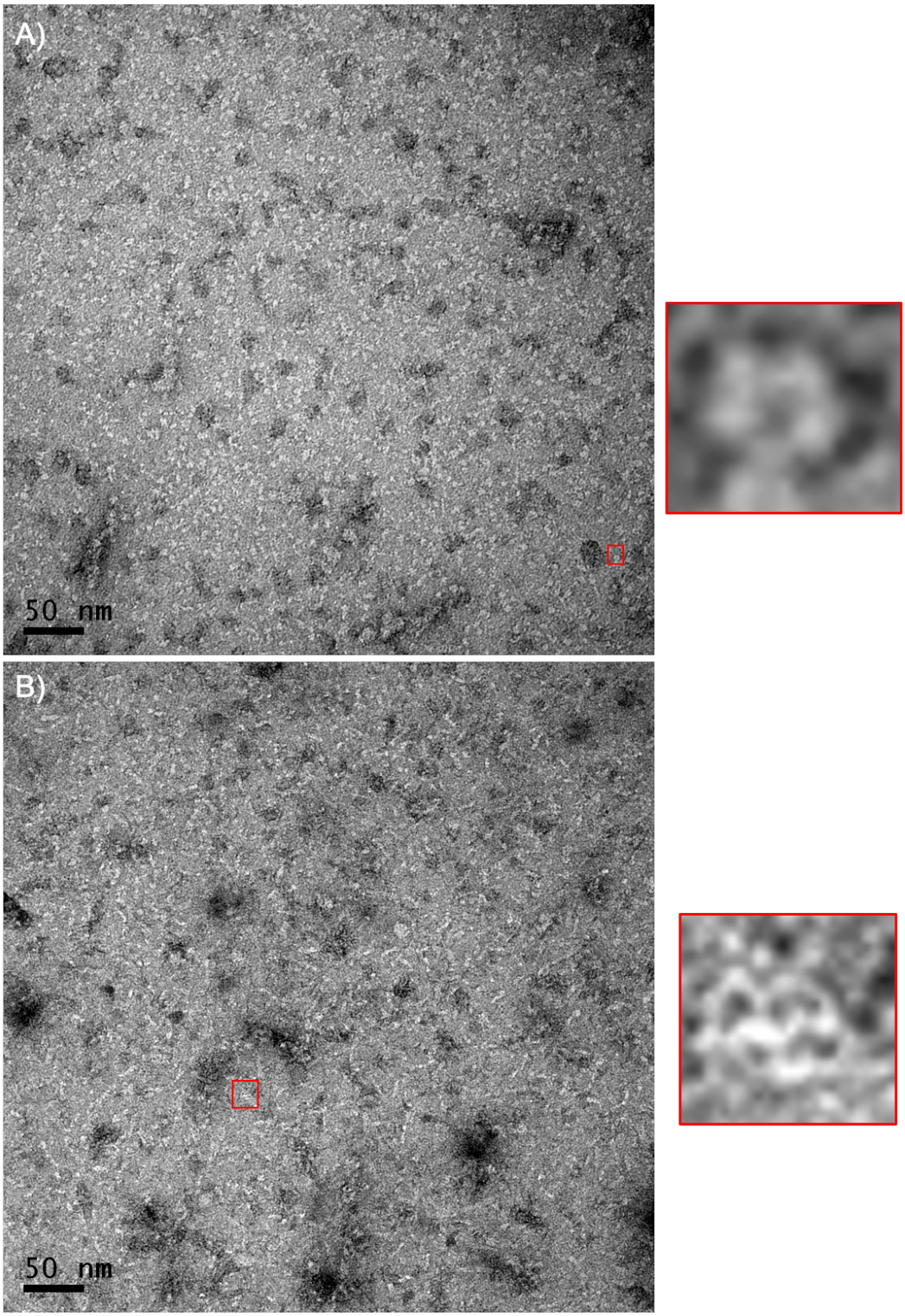


Figure B.9 **TEM negative stain EM images of SmhABC soluble complexes incubated with detergents.** A) SmhABC complex incubated with detergent N-heptyl-thioglucopyranoside. SmhABC separates into more single particles, top views of pores are also visible (right). B) SmhABC complex incubated with detergent Cymal-6, some single particles and many aggregates are visible (right).

B.5 Crystallisation, data collection and structure determination of SmhB

B.5.1 Crystallisation trials of SmhB

SmhB was purified and concentrated before being subject to crystallisation trials as described in Section 2.5.5 and Paper 3 (Section 5.3). Trials were carried out in a 96 well sitting drop plate. Plates were then incubated at 17 ^{o}C and checked regularly. Crystals grew in three conditions described in Paper 3 (Section 5.3). All crystals were of sufficient size for subsequent data collection (Figure B.10), and as such were mounted on litholoops (molecular dimensions) and preserved in liquid nitrogen as described in Section 2.5.4.

B.5.2 X-ray data collection of SmhB

Three different crystal structures of SmhB were determined, these were SmhB crystal Form 1, SmhB crystal Form 2 and SmhB pore form. Data were collected from a single crystal of each crystal form as described in Section 2.5.7 and Paper 3 (Section 5.3), with full data processing statistics given in Section 5.3, Paper 3, Table 1. Table B.1 gives data collection parameters, AU and solvent content for each crystal, with an example of diffraction images provided in Figures B.11-B.13.

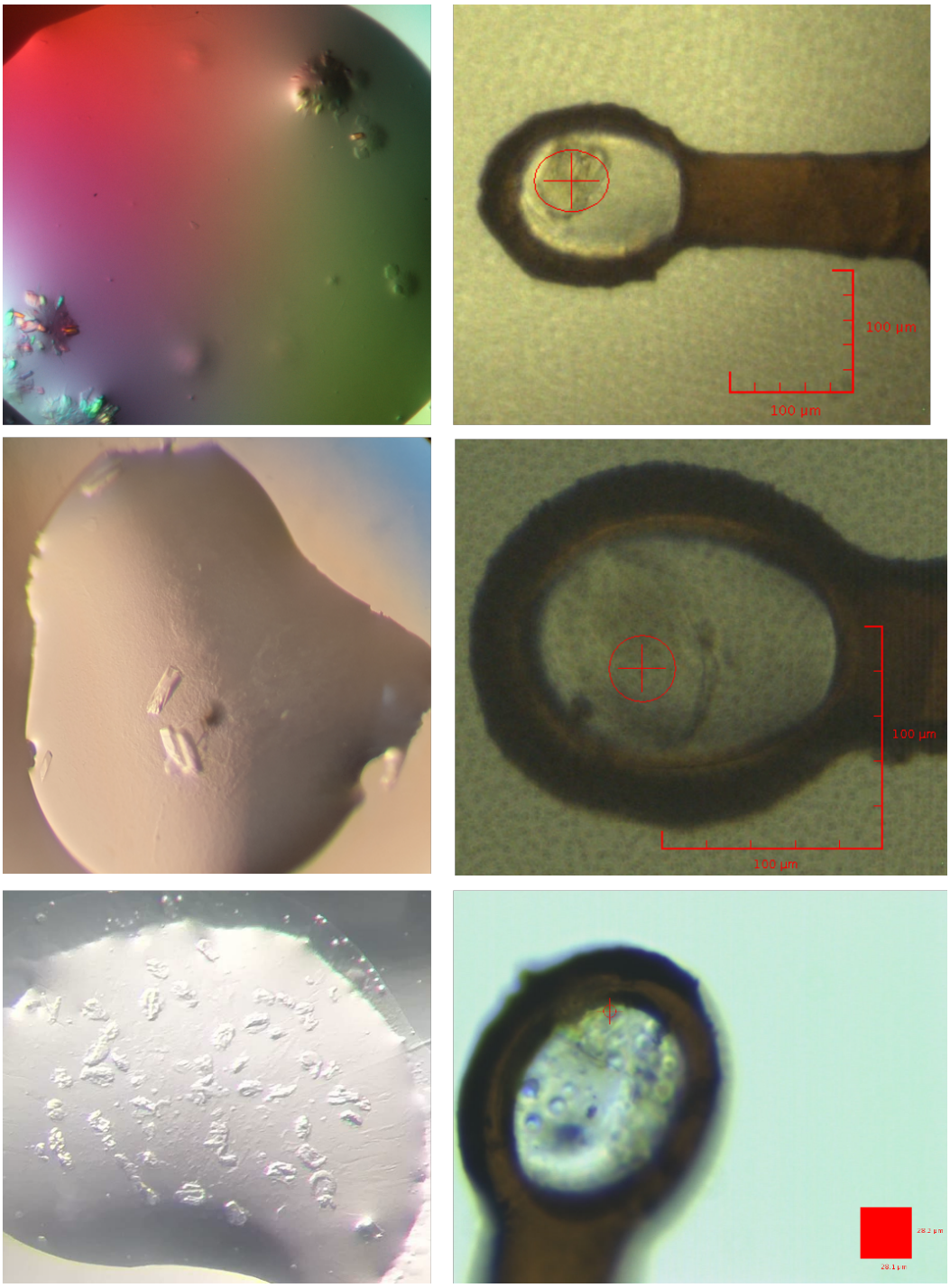


Figure B.10 **Crystals of SmhB (left) and mounted crystal in litholoop (right).** Top, Crystals of soluble SmhB crystal Form 1. Centre, crystals of soluble SmhB crystal Form 2. Bottom, crystals of SmhB pore form.

Table B.1 Data collection parameters for SmhB crystal forms.

crystal Form	1	2	Pore
Oscillation (°)	0.1	0.1	0.1
images	2000	3600	2000
Exposure (s)	0.1	0.05	0.05
Transmission (%)	100	100	100
Space group	$P2_12_12_1$	$P2_1$	C2
Cell dimension			
a, b, c = (Å)	50.94, 113.95, 130.37	95.93, 49.73 99.49	224.5, 118 209.5
α , β, γ = (°)	90, 90, 90	90, 118.43, 90	90, 109, 90
Resolution (Å)	1.84 Å	2.57 Å	6.98Å
Solvent (%)	50	55	64
Molecules in AU	2	2	10

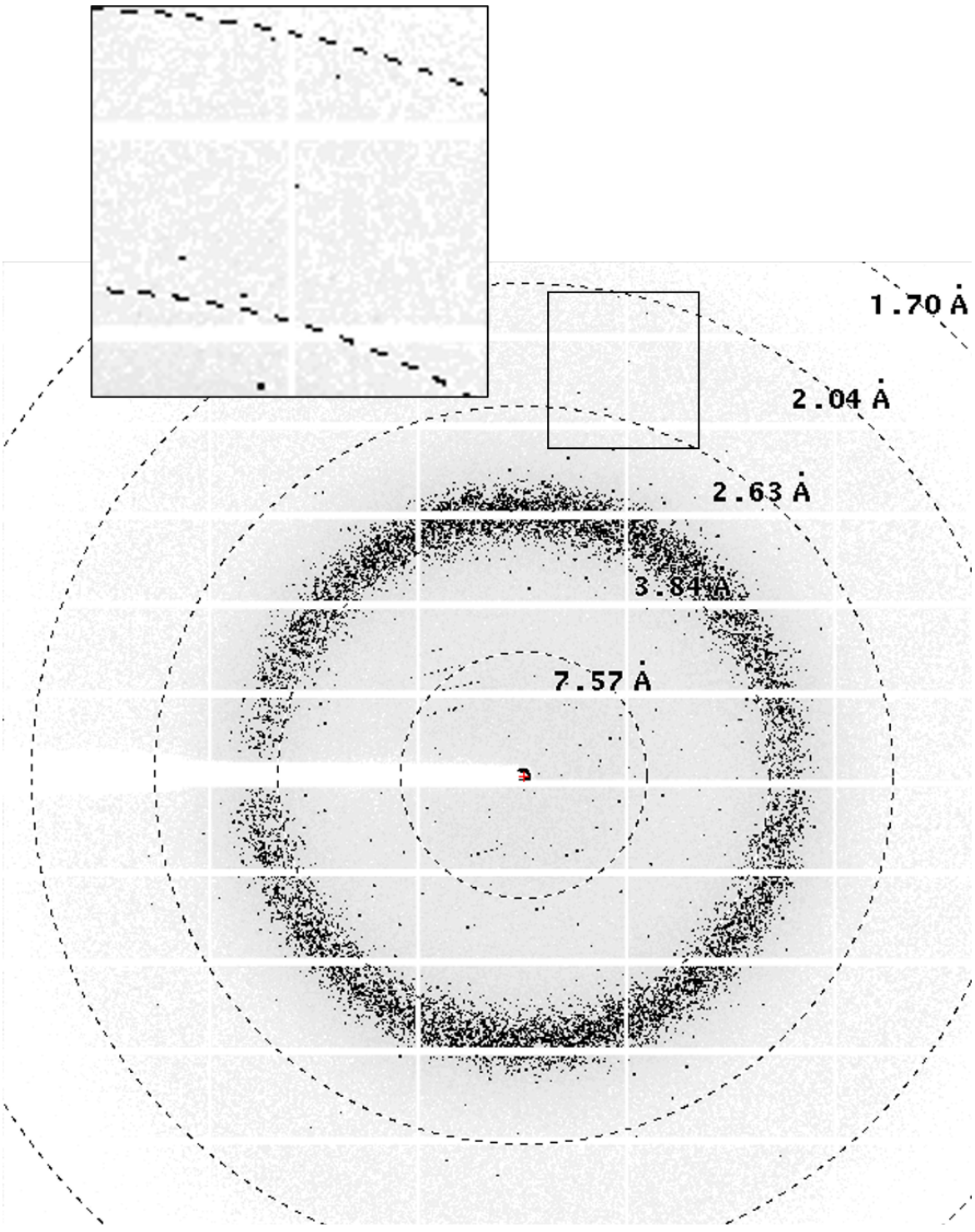


Figure B.11 **A representative 0.1 º oscillation image from SmhB crystal Form 1.** An enlarged view of the region highlighted by the square shows diffraction extends to around 2 Å.

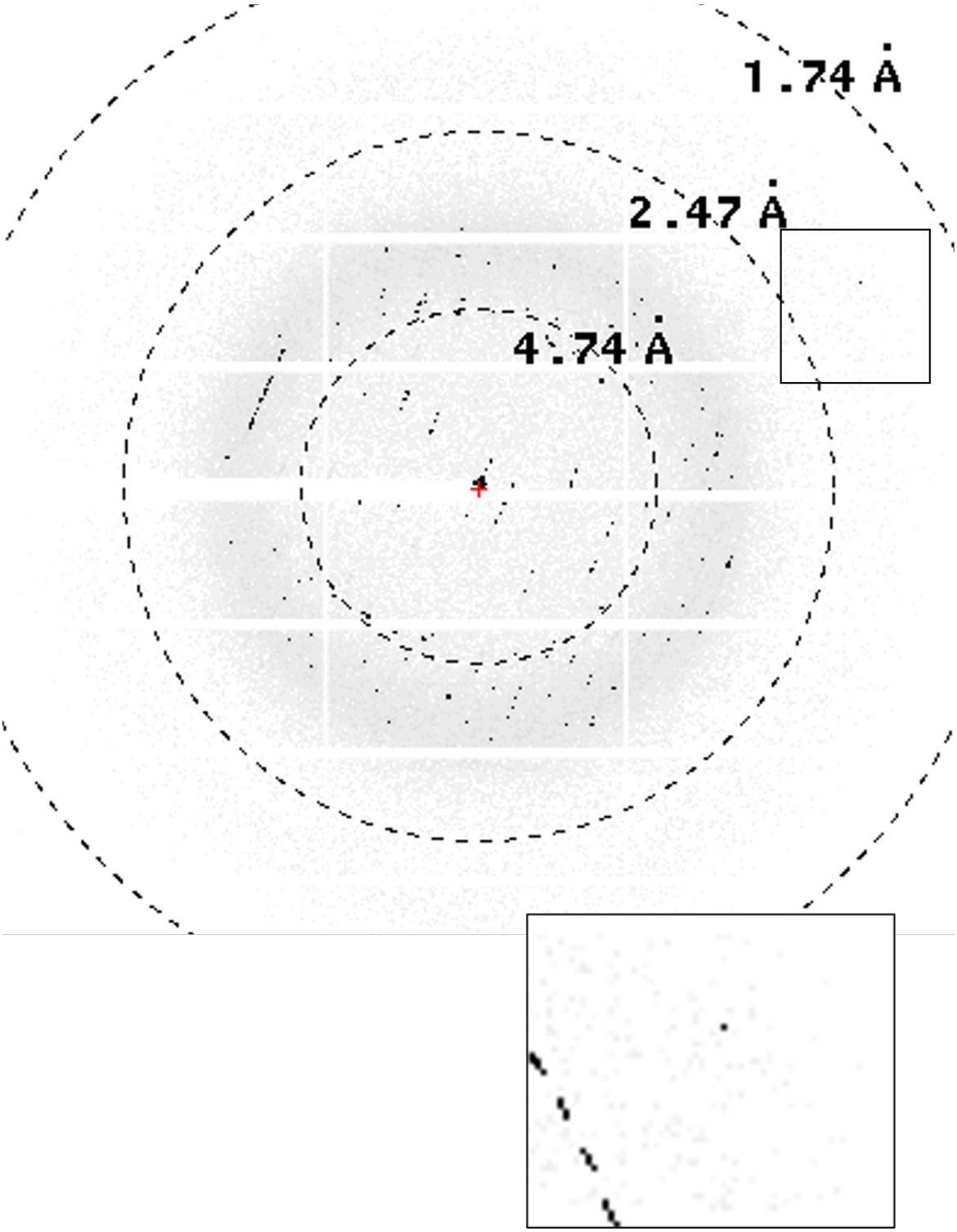


Figure B.12 **A representative 0.1 º oscillation image from SmhB crystal Form 2.** An enlarged view of the region highlighted by the square shows diffraction extends to around 2 Å.

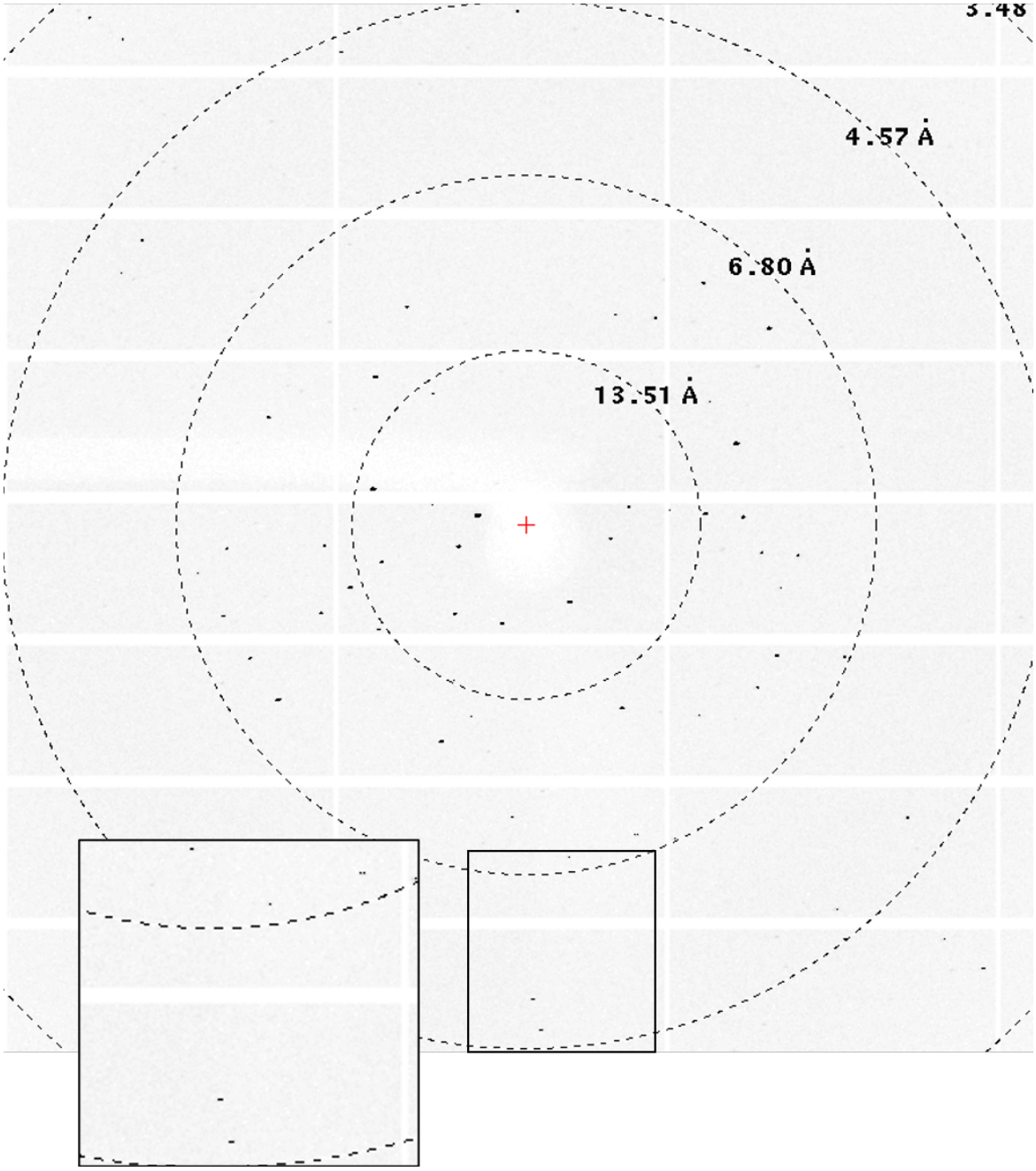


Figure B.13 **A representative 0.1 ^{o} oscillation image from SmhB pore Form.** An enlarged view of the region highlighted by the square shows diffraction extends to around 5 Å.

B.5.3 Structure determination, model building and validation of SmhB crystal Form 1 and crystal Form 2

Initial phases for crystal Form 1 were calculated using Phaser MR (McCoy et al., 2007) using a poly-Ala model of soluble AhlB, as described in Section 2.5.10 and Paper 3 (Section 5.3). This gave a single unique solution with a refined TFZ=38.24 and LLG=1609.78 and placed two molecules in the AU. Refinement of this model using REFMAC5 (Murshudov et al., 1997) gave a R-factor of 0.42 and Free R of 0.44. Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.19 and Free R of 0.22 (Figure B.14), full validation details can be found in Section 5.3, Paper 3, Table 1. SmhB crystal Form 1 is composed of 2 monomeric chains of the soluble conformation of SmhB, in agreement with the gel filtration analysis. Both chains of SmhB are almost identical apart from helix α7, which in Chain A is a single kinked helix (the kink is located around residues S320 and T321), while in chain B helix α7 is broken into two helices with residues S320-T324 in a loop between these helices. This variation is caused by differences in crystal packing between each chain and the neighbouring symmetry-related molecule. In Chain A, α7 packs against α3 in the tail domain of a symmetry-related chain B. However in chain B α7 interacts with the loop between α1 and α2 and β4 and α8 in the head domain of a symmetry-related chain A, resulting in the unfolding of the helix (Figure B.15).

Initial phases for Crystal Form 2 were also calculated using Phaser MR (McCoy et al., 2007) using Chain A of SmhB crystal Form 1 as a model, as described in Section 2 and Paper 3. This gave a single unique solution with a refined TFZ=63.57 and LLG=9095.86 and placed two molecules in the AU. Refinement of this model using REFMAC5 (Murshudov et al., 1997) gave a R-factor of 0.29 and Free R of 0.33. Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.21 and Free R of 0.25, full validation details

can be found in Paper3, Table 1. As with crystal Form 1, crystal Form 2 contained two monomeric chains of the soluble conformation of SmhB (Figure B.16A). This second crystal form provides a third conformation for α7. In chain A, α7 on one side packs against α7 on a symmetry-related chain B and is exposed to the solvent on the opposite side. This results in the same kink as seen in crystal form 1 chain A, but with a distorted helix at the N-terminus (Figure B.16). In chain B α7 packs against α7 and α2 of a symmetry-related chain A stabilising the helix and is structurally the same as chain A of crystal form1 (Figure B.16C). As described in Paper 3, this reveals flexibility in α7, an important helix likely involved in the movement of the C-terminus and ultimate transformation from soluble to pore conformations.

B.5.4 Structure determination, model building and validation of SmhB pore form.

Initial phases for the SmhB pore were calculated using Phaser MR (McCoy et al., 2007) using a poly-Ala model of the AhlB pore, as described in Section 2 and Paper 3. This gave a single unique solution with a refined TFZ=19.45 and LLG=519.59 and placed 1 molecule in the AU. Refinement of this model using REFMAC5 (Murshudov et al., 1997) gave a R-factor of 0.32 and Free R of 0.34. Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.33 and Free R of 0.34. Full validation details can be found in Paper 3, Table 1. Due to the low resolution of the data side chains could not be refined and were left as poly-Ala. Due to the low resolution, checks were carried out to ensure the structure solution by Phaser was correct. The head domain (residues 157-251 in chains A and B) was removed from the Phaser output model and the remaining model refined against the experimental data in REFMAC5. Clear positive density could be seen for the head domains of both chains, this provided convincing evidence that the structure was correct (Supplementary figure Paper 3, Section 5.3).

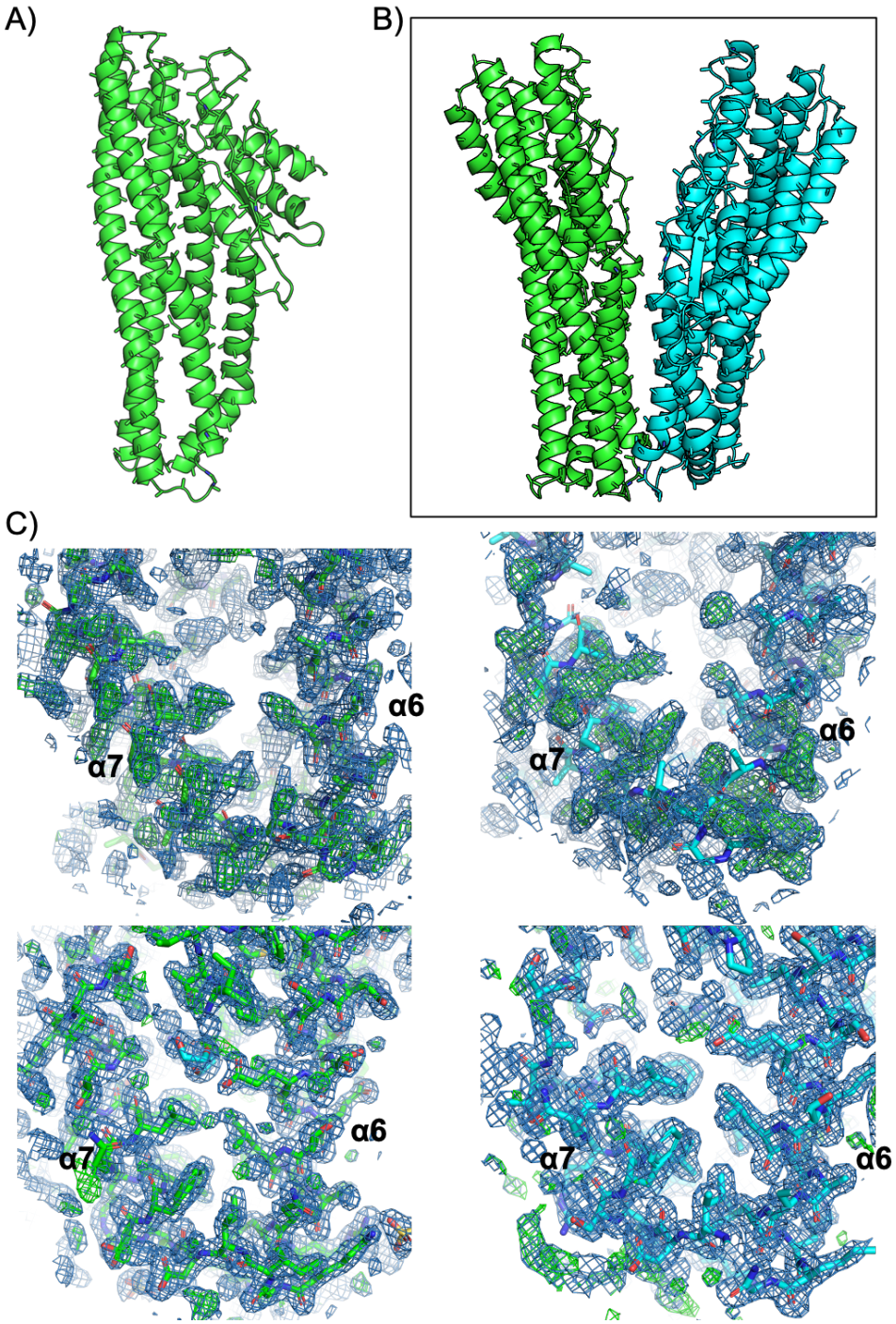


Figure B.14 **Molecular replacement, model building and refinement of SmhB crystal Form 1.** A) Poly-ala model of AhlB used for molecular replacement. B) PhaserMR placed two molecules in the asymmetric unit (green – chain A, blue –chain B). C) Models and electron density around $\alpha 6$ and $\alpha 7$ for chain A (left) and chain B (right) of soluble SmhB, Form 1. Top, after PhaserMR, bottom, after rebuilding and refinement. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh)

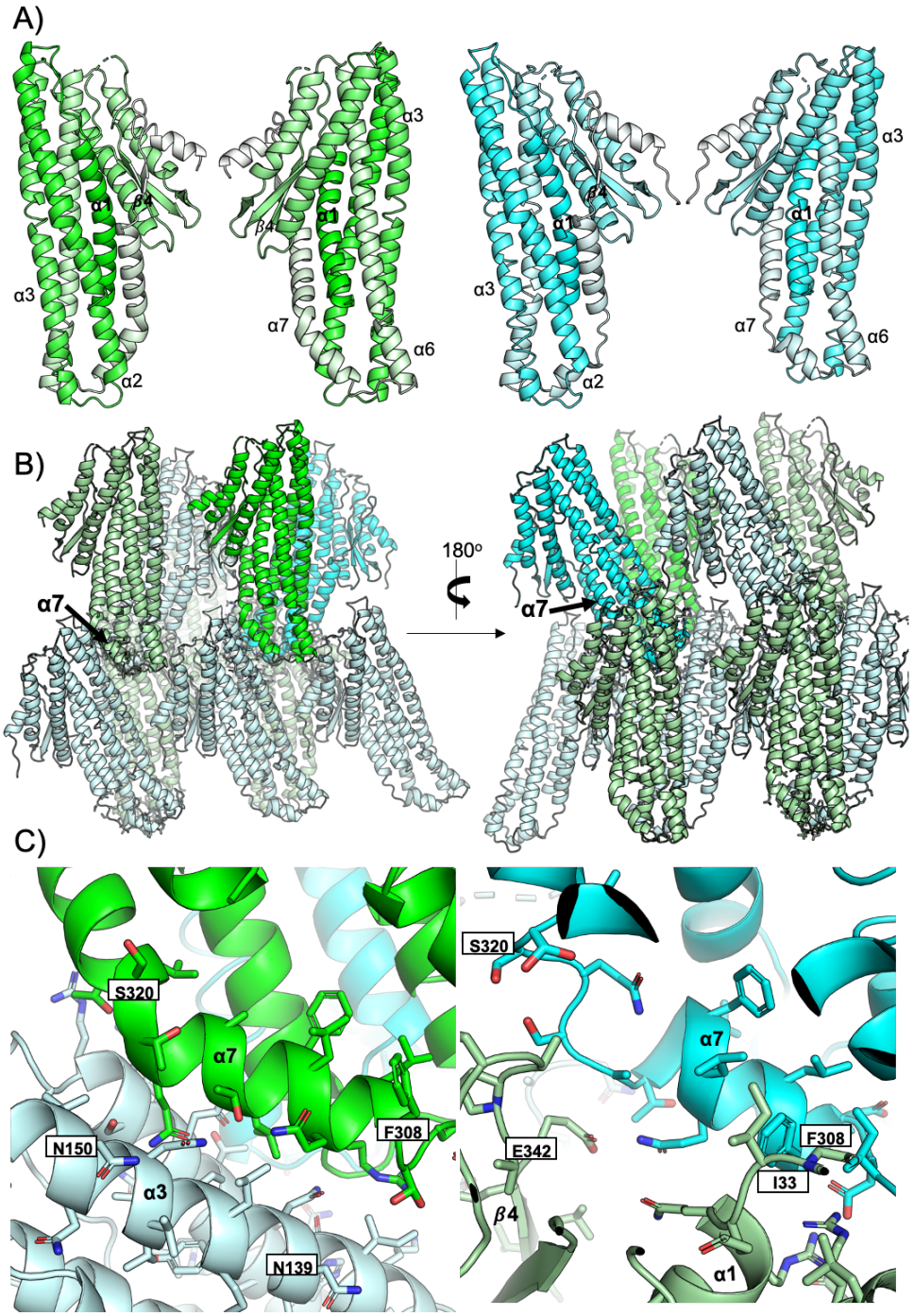


Figure B.15 **The final model of SmhB crystal Form 1 shows differences in α7 between chain A and B caused by crystal packing interactions.** A) Chain A (green) has a single kinked α7, while in chain B (blue) α7 is split into two helices. Both chains coloured dark (N-terminus) to light (C-terminus). B) Crystal packing (symmetry-related molecules shown as light green and light blue, chain A and chain B, respectively) shows α7 in chain A and B pack against different regions of symmetry-related molecules. C) Close up showing the interactions in the same view as B). In chain A (left) α7 packs against α3 of a symmetry-related chain B, while in chain B (right) α7 packs against the loops between α1 and α2 and α4 and α8 of a symmetry-related chain A.

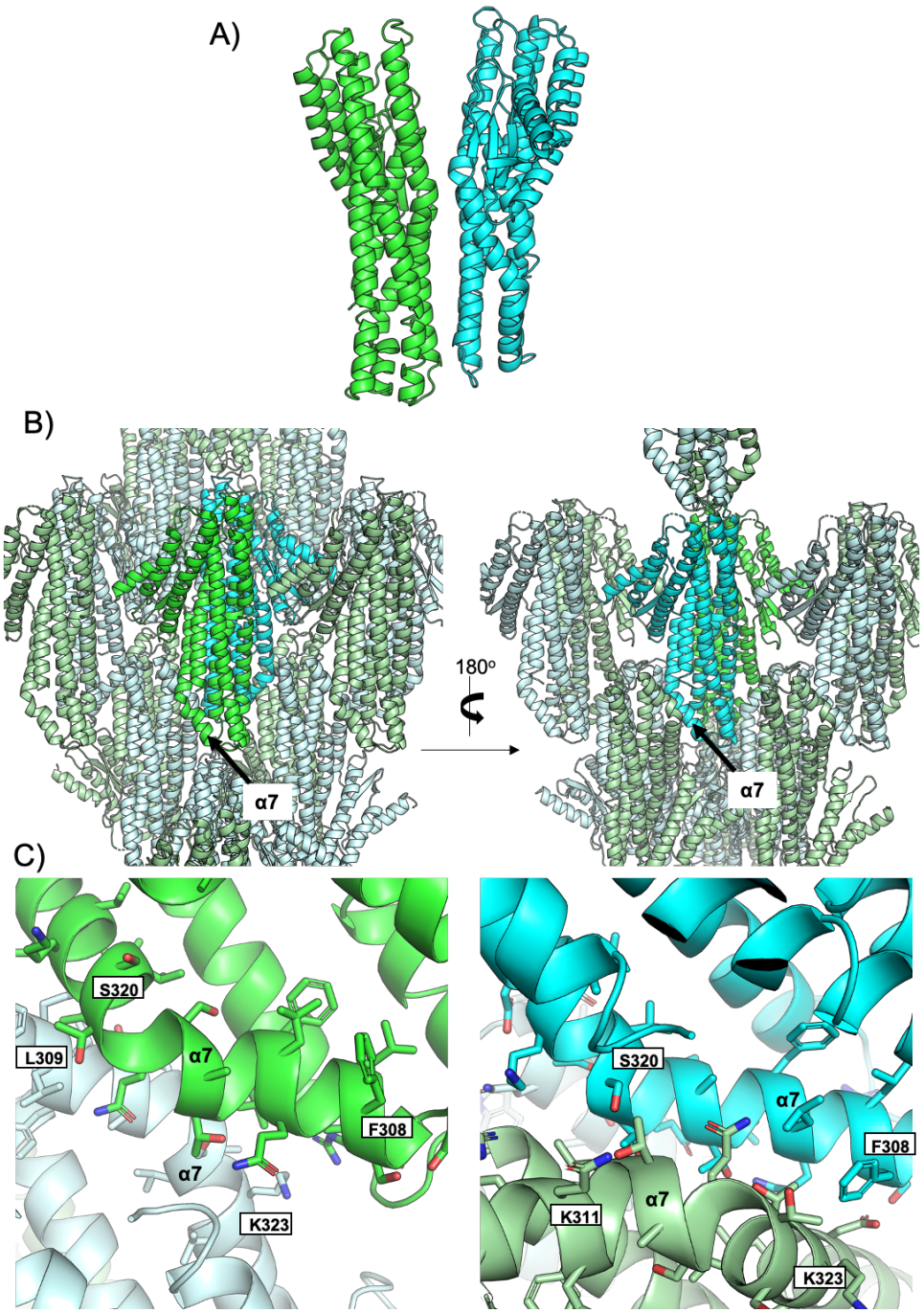


Figure B.16 **SmhB crystal Form 2 shows further differences in α7 between chain A and B caused by crystal packing interactions.** A) Final model of SmhB crystal form 2, green, chain A. blue, chain B. B) Crystal packing (symmetry-related molecules shown as light green and light blue, chain A and chain B respectively) shows α7 in chain A and B pack against different regions of symmetry-related molecules. C) Close up showing the interactions in the same view as B). In chain A (left) α7 packs partially against α7 of a symmetry-related chain B and is also exposed to solvent, while in chain B (right) α7 packs between α7 and α2 of a symmetry-related chain A.

B.6 Crystallisation, data collection and structure determination of SmhA

B.6.1 Crystallisation trials of SmhA

SmhA was purified and concentrated before being subject to crystallisation trials as described in Section 2.5.5, and Paper 2 (Section 4.3). Trials were carried out in a 96 well sitting drop plate. Plates were then incubated at 17 ^{o}C, which resulted in a large number of hits containing a multitude of microcrystals after 1 day. As a result, screens were repeated at 7 ^{o}C to slow down crystal growth and hopefully result in larger crystals.

Crystals grew in five conditions shown in Table B.2. All crystals were of sufficient size for subsequent data collection (Figure B.17), and as such were mounted on litholoops (Molecular Dimensions) and preserved in liquid nitrogen as described in Section 2.5.4.

B.6.2 X-ray data collection of SmhA

Data were collected from a single crystal of each SmhA crystal form, as detailed in Table B.2 and B.3. Crystal form 1 was successfully phased by SAD and the model of this structure used to determine the other crystal Forms using MR. All of these structures were in the soluble conformation of SmhA. As crystal Form 2 had the highest resolution and lowest number of molecules in the AU (4 vs 16 for Form 3, 4 and 5) this crystal form was used in subsequent analysis. Attempts were made to refine and build the structures of crystal Forms 3, 4 and 5, all of which had 16 copies in the AU, but the resolution of the data was such that many side chains could not be placed confidently and thus little extra information could be gleaned. Crystal Forms 1 and 2 will be described from this point onwards.

Crystal form 1 contained Selenomethionine derivatised SmhA, and thus a SAD data set of 7200 images was collected as described in Section 2.5.9 (full data processing statistics given in Section 5.3, Paper 3, Table1), with an oscillation per image of 0.1 ^{o} (total 720 ^{o}),

Table B.2 Crystallisation conditions and data collection statistics for different crystal forms of SmhA.

Model name	Crystallisation condition	Space group	Resolution	Molecules in AU
Crystal form 1 (Semet)	0.16 M $CaCl_2$, 0.1 M MES pH 6.5, 20 % PEG 6000	$P4_2$	2.98 Å	8
Crystal form 2	0.2 M KNO_3, 20% PEG 3350	$P2_12_12_1$	2.57 Å	4
Crystal form 3	0.12 M $CaCl_2$, 0.1 M MES pH 5.8, 20 % PEG 6000	$P4_3$	3 Å	16
Crystal form 4	0.3 M $NaNO_3$, 0.3 M Na_2HPO_4, 0.3 M $(Na_4)_2SO_4$ 0.1 M MES/imidazole pH 6.5 10 % PEG 20,000, 20 % PEG MME 550,	$P12_11$	2.96 Å	16
Crystal form 5	0.2 M KI, 0.1 M MES pH 6.5, 25 % PEG 4000	P4	3.36 Å	16

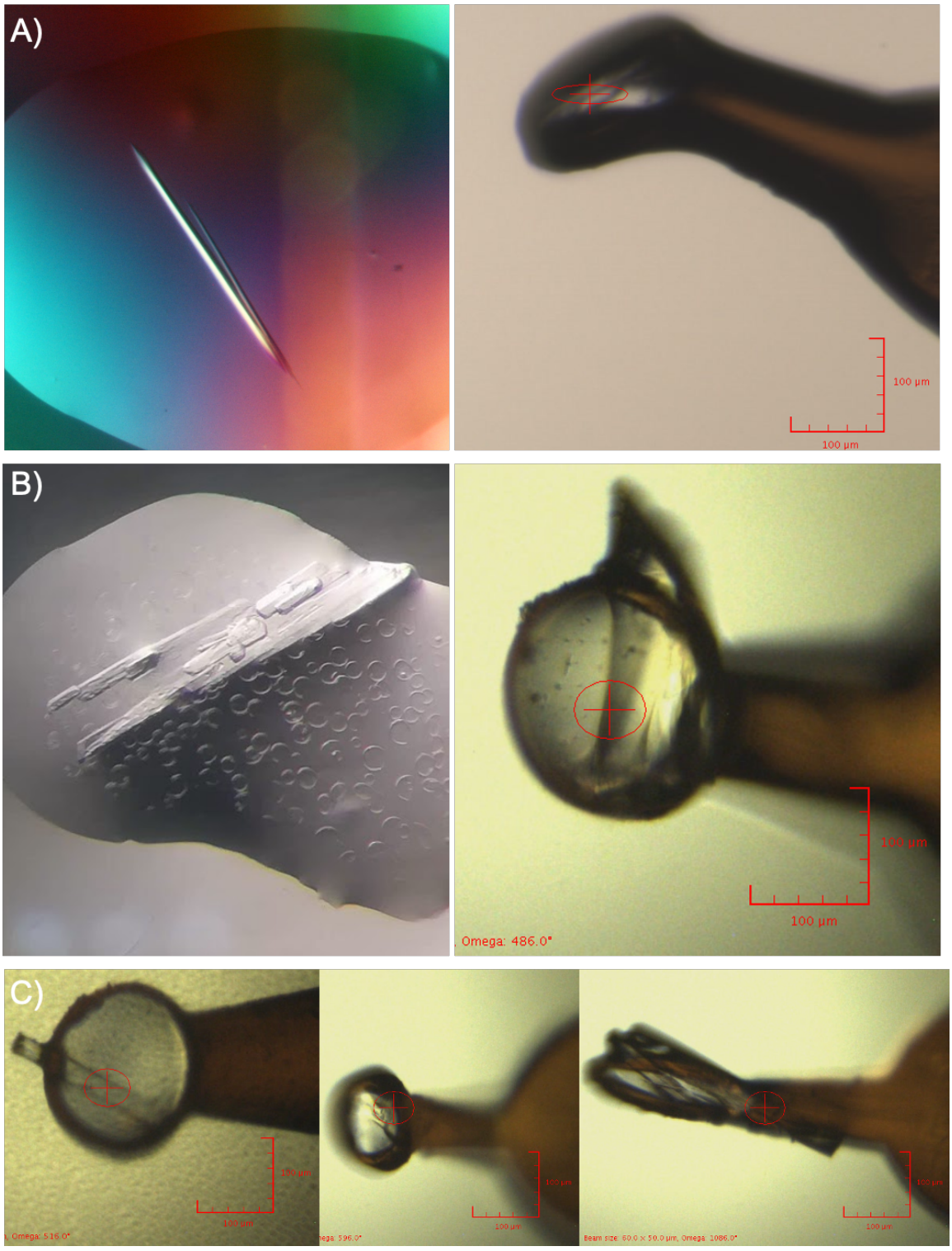


Figure B.17 **Crystals of SmhA.** A) Crystals of soluble SmhA crystal Form 1 (left) and mounted crystal in litholoop (right). B) Crystals of soluble SmhA crystal Form 2 (left) and mounted crystal in litholoop (right). C) Crystals mounted in litholoops of soluble SmhA crystal Form 3 (left), crystal Form 4 (centre), crystal Form 5 (right).

Table B.3 Crystallisation conditions and data collection statistics for different crystal forms of SmhA.

	SmhA crystal Form 3	SmhA crystal Form 4	SmhA crystal Form 5
Beamline	I03	I04-1	I04-1
Wavelength (Å)	0.9792	0.9159	0.9159
Space group	$P4_3$	$P12_11$	P4
Cell parameters			
a (Å) =	151.3	92.18	149.73
b (Å) =	151.3	219.69	149.73
c (Å) =	266.9	162.08	62.68
α (°) =	90	90	90
β (°) =	90	90.52	90
γ (°) =	90	90	90
Molecules per asymmetric unit	16	16	16
Resolution (Å)	3.00 - 57.63 (3.00 - 3.00)	2.96 – 162.1 (2.96 - 3.01)	3.36 – 149.7 (3.36 – 3.42)
Total reflections[a]	932905 (45663)	507088 (22985)	262635 (12628)
Unique reflections[a]	119577 (5886)	133510 (6521)	20195 (1077)
R_{merge}[a,b]	0.292 (1.346)	0.301 (1.314)	0.767 (4.123)
R_{pim}[a,c]	0.111 (0.514)	0.178 (0.811)	0.221 (1.172)
Mean I/**σ**(I)[a]	4.6 (1.3)	3.5 (0.9)	2.9 (0.8)
Completeness (%)[a]	100 (99.8)	99.9 (98.1)	100 (99.8)
Multiplicity[a]	7.8 (7.8)	3.8 (3.5)	13 (11.7)

[a]Values in brackets are for data in the high-resolution shell
[b]$R_{merg} = \Sigma_{hkl} \Sigma_i | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$
[c]$R_{pim} = \Sigma_{hkl} \sqrt{1/n - 1} \Sigma_{i=1} | I_i - I_m | / \Sigma_{hkl} \Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively

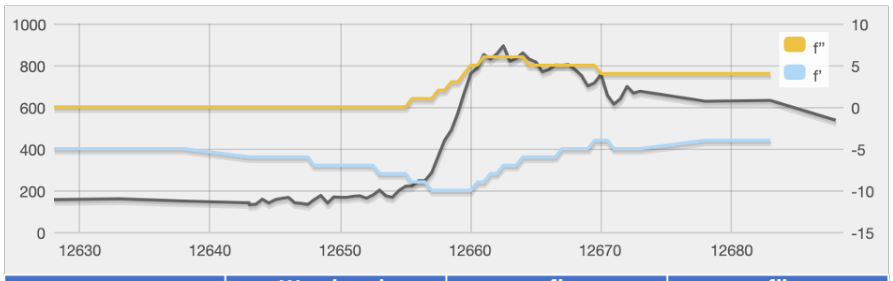
exposure of 0.008 s, and transmission of 20 %. An example diffraction image is provided in (Figure B.18), showing diffraction out to 2.6 Å. Cell parameters (a=151.36 Å, b=151.36 Å, c=133.46 Å, α=90 ^{o}, β=90 ^{o}, γ=90 ^{o}), suggested a solvent content of 48% would give an asymmetric unit content of 8 molecules, as determined by the Matthews calculation (Matthews, 1968)(Figure B.19). As SmhA contains 9 methionine residues in its sequence these values indicate that 72 Selenomethionines were present in the asymmetric unit, and therefore 72 Se atoms were searched for in the initial substructure determination.

A dataset of 2000 images was collected for crystal Form 2 as described in section 2.5.7 (full data processing statistics given in Section 5.3, Paper 3, Table1), with an oscillation per image of 0.1 ^{o} (total 200 ^{o}), exposure of 0.1 s, and transmission of 100 %. An example diffraction image is provided in (Figure B.20), showing diffraction out to 2.8 Å. Cell parameters suggested 4 molecules in the asymmetric unit with a solvent content of 53 %, as determined by the Matthews calculation (Matthews, 1968)(Figure B.21).

B.6.3 Structure determination, model building and validation of SmhA

The SmhA crystal Form 1 Selenomethionine SAD data set had an anomalous mid-slope of 1.045 as shown in Section 5.3, Paper 3, Table1. The CRANK2 pipeline (Skubák and Pannu, 2013) was used for phasing and initial model building, a brief outline of this pipeline is given. SHELXC showed that the anomalous signal was present (>0.8 d"/sig) out to 3.19 Å. SHELXD found 92 Selenium atom sites, within the range expected for 8 molecules of SmhA. Phase calculation, density modification and heavy atom refinement using REFMAC5 and PEAKMAX reduced this to 79 Selenium atom sites (Figure B.22A). A solvent content of 48 %, was used to calculate phases and run density modifications. Inspection of the sites showed strong density for 72 of the 79 sites, accounting for all of the expected Selenium atoms. The best density map was produced for the original hand (density modification FOM = 0.56 and phasing CLD = 2.93, vs FOM=0.55 and CLD = 1.4 for inverted hand),

A)



	Wavelength	f'	f''
Peak	0.9792 Å	-8.06e	6.18
Inflection	0.9795 Å	-10.66e	3.33

B)

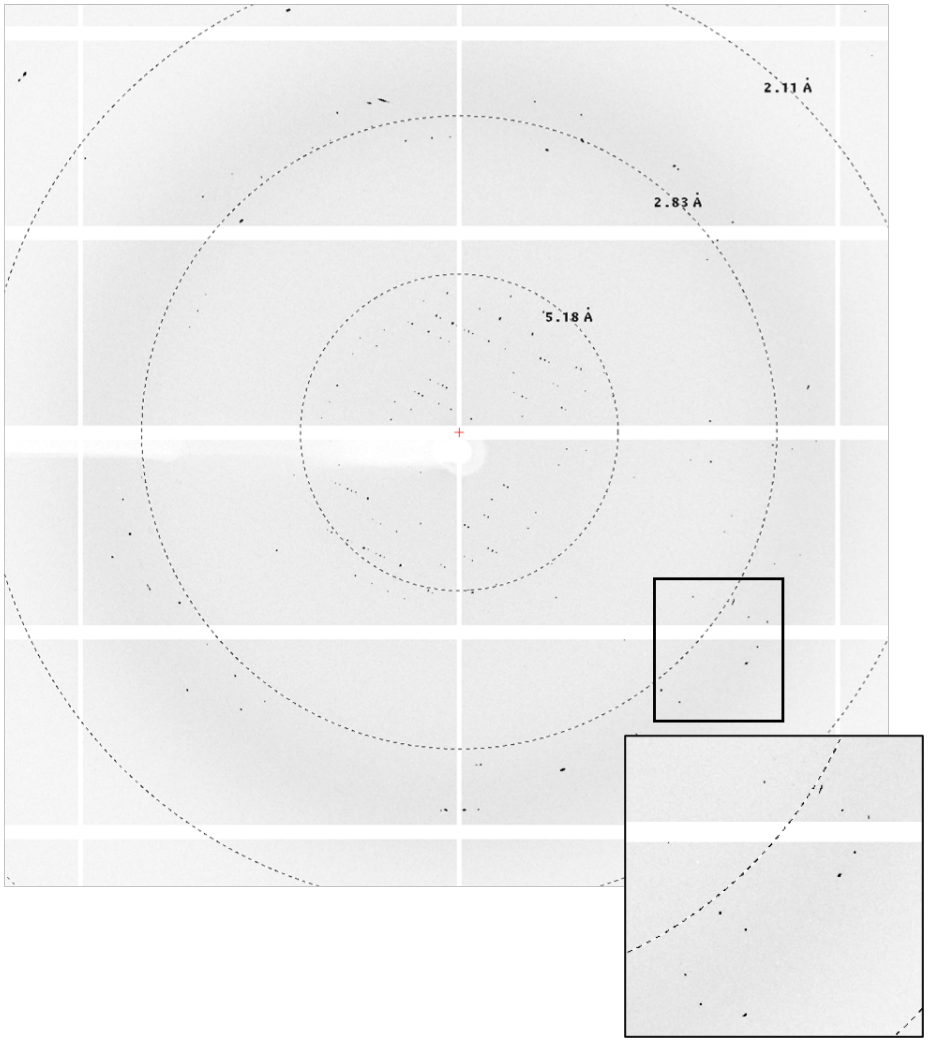


Figure B.18 **SmhA SeMet SAD data collection.** A) Fluorescence scan of SmhA SeMet crystal at Se K edge carried out with 1 s exposure and 3.2 % transmission. The black line shows fluorescence at each wavelength, the blue line plots f' and yellow f''. (B) Diffraction image from data collected on a single crystal of SmhA Semet crystal Form 2 at phi=0.4o with an oscillation of 0.1 ^{o} from SeMet A11 AhlB crystal. Magnified image shows diffraction to 2.6 Å

A)

N(mol)	Prob(N) for resolution	Prob(N) overall	Vm A**3/Da	Vs % solvent	Mw Da
1	0.0000	0.0000	19.09	93.56	40000.00
2	0.0004	0.0011	9.54	87.11	80000.00
3	0.0045	0.0077	6.36	80.67	120000.00
4	0.0207	0.0264	4.77	74.22	160000.00
5	0.0551	0.0650	3.82	67.78	200000.00
6	0.1230	0.1286	3.18	61.33	240000.00
7	0.2270	0.2269	2.73	54.89	280000.00
8	0.2841	0.2734	2.39	48.44	320000.00
9	0.2121	0.2014	2.12	42.00	360000.00
10	0.0606	0.0575	1.91	35.55	400000.00
11	0.0106	0.0101	1.74	29.11	440000.00
12	0.0020	0.0019	1.59	22.67	480000.00
13	0.0000	0.0000	1.47	16.22	520000.00

B)

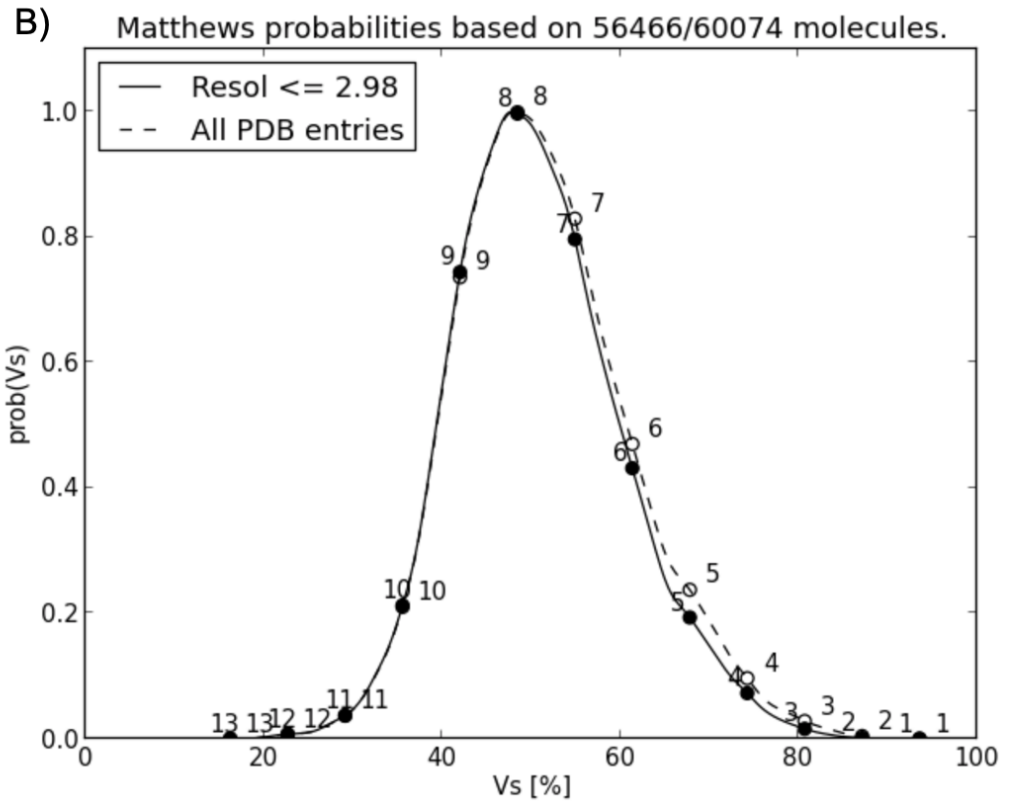


Figure B.19 **SmhA crystal Form 1 Matthews calculation.** A) Table of the possible number of molecules in the AU, with respective probability and percentage solvent content, suggesting 8 molecules in the AU. B) A plot of the probability against the solvent content. Graph and table generated using online software at http://www.ruppweb.org/mattprob/. (Matthews, 1968)

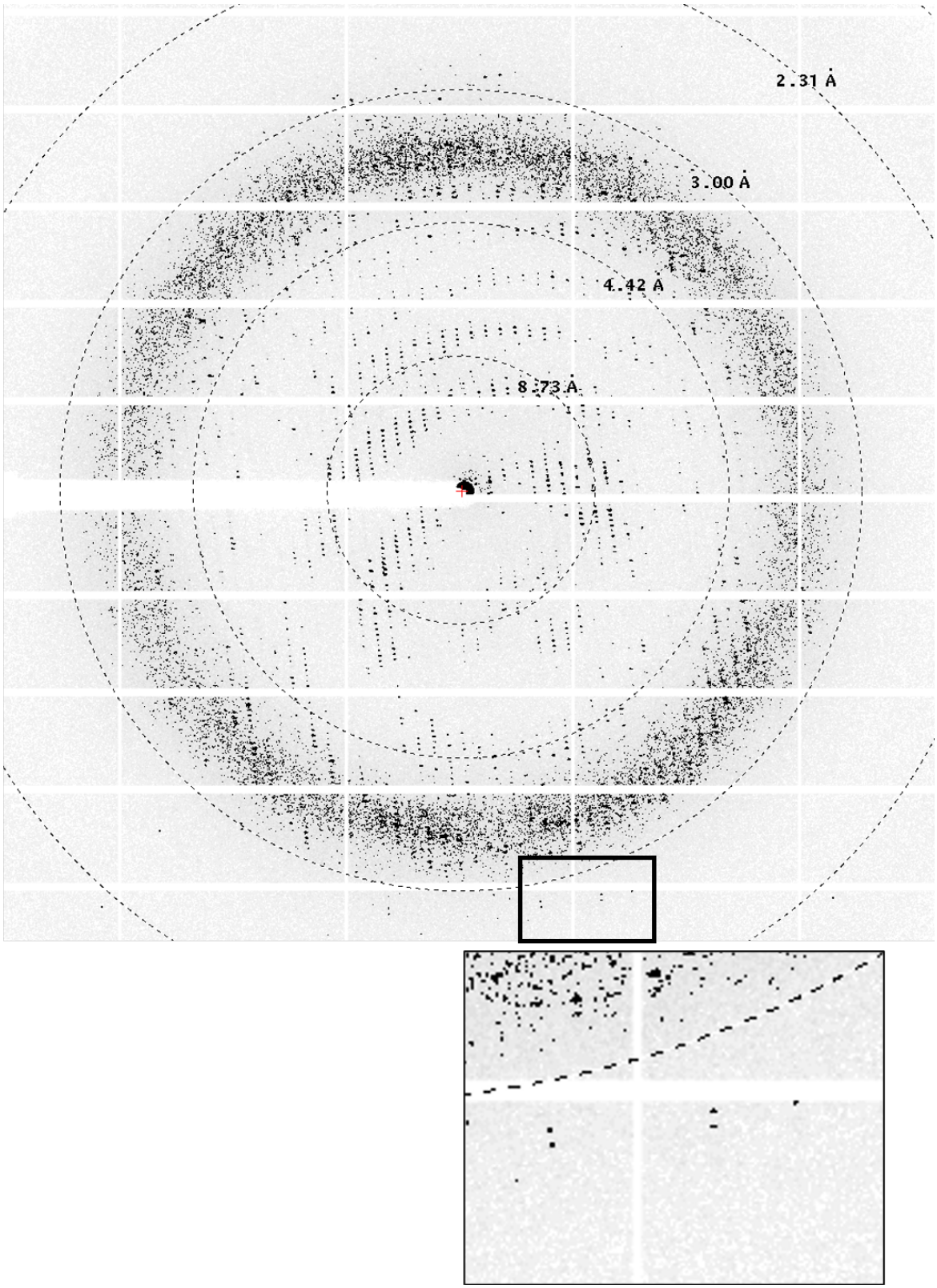


Figure B.20 **SmhA crystal Form 1 data collection.** A representative 0.1 ^{o} oscillation image from SmhA crystal Form 1. An enlarged view of the region highlighted by the square shows diffraction extends to around 2.8 Å.

N(mol)	Prob(N) for resolution	Prob(N) overall	Vm A**3/Da	Vs % solvent	Mw Da
1	0.0001	0.0009	10.39	88.17	40000.00
2	0.0156	0.0309	5.20	76.33	80000.00
3	0.1225	0.1609	3.46	64.50	120000.00
4	0.4739	0.4680	2.60	52.66	160000.00
5	0.3671	0.3212	2.08	40.83	200000.00
6	0.0207	0.0182	1.73	28.99	240000.00
7	0.0000	0.0000	1.48	17.16	280000.00

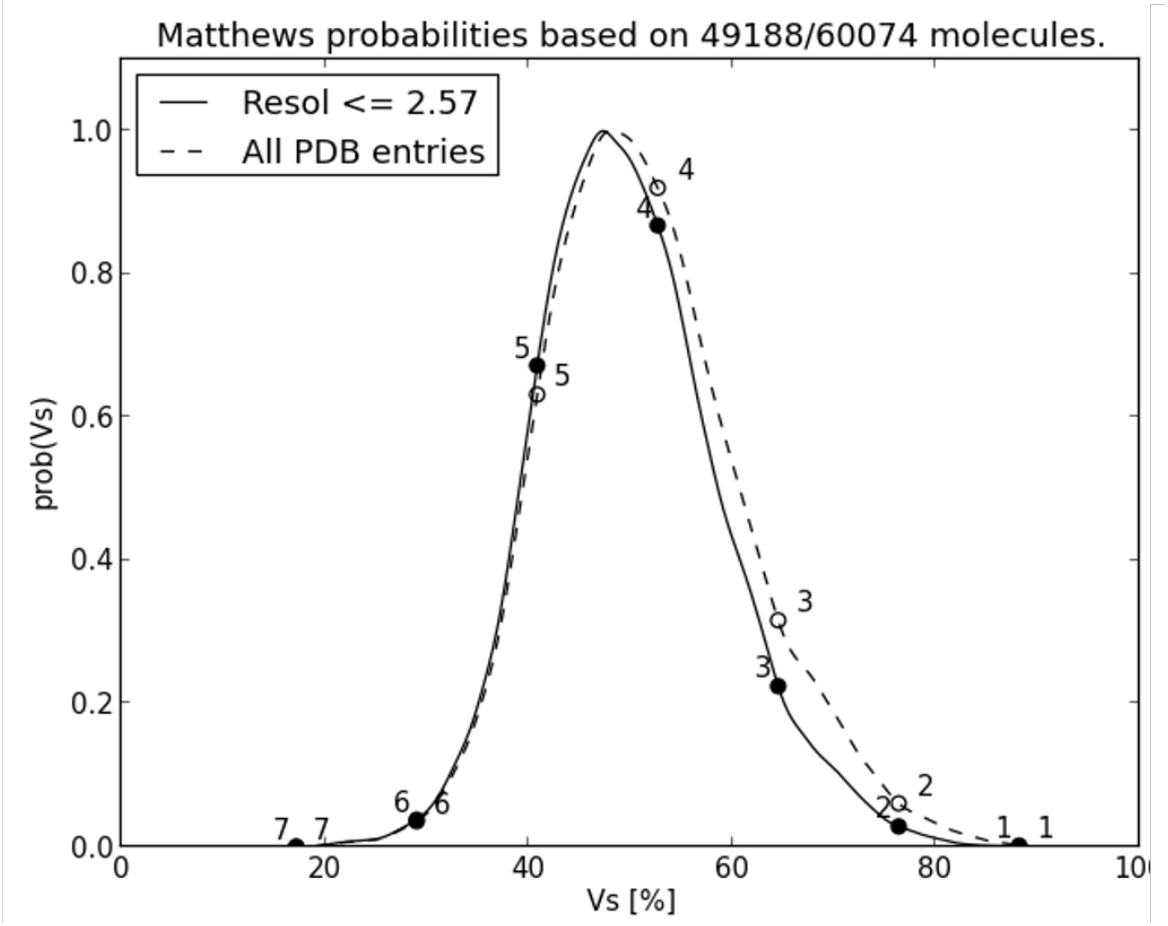


Figure B.21 **SmhA crystal Form 2 Matthews calculation.** A) Table of the possible number of molecules in the AU, with respective probability and percentage solvent content, suggesting 4 molecules in the AU. B) A plot of the probability against the solvent content. Graph and table generated using online software at http://www.ruppweb.org/mattprob/. (Matthews, 1968)

and initial phases were enough to see continuous density of helices within the protein (Figure B.22B). Finally, the CRANK2 pipeline ran Parrot, REFMAC5, and Buccaneer for model building, density modification and phased refinement to produce a model with 8 chains in the AU (Figure B.22C), and gave an R-factor of 0.29 and Free R of 0.35. Subsequent rounds of manual building and refinement using Coot (Emsley et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.21 and Free R of 0.25, with 8 monomeric chains in the AU (Figure B.23A). This is in contrast to gel filtration analysis in Paper 3 (Section 5.3) which gave a molecular weight of 80 kDa equivalent to a dimer of SmhA. As the molecular weight standard used to calibrate the gel filtration column is composed of globular proteins, the estimated molecular weight of SmhA from gel filtration may be inaccurate due to the elongated structure of SmhA, resulting in the discrepancy seen between the crystal structure and gel filtration. Full validation details can be found in Section 5.3, Paper 3, Table 1. Residues 206-228 in chain A, 208-228 in chain B, G and F, 210-227 in chain C, 209-228 in chain E and H, which form part of the loop in the β-tongue (discussed in Paper 3 (Section 5.3)) were omitted from the final model due to poor density. Chain D, however, had poor density throughout the head domain, which unlike the other 7 chains, is not involved in crystal packing interactions (Figure B.23B), and as such residues 40-54, 164-165, 191-198 and 208-228 were omitted from this chain D (Figure B.24). This hints towards the flexibility of the head domain, which, as discussed in Paper 3, must undergo a major conformational change in the soluble to pore transition.

Chain A from crystal Form 1 was used as a molecular replacement model for crystal Form 2. Initial phases for crystal Form 2 were calculated using Phaser MR (McCoy et al., 2007). This gave a single unique solution with a refined TFZ=43.91 and LLG=5397.20, and placed 4 molecules in the AU, as described in Section 2 and Paper 3 (Figure B.25). Refinement of this model using REFMAC5 (Murshudov et al., 1997) gave a R-factor of 0.24 and Free R of 0.3. Subsequent rounds of manual building and refinement using Coot (Emsley

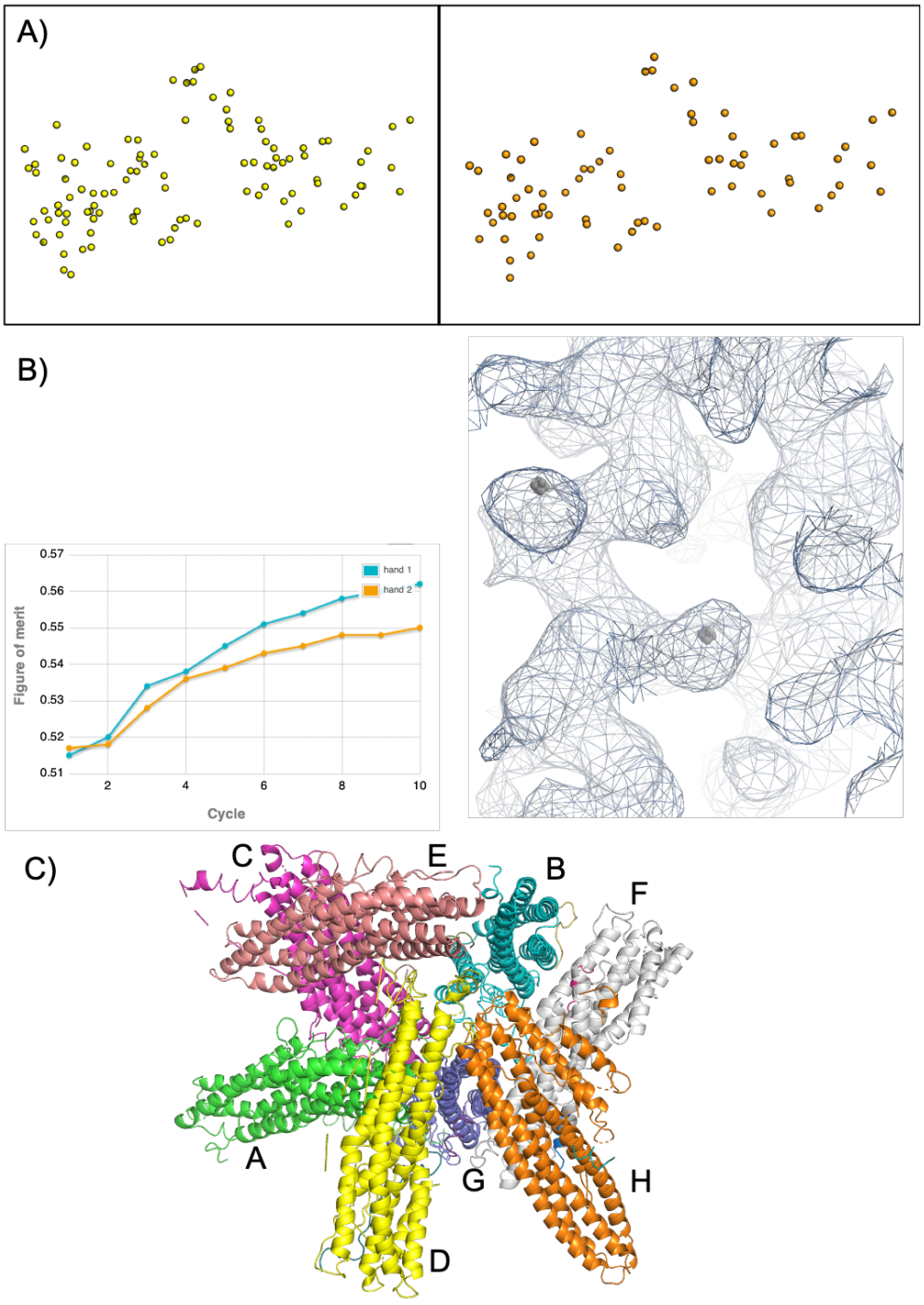


Figure B.22 **Experimental phasing and chain tracing by CRANK2.0** A) Selenium atom co-ordinates. ShelxD placed 92 Se atoms (left), after substructure improvement and phasing by REFMAC5 (Murshudov et al., 1997), PEAKMAX 13 (Skubák and Pannu, 2013) atoms were removed (right). B) Left, Graph of FOM vs density modification cycle for original (hand 1) and inverted hand (hand 2), shows original has the best FOM after 10 cycles. Right, 2Fo-Fc map for the original hand generated from the initial phase determination of SmhA, Se atoms fit density and continuous density for helices was visible. C) Model output by CRANK 2 (Skubák and Pannu, 2013). 8 Chains were built in the AU.

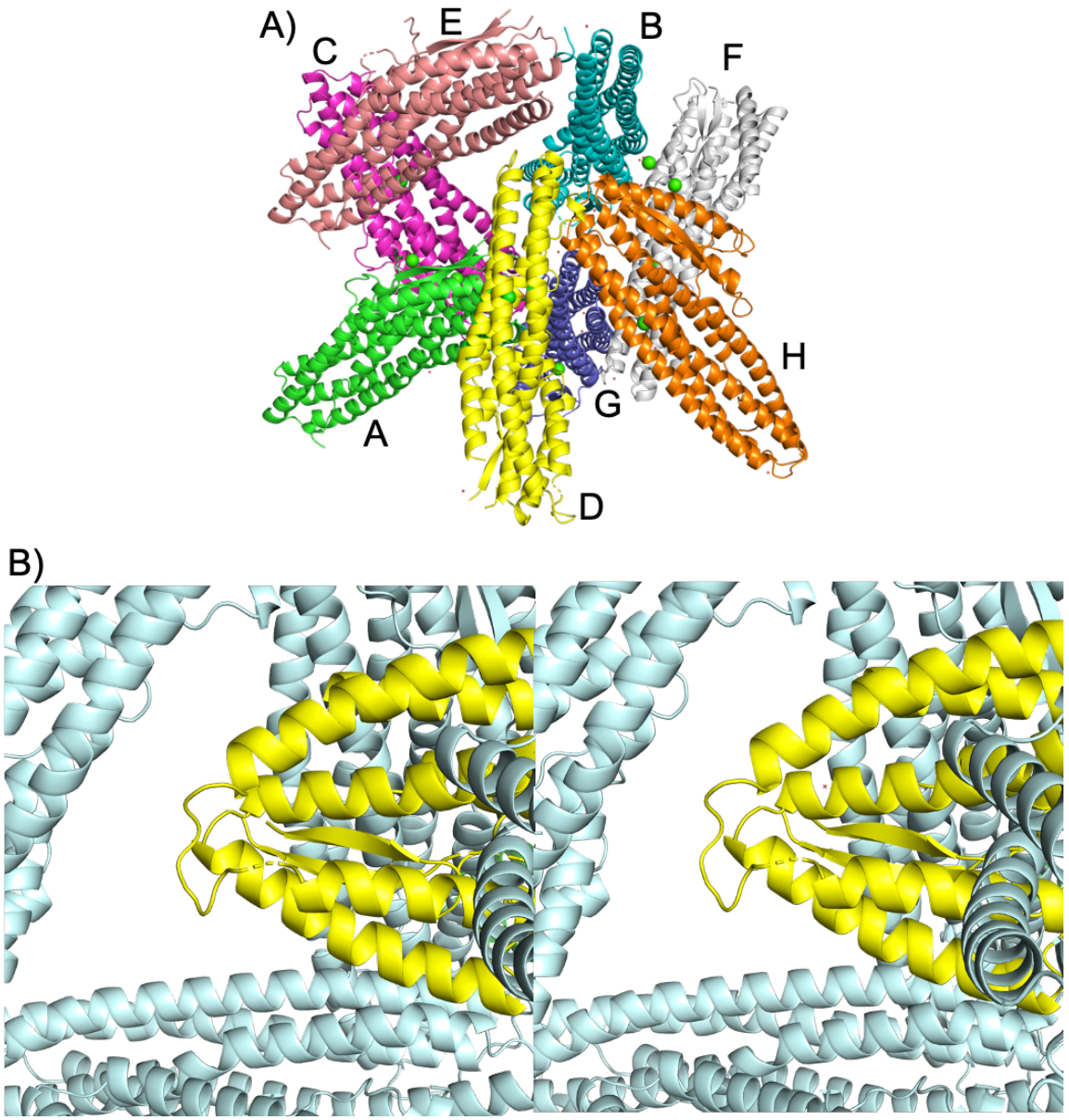


Figure B.23 **Final model of SmhA crystal Form 1.** A) Final validated model of SmhA crystal form 1, 8 chains were successfully built in the AU. B) Stereo image of symmetry-related molecules (cyan) packing around SmhA chain D head (yellow). The head doesn't pack against any other molecules in the crystal and as such is exposed to solvent resulting in poor density for the head domain.

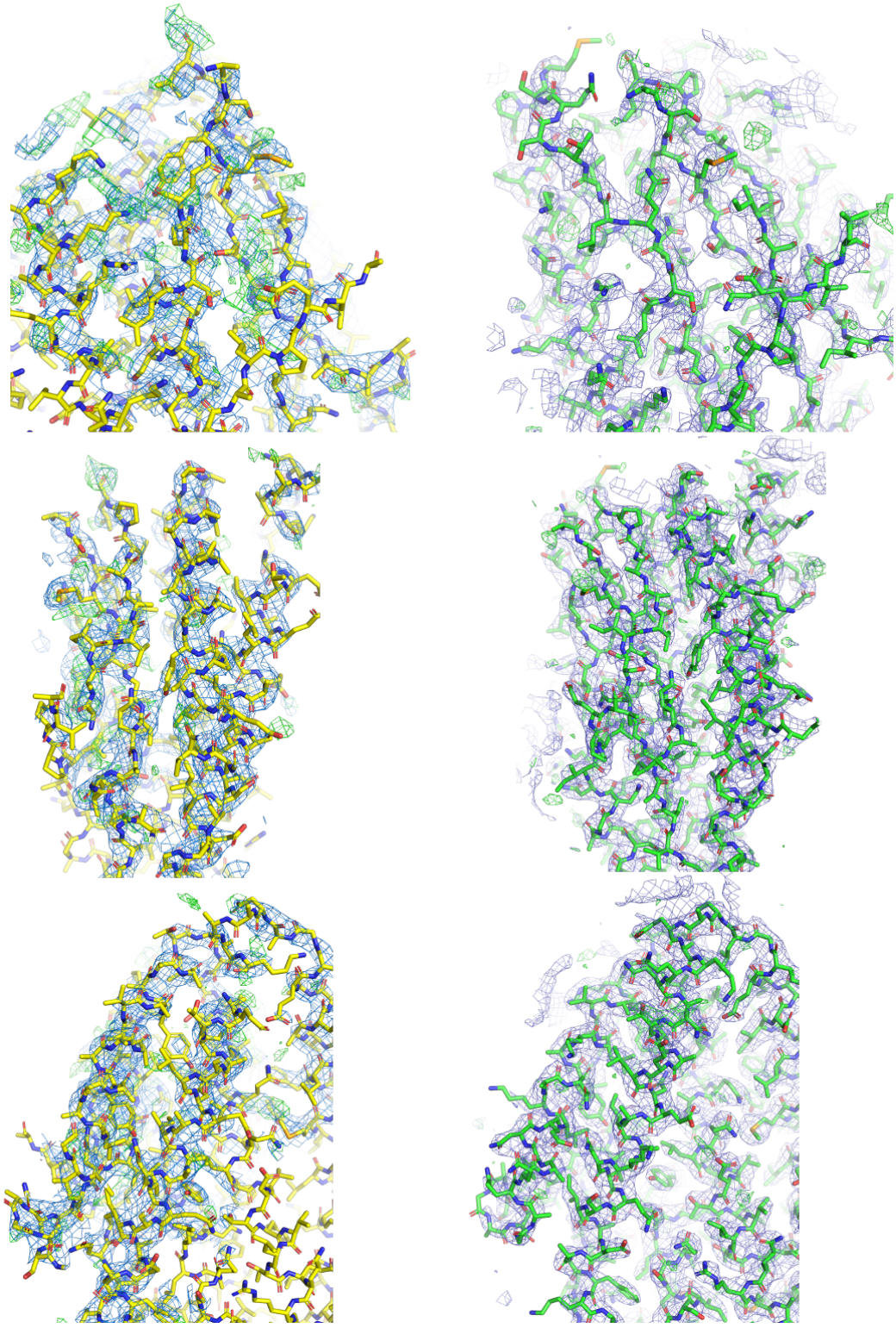


Figure B.24 **Comparison of density around the head domain of Chain D (yellow, left three) and Chain A (green, right three).** Density is significantly worse for chain D with a large amount of undefined density. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and difference map contoured at 2.99 σ (green, positive, and red, negative, mesh).

et al., 2010) and REFMAC5 (Murshudov et al., 1997) produced a final model with R-factor of 0.23 and Free R of 0.25. Full validation details can be found in Paper 3, Table 1. The final model of SmhA crystal Form 2 contained 4 monomeric chains of the soluble conformation of SmhA. All four molecules are almost identical with the only variation seen in the density of the connecting loop between the two β-strands ($\beta 1$ and $\beta 2$) in the β-tongue (Figure B.26). Chain C is most complete with only density for residues 210-225 missing.

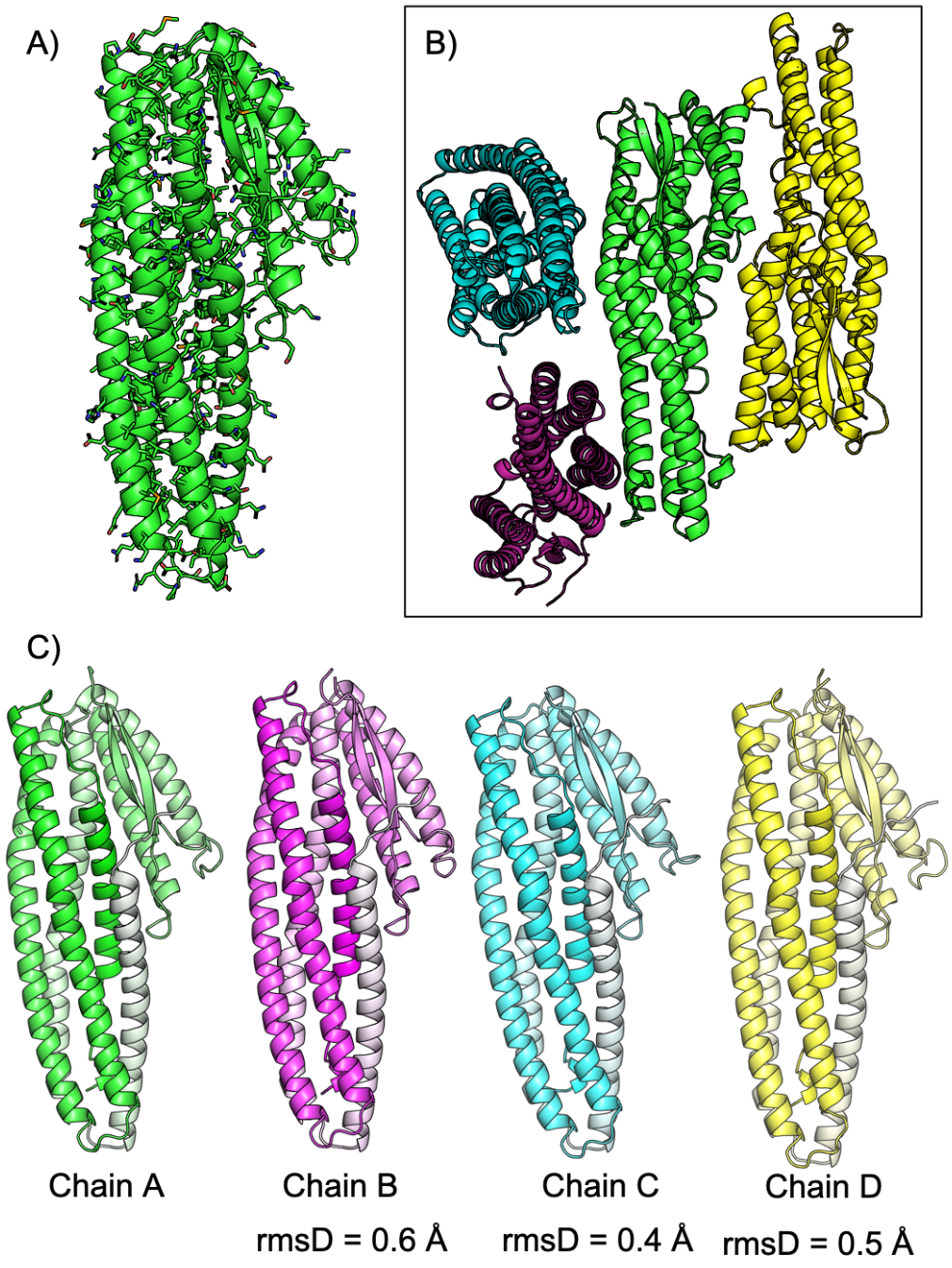
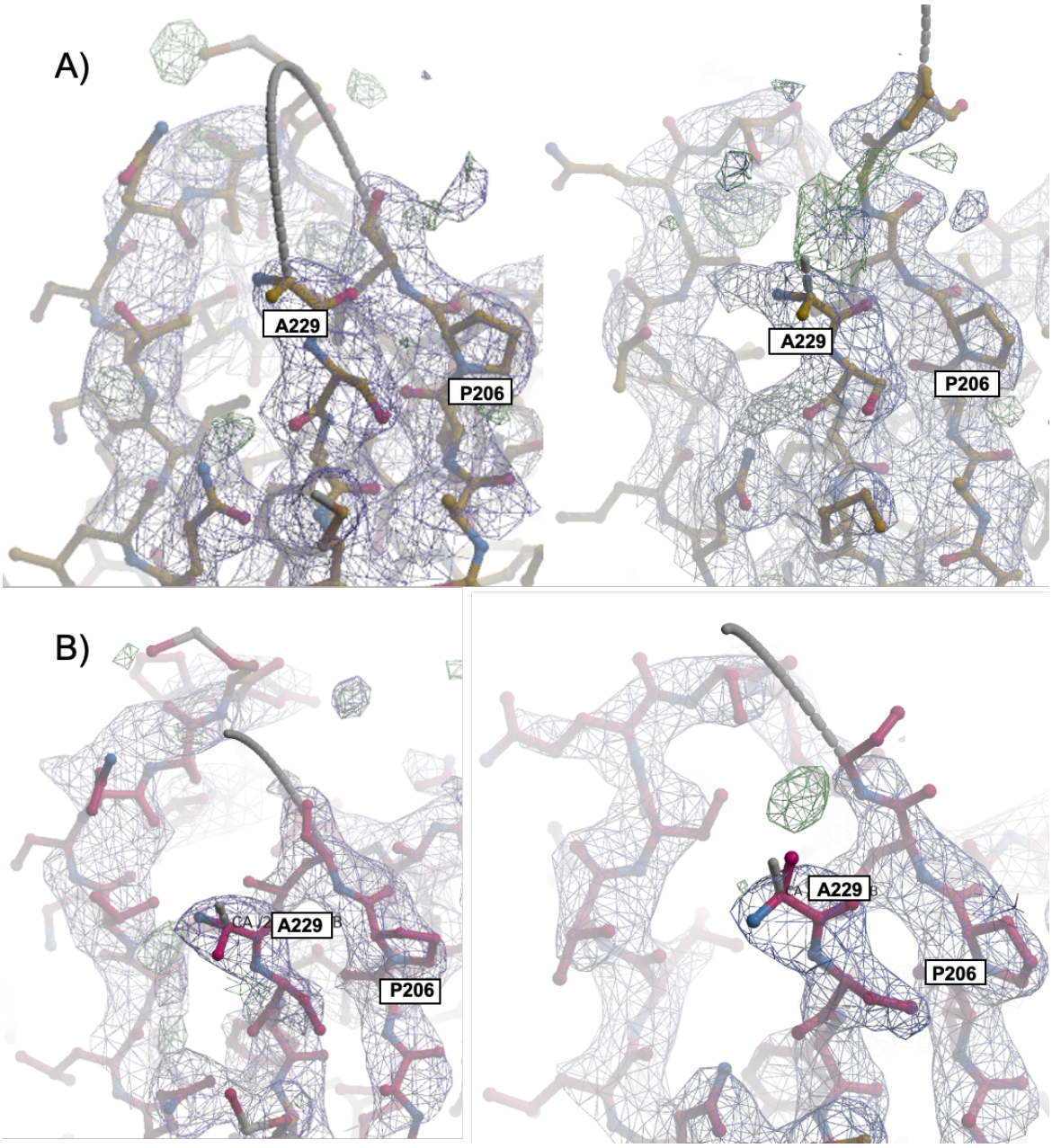


Figure B.25 **Molecular replacement and model building for SmhA crystal Form 2.** A) model of chain A SmhA crystal Form 1 used for molecular replacement. B) PhaserMR (McCoy et al., 2007) placed four molecules in the asymmetric unit. C) Models for all four chains of SmhA crystal Form 2, coloured dark (N-terminus) to light (C-terminus), all four chains are almost identical, rmsD Cα for each chain against chain A is given below.


A)
A229
P206
A229
P206
B)
A229
P206
A229
P206

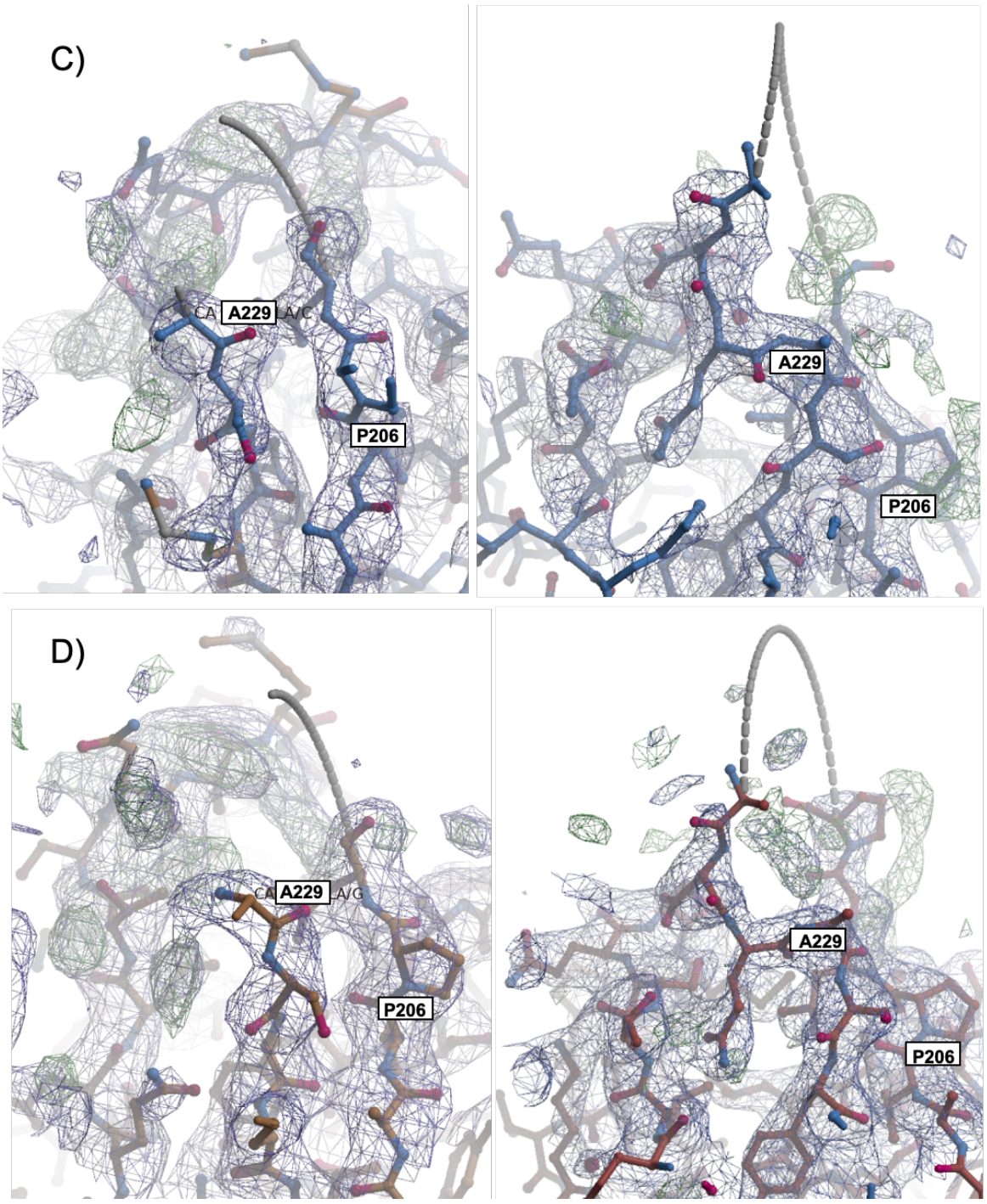


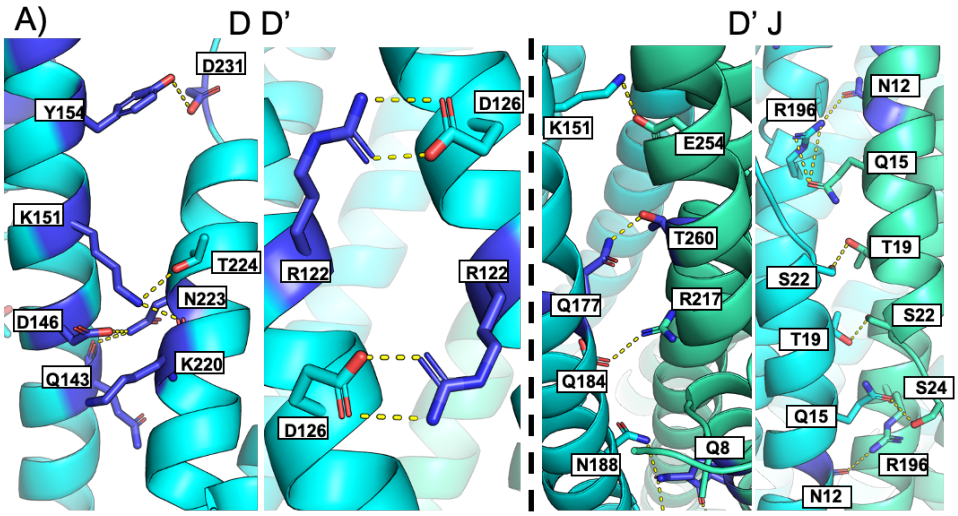
Figure B.26 **Difference in electron density for each chain of SmhA crystal Form 2.** Phaser MR (McCoy et al., 2007) (left) and validated (right) models and electron density around A229 and P206, for all four chains of SmhA crystal form 1. 2Fo-Fc map contoured at 1.0 σ (blue mesh) and positive difference map contoured at 2.99 σ (green mesh). Panels A-D show chain A-D, respectively.

B.7 Bioinformatics analysis of SmhC

Although SmhC was successfully expressed and purified, crystallisation trials proved unsuccessful. Therefore to try and determine if SmhC is likely to share any structural similarities with AhlC, bioinformatics analyses were carried out. A BlastP (Altschul et al., 1990) sequence alignment of SmhC showed that all SmhC homologues shared the same short hydrophobic stretch between α3-α4 (residues 155-173) seen in AhlC (Figure B.27) with sequence identities ranging from 25-76 % (Figure B.27). Further to this Y154 and D231, which have been proposed to be involved in tetramer assembly in AhlC (Section A.5.4), are also present in SmhC and well conserved across the homologues. Gel filtration analysis described in Section 5.3 (Paper 3, supplementary figure 1) shows that SmhC elutes at an approximate molecular weight equivalent to a dimer of SmhC. The AhlC head mutant structure (AhlCHM), which has the most complete tetrameric structure for all AhlC structures, was used for analysis of all residues involved in hydrogen bonding or salt bridge formation at the interface as identified by ePISA (Krissinel and Henrick, 2007). The interface between chains D and D' in the tetramer of AhlCHM (PDB code 6R1J), showed that all but 1 residue (T224 to Ala in SmhC) are conserved in SmhC (Figure B.27 -B.28A). However, analysis of interface residues between chains J and D show that only 4 of 22 residues in AhlC are conserved in SmhC (Figure B.27 - B.28A). It therefore seems possible that SmhC could indeed form a dimer, using an interface similar to that seen in chain D and D' in AhlCHM (Figure B.28B).

Figure B.27 **Sequence analysis of SmhC homologues identified by BlastP(Altschul et al., 1990).** Sequence alignment using Tcoffee \citep{Notredame2000} of homologues of SmhC. Conserved residues are shown in blue, light to dark least to most well conserved. The hydrophobic head is highlighted in green, red shows conserved residues Y154, and D231. yellow highlights residues involved in interface 1 (DD'), pink highlights residues involved in interface 2 (D'D)

Figure B.27



B)

ePISA results	Interface 1 (D D')	Interface 2 (D' J)
Interface surface area (Å²)	1500	1700
Interface residues (%)	18	19
Delta G	-16	-9.1
Binding energy	-21.7	-18.6

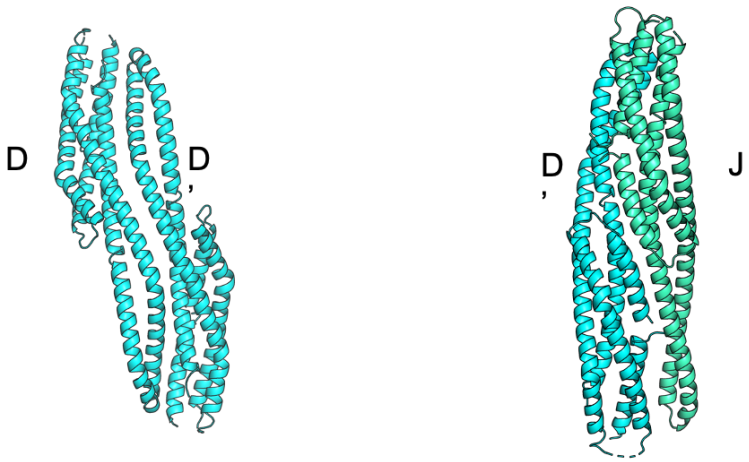
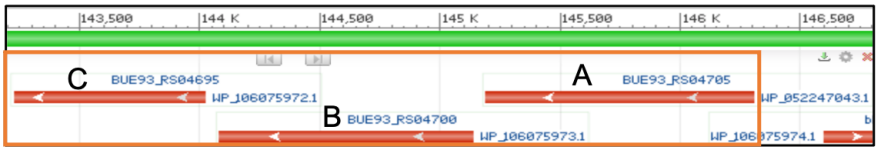


Figure B.28 **Location of interfaces and conserved residues in the C tetramer.** A)AhlCHM (6R1J) DD' And D'J interfaces of the AhlCHM, with residues identical between AhlC and SmhC highlighted in blue, bonds are shown as yellow dashes, with chain D and D' shown in cyan and chain J in green. Many more identical residues are found in the DD' interface than the D'J interface. B)ePISA (Krissinel and Henrick, 2007) analysis for each interface in the tetramer, with dimers of DD' and D'J shown below.

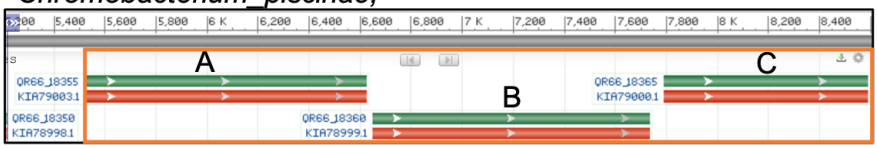
B.8 Homology searches and sequence alignments with SmhA and SmhB

Homologues of both SmhA and SmhB were identified using BlastP (Altschul et al., 1990) using a cut off of E<0.01. All homologues identified were used in sequence alignments and analysis described in Paper 3 (Section 5.3). To determine if these SmhA and SmhB homologues were members of a tripartite α-PFT, regions of the host genome were analysed upstream and downstream of the homologous gene (Figure B.29). All SmhA homologues seemed to be located directly upstream of a further two genes of the appropriate length expected for a B and C protein, and therefore part of a tripartite α-PFT. In many cases, an SmhB homologue was located directly downstream of a SmhA homologue (Figure B.29). In the SmhB sequence alignments, four SmhB homologues were not obviously part of a tripartite toxin, instead, being an orphan protein (*Photobacterium profundum*), a bipartite toxin (*Paraburkholderia*), or part of a four protein operon (*Aquimarina megaterium, Tetrasaphera japonica*), where the fourth protein could perhaps act as regulator or chaperone. This observation is consistent with observations that SmhB and AhlB are both able to form pores in isolation, and therefore B homologues resulting from gene duplication events are not discarded. SmhA, however requires the presence of the other two proteins in the tripartite toxin for activity, and therefore homologues of SmhA are only found as part of a tripartite α-PFT, as any gene duplication event will ultimately be lost due to inactivity of the orphan gene. This could prove a useful tool in the identification of further tripartite α-PFT toxins, as results from Blast searches with the A component will always show proteins in tripartite α-PFT, while searches with the B component might not.

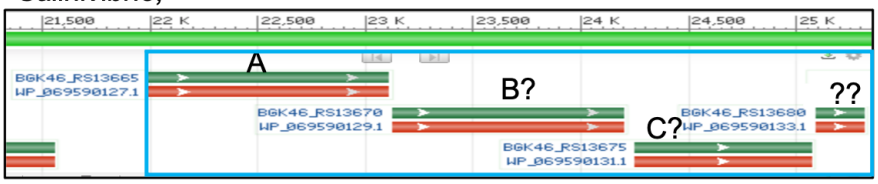
Chromobacterium_sp,



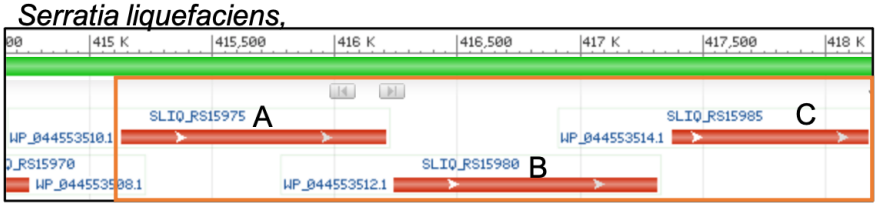
Chromobacterium_piscinae,



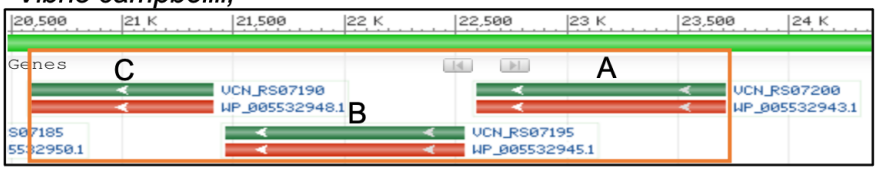
Salinivibrio,



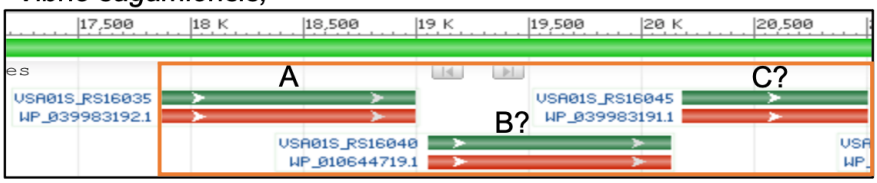
Serratia liquefaciens,



Vibrio campbellii,



Vibrio sagamiensis,



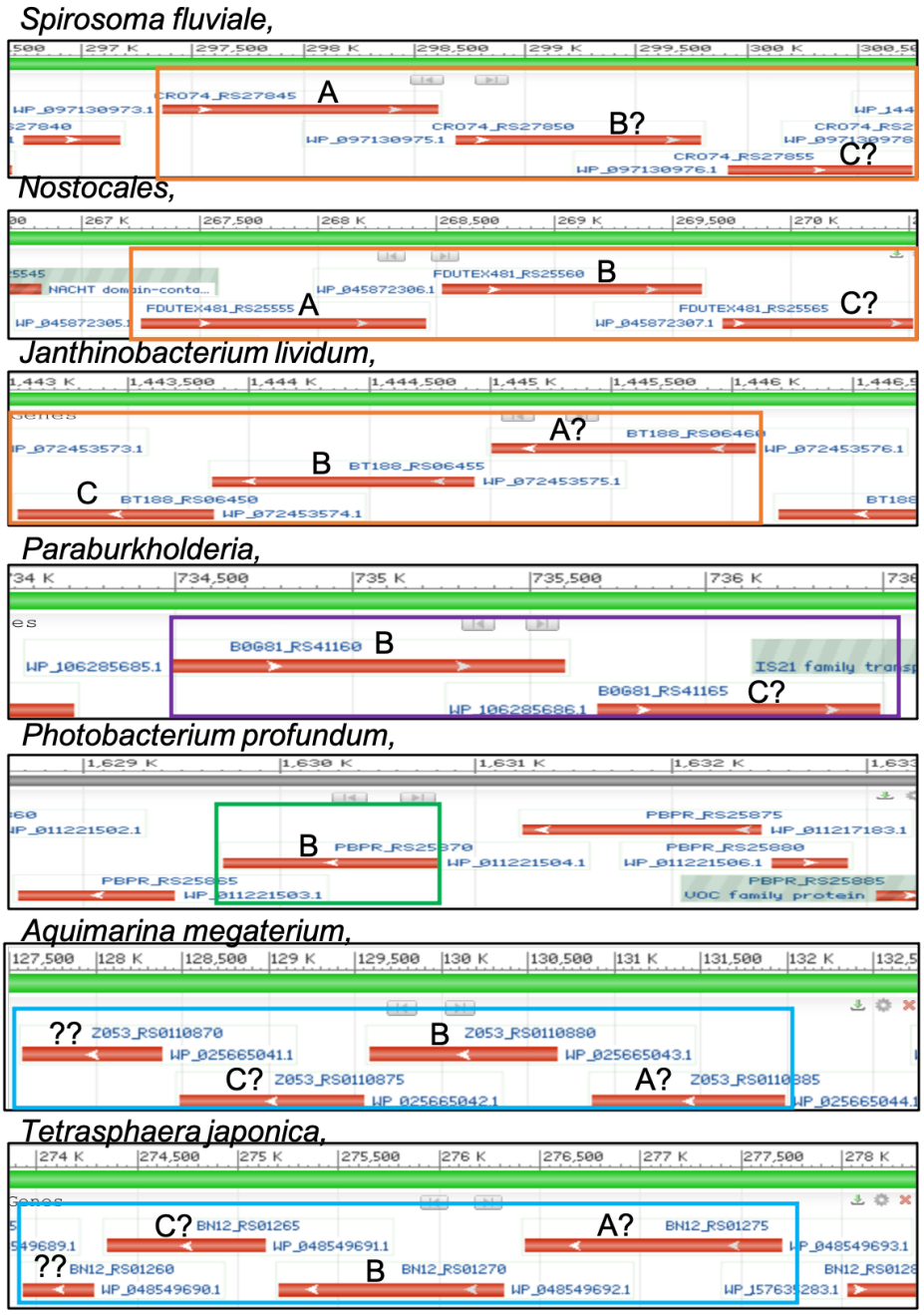


Figure B.29 **Genomic analysis of regions upstream and downstream of SmhA and SmhB homologues identified by BlastP (Altschul et al., 1990).** The homologous gene identified by BlastP is labelled A, B or C (relating to SmhA, SmhB, and SmhC homologues respectively). Possible, but not identified by BlastP, homologues are labelled with a ?. The orange box highlights tripartite operons within the genome, blue, operons with four; purple, two; and green, one gene(s), respectively.

B.9 Electron microscopy of Smh proteoliposomes

As SmhA+SmhB+SmhC and SmhB+SmhC are both able to form pores in erythrocytes (Section 5.3, Paper 3), these components were also incubated with liposomes to see if pore formation could be induced using artificial membranes, with the ultimate aim of using the resultant proteoliposomes for structural studies of pores by Cryo-electron microscopy.

Proteoliposomes were prepared as described in Section 2 by incubating 10 μg of each Smh component with 20 μg of liposomes in 10 μl 10 mM PBS pH 7.4. Samples of these proteoliposomes were then immediately loaded on to carbon grids and stained with uranyl-formate. Negative stain TEM micrographs were collected as described in Section 2.6.1. SmhB+SmhC formed long filaments saturated with pores, while ShmA+SmhB+SmhC formed smaller 'Viking ship' structures like those observed for Ahl pores (Section 3.3(Paper1)) (Figure B.30). These results clearly confirmed that liposomes induced pore formation, and that they could produce possible helical arrangements in membrane filaments, which could be used for helical reconstruction of the individual pores. These samples were therefore taken forward for optimisation.

To produce a cleaner sample and remove unbound soluble protein from the proteoliposome, after incubation samples were loaded on to a sucrose gradient before ultracentrifugation at 200,000 g as described in Section 2.4.4. Proteoliposomes floated to the top of the gradient while soluble protein remained at the bottom. Top fractions containing proteoliposomes were collected and loaded on to carbon grids as described in Section 2.6.1. TEM negative stain micrographs showed much cleaner grids containing only the pore filled filaments for both SmhB+SmhC and ShmA+SmhB+SmhC (Figure B.31). Images were collected at varying defoci for 3D reconstruction. Initial analysis of these filaments showed the holes in the centre to be approximately 130 Å $\pm$ 20 Å wide and spacing between notches on the outside of the filament to be 60 Å $\pm$ 20 Å, with the filament itself having a diameter of 360 Å $\pm$ 20 Å. These measurements were the same for both samples and as so only images for SmhB+SmhC

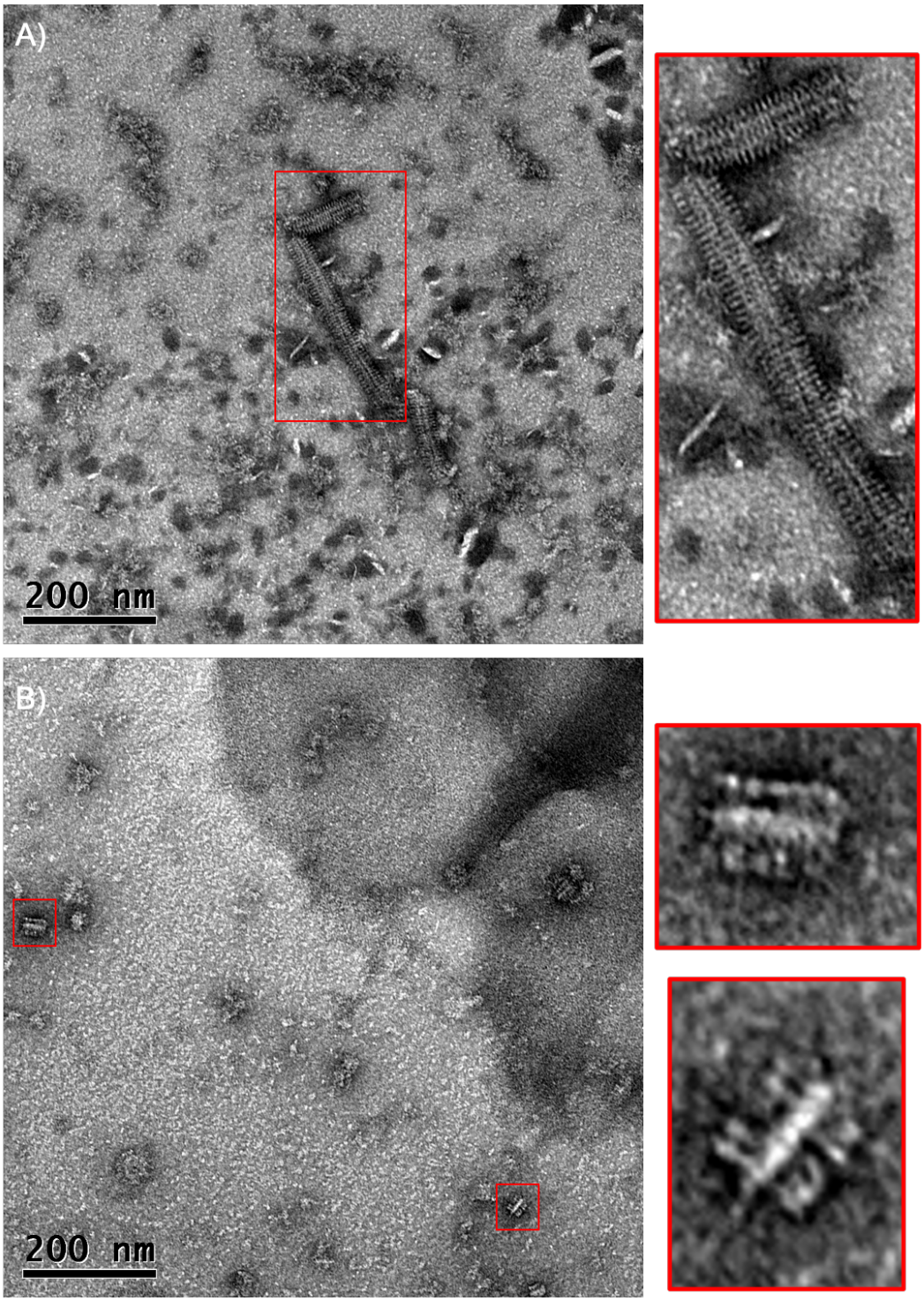


Figure B.30 **TEM negative stain EM images of SmhBC and SmhABC showing pores when incubated with liposomes.** Red boxes highlight regions magnified in images to the right. A) SmhBC incubated with liposomes for 1hr at 37oC. Long filament structures form lined with pores. B) SmhABC, small Viking ship structures form showing side views of pores.

were taken forward for 3D reconstruction. Negative stain micrographs (7.18 Å/pix) were imported to Relion 3.1 (He and Scheres, 2017; Scheres, 2012). Filaments (360 Å in diameter with a 70 Å helical rise) were manually picked and extracted using an 80 pixel box size, resulting in 1213 particles. A soft-edge cylindrical starting model, in a corresponding box size, was generated using the Relion Helix Toolbox (He and Scheres, 2017) with an outer diameter of 360 Å. 25 iterations of 3D classification was carried out with the cylindrical starting model, imposing no symmetry. A 400 Å mask was applied, with angular search ranges of 10 ^{o} in psi and 15 ^{o} in tilt and an angular sampling interval of 7.5 ^{o}. The resulting 3D classes were visually inspected in Chimera (Pettersen et al., 2004). This showed holes in 'pores' (distance between nodes) measured between 96 Å and 152 Å, with a distance between vertical repeats of between 57 and 65 Å. There was a slight tilt between repeats showing a helical nature to the pore arrangement in the filament, making it a good candidate for Cryo-EM and helical reconstructions of the pore (Figure B.32). The size of the holes in the 'pores' observed in these filaments suggests that two pore species are present, with the smaller holes (96 Å) comparable to the SmhBC particles (64 Å $\pm$ 20 Å) in Section B.4 and AhlB and SmhB pores (Paper 1 (Section 3.3) and Paper 3 (Section 5.3)), and the larger holes comparable to the SmhABC soluble complexes (170 Å $\pm$ 20 Å) and AhlABC pores discussed in Paper 3 and Paper 1, respectively.

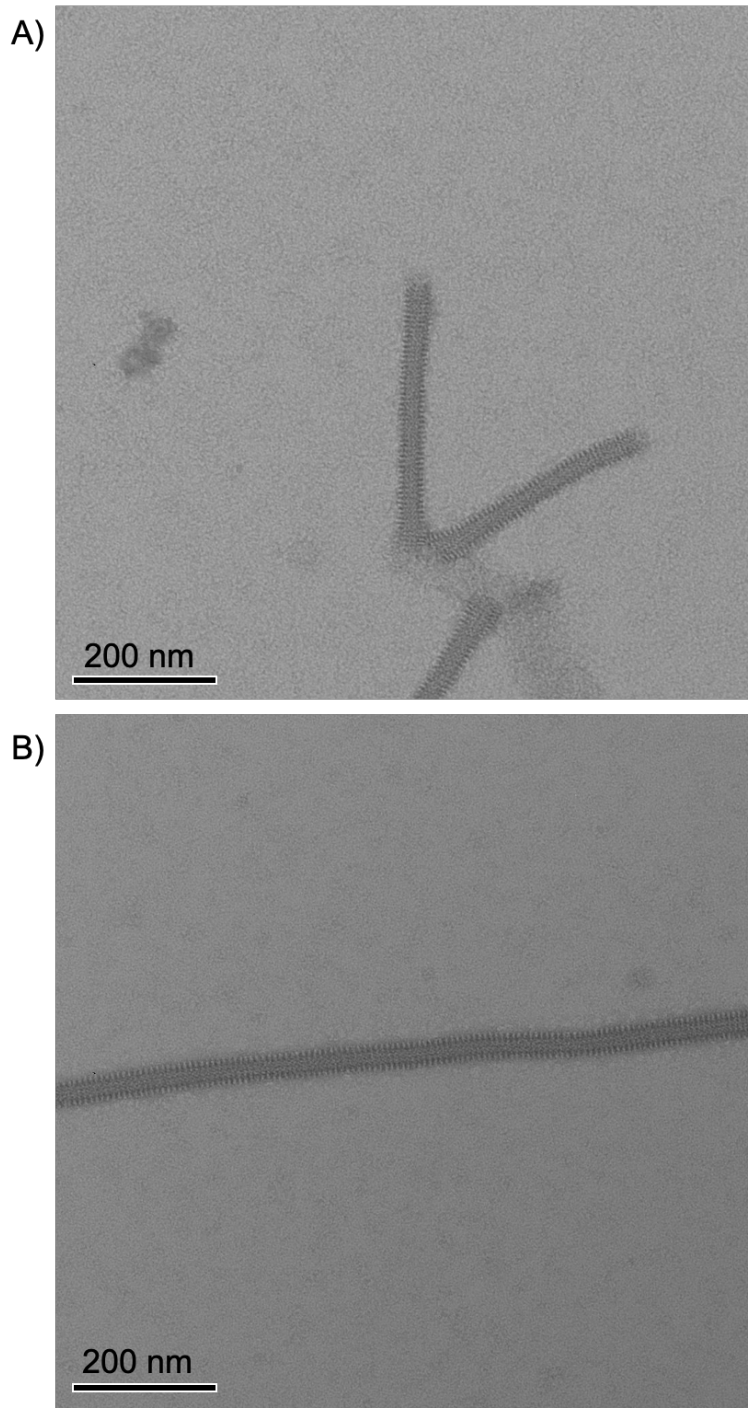


Figure B.31 **TEM negative stain EM images of SmhBC and SmhABC showing pores when incubated with liposomes.** Red boxes highlight regions magnified in images to the right. A) SmhBC incubated with liposomes for 1hr at 37 ^{o}C. Long filament structures form lined with pores. B) SmhABC, small Viking ship structures form showing side views of pores.

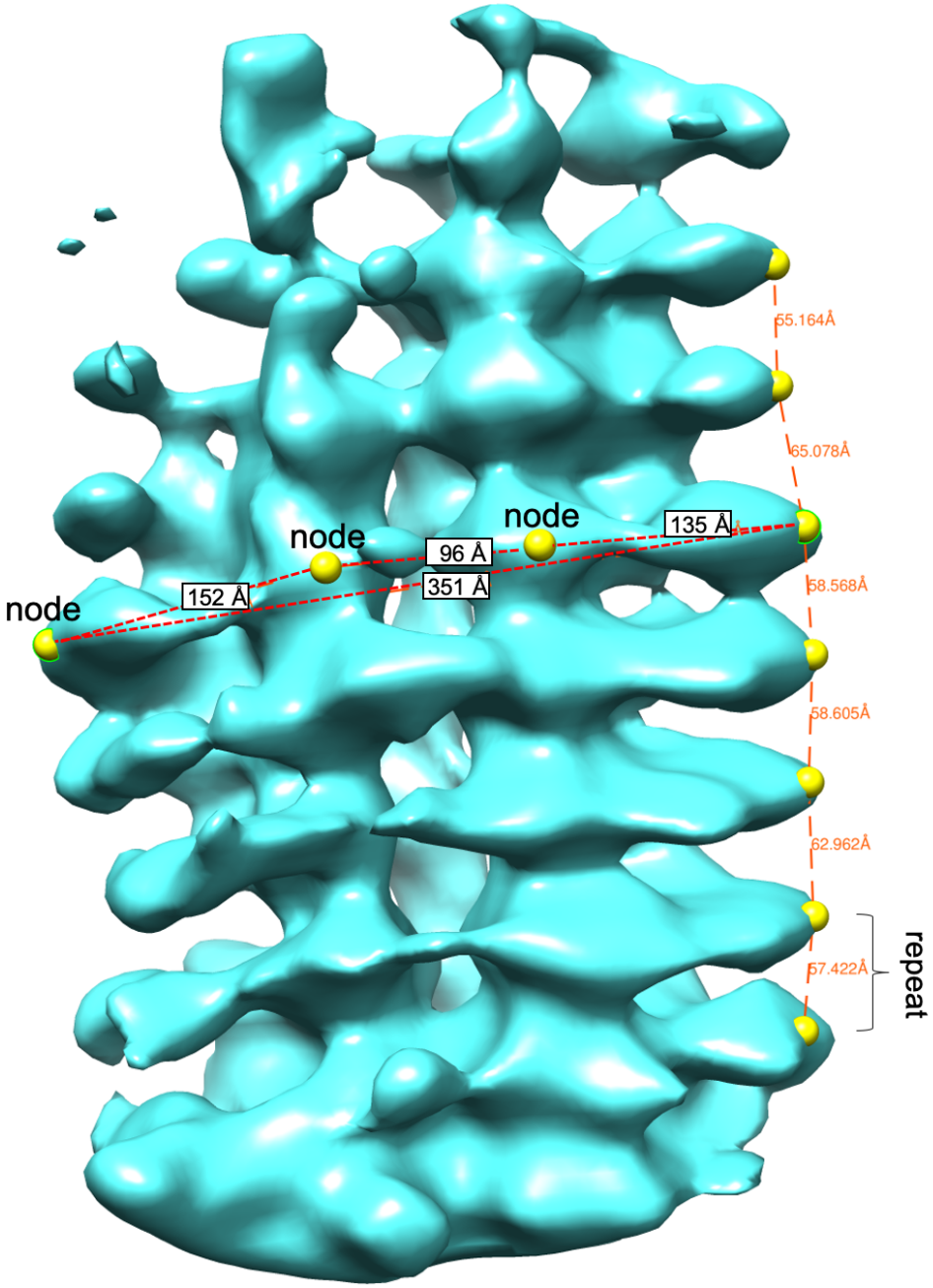


Figure B.32 **3D classification of SmhBC filament.** Distances are shown between vertical repeats and horizontal nodes (yellow). A slight slant is visible between repeats suggesting a possible helical arrangement of 'pores' in the filament.

Chapter 6

Conclusions and future work

6.1 Discussion

This thesis had the aim of expanding knowledge of the ClyA family tripartite α-PFTs through structural and biochemical studies. As shown in the preceding papers the structural determination of the soluble A, B and C components (SmhA, SmhB, AhlB, and AhlC), as well as the first pore structures (AhlB and SmhB) from Gram negative tripartite α-PFTs has been successful. This has enabled the in-depth structural analysis of all three proteins from the tripartite α-PFT family, which has helped to identify their respective roles, and in addition, a conserved mechanism for the transformation from soluble to pore conformations has been identified. Further to this, it has been shown that all ClyA-family α-PFTs require three features to assemble a lytic pore, and these three features are provided by one, two or three proteins, for ClyA, and the bipartite and tripartite families, respectively. These features are (outlined in Figure 6.1):

- Initial insertion through a single leaflet of the membrane by a short hydrophobic helix domain.

- A longer hydrophobic or amphipathic helix domain that penetrates both leaflets of the membrane and forms the pore.
- A hydrophilic lumen.

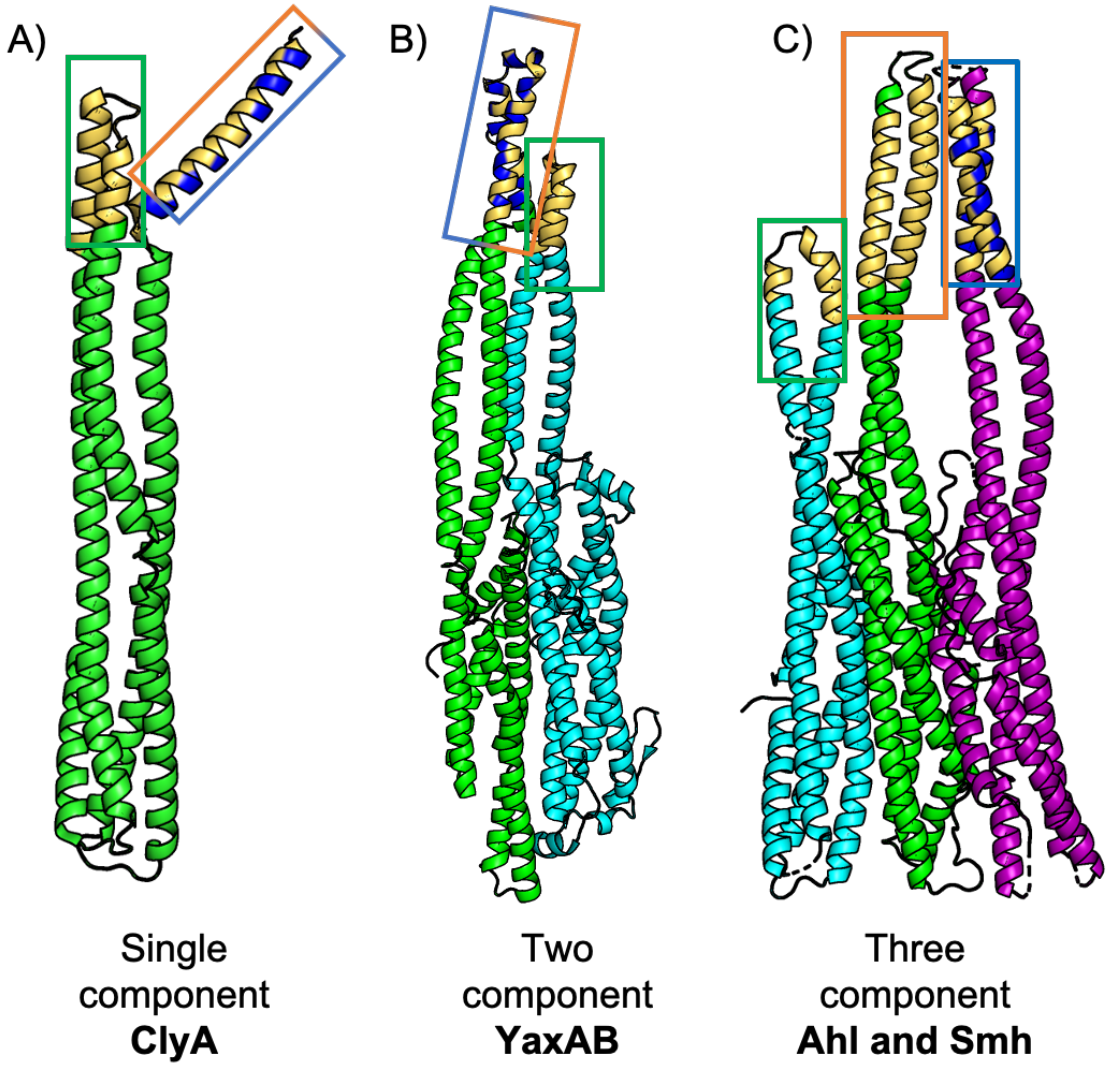


Figure 6.1 **All ClyA family members must have three features for a lytic pore.** Pore conformations for mono, bi and tripartite α-PFTs. The green box highlights the short hydrophobic helical domain, while blue and orange highlight the hydrophobic membrane-spanning domain, and hydrophilic lumen, respectively. A) Single component members (ClyA) contain all three features in a single protein (green). B) Two-component members (YaxAB and XaxAB) split the features between two proteins (A component, light blue, B component green). C) Three-component members (Ahl, Smh, Nhe, Hbl) use three proteins each carrying a single feature (A component, pink; B component, green; C component light blue), for all structures hydrophobic (orange) and hydrophilic (blue) residues are highlighted for the membrane regions.

Finally, analysis of the Gram positive tripartite α-PFTs, (Nhe and Hbl), alongside the biochemical analysis of the two newly identified Gram negative tripartite α-PFTs that formed part of this project, has shown that although these tripartite family members are structurally similar, each varies significantly in their modes of action and pore assembly.

6.1.1 Structures of the A and B components show a conserved mechanism for conformational change

As presented in Paper 1 (Section 3.3) the structure of both the soluble and pore form of AhlB provided the first indication of the conformational transition of a tripartite α-PFT component. These structures were shown to be conserved within the wider Gram negative tripartite α-PFT family, with the structure determination of the soluble and pore forms of SmhB presented in Section 5.3 (Paper 3).

The soluble structures of both AhlB and SmhB shared the same overall compact helical tail domain and β-tongue head domain also observed in the Gram positive NheA (A component) and Hbl-B (C component) tripartite α-PFT proteins from *B. cereus* (Ganash et al., 2013; Madegowda et al., 2008). The structures of the pores of AhlB and SmhB provided the first structural data of a pore conformation from a tripartite α-PFT, and as with the soluble structures, the pore structures of both AhlB and SmhB shared the same overall structure of a decameric pore of two extended helical conformations of the B component. These two conformations of the B pore component have been labelled type 1 and type 2, respectively, with the type 2 conformation adopting a more extended structure than the type 1. This has allowed for a mechanism for the transition from soluble to pore form to be proposed whereby the soluble structure transforms first to the type 1 conformation then to the type 2 conformation, with the β-tongue of the soluble structure refolding to form the distal ends of two hydrophobic extended helices that are of the correct length to span the membrane. This conformational change in the β-tongue is also seen to a smaller extent in ClyA (Roderer and

Glockshuber, 2017). The AhlB and SmhB pores use an extended helix turn helix to span the membrane, like YaxB and XaxB in the YaxAB and XaxAB pores. In the YaxAB and XaxAB pore, however, the tail of B packs against A (equivalent to AhlC and SmhC) rather than a neighbouring B as in AhlB and SmhB pores. This difference in tail packing results in YaxAB and XaxAB having a more funnel-shaped pore with maximum diameters at the head and tail double that of the AhlB and SmhB pores (Bräuning et al., 2018; Schubert et al., 2018). It would be reasonable to assume a similar shape and size for the tripartite pores formed from all three components.

Papers 2 and 3 (Section 4.3 and 5.3) present the structure of SmhA, the A component from the Smh toxin. Although this protein shares very low sequence identity (as determined by BlastP (Altschul et al., 1990)) to NheA (5.2 %), SmhB (no significant similarity) or AhlB (no significant similarity), it was shown to share striking structural similarity to all three, again with the same compact helical tail domain and β-tongue head domain. Sequence analysis described in Section 3.3 (Paper 1) predicted that the A components of the tripartite toxins would use their β-tongue head domain to form two amphipathic helices and pack against the B component to provide a hydrophilic lumen to the pore. The structure of the SmhA component, together with modelling of a pore conformation based on the structural similarities to AhlB and SmhB, did indeed show that SmhA could form two amphipathic helices long enough to span the membrane.

Further structural analysis between SmhA, SmhB, AhlB and NheA identified two conserved latches between the tail and C-terminus, and also between the head and tail domains, at the predicted binding interfaces between the A and B, and B and C components. In the B components, these latches are broken in the pore conformation. It is therefore proposed that these latches would also be broken in the A components, in order to release the C-terminus (which must move to free the head domain) and thus to release the β-tongue to form the extended helices. The high structural similarity between AhlB, SmhB, SmhA and both NheA

and Hbl-B suggests that the β-tongues of these two *B. cereus* proteins will also undergo a similar transformation to form their respective pore conformations. This hypothesis is further supported by monoclonal antibody studies of NheA which have confirmed that a large rearrangement of the structure takes place on binding to NheB in the membrane (Didier et al., 2016). This transition is likely conserved in all β-tongue containing tripartite family member proteins.

6.1.2 Modelling the tripartite pore

Section 5.3, Paper 3 proposes a possible tripartite family pore assembly, where each component is present in equal amounts, with an outer ring of C bound to an inner ring of B and A. This model agrees with both docking predictions and structural analysis of all the ClyA family members, including the known structures of the mono and bipartite pores, and its size agrees with the TEM images of the Ahl, and Smh pores (Figure 1.2A). This structure also fulfils all the requirements for lytic pore formation proposed in Paper 1. This modelled pore structure has a larger diameter than the B pore alone, in order to accommodate the A and C components in the complete assembly. Here the tripartite pore is constructed in a similar way to both the bipartite YaxAB and XaxAB pores (Figure 1.11), except that the A component of the tripartite pore provides the hydrophilic lumen, whereas in the bipartite pores the amphipathic helices of the B component provide both the membrane-spanning and hydrophilic lumen to the pore (Bräuning et al., 2018; Schubert et al., 2018).

During the modelling process, two alternative pore structures were considered, described below, but for each of these alternative structures problems in the assembly occur and both are thus less likely to be correct. A structure determination of the full ABC pore is required to resolve this ambiguity.

Alternative Pore Structure 1

As discussed in Paper 3 the Gram negative A components each contain a 20 amino acid extended loop in the β-tongue. In the pore model proposed in Paper 3, the position of the A component would mean this extended loop, that links the ends of α3 and α4, is located on the intracellular side of the membrane, with the hydrophilic side of the extended amphipathic α3 and α4 helices lining the lumen of the pore (Figure 6.2B). The docking predictions using the Nhe Mab binding sites, the protomer structures and the Haddock server, show that it is possible for an alternative model to be proposed where the A component would bind to B such that the tail would protrude above the rest of the pore (Figure 6.2B). This would place the loop between α3 and α4 within the lumen of the pore, with residues on this loop providing the hydrophilic lining to the pore. However in this position, the hydrophobic surface of the extended amphipathic α3 and α4 helices would have to interact with the hydrophilic residues in the tail domain of B, a very unfavourable situation, and thus this model is perhaps unlikely to occur.

Alternative Pore Structure 2

As the structure of SmhA is very closely related to both SmhB and AhlB, it is also possible that SmhA could adopt either the B type1 or type 2 pore conformations discussed in Section 3.3(paper 1). It is thus possible that the A component could replace one of these conformations in the decameric B pore structure, to create an alternating AB pore lumen constructed from five AB dimers. In this model, the hydrophilic faces of the amphipathic α3 and α4 helices of A would provide the hydrophilic lining to the pore (Figure 6.2C). In this model, the extended loop between α3 and α4 of A would be located on the intracellular side on the membrane, as in the model proposed in Paper 3 (Section 5.3), with a ring of the C component surrounding the alternating AB pore (Figure 6.2C). Although the proposed interfaces between A and B would be correctly located within the AB dimer of such a

structure, the interactions between the AB dimers, which would require a second A binding site on B, would not agree with the Mab binding studies of Nhe, where only a single A binding site was identified on B (Didier et al., 2012), and this model is therefore also unlikely based on the accumulated experimental evidence.

Figure 6.2

Figure 6.2 Models of possible pore assemblies of the tripartite PFT from Gram negative bacteria . A) Model with 10 protomers of components A, B and C as described in Paper 3, component A (pink) provides an inner hydrophilic lumen to the pore. B) Alternative model 1 with 10 protomers of A (pink), B (green) and C (light blue) in which the A component protrudes from the tail of the pore, with the extended loop lining the lumen of the pore. C) Alternative model 2 with the A component replacing a type 2 B component in the B pore to provide the hydrophilic lining, with C bound to B around the exterior of the pore, and with 10 C, 5 A and 5 B protomers. The membrane is represented by a blue line.

6.1.3 Mechanism of pore assembly

Papers 1 and 3 (Section 3.3 and 5.3) propose mechanisms of pore assembly for the Ahl α-PFT and the Smh α-PFT. Biochemical analysis of these two systems has shown that although they share structural and sequence similarities, they adopt mechanisms for assembly of a lytic pore on the membrane, different in crucial aspects, and they also adopt different mechanisms in the regulation of pore formation. These two mechanisms also differ from those of the Gram positive tripartite α-PFTs Nhe and Hbl, however some aspects are shared between the systems.

In paper 1 (Section 3.3) Ahl is proposed to assemble sequentially. Firstly the AhlC tetramer dissociates, binds to and primes the membrane, followed by the binding of AhlB and AhlA, oligomerisation and lytic pore formation. Although no soluble complexes form between these proteins, (due to the tetrameric assembly of AhlC), off-pathway AhlB and AhlBC pores can form if erythrocytes are preincubated with either of these, resulting in a reduced rate of lysis. This may represent a mechanism by which the Ahl toxin regulates lytic pore formation. No reduction in total lysis was observed, however, when AhlB was in excess to AhlA and AhlC, and maximum lysis was reached at a ratio of 1:1:1 for the AhlA, AhlB and AhlC components (Figure 6.3).

Paper 3 (Section 5.3) describes a similar mechanism of assembly for Smh, where initial binding of SmhC enhances the rate of pore formation and then SmhB and SmhA can rapidly assemble at the membrane to form a lytic pore. Lysis is increased by preincubation with SmhC in combination with SmhB or SmhA, and, unlike with Ahl, the formation of SmhBC pores does not reduce lysis. Also, in contrast to Ahl, the Smh components are able to form inhibitory soluble complexes, with SmhA+SmhB+SmhC forming 1MDa soluble SmhABC complexes and SmhB+SmhC forming 1.5MDa soluble SmhBC complexes (Paper 3 (Section 5.3) and Appendix B Section B.4). Optimum lysis and rate are achieved when SmhC is in excess to SmhB and SmhA (Appendix B Section B.3). These observations lead

to a mechanism of assembly with two possible pathways, the first, where trimeric SmhABC structures can form in the membrane then oligomerise and lyse the cell, and the second where SmhBC pores can form in the membrane, allowing the rapid recruitment of SmhA to make the active toxin (Figure 6.4). In both cases, SmhC must be present to enhance the rate of pore assembly. Excess of either SmhB or SmhA relative to SmhC results in the formation of inhibitory soluble complexes and acts as a regulating factor to pore formation. This proposed assembly and regulatory mechanism of the Smh toxin shares similarities with both the Nhe and Hbl toxins. The first assembly pathway of Smh is the same as Hbl where each component binds sequentially, and the presence of Hbl-L1 (equivalent to SmhB) enhances binding of Hbl-B (equivalent to SmhC) to the cell improving lysis (Jessberger et al., 2019; Tausch et al., 2017). However, no Hbl-L1B pro-pore can be seen, unlike with SmhBC. In the Nhe system, NheBC pro-pores form before the final recruitment of NheA (Didier et al., 2016; Zhu et al., 2016), this is similar to the second proposed pathway for Smh, described above. Furthermore, the A and B components of Hbl (Hbl- L2, and Hbl- L1) are inhibitory when in excess to Hbl-B and form soluble inhibitory complexes as seen with Smh. This is the opposite to Nhe, where NheC in excess to NheB is inhibitory (Jessberger et al., 2019; Tausch et al., 2017; Zhu et al., 2016) (Figure 6.5).

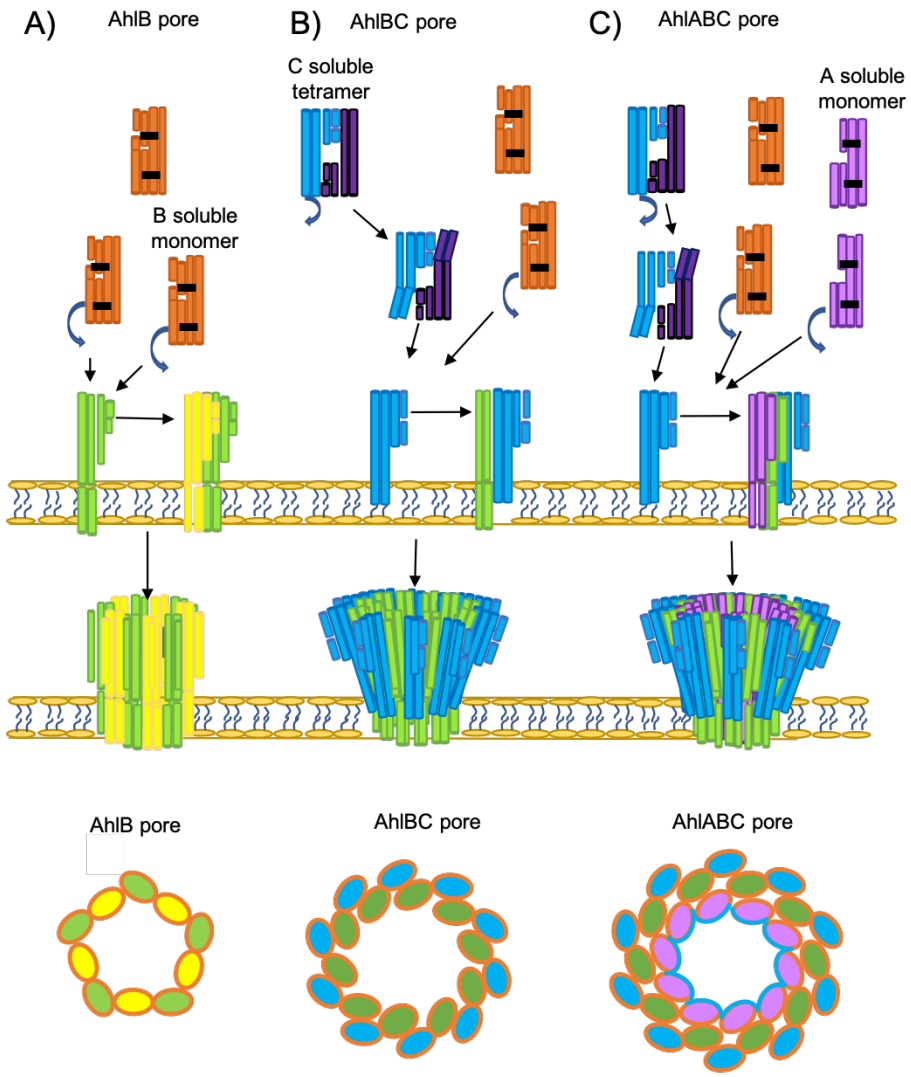


Figure 6.3 **A proposed assembly schematic of AhlB, AhlBC and AhlABC pores.** A) Assembly of AhlB pores; soluble AhlB (orange) reconfigures to the Type 1 pore conformation (yellow) on exposure to the lipid bilayer and recruits more AhlB monomers to form an inactive pore of mixed Type 1, Type 2 (green) conformation. B) Assembly of AhlBC pores; when AhlC (cyan) is present tetramers of AhlC disassemble at the membrane and monomers insert into one leaflet. Soluble AhlB (orange) is recruited to the lipid bilayer where it unpacks (green) to form a hetero dimer with AhlC. Further AhlB and AhlC protomers are recruited until a complete pore is formed with a ring of AhlC on the outside and a hydrophobic ring of AhlB on the inside. C) Assembly of AhlABC pores; AhlC inserts into the membrane as in AhlBC pore assembly, soluble AhlA and AhlB are then recruited to the membrane where they associate with AhlC and further AhlA, AhlB and AhlC are recruited until a complete pore is formed with a hydrophilic lining from AhlA on the inside of the membrane-spanning region. Shown below are cross-section views through the membrane-bound region of each pore, with AhlB type 1 (yellow), AhlB type 2 (green), AhlC (blue) and AhlA (pink) with hydrophobic surfaces (orange) and hydrophilic (light blue) highlighted. Each oval represents the two α3, α4 head helices. Latches in the A and B components are shown as black lines.

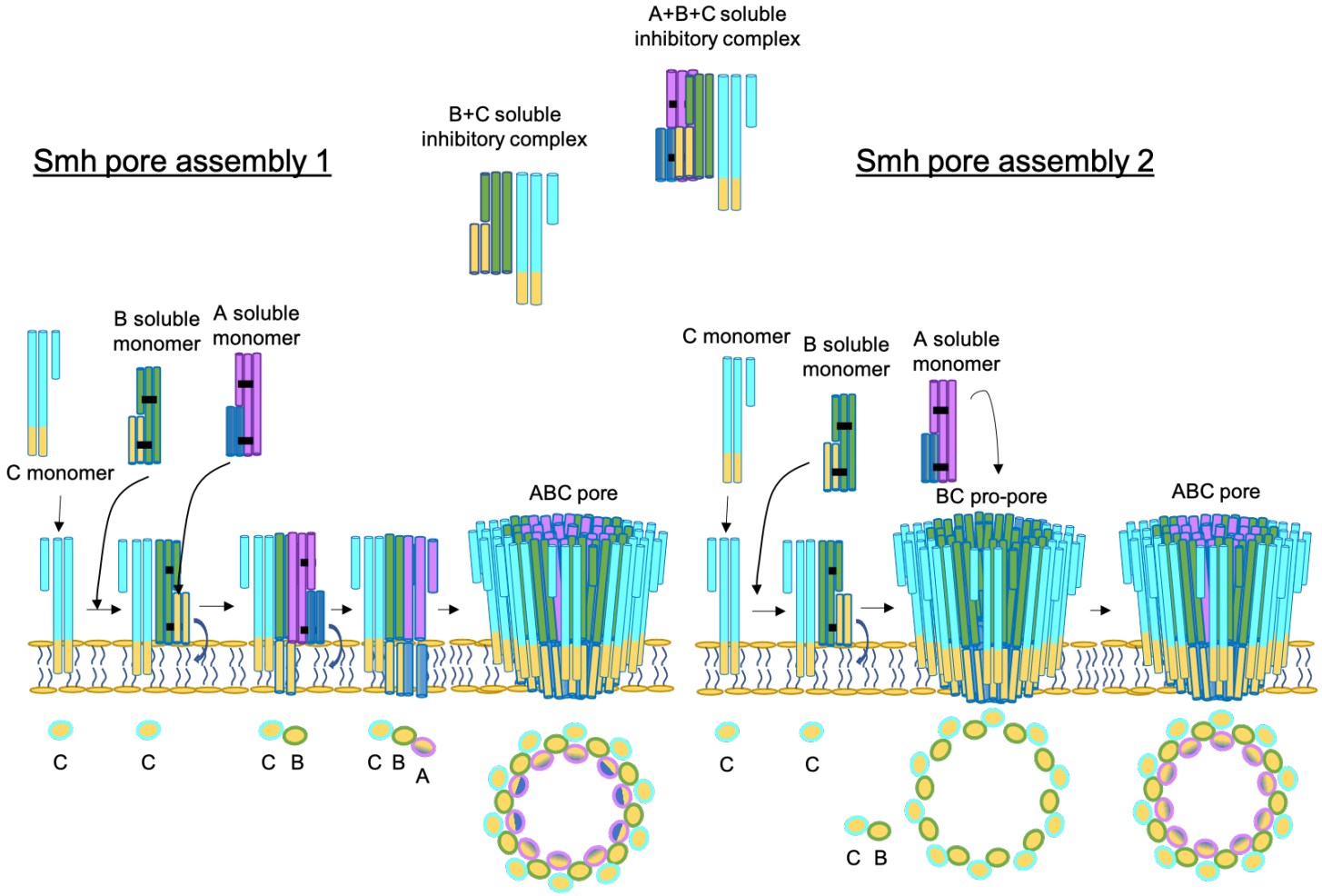


Figure 6.4

Figure 6.4 **Schematic showing pore assembly for the Smh toxin.** . Left – assembly mechanism 1. The C, B and A components assemble sequentially on the membrane to form a trimer which then oligomerises to form the lytic pore. Right - assembly mechanism 2. First a BC pro-pore forms in the membrane before the recruitment of the A component to form a lytic pore. Soluble inhibitory complexes can form between SmhB+SmhC and SmhA+SmhB+SmhC. Latches in the A and B components are shown as black lines.

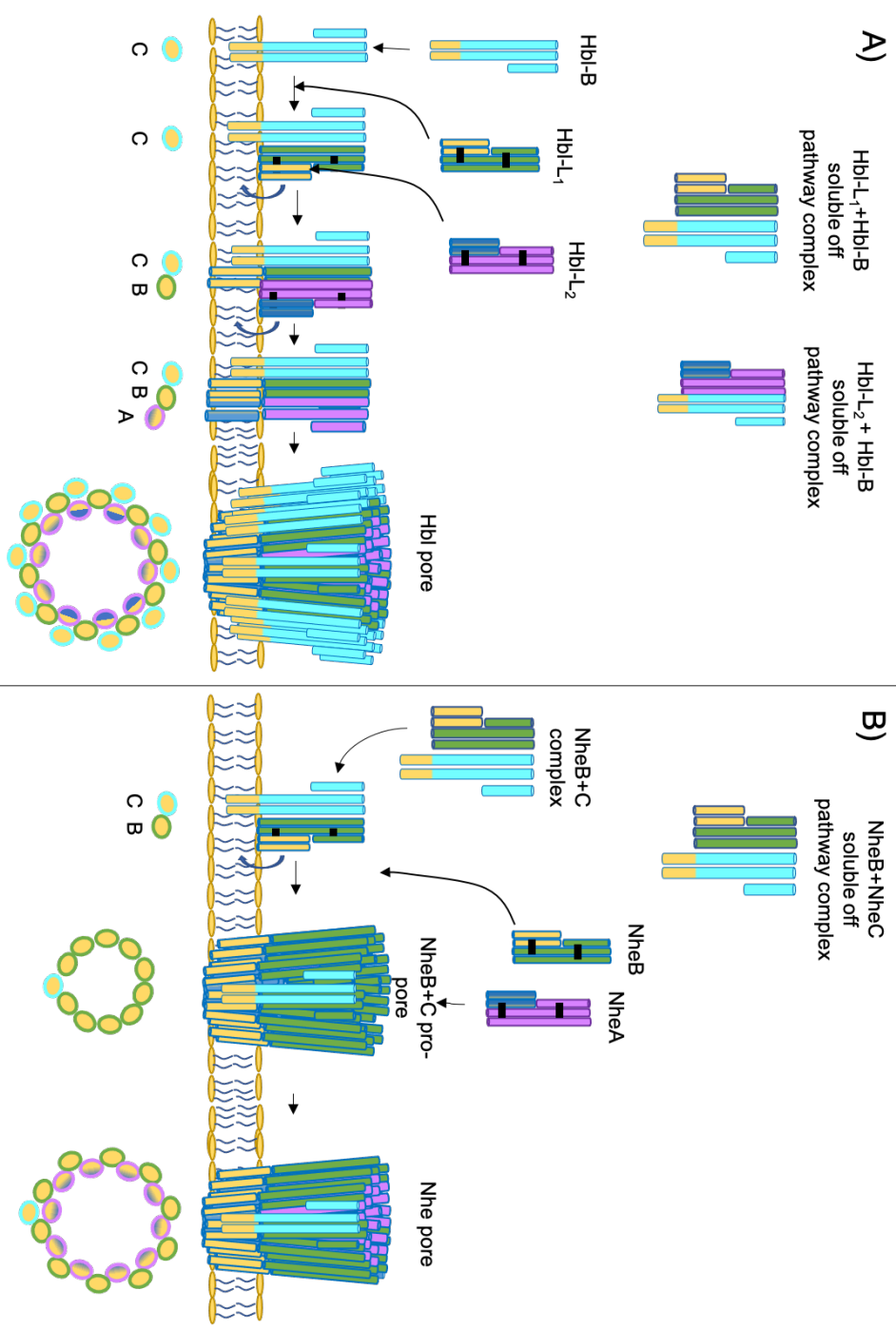


Figure 6.5

Figure 6.5 **Schematic showing pore assembly for the Nhe and Hbl toxins.** A) Hbl toxin assembly, Hbl-B, Hbl-L_1 and Hbl-L_2 assemble sequentially on the membrane to form a trimer which then oligomerises to form the lytic pore. Soluble inhibitory complexes can form between the Hbl-L_1 and Hbl-B components and the Hbl-L_2 and Hbl-B components. B) Nhe assembly, First a BC complex binds the membrane before recruiting further NheB to form a pro-pore in the membrane, this is followed by the rapid recruitment of the NheA component to form a lytic pore. Soluble inhibitory complexes can form between NheB+NheC. Latches in the A and B components are shown as black lines.

In summary, although there is still ambiguity in the exact composition of the tripartite pores based on the structural studies in this thesis, it is probable that the model described in Paper 3 (Section5.3) is a good representation of both the Ahl and Smh pores. Although structural studies have shown many conserved features within the family used to achieve the soluble to pore transition and final pore assembly, the assembly and regulatory mechanisms used prior to cell lysis vary significantly. The Smh toxin seems to have adopted the same regulatory mechanism as Hbl, whereby the ratio of the A and B components relative to the C component is used to control the rate of assembly and lytic activity, while at the same time utilising the assembly mechanisms of both Hbl and Nhe (Jessberger et al., 2019; Tausch et al., 2017; Zhu et al., 2016). In contrast, the only similarity in the assembly and regulatory mechanism of Ahl to the other tripartite pores is the requirement for AhlC to enhance rate and lysis. The tetrameric assembly of AhlC appears to prevent soluble inhibitory complex formation, instead Ahl forms off-pathway pores as a possible regulating step to cell lysis, a mechanism not seen in any of the other tripartite members. Gel filtration and bioinformatics analysis (Appendix B, Section B.7) of SmhC, suggests that it will share the same extended conformation as AhlC but form a dimer rather than the tetramer seen in AhlC. If SmhC does form a dimer then more of the subunit surface area would be exposed compared to the AhlC tetramer and this would mean SmhC could be more available to form BC soluble inhibitory complexes. The variations in regulation of pore formation of all the different tripartite toxins thus appear to be due, in part, to variations in the quaternary structure of the soluble C component (Appendix B Section B.7). However, further structures are required to confirm this hypothesis, not least of the complete pore.

Although the assembly and regulatory mechanism of each system has diverged, structural studies in this thesis have shown that the A and B component structures are well conserved and the mechanism for transformation from soluble to pore is likely also well conserved within the family. Homologues of the A component are only found within tripartite operons,

while B component homologues are often found as orphan genes or part of bipartite operons. The B component has been shown to be the membrane-spanning pore-forming component in all identified tripartite and bipartite toxins and is the most well conserved within the ClyA-family of α-PFTs. It is also the only component able to form a pore alone (Jessberger et al., 2019; Sastalla et al., 2013; Tausch et al., 2017; Zhu et al., 2016). As the A component has a high structural similarity with B, it is possible that A evolved from a gene duplication event of B, but subsequently lost its pore-forming activity and can thus only act as part of a tripartite system. The single leaflet binding components with no β-tongue (YaxA, SmhC, AhlC) (Bräuning et al., 2018), could also have evolved from a reduction in the size of the membrane-spanning characteristics of the B component. As the B component can be found as an orphan gene and is also the only component able to form pores alone suggests the initial protein in the family may have been a B component with a fully hydrophobic membrane-spanning region, this ancestor could have then evolved into the tripartite A, C and bipartite A and B structures. It remains to be seen whether the single component ClyA toxin evolved in a similar way or whether it was the ancestor and all other components evolved from it. Understanding these relationships clearly requires further work and analysis. Nevertheless, these proposed gene duplication events have resulted in the tripartite α-PFT systems being able to separate the three key features required to form a lytic pore, on to three separate proteins. This has enabled divergence in assembly mechanism, and specificity in targeting different cell types, helping different bacterial species to thrive in their unique environments.

6.2 Future work

The work presented in this thesis proffers a number of important discoveries with regard to the understanding of the ClyA-family of tripartite α-PFTs. It has also given rise to several questions which future work could resolve.

6.2.1 Structure of a tripartite pore

The structures presented in this thesis have allowed the creation of a markedly more detailed model of a tripartite pore. To fully understand how the three proteins interact to form a tripartite pore as well as their final pore conformations, a structure of the tripartite pore must be obtained. This project has made significant progress in optimising both the Ahl and Smh pores for structure solution using Cryo-electron microscopy (Appendix A Section A.6 and Appendix B Section B.9). Future work could therefore focus on Cryo-EM of the helical filaments of Smh and single-particle work with Smh and Ahl pores in detergent (N-heptyl-thioglucopyranoside). Further to this LCP crystallisation or crystallisation with detergent mimic additives may also prove fruitful in obtaining a structure of the tripartite pore.

6.2.2 Structures of SmhC and NheC and Soluble complexes

Papers 1 and 3 discuss the role of the C component in pore formation and regulation, highlighting the importance of this component in regulating the tripartite toxins. It seems likely that the quaternary structure of the C component plays a role in the formation of soluble inhibitory complexes. Obtaining structures of SmhC and NheC, which both behave differently from AhlC in lytic assays, would confirm their quaternary structure and whether this is indeed part of the variation in regulatory mechanisms. In addition, structures of inhibitory complexes would help to better understand membrane interactions and pore assembly, as these could be compared with membrane-bound conformations to identify regions important in membrane interaction.

6.2.3 Cell lysis assays

Papers 1 and 3 show that Ahl is 100x less potent than Smh at lysing erythrocytes. This suggests a preference of these systems for certain cell membranes, to date no protein or lipid receptors have been been identified for the tripartite toxins. Assays with a range of different

cell types would help confirm a preference for these toxins to target certain membranes, as has been shown with other PFTs. Assays on different cell types are more involved than simply lysing erythrocytes, but methods such as staining cells with propidium iodide (PI) after incubation with the toxin of interest, then measuring fluorescence signal using FACS could be used. As PI binds to DNA, but cannot traverse the membrane, any cells stained by PI must have severely damaged membranes (dead), similarly unstained cells would have intact membrane (healthy). Live and dead cells can then be sorted and measured depending on the fluorescence signal. This method can be applied to any DNA containing cell line.

6.2.4 Chimeric Haemolytic assays

Paper 3 showed how SmhB and AhlB share both high sequence identity and structural similarities, and how a chimeric pore containing SmhA, AhlB and SmhC was still able to completely lyse erythrocytes. Generation of a range of chimeric Ahl/Smh pores may show if the differences between the Ahl and Smh systems observed and discussed above are due to a single component and whether that characteristic can be transferred between the two pores by switching a single component. If a preference for certain cell membranes is identified, research could be carried out into whether this preference is transferred by swapping a component and thus targeting pores to certain cell types. This type of chimeric assembly could be extended to include components from other potential Gram-negative tripartite α-PFTs identified in the sequence alignments.

6.2.5 Protein engineering

As discussed in Section 1.6, α-PFTs are now of great interest in the biotechnology for applications in DNA and protein sequencing, as well as in biosensors and drug delivery. The work in this thesis has shown tripartite α-PFTs to be highly versatile due to the three components and could make good candidates for biotechnological applications. First, it

must be shown that these pores are stable in an artificial membrane and that a current can be measured across them. Next, engineering these pores by mutation and generation of chimeras could make them adaptable for use in measuring a range of analytes and for use in targeted drug delivery to particular cell types.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010). PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, 66(2):213–221.

Alouf, J., Ladant, D., and Popoff, M. R. (2015). *The Comprehensive Sourcebook of Bacterial Protein Toxins*. Elsevier Inc.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):271–281.

Beecher, D. J., Schoeni, J. L., Lee Wong, A. C., and Wong, A. C. (1995). Enterotoxic activity of hemolysin BL from Bacillus cereus. *Infection and Immunity*, 63(11):4423–4428.

Benke, S., Roderer, D., Wunderlich, B., Nettels, D., Glockshuber, R., and Schuler, B. (2015). The assembly dynamics of the cytolytic pore toxin ClyA. *Nature communications*, 6:6198.

Berg, J., Tymoczko, J., and Stryer, L. (2012). *Biochemistry*. W.H. Freeman, Palgrave Macmillan, seventh ed edition.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., and Zardecki, C. (2002). The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, 58(6 I):899–907.

Bhat, H. B., Kishimoto, T., Abe, M., Makino, A., Inaba, T., Murate, M., Dohmae, N., Kurahashi, A., Nishibori, K., Fujimori, F., Greimel, P., Ishitsuka, R., and Kobayashi, T. (2013). Binding of a pleurotolysin ortholog from Pleurotus eryngii to sphingomyelin and cholesterol-rich membrane domains. *Journal of Lipid Research*, 54(10):2933–2943.

Böhm, M. E., Krey, V. M., Jeßberger, N., Frenzel, E., and Scherer, S. (2016). Comparative bioinformatics and experimental analysis of the intergenic regulatory regions of Bacillus cereus hbl and nhe enterotoxin operons and the impact of CodY on virulence heterogeneity. *Frontiers in Microbiology*, 7(MAY).

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Technical report.

Bräuning, B., Bertosin, E., Praetorius, F., Ihling, C., Schatt, A., Adler, A., Richter, K., Sinz, A., Dietz, H., and Groll, M. (2018). Structure and mechanism of the two-component α-helical pore-forming toxin YaxAB. *Nature Communications*, 9(1):1806.

Castillo, J. P., Rui, H., Basilio, D., Das, A., Roux, B., Latorre, R., Bezanilla, F., and Holmgren, M. (2015). Mechanism of potassium ion uptake by the Na+/K+-ATPase. *Nature Communications*, 6(1):7622.

Churchill-angus, A. M., Sedelnikova, S. E., Schofield, T. H. B., and Patrick, J. (2020). The A component (SmhA) of a tripartite pore forming toxin from Serratia marcescens : expression , purification , and crystallographic analysis. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*.

Cosentino, K., Ros, U., and García-Sáez, A. J. (2016). Assembling the puzzle: Oligomerization of α-pore forming proteins in membranes. *Biochimica et Biophysica Acta - Biomembranes*, 1858(3):457–466.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography*, 62(9):1002–1011.

Cowtan, K. (2010). Recent developments in classical density modification. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):470–478.

Didier, A., Dietrich, R., Gruber, S., Bock, S., Moravek, M., Nakamura, T., Lindbäck, T., Granum, P. E., and Märtlbauer, E. (2012). Monoclonal antibodies neutralize Bacillus cereus Nhe enterotoxin by inhibiting ordered binding of its three exoprotein components. *Infection and immunity*, 80(2):832–8.

Didier, A., Dietrich, R., and Märtlbauer, E. (2016). Antibody Binding Studies Reveal Conformational Flexibility of the Bacillus cereus Non-Hemolytic Enterotoxin (Nhe) A-Component. *PLOS ONE*, 11(10):e0165135.

Dietrich, R., Moravek, M., Bürk, C., Granum, P. E., and Märtlbauer, E. (2005). Production and characterization of antibodies against each of the three subunits of the Bacillus cereus nonhemolytic enterotoxin complex. *Applied and Environmental Microbiology*, 71(12):8214–8220.

Dongre, M., Singh, B., Aung, K. M., Larsson, P., Miftakhova, R., Persson, K., Askarian, F., Johannessen, M., von Hofsten, J., Persson, J. L., Erhardt, M., Tuck, S., Uhlin, B. E., and Wai, S. N. (2018). Flagella-mediated secretion of a novel Vibrio cholerae cytotoxin affecting both vertebrate and invertebrate hosts. *Communications Biology*, 1(1):59.

DuMont, A. L. and Torres, V. J. (2014). Cell targeting by the staphylococcus aureus pore-forming toxins: It's not just about lipids.

Egelman, E. H., Grant, T., Rohou, A., and Grigorieff, N. (2018). friendly software for single-particle image processing. *eLife*, 7.

Eifler, N., Vetsch, M., Ringler, P., Philippsen, A., Fritz, A., Mu, S. A., and Glockshuber, R. (2006). Cytotoxin ClyA from Escherichia coli assembles to a 13-meric pore independent of its redox-state. 25(11):2652–2661.

Elluri, S., Enow, C., Vdovikova, S., Rompikuntal, P. K., Dongre, M., Carlsson, S., Pal, A., Uhlin, B. E., and Wai, S. N. (2014). Outer membrane vesicles mediate transport of biologically active Vibrio cholerae cytolysin (VCC) from V. cholerae strains. *PLoS ONE*, 9(9).

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. Sect. D*, 66(4):486–501.

Escajadillo, T. and Nizet, V. (2018). Pharmacological Targeting of Pore-Forming Toxins as Adjunctive Therapy for Invasive Bacterial Infection. *Toxins*, 10(12):542.

Evans, G. and Pettifer, R. F. (2001). CHOOCH: A program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *Journal of Applied Crystallography*, 34(1):82–86.

Evans, P. (2006). Scaling and assessment of data quality. In *Acta Crystallographica Section D: Biological Crystallography*, volume 62, pages 72–82. International Union of Crystallography.

Evans, P. R. and Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography*, 69(7):1204–1214.

Fagerlund, A., Lindbäck, T., and Granum, P. E. (2010). Bacillus cereus cytotoxins Hbl, Nhe and CytK are secreted via the Sec translocation pathway. *BMC Microbiology*, 10(1):304.

Fagerlund, A., Lindbäck, T., Storset, A. K., Granum, P. E., and Hardy, S. P. (2008). Bacillus cereus Nhe is a pore-forming toxin with structural and functional properties similar to the ClyA (HlyE, SheA) family of haemolysins, able to induce osmotic lysis in epithelia. *Microbiology*, 154(3):693–704.

Fahie, M. A., Liang, L., Avelino, A. R., Pham, B., Limpikirati, P., Vachet, R. W., and Chen, M. (2018). Disruption of the open conductance in the β-tongue mutants of Cytolysin A. *Scientific Reports*, 8(1):3796.

Fox, D., Mathur, A., Xue, Y., Liu, Y., Tan, W. H., Feng, S., Pandey, A., Ngo, C., Hayward, J. A., Atmosukarto, I. I., Price, J. D., Johnson, M. D., Jessberger, N., Robertson, A. A. B., Burgio, G., Tscharke, D. C., Fox, E. M., Leyton, D. L., Kaakoush, N. O., Märtlbauer, E., Leppla, S. H., and Man, S. M. (2020). Bacillus cereus non-haemolytic enterotoxin activates the NLRP3 inflammasome. *Nature Communications*, 11(1).

Franceschini, L., Soskine, M., Biesemans, A., and Maglia, G. (2013). A nanopore machine promotes the vectorial transport of DNA across membranes. *Nature Communications*, 4(1):2415.

Fuentes, J. A., Villagra, N., Castillo-Ruiz, M., and Mora, G. C. (2008). The Salmonella Typhi hlyE gene plays a role in invasion of cultured epithelial cells and its functional transfer to S. Typhimurium promotes deep organ infection in mice. *Research in Microbiology*, 159(4):279–287.

Ganash, M., Phung, D., Sedelnikova, S. E., Lindbäck, T., Granum, P. E., and Artymiuk, P. J. (2013). Structure of the NheA Component of the Nhe Toxin from Bacillus cereus: Implications for Function. *PLoS ONE*, 8(9):e74748.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, pages 571–607.

Geny, B. and Popoff, M. R. (2006). Bacterial protein toxins and lipids: pore formation or toxin entry into cells. *Biology of the Cell*, 98(11):667–678.

Giddings, K. S., Johnson, A. E., and Tweten, R. K. (2003). Redefining cholesterol's role in the mechanism of the cholesterol-dependent cytolysins. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20):11315–11320.

Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L. Y. (2004). A simple method for improving protein solubility and long-term stability. *Journal of the American Chemical Society*, 126(29):8933–8939.

Gonzalez, M. R., Bischofberger, M., Pernot, L., Van Der Goot, F. G., and Frêche, B. (2008). Bacterial pore-forming toxins: The (w)hole story?

Granum, P. E., O'sullivan, K., and Lund, T. (1999). The sequence of the non-haemolytic enterotoxin operon from Bacillus cereus 1. *FEMS Microbiology Letters*, (117):225–229.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, 166(4):557–580.

Hatha, M., Vivekanandhan, A. A., Julie Joice, G., and Christol (2005). Antibiotic resistance pattern of motile aeromonads from farm raised fresh water fish. *International Journal of Food Microbiology*.

He, S. and Scheres, S. H. W. (2017). Helical reconstruction in RELION. *Journal of structural biology*, 198(3):163–176.

Heilkenbrinker, U., Dietrich, R., Didier, A., Zhu, K., Lindbäck, T., Granum, P. E., and Märtlbauer, E. (2013). Complex Formation between NheB and NheC Is Necessary to Induce Cytotoxic Activity by the Three-Component Bacillus cereus Nhe Enterotoxin. *PLoS ONE*, 8(4):e63104.

Holm, L. and Laakso, L. M. (2016). Dali server update. *Nucleic acids research*, 44(W1):351–355.

Huang, G., Voet, A., and Maglia, G. (2019). FraC nanopores with adjustable diameter identify the mass of opposite-charge peptides with 44 dalton resolution. *Nature Communications*, 10(1):835.

Huang, G., Willems, K., Soskine, M., Wloka, C., and Maglia, G. (2017). Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nature Communications*, 8(1):935.

Iacovache, I., De Carlo, S., Cirauqui, N., Dal Peraro, M., van der Goot, F. G., and Zuber, B. (2016). Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. *Nature Communications*, 7(May):12062.

Jessberger, N., Dietrich, R., Schwemmer, S., Tausch, F., Schwenk, V., Didier, A., and Märtlbauer, E. (2019). Binding to the target cell surface is the crucial step in pore formation of hemolysin BL from Bacillus cereus. *Toxins*, 11(5).

Kabsch, W. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography*, 66(2):125–132.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6):845–858.

Krissinel, E. and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, 372(3):774–797.

Leal, R. M. F., Bourenkov, G., Russi, S., and Popov, A. N. (2013). A survey of global radiation damage to 15 different protein crystal types at room temperature: A new decay model. *Journal of Synchrotron Radiation*, 20(1):14–22.

Libby, S. J., Goebel, W., Ludwig, A., Buchmeier, N., Bowe, F., Fang, F. C., Guiney, D. G., Songer, J. G., and Heffron, F. (1994). A cytolysin encoded by Salmonella is required for survival within macrophages. *Proceedings of the National Academy of Sciences of the United States of America*, 91(2):489–493.

Lindbäck, T., Fagerlund, A., Rødland, M. S., Granum, P. E., Lindbä ck, T., Fagerlund, A., Skeie Rødland, M., and Einar Granum, P. (2004). Characterization of the Bacillus cereus Nhe enterotoxin. *Microbiology*, 150(12):3959–3967.

Lindbäck, T., Hardy, S. P., Dietrich, R., Sødring, M., Didier, A., Moravek, M., Fagerlund, A., Bock, S., Nielsen, C., Casteel, M., Granum, P. E., and Märtlbauer, E. (2010). Cytotoxicity of the Bacillus cereus Nhe enterotoxin requires specific binding order of its three exoprotein components. *Infection and Immunity*, 78(9):3813–3821.

Los, F. C. O., Randis, T. M., Aroian, R. V., and Ratner, A. J. (2013). Role of Pore-Forming Toxins in Bacterial Infectious Diseases. *Microbiology and Molecular Biology Reviews*, 77(2):173–207.

Lund, T. and Granum, P. E. (1996). Characterisation of a non-haemolytic enterotoxin complex from Bacillus cereus isolated after a foodborne outbreak. *FEMS Microbiology Letters*, 141(2-3):151–156.

Madegowda, M., Eswaramoorthy, S., Burley, S. K., and Swaminathan, S. (2008). X-ray crystal structure of the B component of Hemolysin BL from Bacillus cereus. *Proteins: Structure, Function and Genetics*, 71(2):534–540.

Maglia, G., Heron, A. J., Stoddart, D., Japrung, D., and Bayley, H. (2010). Analysis of Single Nucleic Acid Molecules with Protein Nanopores. *Methods in Enzymology*, 475(C):591–623.

Makino, A., Abe, M., Murate, M., Inaba, T., Yilmaz, N., Hullin-Matsuda, F., Kishimoto, T., Schieber, N. L., Taguchi, T., Arai, H., Anderluh, G., Parton, R. G., and Kobayashi, T. (2015). Visualization of the heterogeneous membrane distribution of sphingomyelin associated with cytokinesis, cell polarity, and sphingolipidosis. *FASEB Journal*, 29(2):477–493.

Mathur, A., Feng, S., Hayward, J. A., Ngo, C., Fox, D., Atmosukarto, I. I., Price, J. D., Schauer, K., Märtlbauer, E., Robertson, A. A. B., Burgio, G., Fox, E. M., Leppla, S. H., Kaakoush, N. O., and Man, S. M. (2019). A multicomponent toxin from Bacillus cereus incites inflammation and shapes host outcome via the NLRP3 inflammasome. *Nature Microbiology*, 4(2):362–374.

Matthews, B. (1968). Solvent content of protein crystals. *Journal of Molecular Biology*, 33(2):491–497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, 40(4):658–674.

Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R., and Ban, N. (2009). The structure of a cytolytic a -helical toxin pore reveals its assembly mechanism. *Nature*, 459(7247):726–730.

Munro, S. (2003). Lipid Rafts: Elusive or Illusive?

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method.

Nicholls, D. G. and Ferguson, S. J. S. J. (2013). *Bioenergetics*.

Oscarsson, J., Mizunoe, Y., Li, L., Lai, X. H., Wieslander, Å., and Uhlin, B. E. (1999). Molecular analysis of the cytolytic protein ClyA (SheA) from Escherichia coli. *Molecular Microbiology*, 32(6):1226–1238.

Ouldali, H., Sarthak, K., Ensslen, T., Piguet, F., Manivet, P., Pelta, J., Behrends, J. C., Aksimentiev, A., and Oukhaled, A. (2020). Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore.

Peraro, M. D. and van der Goot, F. G. (2015). Pore-forming toxins: ancient, but never really out of fashion. *Nature reviews. Microbiology*, 14(2):77–92.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612.

Piguet, F., Ouldali, H., Pastoriza-Gallego, M., Manivet, P., Pelta, J., and Oukhaled, A. (2018). Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore. *Nature Communications*, 9(1):966.

Read, R. J. and Schierbeek, A. J. (1988). A phased translation function. *Journal of Applied Crystallography*, 21(5):490–495.

Reboul, C. F., Whisstock, J. C., and Dunstone, M. A. (2016). Giant MACPF/CDC pore forming toxins: A class of their own. *Biochimica et Biophysica Acta - Biomembranes*, 1858(3):475–486.

Roderer, D. and Glockshuber, R. (2017). Assembly mechanism of the α-pore-forming toxin cytolysin A from Escherichia coli. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 372(1726):20160211.

Rojko, N. and Anderluh, G. (2015). How Lipid Membranes Affect Pore Forming Toxin Activity. *Accounts of Chemical Research*, 48(12):3073–3079.

Rojko, N., Dalla Serra, M., Maček, P., and Anderluh, G. (2016). Pore formation by actinoporins, cytolysins from sea anemones. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(3):446–456.

Sastalla, I., Fattah, R., Coppage, N., Nandy, P., Crown, D., Pomerantsev, A. P., and Leppla, S. H. (2013). The Bacillus cereus Hbl and Nhe Tripartite Enterotoxin Components Assemble Sequentially on the Surface of Target Cells and Are Not Interchangeable. *PLoS ONE*, 8(10):e76955.

Sathyanarayana, P., Maurya, S., Behera, A., Ravichandran, M., Visweswariah, S. S., Ayappa, K. G., and Roy, R. (2018). Cholesterol promotes Cytolysin A activity by stabilizing the intermediates during pore formation. *Proceedings of the National Academy of Sciences of the United States of America*, 115(31):E7323–E7330.

Scheres, S. H. W. (2012). A Bayesian view on cryo-EM structure determination. *Journal of molecular biology*, 415(2):406–18.

Schrödinger, LLC (2015). The {PyMOL} Molecular Graphics System, Version˜1.8.

Schubert, E., Vetter, I. R., Prumbaum, D., Penczek, P. A., and Raunser, S. (2018). Membrane insertion of α-xenorhabdolysin in near-atomic detail. *eLife*, 7:1–25.

Sheldrick, G. M. (2010). Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):479–485.

Sigworth, F. J. (2004). Classical detection theory and the cryo-EM particle selection problem. In *Journal of Structural Biology*, volume 145, pages 111–122. Academic Press.

Skubák, P. and Pannu, N. S. (2013). Automatic protein structure solution from weak X-ray data. *Nature Communications*, 4(1):2777.

Tamm, L. K., Hong, H., and Liang, B. (2004). Folding and assembly of β-barrel membrane proteins.

Tanaka, K., Caaveiro, J. M., Morante, K., González-Mañas, J. M., and Tsumoto, K. (2015). Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nature Communications*, 6(1):6337.

Tausch, F., Dietrich, R., Schauer, K., Janowski, R., Niessing, D., Märtlbauer, E., and Jessberger, N. (2017). Evidence for complex formation of the Bacillus cereus haemolysin BL components in solution. *Toxins*, 9(9):288.

Thompson, N. E., Ketterhagen, M. J., Bergdoll, M. S., and Schantz, E. J. (1984). Isolation and some properties of an enterotoxin produced by Bacillus cereus. *Infection and Immunity*, 43(3):887–894.

Van Zundert, G. C., Rodrigues, J. P., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S., Van Dijk, M., De Vries, S. J., and Bonvin, A. M. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology*, 428(4):720–725.

Vigneux, F., Zumbihl, R., Jubelin, G., Ribeiro, C., Poncet, J., Baghdiguian, S., Givaudan, A., and Brehélin, M. (2007). The xaxAB genes encoding a new apoptotic toxin from the insect pathogen Xenorhabdus nematophila are present in plant and human pathogens. *Journal of Biological Chemistry*, 282(13):9571–9580.

Wagner, N. J., Lin, C. P., Borst, L. B., and Millera, V. L. (2013). YaxAB, a Yersinia enterocolitica pore-forming toxin regulated by RovA. *Infection and Immunity*, 81(11):4208–4219.

Wai, S. N., Lindmark, B., Söderblom, T., Takade, A., Westermark, M., Oscarsson, J., Jass, J., Richter-Dahlfors, A., Mizunoe, Y., and Uhlin, B. E. (2003). Vesicle-mediated export and assembly of pore-forming oligomers of the enterobacterial ClyA cytotoxin. *Cell*, 115(1):25–35.

Wallace, A. J., Stillman, T. J., Atkins, A., Jamieson, S. J., Bullough, P. A., Green, J., and Artymiuk, P. J. (2000). E. coli hemolysin E (HlyE, ClyA, SheA): X-ray crystal structure of the toxin and observation of membrane pores by electron microscopy. *Cell*, 100(2):265–76.

Wang, S., Zhao, Z., Haque, F., and Guo, P. (2018). Engineering of protein nanopores for sequencing, chemical or protein sensing and disease diagnosis.

Watson, H. (2015). Biological membranes. *Essays in Biochemistry*, 59:43–70.

WHO (2017). GLOBAL PRIORITY LIST OF ANTIBIOTIC-RESISTANT BACTERIA TO GUIDE RESEARCH, DISCOVERY, AND DEVELOPMENT OF NEW ANTIBIOTICS.

Willems, K., Ruić, D., Biesemans, A., Galenkamp, N. S., Van Dorpe, P., and Maglia, G. (2019). Engineering and Modeling the Electrophoretic Trapping of a Single Protein Inside a Nanopore. *ACS Nano*, 13(9):9980–9992.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science*, 27(1):293–315.

Wilson, J. (2016). *Structural and Functional Studies of the AHL Tripartite Pore Forming Toxin Proteins from Aeromonas hydrophila*. PhD thesis, University of Sheffield.

Wilson, J. S., Churchill-Angus, A. M., Davies, S. P., Sedelnikova, S. E., Tzokov, S. B., Rafferty, J. B., Bullough, P. A., Bisson, C., and Baker, P. J. (2019). Identification and structural analysis of the tripartite α-pore forming toxin of Aeromonas hydrophila. *Nature Communications*, 10(1):2900.

Winter, G. (2010). Xia2: An expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, 43(1):186–190.

Winter, G. and McAuley, K. E. (2011). Automated data collection for macromolecular crystallography. *Methods*, 55(1):81–93.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K., and Evans, G. (2018). DIALS: Implementation and evaluation of a new integration package. *Acta Crystallographica Section D: Structural Biology*, 74:85–97.

Wu, Q. and Guo, Z. (2010). Glycosylphosphatidylinositols are potential targets for the development of novel inhibitors for aerolysin-type of pore-forming bacterial toxins.

Wyborn, N. R., Clark, A., Roberts, R. E., Jamieson, S. J., Tzokov, S., Bullough, P. A., Stillman, T. J., Artymiuk, P. J., Galen, J. E., Zhao, L., Levine, M. M., Green, J., and Green, J. (2016). Properties of haemolysin E (HlyE) from a pathogenic Escherichia coli avian isolate and studies of HlyE export. (2004):1495–1505.

Young, J. A. and Collier, R. J. (2007). Anthrax Toxin: Receptor Binding, Internalization, Pore Formation, and Translocation. *Annual Review of Biochemistry*, 76(1):243–265.

Zheng, W., Cao, H., and Yang, X. (2012). Aeromonas veronii infection in the cultured snakehead fish, Ophiocephalus argus (Cantor). *African Journal of Microbiology Research*, 6(44):7218–7223.

Zhu, K., Didier, A., Dietrich, R., Heilkenbrinker, U., Waltenberger, E., Jessberger, N., Märtlbauer, E., and Benz, R. (2016). Formation of small transmembrane pores: An intermediate stage on the way to Bacillus cereus non-hemolytic enterotoxin (Nhe) full pores in the absence of NheA. *Biochemical and biophysical research communications*, 469(3):613–8.