



**MONASH** University

**Structural Aspects of the Pore Forming  
Protein Perforin**

Cyril F Reboul  
B.Sc. (Honours)

A thesis submitted for the degree of Doctor of Philosophy at  
Monash University in 2015

Department of Biochemistry and Molecular Biology  
School of Biomedical Sciences

## **Copyright notice**

### ***Notice***

© The author 2016. Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.



## Table of Contents

Copyright notice.....	ii
Table of Contents.....	iii
Abstract.....	v
Declaration .....	vi
General declaration .....	vii
List of publications during enrolment.....	viii
Acknowledgements.....	x
Index of figures and tables.....	xi
Abbreviations and symbols.....	xii
 <b>Chapter 1 – Giant MACPF/CDC pore forming toxins: a class of their own.....</b>	 <b>1</b>
Declaration for Thesis Chapter 1.....	2
Preface to Chapter 1 .....	3
 <b>Chapter 2 - Predicting giant transmembrane <math>\beta</math>-barrel architecture .....</b>	 <b>57</b>
Declaration for Thesis Chapter 2.....	58
Preface to Chapter 2 .....	59
 <b>Chapter 3 - A new model of pore formation by CDCs .....</b>	 <b>77</b>
Declaration for Thesis Chapter 3.....	78
Preface to Chapter 3 .....	79

<b>Chapter 4 – The Structure of the poly-C9 component of the Complement</b>	
<b>Membrane Attack Complex.....</b>	<b>107</b>
Declaration for Thesis Chapter 4 .....	108
Preface to Chapter 4 .....	110
Additional Methods .....	144
<b>Chapter 5 – General Discussion.....</b>	<b>149</b>
5.1 Introduction .....	150
5.2 A novel $\beta$ -barrel architecture composes the CDC pore .....	153
5.3 Coordinated domains movements best explain the CDC ‘vertical collapse’ ....	154
5.4 Poly-C9 structure provides significant insights into MAC assembly and MACPF pore formation .....	156
5.5 Concluding remarks .....	159
<b>References .....</b>	<b>165</b>
<b>Appendices .....</b>	<b>171</b>
<b>Appendix A.</b> Stepwise visualization of membrane pore formation by suillysin, a bacterial cholesterol-dependent cytolysin .....	172
<b>Appendix B.</b> Conformational changes during pore formation by the perforin-related protein pleurotolysin.....	191

## Abstract

The Membrane Attack Complex/PerForin (MACPF) / (Cholesterol Dependent Cytolysin (CDC) superfamily encompasses two branches of pore forming proteins (PFPs). The CDC branch is a well-described family of pore forming toxins secreted by pathogenic Gram-positive bacteria. The MACPF branch includes effectors of the vertebrate immune system as well as fungal defence toxins and animal venoms. Both MACPF proteins and CDCs share the ability to form giant membrane embedded pores and a common evolutionary ancestry as they all possess the common MACPF/CDC fold that constitutes the pore forming component of the protein.

As such, the common fold shared by CDCs and MACPF PFPs suggests these proteins use a common mechanism to perforate membranes. Structural and biophysical studies have contributed much to our understanding of the basic mechanism of CDC pore formation. In contrast, less is known about the MACPF branch of the family. Furthermore, the lack of sub-nanometer resolution structure of any MACPF/CDC pore limits our knowledge of the structural basis for pore formation.

In order to address these problems, the CDC model of pore formation was first re-examined employing molecular modelling techniques to understand how the toxins change shape to form pores. The CDC pore transmembrane region was suggested to adopt a novel structure in order to span the membrane. In addition, cryo-Electron Microscopy (cryo-EM) guided modelling resulted in an improved model of the CDC pore consistent with new experimental data.

In the second part of this thesis, a structural model of the MACPF immune effector C9 in the oligomeric form was built based on the cryo-EM maps determined by collaborators. These data permitted an improved understanding of the molecular mechanisms underlying MAC assembly and MACPF pore formation.

Collectively, the results presented in this thesis present a new understanding of the common pore forming mechanism utilised by important bacterial toxins and effectors of the vertebrate immune system. This study further highlights the variations of this common mechanism within both branches of the superfamily.

## **Declaration**

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

## Thesis including published works General Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 2 original papers published in peer reviewed journals and 2 unpublished publications. The core theme of the thesis is *Structural aspects of the pore forming protein perforin*. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the *Department of Biochemistry and Molecular Biology* under the supervision of *Professor James Whisstock*.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of *chapters 1, 2, 3 and 4* my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status	Nature and extent of candidate's contribution
1	Giant MACPF/CDC pore forming toxins: a class of their own	Submitted	60%
2	Predicting giant transmembrane $\beta$ -barrel architecture	Published	60%
3	A new model of pore formation by CDCs	Published	60%
4	Structure of the poly-C9 component of the Complement Membrane Attack Complex	Submitted	30%

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

**Student signature:**



**Date:** 13/11/2015

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student and co-authors' contributions to this work.

**Main Supervisor signature:**



**Date:** 13/11/2015

## List of publications during enrolment

### Publications included in this thesis

1. *Giant MACPF/CDC pore forming toxins: a class of their own*

In revision for *BBA Biomembranes* at the time of submission

**CF Reboul**, JC Whisstock & MA Dunstone

2. *Predicting giant transmembrane beta-barrel architecture*

Bioinformatics (2012), 28(10),1299-302

**CF Reboul**, K Mahmood, JC Whisstock & MA Dunstone

3. *A new model for pore formation by cholesterol-dependent cytolysins*

PLoS Computational Biology (2014), 10(8):e1003791

**CF Reboul**, JC Whisstock & MA Dunstone

4. *Structure of the poly-C9 component of the Complement Membrane Attack Complex*

Revised manuscript submitted to *Nature Communications* at the time of submission

NV Dudkina\*, BA Spicer\*, **CF Reboul**\*, PJ Conroy, N Lukoyanova, H Elmlund, RHP Law, SM Ekkel, SC Kondos, RJA Goode, G Ramm, JC Whisstock, HR Saibil & MA Dunstone

\* The authors contributed equally to the work

### Additional publications included in the thesis appendices

5. *Stepwise visualization of membrane pore formation by sullysin, a bacterial cholesterol-dependent cytolysin*

eLife (2014), 3:e04247

C Leung, NV Dudkina, N Lukoyanova, AW Hodel, I Farabella, AP Pandurangan, D Osmanović, **CF Reboul**, MA Dunstone, M Topf, R Lonnen, PW Andrew, HR Saibil, and BW Hoogenboom.

6. *Conformational changes during pore formation by the perforin-related protein pleurotolysin*

PLoS Biology (2015), 13(2):e1002049

N Lukoyanova, S Kondos, I Farabella, R Law, **CF Reboul**, T Caradoc-Davies, B Spicer, O Kleifed, D Traore, S Ekkel, I Voskoboinik, T Hatfaludi, Oliver K, E Hotze, R Tweten, JC Whisstock, M Topf, H Saibil, MA Dunstone

### Additional publications

7. *Stonefish toxin defines an ancient branch of the perforin-like superfamily*

Proceedings of the National Academy of Sciences, accepted

AM Ellisdon, **CF Reboul**, S Panjekar, K Huynh, CA Oellig, KL Winter, MA Dunstone, WC Hodgson, J Seymour, PK Dearden, RK Tweten, JC Whisstock and S McGowan

8. *Cofactor-dependent conformational heterogeneity of GAD65 and its role in autoimmunity and neurotransmitter homeostasis*  
Proceedings of the National Academy of Sciences (2014), 111(25), E2524-E2529  
I Kass, DE Hoke, MGS Costa, **CF Reboul**, BT Porebski, NP Cowieson, H Leh, E Pennacchietti, J McCoey, O Kleifeld, C Borri Voltattorni, D Langley, B Roome, IR Mackay, D Christ, D Perahia, M Buckle, A Paiardini, D De Biase, AM Buckle
9. *X-ray crystal structure of the streptococcal specific phage lysin PlyC*  
Proceedings of the National Academy of Sciences (2012), 109(31), 12752-12757  
S McGowan, AM Buckle, MS Mitchell, JT Hoopes, DT Gallagher, RD Heselpoth, Y Shen, **CF Reboul**, RHP Law, VA Fischetti, JC Whisstock, DC Nelson.
10. *Structural and dynamic requirements for optimal activity of the essential bacterial enzyme dihydrodipicolinate synthase*  
PLoS Computational Biology (2012), 8(6), e1002537  
**CF Reboul**, BT Porebski, MDW Griffin, RCJ Dobson, MA Perugini, JA Gerrard, AM Buckle
11. *Epitope flexibility and dynamic footprint revealed by molecular dynamics of a pMHC-TCR complex*  
PLoS Computational Biology (2012), 8(3), e1002404  
**CF Reboul**, GR Meyer, BT Porebski, NA Borg, AM Buckle

## Acknowledgments

I would first like to thank my main supervisor Professor James Whisstock. His focus, enthusiasm and mentoring have undoubtedly contributed to shape the scientist I am today. Together with his continuous support and guidance, he has provided a unique environment in which to perform exciting research and for me to strive as a scientist. I would also like to thank him for promoting my exposure to other areas of research.

I would also like to thank my co-supervisors Dr Michelle Dunstone and Associate Professor Ashley Buckle for their ongoing support. Their enthusiasm, availability and mentoring have played a great part during my PhD candidature. They have promoted new scientific opportunities and fostered my independent thinking.

A big thank you goes indeed to current and past members of the Whisstock, Dunstone and Buckle labs, as well as collaborators. Through collaborative work, assistance and discussions during lab hours and afterhours, they have made the time spent working on the present thesis an enjoyable scientific and social experience. I would like to thank in no specific order and forgetting numerous people: Ben Porebski, Siew-Siew Pang, Brad Spicer, Stephanie Kondos, Paul Conroy, Andrew Ellisdon, Ruby Law, Daouda Traore, Helen Saibil, Natalya Dudkina, Natalya Lukoyanova, Sheena McGowan, Sue Ekkel, Itamar Kass, Chantelle Linnett, Chris Lupton, David Hoke, Malcolm Buckle, Christina Lucato, David Perahia, Matthew Perugini.

Finally, I would like to thank my partner Bambi and my daughter Claude for their continuous belief, support and care.



## **Index of figures and tables**

### **Chapter 4**

Figure 4.1 Mechanisms of Poly-C9 and MAC assembly.....	111
Figure 4.2 Structure and sequence alignments of MACPF-domain containing proteins of the human complement .....	144-145

### **Chapter 5**

Figure 5.1 Orientation of the MACPF/CDC domain in the pore form.....	152
Figure 5.2 Revised model of CDC membrane insertion.....	156
Figure 5.3 Schematic model of Poly-C9 elongation and assembly .....	158
Table 5.1 Structural features of MACPF/CDC pore formation.....	159
Figure 5.4 Schematic domains composition across the MACPF/CDC superfamily	161

## Abbreviations

MACPF	Membrane Attack Complex / PerForin
MAC	Membrane Attack Complex
CDC	Cholesterol Dependent Cytolysin
C5	Complement Component 5
C6	Complement Component 6
C7	Complement Component 7
C8	Complement Component 8
C9	Complement Component 9
Poly-C9	C9 polymer
SP Cryo-EM	Single Particle Cryo-Electron Microscopy
AFM	Atomic Force Microscopy
FRET	Förster Resonance Energy Transfer
PLY	Pneumolysin
SLY	Suilyisin
SLO	Streptolysin O
LLO	Listeriolysin O
ILY	Intermedilysin
TMH1/2	TransMembrane Helices 1/2
RBC	Red Blood Cell
CCC	Cross-Correlation Coefficient
NCC	Normalised Correlation Coefficient
MD	Molecular Dynamics
MDFF	Molecular Dynamics Flexible Fitting
RMSD	Root Mean Square Deviation
SSE	Secondary Structural Element
HTH	Helix-Turn-Helix
SDS	Sodium Dodecyl Sulphate
TSP1	Thrombospondin-1
EGF	Epidermal Growth Factor
LDLRA	Low Density Lipoprotein Receptor Associated

## Symbols

Å	ångström
cal	calorie
mol	mole
rad	radian
°	degree





## **Chapter 1**

**Giant MACPF/CDC pore forming toxins:**

**A Class of their Own**

Monash University

## Declaration for Thesis Chapter 1

### Declaration by candidate

In the case of Chapter 1, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Wrote manuscript	60

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
JC Whisstock	Wrote manuscript	
MA Dunstone	Wrote manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

Candidate's Signature		Date 13/11/15
Main Supervisor's Signature		Date 16/11/15

## **Preface to Chapter 1**

This chapter is a literature review of the research conducted on the Membrane-Attack Complex/PerForin (MACPF) / Cholesterol Dependent Cytolysin (CDC) superfamily. The chapter focuses mainly on structural and mechanistic aspects that enable pore-forming superfamily members to change shape, from the secreted soluble monomeric state to the oligomeric membrane-embedded pore state.

The present chapter reproduces an invited review submitted to *BBA Biomembranes* (in revision at the time of submission) and provides an introduction to the following chapters of this thesis. This review is presented in the original manuscript format and includes a discussion of the conclusions reached in the present thesis.

# **Giant MACPF/CDC pore forming toxins: a class of their own**

Cyril F. Reboul<sup>1,2</sup> James Whisstock<sup>1,2</sup> and Michelle A. Dunstone<sup>1,2,3</sup>

<sup>1</sup> Department of Biochemistry and Molecular Biology, Monash University, Melbourne,  
Australia

<sup>2</sup> Australian Research Council Centre of Excellence in Advanced Molecular Imaging,  
Monash University, Melbourne, Australia

<sup>3</sup> Department of Microbiology, Monash University, Melbourne, Australia



## Abstract

Pore Forming Toxins (PFTs) represent a key mechanism for permitting the passage of proteins and small molecules across the lipid membrane. These proteins are typically produced as soluble monomers that self-assemble into ring-like oligomeric structures on the membrane surface. Following such assembly PFTs undergo a remarkable conformational change to insert into the lipid membrane. While many different protein families have independently evolved such ability, members of the Membrane Attack Complex PerForin / Cholesterol Dependent Cytolysin (MACPF/CDC) superfamily form distinctive giant  $\beta$ -barrel pores comprised of up to 50 monomers and up to 300 Å in diameter. In this review we focus on recent advances in understanding the structure of these giant MACPF/CDC pores as well as the underlying molecular mechanisms leading to their formation. Commonalities and evolved variations of the pore forming mechanism across the superfamily are discussed.

## **Contents**

### **1. Introduction**

- 1.1 The CDC family of proteins
- 1.2 MACPF proteins
- 1.3 The MACPF/CDC superfamily distribution
- 1.4 The MACPF/CDC superfamily mechanism of pore formation

### **2. CDC pore formation**

- 2.1 Membrane recognition and binding
- 2.2 Oligomerisation
- 2.3 Pore formation
- 2.4 The final pore structure characteristics
- 2.5 Arciform pores

### **3. MACPF: the MAC, perforin and pleurotolysin**

- 3.1 The Membrane Attack Complex (MAC)
- 3.2 Perforin
- 3.3 The bi-component system pleurotolysin
  - 3.3.1 The structures of PlyA and PlyB
  - 3.3.2 Pore formation by the MACPF pleurotolysin
- 3.4 Prepores/Intermediate steps

### **4. Lessons for MACPF/CDC pore formation**

- 4.1 Summary of the common and different mechanisms
- 4.2 Differences between MACPF and CDC clades
- 4.3 Evolution of the pore forming mechanism

### **5. Summary and future perspectives**

## 1. Introduction

Pore forming toxins (PFTs), also referred to as pore forming proteins, have the ability to breach cell membranes by forming pores in the lipid bilayer. These pores can be either lytic to the target cell, e.g. by osmotic flux, or the pores can mediate the translocation of proteins (typically toxins) into the cytoplasm of the target cell. PFTs are commonly categorised into two groups depending on the elements of secondary structure used to span the cell membrane [1, 2].  $\alpha$ -PFTs utilise amphipathic  $\alpha$ -helices to span the bilayer membrane whereas  $\beta$ -PFTs form amphipathic  $\beta$ -barrel pores. Regardless of the mechanism used, all PFTs are more generally characterised by their remarkable ability to change from water-soluble proteins into integral membrane proteins.

Exemplars of well studied  $\beta$ -PFTs include virulence factors produced by pathogenic bacteria such as the cytolytic Pantone-Valentine leukocidins and  $\alpha$ -hemolysin (from *Staphylococcus aureus*) [3, 4], aerolysin-like toxins (from *Aeromonas hydrophila*) [5], and anthrax protective antigen (from *Bacillus anthracis*) [6]. In most  $\beta$ -PFTs, 7 to 9 monomers self-assemble in a ring-shape oligomer on the cell surface transmembrane  $\beta$ -barrel pore typically less than 25 Å in diameter. In contrast, the Membrane-Attack Complex/PerForin (MACPF)/Cholesterol Dependent Cytolysin (CDC) superfamily form giant  $\beta$ -PFT up to 350 Å in diameter. This unusually large and diverse superfamily is found in all kingdoms of life and performs a wide variety of functions. For example, members of this superfamily are involved in vertebrate immunity [7, 8] venom toxicity [9, 10], development [11], neural development [12] and pathogen invasion / egress [13-15].

## 1.1 The CDC family of proteins

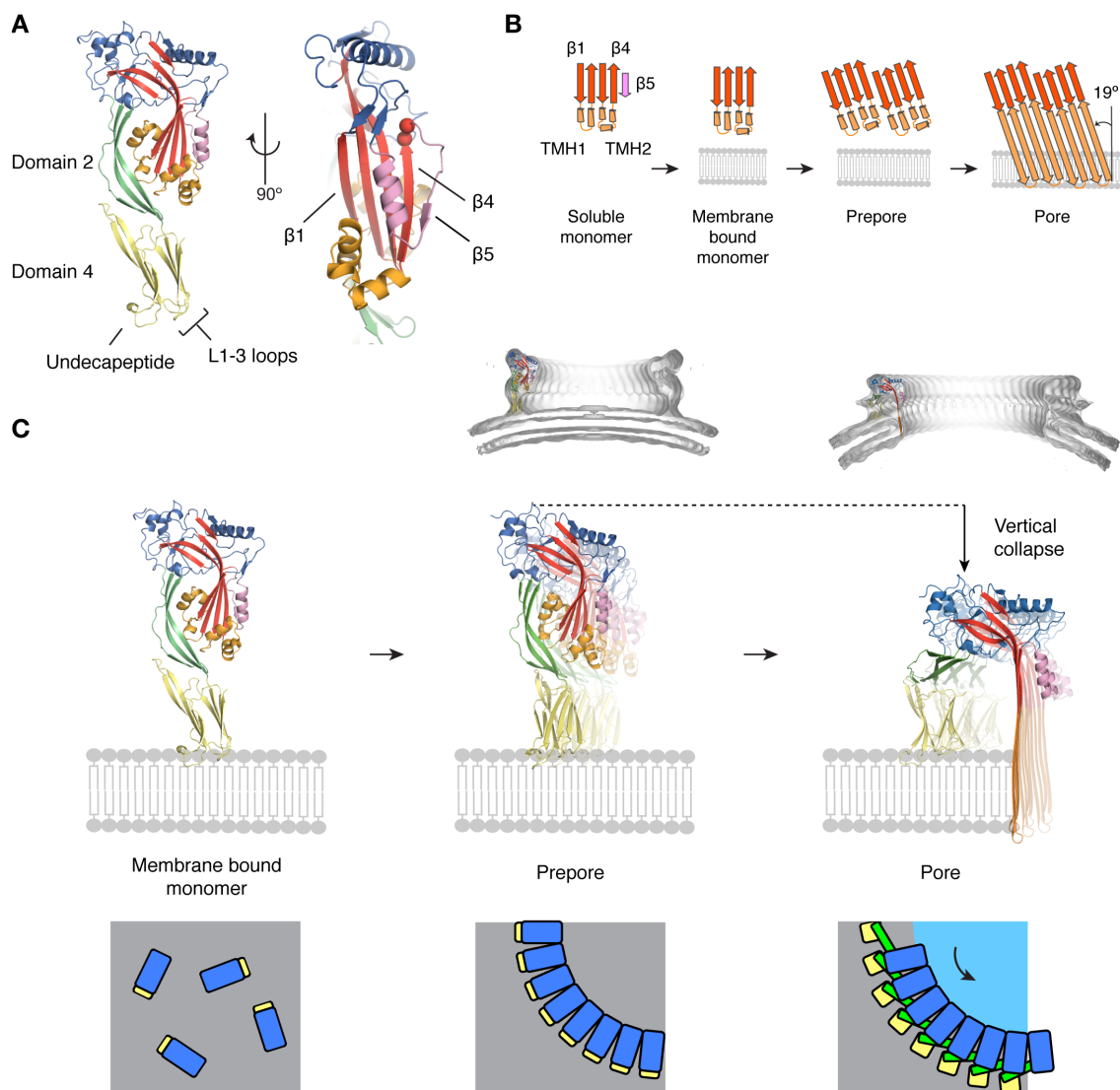
The CDC family of the MACPF/CDC superfamily is a group of toxins produced by a wide range of Gram positive bacteria. Extensive research has determined the role of important CDCs in pathogenicity and also highlighted a range of functions for the CDC pore. The CDC pneumolysin (PLY) is secreted by *Streptococcus pneumoniae*, the major cause of pneumonia and bacterial meningitis, and is suggested to directly mediate cell lysis [16]. Perfringolysin O (PFO, *Clostridia perfringens*) aggravates gas gangrene in synergy with the *C. perfringens*  $\alpha$ -toxin [17]. Listeriolysin O (LLO, *Listeria monocytogenes*) enables phagosomal escape of the bacterium into the macrophage cytosol [18]. Finally, streptolysin O (SLO, *S. pyogenes*) is able to mediate translocation of a cytotoxic effector through the cell membrane without requirement for pore formation [19]. The function of CDCs is not solely limited to pore formation but extends to signalling events leading to a range of effects such as inflammation and adaptive immune responses (for a review see [20]).

Comprehensive studies have shed light on the central conundrum of the cytolysin pore formation: how soluble monomeric CDCs can self-assemble on the target cell surface to form an oligomeric membrane embedded pore. Although, the pores formed by CDCs had been studied using Transmission Electron Microscopy (TEM) since the 1970s [21], a major breakthrough came in 1997 with the X-ray crystal structure of the CDC PFO, which provided the first atomic-resolution insights into the CDC family [22] (Figure 1A). The narrow elongated molecule was identified as being composed of four Domains. Together, the non-contiguous Domains 1 and 3 form a globular domain (Figure 1A). Following the determination of the PFO structure,

biophysical experiments demonstrated that not one but two regions of Domain 3 inserted into the membrane [23, 24].

Over the last 15 years, biochemical and biophysical studies have defined a general three-step mechanism of CDC pore formation [25, 26]. Firstly, the secreted monomeric CDC molecules bind the target cell surface via recognition of a membrane receptor. 30 to 50 monomers then assemble into an oligomeric, ring-shaped, prepore complex [27, 28]. Finally a massive conformational change permits formation of a giant transmembrane  $\beta$ -barrel pore 250 to 350 Å in diameter.

Independent biophysical studies have also demonstrated that a 40 Å vertical collapse of the prepore accompanies conversion of the prepore to the membrane embedded pore [25, 26, 29]. Additionally, two bundles of  $\alpha$ -helices per subunit unfurl from Domain 3 and refold to contribute two amphipathic  $\beta$ -hairpins (4  $\beta$ -strands, Figure 1B) to the pore [23, 24]. Thus, a ring of 30 to 50 CDCs will form a giant  $\beta$ -barrel composed of 120 to 200  $\beta$ -strands.



## Figure 1. The key steps of CDC pore formation

### A. Crystallographic structure of the archetypal CDC PFO (PDB ID: 1PFO) [22].

The MACPF/CDC domain, previously identified as Domain 1 and Domain 3, comprises the characteristic central bent  $\beta$ -sheet (red) sandwiched by the TMH regions (orange). Strand  $\beta$ 5 and the neighbouring  $\alpha$ -helix are in pink. The remainder of the domain is in blue. The pair of conserved glycines is shown as red spheres. Domain 2 is composed of a non-contiguous elongated and twisted  $\beta$ -sheet. Domain 4 (*Ig* fold) is in yellow; the membrane binding L1-3 loops are indicated.

### B. Schematisation of the unfurling and insertion of the TMH regions.

In the soluble state the TMH regions (orange) are present as folded  $\alpha$ -helices. The central  $\beta$ -sheet (red) is expanded by strand  $\beta$ 5 (pink), which is displaced upon membrane binding. In the prepore complex the central  $\beta$ -sheet is postulated to establish oligomeric contacts compatible with the final pore (strands  $\beta$ 1 and  $\beta$ 4). The TMHs also experience some degree of flexibility and unfolding. In the final pore the TMH regions insert into the bilayer membrane as amphipathic  $\beta$ -hairpins with a modest tilt of  $19^\circ$  to form the final pore. The up to 200  $\beta$ -strands that compose the  $\beta$ -barrel extend the central  $\beta$ -sheet.

### C. Mechanism of CDC pore formation.

After membrane anchoring by the Domain 4 loops (left panel), 30 to 50 subunits oligomerise (centre panel) through the flat faces of the MACPF/CDC domain to form the circular prepore complex. The dimensions of the individual subunits in the prepore resemble the monomeric structure. Finally, conversion to the pore (right panel) follows the vertical collapse of the MACPF/CDC domain most likely by coordinated domain movements that see Domain 2 rotate towards the bilayer surface. This brings the MACPF/CDC domains at the required height for TMH insertion as a giant transmembrane  $\beta$ -barrel. The pneumolysin prepore and pore maps are represented as grey surfaces [26], the corresponding protein models are from [30].

Bottom panel: Schematisation of CDC pore formation viewed from the top. The re-orientation of Domain 2 likely results in the coordinated rotation of the ring of MACPF/CDC domains.

Only a quadrant of the ring is represented for clarity.

## 1.2 The MACPF proteins

The MACPF domain was originally characterised as a common pore forming domain present in the immune effectors Membrane Attack Complex (MAC) of the complement system and perforin [31]. The MAC predominantly targets Gram-negative bacteria and is initiated by a complex of proteins C5b, C6, C7 and C8 (C5b-8) on the target membrane. Multiple copies of C9 bind to the C5b-8 complex to form the final MAC pore. Perforin is a cellular immune effector activated by virus-infected cells. In 1999, bioinformatics approaches identified MACPF domains in a diverse range of organisms including *Chlamydia* spp, *Arabidopsis*, sea anemones and fruit flies [32]. Now, the MACPF family is predicted to comprise over 1000 proteins distributed across all kingdoms of life [33].

Prior to any structural information, the MACPF domain was identified as being the site of transmembrane insertion in both the MAC proteins and perforin [8, 34]. In addition, both the MAC and perforin have been well characterised to form giant pores >100 Å in diameter [7, 35-37]. However, a lack of structural information limited our understanding of the detailed mechanism of MACPF pore formation.

In 2007, concomitant elucidations of the structures of the bacterial protein Plu-MACPF (from *Photobacterium luminescens*), and human complement component C8 $\alpha$  MACPF domain revealed the structural similarity between the MACPF domain and the non-contiguous CDC Domains 1 and 3 hereafter referred to as the MACPF/CDC domain (Figures 1A and 2) [38, 39]. This discovery demonstrated that effectors of immune defence and mediators of bacterial attack shared a common ancestor. Moreover, the shared ability to form giant transmembrane pores also suggested a

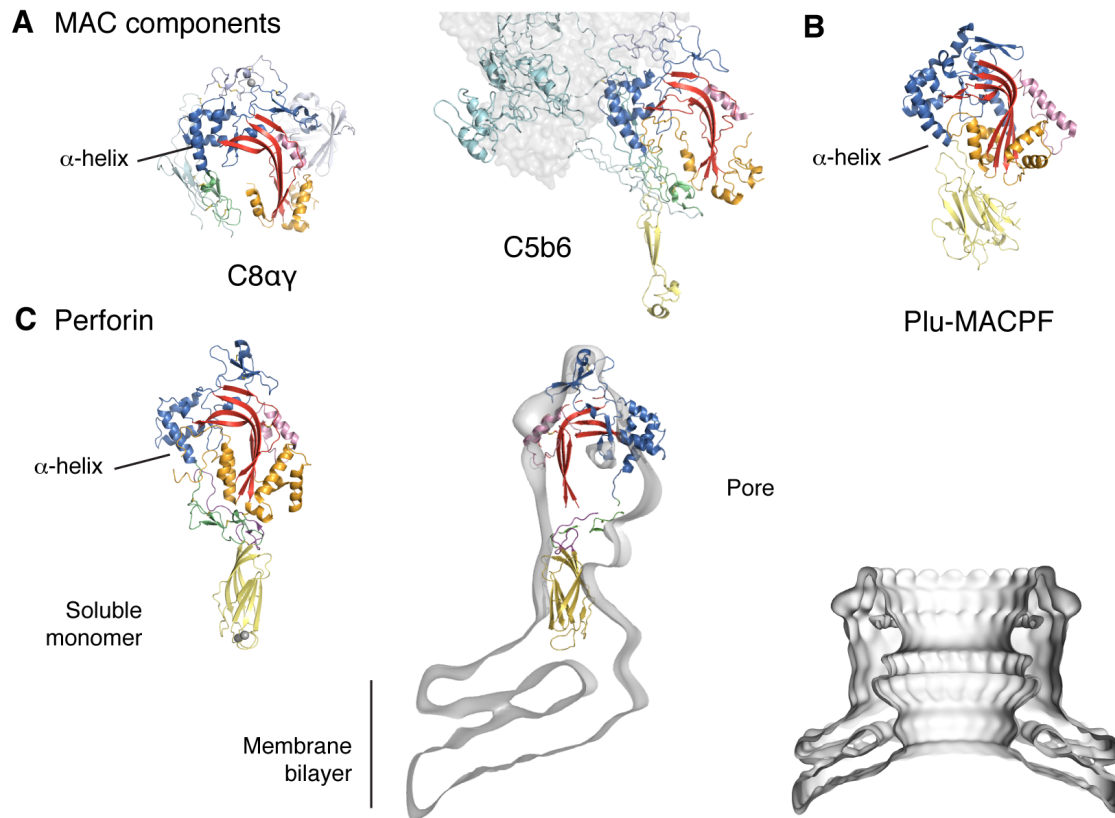


common mechanism of pore formation between both branches of the family. Commonalities in mechanism with the CDC family have now been experimentally supported in via the recent biophysical and structural characterisation of the fungal toxin pleurotolysin pore [40].

### **1.3 The MACPF/CDC superfamily distribution**

Members of the MACPF branch of the MACPF/CDC superfamily are found in a broad range of organisms. These molecules have numerous functions, many of which are poorly understood. To date, MACPF proteins include important effectors of the animal immune system [8, 34, 41], animal venoms (sea anemone [9, 10]) and fungal defence toxins [40, 42]. In all of these systems experimental evidence supports that pore formation is related to their function. Interestingly MACPF proteins have also been identified in the plant immune system [43], fly development [11, 44, 45], vertebrate neural development [12] and parasite ingress/egress across host cell membranes [13, 14]. For these latter proteins it is not known whether pore formation is required for function.

In contrast to the MACPF-like proteins, CDCs are predominantly found in many Gram-positive and several Gram-negative bacteria [46]. It is clear from a structural perspective that the MACPF and CDC members of the MACPF/CDC superfamily share a common evolutionary ancestor. However, at present, it is not clear which of the MACPF or CDC clades most closely resembles the common ancestor. Indeed it is possible that there are other, undiscovered clades of the MACPF/CDC superfamily.



**Figure 2. Structures of the MACPF family**

**A. Representative structures of MAC components.**

**C8 $\alpha\gamma$** , taken from the C8 heterotrimer (PDB ID: 3OJY) [47]. In C8 $\alpha$  four N- and C-terminal ancillary domains decorate the MACPF/CDC domain shown in cyan with the exception of the conserved Epidermal Growth Factor (EGF) domain (green). The disulphide-bridged C8 $\gamma$  (lipocalin fold) is in grey.

**C5b6** (PDB ID: 4A5W) [48]: Nine ancillary domains surround the MACPF/CDC domain of C6 (all coloured light blue except for EGF-like domain (green) and the Thrombospondin 1 (TSP1) domain (yellow)). As observed in the perforin and C8 $\alpha$  structures, C6 has an EGF domain that contacts TMH1 (orange). Several other domains (cyan) form an extensive interface with C5b (shown as a transparent surface).

**B. The bacterial Plu-MACPF** (PDB ID: 2QP2) [38]. The C-terminal membrane-binding domain ( $\beta$ -prism fold; yellow) also packs against the central helical cluster. Plu-MACPF has no reported pore-forming ability.

**C. Perforin** (PDB ID: 3NSJ) [49]. Both putative TMH regions (orange) pack against the ancillary EGF-like domain, followed by a type II C2 domain with calcium-dependent membrane binding ability [50]. The height of the pore is compatible with the dimensions of the monomeric form. Calcium ions are represented as grey spheres.

#### **1.4 The MACPF/CDC superfamily mechanism of pore formation**

Until recently, our knowledge of the whole MACPF/CDC pore forming mechanism was based solely on the well-studied CDC model of pore formation. However, in the past few years structural and biophysical studies have contributed to bridge the gap in our understanding of MACPF pore formation and have highlighted key similarities and differences with CDCs. Furthermore, the unanswered question of how CDC rings collapse towards the membrane has been revisited and new mechanisms proposed. This review aims at describing the three major steps of pore formation across the whole MACPF/CDC superfamily and identifying clade specific variations of the mechanism.

## 2. CDC pore formation

Much of what we know about the MACPF/CDC pore forming mechanism was pioneered by the study of the CDCs PFO, SLO, LLO, PLY and intermedilysin (ILY).

With a total of 10 monomeric crystal structures and two single particle cryo-electron microscopy (SP cryo-EM) reconstructions of both CDC prepores and pores [26, 51], the CDC family is now relatively well structurally characterised. Indeed, the 10 CDC crystal structures currently available have revealed a high degree of structural similarity across the CDC branch of the superfamily [52-56]. At first CDCs were described as being composed of four domains [22], however the non-contiguous Domains 1 and 3 are now better described as the common MACPF/CDC domain.

The core of MACPF/CDC domain consists of a highly twisted and bent three-stranded anti-parallel  $\beta$ -sheet that we term the “central  $\beta$ -sheet”. This central  $\beta$ -sheet is flanked by two clusters of  $\alpha$ -helices termed the TransMembrane Hairpins (TMH) 1 and TMH2 (Figure 1A). The first helical cluster (TMH1) is situated at the core of the domain and packs against Domain 2. The second helical cluster (TMH2) protrudes outwards the MACPF/CDC domain (Figure 1A). The non-contiguous Domain 2 is an elongated and twisted  $\beta$ -sheet that connects the MACPF/CDC domain to the C-terminal Domain 4.

The key pore forming steps that the CDCs undergo include 1) membrane recognition and binding using domain 4, 2) oligomerisation into a prepore and 3) final pore formation (Domain 3). Following membrane binding, 30 to 50 monomers diffuse on the membrane surface to assemble into a membrane-bound ring-shaped oligomer termed the “prepore”. After this oligomerisation step, each  $\alpha$ -helical cluster refolds

into a  $\beta$ -hairpin. The combined unfurling of the TMH1 and TMH2 from all monomers of the prepore ring, together with straightening and untwisting of the central  $\beta$ -sheet, results in a giant, membrane spanning  $\beta$ -barrel pore [24] (Figure 1B). Each step of pore formation is described in details in the next sections.

## **2.1. Membrane recognition and binding**

The CDC C-terminal region, Domain 4, is the highest conserved region of the molecule and displays a  $\beta$ -sandwich fold similar to the immunoglobulin (*Ig*) fold (Figure 1A). A set of four neighbouring loops termed L1-3 and the undecapeptide (Figure 1A) located at the tip of Domain 4 [57] confer CDCs membrane anchoring ability and specificity.

In a majority of CDCs, Domain 4 specifically targets cholesterol-containing lipid bilayers. Indeed, it is suggested that cholesterol constitutes a membrane receptor for most CDCs [58-61]. The specific contacts between the L1-3 loops and cholesterol [62, 63] lead to structural changes in the MACPF/CDC domain that enable oligomerisation of the cytolysin into the prepore complex. The highly conserved undecapeptide motif was also found essential to the conversion of the prepore to the final pore. Mutations in this region result in the abrogation of pore formation [64].

A number of CDCs have been identified to have different membrane receptor preferences. CDCs such as ILY, vaginolysin (VLY) and lectinolysin (LLY), require the cellular receptor human CD59, rather than cholesterol, to recognise and bind the target membrane surface [65-67]. CD59 is a small glycosyl-phosphatidylinositol (GPI)-anchored membrane protein (18-21 kDa). The primary function of CD59 is to

inhibit complement activation, however, in the case of ILY, VLY and LLY, CD59 acts as a receptor. Conversion of these CDCs from the prepore complex to the inserted pore requires both the disengagement of CD59 from Domain 4 and the presence of cholesterol, [65, 68]. The structural basis for the interaction of ILY with CD59 has been recently been determined [69]. Interestingly, it is suggested that CD59 binds not one but two adjacent CDC molecules on non-overlapping sites of Domain 4 [69].

While cholesterol and CD59-dependent CDC membrane binding are now well-documented, recent reports point to the existence of new co-receptors that act to enhance binding to the cell surface. Feil et al. have suggested that LLY uses an additional N-terminal lectin binding domain to target fucose-rich sites cell surfaces [67, 70]. Further, other studies showed that the CDCs PLY and SLO achieve enhanced membrane binding by targeting cellular glycan receptors on RBCs [71, 72]. Therefore, emerging cholesterol-independent modes of membrane recognition seem to co-exist with cholesterol recognition in CDCs. Such alternate ways of membrane addressing are likely to improve membrane affinity and specificity.

## **2.2. Oligomeric assembly: the prepore complex**

After the initial membrane binding event, the CDC molecules oligomerise into a ring-shaped prepore complex [27]. To this end, the CDC molecules undergo an essential structural rearrangement. Direct contact to the membrane triggers the disengagement of strand  $\beta 5$  from the central MACPF/CDC  $\beta$ -sheet [59, 73], a key step in the establishment of oligomeric contacts. The rotation of strand  $\beta 5$  away from the central  $\beta$ -sheet frees the edge of strand  $\beta 4$  (Figure 1A) thereby permitting initial subunits contact and strand-to-strand registration with strand  $\beta 1$  of an adjacent

monomer [59] (Figure 1B). The main-chain hydrogen bonds thus established are required to form the  $\beta$ -barrel formation.

In PFO strand-to-strand registration is exemplified by the formation of an oligomeric Phe-Tyr aromatic interaction. In the prepore complex, Phe318 (strand  $\beta$ 4) pairs with Tyr181 belonging to strand  $\beta$ 1 of an adjacent monomer. Mutation of one of these amino acids abrogates monomer interactions and subsequent prepore formation [59]. The displacement of strand  $\beta$ 5 is likely made possible by the conformational plasticity procured by a pair of strictly conserved glycines directly upstream of strand  $\beta$ 5 [59] (Figure 1A). Thus strand  $\beta$ 5 is hypothesised to occlude unwanted oligomeric contacts in the soluble state, hence preventing aggregation.

Following disengagement of strand  $\beta$ 5, membrane bound CDC subunits assemble to form a large prepore complex prior to membrane insertion [27, 28]. 30 to 50 subunits form this complete ring shape and the monomers interact predominantly via the flat faces of the MACPF/CDC domain. Contacts between the flat faces, that display both shape and charge complementarity, account for most of the prepore subunits interactions [22, 55, 74]. Domains 2 and 4 are thought to contribute only weakly to oligomeric contacts [58]. In the resulting prepore oligomer, the monomers stand upright on the membrane surface with a height of  $\sim 100$  Å, as determined by TEM [26, 51] and Atomic Force Microscopy (AFM) [29]. This height correlates well with the dimension of the X-ray crystal structures of the soluble CDC monomer [22].

Sheppard *et al.* [27] showed that the prepore complex can be further separated into two distinct intermediate states based on its Sodium Dodecyl Sulfate (SDS) resistance: a SDS sensitive early prepore and a SDS resistant late prepore.

Recently, Wade *et al.* [74] suggested that the transition from the early to late prepore is most likely attributable to the formation of the initial oligomeric strands  $\beta 1$ - $\beta 4$  backbone hydrogen bonds that will later become the top of the transmembrane  $\beta$ -barrel pore. Moreover, analysis of the SP cryo-EM structure of the disulphide locked sulysin (SLY) prepore has suggested that the central  $\beta$ -sheet is subjected to slight unbending upon prepore formation while the TMH regions appear partially disordered [51]. This supports previous biochemical experiments on PFO that showed partial unfolding of the TMH regions in the prepore state [75]. The collective picture of the prepore that emerges from these studies is that of a dynamic assembly adopting a continuum of conformations.

This structural heterogeneity present in the prepore is also likely to prime the complex for membrane insertion. Accordingly, the partial untwisting of the central  $\beta$ -sheet together with the oligomeric pairing of strands  $\beta 1$  and  $\beta 4$  constitute a step towards the more open conformation found in the pore form. Similarly, the flexibility and partial unfolding of the TMHs contribute to their final unfurling as  $\beta$ -hairpins. This also highlights the likely lability of the TMHs/Domain 2 interface observed in monomeric crystal structures. The disruption of this interface has been inferred to be one of the first of the series of conformational changes that lead to pore formation in CDCs [30, 76]. Thus, the prepore likely constitutes a nucleation point for the structural changes associated with membrane insertion.

### **2.3. Prepore-to-pore transition**

In the prepore complex the MACPF/CDC domain sits  $\sim 100$  Å atop the membrane surface in a conformation resembling its soluble conformation [26, 51]. Thus, the



MACPF/CDC domain is at such height that even if the two TMH regions were completely unfurled, the corresponding  $\beta$ -hairpins (~30 residues each) can only just reach the membrane surface. In order to pierce the target membrane, the TMHs need to move closer to the membrane. Therefore the prepore-to-pore transition in CDCs must solve the problem of bringing the MACPF/CDC domain closer to the membrane. Accordingly, the prepore to pore transition involves a vertical drop of ~ 40 Å that results in the membrane spanning region being brought closer to the membrane surface [25, 26, 29]. This vertical collapse has previously been attributed to the internal buckling of Domain 2 [26]. However, recent molecular modelling and cryo-EM studies have offered an alternate model for this essential structural change [30, 51]: after disruption of the Domain 2/TMHs interface, intact Domains 2 of the prepore oligomer rotate ~ 60° sideways to change from a perpendicular to near parallel orientation relative to the membrane surface. This has two effects on the CDC prepore assembly. Firstly there is an obvious vertical collapse of the prepore. Secondly the MACPF/CDC domains rotate relative to their associated Domain 4 in an anti-clockwise direction [30] (Figure 1C). Thus the conversion of the prepore to the pore is likely associated with a coordinated domain rotation throughout the oligomer that relocates the MACPF/CDC domain and the TMHs closer to the membrane surface. Only in this position can the folded  $\beta$ -hairpins, extending from the central  $\beta$ -sheet, fully span the target bilayer to form the transmembrane pore.

Concomitant to the prepore-to-pore transition, Domain 4 is also observed to pivot about the membrane surface, and slides away from the centre of the pore [26, 51, 77]. This combined rotation and translation leads to a small expansion of ~ 15 Å of the pore outer diameter [51]. As a result, the MACPF/CDC domain is responsible for the majority of the oligomeric contacts in the final pore. It should be noted that similar

concerted domains rotations are involved in the prepore to pore conversion of the smaller  $\beta$ -PFT aerolysin [78]. This is suggestive that the same pore forming strategy co-evolved to refold from a membrane bound prepore to a membrane-inserted pore state.

The trigger for the prepore-to-pore conversion is to date unclear. Allosteric signals originating from the contact of Domain 4 with the membrane have been proposed as possible trigger to the conversion of the PFO and SLO prepore to the final pore [79, 80]. Additionally, oligomeric contacts of the MACPF/CDC domains in the prepore complex may also promote this conversion in PFO [73, 81].

Finally, it is clear that the transition from the prepore to the pore in CDCs is made possible by a series of potentially energetically costly structural rearrangements. In PFO, the prepore assembles on lipid membranes at 4°C without pore formation. Conversion to the pore is restored after raising the temperature to 37°C [27], suggesting the conversion is an endothermic reaction. The free energy barrier associated with the complex prepore-to-pore structural transitions is estimated to exceed a significant ~19 kcal/mol per monomer [74].

## **2.4 The final pore structure characteristics**

The CDC pore, formed by 120 to 200  $\beta$ -strands and 300-350 Å in diameter, is an order of magnitude larger than the pores deployed by all other  $\beta$ -PFTs, typically comprising 14 to 18  $\beta$ -strands (<25 Å). The large CDC transmembrane channel displays high conductance in planar lipid membrane (PLM) experiments (reviewed in [82]). Consistent with the wide CDC lumen, the SLO and PFO channels allow the

passage of a broad variety of cargos without reported electrostatic selectivity [57, 83, 84].

In addition, the CDC  $\beta$ -barrel pore is characterised by a distinctive structure that sets CDCs further apart from other  $\beta$ -PFTs. Until recently, all  $\beta$ -barrels X-ray crystal structures determined to date were comprised by less than 26  $\beta$ -strands and had a structural architecture ranging from  $S = n$  to  $S = 2n$  [85]. This results in a characteristic tilt of the  $\beta$ -strands with respect to the  $\beta$ -barrel axis ranging from  $37^\circ$  to  $56^\circ$ , respectively [86]. In contrast, the giant CDC  $\beta$ -barrel belongs to a structural architecture previously unidentified ( $S = n/2$ ). This unique architecture, first hypothesised by molecular modelling and confirmed by systematic disulphide scanning, is associated with a modest tilt of the strands of  $19^\circ$  (figure 1B) [75, 87]. Although there is no atomic resolution structures of a CDC pore as yet, the unique  $S = n/2$  architecture has since been observed in the X-ray crystal structure of the bacterial secretin, CsgG, that forms a 36-stranded transmembrane channel with a lumen diameter of  $\sim 40$  Å [88].

## 2.5. Arciform pores

Aside from the observation of prepores and pores, another notable phenomenon is the formation transmembrane channels by incomplete rings known as “arcs”. Bhakdi and colleagues originally proposed that SLO arcs could form membrane inserted crescent shaped oligomers and further proposed that these arcs could be physiologically relevant [89]. These unusual transmembrane channels were thought to be composed of a variable number of CDC subunits lined by two edges: one proteinaceous edge (the incomplete  $\beta$ -barrel) and one lipidic edge formed by the

curled-up membrane bilayer. A cohort of independent studies have supported the existence of incomplete inserted CDCs oligomers with a wide range of techniques such as AFM [29], size exclusion [90], EM [79, 91, 92] and PLM experiments [82].

The existence and role of inserted arcs has remained debated for a number of years. Recent evidence that CDC arcs can disrupt membranes comes from advances in time-resolved AFM applied to SLY [51] and visualisation of such arciform pores by sub-tomogram averaging (PLY) [93]. SLY membrane inserted arcs (e.g. after vertical collapse) composed of a minimum of 5 subunits were directly observed using AFM [51]. It was found that the monomers can elongate a nascent incomplete prepore. Conversion of the crescent-shaped oligomers into membrane-embedded channels was determined to follow the depletion of available membrane-bound subunits. After these arcs have experienced the vertical collapse they can no longer be expanded by addition of SLY monomers and oligomers [51]. Much remains however to be determined about the activity and role of such arciform pores. After phagosomal internalization of the bacterium *L. monocytogenes*, arcs formed by the CDC LLO on the phagosomal membrane delay phagosome maturation and degradation of the bacterium [94].

### 3. MACPF proteins

As the first structures of the MACPF proteins were determined, an outstanding question was whether MACPF proteins formed pores in a CDC-like fashion. Despite numerous studies of MACPF proteins such as perforin and components of the MAC, the structural basis for pore formation in the MACPF branch remained to be determined.

Nevertheless, the first crystallographic structures of the MAC component, C8 $\alpha$ , and the bacterial Plu-MACPF (Figure 2) defined the boundaries of the MACPF/CDC domain and revealed key structural differences between MACPF and CDC proteins [38, 39]. These differences include the lack of strand  $\beta$ 5, substitution of the CDC Domain 2  $\beta$ -sheet with a  $\alpha$ -helix and a different composition and arrangement of domains ancillary to the MACPF/CDC domain (Figures 2 and 3). However, these structural disparities did not permit elucidating the mode of MACPF pore formation.

In the following sections we present current findings that have improved our understanding of the MACPF pore forming mechanism, focusing on the vertebrate immune effectors MAC and perforin. We also focus on the fungal MACPF pleurotolysin, whose recent study has provided the most complete structural understanding of MACPF pore formation [40]. Other members of the MACPF family have relatively little known about their structure and mechanism.

### 3.1. The Membrane-Attack Complex (MAC)

The MAC is a key component of the vertebrate immune system that can target bacteria and parasites [95, 96]. Activation of the complement system leads to the induction of multiple effector pathways including formation of the MAC. The MAC is deposited onto the outer-membrane (in the case of Gram-negative bacteria), which can result in pore formation. Little is known about how the MAC pore lyses Gram-negative bacteria however it is known that the MAC can facilitate the passage of the glycoside hydrolase lysozyme into the peptidoglycan-rich periplasm [97].

The MAC is formed by the assembly of 6 proteins, 5 of which have been originally identified as containing a MACPF/CDC domain by sequence analysis [31]. These proteins are C6, C7, C8 $\alpha$ , C8 $\beta$  and C9. All are separate plasma proteins with the exception of the C8 which is a heterotrimer composed of C8 $\alpha$  covalently bound to C8 $\gamma$ , and non-covalently bound to C8 $\beta$ . Three to nine ancillary modules decorate the common MACPF/CDC domain at both termini, with three of them conserved amongst all five MACPF domain components (C6 – C9). These conserved core domains are the Thrombospondin 1 (TSP 1) domain and Low-Density Lipid Receptor class A (LDLRA) calcium binding region, positioned at the N-terminus of the MACPF/CDC domain, and an EGF domain, positioned at the C-terminus (Figures 2 and 3). The sixth protein, C5b, is the result of the proteolytic cleavage of C5 by the C5 convertase and is critical in initiating MAC assembly [98].

Early TEM observations identified MAC-induced membrane lesions that were estimated to be over 100 Å in diameter and to protrude by ~120 Å from the membrane surface (excluding C5b) [36, 37, 99]. The dimensions of the MAC pore are thus consistent with those observed for CDCs. Bioinformatics studies, based on

X-ray crystal structures, also identified two amphipathic TMH-like regions in the C6-9 components [39]. Only recently have Weiland *et al.* demonstrated that TMH2 from C8 $\alpha$  and C9 is capable of transmembrane insertion [100]. Together these data demonstrated that the MAC forms genuine MACPF/CDC pores. The giant size of the  $\beta$ -barrel indicates that each MAC protein would probably need to contribute two  $\beta$ -hairpins per subunit to form the final pore, and therefore it is possible, but not known, that the MAC also use the  $S = n/2$   $\beta$ -barrel architecture. The MAC  $\beta$ -barrel is yet to be structurally characterised.

The MAC follows a unique and seemingly complex assembly pathway. In the first step of MAC deposition, C5b binds C6 to form C5b6. Subsequent addition of C7 results in the trimeric C5b7, which has been suggested to initiate membrane binding [101]. C8 is then recruited to form C5b8. At this stage, C8 $\alpha$  TMH regions are proposed to penetrate the target membrane outer leaflet for final membrane anchoring, as suggested by fluorescent labelling experiments [102]. Therefore the pore forming regions TMH1 and TMH2 in the MAC seem capable of contacting the target bilayer prior to completion of a ring. Then, an estimated 12 to 18 C9 subunits expand C5b8 to further complete the full MAC pore.

The last decade has seen breakthroughs in the structural biology of the MAC. X-ray crystal structures of C6, C5b6, C8 $\alpha$  and C8 have characterised the structure of the MACPF/CDC fold in these proteins [47, 48, 103, 104]. In addition, 24 Å resolution SP cryo-EM structures of C8 and the soluble form of the C5b-9 complex (sC5b-9, a multimer comprised of C5b8 and several copies of C9) have been determined [48, 105]. Overall, the orientation of C8 $\alpha$  and C8 $\beta$  in the C8 X-ray structure, together with the cryo-EM maps of sC5b-9, strongly supported that the orientation of the

MACPF/CDC domain adopted in the MAC is the same as observed in CDC pores [47, 105].

The combined analysis of C8, C6 and C5b6 crystallographic structures [47, 103, 104] may give some additional insights into the MAC assembly. Initially, the binding of C5b to C6 induces conformational changes throughout the molecules [103]. While some of these changes affect the C5b binding regions, the central  $\beta$ -sheet in C5b6 is found modestly less twisted than in C6 alone [48, 103]. This partial straightening is accompanied by the rotation of four C6 ancillary domains, including the EGF domain and the TSP2-3 domains that flank the MACPF/CDC domain. These structural changes that mirror the first steps of the CDC prepore-to-pore transition have been speculated to prime subsequent MAC assembly [103]. To date, there has been no report of a MAC prepore-like intermediate.

Thus, these recent studies have highlighted features common to both the MAC and CDC pores: MACPF/CDC domain orientation, TMH2 insertion and putative central  $\beta$ -sheet unbending. However, structural discrepancies also suggest major differences in the MAC mechanism of pore formation. The X-ray crystal structures of C6, C5b6 and C8 have shown the presence of a linker  $\alpha$ -helix, as observed in all MACPF proteins structures solved to date, in place of the CDC Domain 2 [47, 48, 103, 104] (Figure 2). In addition, the MAC component proteins lack the same *Ig*-like ancillary domain observed in CDCs that is responsible for binding to membranes. Whilst the MAC component proteins are flanked by between three and nine ancillary domains, little is known about how these contribute, if at all, to membrane binding (Figure 3). In CDCs both Domains 2 and 4 are intrinsically linked to the mechanism of pore formation and condition the final height of the CDC pore. In the MAC, however, this

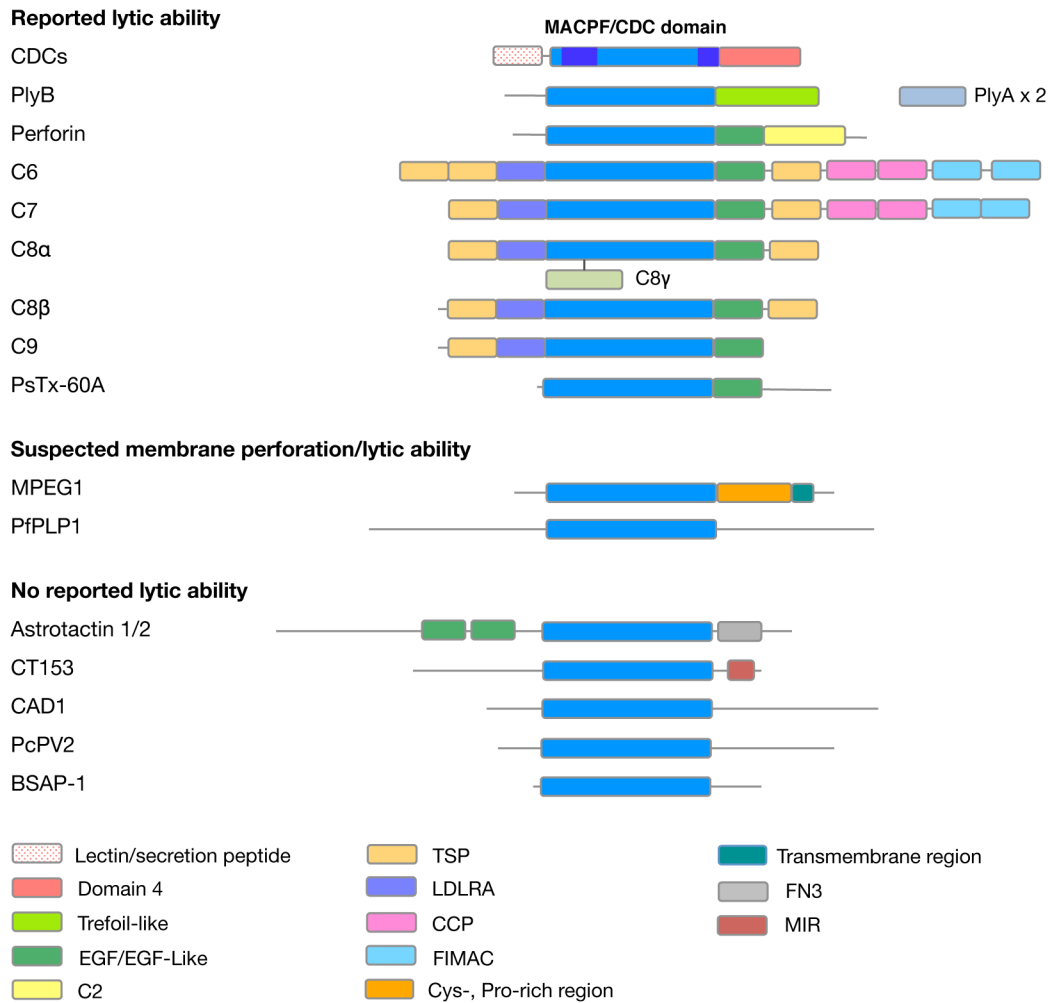


common  $\alpha$ -helix and the lack of an *Ig*-like domain would suggest a pore-forming mechanism exempt of vertical collapse. Additionally, the analogous heights of the C5b6 complex together with the dimension of the inserted MAC further argue against a CDC-like vertical collapse [36, 104].

### **3.2. Perforin**

Perforin, a key effector of the cellular immune system, is secreted by either cytotoxic T lymphocytes or Natural Killer cells. Perforin is stored in specialised granules until activation of the effector cell by virus-infected or neo-cancerous cells. The granules are then transported to the immune synapse where perforin is released. Perforin targets the membrane surface whereupon 18 to 25 monomers self-assemble into large pores ( $\sim 120$  Å inner diameter) in a calcium dependent fashion [35, 49]. There, the perforin pores allow the delivery of a cocktail of pro-apoptotic proteases, called granzymes, into the target cell cytosol. The granzymes activate the apoptosis pathway thereby ensuring removal of the target cell from the host (reviewed in [106]).

A short N-terminal module decorates perforin MACPF/CDC domain. At the C-terminus an EGF-like domain packs against TMH1 followed by a C2 domain [49] and an elongated C-terminal loop (Figure 2). The C2 domain is solely responsible for perforin membrane recognition and binding. Unlike CDCs, perforin does not require a specific membrane surface receptor for efficient anchoring. Little is known about perforin lipid specificity and membrane composition requirements. Although it was initially reported that perforin preferentially targets phosphorylcholine lipid headgroups [107], the prevalence of phospholipid spacing for membrane anchoring was later suggested [108].



**Figure 3. Ancillary domains associated with MACPF/CDC domains**

In CDCs the MACPF/CDC domain includes Domain 2 (dark blue) to form a folded continuous 350 amino-acid structural entity usually referred to as Domains 1 and 3 [52]. CDCs possess a secretion signal anterior to the MACPF/CDC domain up to 150 amino-acids in length. In the CDC LLY the N-terminal lectin domain is suggested to enhance pore formation [70]. Perforin, MAC members C6, C7 C8α, C8β, C9 and the sea anemome toxin PsTx-60A/B [9, 10] have reported lytic ability. MPEG1 (found in animal macrophages) [41] and PfPLP1 (secreted by the parasite *Plasmodium falciparum*) [13] are suspected to exhibit lytic activity. CT153 (of chlamydial origin) [109], PcPV2 (snail egg neurotoxin) [110], CAD1 (part of the plant immune system) [43], BSAP-1 (an antimicrobial protein secreted by the gut symbiot *Bacteroides fragilis*) [111] and astrotactin1/2 (involved in neural development) [12] have no reported pore-forming ability. PcPV2 is disulphide linked to a lectin-like domain susceptible to address the protein to the membrane [110].

A better characterised property of perforin is its strict requirement for calcium to achieve membrane anchoring. The membrane binding C2 domain, specifically its calcium binding region (CBR), is solely responsible for recruiting calcium ions [49, 50, 112, 113], found in abundance at the immune synapse. The CBR is formed by four loops in a position equivalent to the CDC Domain 4 membrane binding loops, and is composed by a network of aspartate amino acids. The CBR is only able to recruit calcium ions at neutral pH when the aspartate residues are fully deprotonated, resulting in perforin membrane binding [50]. Interestingly, prior to secretion into the immune synapse, perforin is stored in specialised secretory granules. The low pH of the granules and low calcium concentration prevents disruption of the vesicle membrane and prevents perforin potentially harmful escape to the host cytosol. Consequently, perforin pH and calcium dependent membrane binding offers the host organism a self-protective mechanism against perforin-mediated cytolysis.

Critical aspects of perforin oligomerisation and pore formation remain elusive. Early TEM imaging showed that perforin forms giant ring-shaped pores (50-160 Å inner diameter), similar to that observed for the MAC [35, 114]. Identification of amphipathic sequences in the TMH1/2 regions and the presence of the MACPF/CDC fold suggested that perforin could form giant  $\beta$ -barrel pores similar to that observed for CDCs [49]. However, the cryo-EM structure of the 20-mer perforin pore together with labelling experiments supported that the MACPF/CDC domain in perforin adopts an orientation reverse to that of CDCs (with the bent portion of the central  $\beta$ -sheet facing outside the pore, Figure 2). The low resolution of the structure however precluded determination of the molecular details of the assembly [49].

To achieve pore formation perforin assembles through the flat faces of the MACPF/CDC domain, as evidenced by the identification of critical salt bridges on these faces [115]. Whether a prepore-like intermediate is formed remains still unclear, although the existence of a CDC-like prepore is suspected [116]. However, what can be concluded with confidence is that the height of the pore was found similar to that of the monomer. This suggests that perforin pore formation is not associated with a CDC-like vertical collapse. This lack of collapse is consistent with the presence of the conserved linker  $\alpha$ -helix in the perforin structure, similarly to MAC components structures, and the lengths of TMH1 and TMH2. In perforin the TMH regions each comprise ~55 residues, while in CDCs they are composed of ~30 residues. These additional residues in perforin could thus allow formation of a longer, higher,  $\beta$ -barrel pore, without necessitating a CDC-like collapse to bring the TMH regions closer to the membrane surface. Although the molecular details of the perforin pore remain to be established, the pore displays selectivity for positively charged proteins [84]. This observation is indeed in agreement with the ability of perforin to deliver positively charged granzymes into the target cell cytosol [117, 118].

Finally, it is suggested that perforin displays, like CDCs, the ability to form transmembrane channels from incomplete rings. Conductance measurements on PLMs [116, 119] were found to correlate with arciform structures as imaged by AFM and deep-etch electron microscope [119].

### 3.3. The bi-component system pleurotolysin

The fungal MACPF pleurotolysin is a bi-component system composed of PlyA (15 kDa) and PlyB (59 kDa) from the mushroom *Pleurotus ostreatus* [42, 120]. Of the two components it is PlyA that was originally identified as a member of the aegerolysin family of PFTs found in fungi and *Clostridia spp* and known to bind to sphingomyelin rich membranes [121]. PlyA is responsible for membrane recognition and binding although PlyA lacks the ability to perforate the lipid membrane compared to aegerolysin homologues [42, 122, 123]. PlyA thus plays an analogous role to that discussed for the CDCs Domain 4. PlyB contains the MACPF/CDC fold essential to oligomerisation and pore formation. Only in concert do PlyA and PlyB exhibit cytolytic activity [42, 120].

The combined use of biophysical, crystallographic and SP cryo-EM studies have recently elucidated the monomeric, pore intermediates and pore structures at an unprecedented level of details in the MACPF/CDC superfamily [40]. Pleurotolysin therefore provides a model to further our understanding of MACPF pore formation and has the potential to shed further light on the CDC pore-forming mechanism.

#### 3.3.1. The structures of PlyA and PlyB

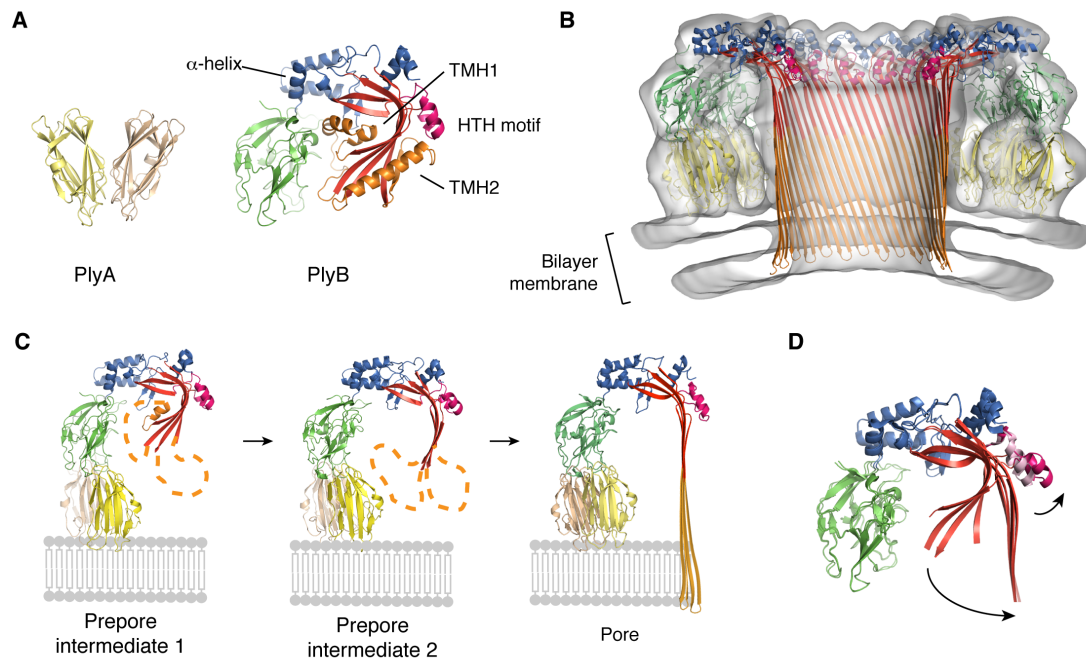
PlyA targets specifically sphingomyelin lipids of cholesterol-rich membranes. PlyA crystal structure shows a  $\beta$ -sandwich bearing structural similarity to the actinoporin family of  $\alpha$ -PFTs, originally identified as sea anemone toxins. However PlyA lacks the N-terminal  $\alpha$ -helical segment known to insert into the membrane [124]. Interestingly the structure of PlyA conclusively unites the aegerolysin family with the

actinoporin family of pore forming toxins. The cryo-EM structure of the pleurotolysin pore also revealed that PlyA functions as a dimer where two PlyA monomers associate to form a V-shaped head-to-head dimer (Figure 4) [40].

PlyB presents the typical MACPF/CDC fold; a bent and twisted four stranded  $\beta$ -sheet surrounded on both sides by one helical cluster (TMH1/2) (Figure 4A). Unlike all other MACPF family protein structures determined to date, PlyB has a fifth  $\beta$ -strand that extends the four stranded central  $\beta$ -sheet. This is similar to CDCs, however, this  $\beta$ -strand is part of TMH2 and contributes directly to the transmembrane  $\beta$ -barrel [40]. TMH1 packs against the central bent  $\beta$ -sheet and the C-terminal domain, a diverged trefoil-like fold.

### **3.3.2. Pore formation by the MACPF pleurotolysin**

The majority of pleurotolysin pores are formed by 13 membrane bound PlyA dimers and 13 PlyB subunits nestled on top of the PlyA dimers. This complex results in a transmembrane channel with an  $\sim 80$  Å inner diameter and is the smallest MACPF/CDC pore identified to date [40, 125]. Interestingly the pleurotolysin pore appears to selectively allow the passive diffusion of small cationic proteins such as granzyme B over small neutral and anionic proteins [84].



**Figure 4. Structural transitions of the fungal MACPF pleurotolysin upon membrane insertion**

**A.** Crystallographic structures of PlyA (PDB ID: 4OEB) and PlyB (PDB ID: 4OEJ). PlyA forms a dimer (yellow and light orange) in the crystallographic and pore cryo-EM structures. The TMH regions of PlyB are highlighted in orange and pack against the central MACPF/CDC  $\beta$ -sheet. The HTH motif (magenta) interacts with the top of TMH2 and the central  $\beta$ -sheet. The remainder of the MACPF/CDC domain is in blue. The  $\alpha$ -helix that connects the C-terminal domain (green) is indicated.

**B.** Thirteen PlyA dimers bind the membrane in a ring shape onto which the thirteen PlyB monomers dock to form the final pore (PDB ID: 4V2T). The central  $\beta$ -sheet is largely open and the TMH regions are inserted into the membrane as amphipathic  $\beta$ -hairpins. The pore fitted model is in cartoon representation within the cryo-EM map (grey surface).

**C.** After membrane binding (central panel) and initial oligomerisation, the TMHs partially unfold (dashed line) and the central sheet opens slightly. Progressive opening of the central sheet is associated with the displacement of the HTH motif and leads the insertion of the  $\beta$ -hairpins into the bilayer. Only the MACPF/CDC domain undergoes conformational changes during pore formation. Prepore intermediate 1 (PDB ID: 4V3M), prepore intermediate 2 (PDB ID: 4V3A),

**D.** Upon pore formation, the central  $\beta$ -sheet opens largely. The HTH motif rotates from the crystallographic (pink) to the pore (magenta) conformations to accommodate such opening. TMH1/2 are omitted for clarity. Crystallographic, prepore intermediates and pore structures are from [40].

The PlyB ancillary C-terminal domain accounts for all the contacts with PlyA and sits in the centre of PlyA V-shaped dimer. This locates the MACPF/CDC domain  $\sim 90$  Å above the membrane surface. The MACPF/CDC domain adopts an orientation similar to that of CDCs, that is, with the bent portion of central  $\beta$ -sheet lining the pore lumen. Single cysteine fluorescent labelling studies further confirmed that TMH1 spans the membrane as an amphipathic  $\beta$ -hairpin and showed that the TMH2 can penetrate the membrane [40].

The SP-cryo EM map of the final pore revealed a remarkably tall  $\beta$ -barrel ( $\sim 100$  Å, Figure 4B) [40], which is much taller than that of CDCs ( $\sim 65$  Å). The height of this  $\beta$ -barrel is comparable in height with the anthrax protective antigen  $\beta$ -barrel pore [6]. Moreover, just as was experimentally elucidated for CDCs, the  $\beta$ -barrel displays the  $S = n/2$  architecture suggesting it is a shared feature of the MACPF/CDC superfamily. Interestingly, most of the interfacial contacts between the proteins within the oligomer are contributed primarily by the  $\beta$ -barrel itself and to a lesser degree by the flat faces of the MACPF/CDC domain [40]. PlyA and the ancillary domains contribute marginally to oligomeric contacts. It is therefore proposed that the MACPF/CDC domain drives oligomerisation.

Another remarkable feature of the pleurotolysin pore is that the major conformational changes required for pore formation are exclusively limited to the MACPF/CDC domain. Comparison of the monomeric and pore conformations identified unambiguously the structural changes associated with pore formation. Like CDCs, these changes include untwisting of the MACPF/CDC central  $\beta$ -sheet and refolding of the TMH regions into  $\beta$ -hairpins. However, in contrast to CDC pore formation where Domains 2 and 4 are subjected to structural changes, the domains ancillary to the



MACPF/CDC domain in pleurotolysin are unaffected by pore formation. Indeed, this demonstrates that pore formation in the model MACPF pleurotolysin deviates significantly from the CDC pore-forming mechanism.

### **3.4 Prepores/Intermediate steps**

In addition to the pleurotolysin pore complex, three SP cryo-EM structures of disulphide linked pore intermediates, akin to the CDC prepore, were further determined [40]. All intermediates exhibited the same stoichiometry and dimensions as the pore. Comparative analysis suggested a pathway towards membrane insertion with the progressive opening of the central  $\beta$ -sheet, oligomeric strand-to-strand registration and gradual unfolding of the TMH regions to finally follow a top-down zippering of the  $\beta$ -barrel. Interestingly, partial unfolding of the TMH regions were identified at an early prepore stage, concomitant with modest unbending of the central  $\beta$ -sheet (Figure 4C). Overall, these findings mirror the plasticity of the two TMH regions in the CDC prepore as discussed in section 2.2.

Further analysis of the prepores dimensions also shed more light on how pleurotolysin inserts into membranes. The observed heights of the prepores and pore are identical thereby excluding a CDC-like collapse. In CDCs, the  $\sim 40$  Å vertical collapse of the  $\sim 100$  Å high prepore provides a mean to insert the otherwise too short  $\beta$ -barrel ( $\sim 65$  Å) into the lipid bilayer. Notably, the pleurotolysin TMH regions are composed of  $\sim 45$  amino acids each, while in CDCs they are composed of  $\sim 30$  amino-acids. Thus the longer TMH1 and TMH2 in pleurotolysin allow both regions to reach further and fully span the membrane bilayer as  $\beta$ -hairpins. Therefore, there is

no requirement for a vertical collapse in pleurotolysin. The MACPF/CDC domain alone undergoes the conformational changes necessary to membrane insertion.

A notable structural change is the displacement of a helix-turn-helix (HTH) motif upon pore formation (Figure 4D). Situated directly above the hinge of sheet unbending, the motif caps the hydrophobic pocket formed by TMH2 and central  $\beta$ -sheet itself in the monomer. In the pore conformation, the HTH motif accommodates the opening of the sheet with a rotation of  $\sim 15^\circ$  to form a circular oligomeric ring directly exposed to the pore lumen. Further molecular modelling of the prepores cryo-EM structures suggested that the motif is involved in triggering the MACPF/CDC domain structural changes. The authors proposed that disruption of the intimate contacts at the HTH/TMH2 interface trigger the release of TMH2 that subsequently enables the essential opening of the central  $\beta$ -sheet [40]. Of note, this structural motif is present in all MACPF proteins structures solved to date (Figure 2). The HTH motif in all CDCs is substituted by a strand  $\beta$ 5-turn-helix motif (Figure 1) where the major  $\alpha$ -helix also forms hydrophobic contacts with TMH2 and the central  $\beta$ -sheet. Thus we hypothesise that the displacement of this motif in CDCs may play a dual role: enabling oligomeric registration of the monomers as well as triggering the release of TMH2 and unbending of the central  $\beta$ -sheet.

## 4. Lessons for MACPF/CDC pore formation

### 4.1. Common features of pore formation

Comparison of proteins of the CDC and MACPF families allows us to identify and discuss structural and mechanistic features shared across the superfamily. These are summarized in Table 1. Arguably, analysis of the commonalities of pore formation by CDCs and the fungal MACPF pleurotolysin are of particular interest as they represent the best described members of the MACPF/CDC superfamily. Such common features are as follows:

- The analogous orientation of the MACPF/CDC domain in the pore, that is, the bent region of the central  $\beta$ -sheets faces the pore lumen.
- The central  $\beta$ -sheet opens and contributes two  $\beta$ -hairpins per monomer (TMH1 and TMH2) to the final  $\beta$ -barrel pore.
- The giant  $\beta$ -barrel is characterised by the distinctive  $S = n/2$  architecture.
- The existence of a membrane bound, ring-shaped, prepore intermediate.

Furthermore, the common MACPF/CDC domain accounts for the majority of interactions between subunits both in the prepore and pore states.

Remarkably, these analogous features are all attributable to the MACPF/CDC domain, which, alone, undergoes similar structural transitions in CDCs, pleurotolysin and the MAC. Although no MAC prepore intermediate has been reported, experimental evidence supports the analogous orientation of the MACPF/CDC

domain and unbending of its central  $\beta$ -sheet [48, 103, 104]. Accordingly, these data reveal the hypothesised common mechanism of MACPF/CDC pore formation.

Despite these comparable features, discrepancies are indeed apparent. The CDC vertical collapse is most likely absent from all the MACPF PFTs described in this review. Thus, an interesting question arises from our comparison: why do these molecules, where the MACPF/CDC domain plays a similar role, adopt such divergent ways to alter membrane integrity? Examination of the singular differences between CDCs and MACPF family members may give us clues to answer this question.

#### **4.2 Differences between MACPF and CDC clades**

Clear differences are the domain composition and their arrangement. In CDCs, the Domains 2 and 4 pair is undoubtedly critical to the well-documented vertical collapse, as discussed in section 2.3. This pair of domains is absent from the hetero-oligomeric MAC and the bi-component pleurotolysin. These two MACPF systems adopt a pore forming mechanism largely different from CDCs. Thus, we speculate that the variations in ancillary domain composition are intrinsically linked to the variations in the pore-forming mechanism. In other words, the most logical hypothesis is that the MACPF/CDC domain encapsulates the core conformational tools for oligomerisation and membrane insertion (overall shape, orientation, central  $\beta$ -sheet unbending, TMHs). In support of this hypothesis, the common MACPF/CDC domain accounts for the majority of the interactions between subunits both in the prepore and pore states. The role of the ancillary domains may be to supply specific membrane binding (e.g. pleurotolysin) or scaffolding (e.g. the MAC) to allow pore formation.

In this light, we note in MACPF PFTs the presence of a common  $\alpha$ -helix in place of the CDC Domain 2 in all crystal structures determined to date (Figure 2). We also note in place of the CDC Domain 4 the variability in domain composition determined from structural and bioinformatics analysis (Figures 2, 3 and 4). We thus further postulate that this range of domain composition reflects possible deviations from the pore-forming mechanisms observed in CDCs and pleurotolysin. MACPF PFTs may thus display variations on the common MACPF/CDC pore-forming mechanism co-evolved with a specific structural scaffold, provided by precise ancillary domains.

Characterising the discrete structural steps and events of the conversion from the membrane-bound oligomer to the final pore has defined a common conformational trajectory of the MACPF/CDC domain in CDCs, pleurotolysin and to some extent the MAC. However, the question of how this conversion is triggered remains. Interestingly, this transition to the pore form occurs, at the superfamily level, regardless of the ancillary domains. Thus, we argue that the oligomeric context provides the MACPF/CDC domain the ability to drive membrane insertion in a top-down direction and dictates the timing of insertion into the target bilayer.

This suggested role of the MACPF/CDC domain in triggering the conversion of the CDC prepore to the pore is in opposition with previous hypothesis. It has been proposed that allosteric signals emerging from contacts of Domain 4 with the membrane surface would act as the CDC membrane insertion trigger [79, 80]. However, further support for this top-down insertion comes from the already established role of the MACPF/CDC domain oligomeric contacts in driving CDC  $\beta$ -barrel insertion [73, 81]. While we cannot rule out a role for allosteric signalling in

triggering the CDC prepore-to-pore conversion, we speculate that the transition is simply initiated by the oligomerisation step.

### 4.3 Evolution of the pore forming mechanism

This comparison leads us to consider the evolution of the MACPF/CDC superfamily. MACPF-like protein structures of the Gram negative *P. luminescens* and *B. thetaioamicron* both have an  $\alpha$ -helix and resemble the MACPF-like toxins in overall sequence and structure [38, 126] (Figure 2). Moreover, there are predicted MACPF-like sequences in bacteria such as *Chlamydia spp* [32]. Likewise, CDC-like sequences (including conservation of Domain 4) have been characterised not just in Gram positive bacteria but also in Gram negative bacteria, such as *Desulfobulbous propionicus*, and in the diatom *Thalassiosira oceanica*, a eukaryotic organism [33]. While it seems likely that these proteins function as pore forming proteins, to date no such activity has been identified for bacterial MACPF proteins.

Overall, this suggests an early duplication of the family into MACPF-like and CDC-like genes that evolved separate mechanisms of pore formation. A key question is did the common evolutionary MACPF/CDC ancestor have a MACPF-like “non-collapsing” mechanism or a CDC-like “collapsing” mechanism? It is too preliminary to conclude either mechanism until there is more research on the structure and mechanism of the bacterial MACPF-like PFTs.

**Table 1. Key features of MACPF/CDC pores and pore-forming mechanisms**

<b>Molecule Feature</b>	<b>CDCs</b>	<b>MAC</b>	<b>Perforin</b>	<b>Pleurotolysin</b>
MACPF/CDC domain orientation: bent $\beta$ -sheet line the pore lumen	Yes	Yes	Proposed “inside-out” orientation	Yes
Two $\beta$ -hairpins per protein	Yes	Possible	Possible	Yes
$S = n/2$ $\beta$ -barrel architecture	Yes	?	?	Yes
Prepore intermediate	Yes	?	Suspected	Yes
Vertical collapse	Yes	Unlikely	Unlikely	No
Linker region/domain	Domain 2 (non-contiguous $\beta$ -sheet)	Linker $\alpha$ -helix	Linker $\alpha$ -helix	Linker $\alpha$ -helix
Ancillary and membrane binding regions	Ig domain (Domain 4)	3 – 9 ancillary domains	EGF-like and C2 domains	Separate protein (PlyA)

## 5. Conclusion and future perspectives

The shared ability of MACPF/CDC members to form large pores together with their common ancestry suggested that MACPF PFTs formed pores in a fashion similar to the well-characterised CDCs. However, recent studies made possible by the advance of biophysical techniques have demonstrated that members of the superfamily have evolved variations of the common conformational strategy (summarized in Table 1).

Features common to the superfamily are conferred by the shared MACPF/CDC fold. They include an analogous domain orientation, mechanism of sheet opening and insertion of two  $\beta$ -hairpins to form a giant  $\beta$ -barrel of distinct architecture, unique among all  $\beta$ -PFTs.

Structural and mechanistic variations have also come to light. The ancillary domains in the MACPF/CDC proteins seem to form an integral component of the clade-specific pore forming mechanism. With CDCs subjected to complex coordinated domains movements compared to the essentially static MACPF pleurotolysin we may have observed both ends of the MACPF/CDC spectrum of pore-forming mechanisms. Given the diverse domain composition in the superfamily and the already suspected role of such domains (MAC) we suggest that pore formation may follow a case-by-case scenario within the superfamily.

An unresolved question common to all MACPF/CDC PFTs is what triggers the insertion of the pore-forming regions? When considering the superfamily as a whole, it is apparent that pore formation occurs either from a prepore intermediate or an incomplete oligomeric ring and is independent of a specific combination of ancillary



or membrane binding domains. We have thus put forward the idea that a possible trigger to pore formation is encapsulated in the common MACPF/CDC domain itself.

The best understood MACPF/CDC proteins form characteristic giant pores, unique in structure and size among all PFTs. Other superfamily members remain to be characterised and still to be determined whether they form pores and how these relate to their respective physiological role, including immunity, toxicity, development and neural development. It is however, suggested from the common structure of the MACPF/CDC fold that if the proteins do form pores these pores will be giant  $\beta$ -PFT typical of the MACPF/CDC superfamily.

## Acknowledgments

CFR is supported by the Monash University FMNHS Postgraduate Bridging Fellowship Award.

MAD acknowledges support from the Australian Research Council (ARC) [DP120104058, CE140100011] and the National Health and Medical Research Council (NHMRC) [606471].

JCW acknowledges the ARC and NHMRC [FF0883418, DP120104058, CE140100011]

## References

1. Lesieur, C., et al., *Membrane insertion: The strategies of toxins (review)*. Mol Membr Biol, 1997. **14**(2): p. 45-64.
2. Gouaux, E., *Channel-forming toxins: tales of transformation*. Curr Opin Struct Biol, 1997. **7**(4): p. 566-73.
3. Vandenesch, F., G. Lina, and T. Henry, *Staphylococcus aureus hemolysins, bi-component leukocidins, and cytolytic peptides: a redundant arsenal of membrane-damaging virulence factors?* Front Cell Infect Microbiol, 2012. **2**: p. 12.
4. Menestrina, G., et al., *Ion channels and bacterial infection: the case of beta-barrel pore-forming protein toxins of Staphylococcus aureus*. FEBS Lett, 2003. **552**(1): p. 54-60.
5. Iacovache, I., et al., *A rivet model for channel formation by aerolysin-like pore-forming toxins*. EMBO J, 2006. **25**(3): p. 457-66.
6. Jiang, J., et al., *Atomic structure of anthrax protective antigen pore elucidates toxin translocation*. Nature, 2015.
7. Young, J.D., et al., *Extracellular release of lymphocyte cytolytic pore-forming protein (perforin) after ionophore stimulation*. Proc Natl Acad Sci U S A, 1986. **83**(15): p. 5668-72.

8. Lichtenheld, M.G., et al., *Structure and function of human perforin*. Nature, 1988. **335**(6189): p. 448-51.
9. Nagai, H., et al., *Novel proteinaceous toxins from the nematocyst venom of the Okinawan sea anemone *Phyllodiscus semoni* Kwietniewski*. Biochem Biophys Res Commun, 2002. **294**(4): p. 760-3.
10. Satoh, H., et al., *Characterization of PsTX-60B, a new membrane-attack complex/perforin (MACPF) family toxin, from the venomous sea anemone *Phyllodiscus semoni**. Toxicon, 2007. **49**(8): p. 1208-10.
11. Martin, J.R., A. Raibaud, and R. Olo, *Terminal pattern elements in *Drosophila* embryo induced by the torso-like protein*. Nature, 1994. **367**(6465): p. 741-5.
12. Zheng, C., N. Heintz, and M.E. Hatten, *CNS gene encoding astrotactin, which supports neuronal migration along glial fibers*. Science, 1996. **272**(5260): p. 417-9.
13. Garg, S., et al., *Calcium-dependent permeabilization of erythrocytes by a perforin-like protein during egress of malaria parasites*. Nat Commun, 2013. **4**: p. 1736.
14. Kadota, K., et al., *Essential role of membrane-attack protein in malarial transmission to mosquito host*. Proc Natl Acad Sci U S A, 2004. **101**(46): p. 16310-5.
15. Wade, K.R. and R.K. Tweten, *The Apicomplexan CDC/MACPF-like pore-forming proteins*. Curr Opin Microbiol, 2015. **26**: p. 48-52.
16. Hirst, R.A., et al., *The role of pneumolysin in pneumococcal pneumonia and meningitis*. Clin Exp Immunol, 2004. **138**(2): p. 195-201.
17. Awad, M.M., et al., *Synergistic effects of alpha-toxin and perfringolysin O in *Clostridium perfringens*-mediated gas gangrene*. Infect Immun, 2001. **69**(12): p. 7904-10.
18. Hamon, M.A., et al., *Listeriolysin O: the Swiss army knife of *Listeria**. Trends Microbiol, 2012. **20**(8): p. 360-8.
19. Madden, J.C., N. Ruiz, and M. Caparon, *Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria*. Cell, 2001. **104**(1): p. 143-52.
20. Cassidy, S.K. and M.X. O'Riordan, *More than a pore: the cellular response to cholesterol-dependent cytolysins*. Toxins (Basel), 2013. **5**(4): p. 618-36.
21. Smyth, C.J., J.H. Freer, and J.P. Arbuthnott, *Interaction of *Clostridium perfringens* theta-haemolysin, a contaminant of commercial phospholipase C, with erythrocyte ghost membranes and lipid dispersions. A morphological study*. Biochim Biophys Acta, 1975. **382**(4): p. 479-93.
22. Rossjohn, J., et al., *Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form*. Cell, 1997. **89**(5): p. 685-92.

23. Shepard, L.A., et al., *Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin Clostridium perfringens perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy*. Biochemistry, 1998. **37**(41): p. 14563-74.
24. Shatursky, O., et al., *The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins*. Cell, 1999. **99**(3): p. 293-9.
25. Ramachandran, R., R.K. Tweten, and A.E. Johnson, *The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation*. Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7139-44.
26. Tilley, S.J., et al., *Structural basis of pore formation by the bacterial toxin pneumolysin*. Cell, 2005. **121**(2): p. 247-56.
27. Shepard, L.A., et al., *The mechanism of pore assembly for a cholesterol-dependent cytolysin: formation of a large prepore complex precedes the insertion of the transmembrane beta-hairpins*. Biochemistry, 2000. **39**(33): p. 10284-93.
28. Hotze, E.M., et al., *Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane beta-sheet from a prepore intermediate*. J Biol Chem, 2001. **276**(11): p. 8261-8.
29. Czajkowsky, D.M., et al., *Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane*. EMBO J, 2004. **23**(16): p. 3206-15.
30. Reboul, C.F., J.C. Whisstock, and M.A. Dunstone, *A new model for pore formation by cholesterol-dependent cytolysins*. PLoS Comput Biol, 2014. **10**(8): p. e1003791.
31. Tschopp, J., D. Masson, and K.K. Stanley, *Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytotoxicity*. Nature, 1986. **322**(6082): p. 831-4.
32. Ponting, C.P., *Chlamydial homologues of the MACPF (MAC/perforin) domain*. Curr Biol, 1999. **9**(24): p. R911-3.
33. Anderluh, G., et al., *Distribution of MACPF/CDC proteins*. Subcell Biochem, 2014. **80**: p. 7-30.
34. Stanley, K.K. and J. Herz, *Topological mapping of complement component C9 by recombinant DNA techniques suggests a novel mechanism for its insertion into target membranes*. EMBO J, 1987. **6**(7): p. 1951-7.
35. Young, J.D., et al., *Purification and characterization of a cytolytic pore-forming protein from granules of cloned lymphocytes with natural killer activity*. Cell, 1986. **44**(6): p. 849-59.

36. Tschopp, J., E.R. Podack, and H.J. Muller-Eberhard, *Ultrastructure of the membrane attack complex of complement: detection of the tetramolecular C9-polymerizing complex C5b-8*. Proc Natl Acad Sci U S A, 1982. **79**(23): p. 7474-8.
37. Bhakdi, S. and J. Tranum-Jensen, *Molecular nature of the complement lesion*. Proc Natl Acad Sci U S A, 1978. **75**(11): p. 5655-9.
38. Rosado, C.J., et al., *A common fold mediates vertebrate defense and bacterial attack*. Science, 2007. **317**(5844): p. 1548-51.
39. Hadders, M.A., D.X. Beringer, and P. Gros, *Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense*. Science, 2007. **317**(5844): p. 1552-4.
40. Lukyanova, N., et al., *Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin*. PLoS Biol, 2015. **13**(2): p. e1002049.
41. McCormack, R., et al., *Inhibition of intracellular bacterial replication in fibroblasts is dependent on the perforin-like protein (perforin-2) encoded by macrophage-expressed gene 1*. J Innate Immun, 2013. **5**(2): p. 185-94.
42. Sakurai, N., et al., *Cloning, expression, and pore-forming properties of mature and precursor forms of pleurotolysin, a sphingomyelin-specific two-component cytolytic from the edible mushroom Pleurotus ostreatus*. Biochim Biophys Acta, 2004. **1679**(1): p. 65-73.
43. Morita-Yamamuro, C., et al., *The Arabidopsis gene CAD1 controls programmed cell death in the plant immune system and encodes a protein containing a MACPF domain*. Plant Cell Physiol, 2005. **46**(6): p. 902-12.
44. Henstridge, M.A., et al., *Trunk cleavage is essential for Drosophila terminal patterning and can occur independently of Torso-like*. Nat Commun, 2014. **5**: p. 3419.
45. Johnson, T.K., et al., *Torso-like functions independently of Torso to regulate Drosophila growth and developmental timing*. Proc Natl Acad Sci U S A, 2013. **110**(36): p. 14688-92.
46. Hotze, E.M., et al., *Identification and characterization of the first cholesterol-dependent cytolyticins from Gram-negative bacteria*. Infect Immun, 2013. **81**(1): p. 216-25.
47. Lovelace, L.L., et al., *Structure of human C8 protein provides mechanistic insight into membrane pore formation by complement*. J Biol Chem, 2011. **286**(20): p. 17585-92.
48. Hadders, M.A., et al., *Assembly and regulation of the membrane attack complex based on structures of C5b6 and sC5b9*. Cell Rep, 2012. **1**(3): p. 200-7.
49. Law, R.H., et al., *The structural basis for membrane binding and pore formation by lymphocyte perforin*. Nature, 2010. **468**(7322): p. 447-51.

50. Voskoboinik, I., et al., *Calcium-dependent plasma membrane binding and cell lysis by perforin are mediated through its C2 domain: A critical role for aspartate residues 429, 435, 483, and 485 but not 491*. J Biol Chem, 2005. **280**(9): p. 8426-34.
51. Leung, C., et al., *Stepwise visualization of membrane pore formation by suilysin, a bacterial cholesterol-dependent cytotoxin*. Elife, 2014. **3**.
52. Bourdeau, R.W., et al., *Cellular functions and X-ray structure of anthrolysin O, a cholesterol-dependent cytotoxin secreted by Bacillus anthracis*. J Biol Chem, 2009. **284**(21): p. 14645-56.
53. Xu, L., et al., *Crystal structure of cytotoxin protein suilysin from Streptococcus suis*. Protein Cell, 2010. **1**(1): p. 96-105.
54. Polekhina, G., et al., *Insights into the action of the superfamily of cholesterol-dependent cytotoxins from studies of intermedilysin*. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 600-5.
55. Koster, S., et al., *Crystal structure of listeriolysin O reveals molecular details of oligomerization and pore formation*. Nat Commun, 2014. **5**: p. 3690.
56. Feil, S.C., et al., *Structural studies of Streptococcus pyogenes streptolysin O provide insights into the early steps of membrane penetration*. J Mol Biol, 2014. **426**(4): p. 785-92.
57. Heuck, A.P., R.K. Tweten, and A.E. Johnson, *Assembly and topography of the prepore complex in cholesterol-dependent cytotoxins*. J Biol Chem, 2003. **278**(33): p. 31218-25.
58. Ramachandran, R., et al., *Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytotoxin*. Nature Structural Biology, 2002. **9**(11): p. 823-827.
59. Ramachandran, R., R.K. Tweten, and A.E. Johnson, *Membrane-dependent conformational changes initiate cholesterol-dependent cytotoxin oligomerization and intersubunit beta-strand alignment*. Nature structural & molecular biology, 2004. **11**(8): p. 697-705.
60. Flanagan, J.J., et al., *Cholesterol exposure at the membrane surface is necessary and sufficient to trigger perfringolysin O binding*. Biochemistry, 2009. **48**(18): p. 3977-87.
61. Bavdek, A., et al., *Sterol and pH interdependence in the binding, oligomerization, and pore formation of Listeriolysin O*. Biochemistry, 2007. **46**(14): p. 4425-37.
62. Soltani, C.E., et al., *Structural elements of the cholesterol-dependent cytotoxins that are responsible for their cholesterol-sensitive membrane interactions*. Proc Natl Acad Sci U S A, 2007. **104**(51): p. 20226-31.
63. Farrand, A.J., et al., *Only two amino acids are essential for cytolytic toxin recognition of cholesterol at the membrane surface*. Proc Natl Acad Sci U S A, 2010. **107**(9): p. 4341-6.

64. Dowd, K.J. and R.K. Tweten, *The cholesterol-dependent cytolysin signature motif: a critical element in the allosteric pathway that couples membrane binding to pore assembly*. PLoS Pathog, 2012. **8**(7): p. e1002787.
65. Giddings, K.S., et al., *Human CD59 is a receptor for the cholesterol-dependent cytolysin intermedilysin*. Nat Struct Mol Biol, 2004. **11**(12): p. 1173-8.
66. Zilnyte, M., et al., *The Cytolytic Activity of Vaginolysin Strictly Depends on Cholesterol and Is Potentiated by Human CD59*. Toxins (Basel), 2015. **7**(1): p. 110-28.
67. Wickham, S.E., et al., *Mapping the intermedilysin-human CD59 receptor interface reveals a deep correspondence with the binding site on CD59 for complement binding proteins C8alpha and C9*. J Biol Chem, 2011. **286**(23): p. 20952-62.
68. LaChapelle, S., R.K. Tweten, and E.M. Hotze, *Intermedilysin-receptor interactions during assembly of the pore complex: assembly intermediates increase host cell susceptibility to complement-mediated lysis*. J Biol Chem, 2009. **284**(19): p. 12719-26.
69. Johnson, S., et al., *Structural basis for recognition of the pore-forming toxin intermedilysin by human complement receptor CD59*. Cell Rep, 2013. **3**(5): p. 1369-77.
70. Feil, S.C., et al., *Structure of the lectin regulatory domain of the cholesterol-dependent cytolysin lectinolysin reveals the basis for its lewis antigen specificity*. Structure, 2012. **20**(2): p. 248-58.
71. Shewell, L.K., et al., *The cholesterol-dependent cytolysins pneumolysin and streptolysin O require binding to red blood cell glycans for hemolytic activity*. Proc Natl Acad Sci U S A, 2014. **111**(49): p. E5312-20.
72. Mozola, C.C., N. Magassa, and M.G. Caparon, *A novel cholesterol-insensitive mode of membrane binding promotes cytolysin-mediated translocation by Streptolysin O*. Mol Microbiol, 2014. **94**(3): p. 675-87.
73. Hotze, E.M., et al., *Monomer-monomer interactions propagate structural transitions necessary for pore formation by the cholesterol-dependent cytolysins*. J Biol Chem, 2012. **287**(29): p. 24534-43.
74. Wade, K.R., et al., *An intermolecular electrostatic interaction controls the prepore-to-pore transition in a cholesterol-dependent cytolysin*. Proc Natl Acad Sci U S A, 2015. **112**(7): p. 2204-9.
75. Sato, T.K., R.K. Tweten, and A.E. Johnson, *Disulfide-bond scanning reveals assembly state and beta-strand tilt angle of the PFO beta-barrel*. Nat Chem Biol, 2013. **9**(6): p. 383-9.
76. Rossjohn, J., et al., *Structures of perfringolysin O suggest a pathway for activation of cholesterol-dependent cytolysins*. J Mol Biol, 2007. **367**(5): p. 1227-36.

77. Ramachandran, R., et al., *Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin*. Nat Struct Biol, 2002. **9**(11): p. 823-7.
78. Degiacomi, M.T., et al., *Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism*. Nat Chem Biol, 2013. **9**(10): p. 623-9.
79. Palmer, M., et al., *Streptolysin O: a proposed model of allosteric interaction between a pore-forming protein and its target lipid bilayer*. Biochemistry, 1998. **37**(8): p. 2378-83.
80. Heuck, A.P., et al., *Mechanism of membrane insertion of a multimeric beta-barrel protein: perfringolysin O creates a pore using ordered and coupled conformational changes*. Mol Cell, 2000. **6**(5): p. 1233-42.
81. Hotze, E.M., et al., *Monomer-monomer interactions drive the prepore to pore conversion of a beta-barrel-forming cholesterol-dependent cytolysin*. J Biol Chem, 2002. **277**(13): p. 11597-605.
82. Marchiorretto, M., et al., *What planar lipid membranes tell us about the pore-forming activity of cholesterol-dependent cytolysins*. Biophys Chem, 2013. **182**: p. 64-70.
83. Menestrina, G., C.L. Bashford, and C.A. Pasternak, *Pore-forming toxins: experiments with S. aureus alpha-toxin, C. perfringens theta-toxin and E. coli haemolysin in lipid bilayers, liposomes and intact cells*. Toxicon, 1990. **28**(5): p. 477-91.
84. Stewart, S.E., et al., *The Perforin Pore Facilitates the Delivery of Cationic Cargos*. J Biol Chem, 2014.
85. Nagano, N., E.G. Hutchinson, and J.M. Thornton, *Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels*. Protein Sci, 1999. **8**(10): p. 2072-84.
86. Murzin, A.G., A.M. Lesk, and C. Chothia, *Principles determining the structure of beta-sheet barrels in proteins. I. A theoretical analysis*. J Mol Biol, 1994. **236**(5): p. 1369-81.
87. Reboul, C.F., et al., *Predicting giant transmembrane beta-barrel architecture*. Bioinformatics, 2012. **28**(10): p. 1299-302.
88. Goyal, P., et al., *Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG*. Nature, 2014. **516**(7530): p. 250-3.
89. Bhakdi, S., J. Tranum-Jensen, and A. Sziegoleit, *Mechanism of membrane damage by streptolysin-O*. Infect Immun, 1985. **47**(1): p. 52-60.
90. Palmer, M., et al., *Assembly mechanism of the oligomeric streptolysin O pore: the early membrane lesion is lined by a free edge of the lipid membrane and is extended gradually during oligomerization*. EMBO J, 1998. **17**(6): p. 1598-605.



91. Morgan, P.J., et al., *Modeling the bacterial protein toxin, pneumolysin, in its monomeric and oligomeric form*. J Biol Chem, 1994. **269**(41): p. 25315-20.
92. Olofsson, A., H. Hebert, and M. Thelestam, *The projection structure of perfringolysin O (Clostridium perfringens theta-toxin)*. FEBS Lett, 1993. **319**(1-2): p. 125-7.
93. Sonnen, A.F., J.M. Plitzko, and R.J. Gilbert, *Incomplete pneumolysin oligomers form membrane pores*. Open Biol, 2014. **4**: p. 140044.
94. Shaughnessy, L.M., et al., *Membrane perforations inhibit lysosome fusion by altering pH and calcium in Listeria monocytogenes vacuoles*. Cell Microbiol, 2006. **8**(5): p. 781-92.
95. Pearson, R.D. and R.T. Steigbigel, *Mechanism of lethal effect of human serum upon Leishmania donovani*. J Immunol, 1980. **125**(5): p. 2195-201.
96. Joiner, K.A., E.J. Brown, and M.M. Frank, *Complement and bacteria: chemistry and biology in host defense*. Annu Rev Immunol, 1984. **2**: p. 461-91.
97. Martinez, R.J. and S.F. Carroll, *Sequential metabolic expressions of the lethal process in human serum-treated Escherichia coli: role of lysozyme*. Infect Immun, 1980. **28**(3): p. 735-45.
98. Morgan, B.P., *The membrane attack complex as an inflammatory trigger*. Immunobiology, 2015.
99. Steckel, E.W., B.E. Welbaum, and J.M. Sodetz, *Evidence of direct insertion of terminal complement proteins into cell membrane bilayers during cytolysis. Labeling by a photosensitive membrane probe reveals a major role for the eighth and ninth components*. J Biol Chem, 1983. **258**(7): p. 4318-24.
100. Weiland, M.H., Y. Qian, and J.M. Sodetz, *Membrane pore formation by human complement: functional importance of the transmembrane beta-hairpin (TMH) segments of C8alpha and C9*. Mol Immunol, 2014. **57**(2): p. 310-6.
101. Marshall, P., et al., *Interaction between complement proteins C5b-7 and erythrocyte membrane sialic acid*. J Exp Med, 1996. **184**(4): p. 1225-32.
102. Preissner, K.T., E.R. Podack, and H.J. Muller-Eberhard, *The membrane attack complex of complement: relation of C7 to the metastable membrane binding site of the intermediate complex C5b-7*. J Immunol, 1985. **135**(1): p. 445-51.
103. Aleshin, A.E., et al., *Crystal structure of C5b-6 suggests structural basis for priming assembly of the membrane attack complex*. J Biol Chem, 2012. **287**(23): p. 19642-52.
104. Aleshin, A.E., et al., *Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC)*. J Biol Chem, 2012. **287**(13): p. 10210-22.

105. Bubeck, D., et al., *Structure of human complement C8, a precursor to membrane attack*. J Mol Biol, 2011. **405**(2): p. 325-30.
106. Voskoboinik, I., J.C. Whisstock, and J.A. Trapani, *Perforin and granzymes: function, dysfunction and human pathology*. Nat Rev Immunol, 2015.
107. Tschopp, J., et al., *Phosphorylcholine acts as a Ca<sup>2+</sup>-dependent receptor molecule for lymphocyte perforin*. Nature, 1989. **337**(6204): p. 272-4.
108. Antia, R., R.A. Schlegel, and P. Williamson, *Binding of perforin to membranes is sensitive to lipid spacing and not headgroup*. Immunol Lett, 1992. **32**(2): p. 153-7.
109. Taylor, L.D., et al., *Biological characterization of Chlamydia trachomatis plasticity zone MACPF domain family protein CT153*. Infect Immun, 2010. **78**(6): p. 2691-9.
110. Dreon, M.S., et al., *Novel animal defenses against predation: a snail egg neurotoxin combining lectin and pore-forming chains that resembles plant defense and bacteria attack toxins*. PLoS One, 2013. **8**(5): p. e63782.
111. Chatzidaki-Livanis, M., M.J. Coyne, and L.E. Comstock, *An antimicrobial protein of the gut symbiont Bacteroides fragilis with a MACPF domain of host immune proteins*. Mol Microbiol, 2014. **94**(6): p. 1361-74.
112. Fraser, S.A., et al., *Perforin lytic activity is controlled by calreticulin*. J Immunol, 2000. **164**(8): p. 4150-5.
113. Traore, D.A., et al., *Defining the interaction of perforin with calcium and the phospholipid membrane*. Biochem J, 2013. **456**(3): p. 323-35.
114. Podack, E.R. and P.J. Konigsberg, *Cytolytic T cell granules. Isolation, structural, biochemical, and functional characterization*. J Exp Med, 1984. **160**(3): p. 695-710.
115. Baran, K., et al., *The molecular basis for perforin oligomerization and transmembrane pore assembly*. Immunity, 2009. **30**(5): p. 684-95.
116. Praper, T., et al., *Human perforin employs different avenues to damage membranes*. J Biol Chem, 2011. **286**(4): p. 2946-55.
117. Shi, L., et al., *Granzyme B binds to target cells mostly by charge and must be added at the same time as perforin to trigger apoptosis*. J Immunol, 2005. **174**(9): p. 5456-61.
118. Bird, C.H., et al., *Cationic sites on granzyme B contribute to cytotoxicity by promoting its uptake into target cells*. Mol Cell Biol, 2005. **25**(17): p. 7854-67.
119. Metkar, S.S., et al., *Perforin oligomers form arcs in cellular membranes: a locus for intracellular delivery of granzymes*. Cell Death Differ, 2015. **22**(1): p. 74-85.
120. Tomita, T., et al., *Pleurotolysin, a novel sphingomyelin-specific two-component cytolysin from the edible mushroom Pleurotus ostreatus*,

- assembles into a transmembrane pore complex*. J Biol Chem, 2004. **279**(26): p. 26975-82.
121. Berne, S., L. Lah, and K. Sepcic, *Aegerolysins: structure, function, and putative biological role*. Protein Sci, 2009. **18**(4): p. 694-706.
  122. Ota, K., et al., *Membrane cholesterol and sphingomyelin, and ostreolysin A are obligatory for pore-formation by a MACPF/CDC-like pore-forming protein, pleurotolysin B*. Biochimie, 2013. **95**(10): p. 1855-64.
  123. Bhat, H.B., et al., *Binding of a pleurotolysin ortholog from Pleurotus eryngii to sphingomyelin and cholesterol-rich membrane domains*. J Lipid Res, 2013. **54**(10): p. 2933-43.
  124. Bakrac, B. and G. Anderluh, *Molecular mechanism of sphingomyelin-specific membrane binding and pore formation by actinoporins*. Adv Exp Med Biol, 2010. **677**: p. 106-15.
  125. Ota, K., et al., *Fungal MACPF-like proteins and aegerolysins: bi-component pore-forming proteins?* Subcell Biochem, 2014. **80**: p. 271-91.
  126. Xu, Q., et al., *Structure of a membrane-attack complex/perforin (MACPF) family protein from the human gut symbiont Bacteroides thetaiotaomicron*. Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 10): p. 1297-305.



## **Chapter 2**

### **Predicting Giant Transmembrane**

### **$\beta$ -Barrel Architecture**

Monash University

## Declaration for Thesis Chapter 2

### Declaration by candidate

In the case of Chapter 2, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, generated data and structural models, analysed data and wrote manuscript	60

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Michelle A Dunstone	Designed experiments, analysed data, wrote manuscript and led the research	
James C Whisstock	Analysed data, provided intellectual input and reviewed manuscript	
Khalid Mahmood	Analysed data	5

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

Candidate's Signature		Date	13/11/15
Main Supervisor's Signature		Date	16/11/15

## Preface to Chapter 2

This chapter describes our work predicting how CDCs span the membrane via a novel  $\beta$ -barrel architecture. The  $\beta$ -barrel architecture is unusual and we suggest it is unique to giant exemplars of such structures. The chapter is presented in the journal article format and was published in *Bioinformatics* [1]. The supplementary material is reproduced at the end of the chapter.

In order to investigate the architecture of the CDC transmembrane channel, we first investigated the structural features of small homo-oligomeric  $\beta$ -barrels (<25  $\beta$ -strands) using available X-ray crystal structures. These data were then applied and extrapolated to the CDC  $\beta$ -barrel. The latter structures are an order of magnitude larger (120-200  $\beta$ -strands), than known, experimentally described  $\beta$ -barrels [2].

We found that the previous model put forward by Tilley and colleagues [3] where the  $\beta$ -strands are perpendicular to the membrane, is incompatible with the dimensions of the CDC pore. In addition, we found that the range of architectures observed in small  $\beta$ -barrels is also incompatible with the dimensions of the CDC pore.

Here, a single alternative and novel  $\beta$ -barrel architecture is described that satisfies both the dimensions of the CDC pore and general principles of  $\beta$ -barrel atomic structure. Moreover, we discuss why this previously uncharacterized structure is not found in smaller proteins and is specific to giant  $\beta$ -barrels.

Following publication of these data, another group experimentally validated our work through systematic disulphide bond scanning and spectroscopic quenching

[4]. Accordingly, this finding raises the question of whether or not the same  $\beta$ -barrel architecture is deployed in the MACPF family. This question is addressed in Chapter 4.

Finally, it is notable that the unique giant  $\beta$ -barrel architecture we have identified has been recently observed at near-atomic resolution in the transmembrane segment of a mitochondrial bacterial secretin (GsgG, a molecule that contains 36  $\beta$ -strands) [5, 6]. Consequently, this demonstrates that the  $\beta$ -barrel architecture deployed by CDCs is not exclusive to CDC pores and highlights the relevance of this work across protein biology.



## Predicting giant transmembrane $\beta$ -barrel architecture

Cyril F. Reboul<sup>1,2</sup>, Khalid Mahmood<sup>1,2</sup>, James C. Whisstock<sup>1,2</sup> and Michelle A. Dunstone<sup>1,3,\*</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology, <sup>2</sup>ARC Centre of Excellence in Structural and Functional Microbial Genomics and <sup>3</sup>Department of Microbiology, Monash University, Clayton, VIC 3800, Australia

Associate Editor: Anna Tramontano

### ABSTRACT

**Motivation:** The  $\beta$ -barrel is a ubiquitous fold that is deployed to accomplish a wide variety of biological functions including membrane-embedded pores. Key influences of  $\beta$ -barrel lumen diameter include the number of  $\beta$ -strands ( $n$ ) and the degree of shear ( $S$ ), the latter value measuring the extent to which the  $\beta$ -sheet is tilted within the  $\beta$ -barrel. Notably, it has previously been reported that the shear value for small antiparallel  $\beta$ -barrels ( $n \leq 24$ ) typically ranges between  $n$  and  $2n$ . Conversely, it has been suggested that the  $\beta$ -strands in giant antiparallel  $\beta$ -barrels, such as those formed by pore forming cholesterol-dependent cytolysins (CDC), are parallel relative to the axis of the  $\beta$ -barrel, i.e.  $S=0$ . The  $S=0$  arrangement, however, has never been observed in crystal structures of small  $\beta$ -barrels. Therefore, the structural basis for how CDCs form a  $\beta$ -barrel and span a membrane remains to be understood.

**Results:** Through comparison of molecular models with experimental data, we are able to identify how giant CDC  $\beta$ -barrels utilize a 'near parallel' arrangement of  $\beta$ -strands where  $S=n/2$ . Furthermore, we show how side-chain packing within the  $\beta$ -barrel lumen is an important limiting factor with respect to the possible shear values for small  $\beta$ -barrels ( $n \leq 24$   $\beta$ -strands). In contrast, our models reveal no such limitation restricts the shear value of giant  $\beta$ -barrels ( $n > 24$   $\beta$ -strands). Giant  $\beta$ -barrels can thus access a different architecture compared with smaller  $\beta$ -barrels.

**Contact:** michelle.dunstone@monash.edu

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

Received on January 26, 2012; revised on March 14, 2012; accepted on March 25, 2012

### 1 INTRODUCTION

The  $\beta$ -barrel structure is central to a large number of functionally diverse proteins.  $\beta$ -Barrels are deployed as enzymes [e.g. the outer membrane protease, OmpT (Vandeputte-Rutten *et al.*, 2001)], as water-soluble transport proteins [e.g. lipocalins (Breustedt *et al.*, 2005)], as fluorescent proteins [e.g. green fluorescent protein (Yang *et al.*, 1996)] and as membrane-embedded channels [e.g. porins (Weiss and Schulz, 1992)].

The properties of small  $\beta$ -barrels have been well-characterized. In particular, the shear value ( $S$ ) of the  $\beta$ -barrel describes the stagger of individual  $\beta$ -strands. For a given number of  $\beta$ -strands ( $n$ ), the

greater the shear value, the greater the slope each  $\beta$ -strand is relative to the axis of the  $\beta$ -barrel lumen. Conversely, the lower the  $S$ -value, the closer to parallel the  $\beta$ -strand is relative to the axis of the  $\beta$ -barrel lumen (Supplementary Fig. S1). (A shear value of 0 results in  $\beta$ -strands that lie parallel to the barrel lumen.) Most importantly, when the number of  $\beta$ -strands and the shear value is known for a  $\beta$ -barrel, the diameter can be calculated for cylindrical  $\beta$ -barrels with a high degree of accuracy (McLachlan, 1979; Murzin *et al.*, 1994a).

For small  $\beta$ -barrels ( $n \leq 24$ ) that are structurally characterized, the  $S$ -value of these structures lies between  $n$  and  $2n$  (Murzin *et al.*, 1994b; Nagano *et al.*, 1999; Schulz, 2002). To date, the structures of over 190 transmembrane  $\beta$ -barrels have been determined (Fairman *et al.*, 2011); analysis of these data reveal that membrane  $\beta$ -barrels obey the same structural rules observed for their soluble counterparts. A subset of transmembrane  $\beta$ -barrels includes structures that are formed by homo-oligomerization. Examples of such  $\beta$ -barrels include TolC (Koronakis *et al.*, 2000) and  $\beta$ -pore forming toxins such as  $\alpha$ -haemolysin (Song *et al.*, 1996).

Although membrane  $\beta$ -barrels have contributed to our understanding of small  $\beta$ -pore forming toxins, less is known about the giant homo-oligomeric  $\beta$ -barrel pores formed by members of the cholesterol-dependent cytolysins (CDC)/Membrane Attack Complex/Perforin (MACPF) superfamily (Hotze and Tweten, 2011; Rosado *et al.*, 2008). This superfamily includes key bacterial virulence factors produced by Gram positive bacteria as well as the major immune effectors of mammals (Kondos *et al.*, 2010). No crystal structure of a MACPF or CDC pore has been determined to date, however, the 28 Å resolution cryo-electron microscopy (EM) structure of pneumolysin reveals that this CDC forms a giant 240 Å diameter  $\beta$ -barrel that contains 38 monomers (Tilley *et al.*, 2005). To make this  $\beta$ -barrel, each pneumolysin monomer contributes two amphipathic  $\beta$ -hairpins, i.e. four antiparallel  $\beta$ -strands (Shatursky *et al.*, 1999; Shepard *et al.*, 1998).

Initial models presented through EM studies of the CDC, pneumolysin, suggested that the  $\beta$ -strands are parallel to the lumen in order to reach and fully span the lipid bilayer as a  $\beta$ -barrel (Tilley *et al.*, 2005). However, such an arrangement would not be consistent with a shear value between  $n$  and  $2n$  as observed in the known crystal structures of transmembrane  $\beta$ -barrels (Pali and Marsh, 2001; Schulz, 2002). Here, using molecular models, we show that the optimal shear value of these giant  $\beta$ -barrel structures is  $S=n/2$ , i.e. arranged in a near parallel fashion with respect to the  $\beta$ -barrel lumen. We further demonstrate that  $\beta$ -barrels with  $<24$   $\beta$ -strands are unable to readily access such architecture as a consequence of luminal side-chain packing constraints.

\*To whom correspondence should be addressed.

## 2 METHODS

### 2.1 $\beta$ -Barrel parameter refinement

To build atomic models of  $\beta$ -barrels, an initial  $C\alpha$  template geometry was generated based on known  $\beta$ -barrel structures (McLachlan, 1979; Murzin et al., 1994a). For a given pair of  $n$  and  $S$ , idealized cylindrical  $\beta$ -barrels were built with a radius  $R$  and a  $\beta$ -strand slope of  $\alpha$  with respect to the cylinder axis, following McLachlan (1979) and in a fashion reminiscent of research by Sansom and Kerr (1995). Each residue is represented by its  $C\alpha$  and placed on the surface of the cylinder. Positions of the  $C\alpha$  are fully determined by the values of  $n$  and  $S$  and by the choice of the grid parameters  $a$  and  $b$  (Murzin et al., 1994a).  $a$  is the  $C\alpha$ – $C\alpha$  distance between adjacent residues on the same  $\beta$ -strand (intra-strand), while  $b$  is the  $C\alpha$ – $C\alpha$  distance across hydrogen-bonded  $\beta$ -strands (inter-strand).

When building such  $\beta$ -barrels, the values of the parameters  $a$  and  $b$  influence the local (e.g. intra- and inter-strand distances) as well as global (e.g. radius) structural features of the resulting models. In this study, we optimized the  $a$  and  $b$  parameters using a training set of antiparallel membrane-embedded  $\beta$ -barrels that are formed by homo-oligomers. We searched protein structure databases using the MPTopo and PDBtools (Jayasinghe et al., 2001; Velankar et al., 2011) and found eight suitable structures following these strict criteria. Of these eight chosen structures (Supplementary Table S1), the transmembrane  $\beta$ -barrels were isolated and their  $C\alpha$  coordinates projected on the surface of a fitted cylinder (based on least-square fitting) and the grid parameters were derived. All structures showed a clear cylindrical and symmetric shape and our analysis found that  $a=3.48$  Å and  $b=4.83$  Å. The latter values are found to be in very good agreement with previously reported ones for antiparallel water-soluble  $\beta$ -barrels ( $a=3.53$  Å;  $b=4.87$  Å) (Chou and Scheraga, 1982). The ratio  $a/b=0.721$  is also in excellent agreement with the 0.719 reported for a non-overlapping set of transmembrane  $\beta$ -barrels (Pali and Marsh, 2001).

Using these refined parameters, the  $\beta$ -barrel  $C\alpha$  scaffold was built: first the  $C\alpha$  ( $\beta$ -strand 1) is placed at a distance  $R$  to the arbitrary cylinder axis. The  $C\alpha$  of the next residue is then positioned at a distance  $a$  on the cylinder surface such that the angle between the  $C\alpha$ – $C\alpha$  vector and the  $\beta$ -barrel axis is  $\alpha$ . The remaining residues of the  $\beta$ -strand are iteratively positioned the same way. The first  $C\alpha$  of the adjacent  $\beta$ -strand (strand 2) is positioned on the cylinder surface, at a distance  $b$  from an adjacent  $C\alpha$  belonging to  $\beta$ -strand 1 in such a way that the  $C\alpha$ – $C\alpha$  vector is perpendicular to  $\beta$ -strand 1. The other residues of  $\beta$ -strand 2 are placed as was done for  $\beta$ -strand 1. The remaining  $\beta$ -strands are built iteratively using the same method.

### 2.2 Atomistic $\beta$ -barrel models

Based on the  $C\alpha$  scaffold, full atomistic models were first built as polyaniline  $\beta$ -strands. The  $C\beta$  atoms were positioned 1.5 Å apart from the  $C\alpha$  alternately facing towards the pore lumen and towards the membrane, thus ensuring proper orientation of the side chains and the antiparallel nature of the  $\beta$ -barrels. Amino acid sequences of the  $\beta$ -hairpins and  $\beta$ -turns were subsequently mapped onto this framework using MODELLER (Eswar et al., 2006). The resulting structures were subjected to three cycles of energy minimization employing NAMD 2.8 (Phillips et al., 2005), the CHARMM22 forcefield with CMAP correction (MacKerell et al., 1998; Mackerell et al., 2004) and the GBIS solvent model (Tanner et al., 2011). In the first step, harmonic restraints on heavy atoms are smoothly relaxed from 200 to 5 kJ (mol Å)<sup>−1</sup> to resolve potential steric clashes. In the second step, inter-strand backbone hydrogen bonds constraints are employed (N–H...O=C) with a reference value of 1.9 Å (McLachlan, 1979) and relaxed from 5 to 1 kJ (mol Å)<sup>−1</sup>. The hydrogen bonds correspond to the canonical antiparallel  $\beta$ -strand hydrogen bonding pattern. In the final step, all constraints are relaxed. Energy minimizations involved 15 000 steps in total.

Based on the values of  $n$  and  $S$  in our dataset, the corresponding  $\beta$ -barrels atomistic models were built and found to reproduce closely the crystallographic structure with an average  $C\alpha$  RMSD  $\sim 1.1$  Å (Supplementary Table S1), thus validating our approach and methodology. Model pore radii were

found to be in excellent agreement with radii calculated from the X-ray structures.

Transmembrane  $\beta$ -barrel models were subsequently built for pneumolysin (38 subunits;  $n=152$ ) (Tilley et al., 2005). Five models were built with shear values  $S=0$ ,  $n/2$ ,  $n$ ,  $3n/2$  and  $2n$ . In addition, the estimate of error of  $\beta$ -barrel diameter estimation is calculated (see Supplementary Material).

### 2.3 Side-chain packing in the $\beta$ -barrel interior

To probe the influence of side-chain packing on  $\beta$ -barrel architecture (e.g. the choice of  $S$  given  $n$ ), we positioned pseudo-side chains in the  $\beta$ -barrel interior of our  $C\alpha$  scaffold and calculated the van der Waals (vdW) interaction energy ( $E_{vdW}$ ) of one pseudo-side chain with its closest eight neighbours (Supplementary Fig. S2). To achieve this, a pseudo-methyl group per residue was positioned in the  $\beta$ -barrel interior 3.0 Å away from the  $C\alpha$ , thus mimicking an amino acid side chain with an aliphatic group at position  $C\gamma$ . As a result, a methyl group is placed in the interior every second residue (i.e. following the alternate orientation of side chains in  $\beta$ -strands). The resulting semi-quantitative  $E_{vdW}$  calculated for each combination of ( $n$ ,  $S$ ) are presented in Figure 2 and compared with known structures (Supplementary Table S2). Methyl group vdW parameters were taken from the GROMOS forcefield (Oostenbrink et al., 2004).

## 3 RESULTS

### 3.1 Predicting shear values of CDC $\beta$ -barrels

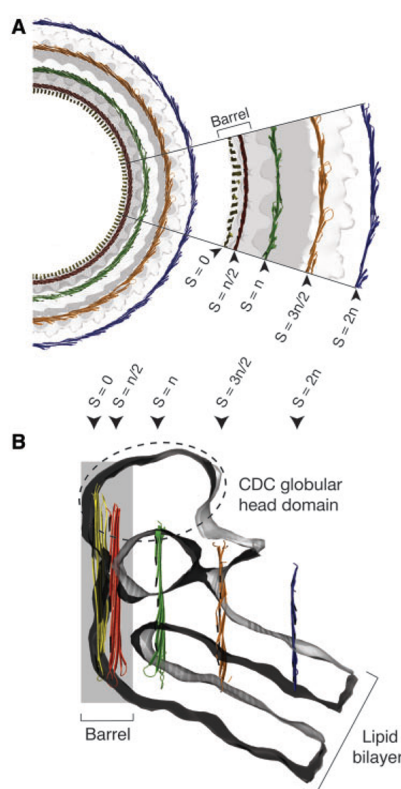
The 28 Å resolution cryo-EM structure of pneumolysin contains 152  $\beta$ -strands and forms a pore 240 Å in diameter [PLY; EMD ID: EMD-1107 (Tilley et al., 2005)]. We compared a range of models and identified the  $\beta$ -barrel with a shear value that most closely agreed with these experimental data. Homo-oligomeric antiparallel  $\beta$ -barrels typically have an  $S$ -value that is an integer multiple of the number of  $\beta$ -strands, therefore,  $S=n$  and  $S=2n$  models were built together with the  $S=0$  model. Moreover, since two hairpins (i.e. four  $\beta$ -strands) are contributed per monomer, it is possible for CDC  $\beta$ -barrels to have shear values of  $S=n/2$  or  $S=3n/2$  (Supplementary Fig. S1).

The model that fitted best within the experimental EM data, based upon predicted radius, possessed a shear value of  $S=n/2$  (Fig. 1, Supplementary Table S3). Furthermore, our analysis suggests that a  $\beta$ -barrel with  $S=n/2$  would also permit a right-handed twist of the  $\beta$ -strands that is characteristic of  $\beta$ -barrels formed by L-amino acids.

Given the 28 Å resolution of the EM data, however, the model with  $S=0$  may be possible. Neither the  $S=0$  nor the  $S=n/2$  model can be favoured in terms of the overall height of the  $\beta$ -barrels with respect to the position of the CDC globular head domain since there is only a 7 Å reduction in height between the  $S=0$  and the  $S=n/2$  models (over a total distance of 88 Å). In both models, the globular head domains can fit well within the density (Fig. 1B).

In contrast,  $S$ -values of  $n$ ,  $3n/2$  or  $2n$  resulted in pore sizes that are far larger than the observed dimensions of the pore (Fig. 1). Importantly, the calculated minimum diameter for the  $S=n$  model and the calculated maximum diameter for the  $S=n/2$  model do not overlap (Supplementary Table S3). The minimum diameter for the  $S=n$  model does not overlap with the EM density of the  $\beta$ -barrel. We were, therefore, confident in ruling out the  $S=n$  model for this CDC.

Perfringolysin O (PFO), while it has no 3D EM data of the pore, has contributed to our understanding of the mechanism of CDC pore formation (Rossjohn et al., 1997; Shatursky et al., 1999; Shepard et al., 1998). Importantly, pyrene labelling studies of adjacent side



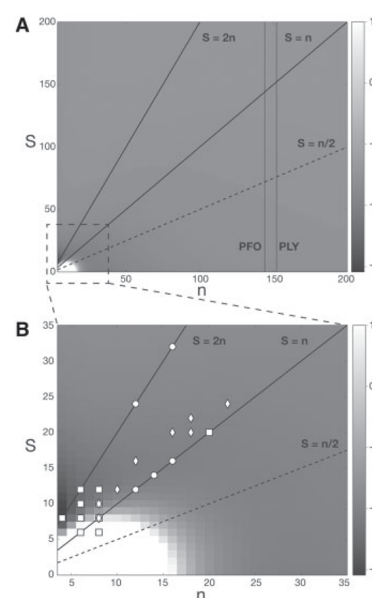
**Fig. 1.** Superposition of the pneumolysin 38-mer pore with atomic pore models of various shear ( $S$ ) values. **(A)** Top view. Left: models and 3D-EM map; right: models with shear values indicated by arrows. **(B)** Blow-out view. The  $S$ -values presented are: 0 (yellow),  $n/2$  (76; red),  $n$  (152; green),  $n/2$  (228; orange) and  $2n$  (304; blue).

chains in PFO suggest that the  $\beta$ -strands are slightly staggered and, therefore, favour the slightly tilted  $S=n/2$  model over the parallel  $S=0$  model (Ramachandran *et al.*, 2004; Supplementary Fig. S3).

### 3.2 Constraints governing $\beta$ -barrel size and shape

For small  $\beta$ -barrels ( $n < 24$ ) the shear value lies within the  $n \leq S \leq 2n$  range (Murzin *et al.*, 1994b; Nagano *et al.*, 1999; Schulz, 2002). We, therefore, investigated whether there were any constraints that limited the  $S$ -value for a given number of  $\beta$ -strands.

A major constraint for small  $\beta$ -barrel architecture is the packing of the side chains in the interior of the  $\beta$ -barrel (Lesk *et al.*, 1989; Murzin *et al.*, 1994a). Typically,  $\beta$ -sheets in water-soluble  $\beta$ -barrels bury hydrophobic residues in the interior. In contrast, in membrane-embedded pores, the interior channel is lined with hydrophilic residues, whereas the exterior surface of the  $\beta$ -barrel is hydrophobic. The packing of the side chains in the lumen of the  $\beta$ -barrel plays an important role in the overall architecture of the  $\beta$ -barrel.



**Fig. 2.** Interior side-chain packing as a function of  $\beta$ -barrel architecture. In **(A)** and **(B)**, the interior side-chain packing is semi-quantitatively estimated by side-chain van der Waals energies (in  $\text{kJ mol}^{-1}$ ). White regions ( $E_{\text{vdW}} > 1.0 \text{ kJ mol}^{-1}$ ) are considered unfavourable, grey regions ( $E_{\text{vdW}} < 1.0 \text{ kJ mol}^{-1}$ ) are considered favourable. Squares represent water-soluble  $\beta$ -barrels, diamonds represent membrane  $\beta$ -barrels and circles represent oligomeric antiparallel membrane  $\beta$ -barrels.

We calculated the theoretical side-chain van der Waals energies for all theoretical combinations of  $S$  and  $n$  in antiparallel  $\beta$ -barrels (Fig. 2). It can be predicted from the plot that it is not energetically favourable for small antiparallel  $\beta$ -barrels to achieve  $0 \leq S \leq n$  structural arrangement due to steric clashes made by the lumen-orientated side chains. However, side-chain packing no longer becomes a constraint for  $\beta$ -barrel architecture in  $\beta$ -barrels  $> 20$   $\beta$ -strands. Therefore, giant  $\beta$ -barrels formed by CDCs are capable of forming ‘near parallel’ barrels, whereas small  $\beta$ -barrels are not.

### 3.3 Validation of the side-chain constraints for small soluble $\beta$ -barrels

To validate the side-chain packing plot (Fig. 2), we analysed the range in known structures of both soluble  $\beta$ -barrels and membrane-embedded  $\beta$ -barrel solved to date. Most structures lie within the energetically favourable region of the plot (Fig. 2B, Supplementary Table S2). In addition, the structures in general lie within the  $n \leq S \leq 2n$  (Murzin *et al.*, 1994b; Nagano *et al.*, 1999; Schulz, 2002) region.

There are two known exceptions to this rule: the first is oxidoreductase (PDB ID 3sod), which has eight  $\beta$ -strands and a shear value of 6 ( $n=8, S=6$ ) and thus lying within the energetically unfavourable region. However, oxidoreductase has only a single hydrogen bond between two of the  $\beta$ -strands due to the constraints

of the  $\beta$ -sheet and the stacking of side chains within the  $\beta$ -barrel and thus resembles a 'burst' or overstuffed  $\beta$ -barrel (Supplementary Fig. S4). The second  $\beta$ -barrel type lies in the unfavourable region is tetrahydridipicolinate N-succinyltransferase ( $n=6$ ,  $S=6$ ) (PDB ID 1tdt). This is a very short  $\beta$ -barrel that contains only a single side chain per  $\beta$ -strand stacked inside the  $\beta$ -barrel lumen (Supplementary Fig. S5). Lumen-directed side chains in this structure can, therefore, readily access space at the top and the bottom of the  $\beta$ -barrel.

Together, these exceptions support the idea that steric clashes in the interior of  $\beta$ -barrels affect the possible combinations of shear ( $S$ ) and  $\beta$ -strand number ( $n$ ) in small antiparallel 'near-ideal'  $\beta$ -barrels that contain extensive  $\beta$ -sheet hydrogen bonds between all  $\beta$ -strands.

It is also important to note that our calculations suggest that the constraints imposed by side chains on the  $S$  to  $n$  ratio disappear as  $\beta$ -barrel sizes grow over 20  $\beta$ -strands. For giant  $\beta$ -barrels such as CDC pore forming proteins, it is predicted that there are no limitations to the  $S/n$  ratio imposed by luminal side-chain packing. In summary, our data show that, as  $\beta$ -barrels grow larger, the interior lumen becomes less curved and more amenable to a smaller shear to number of  $\beta$ -strands ratio.

#### 4 DISCUSSION

Comparison of  $\beta$ -barrel models with published 3D-EM structures (Tilley et al., 2005), reveals that for CDCs, a  $\beta$ -barrel formed with a shear value of  $S=n/2$  provides the best fit. This is supported by the biophysical research performed upon PFO (Ramachandran et al., 2004). We further found that giant  $\beta$ -barrels have no apparent limitation with respect to shear values. In contrast, the architecture of smaller  $\beta$ -barrels is severely constrained by lumen side-chain packing. To the best of our knowledge,  $\beta$ -barrels that are long enough to span a membrane and possess a shear value of  $S=n/2$  or 0 have not been observed at atomic resolution. However, our data lend support to the idea that CDCs can indeed form such structures. These ideas may be of further utility in modelling large  $\beta$ -barrels formed by other supermolecular complexes.

#### ACKNOWLEDGEMENTS

The authors thank Helen Saibil, Rodney Tweten and Bosco Ho for helpful discussions.

**Funding:** Australian Research Council [DP0986811, FF0883418 to J.C.W.]; National Health and Medical Research Council [606471 to M.A.D.]; and C.F.R. is supported by the Australian Postgraduate Award.

**Conflict of Interest:** none declared.

#### REFERENCES

- Breustedt, D.A. et al. (2005) The 1.8-Å crystal structure of human tear lipocalin reveals an extended branched cavity with capacity for multiple ligands. *J. Biol. Chem.*, **280**, 484–493.
- Chou, K.C. and Scheraga, H.A. (1982) Origin of the right-handed twist of beta-sheets of poly(LVal) chains. *Proc. Natl Acad. Sci. USA*, **79**, 7047–7051.
- Eswar, N. et al. (2006) Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics*, **Chapter 5**, Units 5–6.
- Fairman, J.W. et al. (2011) The structural biology of  $\beta$ -barrel membrane proteins: a summary of recent reports. *Curr. Opin. Struct. Biol.*, **21**, 523–531.
- Hotze, E.M. and Tweten, R.K. (2012) Membrane assembly of the cholesterol-dependent cytolysin pore complex. *Biochim. Biophys. Acta.*, **1818**, 1028–1038.
- Jayasinghe, S. et al. (2001) MPTopo: a database of membrane protein topology. *Protein Sci.*, **10**, 455–458.
- Kondos, S.C. et al. (2010) The structure and function of mammalian membrane-attack complex/perforin-like proteins. *Tissue Antigens*, **76**, 341–351.
- Koronakis, V. et al. (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*, **405**, 914–919.
- Lesk, A.M. et al. (1989) Structural principles of alpha/beta barrel proteins: the packing of the interior of the sheet. *Proteins*, **5**, 139–148.
- MacKerell, A.D. et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
- MacKerell, A.D. Jr. et al. (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.*, **25**, 1400–1415.
- McLachlan, A.D. (1979) Gene duplication in the evolution of the yeast hexokinase active site. *Eur. J. Biochem.*, **100**, 181–187.
- Murzin, A.G. et al. (1994a) Principles determining the structure of beta-sheet barrels in proteins. I. A theoretical analysis. *J. Mol. Biol.*, **236**, 1369–1381.
- Murzin, A.G. et al. (1994b) Principles determining the structure of beta-sheet barrels in proteins. II. The observed structures. *J. Mol. Biol.*, **236**, 1382–1400.
- Nagano, N. et al. (1999) Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels. *Protein Sci.*, **8**, 2072–2084.
- Oostenbrink, C. et al. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.*, **25**, 1656–1676.
- Pali, T. and Marsh, D. (2001) Tilt, twist and coiling in beta-barrel membrane proteins: relation to infrared dichroism. *Biophys. J.*, **80**, 2789–2797.
- Phillips, J.C. et al. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
- Ramachandran, R. et al. (2004) Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment. *Nat. Struct. Mol. Biol.*, **11**, 697–705.
- Rosado, C.J. et al. (2008) The MACPF/CDC family of pore-forming toxins. *Cell Microbiol.*, **10**, 1765–1774.
- Rossjohn, J. et al. (1997) Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form. *Cell*, **89**, 685–692.
- Sansom, M.S. and Kerr, I.D. (1995) Transbilayer pores formed by beta-barrels: molecular modeling of pore structures and properties. *Biophys. J.*, **69**, 1334–1343.
- Schulz, G.E. (2002) The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta.*, **1565**, 308–317.
- Shatursky, O. et al. (1999) The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell*, **99**, 293–299.
- Shepard, L.A. et al. (1998) Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy. *Biochemistry*, **37**, 14563–14574.
- Song, L. et al. (1996) Structure of staphylococcal  $\alpha$ -hemolysin, a heptameric transmembrane pore. *Science*, **274**, 1859–1865.
- Tanner, D.E. et al. (2011) Parallel generalized born implicit solvent calculations with NAMD. *J. Chem. Theory Comput.*, **7**, 3635–3642.
- Tilley, S.J. et al. (2005) Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell*, **121**, 247–256.
- Vandeputte-Rutten, L. et al. (2001) Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. *EMBO J.*, **20**, 5033–5039.
- Velankar, S. et al. (2012) PDB: protein data bank in Europe. *Nucleic Acids Res.*, **40**, D445–452.
- Weiss, M.S. and Schulz, G.E. (1992) Structure of porin refined at 1.8 Å resolution. *J. Mol. Biol.*, **227**, 493–509.
- Yang, F. et al. (1996) The molecular structure of green fluorescent protein. *Nat. Biotechnol.*, **14**, 1246–1251.

## Supplementary Material

### Contents:

**Table S1.** Structural aspects of transmembrane, circular, oligomeric  $\beta$ -barrels used for analysis and modelling in this study

**Table S2.**  $\beta$ -barrel topologies for a subset of proteins

**Table S3.** Error estimates of the pneumolysin pore radius

**Figure S1.** Schematic topologies for oligomeric  $\beta$ -pores

**Figure S2.** Schematic representation of the neighbouring sidechains taken into account in  $E_{vdw}$ .

**Figure S3.** Inter-strand arrangement of sidechains for the 5 different PFO models.

**Figure S4.** Cartoon representation of the oxidoreductase Cu, Zn superoxide dismutase with PDB ID 3sod

**Figure S5.** Cartoon representations of the  $\beta$ -barrel of tetrahydrodipicolinate N-succinyltransferase

**Supplementary Methods:** Pores radii errors estimation

**Supplementary references**

**Table S1**

**Structural aspects of transmembrane, circular, oligomeric  $\beta$ -barrels used for analysis and modelling in this study**

	<b>n</b>	<b>S</b>	<b>Oligomers/pore</b>	<b>PDB ID /EMD ID</b>	<b>RMSD</b>
<b>Pneumolysin</b>	140-160	-	35-40	EMD-1107	-
<b>Perfringolysin O</b>	140-160	-	35-40	-	-
<b>HiA</b>	12	12	3	2gr8	1.1
<b>Hyl-A</b>	14	14	7	3o44	1.2
<b><math>\alpha</math>-Hemolysin</b>	14	14	7	7ahl	0.9
<b><math>\gamma</math>-Hemolysin</b>	16	16	4 dimers	3b07	0.9
<b>CusC</b>	12	24	3	3pik	1.2
<b>ToIC</b>	12	24 *	3	1ek9	0.9
<b>MspA</b>	16	32 **	8	1uun	1.3
<b>VCEC</b>	12	24	3	1yc9	1.3

\* After careful examination we find S = 24 as opposed to the previously reported S = 20 (Schulz, 2002).

\*\* S = 32 in the main transmembrane region. A second, shorter transmembrane section is S = 16 thereby creating a goblet shape (Faller, et al., 2004).

Note, a ninth oligomeric transmembrane barrel, OprM (Akama, et al., 2004), was discarded as its  $\beta$ -barrel exhibits a near triangular shape.



**Table S2**

**$\beta$ -barrel topologies for a subset of proteins (Murzin, et al., 1994; Nagano, et al., 1999; Pali and Marsh, 2001; Schulz, 2002)**

S values were manually validated.

<i>Water soluble proteins</i>		
<b>n</b>	<b>S</b>	<b>Example PDB ID</b>
4	8	1bnd
6	6	1tdt
6	8	1dsu
6	10	1fmb
6	12	1tie
8	6	3sod
8	8	1zdy
8	10	8cat
8	12	1brp
20	20	1gtp
<i>Transmembrane proteins</i>		
8	8	1qj8
8	10	1qjp
10	12	1i78
12	16	1qd5
16	20	2por
18	20	1mpm
18	22	1mal
22	24	1by3

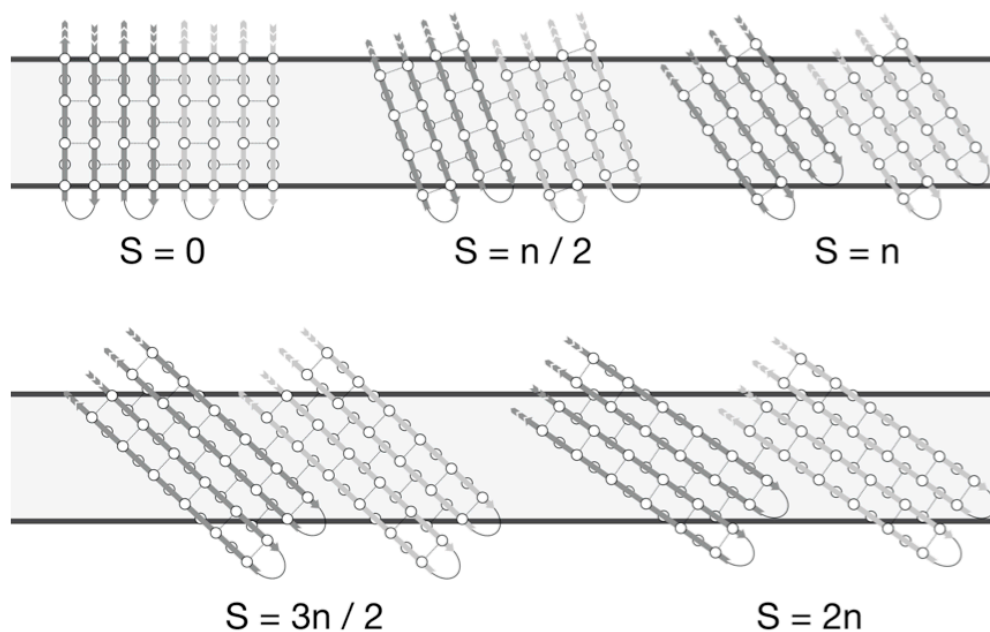
**Table S3****Error estimates of the pneumolysin pore radius (in Å)**

	Models				
S =	0	n/2 = 76	n = 152	3n/2 = 226	2n = 304
Minimum radius limit	107.8	115.5	135.9	164.6	198.8
Optimal radius	117.3	124.1	144.3	172.3	205.2
Maximum radius limit	126.1	133.1	151.7	179.4	211.2



**Figure S1**

**Schematic topologies for oligomeric  $\beta$ -pores**



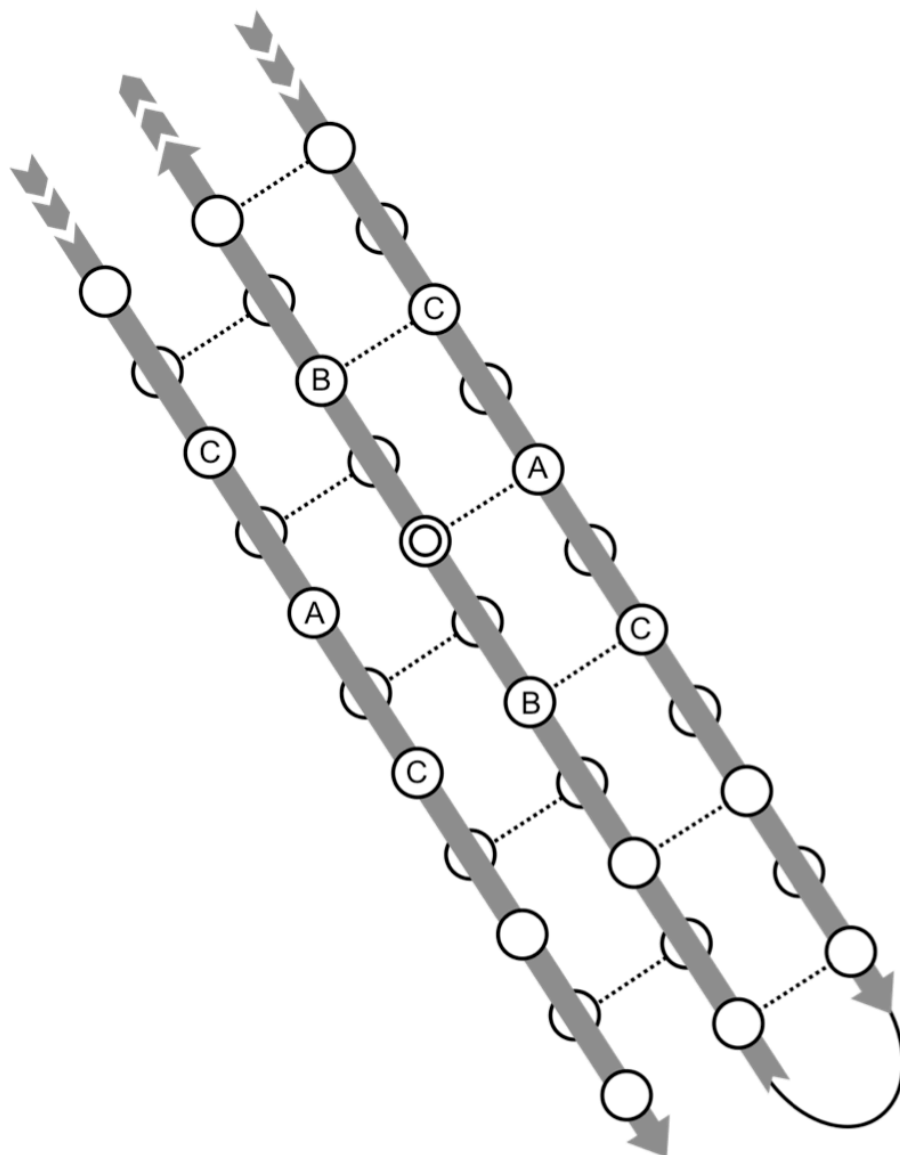
For each possible combination of  $(n, S)$  the regular stagger of the individual subunits and their membrane lining is represented. Two monomers are displayed for each topology in light and dark grey. White circles represent sidechains alternately pointed towards the pore lumen or the membrane. Dashed lined represent the alternate hydrogen bonding pattern. Thick black lines schematically represent the upper and lower borders of the membrane bilayer (in light grey).

**Figure S2**

**Schematic representation of the neighbouring sidechains taken into account in  $E_{\text{vdw}}$ .**

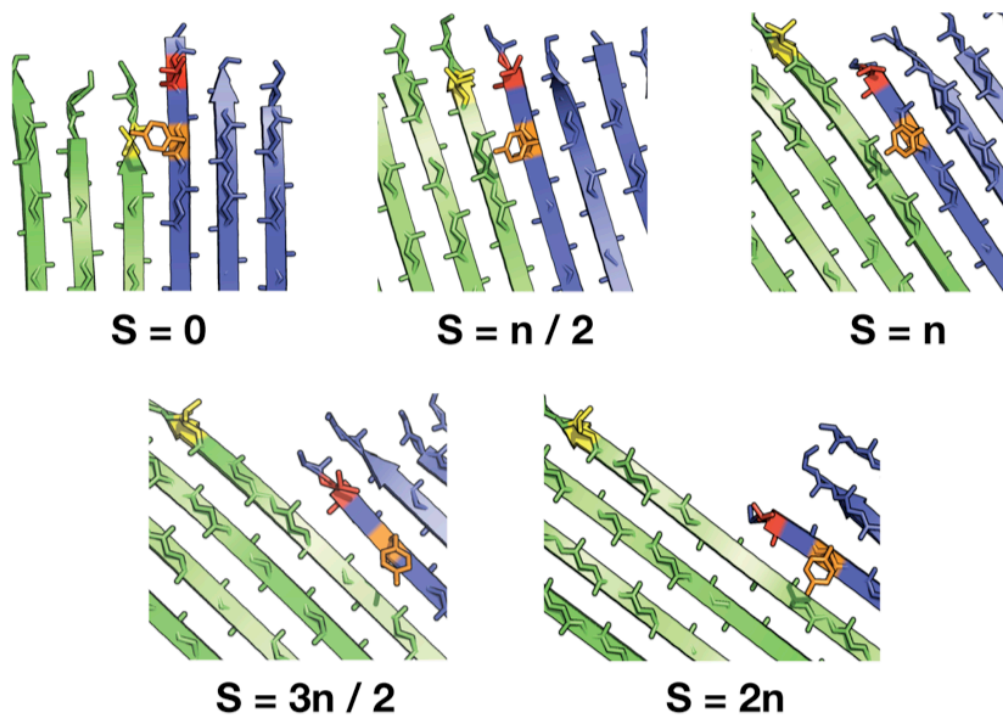
Sheets are represented in grey. Side chain are symbolized by white circles alternatively pointing towards the barrel interior and exterior.

The central central residue (O) is in the vicinity of 8 residues at 3 different distances, represented as A, B and C.



**Figure S3**

**Inter-strand arrangement of sidechains for the 5 different PFO models.**



5 atomic models were created for PFO in a fashion similar to PLY (see Methods) with a ring of 36 monomers (Dang, et al., 2005) and  $S = 0, 72, 144, 216$  and  $288$ . For each shear value  $S$ , two subunits from our models are represented as cartoons in green and blue. The positions the pyrene-labelled Val322 (yellow), Tyr181 (orange) and Thr179 (red) are highlighted. Interaction between Val322 and Thr179 was observed in a pyrene experiment by Tweten and colleagues (Ramachandran, et al., 2004).

**Figure S4**

**Cartoon representation of the oxidoreductase Cu, Zn superoxide  
dismutase with PDB ID 3sod**

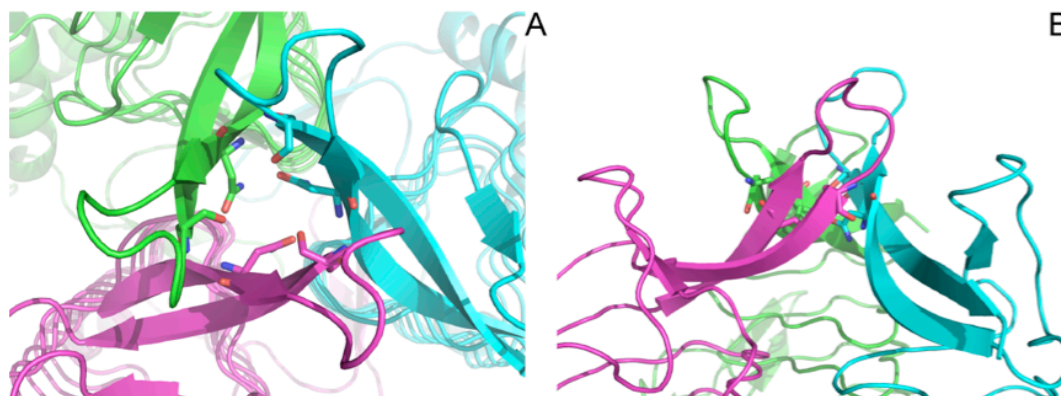
The portion of the mainchain discussed in the manuscript is represented as sticks. The orange dashed line displays the hydrogen bond discussed in the text.



**Figure S5**

**Cartoon representations of the  $\beta$ -barrel of tetrahydrodipicolinate N-succinyltransferase, PDB ID 1tdt**

A. Top view, only one side chain per strand participates to the barrel interior (shown in sticks). B. Side view. The three hairpins forming the  $\beta$ -barrel are largely deformed.



## Supplementary Methods

### Pores radii errors estimation

To calculate the error in estimating giant barrels dimensions we build lower and higher limits models to the pneumolysin pore, which allows to define lower and higher limits to the pore radii (for each value of  $S$  presented).

The parameter  $b$  used (see Methods) is the average between the hydrogen bonded interstrand  $C\alpha$  distance and the non-hydrogen bonded interstrand  $C\alpha$  distance. Here we find the former to be 4.39 Å and the latter 5.26 Å. One can notice that hydrogen bonded residue pairs are indeed found closer than the non-hydrogen bonded  $C\alpha$  pairs.

We modelled for each value of  $S$  two additional pores providing a lower and higher limit to the radius estimation. The lower limit value corresponds to a barrel where all strands are considered hydrogen bonded. As a consequence all strands are initially placed at an, on average, unreasonable short distance ( $b = 4.39$ ), giving a smaller pore radius than when using the more realistic value  $b = 4.83$ . The higher limit value corresponds to the radius of a barrel where all strands are considered non hydrogen-bonded ( $b = 5.26$ ). Consequently, this gives a radius unrealistically large.

Both small and large models were then subjected to the same minimisation procedure (see Methods section). Resulting radii are presented in table S3. We suggest the reader should regard them as generous limits to the pores radii.

## Supplementary references

- Dang, T.X., *et al.* (2005) Prepore to pore transition of a cholesterol-dependent cytolysin visualized by electron microscopy, *J Struct Biol*, **150**, 100-108.
- Faller, M., Niederweis, M. and Schulz, G.E. (2004) The structure of a mycobacterial outer-membrane channel, *Science*, **303**, 1189-1192.
- Murzin, A.G., Lesk, A.M. and Chothia, C. (1994) Principles determining the structure of beta-sheet barrels in proteins. II. The observed structures, *J Mol Biol*, **236**, 1382-1400.
- Nagano, N., Hutchinson, E.G. and Thornton, J.M. (1999) Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels, *Protein Sci*, **8**, 2072-2084.
- Pali, T. and Marsh, D. (2001) Tilt, twist, and coiling in beta-barrel membrane proteins: relation to infrared dichroism, *Biophys J*, **80**, 2789-2797.
- Ramachandran, R., Tweten, R.K. and Johnson, A.E. (2004) Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment, *Nat Struct Mol Biol*, **11**, 697-705.
- Schulz, G.E. (2002) The structure of bacterial outer membrane proteins, *Biochim Biophys Acta*, **1565**, 308-317.





## **Chapter 3**

### **A New Model for Pore Formation**

#### **by Cholesterol Dependent Cytolysins**

## Declaration for Thesis Chapter 3

### Declaration by candidate

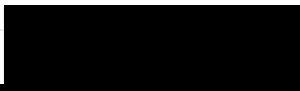
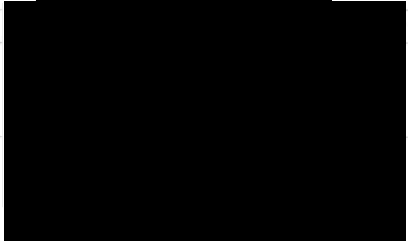
In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, generated data, analysed data and wrote manuscript	60

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Michelle A Dunstone	Designed experiments, analysed data, wrote manuscript and led the research	
James C Whisstock	Analysed data, provided intellectual input and reviewed manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

Candidate's Signature		Date 13/11/15
Main Supervisor's Signature		Date 16/11/15

## Preface to Chapter 3

This chapter describes our analyses of CDC structures and the predicted structural rearrangements that take place in these molecules during transition from the prepore to the pore form. Previous published experiments revealed that the CDC monomers assemble together into a prepore complex following interaction with the lipid membrane. Atomic force microscopy (AFM) [7] experiments suggested that the CDC prepore subsequently undergoes a major conformational change in order to form the final pore. Remarkably, the prepore to pore transition includes a 40 Å vertical collapse of the bulk of the structure towards the membrane surface, later confirmed by Förster resonance energy transfer [8] and single particle cryo-electron microscopy (SP cryo-EM) [3].

Here, the SP cryo-EM 3D reconstructions of the CDC prepore and pore conformations were used to develop a structural model for the vertical transition of the prepore to pore form. We were particularly interested in the movements that take place at the putative crucial “hinge” region (Domain 2) [3]. Previously it was proposed Domain 2 crumples in order to accommodate the conformational change. In this publication we propose an alternative model for this collapse. Our new model is presented in the journal article format and was published in *PLoS Computational Biology* [9]. The supplementary material is reproduced at the end of the chapter.

A systematic structural analysis of the entire cohort of CDC crystallographic structures available at the time (early 2014) was used to investigate the potential trajectory of CDC domains movement. Based on these analyses and using molecular modelling guided by cryo-EM density we suggest that the prepore

collapse involves a simple rotation of Domain 2 [3]. Our new model is consistent with the 29 Å resolution SP cryo-EM reconstruction of the CDC pneumolysin pore. Further, it is suggested that the new model is more structurally plausible in comparison to the previous model [3]. The models have been made publicly available to the community and experiments to validate this new mechanism are discussed.

Since the publication of the present study in August 2014, Leung and colleagues [10] have obtained an improved 15 Å resolution SP cryo-EM reconstruction of the CDC suilysin pore. The higher resolution data obtained directly supports the proposed mechanism of pore formation presented in this chapter.

# A New Model for Pore Formation by Cholesterol-Dependent Cytolysins

Cyril F. Reboul<sup>1,2</sup>, James C. Whisstock<sup>1,2</sup>, Michelle A. Dunstone<sup>1,2,3\*</sup>

**1** Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, Victoria, Australia, **2** ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton Campus, Melbourne, Victoria, Australia, **3** Department of Microbiology, Monash University, Clayton Campus, Melbourne, Victoria, Australia

## Abstract

Cholesterol Dependent Cytolysins (CDCs) are important bacterial virulence factors that form large (200–300 Å) membrane embedded pores in target cells. Currently, insights from X-ray crystallography, biophysical and single particle cryo-Electron Microscopy (cryo-EM) experiments suggest that soluble monomers first interact with the membrane surface *via* a C-terminal Immunoglobulin-like domain (Ig; Domain 4). Membrane bound oligomers then assemble into a prepore oligomeric form, following which the prepore assembly collapses towards the membrane surface, with concomitant release and insertion of the membrane spanning subunits. During this rearrangement it is proposed that Domain 2, a region comprising three  $\beta$ -strands that links the pore forming region (Domains 1 and 3) and the Ig domain, must undergo a significant yet currently undetermined, conformational change. Here we address this problem through a systematic molecular modeling and structural bioinformatics approach. Our work shows that simple rigid body rotations may account for the observed collapse of the prepore towards the membrane surface. Support for this idea comes from analysis of published cryo-EM maps of the pneumolysin pore, available crystal structures and molecular dynamics simulations. The latter data in particular reveal that Domains 1, 2 and 4 are able to undergo significant rotational movements with respect to each other. Together, our data provide new and testable insights into the mechanism of pore formation by CDCs.

**Citation:** Reboul CF, Whisstock JC, Dunstone MA (2014) A New Model for Pore Formation by Cholesterol-Dependent Cytolysins. PLoS Comput Biol 10(8): e1003791. doi:10.1371/journal.pcbi.1003791

**Editor:** Nir Ben-Tal, Tel Aviv University, Israel

**Received:** March 18, 2014; **Accepted:** July 1, 2014; **Published:** August 21, 2014

**Copyright:** © 2014 Reboul et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the Australian Research Council (FF0883418 to JCW, DP0986811 to JCW and MAD, DP120104058 to MAD, CE140100011 to JCW) and the National Health and Medical Research Council (606471 to MAD). CFR is supported by the Australian Postgraduate Award. The authors acknowledge the support of the Victorian Life Sciences Computation Initiative (VLSI, Melbourne, Australia) and the Multi-modal Australian Science Imaging and Visualisation Environment (MASSIVE) (www.massive.org.au). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: Michelle.Dunstone@monash.edu

## Introduction

Cholesterol dependent cytolysins (CDCs) represent a major branch of the CDC/membrane attack complex/perforin-like (MACPF) protein superfamily. Originally identified as virulence factors produced by Gram positive pathogens, CDC toxins have recently been identified in Gram negative bacteria such as *Desulfobulbus propionicus* and *Enterobacter lignolyticus* [1,2]. Well-studied family members include perfringolysin O (PFO), pneumolysin (PLY), listeriolysin O (LLO), streptolysin O (SLO) and intermedilysin (ILY). A unifying feature of these toxins is the ability to assemble into giant, membrane embedded pores [1]. Pore formation is associated with a variety of toxic functions, including escape from the intracellular phagolysosome (LLO) [3] and the delivery of folded toxins such as nicotinamide adenine-dinucleotide-glycohydrolase by SLO [4].

The structure of CDCs has been well studied. The first crystal structure of a monomeric CDC (PFO) suggested that the molecule comprises four distinct domains. Domains 1 and 3 are non-contiguous regions forming a head region that is linked via Domain 2 to the Ig-like Domain 4 (Figure 1A) [5]. The

mechanism of CDC membrane insertion has also been well characterized and mapped to the structure. Briefly, during pore formation two clusters of helices (Transmembrane Helix 1 (TMH1) and 2 (TMH2)) within Domain 3 unwind and insert into the membrane as two amphipathic  $\beta$ -hairpins. Together, Domains 1 and 3 are homologous to the distantly related MACPF proteins [6]. Domain 2, a region unique to CDCs, essentially comprises an elongated three-stranded  $\beta$ -sheet that links the pore forming head domain (Domains 1 and 3) to Domain 4. Finally, Domain 4 contains the determinants for interacting with the membrane, including a key conserved sequence that is important for binding cholesterol [7].

Single particle cryo-Electron Microscopy (SP-cryo-EM), Atomic Force Microscopy and Förster Resonance Energy Transfer (FRET) studies [8–10] have provided key insights into the transition from the prepore to the pore structure. Following interaction with the membrane surface *via* Domain 4, CDC monomers assemble into a prepore form. In this conformation, SP cryo-EM data suggest that the conformation of each subunit broadly resembles that seen in crystal structures (i.e. no major conformational change is apparent). Biophysical and microscopy

## Author Summary

Pore formation is central to the ability of cholesterol dependent cytolytins (CDCs) to act as important bacterial virulence factors. Secreted by numerous pathogens the toxins assemble into a circular ring and then perforate the target membrane to form the largest self-assembling proteinaceous pores known. In this paper we investigated computationally the conformational properties of the CDC molecule and deduced a new structural model of pore formation and membrane insertion that reconciles all experimental data. The mechanism of membrane perforation by CDCs put forward here reveals concerted and unsuspected domains motion of large amplitude, which conflicts with the currently proposed model. The work presented here procures a plausible structural mechanism of CDC oligomeric transition and furthers our understanding of pore formation by these important toxins.

data reveal that following prepore assembly, and in order to form a transmembrane pore, Domains 1 and 3 undergo a significant 40 Å movement towards the membrane surface [9,10]. Further, the cryo-EM structure of the pneumolysin pore [8] shows that the central four-stranded  $\beta$ -sheet opens, an event that separates Domains 2 and 3. Concomitant with these events, the two small clusters of  $\alpha$ -helices TMH1 and TMH2 on either side of the central sheet unwind and insert into the membrane as amphipathic  $\beta$ -strands (Figure 1B).

The conformational changes that surround Domains 1 and 3 are relatively well understood. However, a key question remains about how the prepore form collapses towards the membrane surface. Interpretation of cryo-EM data strongly suggests that Domain 2 “buckles” or “doubles over” itself. However, these data are of low resolution (29 Å) and to date it has not been possible to unambiguously model the position and conformation of Domain 2 [8]. Furthermore, attempts at conformationally trapping Domain 2 to prevent buckling have been unsuccessful [11]. Therefore understanding the structural perturbations that take place in Domain 2 remains central to understanding the mechanism of membrane insertion in CDCs.

Previous crystallographic studies have demonstrated wide variability in the position of the membrane binding Domain 4 with respect to Domains 1, 2 and 3 [12,13]. It has also been suggested that Domain 2 distortion governs different orientations of Domain 4 [12]. In contrast, a second hypothesis postulates that movement in Domain 4 is entirely attributable to a hinge bending motion located at the Domain 2/4 interface [11,14]. However, to date, there has been no family-wide description of the regions of rigidity and plasticity of the CDCs.

Here, we characterize the variability between the fifteen available CDC crystal structures and use this information to revisit the role of Domain 2 in conformational change using the published cryo-EM maps [8]. This analysis allowed a novel and methodical molecular model building strategy. Our data suggest that a rotational collapse involving Domain 2 provides the most logical mechanistic model for CDC pore formation with the current available data.

## Results and Discussion

### The CDC monomers: Rigid fragments and regions of deformation

To characterise the rigid fragments we performed superposition experiments [15] on all known CDC crystal structures (Table 1).

By first aligning the whole molecules we identified a major rigid body consisting of Domains 1, 3 (excluding the TMH2 region) and the upper part of Domain 2 close to Domain 1 and packing against TMH1 (positions 53–56; 81–90; 380–384) (Figure 2A,B). The alignment highlights the structural variability of the base of Domain 2 as well as the variability of Domain 4 orientation across the family (Figure S1). Further, structural alignments of Domains 2 and 4 either separately or together (Figure 2C,D) demonstrate that Domain 4 is to be treated as a rigid body and identify the base of Domain 2 as a region of plastic deformation.

Closer inspection of the structural alignments identified a direct relationship between the deformation of residues in the base of Domain 2 (positions 69–76; 387–390) and the orientation of Domain 4. As the overall twist in Domain 2  $\beta$ -sheet increases Domain 4 rotates away from the body of the molecule by up to 35° (Figure 3). This defines that the orientation of Domain 4 in CDCs is in part attributable to the plasticity at the base of Domain 2.

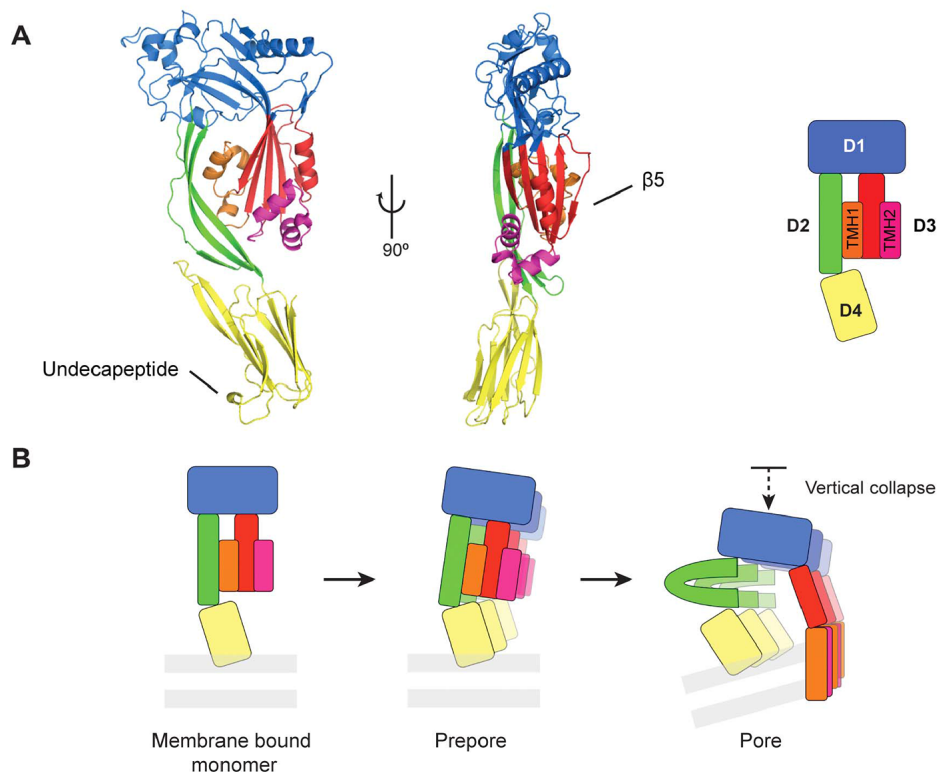
### The conformational properties of Domain 2 in PFO

We next investigated the deformation in PFO Domain 2. This is made possible by the availability of multiple structures that allow us to define the conformational space accessible to one molecule. We identified seven conformations derived from three crystal forms (Table 1 and [16]). In two conformations (PFO IIIA and C) the interface between Domains 2 and 3 is partially disrupted. PFO IIIA and IIIC conformers display a loss of contacts between the TMH2 region and Domain 2, corresponding to a loss of  $\sim 180 \text{ Å}^2$  (30%) surface area. This loss of surface area is associated with an increased distance between the pair of residues Ser287 (TMH2) and Glu388 (Domain 2) compared to the other five conformers (Table 2 and [16]). Notably these contacts all involve amino acids located at the base of Domain 2 (Ile76, Ser386 & Glu388) that we have identified as a region of plasticity (Figure 4D).

Although it has been identified that the loss of contacts at the Domains 2/4 interface is associated with distortion of the elongated Domain 2  $\beta$ -sheet [16], little characterization of this region has been performed. Consequently we focused on the deformation in PFO Domain 2 by measuring the twist of the  $\beta$ -sheet in terms of inter-strand pairing of amino acids upon partial loss of this interface (Figure 4C,D). The twist values represented in Figure 4C show moderate differences as they mostly overlap and are characterized by two peaks at positions 382 (high twist central to the sheet) and 390 (C-terminal region of the sheet) (Figure 4D). However, we note a reduced twist in conformations with a weaker interface (orange) with differences of 6° and 11° at two consecutive positions: Thr384 and Thr385. A lower twist of the sheet at these positions immediately comprised between the major central twist and residues involved in the TMH2-Domain 2 interface leads to a significant shift of  $\sim 8 \text{ Å}$  of the C-terminal segment of Domain 2 (Figure 4D). This is indicative that the partial loss of the interface is structurally coupled with un-twisting of Domain 2.

As crystallographic packing artifacts may limit our structural analysis we performed a series of molecular dynamics (MD) simulations starting from conformations with either a full (PFO I $\times$ 2, PFO IIIB $\times$ 1, summarized in Table 2) or partial (PFO IIIA $\times$ 1; PFO IIIC $\times$ 1) Domains 2/TMH2 interface. We observed that the molecule has the ability to transition between full and partial interface (PFO I Simulation 2, PFO IIIB, Figure 5 and Videos S1, S2, S3, S4, S5). Conversely PFO IIIC was able to transition from a partial interface to a full interface. Associated with the fluctuations of the interface, we observed a clustering of twist values at the positions outlined by our structural analysis





**Figure 1. CDC domain organisation and mechanism of pore formation.** A. Crystal structure of the archetypal CDC PFO and its schematic representation. Domain 1 is coloured blue, Domain 2 coloured green, Domain 3 coloured red, orange and pink and Domain 4 coloured yellow. Together Domains 1 and 3 form the ‘head’ domain distantly related to the MACPF domain. Specific transmembrane regions include the TransMembrane Helices (TMH) 1 coloured orange and TMH2 coloured pink; the strand  $\beta 5$  and the undecapeptide loop are indicated. B. Current model of CDC pore formation. After the membrane binding event, monomers oligomerize into a ring-like structure (30 to 50 monomers; prepore). Upon formation of the oligomeric pore, both helical clusters insert into the transmembrane bilayer (grey bars) as two  $\beta$ -hairpins (orange and pink) part of a giant  $\beta$ -sheet barrel. Concomitantly Domain 1 is subject to a vertical collapse associated with a proposed “buckling” of Domain 2 (reviewed in [1,58]).

doi:10.1371/journal.pcbi.1003791.g001

(Thr384-Arg80 and Thr385-Glu79, Figure 4C, Table 2). An increase in the distance between TMH2 and Domain 2 (i.e. from full to partial interface) is associated with a decrease in the twist values (Figure 5). We therefore conclude from MD simulations that PFO has the potential to fluctuate between discrete states independent of crystal packing. In addition, structural analysis of the MD simulations identified a common pattern whereby the partial loss of the interface is structurally coupled with un-twisting, or straightening, of Domain 2.

Next we analyzed our MD simulations in terms of the positions adopted by Domain 4. In all simulations performed, Domain 4 exhibited a large range of orientations (Figure 6A, Figure S2). Principal component analysis of the MD simulations identified such domain movement as the slowest ‘breathing’ mode of motion with a minimum of the total variance explained of 40% and a collectivity of typically 0.6 across all simulations. Furthermore, in simulations where the Domain2/TMH2 interface transitions between full and partial states (PFO I simulation 2, PFO IIIB) we found that a single mode (the third slowest mode in both cases) best described the departure of the base of Domain 2 from TMH2. Notably, this mode also

encapsulated accompanying motion of Domain 4, rotating away from the body of the molecule (Figure 6B). Thus, MD simulations further support that the orientation of Domain 4 in PFO can be ascribed to the plasticity at the base of Domain 2.

### The conformation properties of CDCs

To extend our investigation of the conformational properties of CDCs we performed additional MD simulations of other members of the family including intermedilysin, suilysin, anthrolysin and streptolysin. The major observation is the consistent flexibility of Domain 4 with respect to domains 1–3, similar to the flexibility observed for PFO (Figure 6A, Figure S2B, Videos S6,S7).

Secondly, for ILV IA and to a smaller extent for SLV simulations we observed a reduction of twist where the Domain 2/TMH2 interface is partially lost (Figure S2A). This correlates with the observations from the PFO simulations. We also noted that partial loss of the interface was associated with a similar decrease in twist at nearby positions (ILV IA (105–412) and SLV (73–383), not shown). This data suggests the torsion of the  $\beta$ -sheet may be subtly modulated in a toxin-dependent fashion.

**Table 1.** CDC crystal structures used in this study.

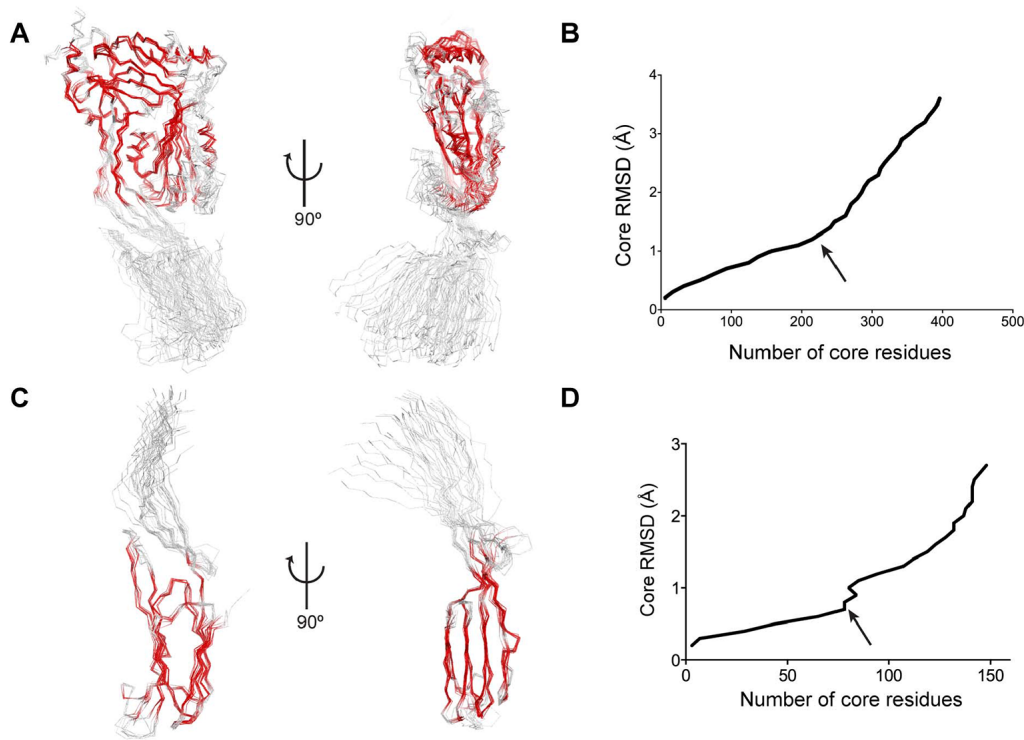
CDC (abbreviation)	PDB ID	Resolution (Å)	Molecules in ASU	Abbreviation used in this study
Perfringolysin O (PFO)	1pfo	2.20	1	PFO I
	1m3j	2.90	2	PFO II A, B
	1m3i	3.00	4	PFO III A, B, C, D
Intermedilysin (ILY)	1s3r	2.60	2	ILY I A, B
	4bik	3.49	2	ILY II A, B
Anthrolysin O (ALO)	3cqf	3.10	2	ALO A, B
Streptolysin O (SLO)	4hsc	2.10	1	SLO
Suilylin (SLY)	3hvn	2.85	1	SLY

ASU stands for asymmetric unit.  
doi:10.1371/journal.pcbi.1003791.t001

Overall these data, although non-exhaustive, support that two characteristics emerge as unifying features across the CDC family: the substantial relative flexibility of Domain 4 and the ability of Domain 2 to straighten upon weakening of the Domain 2/TMH2 interface.

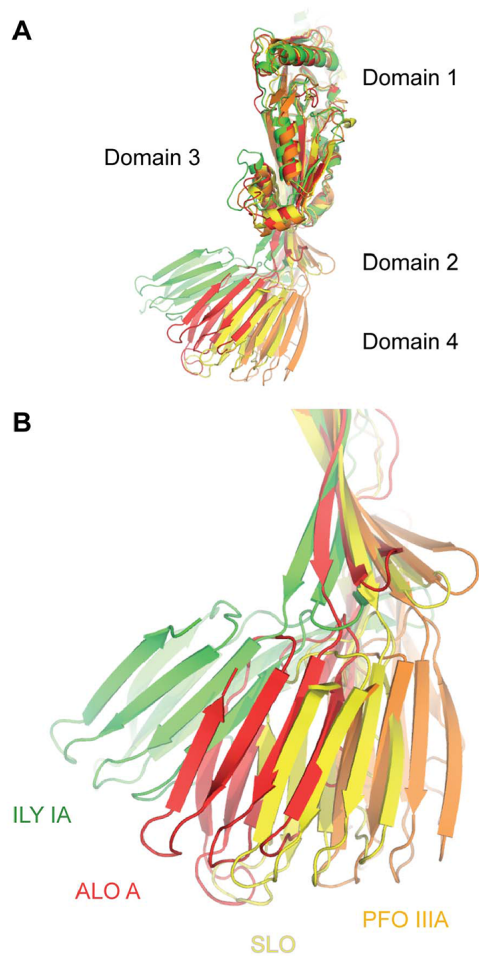
Taken together, our cross-comparison of all available CDC structures and MD simulations analysis allow us to define the conformational properties of the molecule in the monomeric form.

Domain 2 wraps around TMH1 and the base of TMH2 and most likely prevents their premature release and aggregation of the molecule (Figure 1A and as demonstrated in the case of LLO [17]). Our data suggest that preservation of the interface of TMH1 and TMH2 with Domain 2 is accompanied by conformational torsion in Domain 2 and twisting of the elongated sheet. We suggest that release of this interface, an early and critical step of membrane insertion [18], results in the straightening of Domain 2.



**Figure 2. Identification of rigid fragments in CDC structures.** A & B. Structural alignment of all CDC structures and corresponding Lesk-Hubbard plot. 3.7 Å C $\alpha$ -RMSD for 429 conserved positions; the common core at 1.3 Å sieving RMSD (indicated by the arrow) in B in red (228 positions). In Domain 3, the  $\beta$ 5 strand and the TMH2 helical bundle are regions of plastic deformation. C & D. Structural alignment of all CDC Domain 2 and 4 structures and corresponding Lesk-Hubbard plot. 1.9 Å C $\alpha$ -RMSD for 157 residues; common core at 0.8 Å sieving RMSD (indicated by the arrow) in red (78 positions). The undcapeptide and the protruding  $\beta$ -hairpin are regions of plasticity.  
doi:10.1371/journal.pcbi.1003791.g002





**Figure 3. Twist of Domain 2 and its influence on Domain 4 orientation.** A. Variations in Domain 4 orientation across the CDC family. Superposition is the same as Figure 2A. Only four representative structures are shown, which cover the entire range of Domain 4 orientation in CDCs. B. Coupling between Domain 2 twist and Domain 4 orientation. Domain 3 is omitted for clarity. doi:10.1371/journal.pcbi.1003791.g003

Domain 2 has been first proposed to undergo some conformational change after loss of TMH1/2 contacts in order to account for the 40 Å vertical collapse observed upon prepore-to-pore transition [9]. Tilley and coworkers [8] hypothesized that one way to account for this collapse was for the triple stranded  $\beta$ -sheet of Domain 2 to fold sharply in half. Here, and in contrast to this idea, our combined structural analysis and MD simulations suggest that Domain 2 has the propensity to straighten upon the loss of TMH1/2 contacts. Moreover, we argue that an energetic requirement to preserve inter-strand hydrogen bonds and local packing in anti-parallel  $\beta$ -sheets favours continuous deformation of the Domain 2 region rather than a major collapse [19]. Given this analysis, we re-visited the conformational states of prepore and pore in CDCs with improved knowledge of their conformational properties.

### The domain architecture of the prepore oligomer pre-exists in the isolated monomer

Next we examined the conformation of the PLY monomer within the prepore oligomer. We modeled the prepore conformation within the available cryo-EM map with 31-fold circular symmetry (C31) [8] (Figure 7A,B; detailed in Methods). After the flexible fitting step the cross-correlation coefficient (CCC) for the oligomer improved from 0.57 to 0.61. The structural transitions accompanying the assembly of the prepore are well described. First the conformationally labile  $\beta 5$  strand rotates away from the  $\beta 4$  strand leaving its edge exposed to the formation of mainchain hydrogen bonds with the  $\beta 1$  strand of an adjacent monomer [20]. Our model takes this structural change into account. The  $\beta 5$  strand is modeled here as a short helix by analogy with the structurally equivalent position in the complement component C6 [21] (Figure S3), a member of the distantly but structurally and functionally related MACPF family [6,22]. Secondly, the oligomer transitions to a SDS-resistant prepore upon the formation of specific  $\beta 1$ – $\beta 4$  contacts [20,23]. Our prepore model displays some, but not all, oligomeric  $\beta 1$ – $\beta 4$  mainchain hydrogen bonds compatible with the pattern later displayed by the pore form [24,25] (Figure S4).

In agreement with previous modeling we find that the prepore conformation can be explained solely by a tilt ( $\sim 40^\circ$ ) of Domain 4 with respect to the long axis of the molecule [8] (Figure 7B,C). Such orientation is supported by the determined solvent exposure of amino acids of Domain 4 [9] (Figure S5). Based on our previous structural and MD simulations analyses we hypothesize that this minor domain rearrangement may be attributable to the intrinsic flexibility of Domain 4 and deformation of Domain 2, although the low resolution of the cryo-EM data (28 Å) prevents further interpretation. In addition, the relative orientation of Domain 4 in the prepore conformation is broadly similar to the crystallographic conformation of ILY IB (Figure 7C), and within the range of observed CDC crystallographic conformations (Figure 7D). Stereochemical features of the model are reported in Table 3. Therefore this indicates that the CDC monomer in the prepore form adopts an arrangement of domains readily accessible in the soluble form, which is conformationally trapped upon oligomerization.

To validate this prepore model a suggested site of interaction would be between strand 2 of Domain 2 of one molecule with the TMH1 of the adjacent molecule (respectively Thr86-Ser88 and Lys201-Asn205, PFO numbering). This could be performed using either disulphide bond formation or short crosslinkers. Disulphide bond formation experiments have successfully characterized Domain 3 oligomeric interactions in PFO [24].

### A novel domain and subdomain re-organization in the pore form

Finally we investigated the conformation of CDCs in the pore form to address how CDCs change conformation particularly with respect to Domain 2. Our CDC pore model is presented in Figure 8. Domain 1 can be fitted intact (see Methods, CCC for the individual domain of 0.70) into the cryo-EM density in agreement with Tilley et al. [8]. Together with Domains 1 of adjacent subunits the domains exhibit packing similar to the prepore complex. TMH1/2 are entirely restructured from bundles of  $\alpha$ -helices to a giant transmembrane  $\beta$ -barrel, concomitantly with the opening of Domain 3. It has been established that the amino acids forming the  $\beta$ -barrel adopt a novel  $\beta$ -barrel architecture specific to CDCs [24,25]: the membrane-embedded  $\beta$ -hairpins adopt a  $20^\circ$  tilt to the axis normal to the membrane (Figure 8B). This departs

**Table 2.** Domain 2-Domain 3 interface features and Molecular simulations performed.

Molecule/Isoform	Domains 2/3 Interface Area (Å <sup>2</sup> )	TMH2/Domain 2 C $\alpha$ distance (Å) <sup>a</sup>	MD simulation performed <sup>b</sup>
PFO I	578	4.8 (S287-E388)	2 (1)
PFO IIA	609	5.2	-
PFO IIB	641	5.3	-
<b>PFO IIIA</b>	<b>416</b>	<b>10.2</b>	1 (0)
PFO IIIB	592	5.5	1 (1)
<b>PFO IIIC</b>	<b>430</b>	<b>8.7</b>	1 (1)
PFO IIID	611	5.4	-
ILY IA	577	6.1 (K313-V413)	2 (1)
ILY IB	610	6.3	-
ILY IIA	612	6.3	-
ILY IIB	590	7.1	-
ALO A	651	5.6 (K299-T399)	2 (0)
ALO B	679	5.6	-
SLO	595	5.2 (K357-T458)	2 (0)
SLY	593	6.4 (S284-S383)	2 (1)

<sup>a</sup>Pairs of residues with corresponding numbering are given in brackets.

<sup>b</sup>The number of simulations where transitions of the Domain 2/TMH2 interface occurs is given in brackets. Except for PFO I MD simulations, only one simulation is presented in this work.

doi:10.1371/journal.pcbi.1003791.t002

from a pore model where the  $\beta$ -hairpins are proposed to stand perpendicular (0° tilt) to the membrane surface [8]. The modeled  $\beta$ -barrel is in full agreement with the experimentally established amphipathic pattern of the membrane-spanning PFO  $\beta$ -hairpins [26,27] (CCC of 0.57, Figure S6).

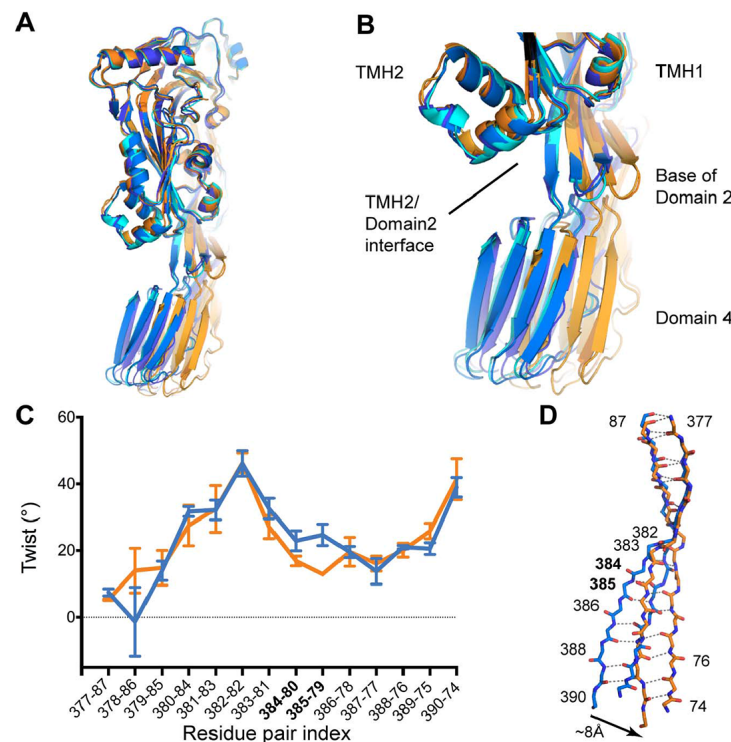
The position of the membrane binding Domain 4 is readily identifiable in the density [8] with only the region surrounding the undecapeptide loop inserting into the upper leaflet of the membrane bilayer [28,29]. Domain 4 loosely packs against its adjacent counterparts as demonstrated [29]. Its orientation, tilted but not lying on the membrane surface, is further supported by the pattern of solvent exposure of amino acids distributed on the domain's surface [29] (CCC of 0.48, Figure S7).

Taken together, our structural and MD analyses suggest that Domain 2 does not favour the proposed bending [8]. Instead our data suggest that the  $\beta$ -sheet Domain 2 simply untwists and rotates with respect to Domain 4. Motivated by this finding we modeled the elongated  $\beta$ -sheet without altering its structural integrity (see Methods). We found that Domain 2 can be fitted in the density linked to Domain 1 and the Domain 4 of the adjacent monomer (clockwise when viewed from the top; Figure 8C). This dramatic sideways rotation has not been postulated to date, yet it provides a good fit of the domains within the cryo-EM maps. The Domain 2 region in our model has a CCC of 0.52 with no clashes observed between residues, which is excellent for this resolution of cryo-EM data. This is an improvement on the existing unrefined model, which has a CCC of 0.45 for the Domain 2 region (PDB ID 2bk1). Following the flexible fitting step the 38-mer exhibited a best CCC value of 0.67, an improvement on the initial 0.62.

In support of this model we investigated the Domain 2 boundaries with Domain 1 and Domain 4. Our new model preserves the hydrophobic Domain 2/4 interface, with Domain 2 linked to Domain 4 by a glycine linker. There is an introduction of a kink of ~40° at the Domain 1/2 interface. The Domain 1/2 interface is constituted by three mainchain covalent links and contains no secondary structure elements or specific contacts. In

the pore conformation, Domain 2 orientation is at a ~25° angle to the bilayer surface and extends the range of orientations observed in crystallographic structures (Figure 8D, Figure S8). Our model also suggests that the orientation of Domain 2 is constrained by the packing of its adjacent counterparts with the possibility of mainchain parallel hydrogen bonds between positions 54–56 and 384–386 (PFO numbering) of an adjacent monomer. In analyzing conformations fitted at such resolution (29 Å) it should be noted that the predicted interaction is indicative of the close proximity of individual Domains 2 in the pore form (Figure S9). Therefore experiments designed to test this hypothesis should take this aspect into account and may include the use of techniques such as FRET, disulfide bond formation and chemical crosslinking. If strands of Domain 2 are close enough to establish parallel hydrogen bonds then the close proximity can be tested by formation of disulfide bonds. Alternatively in the case of less intimate contacts, chemical crosslinking would be more appropriate. We suggest that probing the Domain 2 oligomeric contact is well suited to distinguishing between the ‘buckling’ model and our proposed model of CDC pore assembly.

To assess the stability of our pore model the fitted conformation was energy-minimized and subjected to a brief MD simulation in a membrane bilayer environment and free of all constraints. After 15 ns of simulation the assembly reached a plateau at 4.4 Å over the last 5 ns (Figure S10). The general subunits arrangement remained stable with little deviation from the initial conformation. Minor structural deviations included a difference in the orientation of Domain 4 as well as its penetration into the membrane bilayer. Since the details of its contacts with the bilayer are currently lacking slight deviations are not unexpected. We also noted in the monomer situated at the clockwise end of the tetramer an increased divergence of Domain 2 and the transmembrane Domain 3 (Figure S10). Given their positions at one extremity of the assembly we concluded that this is likely to be attributable to the lack of explicit circular symmetry in our setup (see Methods) and the absence of buttressing/specific contacts with the adjacent



**Figure 4. PFO Domain 4 and twist of Domain 2.** A. Variations in Domain 4 orientation for the PFO structures. Superposition is the same as in Figure 2A. Only five structures are shown for clarity: PFO I (blue); PFO IIB (dark blue); PFO IIID (cyan); PFO IIIA (orange); PFO IIIC (light orange). B. Coupling between PFO Domain 2 twist and Domain 4 orientation. Domain 3 is omitted for clarity. C. Twist values as a function of amino acid pairs for the continuous strands forming Domain 2. The seven crystal conformations are classified in two groups: conformations with partial loss of Domain 2/4 interface (interface  $\sim 425 \text{ \AA}^2$ ; PFO IIIA & IIIC) in orange and all the others in blue (interface  $\sim 605 \text{ \AA}^2$ ; Table 2). Minimum and maximum values are displayed for each group. D. Representative mainchains of Domain 2  $\beta$ -sheet in stick representation. Blue: PFO I; orange: PFO IIIA. Superposition is the same as in Figure 2A. Numbers indicate amino-acid positions, dashed lines indicate mainchain hydrogen bonds.  
doi:10.1371/journal.pcbi.1003791.g004

monomer. We found the MD simulation demonstrates the overall stability of our pore model and reflects the quality of its stereochemistry (Table 3).

### Conclusions. Prepore-to-pore transition: A new model of pore formation

This study has also allowed us to map the rigid bodies present in CDCs and their spectrum of flexibility relative to each other. Following this analysis we have revisited both prepore and pore conformations employing available cryo-EM data.

Importantly we find the pore conformation can be modeled without potentially energetically costly restructuring of Domain 2. Instead our modeling suggests that simple domains rotations can account for the well-documented CDC vertical collapse [8–10].

In addition, both our prepore and pore conformations define a pathway for the most logical mechanism of pore formation. Only a coordinated vertical collapse together with rotations of Domains 1/3 ( $\sim 10^\circ$ ) is compatible with the extent of Domain 2 rotation ( $\sim 60^\circ$ ) from a nearly perpendicular to the membrane surface conformation (prepore) to nearly parallel (pore). Furthermore, the oligomeric packing and specific contacts established in the prepore form [23] are likely to impose the constraints that result in a downward spiral rotation of Domains 1/3 (counter-clockwise

rotation corresponding to a monomer and vertical collapse; Figure 9A–C) within the entire oligomeric assembly. This movement defines an unprecedented and orchestrated global motion whereby the prepore transitions to the pore form by rotation of Domain 2 of all subunits, which brings the CDC head domains (Domains 1/3) closer to the membrane surface.

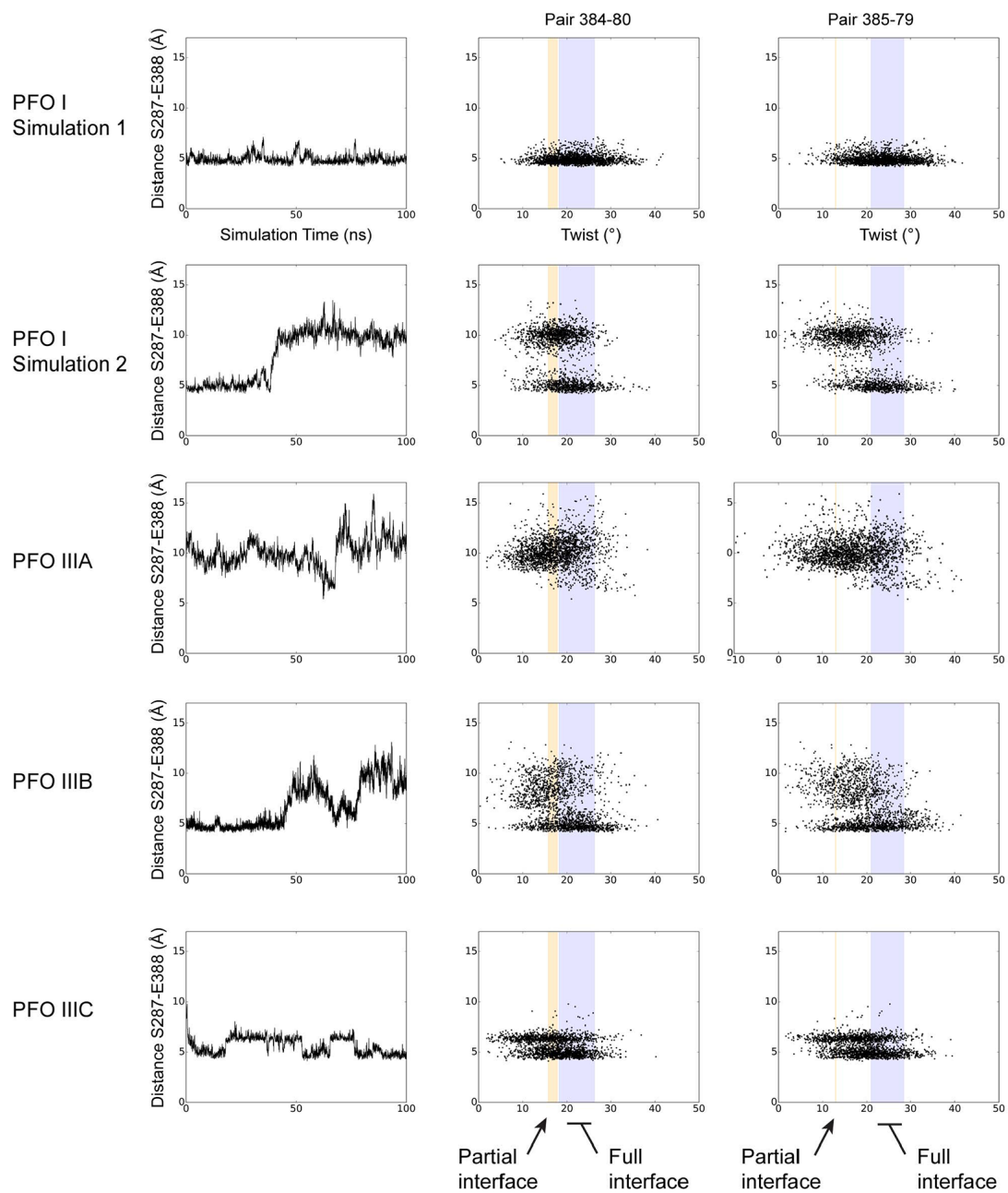
Interestingly, a recent study on aerolysin proposed that a swirling-like motion is central to the mechanism of pore formation by this toxin. While aerolysin and the CDCs are not related, the mechanism we propose is somewhat mechanically analogous [30].

Thus, to conclude, we propose that CDCs achieve pore formation by employing large, concerted domains rotations (schematically summarized in Figure 9D). Our work supports a new model of membrane insertion for CDCs in considerable departure from the currently accepted model. This mechanism presents mechanical similarities to other  $\beta$ -pore forming toxins and a new, testable model of pore formation for CDCs.

### Methods

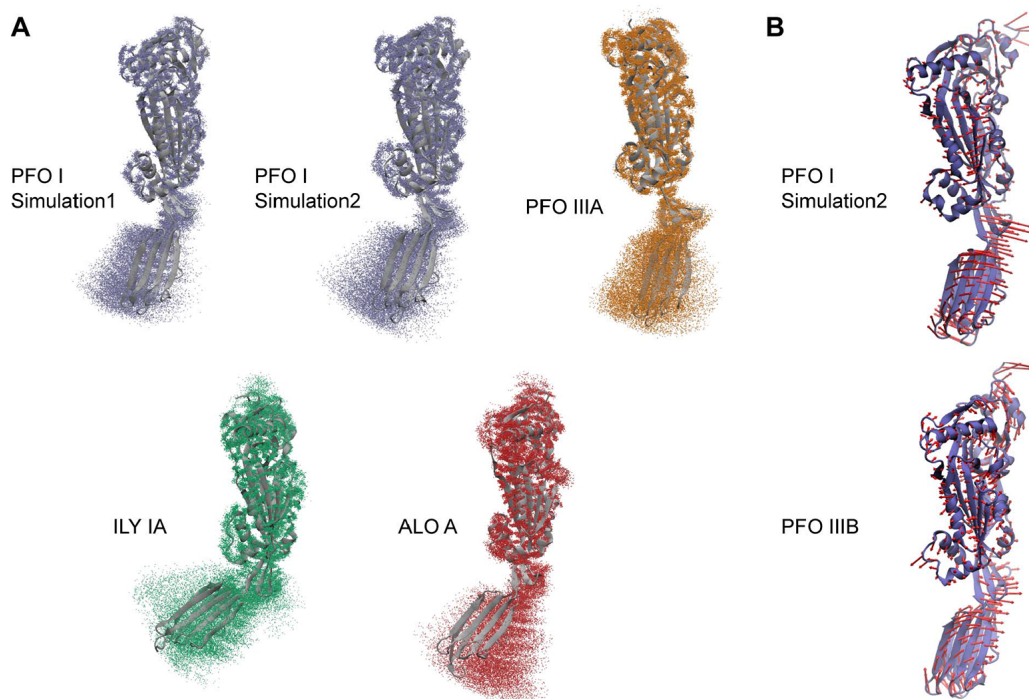
#### Structural alignments and analysis

The identification of common substructures from structural alignments is a powerful approach that allows us to extract the



**Figure 5. Plasticity and deformation of Domain 2 from PFO MD simulations.** Analysis of the different MD simulations for PFO I, PFO IIIA, PFO IIIB, PFO IIIC (indicated at the left of each panel). Left panel: distance representative of the Domain 2/TMH2 interface between residues 287 and 388 (see also Table 2). Centre and right panels: values of twist at positions discussed in the text and illustrated in Figure 4C plotted versus the distance between residues in left panel. The coloured vertical bars correspond to the range of twist values derived from the analysis of the PFO crystal structures (Figure 4B). The pairs of residues considered are indicated at the top of the figure.  
doi:10.1371/journal.pcbi.1003791.g005





**Figure 6. Flexibility of Domain 4 with respect to Domains 1–3 in MD simulations.** A. Each panel corresponds to an MD simulation. The starting conformer (cartoon representation, grey) is indicated at the left of the molecule. C $\alpha$  positions taken from snapshots of the simulations are represented as dots after alignment on Domains 1–3. B. Arrows (red) indicate the directions of motion of the third slowest mode (scaled to 3.5 Å for clarity), which best describes the increased distance between Domain 2 and TMH2. The starting conformer is indicated at the left of its cartoon representation. The reasonably small fractional variance explained by the modes (11% (PFO I) and 8% (PFO IIIB), collectivity of 0.34 and 0.45 respectively) identifies a more localized nature of the motion. This suggests that such movement participates to, or enhances, the observed flexibility of Domain 4.

doi:10.1371/journal.pcbi.1003791.g006

rigid fragments and conformational properties [15,31] of the CDC molecule. The method provides a standardized way to identify common structural cores of homologous proteins (the “sieving” procedure) [32,33] although it cannot unambiguously distinguish between conformational change and structural divergence. It is, however, important to note that analysis of structures of one protein, and/or homologous proteins, determined in different conditions captures the conformational features across a family [15,34–36].

All structural alignments were performed with Mustang-MR [37] with domains definition as reported by Rossjohn et al. [5]. Examination and analysis were undertaken using Prody 1.4 [38], Pymol 1.3 [39] and VMD 1.9 [40].  $\beta$ -sheet twist values are the angles between mainchain vectors of residues in an inter-strand pair and were calculated following Ho and Curmi [41]. Accessible surface areas are as reported by PISA [42]. MD simulations were analyzed with VMD and Prody. PCA were performed with Prody employing 10 K snapshots collected every 10 ps (C $\alpha$  coordinates).

#### Monomeric CDC molecular dynamics simulations

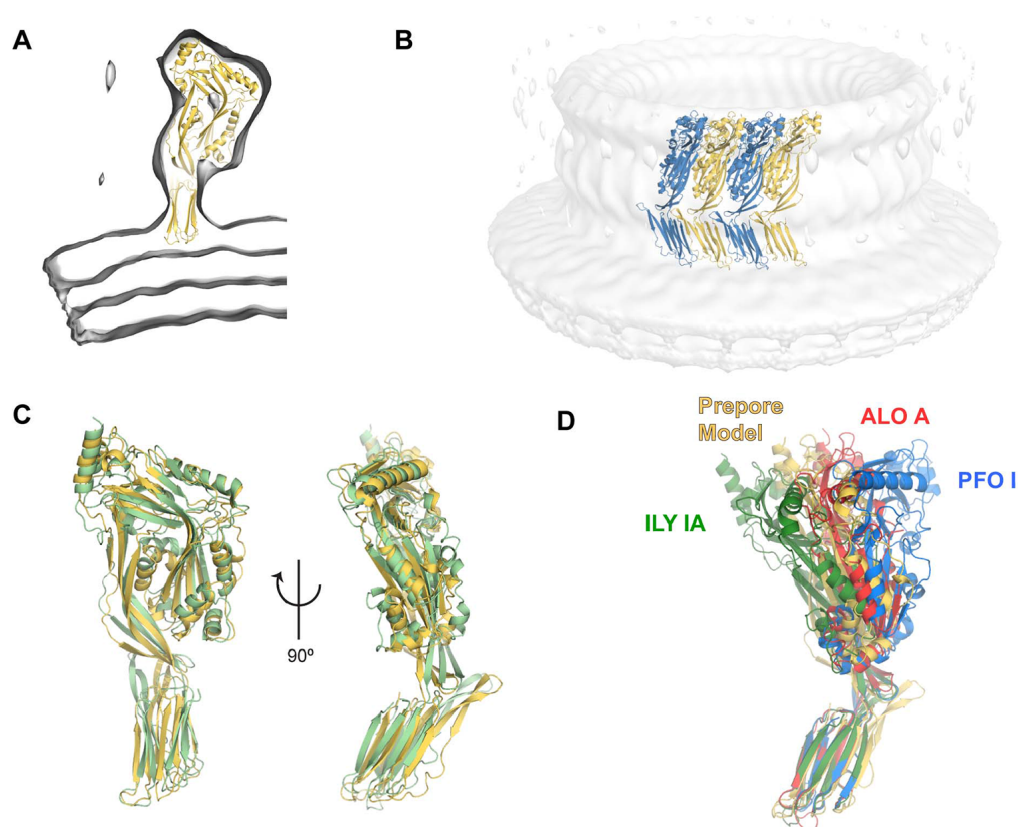
Initial conformations and MD simulations performed are reported in Table 2. In all cases topologies were built and solvated using *leap* [43] and the Amber ff99SB force field [44]. MD simulations employed truncated octahedron water boxes (TIP3P, 12 Å padding), sodium and chloride ions were added to charge

neutrality. The systems were typically comprised of 164,000 to 171,000 atoms. Temperature was maintained at 300 K using Langevin dynamics with a damping constant of 5 ps<sup>-1</sup>. Pressure was maintained at 1 atm with a Nosé-Hoover-Langevin piston and Periodic Boundary Conditions (PBC) were used. An integration time step of 2 fs was used, short-range forces and long-range electrostatics were calculated every time step. Non-bonded interactions employed a 10 Å cut-off and long-range electrostatic forces were computed by the particle-mesh Ewald (PME) summation method (grid spacing smaller than 1 Å). All systems were subjected to equilibration steps with harmonic restraints first applied to all heavy protein atoms (100 ps, 1 fs time step for this step only), followed by restraints applied only to mainchain atoms (250 ps) and finally C $\alpha$  atoms (500 ps). 100 ns MD simulations were then produced and analyzed after typically removing the initial 2 to 3 ns. All the simulations were conducted with NAMD v2.9 [45].

#### Prepore model

To obtain initial models of the prepore conformation we first performed rigid body docking into the cryo-EM density of representative CDC crystallographic structures in order to leverage the conformational variability highlighted in our analysis.

Four crystallographic structures were used: PFO I, PFO IIIA, ILY IA, ILY IB (cf. Table 1). The PLY sequence (UniProt ID:

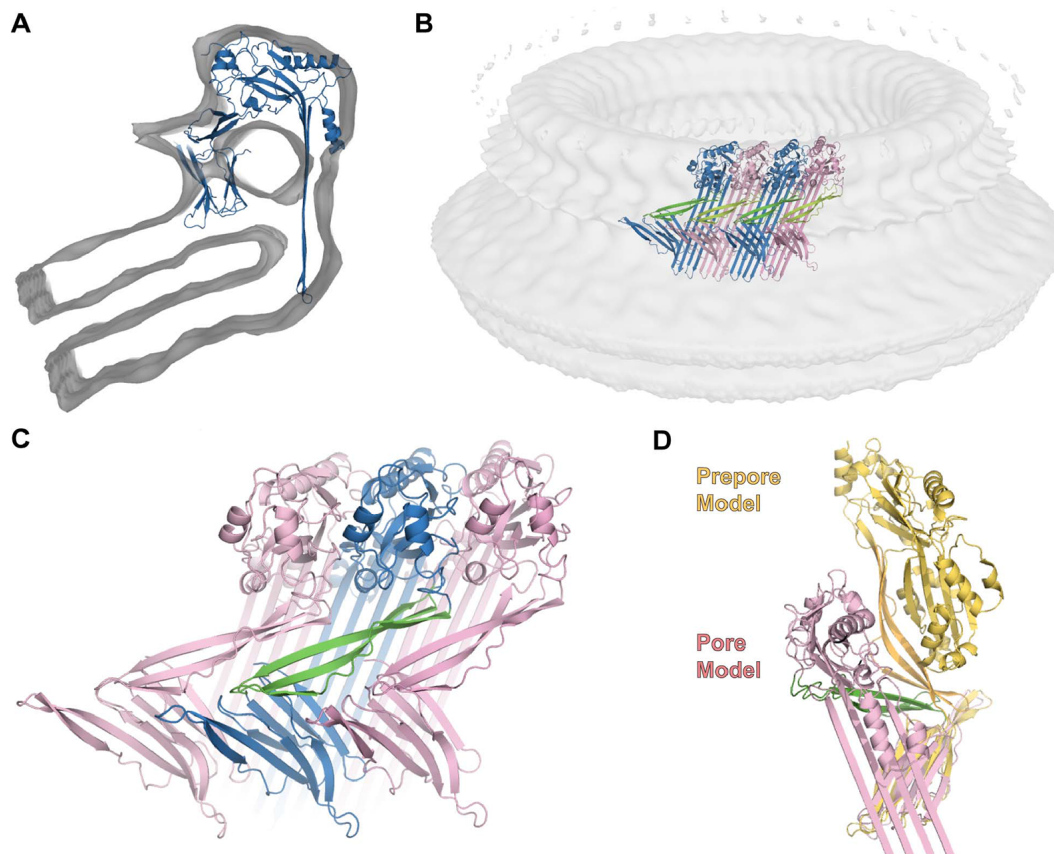


**Figure 7. Prepore conformation of the CDC molecule.** A. View of the CDC monomer in the prepore conformation within the cryo-EM map (transparent surface). B. Overall oligomeric arrangement within the cryo-EM map. Only the tetramer used in the modeling is shown (see Methods). C. Structural alignment of the ILY IB conformation (pale green) to the prepore conformation (pale orange) (4.4 Å; 448 positions). D. Prepore conformation in the context of representative CDC crystallographic structures. doi:10.1371/journal.pcbi.1003791.g007

Q04IN8) was first threaded onto each structural template employing Modeller 2.11 [46] and amino acids corresponding to β5 were simply discarded. The coordinates of the loop at positions 95 to 101 (PFO numbering) were also discarded as the corresponding amino acids were found to produce steric hindrance upon oligomeric assembly. The coordinates were then docked using Situs [47] into the cryo-EM map reconstruction of the PLY prepore (EMDB ID: 1106) with the density of the

Table 3. Stereochemistry indicators for the atomic models.					
Indicator	Prepore		Pore		
	Initial <sup>a</sup>	Final	Initial	Final	Minimized
Clashscore	0.37	0.00	1.69	0.00	0.00
Poor rotamers (%)	2.08	6.12	1.77	7.59	4.28
Ramachandran outliers (%)	1.02	4.28	1.23	2.68	2.14
Ramachandran favored (%)	93.68	91.43	94.87	89.88	91.97
Cβ violations	0.84	0.68	0.51	1.24	0.39
Bad bonds (%)	0.00	0.00	0.01	0.01	0.00
Bad angles (%)	0.34	0.90	0.35	1.49	0.43

Stereochemistry indicators are as reported by Molprobit v4.1 [59].  
<sup>a</sup>Values are only reported for the model with optimal CCC.  
doi:10.1371/journal.pcbi.1003791.t003



**Figure 8. Pore conformation of the CDC molecule.** A. Cut view of the CDC monomer in the pore conformation within the cryo-EM map (transparent surface). B. Subunits arrangement in the pore (cut view). The tetramer shown is the symmetrically modeled tetramer (see Methods). Subunits have alternate colouring with Domain 2 coloured green. C. Domains 2 arrangement in the pore viewed from outside of the ring. A Domain 2 is highlighted in green sandwiched by two adjacent monomers. D. Prepore and pore conformations aligned on Domain 4. Prepore Domain 2 is in orange; pore Domain 2 is in green.  
doi:10.1371/journal.pcbi.1003791.g008

membrane discarded for this step only as including the bilayer density produced unrealistic placements. The docked individual subunits displayed CCCs ranging from 0.66 to 0.76 (truncated map) and 0.63 to 0.71 (no truncation).

Secondly, for each template the docked structure with the highest CCC was then replicated with C31 symmetry and the positions and orientations of all subunits refined against the density with Situs. Finally, four consecutive subunits (referred to as tetramer in the following) were selected and the missing loop modeled with Modeller, with amino acids corresponding to the  $\beta 5$  region modeled as an  $\alpha$ -helix (see text). This led to the production of five initial PLY conformations with a best CCC of 0.57.

Each tetramer was then subjected to a step of flexible fitting into the cryo-EM density with C31 symmetry restraints (see relevant section). A total of 12 models were thus produced with an average root mean square deviation (rmsd) of  $2.4 \pm 0.9$  Å and an average CCC of  $0.60 \pm 0.01$ . Although small differences in orientation of the domains were observed (as reflected by the rmsd) all models presented the same domain architecture (discussed in the text). C $\alpha$  coordinates of the model with optimal CCC (0.61) are provided as Dataset S1.

### Pore model

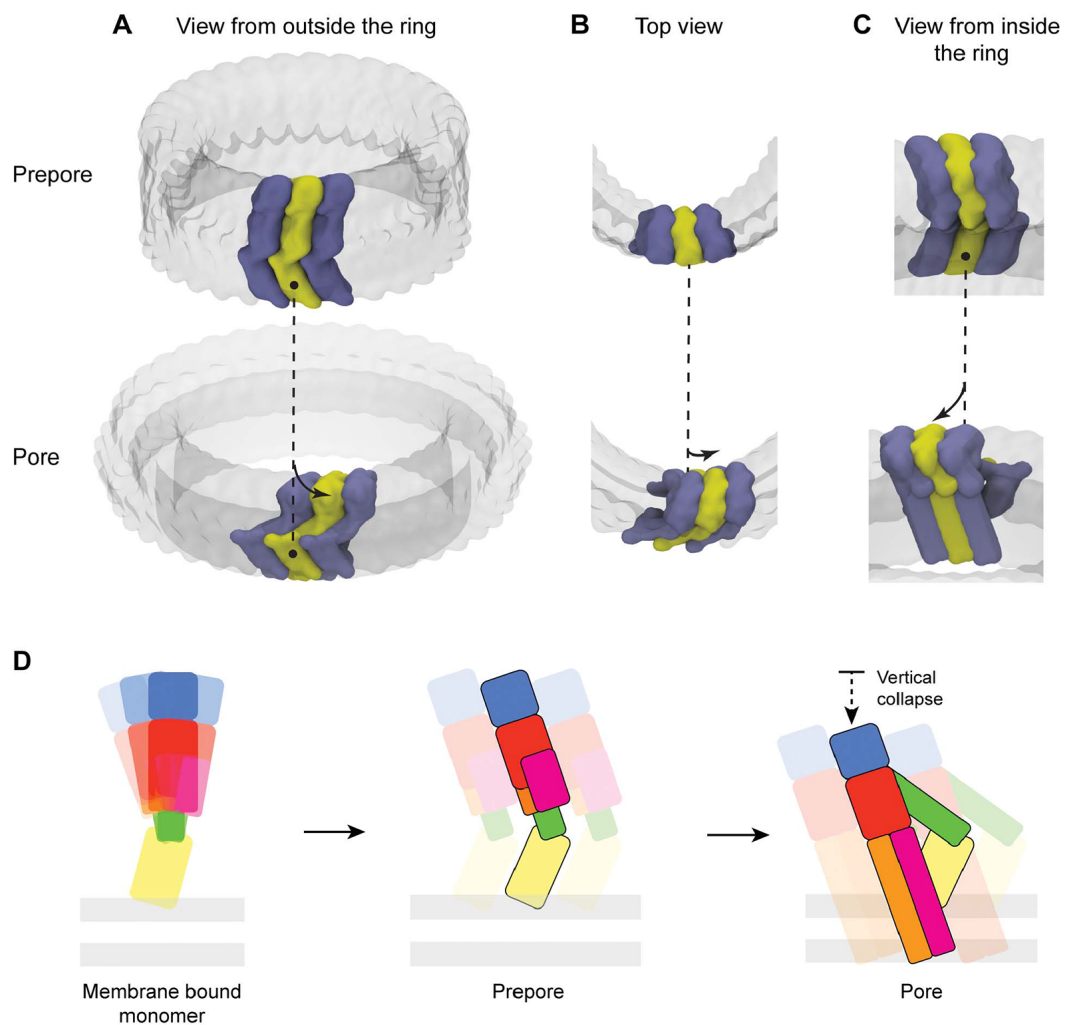
Domain 1 was initially placed manually in the cryo-EM density (EMD ID: 1107; C38 symmetry; 29 Å resolution) and its position refined locally in the presence of symmetric subunits employing Chimera 1.8 [48]. Here we modeled Domain 3 as a  $\beta$ -barrel with architecture  $S=n/2$  as detailed in our previous work [25]. Domain 1 orientation was then adjusted to satisfy both a reasonable fit to the density as well as the distance constraints from the four covalent bonds linking Domain 1 and the four  $\beta$ -strands forming the  $\beta$ -barrel (Domain 3).

Before refinement into the cryo-EM density the altitude of Domain 4 was adjusted so as to place the amino acids identified as exposed to solvent and buried in the membrane [29] (Figure S6).

Domain 2 was placed manually in the density without modifications to its internal structure, consistent with the conclusions of our structural analysis. Its initial placement also satisfies the distance constraints imposed by the covalent bonds to Domain 1 and 4 and a reasonable fit to the cryo-EM map.

Our pore model was built as a tetramer with C38 symmetry. Initial coordinates of PFOIII-A were employed with the PLY





**Figure 9. Orchestrated domain movement and proposed mechanism of pore formation.** Schematic representation of the proposed new model of prepore (top row) to pore (bottom row) transition shown from three point of views: A) from outside the ring formed by the oligomeric complex, B) Top view and C) from inside the ring. The dashed line symbolizes the position of Domain 4 in the prepore complex. The arrows symbolize the concerted movement experienced by the globular head domain (Domains 1 & 3) upon membrane insertion. Alternate monomers are represented with alternate colors; only three monomers are displayed for clarity; remaining monomers are represented as a transparent surface. D. Schematic representation of proposed CDC pore formation. The monomer displays flexibility in the orientation of Domains 1–3 versus the membrane binding Domain 4. Upon self-oligomerization into the prepore complex the monomer is trapped in a monomer accessible conformation. Upon membrane insertion, the orientation of Domain 2 flattens with respect to the membrane surface. This is accompanied by a vertical collapse of Domain 1 and 3, which brings them closer to the bilayer surface and allows insertion of TMH1/2 as  $\beta$ -hairpins. Colours are identical to Figure 1. doi:10.1371/journal.pcbi.1003791.g009

sequence threaded. Initial positions of all domains were adjusted to remove inter-subunit steric clashes. Furthermore, the initial coordinates were perturbed by  $1^\circ$  clockwise and anti-clockwise rotations around the pore axis thus producing three starting points for the flexible fitting step. The best CCC was 0.62 for the initial conformation.

The three tetramers were subjected to a step of flexible fitting into the cryo-EM density with C38 symmetry (see relevant section). Final average rmsd was  $1.4 \pm 0.4$  per monomer and the average CCC for the 38-mer assembly was  $0.66 \pm 0.01$ . C $\alpha$

coordinates of the model with optimal CCC (0.67) are provided as Dataset S2.

#### Flexible fitting into cryo-EM maps

Flexible fitting was performed following the MDFF methodology [49] with NAMD 2.9. Symmetry restraints were employed for the prepore and pore conformation with the corresponding circular symmetries [50]. All MDFF simulations employed the CHARMM36 forcefield [51] *in vacuum* (long range interactions were cut off at 12 Å; dielectric constant of 80; 1 fs time-step;



298 K). Additional restraints were applied to preserve correct stereochemistry and prevent structural distortions [52] (secondary structure restraints force constant of 200 kcal mol<sup>-1</sup> rad<sup>-2</sup>).

In each MDFF simulation minimization and equilibration steps were as follows: 10,000 steps of minimization with non-hydrogen atoms harmonically constrained, 100,000 steps of equilibration with protein main-chain constrained and 100,000 steps with C $\alpha$  atoms constrained. Three 5,000,000 steps MDFF runs were then performed with linearly increasing symmetry restraints to a final force constant of 5 kcal mol<sup>-1</sup> Å<sup>-2</sup> driving the system to the desired circular symmetry, and each with different grid force scaling parameter  $\xi = 0.2; 0.3; 0.5$ . Convergence was reached in all cases. Each of the three conformations obtained was subjected to 10,000 steps of minimization with  $\xi = 1.0$ . Therefore each starting conformation produced three final conformations.

#### Pore model molecular dynamics simulation

A mixed square bilayer membrane patch was generated with DMPC and cholesterol (50/50 ratio, 1050 molecules each) with CHARMM-GUI 1.4 [53,54] and the CHARMM36 forcefield [55] and equilibrated for 4 ns following the CHARMM-GUI provided settings in the presence of a TIP3P water layer of 15 Å thickness, 0.15 M sodium/chloride ions and planar constraints with the NAMD 2.9 software. The dimensions of the equilibrated bilayer system were 205 Å × 205 Å × 71 Å.

The tetramer pore model was energy-minimized for 2,500 steps free of cryo-EM restraints employing the Generalized Born/Solvent Accessible Surface Area implicit solvation [56]. Bilayer and solvent were then added and their height manually adjusted to match the position of the bilayer as judged from the cryo-EM density. Waters and lipids within 1.4 Å of the protein assembly were discarded. TIP3P waters were then added to a system of initial dimensions 205 Å × 205 Å × 175.5 Å. Ions were added to 0.15 M and system charge neutrality. The system was heated to 300 K and equilibrated in steps for 4 ns, first melting the lipid tails and cholesterol, then the headgroups and solvent and finally smoothly relaxing harmonic restraints on the tetramer. Care was taken to keep water molecules outside of the bilayer in the first steps of equilibration.

The system was then simulated for 15 ns in the NPAT ensemble free of constraints with PBC. Temperature was maintained at 300 K using Langevin dynamics with a damping constant of 1 ps<sup>-1</sup>. Pressure was maintained at 1 atm with a Nosé-Hoover-Langevin piston. An integration time step of 1 fs was used, short-range forces and long-range electrostatics were calculated every 1 and 2 fs respectively. Non-bonded interactions employed a 12 Å cut-off with a shorter Lennard-Jones switching function (11 to 12 Å) [55], long-range electrostatic forces were computed by the PME summation method (grid spacing smaller than 1 Å). The final dimensions of the system were 204 Å × 204 Å × 164 Å (709,660 atoms). Simulations were performed with the Multimodal Australian ScienceS Imaging and Visualisation Environment (MASSIVE) [57].

#### Supporting Information

**Dataset S1** PLY prepore model C $\alpha$  coordinates.  
(PDB)

**Dataset S2** PLY pore model C $\alpha$  coordinates.  
(PDB)

**Figure S1** Structural alignment illustrating the variability of CDC structures. Conformers of PFO with a tight

Domain2/TMH2 interface (see Table 2) are in blue, conformers with a weaker interface are in orange. ILY conformers are in green; ALO in red (both conformers are represented and have an overall rmsd <0.1 Å); SLO in yellow and SLV in teal.  
(PDF)

**Figure S2** Domain 2 plasticity and Domain 4 flexibility in CDCs. A. Left panel: distance of the Domain 2/TMH2 interface (see also Table 2). Center and right panels: values of twist at positions discussed in the text. The coloured vertical bars correspond to the range of twist values derived from the structural analysis. The pairs of residues considered are indicated at the top of each plot. The starting conformation for each MD simulation is indicated on the left of each panel. B. Each panel corresponds to an MD simulation whose starting conformation (cartoon representation, grey) is indicated at the left of the molecule. C $\alpha$  positions taken from snapshots of the simulations are represented as dots after alignment on Domains 1–3 of each CDC molecule.  
(PDF)

**Figure S3** Proposed conformational change involving residues of the  $\beta 5$  strand. Coloring of C6 (pdb id: 3t5o) mimics CDC structurally equivalent positions.  
(PDF)

**Figure S4** Monomer-Monomer  $\beta 1$ – $\beta 4$  Hydrogen bonds in the prepore model. Mainchain atoms are represented in stick with one monomer in yellow and the adjacent monomer in blue. Dashed lines display the hydrogen bonds present in the model. The residues corresponding to the  $\beta 5$  strand of the yellow monomer are not displayed for clarity.  
(PDF)

**Figure S5** Domain 4 residues exposure in the prepore conformation of PLY. The position of residues is indicated by spheres at their C $\alpha$  position. Asn402 (blue, Asn433 PFO numbering) was quenched by a collisional quencher in the prepore complex [9]. Lys395 (green, Lys426 PFO numbering) was not quenched. The membrane surface is defined by the cryo-EM map.  
(PDF)

**Figure S6** Exposure and location of  $\beta$ -barrel forming residues. A. Location of residues in the barrel overlaid with the cryo-EM density of the PLY pore. Residues in red have been determined to be located near the surface, in grey to be near the centre of the bilayer and in orange to be part of the hairpin turns [27]. B. Amphipathic pattern of membrane spanning amino-acids. Residues in grey have been determined to be exposed to the membrane bilayer, residues in blue have been determined to be exposed to the aqueous milieu [26,27].  
(PDF)

**Figure S7** Domain 4 residues exposure in the pore conformation. Spheres at the position of their C $\alpha$  indicate the position of residues. Only Domain 4 is shown. The residues shown are at position equivalent to PFO and in three categories: exposed (blue), interfacial (yellow) and buried (grey) as determined by Ramachandran et al. [29].  
(PDF)

**Figure S8** Pore conformation in the context of representative CDC crystallographic structures and the prepore model. The prepore model is in yellow, the pore conformation in pink. PFO I is in blue and ILY IA in green.  
(PDF)

**Figure S9 Proximity of Domain 2 to adjacent subunits in the pore form.** The regions of potential interactions (red; Ala54-Asn56 and Thr384-Ser386, PFO numbering) are discussed in the text.  
(PDF)

**Figure S10 Molecular dynamics simulation of the pore conformation.** A. Final snapshot of the simulation. The tetramer conformation is in cartoon presentation. Only the cholesterol oxygen (pink) and DMPC phosphate atoms (orange) are represented for clarity. The periodic box is depicted in grey. The regions indicated by arrows are discussed in the text. B. Tetramer rmsd plot versus simulation time.  
(PDF)

**Video S1 Molecular dynamics simulation starting from PFO I (simulation 1).** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S2 Molecular dynamics simulation starting from PFO I (simulation 2).** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S3 Molecular dynamics simulation starting from PFO IIIA.** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S4 Molecular dynamics simulation starting from PFO IIIB.** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S5 Molecular dynamics simulation starting from PFO IIIC.** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S6 Molecular dynamics simulation starting from ILY IA.** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

**Video S7 Molecular dynamics simulation starting from ALO A.** The pair of relevant residues is represented as C $\alpha$  spheres.  
(MP4)

## Acknowledgments

The authors thank Dr. Bosco K. Ho for helpful discussion.

## Author Contributions

Conceived and designed the experiments: CFR JCW MAD. Performed the experiments: CFR. Analyzed the data: CFR MAD. Contributed reagents/materials/analysis tools: CFR JCW. Contributed to the writing of the manuscript: CFR JCW MAD.

## References

- Hotze EM, Tweten RK (2012) Membrane assembly of the cholesterol-dependent cytolysin pore complex. *Biochimica et biophysica acta* 1818: 1028–1038.
- Hotze EM, Le HM, Sieber JR, Bruxvoort C, McInerney MJ, et al. (2013) Identification and characterization of the first cholesterol-dependent cytolysins from Gram-negative bacteria. *Infect Immun* 81: 216–225.
- Gaillard JL, Berche P, Mounier J, Richard S, Sansonetti P (1987) In vitro model of penetration and intracellular growth of *Listeria monocytogenes* in the human enterocyte-like cell line Caco-2. *Infect Immun* 55: 2822–2829.
- Madden JC, Ruiz N, Caparon M (2001) Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria. *Cell* 104: 143–152.
- Rosjohn J, Feil SC, McKinstry WJ, Tweten RK, Parker MW (1997) Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form. *Cell* 89: 685–692.
- Rosado CJ, Buckle AM, Law RH, Butcher RE, Kan WT, et al. (2007) A common fold mediates vertebrate defense and bacterial attack. *Science* 317: 1548–1551.
- Soltani CE, Hotze EM, Johnson AE, Tweten RK (2007) Structural elements of the cholesterol-dependent cytolysins that are responsible for their cholesterol-sensitive membrane interactions. *Proc Natl Acad Sci U S A* 104: 20226–20231.
- Tilley SJ, Orlova EV, Gilbert RJ, Andrew PW, Saibil HR (2005) Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* 121: 247–256.
- Czajkowsky DM, Hotze EM, Shao Z, Tweten RK (2004) Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. *EMBO J* 23: 3206–3215.
- Ramachandran R, Tweten RK, Johnson AE (2005) The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation. *Proc Natl Acad Sci U S A* 102: 7139–7144.
- Oloo EO, Yethon JA, Ochs MM, Carpick B, Oomen R (2011) Structure-guided antigen engineering yields pneumolysin mutants suitable for vaccination against pneumococcal disease. *J Biol Chem* 286: 12133–12140.
- Polekhina G, Giddings KS, Tweten RK, Parker MW (2005) Insights into the action of the superfamily of cholesterol-dependent cytolysins from studies of intermedilysin. *Proc Natl Acad Sci U S A* 102: 600–605.
- Bourdeau RW, Malito E, Chenal A, Bishop BL, Musch MW, et al. (2009) Cellular functions and X-ray structure of anthrolysin O, a cholesterol-dependent cytolysin secreted by *Bacillus anthracis*. *J Biol Chem* 284: 14645–14656.
- Xu L, Huang B, Du H, Zhang XC, Xu J, et al. (2010) Crystal structure of cytotoxin protein suliyisin from *Streptococcus suis*. *Protein Cell* 1: 96–105.
- Whistock JC, Skinner R, Carrell RW, Lesk AM (2000) Conformational changes in serpins: I. The native and cleaved conformations of alpha(1)-antitrypsin. *J Mol Biol* 296: 685–699.
- Rosjohn J, Polekhina G, Feil SC, Morton CJ, Tweten RK, et al. (2007) Structures of perfringolysin O suggest a pathway for activation of cholesterol-dependent cytolysins. *J Mol Biol* 367: 1227–1236.
- Schuerch DW, Wilson-Kubalek EM, Tweten RK (2005) Molecular basis of listeriolysin O pH dependence. *Proc Natl Acad Sci U S A* 102: 12537–12542.
- Hotze EM, Wilson-Kubalek EM, Rosjohn J, Parker MW, Johnson AE, et al. (2001) Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane beta-sheet from a prepore intermediate. *J Biol Chem* 276: 8261–8268.
- Salemme FR, Weatherford DW (1981) Conformational and geometrical properties of beta-sheets in proteins. II. Antiparallel and mixed beta-sheets. *J Mol Biol* 146: 119–141.
- Ramachandran R, Tweten RK, Johnson AE (2004) Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment. *Nature structural & molecular biology* 11: 697–705.
- Aleshin AE, Schraufstatter IU, Stec B, Bankston LA, Liddington RC, et al. (2012) Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC). *J Biol Chem* 287: 10210–10222.
- Rosado CJ, Kondos S, Bull TE, Kuiper MJ, Law RH, et al. (2008) The MACPF/CDC family of pore-forming toxins. *Cell Microbiol* 10: 1765–1774.
- Hotze EM, Wilson-Kubalek E, Farrand AJ, Bentsen L, Parker MW, et al. (2012) Monomer-monomer interactions propagate structural transitions necessary for pore formation by the cholesterol-dependent cytolysins. *J Biol Chem* 287: 24534–24543.
- Sato TK, Tweten RK, Johnson AE (2013) Disulfide-bond scanning reveals assembly state and beta-strand tilt angle of the PFO beta-barrel. *Nat Chem Biol* 9: 383–389.
- Reboul CF, Mahmood K, Whistock JC, Dunstone MA (2012) Predicting giant transmembrane beta-barrel architecture. *Bioinformatics* 28: 1299–1302.
- Shatursky O, Heuck AP, Shepard LA, Rosjohn J, Parker MW, et al. (1999) The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* 99: 293–299.
- Shepard LA, Heuck AP, Hamman BD, Rosjohn J, Parker MW, et al. (1998) Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy. *Biochemistry* 37: 14563–14574.
- Heuck AP, Hotze EM, Tweten RK, Johnson AE (2000) Mechanism of membrane insertion of a multimeric beta-barrel protein: perfringolysin O creates a pore using ordered and coupled conformational changes. *Mol Cell* 6: 1233–1242.

29. Ramachandran R, Heuck AP, Tweten RK, Johnson AE (2002) Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin. *Nature Structural Biology* 9: 823–827.
30. Degiacomi MT, Iacovache I, Pernot L, Chami M, Kudryashev M, et al. (2013) Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nat Chem Biol* 9: 623–629.
31. Perica T, Chothia C, Teichmann SA (2012) Evolution of oligomeric state through geometric coupling of protein interfaces. *Proc Natl Acad Sci U S A* 109: 8127–8132.
32. Irving JA, Whisstock JC, Lesk AM (2001) Protein structural alignments and functional genomics. *Proteins* 42: 378–382.
33. McPhalen CA, Vincent MG, Picot D, Jansonius JN, Lesk AM, et al. (1992) Domain closure in mitochondrial aspartate aminotransferase. *J Mol Biol* 227: 197–213.
34. Best RB, Lindorff-Larsen K, DePristo MA, Vendruscolo M (2006) Relation between native ensembles and experimental structures of proteins. *Proc Natl Acad Sci U S A* 103: 10901–10906.
35. Whisstock JC, Pike RN, Jin L, Skinner R, Pei XY, et al. (2000) Conformational changes in serpins: II. The mechanism of activation of antithrombin by heparindaggar. *J Mol Biol* 301: 1287–1305.
36. Friedland GD, Lakomek NA, Griesinger C, Meiler J, Kortemme T (2009) A correspondence between solution-state dynamics of an individual protein and the sequence and conformational diversity of its family. *PLoS Comput Biol* 5: e1000393.
37. Konagurthu AS, Reboul CF, Schmidberger JW, Irving JA, Lesk AM, et al. (2010) MUSTANG-MR structural sieving server: applications in protein structural analysis and crystallography. *PLoS One* 5: e10048.
38. Bakan A, Meireles LM, Bahar I (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* 27: 1575–1577.
39. Schrödinger LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
40. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33–38, 27–38.
41. Ho BK, Curmi PM (2002) Twist and shear in beta-sheets and beta-ribbons. *J Mol Biol* 317: 291–308.
42. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797.
43. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, et al. (2012) AMBER 13. University of California, San Francisco.
44. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65: 712–725.
45. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802.
46. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics/editorial board, Andreas D Baxevanis [et al]* Chapter 5: Unit 5.6.
47. Wriggers W (2010) Using Situs for the integration of multi-resolution structures. *Biophys Rev* 2: 21–27.
48. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612.
49. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16: 673–683.
50. Chan KY, Gumbart J, McGreevy R, Watermeyer JM, Sewell BT, et al. (2011) Symmetry-restrained flexible fitting for symmetric EM maps. *Structure* 19: 1211–1218.
51. Huang J, MacKerell AD, Jr. (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* 34: 2135–2145.
52. Schreiner E, Trabuco LG, Freddolino PL, Schulten K (2011) Stereochemical errors and their implications for molecular dynamics simulations. *BMC Bioinformatics* 12: 190.
53. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 29: 1859–1865.
54. Jo S, Lim JB, Klauda JB, Im W (2009) CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* 97: 50–58.
55. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, et al. (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114: 7830–7843.
56. Sitkoff D, Sharp KA, Honig B (1994) Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *The Journal of Physical Chemistry* 98: 1978–1988.
57. Goscinski WJ, McIntosh P, Felzmann U, Maksimenko A, Hall CJ, et al. (2014) The multi-modal Australian ScienceS Imaging and Visualization Environment (MASSIVE) high performance computing infrastructure: applications in neuroscience and neuroinformatics research. *Front Neuroinform* 8: 30.
58. Dunstone MA, Tweten RK (2012) Packing a punch: the mechanism of pore formation by cholesterol dependent cytolysins and membrane attack complex/perforin-like proteins. *Curr Opin Struct Biol* 22: 342–349.
59. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66: 12–21.

## **Supplementary Material**

### **Dataset S1.**

**PLY prepore model Ca coordinates.**

doi:10.1371/journal.pcbi.1003791.s001  
(PDB)

### **Dataset S2.**

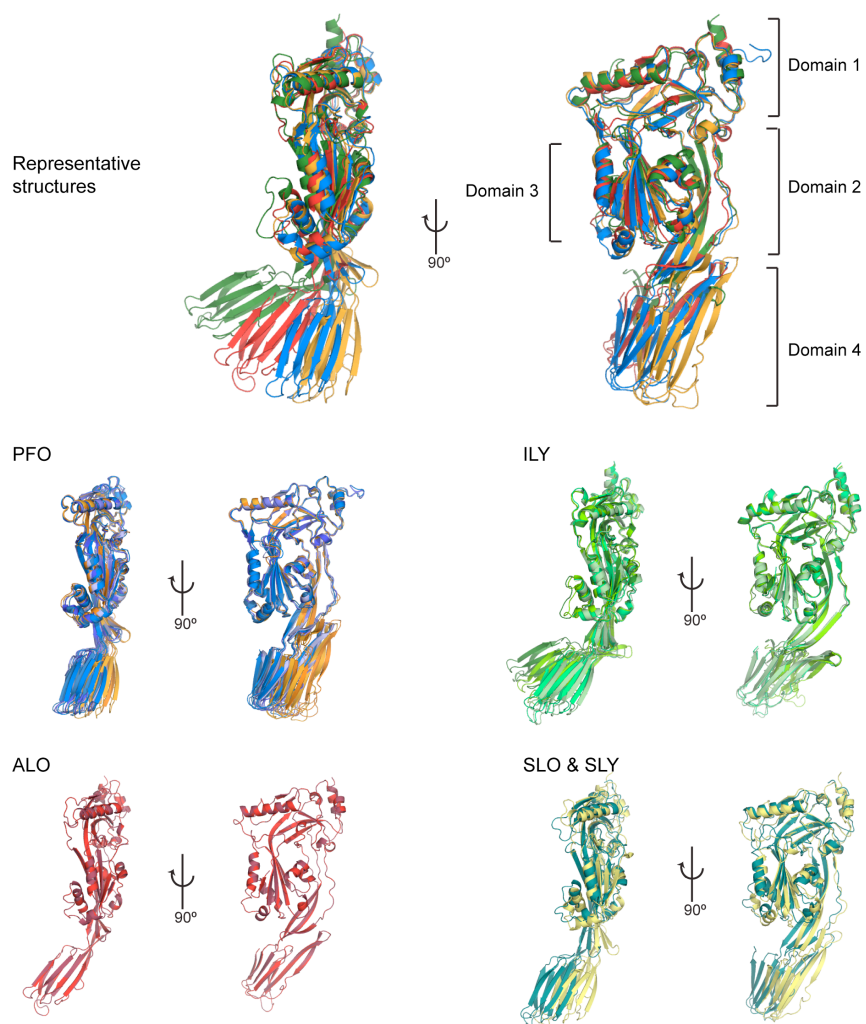
**PLY pore model Ca coordinates.**

doi:10.1371/journal.pcbi.1003791.s002  
(PDB)

### Figure S1.

**Structural alignment illustrating the variability of CDC structures.** Conformers of PFO with a tight Domain2/TMH2 interface (see Table 2) are in blue, conformers with a weaker interface are in orange. ILY conformers are in green; ALO in red (both conformers are represented and have an overall rmsd <0.1 Å); SLO in yellow and SLY in teal.

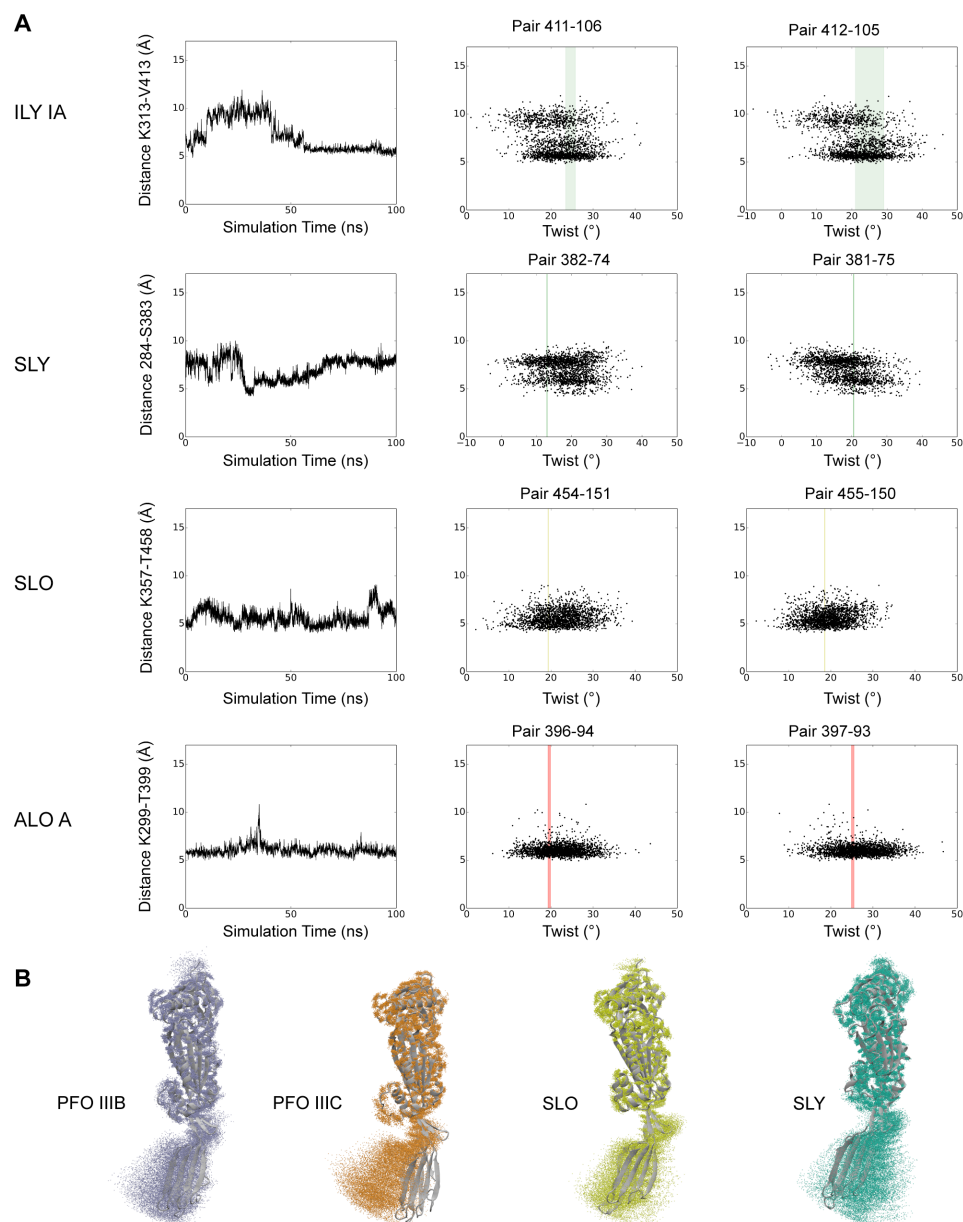
doi:10.1371/journal.pcbi.1003791.s003



## Figure S2.

**Domain 2 plasticity and Domain 4 flexibility in CDCs.** A. Left panel: distance of the Domain 2/TMH2 interface (see also Table 2). Center and right panels: values of twist at positions discussed in the text. The coloured vertical bars correspond to the range of twist values derived from the structural analysis. The pairs of residues considered are indicated at the top of each plot. The starting conformation for each MD simulation is indicated on the left of each panel. B. Each panel corresponds to an MD simulation whose starting conformation (cartoon representation, grey) is indicated at the left of the molecule. Ca positions taken from snapshots of the simulations are represented as dots after alignment on Domains 1–3 of each CDC molecule.

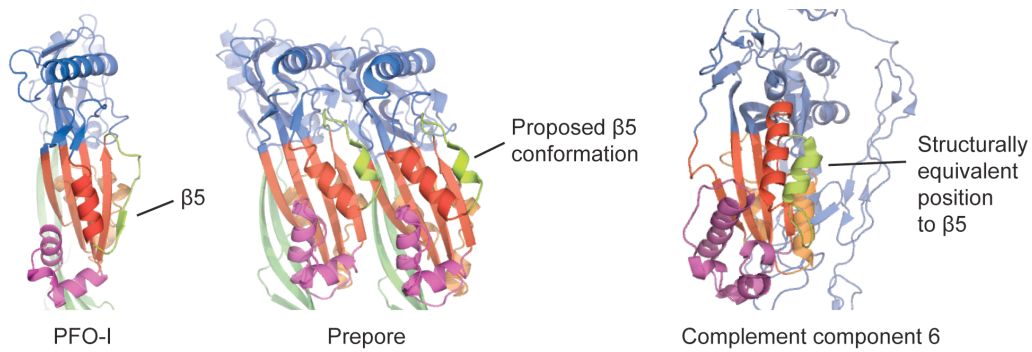
doi:10.1371/journal.pcbi.1003791.s004



**Figure S3.**

**Proposed conformational change involving residues of the  $\beta 5$  strand.** Coloring of C6 (pdb id: 3t5o) mimics CDC structurally equivalent positions.

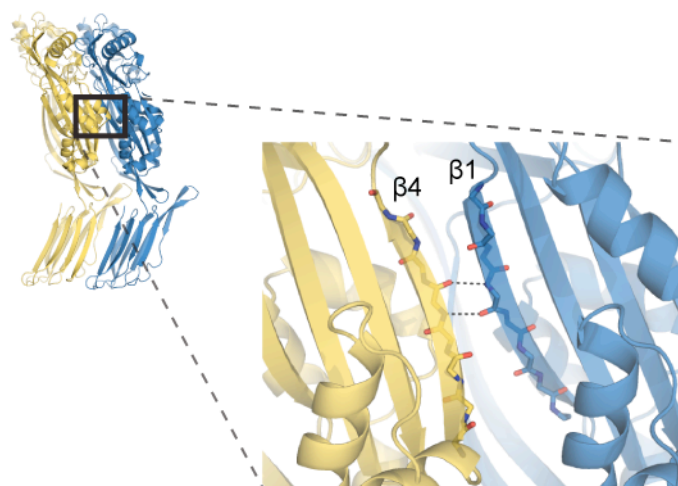
doi:10.1371/journal.pcbi.1003791.s005



**Figure S4.**

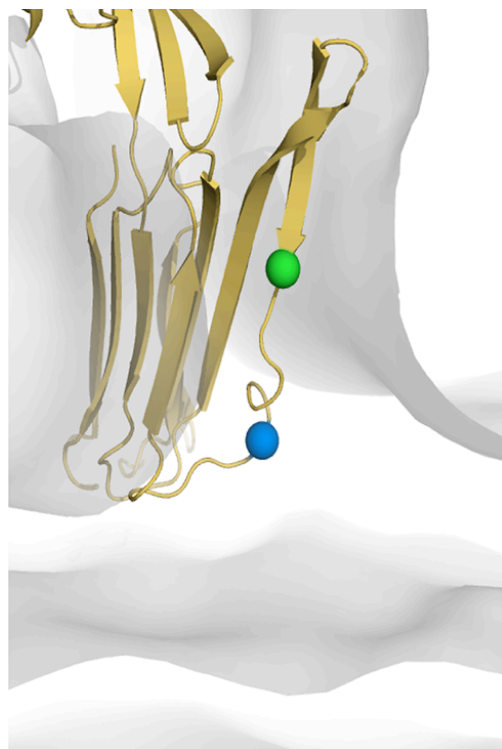
**Monomer-Monomer  $\beta 1$ – $\beta 4$  Hydrogen bonds in the prepore model.** Mainchain atoms are represented in stick with one monomer in yellow and the adjacent monomer in blue. Dashed lines display the hydrogen bonds present in the model. The residues corresponding to the  $\beta 5$  strand of the yellow monomer are not displayed for clarity.

doi:10.1371/journal.pcbi.1003791.s006



**Figure S5.**

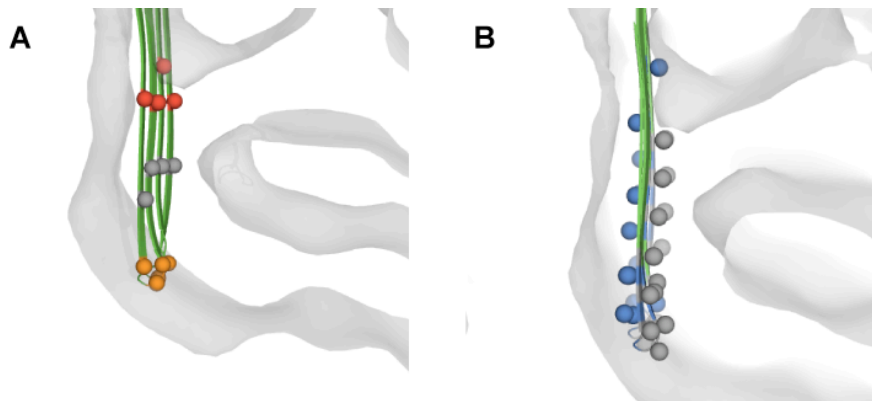
**Domain 4 residues exposure in the prepore conformation of **PLY**.** The position of residues is indicated by spheres at their Ca position. Asn402 (blue, Asn433 PFO numbering) was quenched by a collisional quencher in the prepore complex [9]. Lys395 (green, Lys426 PFO numbering) was not quenched. The membrane surface is defined by the cryo-EM map.  
doi:10.1371/journal.pcbi.1003791.s007





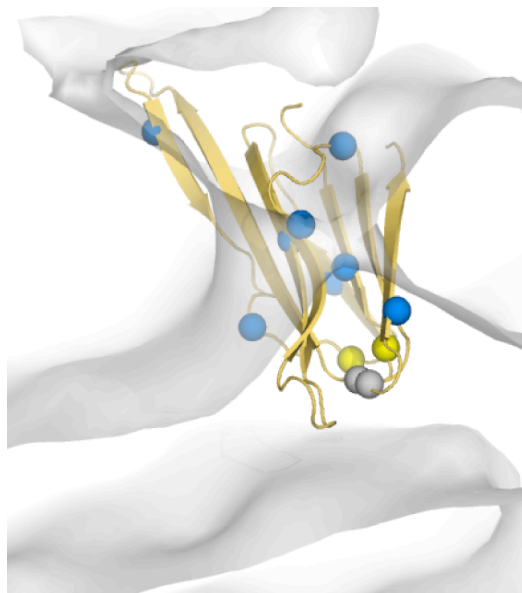
**Figure S6.**

**Exposure and location of  $\beta$ -barrel forming residues.** A. Location of residues in the barrel overlaid with the cryo-EM density of the PLY pore. Residues in red have been determined to be located near the surface, in grey to be near the centre of the bilayer and in orange to be part of the hairpin turns [27]. B. Amphipathic pattern of membrane spanning amino-acids. Residues in grey have been determined to be exposed to the membrane bilayer, residues in blue have been determined to be exposed to the aqueous milieu [26], [27].  
doi:10.1371/journal.pcbi.1003791.s008



**Figure S7.**

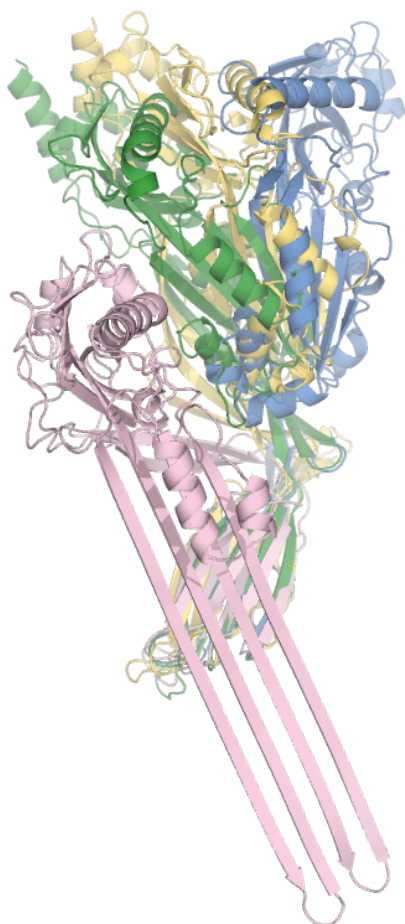
**Domain 4 residues exposure in the pore conformation.** Spheres at the position of their Ca indicate the position of residues. Only Domain 4 is shown. The residues shown are at position equivalent to PFO and in three categories: exposed (blue), interfacial (yellow) and buried (grey) as determined by Ramachandran et al. [29].  
doi:10.1371/journal.pcbi.1003791.s009



**Figure S8.**

**Pore conformation in the context of representative CDC crystallographic structures and the prepore model.** The prepore model is in yellow, the pore conformation in pink. PFO I is in blue and ILY IA in green.

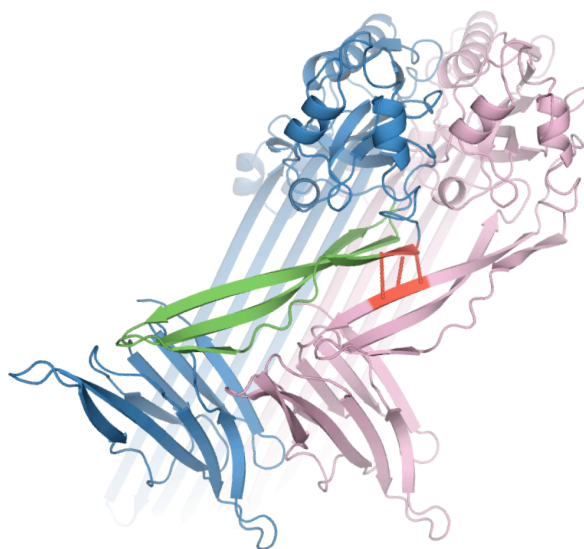
doi:10.1371/journal.pcbi.1003791.s010



**Figure S9.**

**Proximity of Domain 2 to adjacent subunits in the pore form.** The regions of potential interactions (red; Ala54-Asn56 and Thr384-Ser386, PFO numbering) are discussed in the text.

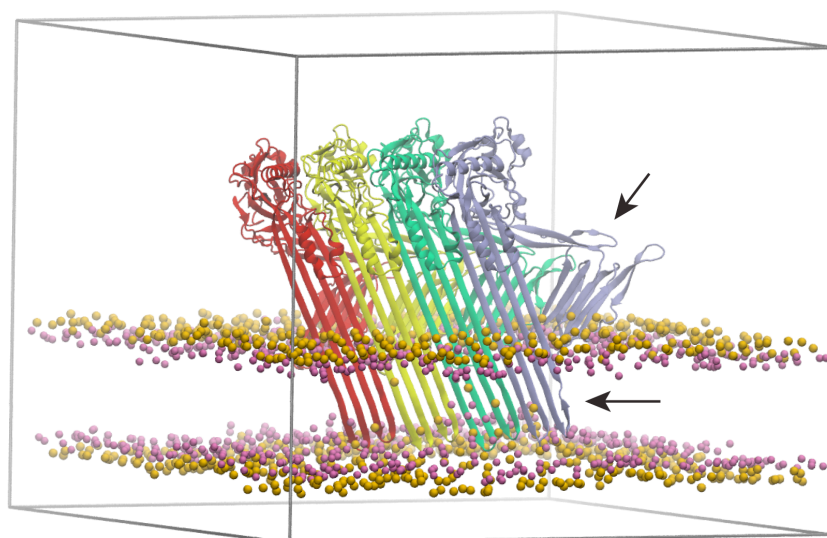
doi:10.1371/journal.pcbi.1003791.s011



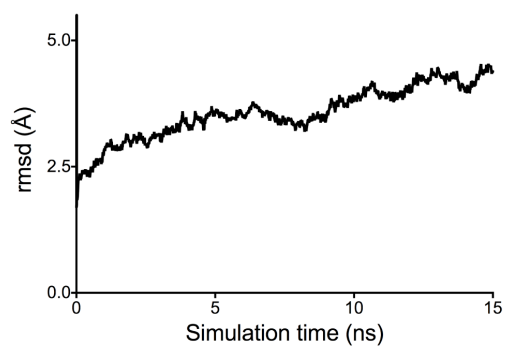
**Figure S10.**

**Molecular dynamics simulation of the pore conformation.** A. Final snapshot of the simulation. The tetramer conformation is in cartoon presentation. Only the cholesterol oxygen (pink) and DMPC phosphate atoms (orange) are represented for clarity. The periodic box is depicted in grey. The regions indicated by arrows are discussed in the text. B. Tetramer rmsd plot versus simulation time.  
doi:10.1371/journal.pcbi.1003791.s012

**A**



**B**



**Video S1.**

**Molecular dynamics simulation starting from PFO I (simulation 1).** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s013

(MP4)

**Video S2.**

**Molecular dynamics simulation starting from PFO I (simulation 2).** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s014

(MP4)

**Video S3.**

**Molecular dynamics simulation starting from PFO IIIA.** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s015

(MP4)

**Video S4.**

**Molecular dynamics simulation starting from PFO IIIB.** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s016

(MP4)

**Video S5.**

**Molecular dynamics simulation starting from PFO IIIC.** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s017

(MP4)

**Video S6.**

**Molecular dynamics simulation starting from ILY IA.** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s018

(MP4)

**Video S7.**

**Molecular dynamics simulation starting from ALO A.** The pair of relevant residues is represented as Ca spheres.

doi:10.1371/journal.pcbi.1003791.s019

(MP4)



## **Chapter 4**

### **Structure of the poly-C9 component of the Complement Membrane Attack Complex**

## Declaration for Thesis Chapter 4

### Declaration by candidate

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Developed the structural model of the poly-C9 oligomer based on the 3D cryo-EM reconstruction and performed structural analysis	30

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
NV Dudkina	Collected EM data and generated 3D cryo-EM reconstruction	
BA Spicer	Produced, analysed and optimised samples for EM experiments	30
PJ Conroy	Assisted protein production and sample optimisation	
N Lukoyanova	Assisted EM data collection	
H Elmlund	Performed computational analysis	
RHP Law	Assisted protein production and sample optimisation	
SM Ekkel	Assisted protein production and sample optimisation	
SC Kondos	Assisted protein production and sample optimisation	
RJA Goode	Performed and analysed Mass Spectrometry experiments	
G Ramm	Assisted sample preparation for EM experiments	
JC Whisstock	Designed experiments, analysed data, wrote manuscript and led the research	



<b>HR Saibil</b>	Designed experiments, analysed data, wrote manuscript and led the research	
<b>MA Dunstone</b>	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

<b>Candidate's Signature</b>		<b>Date</b> 13/11/15
<b>Main Supervisor's Signature</b>		<b>Date</b> 16/11/15

## Preface to Chapter 4

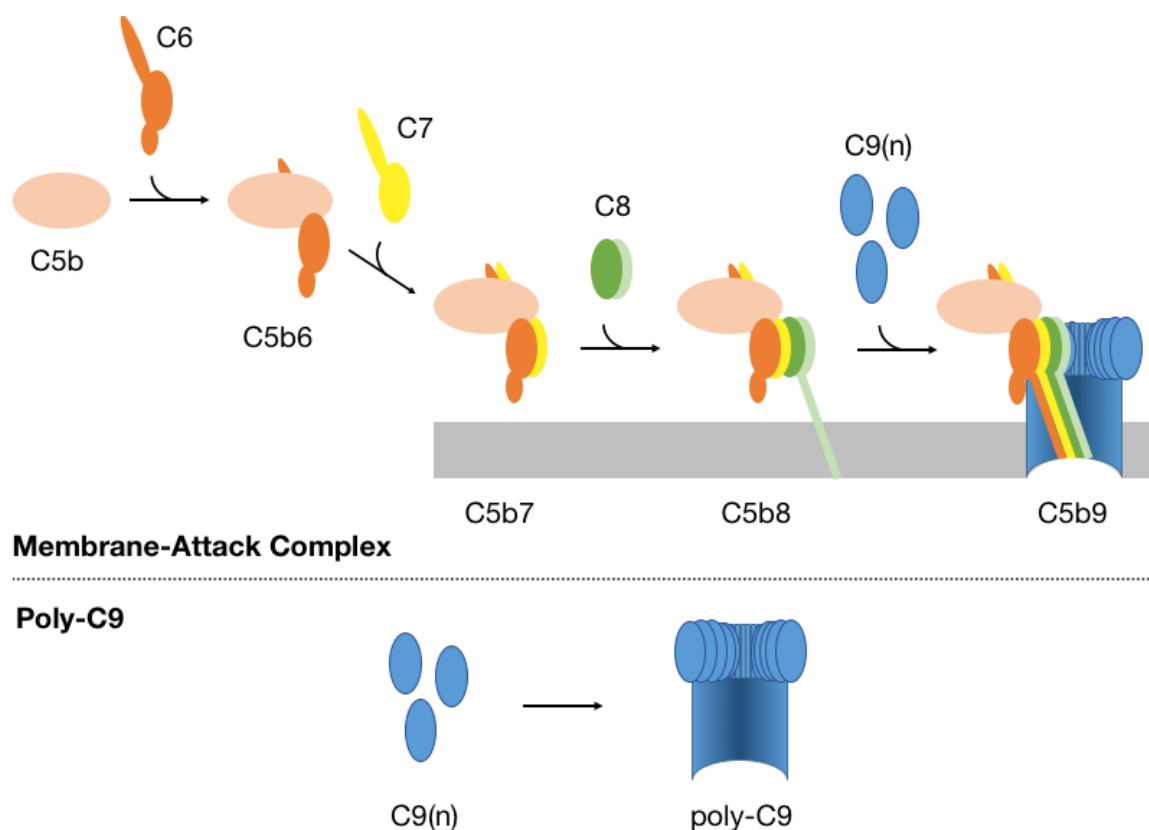
This chapter presents our structural model and analysis of the C9 polymer (poly-C9) in the context of the Membrane Attack Complex (MAC) assembly and the broader field of MACPF pore formation. The MAC is an important element of the vertebrate innate immune system that forms pores on the cell membrane of target gram-negative bacteria and some pathogenic parasites [11, 12]. C9 is a major component of the MAC forming 65% of its mass and the major component of transmembrane pore formation [13].

The MAC is formed by the assembly of 7 proteins, 5 of which contain a MACPF/CDC domain [14]. These proteins are C6, C7, C8 $\alpha$ , C8 $\beta$  and C9. All are separate plasma proteins with the exception of the C8 hetero-trimer, composed of C8 $\alpha$ , C8 $\beta$  and C8 $\gamma$ . C5b, the seventh component, results from the proteolytic cleavage of C5 by the C5 convertase and is critical in initiating MAC formation [12].

The MAC follows a unique and complex assembly pathway (Figure 4.1). In the first step of MAC deposition, C5b binds C6 to form C5b6. Subsequent addition of C7 results in the trimeric C5b7 [15]. C8 is then recruited to form C5b8. Then, an estimated 12 to 18 C9 subunits expand C5b8 to further complete the full MAC pore (Figure 4.1) [13, 16].

It has long been observed that C9 can be induced to oligomerise into a ring shape oligomer with dimensions similar to the MAC and in the absence of other MAC components (Figure 4.1) [17, 18]. Interestingly, this oligomer referred to as poly-C9 is a water soluble complex [17-19]. Observations of poly-C9 tubules by transmission

Electron Microscopy (EM) span the last 35 years [13, 18, 20, 21] yet the structural basis for poly-C9 assembly has remained elusive.



**Figure 4.1. Mechanisms of poly-C9 and MAC assembly**

*Top. General Mechanism of MAC assembly.* Following the initial binding of C6 (orange) by C5b (pink) to form C5b6, C7 (yellow) and C8 (green, C8 $\gamma$  is omitted for clarity) are subsequently recruited to form C5b7 and C5b8, respectively. At this stage, C8 $\alpha$  TMH regions are proposed to penetrate the target membrane [22] (grey). An estimated 12 to 18 C9 subunits (blue) are added to the nascent ring. Formation of the final MAC pore is thought to result from the insertion of the TMH regions of all MACPF/CDC domain-containing components into the bilayer membrane.

*Bottom. Poly-C9 assembly.* Formation of poly-C9 follows the circular oligomerisation of an estimated 14 to 18 C9 subunits in an aqueous environment. The TMH regions are thought to form the walls of the poly-C9 ‘tubule’.

Here, the Single Particle cryo-EM reconstruction of poly-C9 was used to develop a structural model of this soluble MAC mimic. This 3D reconstruction constitutes the first sub-nanometer resolution (8 Å) structure of poly-C9 to date and the highest

resolution SP cryo-EM reconstruction in the entire MACPF/CDC superfamily. The unprecedented level of details allows for an unambiguous assignment of C9 individual domains and sub-domains. The work performed as part of this thesis and presented in this chapter includes the molecular modelling used to build the structural model of poly-C9 within the 3D reconstruction. Methods used and the derived structural basis for C9 polymerisation are presented in the revised manuscript format submitted to *Nature Communications*. The supplementary material is reproduced at the end of the chapter. Additional methods details not included in the manuscript are presented in the section *Additional Methods*.

Our structural model reveals a ring of 22 C9 MACPF/CDC domains from which protrudes an unusually long  $\beta$ -barrel forming the tubular walls of the oligomer. This  $\beta$ -barrel adopts the  $S = n/2$  architecture consistently with others MACPF/CDC PFPs. Unexpectedly, the N-terminal ancillary Thrombospondin-1 domain (TSP1) is positioned on the outer edge of the ring and found to contribute significantly to C9 oligomeric interface. The poly-C9 structure further suggests that the TSP1 region contributes to the sequential and uni-directional assembly of poly-C9 and the MAC.

# Structure of the poly-C9 component of the Complement Membrane Attack Complex

Natalya V. Dudkina<sup>1,\*</sup>, Bradley A. Spicer<sup>2,3,\*</sup>, Cyril F. Reboul<sup>2,3,\*</sup>, Paul J. Conroy<sup>2,3</sup>,  
Natalya Lukoyanova<sup>1</sup>, Hans Elmlund<sup>2,3</sup>, Ruby H.P. Law<sup>2,3</sup>, Susan M. Ekkel<sup>2,3</sup>,  
Stephanie C. Kondos<sup>3</sup>, Robert J.A. Goode<sup>3</sup>, Georg Ramm<sup>2,3</sup>, James C. Whisstock<sup>2,3,#</sup>,  
Helen R. Saibil<sup>1,#</sup>, Michelle A. Dunstone<sup>2,3,4,#</sup>

<sup>1</sup>Crystallography, Institute of Structural and Molecular Biology, Birkbeck College,  
London, United Kingdom

<sup>2</sup>ARC Centre of Excellence in Advanced Molecular Imaging, Monash University,  
Clayton Campus, Melbourne VIC, 3800, Australia.

<sup>3</sup>Department of Biochemistry and Molecular Biology, Biomedicine Discovery  
Institute, Monash University, Clayton Campus, Melbourne VIC, 3800, Australia.

<sup>3</sup>Department of Microbiology, Biomedicine Discovery Institute, Monash University,  
Clayton Campus, Melbourne VIC, 3800, Australia.

\*These authors contributed equally to this work

<sup>#</sup>Joint senior and corresponding authors;   


## **Abstract**

The Membrane Attack Complex / Perforin-like (MACPF) protein Complement component 9 (C9) is the major component of the MAC, a multi-protein complex that forms pores in the membrane of target pathogens. In contrast to homologous proteins such as perforin and the cholesterol dependent cytolytins (CDCs), all of which require the membrane for oligomerisation, C9 assembles directly onto the nascent MAC from solution. However, the molecular mechanism of MAC assembly remains to be understood. Here we present the 8 Å cryo-EM structure of a soluble form of the poly-C9 component of the MAC. These data reveal a 22-fold symmetrical arrangement of C9 molecules that yield an 88-strand pore-forming  $\beta$ -barrel. The N-terminal thrombospondin-1 (TSP1) domain forms an unexpectedly extensive part of the oligomerisation interface, thus likely facilitating solution-based assembly. These TSP1 interactions may also explain how additional C9 subunits can be recruited to the growing MAC subsequent to membrane insertion.

## Introduction

Paul Ehrlich originally characterised the haemolytic properties of human blood over 100 years ago <sup>1</sup>. Subsequent work revealed that the terminal “Membrane Attack Complex” (MAC) portion of Complement represents the lytic, pore forming part of the system <sup>2,3</sup>. This structure is responsible for eliminating Gram-negative bacteria and other pathogens.

The MAC comprises seven components: C5b, C6, C7, C8 (a heterotrimer composed of C8 $\alpha$ , C8 $\beta$  and C8 $\gamma$ ) and multiple copies of C9 (Supplementary Fig. 1). *In vitro* studies reveal that multiple C9 subunits are recruited to the C5b678 complex, whereupon it self-assembles to form large, ring-shaped pores with a lumen over 100 Å in diameter embedded in the membrane of target cells <sup>4</sup>. C9 can also be induced to form poly-C9 pore-like structures in solution that closely resemble the MAC pore <sup>5</sup>. C6, C7, C8 $\alpha$ , C8 $\beta$  and C9 all belong to the MACPF/CDC superfamily <sup>6,7</sup> and include a common set of 4 core domains; a TSP1 domain followed by a Low Density Lipoprotein Receptor Associated (LDLRA) domain, a MACPF domain and an Epidermal Growth Factor (EGF) domain (Supplementary Fig. 1).

Much of our understanding of the MACPF/CDC superfamily comes from studying CDCs <sup>8–10</sup>. Briefly, soluble CDC monomers bind to and then oligomerise on the membrane surface to form a prepore intermediate <sup>10,11</sup>. Next the assembly undergoes a concerted conformational change that involves significant opening and untwisting of a central, four-stranded  $\beta$ -sheet. This event permits two helical regions (termed transmembrane hairpins TMH-1 and TMH-2) to unravel and insert into the membrane as amphipathic  $\beta$ -hairpins (Supplementary Fig. 2).

Studies on the MAC have revealed mechanistic distinctions from other family members. For example, perforin, pleurotolysin and CDCs bind to membrane lipids or

membrane associated proteins via ancillary domains prior to oligomerisation<sup>9,12,13</sup>. In contrast, C9 does not contain any obvious membrane-binding domain. Thus even when the nascent MAC (C5b678) is associated with the target cell, the assembly process must include the recruitment of C9 from solution (i.e. from plasma, Supplementary Fig. 2B). Consistent with this, a soluble form of the MAC can also assemble independently of the membrane and be detected in blood plasma (Supplementary Fig. 2B).

To understand the mechanism of MAC assembly, we determined the sub-nanometer resolution single particle EM structure of C9 in a polymerized pore-like form. These data reveal the unexpected finding that the TSP1 domain forms a significant portion of the interface between interacting C9 monomers. This finding may explain why the MAC, in contrast to related molecules such as perforin and the CDCs, is able to assemble from monomers directly recruited from the soluble phase. The additional interactions mediated by the TSP1 domain may also explain previous observations<sup>14</sup> that C9 monomers are recruited to a MAC that has already entered the target cell membrane.

## **Results**

### ***The structure of polyC9***

To understand MAC assembly we determined the 8 Å single particle cryo-EM reconstruction of soluble poly-C9 from 5,000 particles (Fig. 1A-D, Supplementary Fig. 3-6). These data revealed a symmetrical assembly of 22 C9 monomers (Fig. 1A-C) that closely resembles the MAC<sup>4</sup>. The structure comprises a ring-shaped assembly of globular domains atop a large β-barrel (Fig. 1A, B). The latter part of the structure



is flexible and is less well resolved than the top half of the structure. However, the diameter of the  $\beta$ -barrel (120 Å) is consistent with the predicted 88-stranded structure and is of sufficient size to permit passage of proteins such as lysozyme<sup>15</sup>. We further observed density consistent with two N-glycosylation sites, one on each TMH sequence (Supplementary Fig. 1 and 7). We observe a bulbous feature at the base of the  $\beta$ -barrel and suggest that this may be a consequence of structural rearrangements to protect the hydrophobic surface that ordinarily contacts the membrane (Fig. 1B). Higher resolution data will be required to validate this suggestion.

In the top, better-resolved portion of the map the position of each of the four domains in C9 can be unambiguously assigned. Although no crystal structure of C9 is available, we were able to interpret the poly-C9 structure using the core TSP1-LDLRA-MACPF-EGF assembly from the crystal structure of C6 (Fig. 1E, F; Supplementary Fig. 1)<sup>16,17</sup>. Indeed, only minor changes in domain orientation are required to dock the C6 structure into the bulk of the poly-C9 density (Fig 1E, F).

### ***The C9 TSP1 domain makes a major contribution to the oligomer interface***

Structural studies on other MACPF/CDC proteins reveal that most interactions within the pre-pore or pore assembly appear to be formed between the relatively flat faces of the MACPF domain<sup>8,11,13</sup>. In contrast the poly-C9 structure reveals that the TSP1 domain packs against the C-terminal  $\alpha$ -helix of the MACPF domain of an adjacent monomer and forms an additional and significant portion of the oligomer interface (Fig. 2). Thus in the pore form, each TSP1 domain is wedged between two C-terminal  $\alpha$ -helices – one contributed *in trans* from an adjacent monomer and one *in cis*. This interaction at the outer edge of the ring-like assembly forms a quarter (~690 Å<sup>2</sup>) of the total (~3000 Å<sup>2</sup>) surface buried in the globular, non-barrel region (Fig. 2).

The remainder of the interacting surface is contributed by interactions between MACPF domains.

In the MAC it is anticipated that the MACPF domain of the related complement components C6, C7 and C8 form part of the overall circular assembly<sup>3</sup>. Like C9, C6-C8 all contain an analogous TSP1 domain that is functionally important (Supplementary Fig. 1)<sup>14</sup>. It is therefore suggested that the TSP1 domain of each protein in the complete MAC will be positioned at the subunit interface. Indeed, we suggest that the specialized TSP1 / MACPF interactions likely explain the unusual ability of the nascent MAC to recruit components directly from solution. In contrast, proteins such as perforin, pleurotolysin and CDCs lack a TSP1 equivalent and do not readily self-assemble in solution. Instead, they require membrane anchoring via ancillary domains in order to oligomerise. Indeed, it is known from the study of receptors that restriction to the membrane plane can favour oligomerisation through weak protein-protein interactions<sup>18</sup>.

### ***Conformational transitions during pore formation***

We next examined the conformational changes that take place in the transition from the soluble monomer to the pore form. Comparison with C6 suggests that the largest conformational rearrangements during the transition from the monomer to the pore form take place within the MACPF domain<sup>19,20</sup>. The bottom half of the central  $\beta$ -sheet is rotated by approximately 10° relative to its position in C6. This movement shifts the lower part of the  $\beta$ -sheet laterally by ~5.5 Å (Fig. 3A, B). Concomitantly with this change, TMH1 and TMH2 unravel to form the  $\beta$ -barrel (Fig. 1B).

The lateral movement in the central sheet of the MACPF domain repositions the conserved helix-turn-helix (HTH) region that sits on top of TMH2 in the soluble

monomeric form. Consistent with this, the top of the poly-C9 pore lumen is lined by pairs of  $\alpha$ -helices (Fig. 3C). Previous mutagenesis and structural studies on the fungal MACPF protein pleurotolysin as well as the CDC suilysin suggest a role of the HTH region in pre-pore assembly and in controlling the transition to the pore<sup>13,21</sup>.

## Discussion

The structure of polyC9 provides mechanistic insight into how components of the MAC may assemble through additional interactions mediated via the TSP1 domain. Furthermore, the structure provides insights into self-association by MACPF domain containing proteins more generally. In particular, our present poly-C9 structure may resolve the controversy regarding the orientation of perforin in the pore assembly. Our previous analysis of the low resolution EM structure of the perforin pore suggested that perforin monomers are orientated in the pore assembly opposite to the CDCs and pleurotolysin<sup>8,11</sup>. The latter two proteins, however, share very limited (<10%) sequence identity in the MACPF domain with perforin, whereas C9 is more closely related (~25% identity). Accordingly, we superposed the perforin structure onto the poly-C9 model. This suggests that perforin most likely oligomerises similarly to C9, following minor rearrangements of the TMH2 and HTH domains (Supplementary Fig. 8). We note that residues shown through mutagenesis studies to interact at the pore interface are brought into close proximity with one another<sup>22</sup>. Further, the absence of the TSP1 domain in perforin at the outer edge of the pore assembly may explain the heterogeneity in perforin pore size and shape. We thus conclude that the present 8-Å-resolution poly-C9 map thus provides a better model for the perforin assembly.

Finally, the new structural insights may help explain how the MAC assembles with respect to target cell membranes. In the current view, C7 and C8 sequentially insert into the membrane, anchoring it in place prior to the recruitment of multiple copies of C9. However, this mechanism contrasts with the current view of the MACPF/CDC pore formation, in which the amphipathic hairpins are proposed to be inserted in a concerted fashion in the context of a complete or incomplete ring<sup>11,23</sup>. The latter mechanism seems more plausible because the conformational change in the MACPF domain during membrane insertion is extensive and would be predicted to disfavor the addition of new subunits. The poly-C9 structure provides new insights into this problem. The additional TSP1 / MACPF interactions involve regions of the molecule that do not undergo significant conformational change. We therefore hypothesise that the TSP1-mediated interactions may permit addition of C8 and C9 to a nascent MAC that has already entered the target membrane.

To conclude, we have determined the structure of poly-C9 at a resolution sufficient to confidently position individual domains and to resolve helical features in density. Our data further reveal an unexpected contribution of domains ancillary to the MACPF domain that likely function to stabilize the overall assembly and the top half of the  $\beta$ -barrel pore.

## **Methods**

### ***Cryo-electron microscopy sample preparation and data acquisition***

Complement C9 was purified from human apheresis plasma as described in Supplementary Methods. This project was deemed by the Monash University Human Research Ethics Committee (project CF14/3761 – 2014001968) to be exempt from ethical review.

Monomeric C9 was polymerized by overnight incubation at 1 mg mL<sup>-1</sup> and 37°C. The resulting polyC9 was applied to lacey carbon coated copper grids (Agar, UK) and frozen with a FEI Vitrobot™ Mark III (FEI, Eindhoven) at 22°C and 100% humidity. Images were recorded manually on a Tecnai G2 Polara microscope (FEI) operating at 300 kV with a Quantum energy filter and K2 Summit detector (Gatan, UK) in counting mode, at a pixel size of 2.76 Å. Exposures were recorded at 1.2 electrons (Å<sup>2</sup>)<sup>-1</sup> s<sup>-1</sup> for 25 s, with defocus values ranging from 1.2-4.9 µm (Supplementary Fig. 3).

### ***3D reconstruction of polyC9***

The detector movies were aligned using IMOD<sup>24</sup>. CTF parameters were determined with CTFFIND4<sup>25</sup>. 10800 particles were extracted manually using Boxer (EMAN 1.9)<sup>26</sup>. Classification and refinement were performed using RELION<sup>27</sup>. 2D classification in IMAGIC (ref 29) revealed mainly end views with 22-fold symmetry, with a small fraction of particles having 21- or 23-fold symmetry (Supplementary Fig. 3). The initial model with 22-fold symmetry was created by angular reconstitution from 2D class averages of particles with all orientations in IMAGIC<sup>28</sup> and refined by projection matching using SPIDER<sup>29</sup>. A subset of ~5000 particles in side and tilted views (homogeneous with respect to diameter of the wide part of the ring, corresponding to the 22-mers) was refined with RELION using the initial model from SPIDER filtered to 20 Å. 22-fold symmetry was applied during refinement. The final map was corrected to the modulation transfer function of the detector and sharpened by applying a B-factor<sup>30</sup> determined by RELION. The final resolution calculation based on gold-standard FSC was estimated at 0.5 and 0.143 FSC in RELION. Local resolution was estimated using the ResMap program<sup>31</sup> (Supplementary Fig. 4).

### ***Determination of handedness***

In order to determine the absolute hand of the 3D reconstruction, the crystallographic structure of C6 was fitted into the map as well as into the map with opposite handedness. Although the fit of the C6 conformation was found to favour one hand over the other, the differences in cross-correlations were too small to conclusively assign the hand of the map (C6: 0.63 vs 0.59; calculated using Chimera software <sup>32</sup>).

To resolve this issue, we examined the fit to both maps (A & B) of conserved structural features of the MACPF domain. The C-terminal  $\alpha$ -helical bundle of the MACPF domain (Supplementary Fig. 5) is likely to be clearly discernable in an 8-Å-resolution cryo-EM density. Its characteristic arrangement of  $\alpha$ -helices is asymmetric and highly conserved in all the crystallographic structures of MAC components <sup>17</sup>. We therefore expected that it should be possible to identify the correct hand from analysis of the fit of this structural motif in the enantiomeric maps (Supplementary Fig. 5 B,C).

Accordingly, we found the map in Supplementary Fig. 5B showed distinct density corresponding to the C-terminal  $\alpha$ -helical bundle. The region of the map identified by rigid body fitting excellently reproduces the topology and length of the  $\alpha$ -helices. Conversely, the map in Supplementary Fig. 5C produces a comparatively poor agreement with the fitted position of C6 (Supplementary Fig. 5C). We concluded that the map in Supplementary Fig. 5B represents the correct hand.

### ***Fitting of atomic models***

A homology model of C9 was fitted into the EM map by employing a combination of manual, rigid body and flexible fitting. The C9 homology model was

generated using the crystallographic structures of C6 (PDB IDs: 3T5O, 4A5W) and C8 (2RD7, 3OJYA, 3OJYB) and Modeller 9.14<sup>33</sup>. The TMH1/2 regions were discarded because these regions form a  $\beta$ -barrel in polyC9.

Five symmetry-related monomers were then subjected to flexible fitting (MDFF methodology as implemented in NAMD 2.10<sup>34</sup>) using symmetry restraints<sup>35</sup>. The protein secondary structure was restrained to avoid overfitting. Oligomeric main chain hydrogen bonds between the  $\beta$ -sheets forming the top of the  $\beta$ -barrel were also restrained to reproduce the pattern conserved in the MACPF/CDC superfamily<sup>13,36</sup>. Two independent 5-ns simulations were performed in vacuo at 310K ( $\gamma=0.3$ ; 1 fs time step; 12 Å cutoff for long range interaction) using the CHARMM36 force field<sup>37</sup> and followed by 5000 steps of energy minimisation ( $\gamma=0.5$ ). The resulting model with the highest CC (0.93; Molprobit score of 1.15) was replicated with C22 symmetry and combined with a structural model of the 88-stranded  $\beta$ -barrel (architecture S = n/2<sup>38</sup>) using Modeller, thus extending the  $\beta$ -strands of the central MACPF  $\beta$ -sheet as performed in<sup>11,13,39</sup>. The final polyC9 22-mer model (CC of 0.94) is shown (Fig. 1).

In order to assess the reliability of the fitting procedure, the flexible fitting step was repeated using cryo-EM maps calculated from randomly partitioned half-sets, independently refined using RELION and used to determine the resolution of the final cryo-EM map (see EM methods). Individual residue RMSDs of both fitted models were calculated with respect to the structural model obtained from the whole dataset (Supplementary Fig. 6). Structural elements displaying an overall high RMSD (e.g. not fitted in a reproducible manner; >3.5 Å) were not included in the final structural model.

## References

1. Kaufmann, S. H. E. Immunology's foundation: the 100-year anniversary of the Nobel Prize to Paul Ehrlich and Elie Metchnikoff. *Nat. Immunol.* **9**, 705–12 (2008).
2. Walport, M. J. Advances in immunology: Complement (first of two parts). *N. Engl. J. Med.* **344**, 1058–1066 (2001).
3. Podack, E. R. Molecular mechanisms of cytolysis by complement and by cytolytic lymphocytes. *J. Cell. Biochem.* **30**, 133–70 (1986).
4. Tschopp, J., Podack, E. R. & Müller-Eberhard, H. J. Ultrastructure of the membrane attack complex of complement: detection of the tetramolecular C9-polymerizing complex C5b-8. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 7474–8 (1982).
5. Tschopp, J., Muller-Eberhard, H. J. & Podack, E. R. Formation of transmembrane tubules by spontaneous polymerization of the hydrophilic complement protein C9. *Nature* **298**, 534–538 (1982).
6. Rosado, C. J. *et al.* A common fold mediates vertebrate defense and bacterial attack. *Science* **317**, 1548–51 (2007).
7. Hadders, M. A., Beringer, D. X. & Gros, P. Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* **317**, 1552–4 (2007).
8. Tilley, S. J., Orlova, E. V., Gilbert, R. J. C., Andrew, P. W. & Saibil, H. R. Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* **121**, 247–256 (2005).
9. Shepard, L. A. *et al.* Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: An  $\alpha$ -helical to  $\beta$ -sheet transition identified by fluorescence spectroscopy. *Biochemistry* **37**, 14563–14574 (1998).
10. Czajkowsky, D. M., Hotze, E. M., Shao, Z. & Tweten, R. K. Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. *EMBO J.* **23**, 3206–3215 (2004).
11. Leung, C. *et al.* Stepwise visualization of membrane pore formation by suilysin, a bacterial cholesterol-dependent cytolysin. *Elife* **3**, e04247 (2014).
12. Law, R. H. P. *et al.* The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* **468**, 447–451 (2010).
13. Lukoyanova, N. *et al.* Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin. *PLoS Biol.* **13**, e1002049 (2015).
14. Scibek, J. J., Plumb, M. E. & Sodetz, J. M. Binding of Human Complement C8 to C9 : Role of the N-Terminal Modules in the. *Biochemistry* 14546–14551 (2002). doi:10.1021/bi026641j
15. Martinez, R. J. & Carroll, S. F. Sequential metabolic expressions of the lethal process in human serum-treated *Escherichia coli*: role of lysozyme. *Infect. Immun.* **28**, 735–45 (1980).
16. Hadders, M. A. *et al.* Assembly and regulation of the membrane attack complex based on structures of C5b6 and sC5b9. *Cell Rep.* **1**, 200–7 (2012).
17. Aleshin, A. E. *et al.* Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC). *J. Biol. Chem.* **287**, 10210–22 (2012).



18. Wu, Y., Vendome, J., Shapiro, L., Ben-Shaul, A. & Honig, B. Transforming binding affinities from three dimensions to two with application to cadherin clustering. *Nature* **475**, 510–3 (2011).
19. DiScipio, R. G. The size, shape and stability of complement component C9. *Mol. Immunol.* **30**, 1097–106 (1993).
20. DiScipio, R. G. & Berlin, C. The architectural transition of human complement component C9 to poly(C9). *Mol. Immunol.* **36**, 575–85 (1999).
21. Ramachandran, R., Tweten, R. K. & Johnson, A. E. Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment. *Nat. Struct. Mol. Biol.* **11**, 697–705 (2004).
22. Baran, K. *et al.* The molecular basis for perforin oligomerization and transmembrane pore assembly. *Immunity* **30**, 684–95 (2009).
23. Sonnen, A. F.-P., Plitzko, J. M. & Gilbert, R. J. C. Incomplete pneumolysin oligomers form membrane pores. *Open Biol.* **4**, 140044 (2014).
24. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–6 (1996).
25. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* (2015). doi:10.1016/j.jsb.2015.08.008
26. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
27. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–30 (2012).
28. van Heel, M., Harauz, G., Orlova, E. V, Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
29. Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–9 (1996).
30. Rosenthal, P. B. & Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
31. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–5 (2014).
32. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).
33. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **Chapter 2**, Unit 2.9 (2007).
34. Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–83 (2008).
35. Chan, K.-Y. *et al.* Symmetry-restrained flexible fitting for symmetric EM maps. *Structure* **19**, 1211–8 (2011).
36. Sato, T. K., Tweten, R. K. & Johnson, A. E. Disulfide-bond scanning reveals assembly state and  $\beta$ -strand tilt angle of the PFO  $\beta$ -barrel. *Nat. Chem. Biol.* **9**, 383–9 (2013).
37. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–45 (2013).

38. Reboul, C. F., Mahmood, K., Whisstock, J. C. & Dunstone, M. A. Predicting giant transmembrane  $\beta$ -barrel architecture. *Bioinformatics* **28**, 1299–302 (2012).
39. Reboul, C. F., Whisstock, J. C. & Dunstone, M. A. A new model for pore formation by cholesterol-dependent cytolysins. *PLoS Comput. Biol.* **10**, e1003791 (2014).

## Acknowledgements

HRS acknowledges support from the ERC (294408), Wellcome Trust Equipment grants 101488, 079605 and 086018 and a BBSRC equipment grant (BB/L014211). We thank D. Houldershaw and R. Westlake for computing support, D. Clare and L. Wang for EM support. We thank the Monash Platforms (Protein Production Unit, MASSIVE, Proteomics, Protein Crystallography and the Clive and Vera Ramaciotti Centre for Structural Cryo-Electron Microscopy) for technical support. MAD acknowledges support from the NHMRC of Australia (Career Development fellowship). JCW acknowledges support from the NHMRC (Senior Principal Research Fellowship) and the Australian Research Council (Federation Fellowship). CFR acknowledges a Monash University Faculty of Medicine Bridging Postdoctoral Fellowship. BAS is supported by an MBO PhD Scholarship. We thank the Australian Red Cross Blood Service for provision of apheresis blood product.

## Author Contributions

MAD, HRS and JCW conceived, co-led the work and co-wrote the paper. NVD collected data, determined the structure and cowrote the paper. BAS, PJC, RHPL, SME, SCK, GR and RJAG produced and analysed protein and optimized sample for EM experiments. CFR and HE performed computational analysis.

### **Author Information**

Correspondence and requests for materials should be addressed to MAD



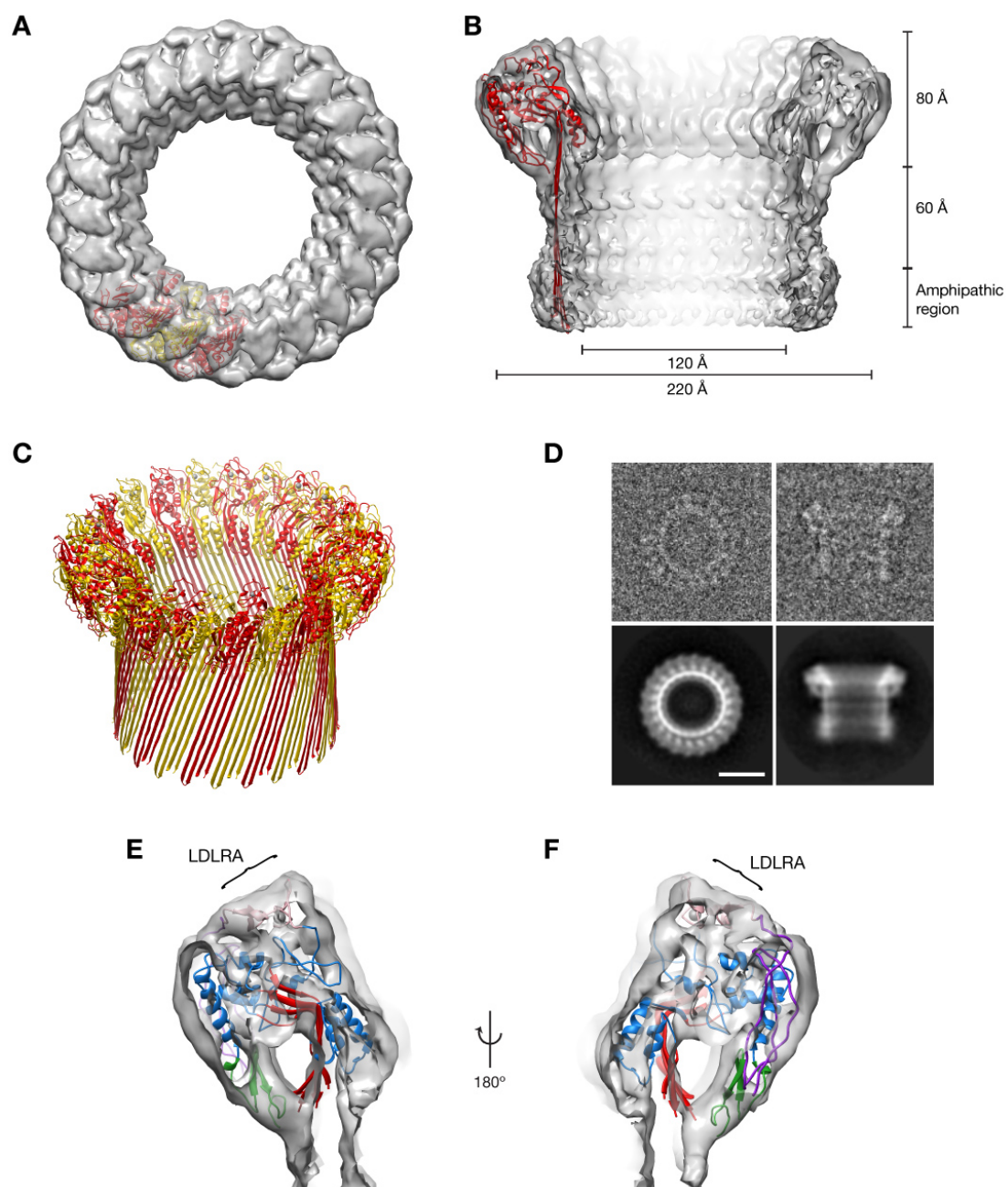
### **Accession codes**

The cryo-EM map of polyC9 is available in the Electron Microscopy Data Bank (accession number EMD-3235). The structure coordinate file for the fitted pore model is available in the Protein Data Bank database (accession number XXXX).

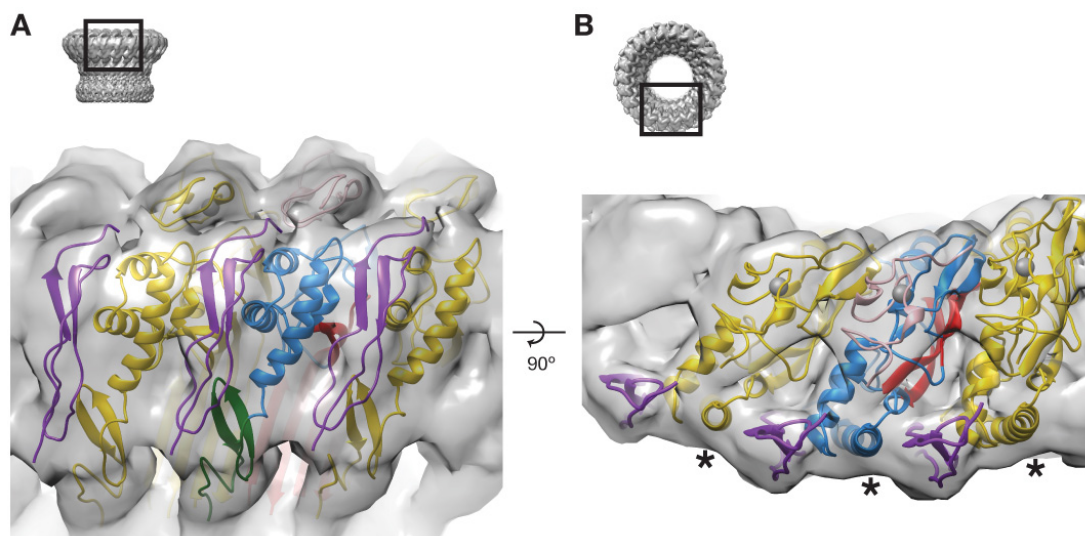
### **Competing Financial Interests**

The authors declare no competing financial interests.

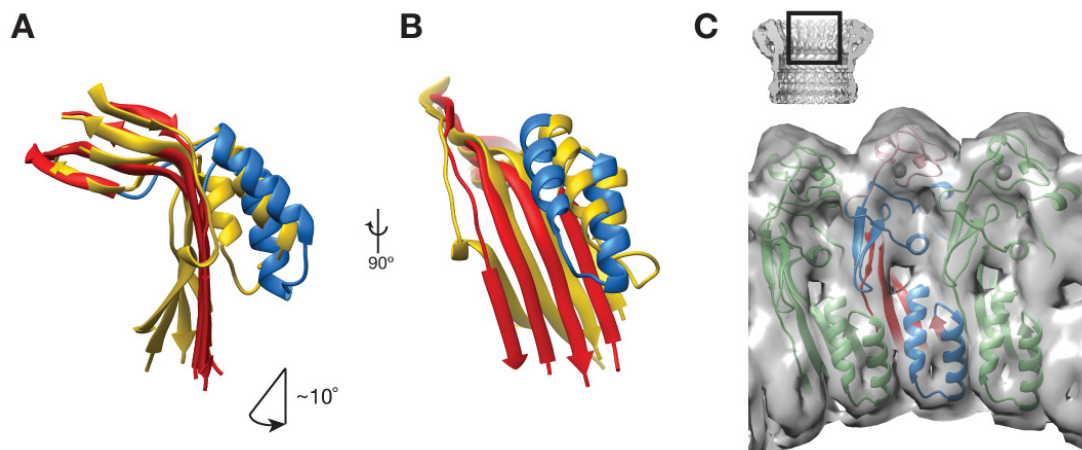
## Figures and Figure Legends



**Fig. 1 The structure of poly-C9.** (A) Top down-view of a C9 trimer in the poly-C9 map and (B) Cut through of the poly-C9 map with cartoon (red) of the poly-C9 model. Approximate dimensions and the predicted amphipathic region are indicated. (C) Cartoon of the full poly-C9 pore (alternating red and yellow monomers). The barrel is best modeled with the architecture  $S = n/2$  <sup>38</sup>. (D) Cryo EM end and side views of poly C9 in individual images (top) and class averages (bottom). (E) and (F) With the exception of the mobile region of the MACPF domain (which in poly-C9 has rearranged in order to form the barrel), the crystal structure of C6 fits well into the map (PDB: 3T5O: TMH1 and TMH2 are omitted for clarity). In this figure the conserved  $\beta$ -sheet of the MACPF domain is in red, the body of the MACPF domain is in blue, the EGF domain in green, the TSP1 domain in purple and the LDLRA domain in pink (labeled).



**Fig. 2 Interactions made by the TSP1 domain:** **(A)** A view of the outside of the globular portion of the poly-C9 map showing the TSP1 domain (purple) located at each subunit interface. The central C9 monomer is colored as in figure 1, with the monomers each side in dark yellow and purple (TSP1 domain). **(B)** A view from the top showing placement of the TSP1 domain between the C-terminal helix (marked with a \*) of each MACPF domain.



**Fig. 3 Comparison between the structure of C6 (PDB ID 3t5o; yellow) and model of poly C9 (red / blue):** The shift of the central bent  $\beta$ -sheet (red) shows **(A)** a  $\sim 10^\circ$  rotation of the bottom half of the sheet together with **(B)** a  $\sim 5.5$  Å lateral movement. **(C)** The HTH region (a pair of  $\alpha$ -helices) lines the pore lumen. A trimer is shown with the central monomer colored red, blue and pink.

**Extended material:**

Structure of the poly-C9 component of the Complement Membrane  
Attack Complex

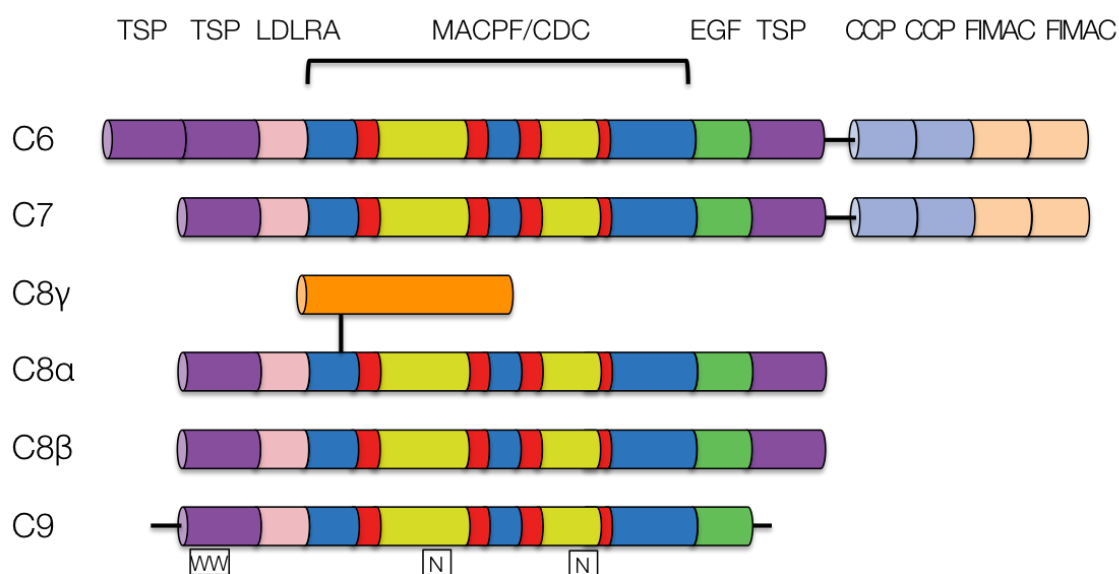
Natalya V. Dudkina<sup>1,\*</sup>, Bradley A. Spicer<sup>2,3,\*</sup>, Cyril F. Reboul<sup>2,3,\*</sup>, Paul Conroy<sup>2,3</sup>,  
Natalya Lukoyanova<sup>1</sup>, Hans Elmlund<sup>2,3</sup>, Ruby H.P. Law<sup>2,3</sup>, Susan Ekkel<sup>2,3</sup>, Stephanie  
C. Kondos<sup>3</sup>, Robert J.A. Goode<sup>3</sup>, Georg Ramm<sup>2,3</sup>, James C. Whisstock<sup>2,3,#</sup>, Helen R.  
Saibil<sup>1,#</sup>, Michelle A. Dunstone<sup>2,3,4,#</sup>

correspondence to:

[REDACTED]

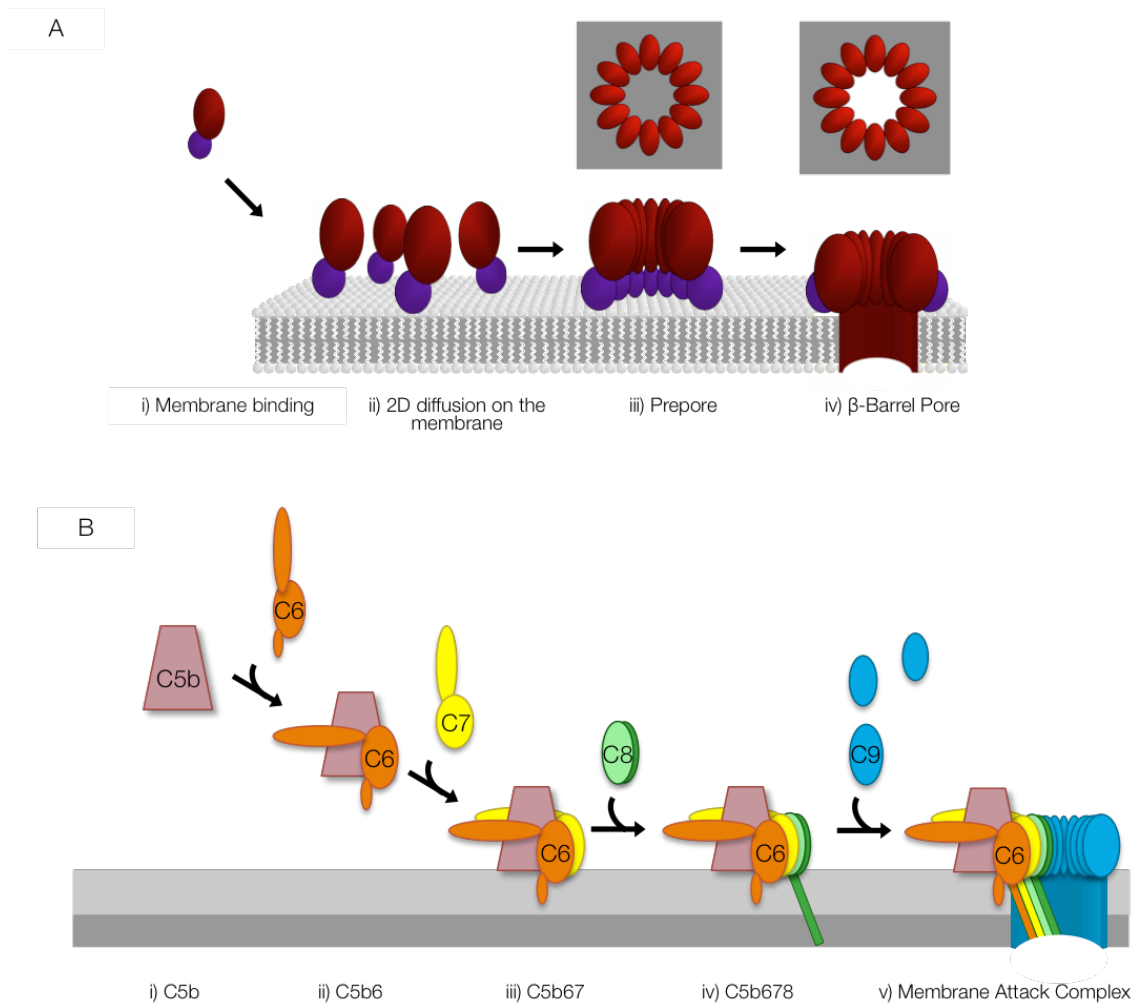
[REDACTED]





**Fig. S1. Schematic of domain composition of the Membrane Attack Complex proteins**

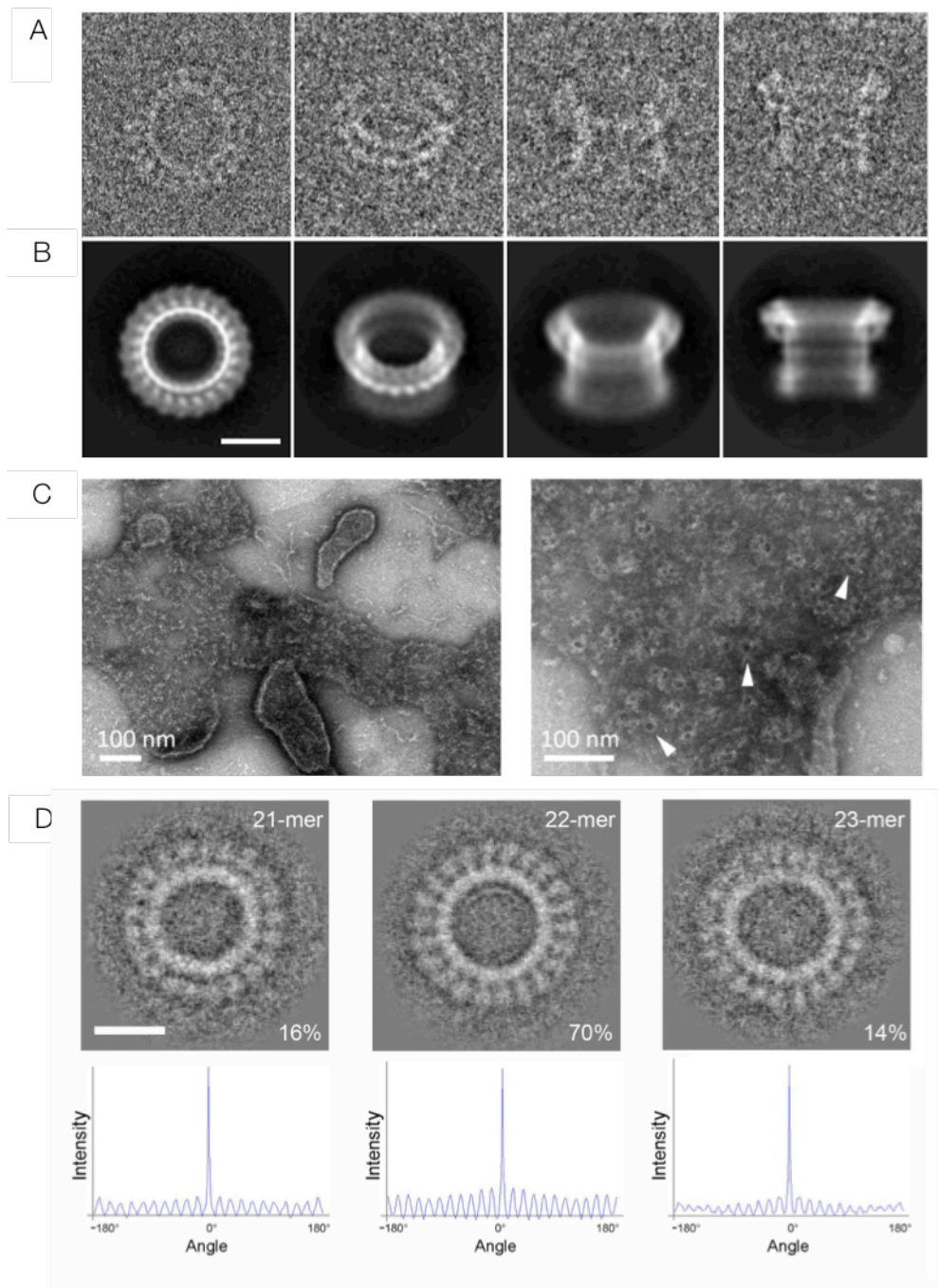
Schematic of the domain composition of C9 in comparison with C6, C7, C8α, C8β. The schema of C9 includes the sites of C-mannosylation (WW) and N-glycosylation (N) of C9 as determined according to methods and in agreement with published data <sup>1,2</sup>. The two predicted TMH regions are also labelled. Colours are the same as used throughout the domain coloured figures. TSP = Thrombospondin Type 1 domain, LDLRA = Low-Density Lipoprotein Receptor Type A, MACPF = Membrane Attack Complex/Perforin / Cholesterol Dependent Cytolysin, EGF = Epidermal Growth Factor-like, CCP = Complement Control Protein, FIMAC = Factor I / Membrane Attack Complex domain. Colours are the same as used in Figs 1, 2 and 3.



**Fig. S2 The general mechanism of MACPF/CDC pore forming proteins (adapted from <sup>3</sup>).**

**A)** The well-defined CDC pore forming mechanism. i) Membrane recognition and binding, ii) two dimensional diffusion of the monomers on the membrane and iii) oligomerisation into the prepore state. iv) β-barrel pore formation, postulated to occur as a concerted, simultaneous insertion of the two TMH regions. Insets at the top show the state of the membrane in the prepore and pore states.

**B)** The general mechanism of the Membrane Attack Complex (MAC) formation. After formation of C5b (pink, panel (i)) there is sequential binding of C6 (orange, panel (ii)), C7 (yellow, panel (iii)) and the C8 heterotrimer (green, C8γ not shown, panel (iv)). It is postulated that the C8α component (dark green) of C8 inserts its TMH2 region into the membrane at this stage. In panel (v) the final pore formation is depicted with the sequential addition of ~18 C9 molecules (blue; initial interface proposed to be with C8).

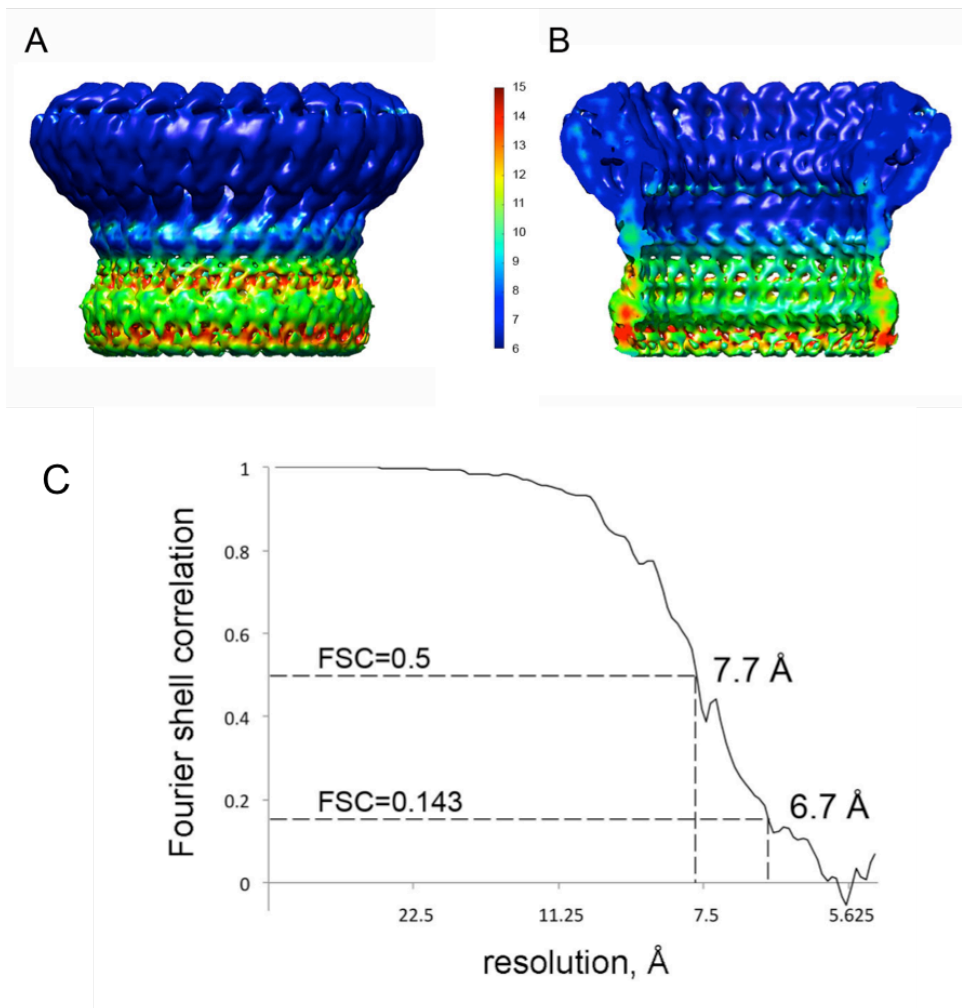


**Fig. S3 Electron microscopy of polyC9 and MAC.**

Representative cryo-EM (A) views of four raw images of polyC9 with the (B) corresponding averaged views. Scale bar, 10 nm.

(C) Rabbit red blood cell ghosts incubated with C9 depleted serum without the addition of purified C9 (left) and with the addition of purified C9 (right). MAC pores are indicated by white arrows.

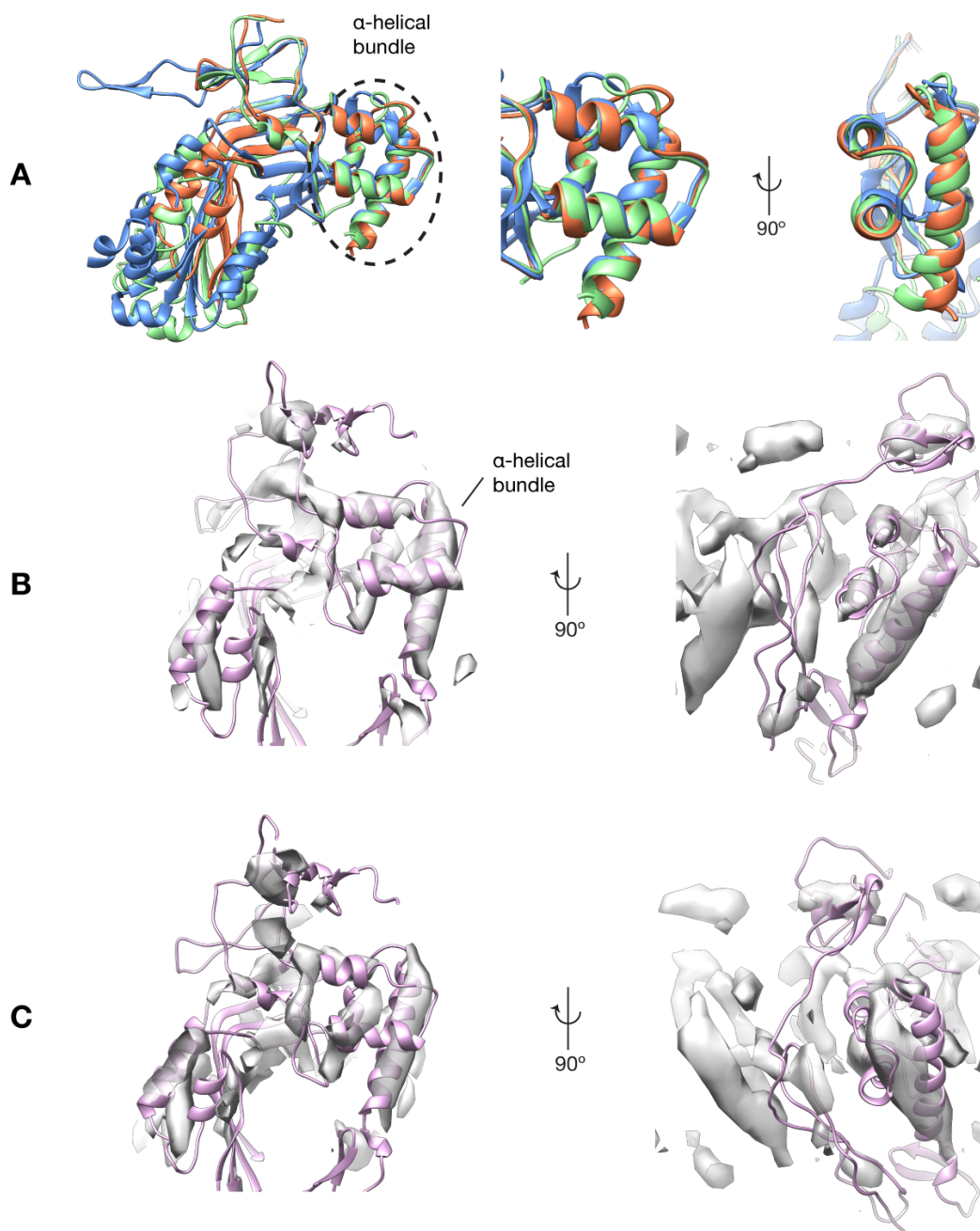
(D) Symmetry of polyC9 pores. Representative averaged views of 21, 22 and 23-fold symmetric pores (upper row) with proportion of particles with each symmetry detected in the data set and corresponding rotational autocorrelations (lower row). The bar is 10 nm.



**Fig. S4 Map quality of poly-C9.**

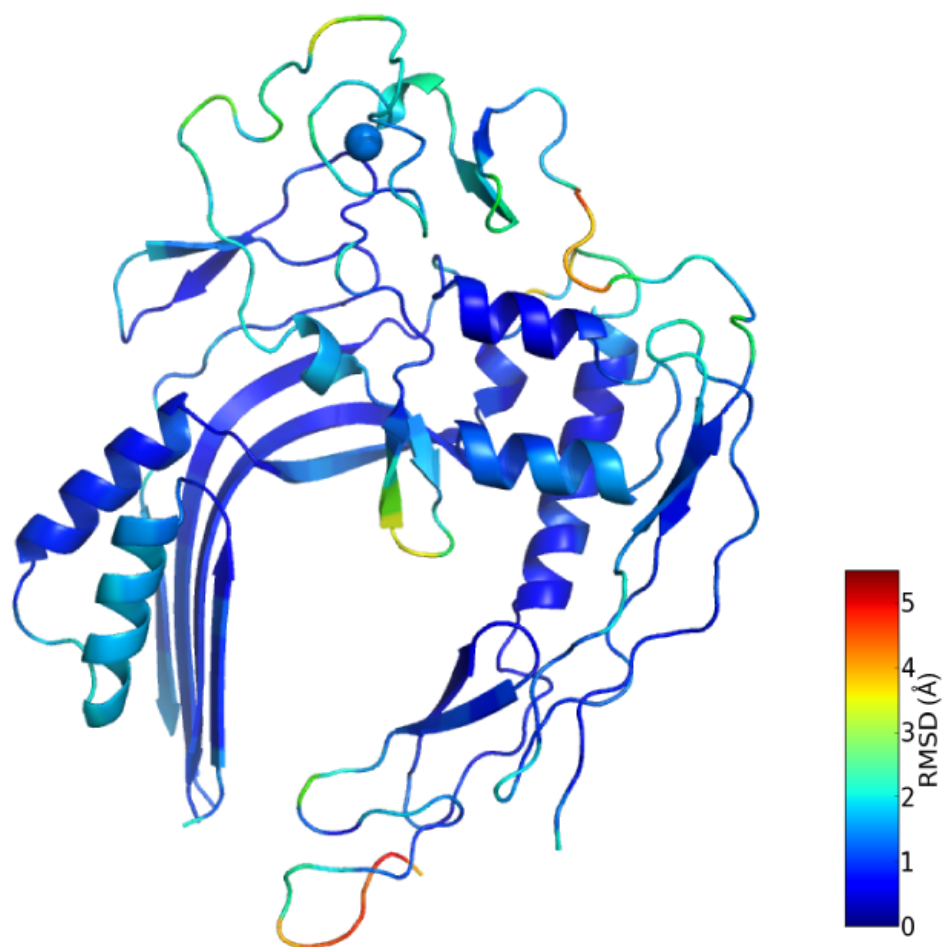
(A) A surface view of the sharpened final map colored according to local resolution.

(B) As in A, but a cut-through view. (C) Fourier-shell correlation curve.



**Fig. S5 Structural alignment of the MACPF domain of MAC components and determination of handedness.**

(A) The alignment identifies two crossed pairs of  $\alpha$ -helices (far left, circled) as a highly structurally conserved region of the MACPF domain. The topology and lengths of the individual  $\alpha$ -helices are identical in MAC components (middle and right panels). Superpositions of the crystallographic structures of C6 (PDB ID 3T5O, orange), C8 $\alpha$  (PDB ID: 2RD7, green) and C8 $\beta$  (PDB ID: 3OJY chain B, blue). (B) The hand of map chosen in this study and (C) the mirrored map. The crystallographic conformation of C6 (PDB ID: 3T5O, pink) is in cartoon representation.



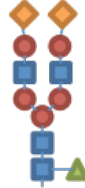


**Fig. S6 Reproducibility of the flexible fitting step.**





Regions displaying a high RMSD are limited to the C-terminus and the loop at positions 74-79.



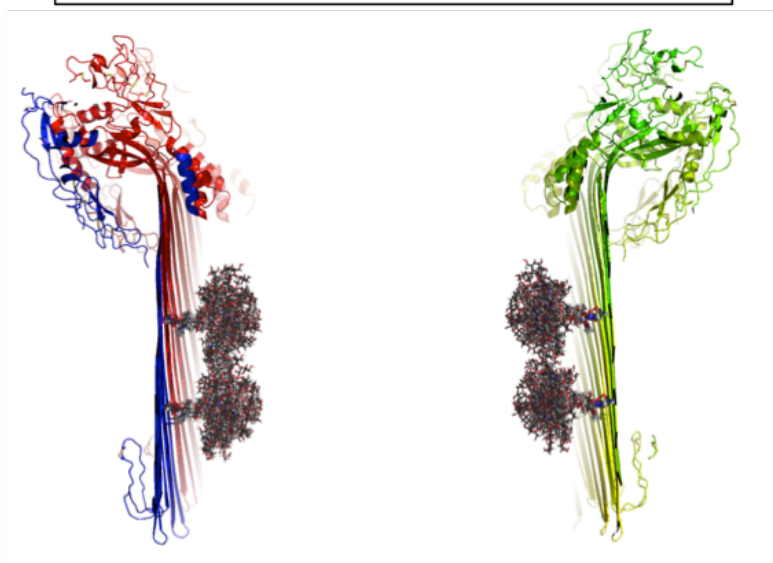
**A**

Structure		Unmodified			
Nominal mass	z	0	2204	1913	2350
N277 Peak Area	3+	N.D.	1.074E9 ~86%	1.786E8 ~14%	N.D.
N415 Peak Area	3+	N.D.	5.894E8 ~86.5%	89.149E7 ~11.8%	1.081E7 ~1.6%
	4+	N.D.	2.179E8 ~86.5%	81.887E7 ~11.8%	N.D.

Legend			
Type	$m_{(nominal)}$	e.g.	
	NeuAc 291	Sialic Acid	
	Hex 162	Mannose, Galactose, Glucose	
	HexNAc 203	GlcNAc, GalNAc	
	dHex 146	Fucose	

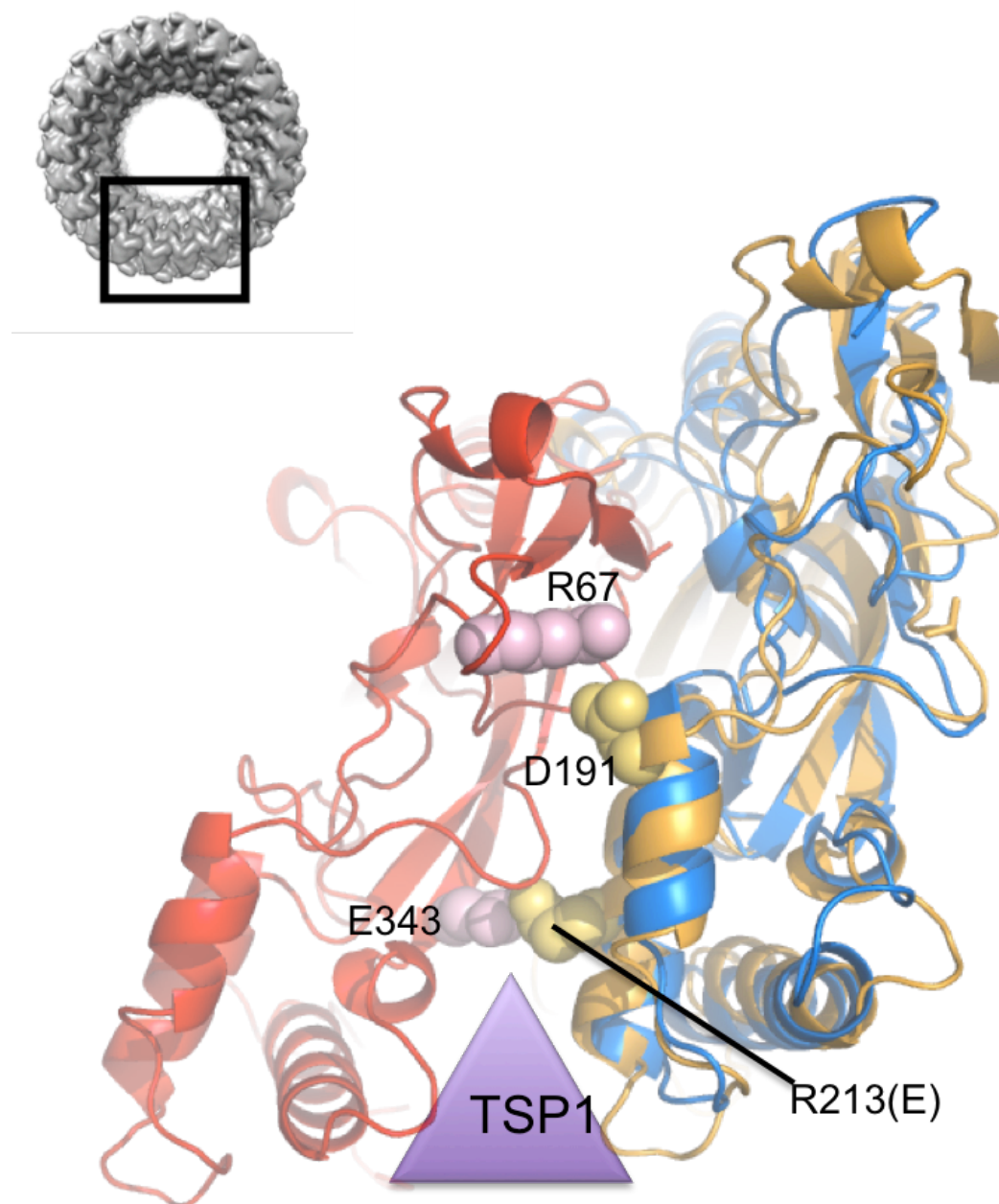
**B**



**Fig. S7 The N-glycosylation of C9**

**(A)** MS-MS characterisation of the two N-glycan sites. Most of the N-glycans have two sialic acid groups each.

**(B)** Superposition of the glycan models from the NMR structure of human chorionic gonadotropin (PDB ID: 1HD4). These N-glycans lack the sialic groups but show the potential degrees of freedom of each of the 44 glycan groups located in the pore lumen.



**Fig. S8 Superposition of perforin on the model of polyC9.**

Illustration of the arrangement of two perforin molecules (red and orange) generated by superposition of the perforin structure (PDB ID: 3NSJ) on the model of poly-C9 (blue). In the superposition R213 (mutated to a Glu in 3NSJ) is in close proximity to E343 as previously predicted <sup>4</sup>. These data further suggest that D191 which has been shown to be important for pore formation forms a salt bridge with R67.



## Supplementary Methods

### ***Protein purification***

Out-of-date apheresis human plasma was supplied by the Australian Red Cross and stored at -80°C until required. Plasma C9 was purified using protocols adapted from established protocols<sup>5,6</sup>. Briefly, apheresis plasma treated with 0.1 mM PMSF, 0.1 mM benzamidine, 0.5 mM EDTA and one protease inhibitor cocktail per 100 mL plasma (Roche) was diluted with 0.4 volumes of ultrapure water at 4° C. Protein was precipitated with between 6-20% (w v<sup>-1</sup>) PEG 4000 and re-suspended in 10 mM sodium phosphate pH 7.4, 45 mM NaCl, 10 mM EDTA. The suspension was passed over I.D. 2.5 cm x 4 cm loosely packed lysine resin (lysine sepharose 4b, GE Healthcare Life Sciences), and the flow through was then passed over I.D. 5 cm x 4 cm of DEAE resin (DEAE sepharose fast flow, GE Healthcare Life Sciences). The protein was eluted with a gradient from 10 mM sodium phosphate pH 7.4, 45 mM NaCl, 10 mM EDTA to 10 mM sodium phosphate pH 7.4, 350 mM NaCl, 10 mM EDTA. C9-containing fractions from DEAE were pooled and loaded onto I.D. 2.5 cm x 4 cm ceramic hydroxyapatite resin (CHT type I, BioRAD) equilibrated in 10 mM sodium phosphate pH 7.0, 100 mM NaCl. Protein was eluted with a sodium phosphate gradient from 45 mM to 350 mM pH 8.1. The fractions containing C9 were further purified by size exclusion chromatography (Superdex 200 resin, GE Healthcare Life Sciences) in I.D. 2.6 cm x 60 cm column in 10 mM Hepes pH 7.2, 200 mM NaCl, and 10 mM EDTA.

The protein underwent an additional chromatography step using MonoQ ion exchange chromatography (GE Healthcare Life Sciences). Here, pooled fractions from the size exclusion chromatography step were concentrated using a 30 kDa MWCO centricon concentrator and buffer exchanged 2-3 times into 10 mM Tris-HCl pH 7.2, 100 mM NaCl. Buffer exchanged protein was loaded onto a 1 mL MonoQ column and eluted over a linear gradient from: 10 mM Tris-HCl pH 7.2, 100 mM NaCl to 10 mM Tris-HCl pH 7.2, 350 mM NaCl.

### ***Haemolytic assays and MAC assembly on ghost cell membranes***

Sheep red blood cells (sRBC) were washed with DGHB++ pH 7.4 (Dextrose Gelatin HEPES Buffer; with 2.5% (w v<sup>-1</sup>) D-glucose, 0.1% (w v<sup>-1</sup>) gelatin, 5 mM HEPES pH 7.4, 71 mM NaCl, 0.15 mM CaCl<sub>2</sub>, 0.5 mM MgCl<sub>2</sub>). 6.5 x 10<sup>8</sup> sRBC were sensitized with an equal volume of anti-sheep antibody to a concentration of 0.75 mg ml<sup>-1</sup> and incubated at 30°C for 30 min to activate the classical pathway. Excess antibody was washed off prior to reactions. The lysis reactions were set up in triplicate with 0.2 ml sRBC (3.75 x 10<sup>6</sup> cells), 1 µl of C9 depleted serum (Complement Technology, USA) and 15 µl of C9 (initial concentration 8 µg ml<sup>-1</sup>) and 2-fold dilutions of purified C9 in thin walled PCR tubes. Reaction tubes were incubated at 37°C on a PCR heat block for 30 minutes and immediately placed at 4°C, then centrifuged for 20 seconds. The supernatant (150 µl) of the lysis reactions was transferred to a 96-well plate and

absorbance at 405 nm was measured. The reactions were normalized to 0% lysis with a buffer control or to a reaction containing PlyA and PlyB <sup>7</sup> for 100% lysis.

Ghost membranes were prepared with rabbit red blood cells (rRBC) washed with DGHB++. Packed red blood cells were pre-incubated with C9 depleted serum for 15 minutes at 37°C. The rRBCs were washed of excess sera and re-suspended in DGHB++. Purified C9 protein was added to rRBCs and incubated at 37°C for 15 min. Reactions were immediately centrifuged for 30 seconds at 16100 rcf and the supernatant transferred to new tubes. The supernatant was further centrifuged for 10 minutes at 16100 rcf and the pellet containing membranes washed once with DGHB++ and then resuspended in 10 mM phosphate buffer pH 8.0, 50 mM NaCl to make ghosts. Carbon coated copper grids containing formvar were glow discharged, then floated over 8  $\mu$ L of re-suspended ghosts followed by staining with 2% (w v<sup>-1</sup>) uranyl acetate for 1 minute. Pores were examined on a Hitachi H7500 TEM at 80 kV.

#### ***Characterisation of the glycosylation state***

Purified C9 was reduced with 2 mM DTT, alkylated with 5 mM iodoacetamide and digested with trypsin (Promega) in 1:40 ratio at 37°C overnight. The digest was desalted with C18-packed tips (OMIX, Agilent) prior to nanoLC-MS/MS (Dionex Ultimate 3000 LC coupled to QExactive Plus, Thermo). Peptides (~1  $\mu$ g) were loaded on a 2 cm trap column (100  $\mu$ m ID, Acclaim PepMap 100, Thermo Scientific) in 2% (v v<sup>-1</sup>) acetonitrile, 0.1% (v v<sup>-1</sup>) trifluoroacetic acid and resolved on a 50 cm column (75  $\mu$ m ID, Acclaim PepMapRSLC, Thermo Scientific) with a non-linear 25 minute gradient from 2% (v v<sup>-1</sup>) to 34% (v v<sup>-1</sup>) acetonitrile in 0.1% (v v<sup>-1</sup>) formic acid. Spectra were acquired in a Top12 strategy with full scans (375-1800 m z<sup>-1</sup>) acquired at 70000 resolution and data-dependent HCD MS2 spectra acquired at 17500 resolution. Peptide assignment was performed with the Preview and Byonic software (Protein Metrics <sup>8</sup>) utilizing Preview-determined modifications and mass tolerances, a focused human database from an initial Byonic search and N- and O-glycosylation databases for assignment of glycan compositions. All glycan composition assignments were manually validated. Skyline software (University of Washington) was employed for semi-quantitative assessment of site-specific glycan compositions, using parent scan extracted ion chromatograms <sup>9</sup> (Supplementary Fig. 7).

## Supplementary References

1. Hofsteenge, J., Blommers, M., Hess, D., Furmanek, A. & Miroshnichenko, O. The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. *J. Biol. Chem.* **274**, 32786–94 (1999).
2. DiScipio, R. G. & Hugli, T. E. The architecture of complement component C9 and poly(C9). *J. Biol. Chem.* **260**, 14802–9 (1985).
3. Kondos, S. C. *et al.* The structure and function of mammalian membrane-attack complex/perforin-like proteins. *Tissue Antigens* **76**, 341–51 (2010).
4. Baran, K. *et al.* The molecular basis for perforin oligomerization and transmembrane pore assembly. *Immunity* **30**, 684–95 (2009).
5. Biesecker, G. & Müller-Eberhard, H. J. The ninth component of human complement: purification and physicochemical characterization. *J. Immunol.* **124**, 1291–6 (1980).
6. Biesecker, G., Lachmann, P. & Henderson, R. Structure of complement poly-C9 determined in projection by cryo-electron microscopy and single particle analysis. *Mol. Immunol.* **30**, 1369–1382 (1993).
7. Lukyanova, N. *et al.* Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin. *PLoS Biol.* **13**, e1002049 (2015).
8. Bern, M., Kil, Y. J. & Becker, C. Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics* **Chapter 13**, Unit13.20 (2012).
9. Schilling, B. *et al.* Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11**, 202–14 (2012).

## Additional Methods

### *Homology model of human C9*

Our structural model of poly-C9 relies on the the flexible-fitting of 22 individual subunits of C9 into the determined cryo-EM map of the poly-C9 assembly. A monomeric C9 homology model was thus generated given the absence of an existing crystallographic structure. The resulting C9 model was then used as the reference structure for the flexible-fitting procedure. The crystallographic structures of the MAC components C6, C8 $\alpha$  and C8 $\beta$  served as structural templates. These proteins that take part in the final MAC assembly and have long been identified as C9 sequence and structure homologues [23-26].

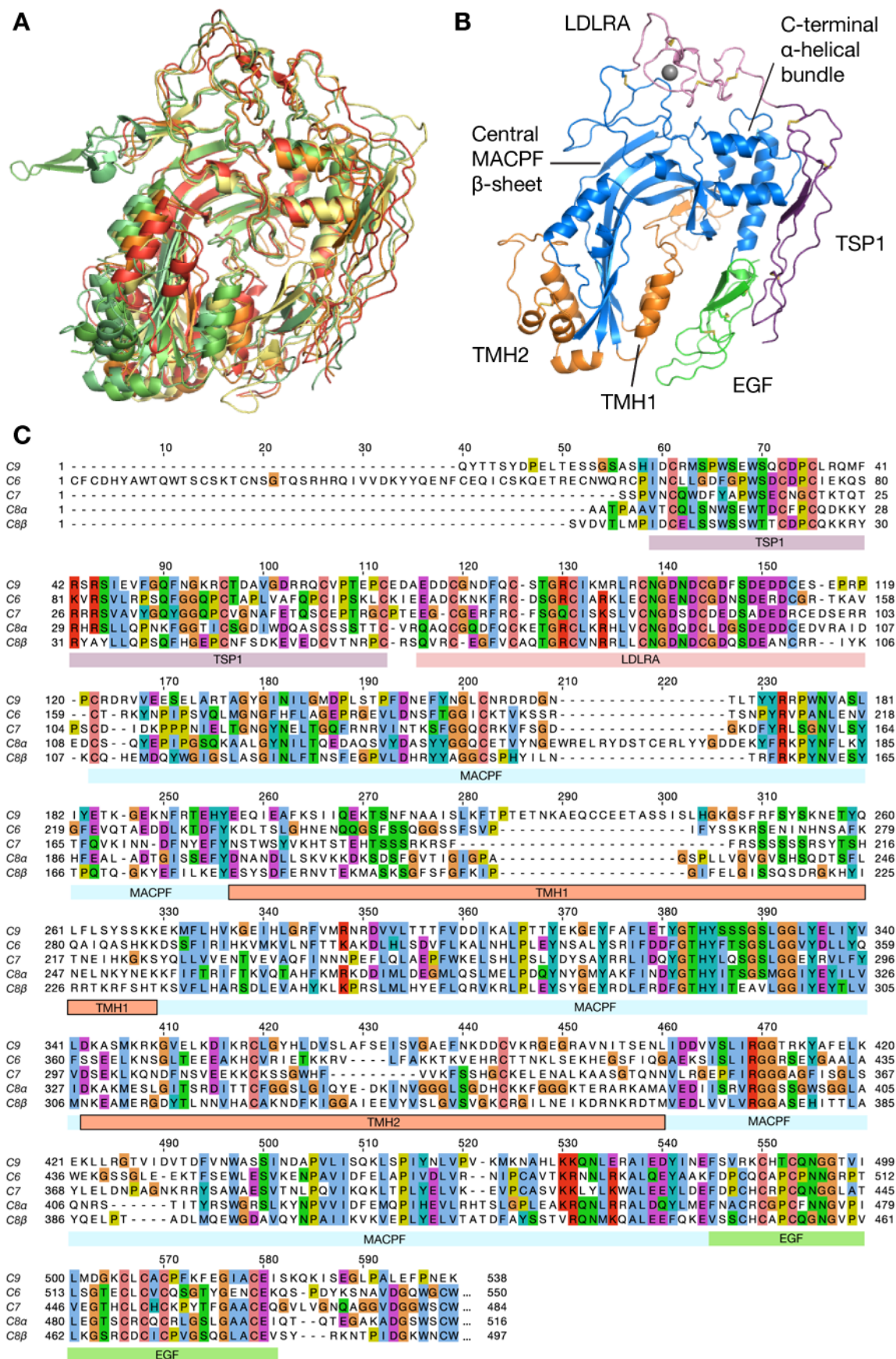
The multiple sequence alignment obtained using ClustalW 2.1 [27] (and also including C7) is presented in Figure 4.2.C. Sequence identities to C9 (over 538 positions; UniProt ID P02748) are as follows: 26.3% (C6; P13671), 26,4% (C7; P10643), 30.9% (C8 $\alpha$ ; P07357), 27.4% (C8 $\beta$ ; P07358). The sequence alignment was further manually edited based on the structural alignment generated with MatchMaker [28] (2.4 Å RMSD over 273 positions; Figure 4.2.A).

### **Figure 4.2. Structure and sequence alignments of MACPF-domain containing proteins of the human Complement**

A. Structure alignment of crystallographic of C6 (PDB ids: 3T5O, red; 4A5W, orange), C8 $\alpha$  (2RD7, light green ; 3OJY:A, dark green) and C8 $\beta$  (3OJY:B, yellow). Only the regions of the structures that display sequence homology to C9 are depicted.

B. Homology model of human C9. The TSP1 domain is in purple; the LDLRA domain in pink and the EGF domain in green. The MACPF domain is in blue, the TMH1/2 regions are highlighted in orange. Regions of the protein discussed in the text are highlighted.

C. Sequence alignment of C6,C7, C8 $\alpha$ , C8 $\beta$  and C9. The sequence and numbering of C9 excludes the 21 residues long signal peptide. Proteins domains and regions are boxed below the alignment with colours identical to B. Amino-acids are coloured using the ClustalX scheme [27] highlighting conservation of physico-chemical properties at the corresponding alignment position.



**Figure 4.2. (Continued) Structure and sequence alignments of MACPF-domain containing proteins of the human Complement**

Despite the relatively modest level of sequence identity, the structural alignment emphasises the significant conservation and arrangement of the core domains, all shared with C9 (LDLRA, TSP1, MACPF and EGF domains). Moreover, the alignment highlights the regions of high structural similarity: the LDLRA motif, the TSP1 domain and the structural core of the MACPF domain including the C-terminal  $\alpha$ -helical bundle and the upper portion of the central MACPF  $\beta$ -sheet, as previously noted [23-26, 29] (Figure 4.2.A-B). It is apparent from the sequence and structural alignments that the most variable positions correspond to the TMH1/2 regions together with the lower portion of the MACPF  $\beta$ -sheet (Figure 4.2.A). This most likely reflects the conformational lability of both the TMH1 and TMH2 regions as well as the ability of the central MACPF  $\beta$ -sheet to open upon pore formation [26, 29, 30].

Modeller 9.14 [31] was then used to generate 2500 homology models of C9 that were further ranked according to the normalised DOPE score [32]. The top-ranking model is shown in Figure 4.2.B and was used as the C9 structural model during the flexible-fitting procedure.







## **Chapter 5**

### **General Discussion**

## 5.1 Introduction

Cholesterol Dependent Cytolysins (CDCs) are a well-studied family of  $\beta$ -pore forming toxins mostly secreted by Gram-positive bacteria. They form giant, oligomeric ring-shaped transmembrane channels 200 to 300 Å in diameter. Such large pores mediate bacterial pathogenicity through direct cell lysis, synergy with other virulence factors (that move through the pore) or by triggering host invasion by the bacterium. Extensive biochemical and biophysical studies spanning over two decades have greatly contributed to our understanding of how CDCs convert from solubly secreted monomeric state to form an oligomeric membrane embedded pore [3, 7, 8, 33-37]. However, because of the limited resolution of published Single Particle cryo-Electron Microscopy (SP cryo-EM) data obtained [3] key aspects of CDC pore structure and the underlying molecular mechanisms associated with CDC pore formation have remained unclear.

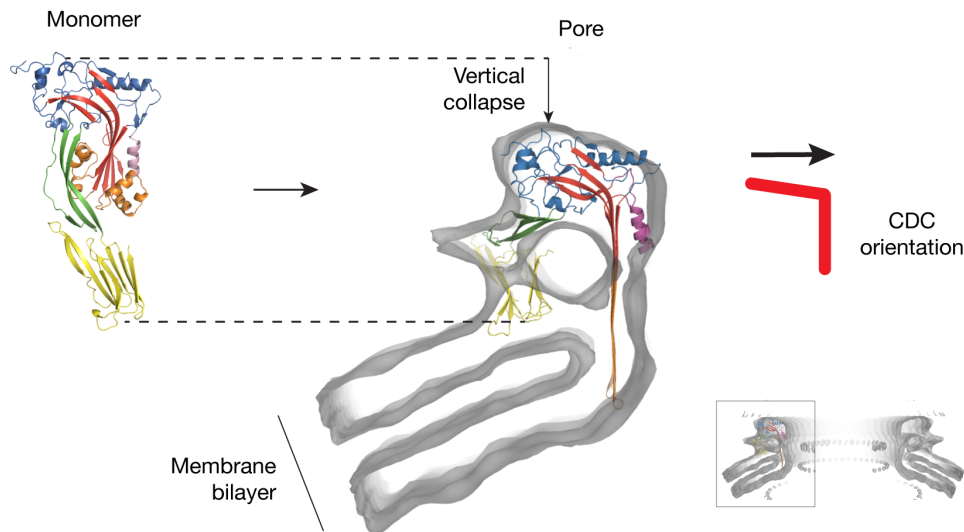
The evolutionary relationship between Membrane Attack Complex / PerForin (MACPF) and CDCs was first identified through structural studies. These data revealed that, despite limited sequence identity, the pore forming region of both families was homologous [14, 38]. It was further suggested that these proteins function in an analogous fashion. Accordingly, proteins sharing this fold have since been referred to as members of MACPF/CDC superfamily [39]. Members of the MACPF family can now be identified in all kingdoms of life and perform functions in animal and plant immunity, developmental biology, venom toxins and parasite invasion and egress. The mechanism of function of these proteins most commonly involves pore formation [39, 40].

In 2010, structural and biochemical studies of perforin, including the first SP cryo-EM structure of a MACPF protein, led to the proposal that perforin would form pores in a manner distinct to CDCs [41]. Specifically, labelling experiments suggested that the MACPF/CDC domain may adopt an orientation reverse to that of CDCs, the “inside-out” orientation (Figure 5.1). However, the low resolution of the data (28.5 Å) precluded further understanding of the structural basis for perforin pore formation [41].

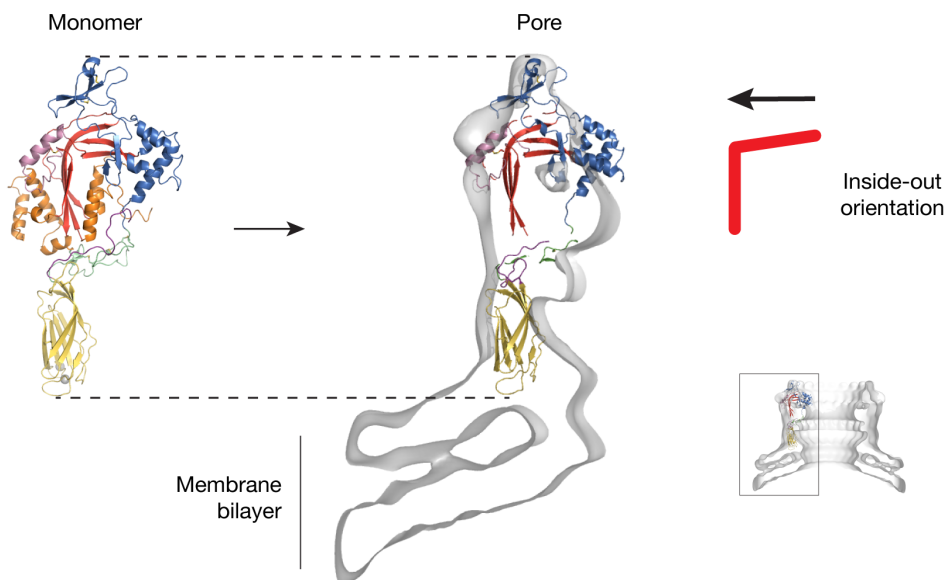
The work presented in this thesis aimed to address two critical questions with respect to the MACPF/CDC pore forming mechanism. Firstly, the pneumolysin pore structure was re-examined using modern molecular modelling methods in order to better understand the pore forming mechanism of the CDC branch of the superfamily (Chapters 2, 3). Secondly, in order to gain further insights into MACPF pore formation a molecular model of the MACPF immune effector C9 oligomer (termed poly-C9) was built using the sub-nanometer SP-cryo-EM maps provided by collaborators (Chapter 4).

Collectively, this research aimed at producing a critical assessment of the existing data and hypothesis, as well as expanding our general understanding of how giant pores are formed in the MACPF/CDC superfamily.

### A. Pneumolysin (CDC)



### B. Perforin (MACPF)



**Figure 5.1 Orientation of the MACPF/CDC domain in the pore form**

A. MACPF/CDC domain orientation in CDCs. Left panel: perfringolysin O X-ray conformation is shown in cartoon representation [34]. The orientation of the central  $\beta$ -sheet characteristic of the MACPF/CDC fold (red) is schematically represented (right panel). The segment of the  $\beta$ -sheet that inserts into the membrane lines the pore lumen and forms the transmembrane  $\beta$ -barrel (middle panel). The remainder of the MACPF/CDC domain is in blue, Domain 2 in green and the membrane binding Domain 4 in yellow. The 38-mer pneumolysin pore cryo-EM density is shown in grey [3]. The change in height ( $\sim 40$  Å) of the monomer upon membrane insertion is often referred to as the CDC ‘vertical collapse’ [3, 7, 8].

B. Proposed orientation of the MACPF/CDC domain in the perforin pore. The crystallographic conformation of perforin is shown (left panel) and found compatible in height with the cryo-EM reconstruction of the 20-mer perforin pore (grey, middle panel) [41]. In the published model of perforin the orientation of the MACPF  $\beta$ -sheet (red, right panel) is reverse to that of CDCs: the region of the bent  $\beta$ -sheet that extends towards the membrane bilayer lines the outer edge of the assembly. The remainder of the MACPF/CDC domain is in blue, the membrane-binding C2 domain in yellow and the EGF-like domain in green.

## 5.2 A novel $\beta$ -barrel architecture composes the CDC pore

As determined by numerous studies, CDCs is proposed to form pores via three discrete steps. Firstly, solubly secreted CDC monomers recognise and bind to cholesterol-rich membranes. Secondly, 30 to 50 subunits then assemble into a ring-shaped oligomer termed the prepore. This prepore complex then undergoes dramatic conformational changes that result in the insertion of the giant  $\beta$ -barrel pore into the bilayer membrane.

It has long been established that CDCs perforate the target membrane by deploying a transmembrane  $\beta$ -barrel [35, 42]. Here, two sets of  $\alpha$ -helices termed TMH1/2 of the MACPF/CDC domain undergo substantial remodelling and refold into two  $\beta$ -hairpins upon membrane insertion. Each monomer thus contributes 2  $\beta$ -hairpins (4  $\beta$ -strands) to the final transmembrane  $\beta$ -barrel formed by 120-200  $\beta$ -strands. This  $\beta$ -barrel originates and protrudes from the MACPF/CDC domain.

CDCs form giant  $\beta$ -barrels that dramatically contrast with other smaller  $\beta$ -pore forming toxins (PFTs). The latter proteins typically form pores less than 25 Å in diameter that comprise a maximum of 18  $\beta$ -strands. In the CDC  $\beta$ -barrel pore, the  $\beta$ -strands had originally been hypothesised to be orientated perpendicular to the membrane surface (tilt of 0° with respect to the barrel axis) [3]. Notably, this  $\beta$ -barrel architecture has never been observed at atomic resolution: in all architectures reported to date the  $\beta$ -strands adopt a tilt greater than 29° [43].

All possible  $\beta$ -barrel architectures were modelled employing established geometric principles from small  $\beta$ -barrels (<25  $\beta$ -strands) and refined structural parameters

(Chapter 2 and [44]). The results presented in Chapter 2 identified that a single  $\beta$ -barrel architecture (tilt of  $19^\circ$ , termed  $S = n/2$  architecture; where  $S$  is the shear number and  $n$  the number of  $\beta$ -strands) best explains the dimensions (diameter and height) of the CDC pneumolysin 38-mer pore. This unique architecture had never been identified in any PFP or other soluble protein.

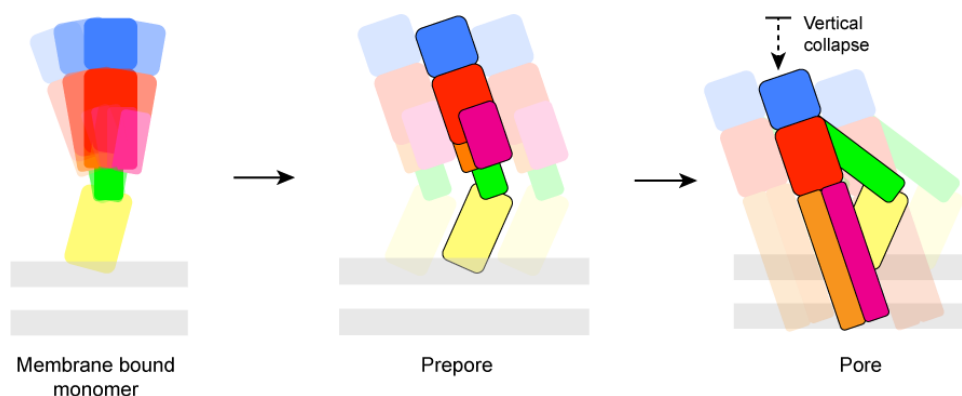
This surprising prediction was experimentally validated by others in 2013 using systematic disulphide link scanning [4]. That CDCs employ a unique  $\beta$ -barrel architecture sets them further apart from other  $\beta$ -PFTs. In addition, this novel architecture raised the question of whether the molecular modelling previously performed on the CDC pneumolysin pore needed re-examination [3]. Moreover, this also asked the question whether MACPF PFPs perforate the target bilayer employing the identified  $\beta$ -barrel structure.

### **5.3 Coordinated domains movements best explain the CDC ‘vertical collapse’**

Another major well-documented structural transition associated with membrane insertion is the ‘vertical collapse’ of the CDC prepore, that sees the MACPF/CDC domain descend onto the membrane by  $\sim 40$  Å. Critically, this collapse brings the domain close to the membrane surface and enables the insertion of the TMHs to reach and span the membranes as  $\beta$ -hairpins. Molecular modelling studies performed in 2005 had proposed that the dramatic change in altitude of the MACPF/CDC domain derives from the crumpling of Domain 2 [3], an elongated twisted  $\beta$ -sheet bridging the common MACPF/CDC domain (Domains 1 and 3 of CDCs) and the membrane binding domain (Domain 4).

In order to better characterise the conformational transitions associated with the prepore-to-pore-transition, systematic structural analysis of all available CDC crystallographic structures (the majority of which were unavailable in 2005) together with molecular dynamics simulations were performed. This study presented in Chapter 3 indicated that Domain 2 did not have the propensity to crumple, in contrast with the previously published model of CDC pore formation.

Motivated by these data, the conformation of the 38-mer pneumolysin pore was re-examined. It was found that pneumolysin pore conformation could be modelled within the published cryo-EM reconstruction with an unaltered Domain 2. The proposed structural rearrangement was predicted to take place *via* a significant rotation of Domain 2 ( $\sim 60^\circ$ ) such that it lies parallel to the membrane surface atop the membrane binding Domain 4. As a result, the suggested rotation of Domain 2 brings the MACPF/CDC domain close enough to the membrane surface for efficient  $\beta$ -barrel insertion. This new model of CDC pore formation therefore suggested that the CDC vertical collapse originated from the coordinated rotations of Domain 2 throughout the prepore oligomer to finally form the membrane embedded pore (Figure 5.2).



**Figure 5.2 Revised model of CDC membrane insertion**

Schematic representation of proposed CDC pore formation. Domain 1 is in blue, Domain 2 in green, Domain 3 in red and Domain 4 in yellow. After membrane binding, the flexible monomers self-oligomerise into the prepore complex (30-50 subunits). Upon membrane insertion, the orientation of Domain 2 flattens with respect to the membrane. This is accompanied by a vertical collapse of Domain 1 and 3, thereby bringing them closer to the bilayer surface and allowing insertion of TMH1/2 as  $\beta$ -hairpins (orange and purple respectively).

This new model of vertical collapse highlighted the complex molecular mechanisms underlying CDC bacterial attack. Further support for this revised model came in 2014 with the 15 Å resolution cryo-EM structures of the 37-mer CDC suilysin prepore and pore. Here it was concluded that the rotation of Domain 2 best explained the suilysin pore conformation [10] (reproduced in appendix A). The improved resolution of the data thus directly substantiated the mechanism of pore formation presented in Chapter 3.

#### **5.4 Poly-C9 structure provides significant insights into MAC assembly and MACPF pore formation**

In order to gain insights into the less understood MACPF pore formation, a structural model of poly-C9 was built using molecular modelling methods and based on an 8 Å resolution SP cryo-EM map (Chapter 4). At this resolution  $\alpha$ -helices are



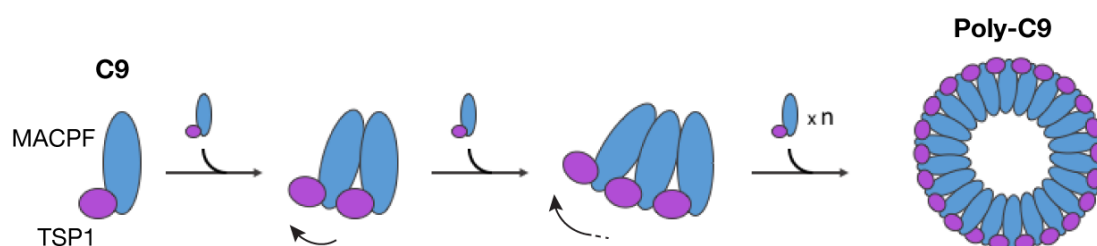
identifiable and the individual domains of C9 can be unambiguously assigned. This work has highlighted key structural aspects of a MACPF pore at the highest resolution obtained to date.

The poly-C9 ring formed by 22 tightly packed C9 molecules is ~120 Å in inner diameter, compatible with the passage of lysozymes through the pore lumen [11]. The overall height of the assembly (~170 Å) is consistent with that of the membrane inserted MAC and suggests that the MAC does not undergo a CDC-like collapse upon membrane insertion. Accordingly, the  $\beta$ -barrel formed from the refolding of C9 TMH regions is unusually high (~100 Å versus ~60 Å in CDCs). Further the structure of the barrel is consistent with the  $S = n/2$  architecture identified for CDCs. Further comparisons of the Poly-C9 conformation with crystallographic structures of other MAC components (C6, C8) suggested that only a modest opening of the central MACPF/CDC  $\beta$ -sheet is required upon  $\beta$ -barrel formation.

In poly-C9, oligomeric contacts are mostly accounted for by the packing of adjacent MACPF/CDC domains. However, the N-terminal TSP1 domain, located at the wedge between MACPF/CDC domains, contributes a notable fraction of the contacts between C9 subunits (23% of the oligomeric interface). Unexpectedly, this ancillary domain is likely to play an architectural role and contribute to the overall poly-C9 stability. This domain is also suggested to confer stability to the MAC, given that the TSP1 domain is uniquely found in all terminal MAC components and no other reported MACPF PFP.

In addition, it is noted that the TSP1 domain is located clockwise with respect to C9 MACPF/CDC domain (top view, Figure 5.3), consistently matching the reported unidirectional assembly of poly-C9 and the MAC [45, 46]. It is thus speculated that

the free TSP1 domain at the edge of the nascent poly-C9 (and likely of the nascent MAC) may contact available C9 subunits and play a role in their sequential recruitment to the final ring (Figure 5.3).



**Figure 5.3 Schematic model of Poly-C9 elongation and assembly**

Poly-C9 elongation is represented in a top view. Contacts between available C9 subunits and the free TSP1 domain (purple, clockwise edge) are suggested to participate in the sequential and unidirectional elongation of Poly-C9. The MACPF/CDC domain is in blue. Arrows indicate the clockwise direction of Poly-C9 elongation.

Finally the MACPF/CDC domain was found to adopt a CDC-like orientation, in contradiction with the proposed perforin ‘inside-out’ orientation. It is thus suggested that poly-C9 procures an adequate template for understanding perforin pore formation since both MACPF vertebrate immune effectors share together a closer evolutionary relationship (25% sequence identity) than with the bacterial CDCs (<10%). Consequently, this work supports that the perforin ‘inside-out’ orientation is not a MACPF-specific feature and, on the contrary, that perforin adopts the common CDC/C9 orientation. However, the molecular details underlying perforin pore formation remain to be determined.

## 5.5 Concluding remarks

The work presented in this thesis has revealed hitherto unsuspected structural and mechanistic insights into the MACPF/CDC superfamily. The new model of pore formation proposed for CDCs and the structure of the MACPF poly-C9 allow for a comparison of the pore forming mechanisms deployed by both branches of the superfamily, summarised in Table 5.1.

Feature Molecule	Prepore intermediate	Membrane Insertion as $\beta$ -Barrel	$\beta$ -Sheet opening	MACPF/CDC domain orientation	Vertical drop	Involvement of ancillary domains
<b>CDCs</b>	Yes	S = n/2 architecture	Yes	TMH2 faces the lumen*	Yes	Coordinated rotations
<b>Poly-C9 / MAC</b>	Unknown	S = n/2 architecture	Yes	TMH2 faces the lumen*	Unlikely	Oligomer stability
<b>Perforin</b>	Suspected [47]	Inferred from sequence	?	TMH2 faces the lumen*§	Unlikely	?

**Table 5.1 Structural features of MACPF/CDC pore formation**

The common MACPF/CDC pore forming mechanism identified in this work is highlighted in grey boxes.

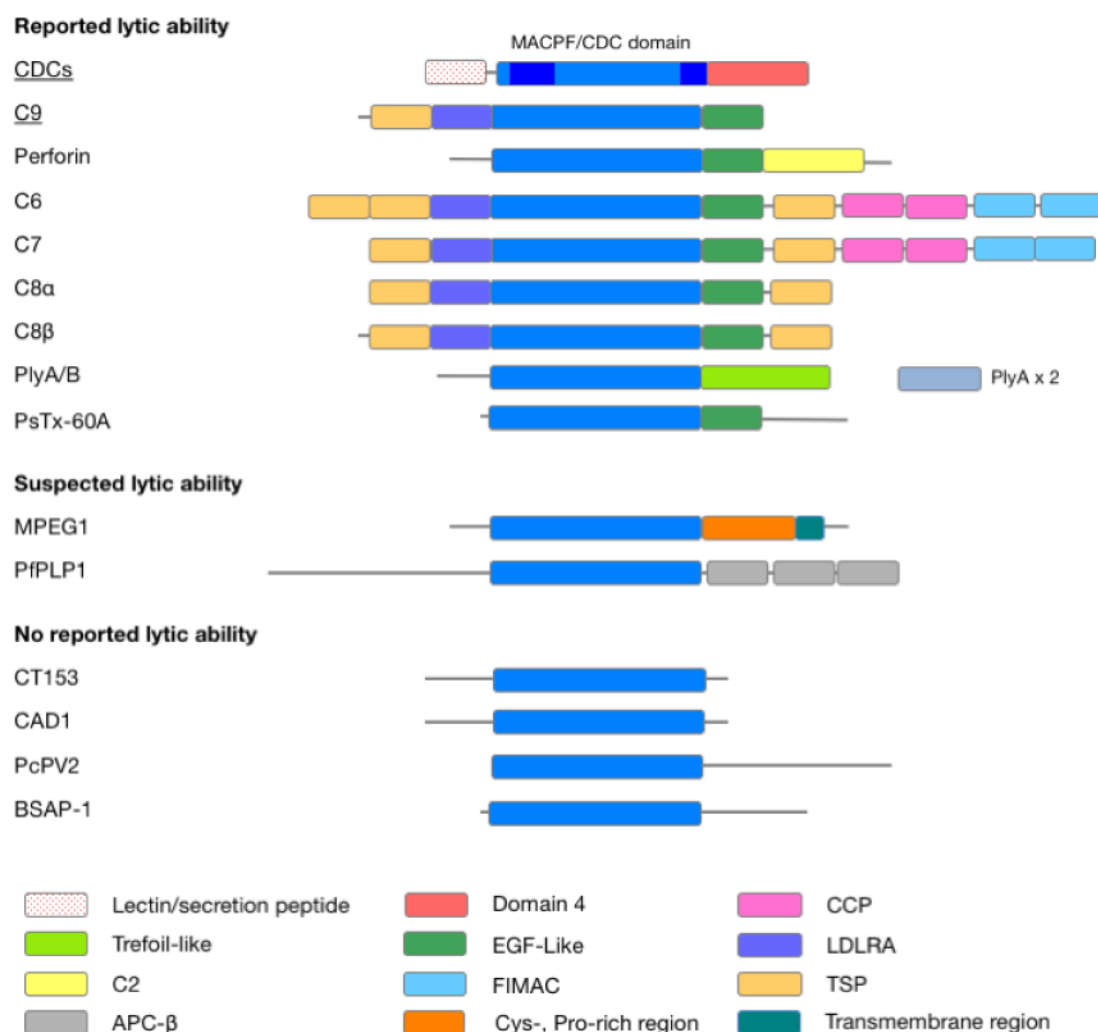
\* Also referred to as the CDC-like orientation.

§ Suggested in this work.

The work presented in this thesis has identified key features that constitute the common MACPF/CDC pore forming mechanism, as originally proposed. They include the orientation of the MACPF/CDC domain and the same transmembrane  $\beta$ -barrel architecture deployed after analogous conformational steps of the MACPF/CDC domain. Although perforin has been proposed to adopt the ‘inside-out’ orientation, our work in contrast supports a CDC/C9-like orientation.

In addition to these commonalities major structural and mechanistic differences between the CDC and MACPF branches of the superfamily have come to light. In contrast to CDCs, which undergo dramatic structural changes (coordinated rotation of Domain 2 and ‘vertical collapse’), the MACPF C9 deploys an unusually high  $\beta$ -barrel rendering a CDC-like vertical collapse unnecessary to reach the target membrane. However, it is important to note that in both cases the domains ancillary to the MACPF/CDC domain are involved in the formation of the final oligomer, albeit differently (Domain 2 enabled CDC ‘vertical collapse’ versus the architectural C9-TSP1). This suggests that the molecular mechanisms underlying pore formation in the MACPF/CDC superfamily may have co-evolved with the ancillary domains themselves.

These hypotheses are further substantiated by recent structural studies of the pleurotolysin pore, a MACPF fungal PFT. Pleurotolysin pore formation was exempt of a ‘vertical collapse’ while the domains ancillary to the MACPF domain were found to play an architectural role in the pore complex [30] (reproduced in appendix B). In summary, given the diverse domain composition in the MACPF family (Figure 5.4), it is hypothesised that each MACPF PFP may have evolved a variation of the general MACPF/CDC pore forming mechanism.



**Figure 5.4. Schematic domains composition across the MACPF/CDC superfamily**

The MACPF/CDC domain is in blue. In CDCs the domain includes Domain 2 (dark blue) to form a folded continuous 350 amino-acid structural entity usually referred to as Domains 1-3 [48]. CDCs possess a secretion signal anterior to the MACPF/CDC domain. In the CDC lectinolysin a lectin domain is present and suggested to enhance pore formation [49, 50]. Perforin, MAC members C6, C7 C8α, C8β, C9, the two-component pleurotolysin (PlyA/B) [30] and the sea anemone toxin PsTx-60A/B [51, 52] have reported lytic ability. MPEG1 (found in animal macrophages) [53] and PfPLP1 (secreted by the parasite *Plasmodium falciparum*) [54] are suspected to exhibit lytic activity. CT153 (of chlamydial origin) [55], PcPV2 (snail egg neurotoxin) [56], CAD1 (part of the plant immune system) [57], BSAP-1 (an antimicrobial protein secreted by the gut symbiot *Bacteroides fragilis*) [58] and astrotactin1/2 (involved in neural migration) [59] have no reported pore-forming ability. PcPV2 is disulphide linked to a lectin-like domain susceptible to address the protein to the membrane [56].

In contrast to the bacterial CDCs, MACPF proteins are found in all kingdoms of life and perform important roles including immune defence, venom toxicity and development. While members of the MACPF branch of the MACPF/CDC superfamily involved in immunity (including perforin and the MAC) have been the focus of numerous studies, it remains to be determined which MACPF proteins form pores, how they form pores and how this relates to their function. Finally, despite the advances presented here, a near atomic resolution structure of a MACPF/CDC pore still remains to be determined. Further understanding of the detailed pore-forming mechanism by MACPF/CDCs awaits such an advance.







## References

1. Reboul, C.F., et al., *Predicting giant transmembrane beta-barrel architecture*. Bioinformatics, 2012. **28**(10): p. 1299-302.
2. Nagano, N., E.G. Hutchinson, and J.M. Thornton, *Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels*. Protein Sci, 1999. **8**(10): p. 2072-84.
3. Tilley, S.J., et al., *Structural basis of pore formation by the bacterial toxin pneumolysin*. Cell, 2005. **121**(2): p. 247-56.
4. Sato, T.K., R.K. Tweten, and A.E. Johnson, *Disulfide-bond scanning reveals assembly state and beta-strand tilt angle of the PFO beta-barrel*. Nat Chem Biol, 2013. **9**(6): p. 383-9.
5. Goyal, P., et al., *Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG*. Nature, 2014. **516**(7530): p. 250-3.
6. Cao, B., et al., *Structure of the nonameric bacterial amyloid secretion channel*. Proc Natl Acad Sci U S A, 2014. **111**(50): p. E5439-44.
7. Czajkowsky, D.M., et al., *Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane*. EMBO J, 2004. **23**(16): p. 3206-15.
8. Ramachandran, R., R.K. Tweten, and A.E. Johnson, *The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation*. Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7139-44.
9. Reboul, C.F., J.C. Whisstock, and M.A. Dunstone, *A new model for pore formation by cholesterol-dependent cytolysins*. PLoS Comput Biol, 2014. **10**(8): p. e1003791.
10. Leung, C., et al., *Stepwise visualization of membrane pore formation by sullysin, a bacterial cholesterol-dependent cytolysin*. eLife, 2014.
11. Martinez, R.J. and S.F. Carroll, *Sequential metabolic expressions of the lethal process in human serum-treated Escherichia coli: role of lysozyme*. Infect Immun, 1980. **28**(3): p. 735-45.
12. Morgan, B.P., *The membrane attack complex as an inflammatory trigger*. Immunobiology, 2015.
13. Tschopp, J., A. Engel, and E.R. Podack, *Molecular weight of poly(C9). 12 to 18 C9 molecules form the transmembrane channel of complement*. J Biol Chem, 1984. **259**(3): p. 1922-8.

14. Tschopp, J., D. Masson, and K.K. Stanley, *Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytolysis*. Nature, 1986. **322**(6082): p. 831-4.
15. Marshall, P., et al., *Interaction between complement proteins C5b-7 and erythrocyte membrane sialic acid*. J Exp Med, 1996. **184**(4): p. 1225-32.
16. Bhakdi, S. and J. Trantum-Jensen, *Molecular nature of the complement lesion*. Proc Natl Acad Sci U S A, 1978. **75**(11): p. 5655-9.
17. Podack, E.R. and J. Tschopp, *Polymerization of the ninth component of complement (C9): formation of poly(C9) with a tubular ultrastructure resembling the membrane attack complex of complement*. Proc Natl Acad Sci U S A, 1982. **79**(2): p. 574-8.
18. Tschopp, J., H.J. Muller-Eberhard, and E.R. Podack, *Formation of transmembrane tubules by spontaneous polymerization of the hydrophilic complement protein C9*. Nature, 1982. **298**(5874): p. 534-8.
19. Steckel, E.W., B.E. Welbaum, and J.M. Sodetz, *Evidence of direct insertion of terminal complement proteins into cell membrane bilayers during cytolysis. Labeling by a photosensitive membrane probe reveals a major role for the eighth and ninth components*. J Biol Chem, 1983. **258**(7): p. 4318-24.
20. DiScipio, R.G. and C. Berlin, *The architectural transition of human complement component C9 to poly(C9)*. Mol Immunol, 1999. **36**(9): p. 575-85.
21. Biesecker, G., P. Lachmann, and R. Henderson, *Structure of complement poly-C9 determined in projection by cryo-electron microscopy and single particle analysis*. Mol Immunol, 1993. **30**(15): p. 1369-82.
22. Preissner, K.T., E.R. Podack, and H.J. Muller-Eberhard, *The membrane attack complex of complement: relation of C7 to the metastable membrane binding site of the intermediate complex C5b-7*. J Immunol, 1985. **135**(1): p. 445-51.
23. DiScipio, R.G., et al., *The structure of human complement component C7 and the C5b-7 complex*. J Biol Chem, 1988. **263**(1): p. 549-60.
24. Chakravarti, D.N., et al., *Structural homology of complement protein C6 with other channel-forming proteins of complement*. Proc Natl Acad Sci U S A, 1989. **86**(8): p. 2799-803.
25. DiScipio, R.G. and T.E. Hugli, *The molecular architecture of human complement component C6*. J Biol Chem, 1989. **264**(27): p. 16197-206.
26. Aleshin, A.E., et al., *Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC)*. J Biol Chem, 2012. **287**(13): p. 10210-22.
27. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-8.

28. Meng, E.C., et al., *Tools for integrated sequence-structure analysis with UCSF Chimera*. BMC Bioinformatics, 2006. **7**: p. 339.
29. Aleshin, A.E., et al., *Crystal structure of C5b-6 suggests structural basis for priming assembly of the membrane attack complex*. J Biol Chem, 2012. **287**(23): p. 19642-52.
30. Lukoyanova, N., et al., *Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin*. PLoS Biol, 2015. **13**(2): p. e1002049.
31. Eswar, N., et al., *Comparative protein structure modeling using Modeller*. Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.], 2006. **Chapter 5**: p. Unit 5 6.
32. Shen, M.Y. and A. Sali, *Statistical potential for assessment and prediction of protein structures*. Protein Sci, 2006. **15**(11): p. 2507-24.
33. Menestrina, G., C.L. Bashford, and C.A. Pasternak, *Pore-forming toxins: experiments with S. aureus alpha-toxin, C. perfringens theta-toxin and E. coli haemolysin in lipid bilayers, liposomes and intact cells*. Toxicon, 1990. **28**(5): p. 477-91.
34. Rossjohn, J., et al., *Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form*. Cell, 1997. **89**(5): p. 685-92.
35. Shatursky, O., et al., *The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins*. Cell, 1999. **99**(3): p. 293-9.
36. Ramachandran, R., R.K. Tweten, and A.E. Johnson, *Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment*. Nature structural & molecular biology, 2004. **11**(8): p. 697-705.
37. Bhakdi, S., J. Tranum-Jensen, and A. Sziegoleit, *Mechanism of membrane damage by streptolysin-O*. Infect Immun, 1985. **47**(1): p. 52-60.
38. Young, J.D., Z.A. Cohn, and E.R. Podack, *The ninth component of complement and the pore-forming protein (perforin 1) from cytotoxic T cells: structural, immunological, and functional similarities*. Science, 1986. **233**(4760): p. 184-90.
39. Rosado, C.J., et al., *The MACPF/CDC family of pore-forming toxins*. Cell Microbiol, 2008. **10**(9): p. 1765-74.
40. Anderluh, G., et al., *Distribution of MACPF/CDC proteins*. Subcell Biochem, 2014. **80**: p. 7-30.
41. Law, R.H., et al., *The structural basis for membrane binding and pore formation by lymphocyte perforin*. Nature, 2010. **468**(7322): p. 447-51.

42. Shepard, L.A., et al., *Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin Clostridium perfringens perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy*. Biochemistry, 1998. **37**(41): p. 14563-74.
43. Murzin, A.G., A.M. Lesk, and C. Chothia, *Principles determining the structure of beta-sheet barrels in proteins. I. A theoretical analysis*. J Mol Biol, 1994. **236**(5): p. 1369-81.
44. Murzin, A.G., A.M. Lesk, and C. Chothia, *Principles determining the structure of beta-sheet barrels in proteins. II. The observed structures*. J Mol Biol, 1994. **236**(5): p. 1382-400.
45. Podack, E.R., et al., *Membrane attack complex of complement: a structural analysis of its assembly*. J Exp Med, 1980. **151**(2): p. 301-13.
46. Tschopp, J., *Ultrastructure of the membrane attack complex of complement. Heterogeneity of the complex caused by different degree of C9 polymerization*. J Biol Chem, 1984. **259**(12): p. 7857-63.
47. Praper, T., et al., *Human perforin employs different avenues to damage membranes*. J Biol Chem, 2011. **286**(4): p. 2946-55.
48. Bourdeau, R.W., et al., *Cellular functions and X-ray structure of anthrolysin O, a cholesterol-dependent cytolysin secreted by Bacillus anthracis*. J Biol Chem, 2009. **284**(21): p. 14645-56.
49. Farrand, S., et al., *Characterization of a streptococcal cholesterol-dependent cytolysin with a lewis y and b specific lectin domain*. Biochemistry, 2008. **47**(27): p. 7097-107.
50. Feil, S.C., et al., *Structure of the lectin regulatory domain of the cholesterol-dependent cytolysin lectinolysin reveals the basis for its lewis antigen specificity*. Structure, 2012. **20**(2): p. 248-58.
51. Nagai, H., et al., *Novel proteinaceous toxins from the nematocyst venom of the Okinawan sea anemone Phyllodiscus semoni Kwietniewski*. Biochem Biophys Res Commun, 2002. **294**(4): p. 760-3.
52. Satoh, H., et al., *Characterization of PsTX-60B, a new membrane-attack complex/perforin (MACPF) family toxin, from the venomous sea anemone Phyllodiscus semoni*. Toxicon, 2007. **49**(8): p. 1208-10.
53. McCormack, R., et al., *Inhibition of intracellular bacterial replication in fibroblasts is dependent on the perforin-like protein (perforin-2) encoded by macrophage-expressed gene 1*. J Innate Immun, 2013. **5**(2): p. 185-94.
54. Garg, S., et al., *Calcium-dependent permeabilization of erythrocytes by a perforin-like protein during egress of malaria parasites*. Nat Commun, 2013. **4**: p. 1736.

55. Taylor, L.D., et al., *Biological characterization of Chlamydia trachomatis plasticity zone MACPF domain family protein CT153*. Infect Immun, 2010. **78**(6): p. 2691-9.
56. Dreon, M.S., et al., *Novel animal defenses against predation: a snail egg neurotoxin combining lectin and pore-forming chains that resembles plant defense and bacteria attack toxins*. PLoS One, 2013. **8**(5): p. e63782.
57. Morita-Yamamuro, C., et al., *The Arabidopsis gene CAD1 controls programmed cell death in the plant immune system and encodes a protein containing a MACPF domain*. Plant Cell Physiol, 2005. **46**(6): p. 902-12.
58. Chatzidaki-Livanis, M., M.J. Coyne, and L.E. Comstock, *An antimicrobial protein of the gut symbiont Bacteroides fragilis with a MACPF domain of host immune proteins*. Mol Microbiol, 2014. **94**(6): p. 1361-74.
59. Zheng, C., N. Heintz, and M.E. Hatten, *CNS gene encoding astrotactin, which supports neuronal migration along glial fibers*. Science, 1996. **272**(5260): p. 417-9.



## **Appendices**

## Appendix A

The following publication results from a collaboration initiated by the work presented in this thesis. This journal article makes a significant contribution to the CDC field of research and is reproduced in the article format.

*Stepwise visualization of membrane pore formation by sullysin, a bacterial cholesterol-dependent cytolysin*

eLife (2014), 3:e04247

C Leung, NV Dudkina, N Lukyanova, AW Hodel, I Farabella, AP Pandurangan, D Osmanović, **CF Reboul**, MA Dunstone, M Topf, R Lonnen, PW Andrew, HR Saibil, and BW Hoogenboom.



# Stepwise visualization of membrane pore formation by suilyisin, a bacterial cholesterol-dependent cytolysin

Carl Leung<sup>1†</sup>, Natalya V Dudkina<sup>2,3†</sup>, Natalya Lukyanova<sup>2,3</sup>, Adrian W Hodel<sup>1</sup>, Irene Farabella<sup>2,3</sup>, Arun P Pandurangan<sup>2,3</sup>, Nasrin Jahan<sup>4</sup>, Mafalda Pires Damaso<sup>4</sup>, Dino Osmanović<sup>1,5</sup>, Cyril F Reboul<sup>6</sup>, Michelle A Dunstone<sup>6,7</sup>, Peter W Andrew<sup>4</sup>, Rana Lonnen<sup>4</sup>, Maya Topf<sup>2,3</sup>, Helen R Saibil<sup>2,3\*‡</sup>, Bart W Hoogenboom<sup>1,5\*‡</sup>

<sup>1</sup>London Centre for Nanotechnology, University College London, London, United Kingdom; <sup>2</sup>Department of Crystallography, Birkbeck College, London, United Kingdom; <sup>3</sup>Institute of Structural and Molecular Biology, Birkbeck College, London, United Kingdom; <sup>4</sup>Department of Infection, Immunity, and Inflammation, University of Leicester, Leicester, United Kingdom; <sup>5</sup>Department of Physics and Astronomy, University College London, London, United Kingdom; <sup>6</sup>Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia; <sup>7</sup>Department of Microbiology, Monash University, Melbourne, Australia

\*For correspondence:

†These authors contributed equally as first authors to this work

‡These authors contributed equally as last authors to this work

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 15

**Received:** 05 August 2014

**Accepted:** 24 November 2014

**Published:** 02 December 2014

**Reviewing editor:** Volker Dötsch, Goethe University, Germany

© Copyright Leung et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

**Abstract** Membrane attack complex/perforin/cholesterol-dependent cytolysin (MACPF/CDC) proteins constitute a major superfamily of pore-forming proteins that act as bacterial virulence factors and effectors in immune defence. Upon binding to the membrane, they convert from the soluble monomeric form to oligomeric, membrane-inserted pores. Using real-time atomic force microscopy (AFM), electron microscopy (EM), and atomic structure fitting, we have mapped the structure and assembly pathways of a bacterial CDC in unprecedented detail and accuracy, focussing on suilyisin from *Streptococcus suis*. We show that suilyisin assembly is a noncooperative process that is terminated before the protein inserts into the membrane. The resulting ring-shaped pores and kinetically trapped arc-shaped assemblies are all seen to perforate the membrane, as also visible by the ejection of its lipids. Membrane insertion requires a concerted conformational change of the monomeric subunits, with a marked expansion in pore diameter due to large changes in subunit structure and packing.

DOI: [10.7554/eLife.04247.001](https://doi.org/10.7554/eLife.04247.001)

## Introduction

The bacterial CDCs and ubiquitous MACPF proteins are expressed as soluble monomers but assemble on membranes to form large, oligomeric pores. They form two branches of the largest superfamily of pore-forming proteins. Proteins of this MACPF/CDC superfamily share a common core topology of a highly bent and twisted  $\beta$ -sheet flanked by two  $\alpha$ -helical regions, though lacking any detectable sequence homology between the two branches (Rosado et al., 2008). Crystal structures of CDCs in their soluble, monomeric form (perfringolysin, Rossjohn et al., 1997; anthrolysin, Bourdeau et al., 2009; suilyisin, Xu et al., 2010; listeriolysin, Köster et al., 2014) revealed extended, key-shaped molecules. Pore-forming domains 1 and 3 (see also below) are linked by a long thin  $\beta$ -sheet (domain 2) to an immunoglobulin fold domain (4) which can bind to the membrane via a tryptophan-rich loop. CDCs form heterogeneous rings and arcs (Dang et al., 2005; Tilley et al., 2005; Sonnen et al., 2014) on cholesterol-rich liposomes and lipid monolayers (for example, the CDC perfringolysin O hardly binds to membranes with <30% molar concentration of cholesterol, Johnson et al., 2012). Extensive biophysical

**eLife digest** Many disease-causing bacteria secrete toxic proteins that drill holes into our cells to kill them. Cholesterol-dependent cytolysins (CDCs) are a family of such toxins, and are produced by bacteria that cause pneumonia, meningitis, and septicaemia.

The bacteria release CDC toxins as single protein molecules, which can bind to the membrane that surrounds the host cell. After binding to the membrane, the toxin molecules assemble in rings to form large pores in the host membrane. There are several stages to this process, but our understanding of what happens at the molecular level is incomplete.

Leung et al. studied suilysin, a CDC toxin produced by a bacterium that has a big impact on the pig farming industry because it causes meningitis in piglets. The bacterium can also cause serious diseases in humans through exposure to contaminated pigs or pig meat.

Leung et al. used a technique called electron microscopy to obtain atomic-scale snapshots of the toxin structures before and after the toxins were inserted into the membrane. In addition, real-time movies of the process were gathered using another technique called atomic force microscopy.

The experiments show that suilysin forms assemblies on the membrane that grow by one molecule at a time, rather than by the merging of larger assemblies of molecules. This results in a mixture of ring-shaped and arc-shaped toxin assemblies on the membrane. The arcs of suilysin are incomplete ring assemblies, but they are still able to make holes in the cell membrane. In order to insert into the membrane, the toxin molecules in the arcs and rings undergo a dramatic change in shape.

Understanding how CDCs assemble in membranes will guide further work into the development of new vaccines that can target these proteins to reduce the damage caused by bacterial infections.

DOI: [10.7554/eLife.04247.002](https://doi.org/10.7554/eLife.04247.002)

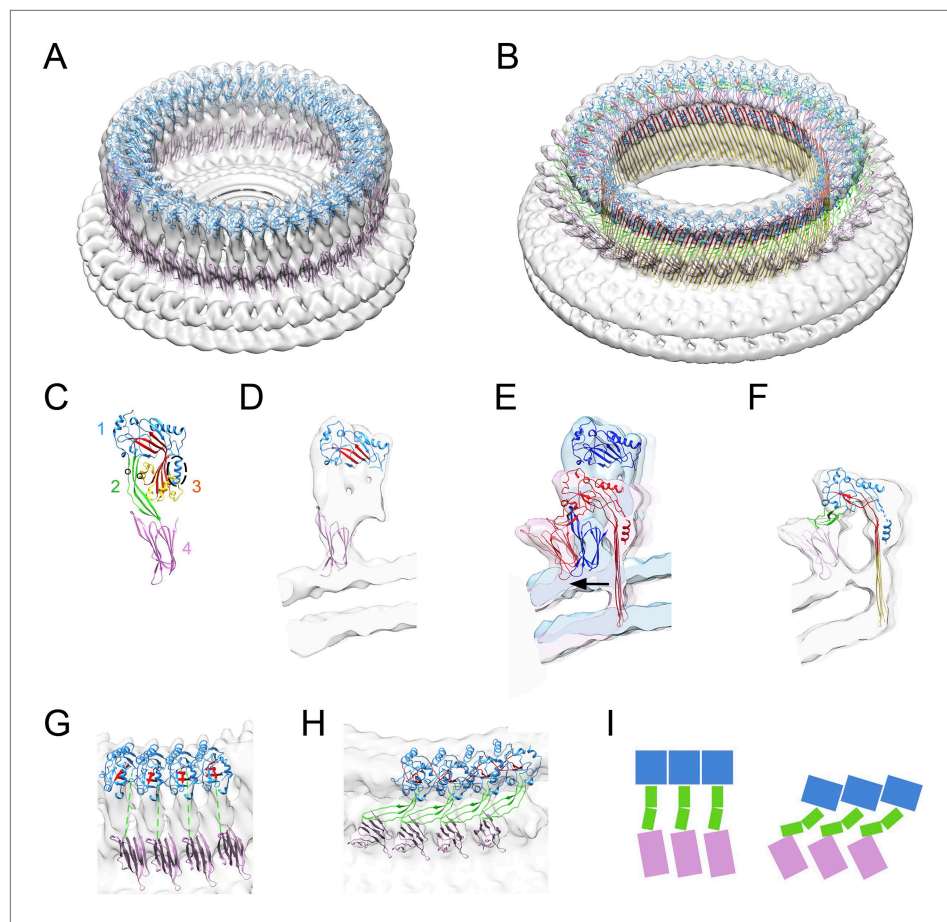
and molecular analysis of CDCs established that, on CDC binding to the membrane (Ramachandran et al., 2004; Hotze et al., 2012),  $\alpha$ -helical regions in domain 3 unfurl to form transmembrane  $\beta$ -hairpins, denoted as TMH1 and TMH2 (Shepard et al., 1998; Shatursky et al., 1999). If the TMH regions are trapped by introducing a disulphide bond (Hotze et al., 2001), prepore oligomers are formed on the membrane surface. Cryo-EM and single particle analysis of liposome-bound CDCs led to low-resolution 3D structures of prepore and pore forms of pneumolysin, a major virulence factor of *Streptococcus pneumoniae* (Tilley et al., 2005). These structures, as well as an AFM study of perfringolysin (Czajkowski et al., 2004), established that the 11 nm high molecule must collapse to a height of 7 nm above the membrane in order to insert the TMH regions. Simple pseudo-atomic models were obtained by fitting domains (broken at plausible hinge points) into the EM density maps. It was proposed that the long, thin  $\beta$ -sheet domain 2 collapses after the molecule opens up to release the TMH regions. However, because of the heterogeneity of the oligomeric assemblies and aggregation of the liposomes upon pore formation, resolution has been limited by the difficulty of obtaining sufficiently large data sets.

After comparing several CDCs (pneumolysin, suilysin, anthrolysin, and listeriolysin), we found that suilysin was less susceptible to these problems and we chose it to pursue new structural and dynamic studies. A disulphide-locked double cysteine mutant of suilysin, designed to prevent TMH1 insertion, enabled us to trap an active prepore state as well as to visualize the pore formation process by AFM in solution. Cryo-EM reconstruction and fitting revealed new details of the  $\beta$ -sheet unbending and changes in subunit packing upon conversion of prepores to pores. AFM images reveal that the prepore state is highly mobile. Following the addition of DTT to trigger insertion of the disulphide-locked prepore, time-lapse AFM yielded real-time movies of its conversion to ring and crescent-shaped pores. The observed distributions of rings and arcs can be explained by a theoretical model for kinetically trapped, noncooperative assembly, fully determined by the relative kinetics of monomer binding to the membrane and monomer assembly on the membrane surface. Together these studies provide substantial new understanding of the structure and dynamics of CDC pore formation.

## Results

### Conformational changes in prepore and pore states determined by cryo-EM

Negative stain EM and rotational symmetry analysis of complete rings of disulphide locked (Gly52Cys/Ser187Cys) suilysin prepores and wild-type pores formed on lipid monolayers revealed that most rings



**Figure 1.** Structural transitions during pore formation. 3D cryo-EM maps of 37-mer prepoire and pore forms of suilysin are shown with fitted atomic structures. (A) Density map of prepoire, surrounded by the extracted disk of membrane, with domains 1 and 4 fitted. (B) Density map of pore with all domains fitted, including the  $\beta$ -barrel with strands at 20° tilt. (C) Suilysin crystal structure with the domains labelled, showing positions of cysteines introduced in the locked form (black circles) and a helical domain adjacent to the bend in the central  $\beta$ -sheet (dashed oval). (D) Cross-section through one side of the prepoire map with the partial fit of atomic structures. (E) Overlay of one side of the prepoire (blue) and pore maps (red), aligned to the same centre, showing the displacement of domain 4 (arrow). (F) Pore section with fit. (G) View of 4 subunits from outside the prepoire. (H) View of 4 subunits from outside the pore. (I) Cartoons of domain packing in prepoire and pore.

DOI: [10.7554/eLife.04247.003](https://doi.org/10.7554/eLife.04247.003)

The following figure supplements are available for figure 1:

**Figure supplement 1.** Symmetry of suilysin prepoires and pores, determined by negative-stain EM.

DOI: [10.7554/eLife.04247.004](https://doi.org/10.7554/eLife.04247.004)

**Figure supplement 2.** Resolution curves for EM maps.

DOI: [10.7554/eLife.04247.005](https://doi.org/10.7554/eLife.04247.005)

**Figure supplement 3.** Electrostatic potential maps and interacting residues.

DOI: [10.7554/eLife.04247.006](https://doi.org/10.7554/eLife.04247.006)

**Figure supplement 4.** Comparison between prepoire and crystal structure conformations.

DOI: [10.7554/eLife.04247.007](https://doi.org/10.7554/eLife.04247.007)

contain 37 subunits (**Figure 1—figure supplement 1**). Unexpectedly, the diameter of the 37-fold suilysin prepoire was smaller than the diameter of the 37-fold pore (see below for quantification), indicating that conformational changes during pore formation are accompanied by changes in subunit packing.

3D reconstruction of suilysin prepoires and pores in liposomes was performed using a pseudo single-particle approach (Tilley et al., 2005), yielding a 15 Å cryo-EM map of the prepoire using the disulphide-locked construct (**Figure 1A,D**), and a 15 Å cryo-EM map of a wild-type suilysin pore (**Figure 1B,F**, see also **Figure 1—figure supplement 2**). The 3D maps, both of 37-mers, confirmed a

significant expansion in ring diameter upon pore formation. In order to interpret the maps, we performed flexible fitting of suilysin domains from the crystal structure (**Figure 1C**; Xu *et al.*, 2010; PDB:3hvn). Domain deformations and hinge movements were identified by normal mode analysis (Lindahl *et al.*, 2006). Additional evidence for the correctness of the fits followed from electrostatic potential maps and analysis of interacting residues at the interfaces of domain 1 (**Figure 1—figure supplement 3**). The results show that both models have extended regions of complementary charge on the predicted interacting surfaces. Although the extent of complementary charge is less in the pore model, in this case the oligomer is stabilized by the  $\beta$ -barrel of the pore. Measured from the fitted position of the base of domain 4, the prepore and pore diameters are 296 Å and 319 Å, respectively (**Figure 1E**), an expansion of 8%. In addition to the 4 nm reduction in height, each pore subunit approximately doubles in width.

The prepore is distorted from the crystal structure, with some collapse of domain 2 and opening of the  $\beta$ -sheet (**Figure 1—figure supplement 4**), despite the presence of the disulphide bridge (**Figure 1D**). However, the map features are not sufficiently defined to guide fitting, most likely owing to the greater flexibility of the prepore state, as described below.

The pore structure is similar to that observed with pneumolysin, but with improved resolution, and showing significant differences in the hinge bending and domain movements. The  $\beta$ -strands are tilted by 20°, in agreement with previous results (Reboul *et al.*, 2012; Sato *et al.*, 2013). The distortion to domain 2 differs from that proposed in the earlier model (Tilley *et al.*, 2005). Seen from outside the ring, domain 2 collapses sideways, to the right, such that domain 1 is aligned above the adjacent domain 4, with a sideways tilt that expands the ring. The expansion is clearly seen in the wider spacing between subunits at domain 4 (**Figure 1G,H,I**). Domain 2 must bend at a central hinge point to fit into the EM density (**Figure 1H,I**; Reboul *et al.*, 2014). As expected, there is a major opening of the bent  $\beta$ -sheet. In addition, a helical subdomain flanking the bend of the central  $\beta$ -sheet (residues 335–347, dashed oval in domain 3, **Figure 1C**) moves as a separate rigid body, as also shown by a spectroscopic study (Ramachandran *et al.*, 2004). Notably, the equivalent region has been implicated in the triggering mechanism for unbending in a recent EM study of a remotely related MACPF protein (Lukyanova *et al.*, *in press*).

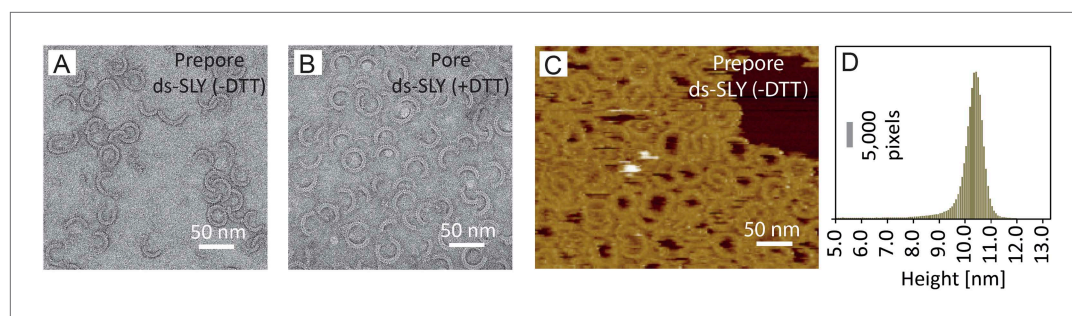
## Real-time visualization of the prepore-to-pore transition and membrane perforation

When imaged by negative-stain EM, both the prepore and pore states appeared in heterogeneous ring- and arc-shaped assemblies (**Figure 2A,B**; Sonnen *et al.*, 2014; Köster *et al.*, 2014). The expansion in ring diameter upon membrane insertion was confirmed by a statistical analysis of the radius of curvature of the arc assemblies (**Figure 2—figure supplement 1**), which also revealed significantly larger variations in arc curvature, that is, larger flexibility, for the prepore than for the pore state.

To facilitate AFM imaging of the disulphide-locked suilysin prepores, the protein was confined to well-defined domains on phase-separated lipid membranes (Connell *et al.*, 2013), showing densely packed suilysin rings and arcs that extended 10–11 nm above the membrane surface (**Figure 2C,D**), consistent with the structural data for the prepore state (**Figure 1A,D**).

At lower packing density on the membrane, suilysin prepores were only resolved when the temperature was lowered to 15°C. Cooling appeared to reduce the prepore mobility such that individual prepore assemblies could be observed while diffusing over the membrane (**Video 1**). When imaged at room temperature, suilysin prepores appeared as streaks in the AFM images—as can be expected for highly mobile proteins—with slightly improved contrast at the lipid phase boundaries (**Figure 3A**). As demonstrated by real-time AFM images of the same area on the membrane, the disulphide-locked suilysin reproducibly converted from the prepore to the pore state upon exposure to DTT (**Figure 3A–E**, **Figure 3—figure supplement 1**): in less than a minute, exposure to DTT triggered the appearance of diffuse rings and arcs that became progressively clearer and more prominent (bottom half of **Figure 3A**), adopting the reduced height typical of the pore state (**Figure 3B–C**). This process was accompanied by a gradual disappearance of the diffuse streaks (i.e., suilysin prepores), while additional high (white) features appeared on the surface. We identify these features as lipid micelles or fragments being ejected from the membrane. This interpretation is supported by the subsequent appearance of larger plateaus that were consistent in height with the collapse of newly formed lipid layers on top of the membrane (**Figure 3D**). After about 20 min, the membrane was cleared of these features, leaving a heterogeneous population of suilysin pore assemblies perforating the membrane





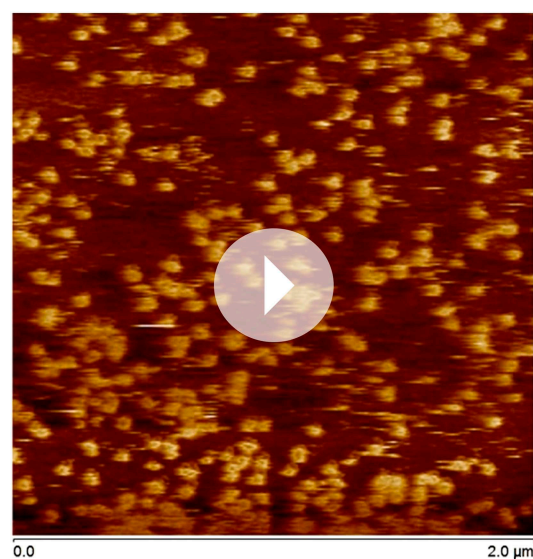
**Figure 2.** Negative-stain EM and AFM of disulphide-locked suilysin. **(A)** Negative-stain EM disulphide-locked suilysin (ds-SLY) on egg PC:cholesterol monolayers (45:55%), locked in the prepore state (–DTT). **(B)** as **(A)**, for disulphide-locked suilysin incubated in the presence of 5 mM DTT in solution to reduce the disulphide bridge, so that the suilysin is rapidly converted to the pore conformation. **(C)** AFM of densely packed suilysin prepores, confined to the egg PC-rich domain of a phase-separated egg PC:DDAB:Cholesterol (33:33:33%) supported lipid bilayer, with its corresponding height distribution **(D)** referenced to the membrane surface.

DOI: [10.7554/eLife.04247.008](https://doi.org/10.7554/eLife.04247.008)

The following figure supplement is available for figure 2:

**Figure supplement 1.** Radius of curvature for arc-shaped suilysin assemblies in the prepore and pore states.

DOI: [10.7554/eLife.04247.009](https://doi.org/10.7554/eLife.04247.009)



**Video 1.** Mobile disulphide-locked suilysin (prepore) assemblies diffusing on the membrane. At a temperature of 15°C, the mobility of disulphide-locked suilysin is sufficiently reduced for the assemblies to be resolved by real-time AFM at 15 s/frame. This sequence of images was captured ~30 min after protein injection and at 384 pixels per line. The timing of the video is accelerated by a factor of ~100. Full z-colour scale = 20 nm.

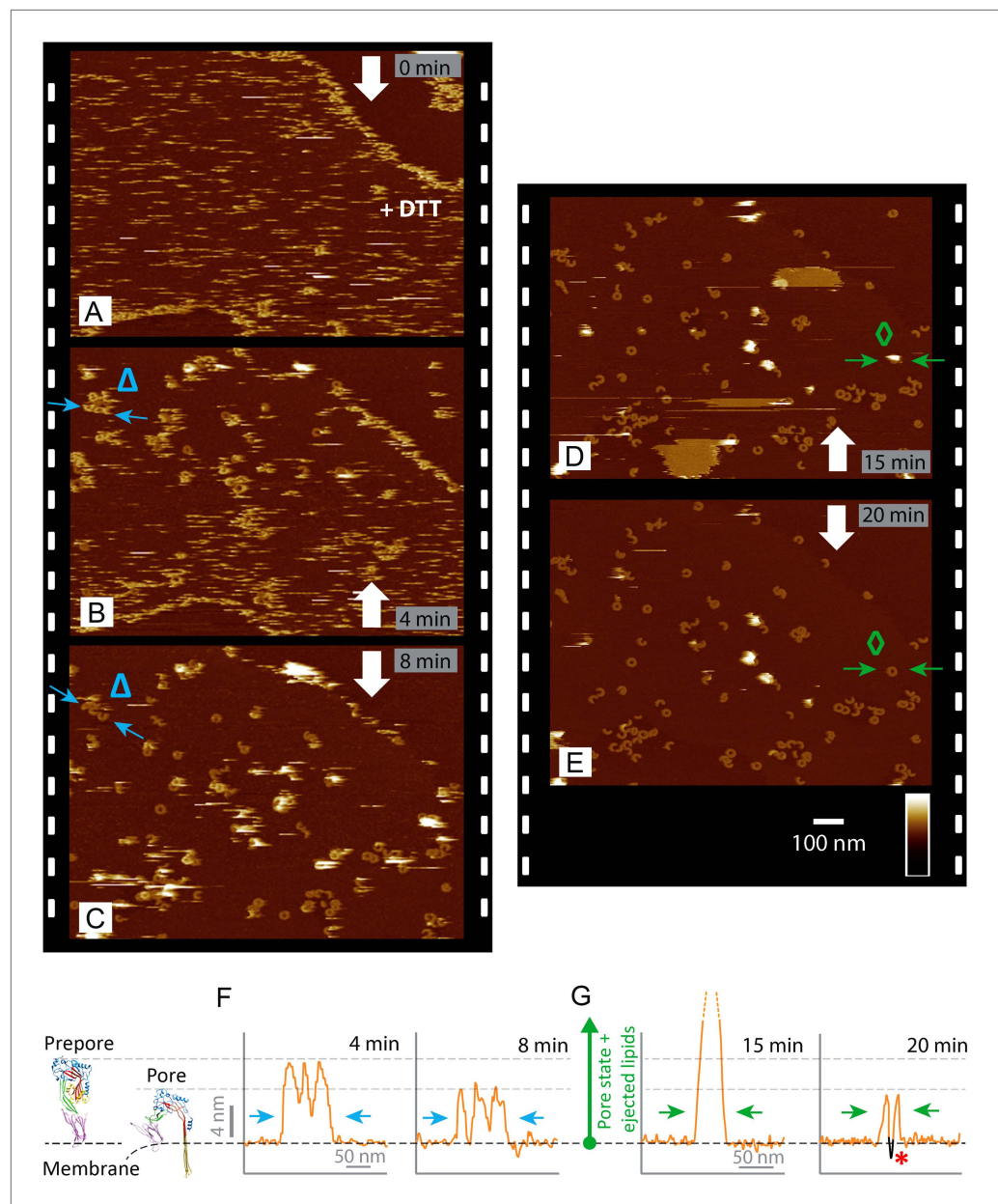
DOI: [10.7554/eLife.04247.010](https://doi.org/10.7554/eLife.04247.010)

(**Figure 3E**). The transition from prepore to pore, as well as the emergence and clearance of lipid aggregates, could also be observed via height profiles taken along various topographic features in these images (**Figure 3F,G**).

### AFM demonstrates that incomplete, arc-shaped assemblies perforate the membrane

The heterogeneity of ring- and arc-shaped assemblies was confirmed for wild-type suilysin by negative-stain EM on egg PC:cholesterol monolayers and by AFM in solution on supported bilayers (**Figure 4A,B**). EM and AFM yielded quantitatively similar arc-length distributions for identical lipid composition, protein concentration, and incubation temperature (**Figure 4—figure supplement 1**). Qualitatively similar behaviour could be observed by negative-stain EM on lipid vesicles (**Figure 4A**, inset). Unlike the disulphide locked construct, wild-type suilysin was converted from its soluble, monomeric state (**Figure 4—figure supplement 2**) to the pore state without the appearance of prepore intermediates, within the time resolution of our AFM experiments. Rings and arcs had a height of 7–8 nm in AFM (**Figure 4B**, inset), in agreement with the cryo-EM data on the pore state (**Figure 1B,F**).

To verify if suilysin rings and arcs perforate the membrane, we used high-aspect ratio AFM tips to probe the membrane in the pore lumen (**Figure 3G**). Both rings (**Figure 4C**) and arcs (**Figure 4D**) enclosed local depressions in the membrane, the depth of which was tip-dependent, but in many cases exceeded the 2.2 nm length of an extended lipid molecule, indicating that the lipid bilayer was locally removed.



**Figure 3.** Real-time imaging of the prepoire-to-pore transition and membrane perforation by suliyisin. Subsequent AFM frames of the same area were alternatively recorded from top to bottom and from bottom to top, as indicated by white arrows. Frame time: 4 min, colour scale: 35 nm. (A) Loosely bound to sphingomyelin-rich domains in the phase-separated lipid mixture (DOPC:sphingomyelin:cholesterol, 33:33:33%), the prepoire intermediates of disulphide-locked suliyisin appear as diffuse streaks. 5 mM of DTT is injected on about 50% completion of the scan. (B) On consecutive scanning, the streaks become more clearly defined as arc-shaped oligomers and complete rings. Towards the top end of the scan, clusters of arc-shaped complexes, mostly in the prepoire intermediate (~10.5 nm high), can be distinguished ( $\Delta$ ). (C) With the scan direction reversed, and the same area scanned again, the cluster of prepoire complexes has converted into the pore state (~7.5 nm high), within ~2 min. The prepoire to pore transition is followed by the ejection of globular features of varying dimensions exceeding 15 nm above the suliyisin in the pore state. We interpret these as ejected lipids. (D) These lipids gradually detach from the surface on the pore state suliyisin assemblies and can be observed as patches of lipids condensing back onto the membrane. The prepoire to pore transition of the suliyisin is now complete. (E) After ~20 min, the surface is almost clear of the ejected lipids. (F) Cross-sectional line profile extracted as indicated ( $\Delta$  in B–C), illustrating the prepoire to pore transition. (G) Cross-sectional line profile extracted as indicated ( $\Delta$  in D–E), illustrating the pore state and the ejection of lipids. *Figure 3. Continued on next page*

Figure 3. Continued

transition. (G) Cross-sectional line profile extracted as indicated ( $\diamond$  in D–E), illustrating the lipid ejection and eventual formation of an aqueous pore in the membrane (\*).

DOI: [10.7554/eLife.04247.011](https://doi.org/10.7554/eLife.04247.011)

The following figure supplement is available for figure 3:

**Figure supplement 1.** Reproducibility of real-time imaging of the prepore-to-pore transition and membrane perforation by sullysin.

DOI: [10.7554/eLife.04247.012](https://doi.org/10.7554/eLife.04247.012)

These depressions were reproducible between the trace and retrace versions of the AFM line scans and were observed for various lengths and orientations of the arcs (Figure 4E). We therefore conclude that sullysin can perforate the membrane irrespective of the completion of the ring assembly, locally removing the lipids to create partial  $\beta$ -barrel pores with an unsealed edge of the lipid bilayer. This is in agreement with recent cryo electron tomography data on pneumolysin assemblies (Sonnen et al., 2014).

Besides isolated rings and arcs, we observed interlocked arcs (Figure 4A,B, arrows), in which one or both ends of the arc contacted another arc or ring. As was the case for the isolated arcs, we found these interlocked arcs capable of perforating the membrane and form membrane lesions that can be smaller, but also larger than those in closed rings (Figure 4F). Once assembled in the pore conformation, the arcs were stable and did not evolve further; even interlocked arcs did not merge into complete rings (Figure 4G).

Interestingly, the prepore-to-pore transition and membrane perforation by the wild-type sullysin was largely prevented, in a dose-dependent manner, by adding the disulphide-locked mutant in the incubation process (Figure 4—figure supplement 3). Subsequent exposure to DTT restored normal pore formation.

## Sullysin assemblies are consistent with kinetically trapped oligomerization

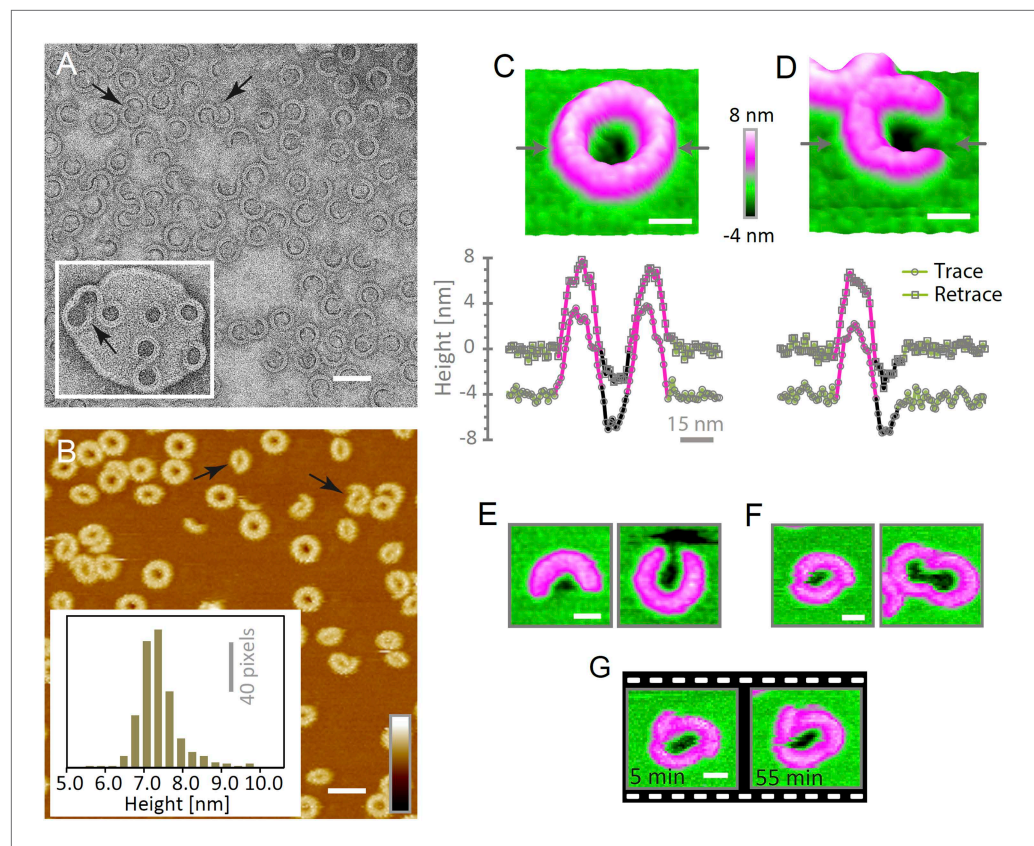
To analyze the oligomerization process, we measured the arc-length distributions of wild-type sullysin in the pore state and of the disulphide-locked mutant in prepore (–DTT) and reduced, pore (+DTT) configurations, based on negative stain EM analysis under the same conditions (Figure 5A–C). The distributions show a broad peak centred between lengths of 10–30 monomers and a smaller, narrow peak corresponding to completed rings of about 37 monomers. The arc-length distributions for disulphide-locked prepores and pores after reduction by DTT are practically identical. Combined with the observation that sullysin does not oligomerize before binding to the cholesterol-containing membrane (Figure 4—figure supplement 2), this demonstrates that assembly is completely determined and terminated in the prepore state, i.e., is not affected by the prepore-to-pore transition. This conclusion is further confirmed by the lack of growth of individual arcs in the pore state upon subsequent, further addition of wild-type sullysin (Figure 5—figure supplement 1), and greatly simplifies the interpretation of the arc-length distributions.

To explain these distributions, we calculated the oligomeric populations for a model in which monomers from the solution irreversibly bind to the membrane with a rate constant  $k_b$ , and in which irreversible oligomerization on the membrane occurs only by monomer addition, with a rate constant  $k_a$  (Figure 5D). In such a simple model, the oligomerization reaction is arrested by depletion of monomers, yielding kinetically trapped assembly intermediates. Since the resulting populations only depend on the ratio  $k_a/k_b$  and on the (experimentally known) total number of monomers per unit area of membrane ( $C$ ), this model can be used to fit the experimental data with a single free parameter ( $k_a/k_b$ ). The broad peak for intermediate arc-lengths (Figure 5A–C and Figure 4—figure supplement 1) can thus be explained by a ratio  $k_a/k_b$  that is sufficiently large to ensure a steady supply of monomers to sustain the oligomerization reaction, but not large enough to yield only completed assemblies (i.e., rings).

## Discussion

CDCs are protein toxins that are potent virulence factors in bacteria. They are part of the major MACPF/CDC superfamily of pore-forming proteins. Our data map the structure (Figure 1) and assembly pathways of membrane pore formation by the CDCs in unprecedented detail and accuracy, as summarized in Figure 6. Using the disulphide-locked sullysin variant, we have resolved the





**Figure 4.** Suilysin assembles into ring- and arc-shaped oligomers that perforate the membrane. **(A)** Negatively stained EM of arc- and ring-shaped assemblies of wild-type suilysin on an egg PC:cholesterol (45:55%) lipid monolayer, and (inset) on a liposome of egg PC:cholesterol (45:55%). **(B)** AFM topography of wild-type suilysin on a supported egg PC:cholesterol (67:33%) lipid bilayer. The wild-type suilysin extends 7–8 nm above the lipid bilayer background, as indicated by the height histogram for 402 individual particles (inset). **(C)** The AFM topography of a complete suilysin ring reveals a circular hole (dark) in its lumen, whereas the lipid bilayer surrounding the ring remains intact (green). **(D)** The topography of a suilysin arc shows a hole (dark) in the membrane only partially enclosed by the suilysin assembly. Images in **C** and **D** are shown in a 15° tilted representation, and height profiles measured across the ring/arc confirm membrane perforation. **(E)** Examples of wild-type suilysin arcs of different lengths. Transmembrane holes are consistently observed. **(F)** Examples of interlocked-arc assemblies. As shown in the right image, the membrane area removed by the two arcs is larger than the hole in the complete ring (**C**). **(G)** Sequence of AFM images of the same interlocked-arc assembly, stable for at least 50 min. Scale bars **A–B**: 50 nm, **C–G**: 15 nm, full z colour scale **B–G**: 12 nm.

DOI: [10.7554/eLife.04247.013](https://doi.org/10.7554/eLife.04247.013)

The following figure supplements are available for figure 4:

**Figure supplement 1.** Suilysin pore assemblies by EM and AFM.

DOI: [10.7554/eLife.04247.014](https://doi.org/10.7554/eLife.04247.014)

**Figure supplement 2.** Suilysin is a monomer in solution.

DOI: [10.7554/eLife.04247.015](https://doi.org/10.7554/eLife.04247.015)

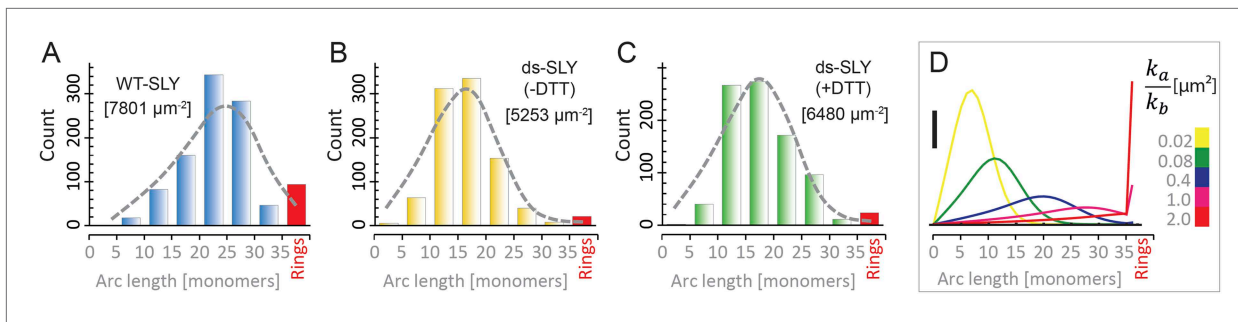
**Figure supplement 3.** AFM assays of wild-type suilysin (WT-SLY) doped with disulphide-locked suilysin (ds-SLY).

DOI: [10.7554/eLife.04247.016](https://doi.org/10.7554/eLife.04247.016)

initial membrane binding of suilysin monomers (**Figure 4—figure supplement 2**), oligomerization (**Figure 5**), and membrane insertion stages (**Figures 2–4**) in pore formation.

Suilysin oligomers exhibit a broad distribution of arc- and ring-shaped assemblies (**Figures 4 and 5**). Under given conditions such as incubation temperature, lipid composition, protein sequence and concentration, these distributions are reproducible between EM and AFM experiments (**Figure 4—figure supplement 1**). The similarity between prepore and pore distributions (**Figure 5B,C**) implies that oligomerization is arrested before pore insertion. Therefore, our work rigorously establishes that the





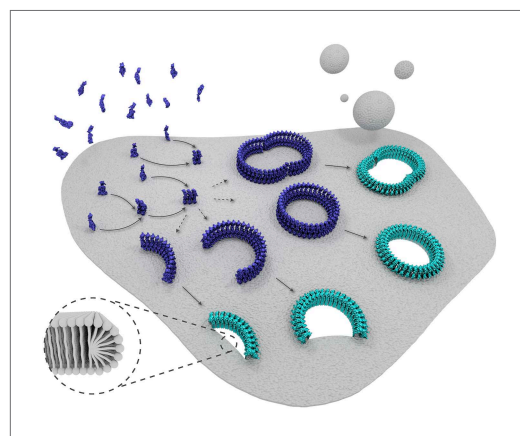
**Figure 5.** Oligomerization states for arc- and ring-shaped assemblies of sulysin. **(A)** The arc-length distribution of wild-type sulysin displays a broad peak for arcs that contain between 15 and 30 monomers, and a smaller, sharp peak for complete rings (37-mers). **(B)** Arc-length distribution for the disulphide-locked sulysin prepore intermediate. **(C)** For the disulphide-locked mutant incubated in the presence of DTT (pore-state), the arc-length distribution is practically identical to the distribution for the prepore-locked intermediate. **(D)** Calculated arc-length distributions for a simple model of kinetically trapped oligomerization, with  $C = 2000$  monomers per square micron (see ‘Materials and methods’). The peak of the arc-length distribution shifts from smaller to larger oligomers on increasing the ratio between the rate constants for monomer association ( $k_a$ ) and monomer binding to the membrane ( $k_b$ ). Vertical scale bar: 40 counts. Grey, dashed lines in **A–C** denote fits of the experimental data with the oligomerization model, yielding  $k_a/k_b = 0.893 \pm 0.008 \mu\text{m}^2$  **(A)**;  $0.438 \pm 0.012 \mu\text{m}^2$  **(B)**;  $0.425 \pm 0.012 \mu\text{m}^2$  **(C)**. Numbers in brackets in **A–C** indicate the estimated total number of monomers per square micron. The experimental data here are based on negative-stain EM images on monolayers of egg PC:cholesterol (45:55%), incubated at 37°C.

DOI: [10.7554/eLife.04247.017](https://doi.org/10.7554/eLife.04247.017)

The following figure supplements are available for figure 5:

**Figure supplement 1.** Sequential addition of wild-type sulysin in the pore state.

DOI: [10.7554/eLife.04247.018](https://doi.org/10.7554/eLife.04247.018)



**Figure 6.** Schematic representation of sulysin membrane binding, assembly, and pore formation. From left to right: monomers bind to the membrane and oligomerize. The assembly of monomers proceeds in the prepore intermediate and results in either complete rings or kinetically trapped arc-shaped oligomers. The arc- and ring-shaped assemblies subsequently collapse to the pore configuration with the transmembrane  $\beta$ -hairpins unfurled and inserted into the lipid bilayer in a concerted conformational change. Lipids are subsequently ejected from the membrane (shown as grey spheres) and aqueous pores of different sizes are formed in the membrane. The inset shows a possible configuration of lipids at the unsealed edges of the bilayer.

DOI: [10.7554/eLife.04247.019](https://doi.org/10.7554/eLife.04247.019)

whole CDC assembly takes place in the prepore state, as was previously suggested for perfringolysin (*Hotze et al., 2001; Hotze and Tweten, 2012*).

As can be deduced from the prevalence of larger but incomplete assemblies of the disulphide-locked sulysin, the association of such larger oligomers ( $\gtrsim 5$  subunits) is not a determining factor in membrane pore formation by CDCs. On the contrary, the oligomerization appears to be dominated by monomer addition (*Figure 5*), although the addition of smaller oligomers ( $\lesssim 5$  subunits) cannot be fully excluded. As demonstrated by our oligomerization model, the measured arc-length distributions are consistent with the kinetically trapped product of a two-stage irreversible and noncooperative reaction. The outcome is largely determined by the ratio of the corresponding rate constants and the density of monomers per unit area of membrane (in addition to the effects of steric hindrance at higher surface densities). The first stage can be the binding of monomers to the membrane, as assumed here, or a rate-limiting nucleation step that triggers the oligomerization reaction, as assumed elsewhere (*Hotze et al., 2001*). The second stage of this reaction is oligomerization by addition of monomers or very small oligomers. As the rate constants can be expected to vary from one protein to another, this model implies that the various CDCs can yield differing oligomeric populations.

Suilysin prepore intermediates appear both more mobile (**Figure 3A–E**) and more flexible than pore assemblies (**Figure 2—figure supplement 1**). This is consistent with the observation that the  $\beta$ -sandwich membrane-binding domain 4 does not significantly penetrate the membrane (**Nakamura et al., 1995; Ramachandran et al., 2002**). In contrast to the earlier observations on pneumolysin (**Tilley et al., 2005**), prepore assembly in suilysin is accompanied by some opening of the central  $\beta$ -sheet. A possible explanation for this difference is that the wild-type pneumolysin preparation in that study formed many stable prepores, apparently in an inactive, dead-end state, whereas wild-type suilysin is extremely active and is not observed in a prepore state.

The activity of wild-type suilysin was greatly impaired, however, by incubating it in the presence of an equal concentration of disulphide-locked suilysin (**Figure 4—figure supplement 3**) and recovered on unlocking the mutant. These results imply that wild-type and mutant co-assemble as expected, and that the prepore-to-pore transition requires a concerted conformational change of all subunits in the suilysin assemblies, suggesting a cooperative insertion of subunits into the membrane.

The mobility of prepore assemblies on the membrane surface makes it possible for the subunits to slide apart upon pore formation, as seen in the 8% diameter expansion. This size difference was seen in the earlier work, but the symmetry measurement was less clear, and it was explained by assigning a lower symmetry to the prepore (31 vs 38 for the pore) (**Tilley et al., 2005**). In view of the present observations, it seems likely that pneumolysin rings also expand and that the previous assignment of different symmetries to pneumolysin pores and prepores was most likely incorrect.

We observed arc-shaped assemblies as small as 5 subunits, with heights corresponding to the suilysin pore state. This gives an estimate of the minimum oligomer size required for membrane insertion. The size of the AFM tip was too large to probe the membrane perforation in arc-shaped complexes smaller than about 15 monomers (see e.g., **Figure 4E**). We have observed membrane perforation for arcs at any size between 15 and 37 monomers, which unambiguously demonstrates that a fully enclosed  $\beta$ -barrel is not essential for the insertion of the transmembrane  $\beta$ -hairpins and membrane perforation, as was previously presumed (**Hotze and Tweten, 2012**). Unsealed lipid edges and incomplete  $\beta$ -barrels are a surprising by-product of this membrane perforation by incomplete CDC oligomers (**Figure 4D,E**), as was recently suggested based on electron tomography of various pneumolysin assemblies (**Sonnen et al., 2014**).

The size of the membrane lesions can thus vary between less than half the lumen in a completed SLY pore to the larger pores formed by interlocked arcs (**Figure 4F,G**), as suggested by conductance measurements on black lipid membranes (**Marchiorretto et al., 2013**).

Regardless of the extent of oligomerization, both arc and ring prepores can convert directly to the pore configuration as the  $\beta$ -hairpins unfurl and insert into the lipid bilayer. This is followed by the ejection of lipids from the membrane as the pore is formed (**Figure 3C,D** and **Figure 3—figure supplement 1C,D**). The results imply that the hydrophilic inner surface of the partially or fully completed  $\beta$ -barrel is sufficient to destabilize the lipid membrane in the pore lumen, leading to ejection of lipid micelles from the pore.

The approach of identifying hinge regions for domain fitting to the suilysin pore map has yielded a pseudo-atomic model that goes beyond the earlier rigid body fitting to pneumolysin (**Tilley et al., 2005**). The mechanism of collapse through buckling of domain 2 involves a sideways movement around the ring, and the bending is in the opposite direction to that proposed in the earlier model. The resulting tilt of domain 1 results in the 8% radial expansion of the ring, clearly seen by the displacement of domain 4 (**Figure 1E**). A helical subdomain thought to be involved in triggering of sheet opening in a MACPF protein (dashed oval in domain 3, **Figure 1C; Lukoyanova et al., in press**) is also likely to move in suilysin, strengthening the notion that the mechanism of unfolding and pore formation is conserved between the remotely related MACPF and CDC subfamilies.

In summary, we have visualized the various stages of membrane pore formation by a CDC at greatly improved spatial and temporal resolution, to provide new insights in domain movements and pathways of assembly for a major superfamily of pore-forming proteins.

## Materials and methods

### SLY expression and purification

#### Cloning of wild-type suilysin gene

Genomic DNA was extracted from *S. suis* (kindly provided by Dr Vanessa Terra, London School of Hygiene and Tropical Medicine; UK), using the PureLink genomic DNA mini kit (Invitrogen, Carlsbad, CA).

The suilysin gene was amplified using the primers: SLY\_F, 5'-CGG CGC CAT GGC TTC CAA ACA AGA TAT TAA TCA GTA TTT TCA AAG -3' and SLY\_R, 5'-GAT AGG ATC CTC ACT CTA TCA CCT CAT CCG CAT ACT GTG-3'. These primers introduced restriction sites for NcoI and BamHI for subsequent cloning into pEHISTEV ([Liu and Naismith, 2009](#)), in-frame with and downstream of nucleotides encoding a 6-histidine tag and a TEV protease cleavage site. The recombinant plasmid expressing suilysin was then transformed into *Escherichia coli* XL-10 Gold competent cells (Agilent, Santa Clara, CA), according to the manufacturer's instructions. Positive clones were identified by blue-white colony screening and were confirmed by sequencing. Plasmid carrying the suilysin gene was subsequently transformed into *E. coli* Rosetta-2 (DE3) (Novagen, Millipore, Watford, United Kingdom).

### Creation of a disulphide-locked construct of suilysin

This mutant was made using a QuikChange Multi Site-Directed Mutagenesis Kit (Agilent). The primers G52C\_F, 5'-CAC AAG AGA TTC TTA CAA ATG AGT GCG AAT ACA TTG ATA ATC CGC CAG C-3' and G52C\_R, 5'-GCT GGC GGA TTA TCA ATG TAT TCG CAC TCA TTT GTA AGA ATC TCT TGT G-3' were used to replace the codon for Gly52 (GGA) in the wild-type with a Cys codon (TGC). The primers S187C\_F, 5'-TGA AAC AAT GGC ATA CAG TAT GTG CCA ATT GAA AAC GAA GTT CGG AAC-3' and S187C\_R, 5'-GTT CCG AAC TTC GTT TTC AAT TGG CAC ATA CTG TAT GCC ATT GTT TCA-3' were used to substitute the Ser187 codon (TCA) by a Cys codon (TGC) in the same construct. Following the digestion of the methylated/hemimethylated parental DNA with DpnI, the construct containing the double mutation was transformed into *E. coli* XL-10 Gold (Agilent). Positive clones were identified by blue-white colony screening and mutations confirmed by sequencing. Purified plasmid from selected clones was subsequently transformed into *E. coli* Rosetta-2 (DE3) (Novagen).

### Expression and purification

Recombinant wild-type suilysin and the cys-locked version were expressed in Overnight Express™ Instant TB medium (Novagen) containing kanamycin (50 µg/ml; Sigma-Aldrich, Dorset, United Kingdom). Cells were harvested by centrifugation and lysed with 1x BugBuster protein extraction reagent (Novagen) supplemented with 10 µg/ml DNase I (Sigma-Aldrich), 5 mM MgCl<sub>2</sub>, and EDTA-free protease inhibitor (Roche Applied Science, Welwyn Garden City, United Kingdom). Soluble cellular extracts were clarified and loaded onto a HisTrap High Performance column (GE Healthcare, Little Chalfont, United Kingdom) in loading buffer (20 mM Tris-HCl, 150 mM NaCl, 20 mM Imidazole, pH 7.5). The column was washed thoroughly with 20 mM Tris-HCl, 150 mM NaCl, 50 mM Imidazole, pH 7.5, and the 6-histidine-tagged wild-type and cys-locked suilysins were eluted using a stepwise gradient of elution buffer (20 mM Tris-HCl, 150 mM NaCl, containing 0 to 500 mM Imidazole, pH 7.5). The purity of the eluted fractions was assessed by SDS-PAGE ([Laemmli, 1970](#)). The haemolytic activity of the suilysins was determined as described previously ([Owen et al., 1994](#)), except that 5 mM DTT was included in some assays as appropriate. Wild-type suilysin had a specific activity of 39,000 HU (haemolysis units)/mg protein. No haemolytic activity was seen with the cys-locked construct, in the absence of DTT, but in 5 mM DTT, its specific activity was 20,000 HU/mg protein.

### Lipid and liposome preparation

All lipid materials (cholesterol, egg PC, DOPC, sphingomyelin, DDAB) were purchased from Avanti Polar Lipids (Alabaster, AL). Small unilamellar lipid vesicles were prepared by the extrusion method ([Hope et al., 1985](#)). Briefly, lipids in powdered form were weighed and dissolved in chloroform to produce a homogeneous mixture with a lipid concentration of ~1 mg/ml. The solvent was then slowly evaporated for at least 5 hr by passing a steady stream of argon in a fume hood, yielding a dry lipid film. The lipid film was resuspended, by vigorous vortexing for 5 min, in 1 ml of 20 mM Tris, 150 mM NaCl, pH 7.8, to form large, multilamellar vesicles. This solution was transferred to a Fisherbrand FB11201 bath sonicator (Fisher Scientific, Loughborough, UK), maintained above the gel-liquid transition temperature of the constituent lipids. The large multilamellar vesicles were disrupted by 15-min sonication treatments at frequencies between 40 and 80 kHz, interspersed by two freeze/thaw cycles. The solution containing the lipid dispersion was loaded into an Avanti mini-extruder kit (Avanti Polar Lipids) and kept above the transition temperature of the lipids. The lipid solution was forced through a Whatman Nucleopore polycarbonate filter (GE Healthcare Lifesciences, Buckinghamshire, UK) with an 80 nm nominal pore diameter. The extrusion process was repeated at least 30 times to yield small unilamellar vesicles with a diameter near the pore size of the filter used, as verified by negative stain EM.

For EM experiments, liposomes were prepared from 5 mM lipids containing ~45 mol% of egg PC and ~55 mol% cholesterol resuspended in 100 mM NaCl, 50 mM HEPES, pH 7.5 by extrusion through an 80-nm filter as previously described (Tilley *et al.*, 2005).

## Electron microscopy sample preparation and data acquisition

### Negative stain

10 µg/ml monomeric wild-type sulysin was negatively stained with 2% wt/vol uranyl acetate. To form prepore and pore complexes on lipid monolayers, a solution of monomeric wild-type or disulphide-locked sulysin (10 µg/ml), or disulphide-locked sulysin reduced by pre-incubation with 10 mM DTT for 10 min, was overlaid with 1 µl of chloroform solution of the lipid mixture described above, at 1 mg/ml, for 25 min at 37°C and the monolayers were transferred to EM grids, as described before (Dang *et al.*, 2005). To image pores on liposomes, 1 µl of a 0.3–0.5 mg/ml solution of wild-type sulysin was incubated with 1 µl of liposomes for 10 min at 37°C. Samples were negatively stained with 2% wt/vol uranyl acetate and imaged on a Tecnai F20 FEG microscope (FEI, Hillsboro, OR) at 200 kV under low dose conditions. Images were taken with a defocus of 0.5 µm on a Gatan 4k × 4k CCD camera giving a final pixel size of 1.85 Å.

### Cryo-EM

For 3D reconstructions of the prepore and the pore, 1 µl of 0.3–0.5 mg/ml solution of either wild-type or disulphide-locked sulysin was incubated with 1 µl of liposomes for 10 min at 37°C. Liposomes were then applied to lacey carbon-coated copper grids (Agar Scientific, Stansted, United Kingdom) and frozen using Vitrobot Mk3 (FEI) at 22°C and 100% humidity. Images were collected on a Tecnai G2 Polara microscope (FEI) at 300 kV, on a Gatan 4k × 4k CCD camera giving a final pixel size of 2 Å, at an electron dose of 20–25 e/Å<sup>2</sup>.

## Image processing of sulysin prepores and pores on lipid monolayers

Ring images were centered and analysed by multivariate statistical analysis (MSA; van Heel, 1984) for classification into subsets of homogeneous diameter and subunit number. Sulysin arc length distributions were determined from negative-stain EM images at 1.85 Å pixel size and AFM images acquired at 26.8 Å pixel size. Using DNA Trace software (Mikhaylov *et al.*, 2013), individual sulysin arcs were manually traced with a step size of 25 Å for both EM and AFM images. The number of monomers within each arc was then calculated by dividing the manually traced arc length by the average size of a monomer in the prepore and pore states, 23.6 Å and 25.7 Å, respectively, as estimated from rotationally averaged negative-stain EM images. This approach yielded an error within ±2 monomers as estimated from averages of rings from the EM monolayer data.

## 3D reconstruction of sulysin prepores and pores on liposomes

The defocus of the cryo-EM images was determined by CTFFIND3 (Mindell and Grigorieff, 2003) and phases were corrected using SPIDER (Frank *et al.*, 1996). Side-view images of prepores (1374) and pores (2700) were extracted using Boxer (EMAN 1.9; Ludtke *et al.*, 1999). Images were aligned in SPIDER to reprojections of pneumolysin prepore and pore maps (Tilley *et al.*, 2005) and the aligned images were sorted by diameter with MSA. Initial reconstructions were calculated by back-projection of either of class sums or aligned raw images, up to 35° from the side view plane, and symmetry estimated by maximising density variance within the maps. These estimates were consistent with the outcomes of statistical analysis for negatively stained prepores and pores formed on lipid monolayers. Most of the pores (~60%) and prepores exhibited 37-fold symmetry. These 37-fold maps were further refined by projection matching with up to 20° out-of-plane tilt. MSA was used to detect and correct for misalignments. Reconstructions were calculated by back-projection in SPIDER. 450 prepore and 600 pore views were selected for the final reconstructions. The final resolution was estimated by 0.5 FSC (Figure 1—figure supplement 2).

## Atomic structure modelling

### Pore map fitting

First, nine different β-barrel models (corresponding to domain 3, residue range 176–225 and 272–346) with architecture  $S = n/2$  were generated as described in Reboul *et al.*, 2012. These nine different β-barrel models were generated with slightly varying  $a$  ( $3.48 \pm 0.1$  Å) and  $b$  ( $4.83 \pm 0.1$  Å) bond lengths, where the value of  $a$  is the distance between Cα of adjacent residues in the same β-strand and

$b$  is the distance between C $\alpha$  of adjacent residues in adjacent  $\beta$ -strands. This resulted in  $\beta$ -barrel models with modest variations of radius and height, in which the  $\beta$ -strands are tilted by 20° from the pore axis (Murzin *et al.*, 1994). All the  $\beta$ -barrels were fitted using the Fit-in-Map tool in Chimera (Pettersen *et al.*, 2004; Goddard *et al.*, 2007). Among the top three best-fitting  $\beta$ -barrels (CCC [cross-correlation coefficient] scores 0.43–0.44, as compared to 0.28–0.40 for the rest), the barrel with the height best matching the membrane was chosen by visual inspection. Next, the missing residues in the N-terminus of the native suliyisin structure (PDB:3HVN) were modelled using MODELLER (Sali and Blundell, 1993). Normal Mode Analysis (NMA) was used to generate rough decoys for domains 1 and 2 using the NOMAD-Ref web server (Lindahl *et al.*, 2006). 50 different decoys were obtained by randomly combining amplitudes of the first 20 modes. The value of the average coordinate root mean square deviation between the native domains and the decoys was set to 5 Å.

The decoy models and the crystal structures of the individual domains were manually fitted as rigid bodies into the pore map. The best fitting model for each domain was selected by a combination of local fit quality and geometric constraints using Chimera. For domains 1 (residue 32–48, 85–175, 226–271 and 347–370) and 2 (residues 49–84 and 371–387), the best fitting models were selected based on CCC from the decoy set. For domain 4 (residues 388–497), the crystal structure was used as the best fitting model (this domain shows very little flexibility based on NMA analysis). The geometric constraints were such that domain 2 is connected to domains 1 and 4, domain 1 is connected to the  $\beta$ -barrel, and domain 4 sits at the membrane surface.

For each domain, 50 models were generated with MODELLER using the above corresponding best fit as a template structure to refine the stereochemistry. Then the models were evaluated using the DOPE statistical potential score (Shen and Sali, 2006). The top models from the individual domains were connected with MODELLER into one partial model (containing domains 1, 2, and 4—but not 3), which was then C37 symmetrized in Chimera. Next, the map was segmented around three asymmetric units of the resulting pore model. Loop refinement was performed on the loops connecting domain 1 and the corresponding strands from the barrel (only on the central asymmetric unit). The resolution was insufficient to include  $\beta$ 5 from domain 3 in the model. The refined asymmetric unit was C37 symmetrized to give the final pore model.

### Prepore map fitting

Only domains 1 and 4 were fitted into the prepore map. The starting structure was the crystal structure of suliyisin monomer. We first rigidly fitted the whole crystal structure and then deleted domain 2 and 3, as their corresponding density was not sufficiently resolved. The fit of domain 4 was further refined to improve the CCC, taking into account the position relative to the membrane. The final fits of domain 1 and 4 were C37 symmetrized to give a final (partial) prepore model without clashes between the monomers.

### Mapping of electrostatic potential and interacting residues

We calculated the electrostatic potential of domain 1 in both prepore and pore models using the APBS method (Baker *et al.*, 2001; available in Chimera) and mapped it onto their molecular surface (Figure 1—figure supplement 3A,B). For the pore model, optimization of the side chain rotamers was done using SCWRL (Krivov *et al.*, 2009) prior to the calculation of the electrostatic potential. We further analysed the contacts in adjacent monomers of domain 1 in the prepore and pore fit (Figure 1—figure supplement 3C,D). We considered two residues as interacting (interface residue) if their corresponding C $\beta$  atoms are within a distance of 7 Å (Malhotra *et al.*, 2014).

### AFM sample preparation

Small unilamellar vesicles were injected onto a freshly cleaved mica surface at a concentration between 5 and 25 nM in the presence of 60  $\mu$ l of 20 mM Tris, 150 mM NaCl, 20 mM MgCl<sub>2</sub>, pH 7.8. Incubation of the vesicles on the mica for 30 min at room temperature allowed them to rupture and adsorb onto the surface, yielding an extended lipid bilayer film. Any remaining vesicles were removed by gently rinsing with 80  $\mu$ l of the adsorption buffer. The rinsing process was repeated 3–7 times to ensure a clean and uniform surface conducive for AFM imaging. Wild-type and disulphide-locked suliyisin were injected into a 150  $\mu$ l fluid cell containing the supported lipid bilayers and allowed to equilibrate for ~10 min prior to imaging. The concentration of suliyisin in the various AFM experiments was 12–180 nM.

For the doping assays (Figure 4—figure supplement 1), wild-type and locked suliyisin were mixed in the desired molar ratios and 60 nM of the protein mixture was incubated on the lipid bilayers for 10 min.



## AFM imaging and data processing

Real-time topographic images of suilysin on the supported lipid bilayers were collected on a Multimode 8 system (Bruker, Santa Barbara, CA) by performing rapid force-distance (PeakForce Tapping) curves. The PeakForce method continuously records force-distance curves with a user-defined force set-point (here about 50 pN) that is referenced to a continuously adjusted baseline. Typically, these force-distance curves were recorded at a frequency of 2 kHz with a maximum tip-sample separation between 5 and 20 nm. The topographic features were verified for consistency between trace and retrace images, as well as for their reproducibility in subsequent scan frames. For imaging, the vertical scan limit was reduced to ~1.5  $\mu\text{m}$ . Typically, images were recorded at 0.2–1 frames/min. Suilysin prepores and pores were also imaged at rates of up to 10 frames/min using a home-built AFM system and miniaturized cantilevers (Leung *et al.*, 2012), but this did not yield information additional to the data presented here. Suilysin prepores were only resolved at high concentration on the membrane (Figure 2C), or when the temperature was lowered to 15°C (Video 1). The real-time, low-temperature measurements were carried out on a Dimension FastScan AFM system in tapping mode with images acquired at 4 frames/min using FastScan Dx probes (Bruker). The AFM probes used for suilysin imaging had nominal spring constants ranging from 0.1 to 0.7 N/m and resonance frequencies between 10 and 130 kHz in liquid. We used silicon nitride AFM probes with batch-processed silicon tips including MSNL E and F (Bruker), ScanAsyst Fluid+ (Bruker), and cantilevers with individually grown carbon tips, for example, Biotool (Nanotools, Munich, Germany). Batches of AFM probes were screened for tip sharpness and appropriate tilt angles prior to data collection.

All AFM imaging was performed in the presence of 20 mM Tris, 150 mM NaCl, 20 mM  $\text{MgCl}_2$ , pH 7.8 with either an E or a J scanner with an integrated temperature control. Images were analysed by either the Nanoscope Analysis software package (Bruker) or using the open-source SPM analysis software, Gwyddion ([www.sourceforge.net](http://www.sourceforge.net)). The raw AFM images were plane-levelled and subsequently line-by-line flattened using the lipid membrane as reference. A Gaussian filter with a full-width-half-maximum of 2-pixels was applied to smooth out high frequency noise where necessary.

## Oligomerization model

The assembly of suilysin (SLY) in the prepore state was described by the irreversible reactions  $\text{SLY}_{n-1}^{(\text{pre})} + \text{SLY}_1^{(\text{pre})} \rightarrow \text{SLY}_n^{(\text{pre})}$  for oligomerization via monomer-association with a rate constant  $k_a$ . Here  $n$  denotes the number of monomeric subunits in an oligomer, ranging from 1 to the maximum number of monomers in a complete ring,  $N = 37$ . The prepore monomers originated from the binding of soluble suilysin monomers to the membrane,  $\text{SLY}_1^{(\text{sol})} \rightarrow \text{SLY}_1^{(\text{pre})}$ , here assumed to occur with a rate constant  $k_b$ .

$\sigma_n(t)$  was defined as the number of suilysin prepore  $n$ -mers per unit area on the membrane, and  $C$  as the number of monomers in solution above a unit membrane area, immediately after injection of suilysin at time  $t = 0$ . With these definitions, the oligomerization reactions can be modelled by the rate equations

$$\begin{aligned} \frac{d\sigma_n(t)}{dt} = & \delta_{n,1} k_b C e^{-k_b t} + \sum_{m=1}^{n-1} \frac{1}{2} k_a (\delta_{m,1} + \delta_{n-m,1}) \sigma_m(t) \sigma_{n-m}(t) \\ & - \sum_{m=1}^{N-n} k_a (\delta_{m,1} + \delta_{n,1}) \sigma_m(t) \sigma_n(t). \end{aligned}$$

These reactions lead to kinetically trapped prepore assemblies on depletion of free monomers on the membrane, that is, when  $\sigma_1(t) \rightarrow 0$  as time evolves. With the substitutions  $t = \tau/k_b$  and  $\sigma_n(t) = s_n(\tau)k_b/k_a$ , the rate equations can be rewritten in terms of a dimensionless surface density  $s_n(\tau)$  and a dimensionless time  $\tau$ , to demonstrate that the shape of the solution for  $s_n(\tau \rightarrow \infty)$  versus  $n$ , and thus of the resulting arc length distribution ( $\sigma_n(t \rightarrow \infty)$  versus  $n$ ), is a function of the parameter  $Ck_a/k_b$  only.

The coupled and nonlinear differential equations for  $s_n(\tau)$  were integrated numerically for different  $Ck_a/k_b$  using the Runge-Kutta method, until a stationary solution was reached. For fitting experimental data,  $C$  was determined from the accumulated length of all measured oligomers, normalized to the measured membrane area. The best  $k_a/k_b$  then followed from the numerical solution that yielded the lowest sum of squared residues.

# Acknowledgements

This work has been funded by the BBSRC (BB/G011729/1, BB/J005932/1, BB/J006254 and BB/K01692X/1), the ERC (advanced grant 294408), and the Leverhulme Trust (RPG-2012-519). We thank Richard Thorogate for technical support, Dan Clare, Elena Orlova, and Luchun Wang for help with EM and image processing, Dave Houldershaw and Alan Lowe for computing support, Chanmin Su and Khaled Kaja (Bruker) for AFM probes and access to equipment, and James Whisstock for comments on the manuscript.

Accession codes for EM maps: prepore EMD-12698, pore EMD-12711.

# Additional information

## Funding

Funder	Grant reference number	Author
Biotechnology and Biological Sciences Research Council	BB/G011729/1, BB/J005932/1, BB/J006254 and BB/K01692X/1	Helen R Saibil, Bart W Hoogenboom
European Research Council	294408	Helen R Saibil
Leverhulme Trust	RPG-2012-519	Maya Topf

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## Author contributions

CL, NVD, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; NL, BWH, Conception and design, Analysis and interpretation of data, Drafting or revising the article; AWH, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; IF, APP, CFR, MAD, Development of atomic models, Analysis and interpretation of data, Drafting or revising the article; NJ, MPD, Expression, Purification, Testing of wild-type and mutant proteins, Drafting or revising the article, Contributed unpublished essential data or reagents; DO, Development of oligomerization model, Analysis and interpretation of data, Drafting or revising the article; PWA, RL, Conception and design, Drafting or revising the article, Contributed unpublished essential data or reagents; MT, Analysis and interpretation of data, Drafting or revising the article; HRS, Conception and design, Analysis and interpretation of data, Drafting or revising the article, Contributed unpublished essential data or reagents

# Additional files

## Major datasets

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Dudkina NV, Leung C, Lukoyanova N, Hodel AW, Farabella I, Pandurangan AP, Jahan N, Damaso MP, Osmanovic D, Reboul C, Dunstone MA, Topf M, Andrew PW, Lonnen R, Saibil HR, Hoogenboom BW	2014	Cryo EM structure of suilysin prepore	EMD-12698	Publicly available at EMDDataBank.
Dudkina NV, Leung C, Lukoyanova N, Hodel AW, Farabella I, Pandurangan AP, Jahan N, Damaso MP, Osmanovic D, Reboul C, Dunstone MA, Topf M, Andrew PW, Lonnen R, Saibil HR, Hoogenboom BW	2014	Cryo EM structure of suilysin pore	EMD-12711	Publicly available at EMDDataBank.

The following previously published dataset was used:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Xu L, Huang B, Du H, Zhang CX, Xu J, Li X, Rao Z	2010	Crystal structure of cytotoxin protein suilyisin from <i>Streptococcus suis</i>	<a href="http://www.pdb.org/pdb/explore/explore.do?structureId=3hvn">http://www.pdb.org/pdb/explore/explore.do?structureId=3hvn</a>	Publicly available at RCSB Protein Data Bank.

## References

- Baker NA**, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of USA* **18**:10037–10041. doi: [10.1073/pnas.181342398](https://doi.org/10.1073/pnas.181342398).
- Bourdeau RW**, Malito E, Chenal A, Bishop BL, Musch MW, Villereal ML, Chang EB, Mosser EM, Rest RF, Tang WJ. 2009. Cellular functions and X-ray structure of anthrolysin O, a cholesterol-dependent cytolysin secreted by *Bacillus anthracis*. *The Journal of Biological Chemistry* **284**:14645–14656. doi: [10.1074/jbc.M807631200](https://doi.org/10.1074/jbc.M807631200).
- Connell SD**, Heath G, Olmsted PD, Kisil A. 2013. Critical point fluctuations in supported lipid membranes. *Faraday Discussions* **161**:91–111. doi: [10.1039/c2fd20119d](https://doi.org/10.1039/c2fd20119d).
- Czajkowsky DM**, Hotze EM, Shao ZF, Tweten RK. 2004. Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. *The EMBO Journal* **23**:3206–3215. doi: [10.1038/sj.emboj.7600350](https://doi.org/10.1038/sj.emboj.7600350).
- Dang TX**, Hotze EM, Rouiller I, Tweten RK, Wilson-Kubalek EM. 2005. Prepore to pore transition of a cholesterol-dependent cytolysin visualized by electron microscopy. *Journal of Structural Biology* **150**:100–108. doi: [10.1016/j.jsb.2005.02.003](https://doi.org/10.1016/j.jsb.2005.02.003).
- Frank J**, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A. 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology* **116**:190–199. doi: [10.1006/jsbi.1996.0030](https://doi.org/10.1006/jsbi.1996.0030).
- Goddard TD**, Huang CC, Ferrin TE. 2007. Visualizing density maps with UCSF Chimera. *Journal of Structural Biology* **157**:281–287. doi: [10.1016/j.jsb.2006.06.010](https://doi.org/10.1016/j.jsb.2006.06.010).
- Hope MJ**, Bally MB, Webb G, Cullis PR. 1985. Production of large unilamellar vesicles by a rapid extrusion procedure - characterization of size distribution, trapped volume and ability to maintain a membrane-potential. *Biochimica Et Biophysica Acta* **812**:55–65. doi: [10.1016/0005-2736\(85\)90521-8](https://doi.org/10.1016/0005-2736(85)90521-8).
- Hotze EM**, Tweten RK. 2012. Membrane assembly of the cholesterol-dependent cytolysin pore complex. *Biochimica et Biophysica Acta* **1818**:1028–1038. doi: [10.1016/j.bbame.2011.07.036](https://doi.org/10.1016/j.bbame.2011.07.036).
- Hotze EM**, Wilson-Kubalek EM, Rossjohn J, Parker MW, Johnson AE, Tweten RK. 2001. Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane ss-sheet from a prepore intermediate. *The Journal of Biological Chemistry* **276**:8261–8268. doi: [10.1074/jbc.M009865200](https://doi.org/10.1074/jbc.M009865200).
- Hotze EM**, Wilson-Kubalek E, Farrand AJ, Bentsen L, Parker MW, Johnson AE, Tweten RK. 2012. Monomer-monomer interactions Propagate structural transitions necessary for pore formation by the cholesterol-dependent cytolysins. *The Journal of Biological Chemistry* **287**:24534–24543. doi: [10.1074/jbc.M112.380139](https://doi.org/10.1074/jbc.M112.380139).
- Johnson BB**, Moe PC, Wang D, Rossi K, Trigatti BL, Heuck AP. 2012. Modifications in perfringolysin O domain 4 alter the cholesterol concentration threshold required for binding. *Biochemistry* **51**:3373–3382. doi: [10.1021/bi3003132](https://doi.org/10.1021/bi3003132).
- Köster S**, van Pee K, Hudel M, Leustik M, Rhinow D, Kühlbrandt W, Chakraborty T, Yildiz Ö. 2014. Crystal structure of listeriolysin O reveals molecular details of oligomerization and pore formation. *Nature Communications* **5**:3690. doi: [10.1038/ncomms4690](https://doi.org/10.1038/ncomms4690).
- Krivov GG**, Shapovalov MV, Dunbrack RL Jr. 2009. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**:778–795. doi: [10.1002/prot.22488](https://doi.org/10.1002/prot.22488).
- Laemmli UK**. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685. doi: [10.1038/227680a0](https://doi.org/10.1038/227680a0).
- Leung C**, Bestembayeva A, Thorogate R, Stinson J, Pyne A, Marcovich C, Yang J, Drechsler U, Despont M, Jankowski T, Tschöpe M, Hoogenboom BW. 2012. Atomic force microscopy with Nanoscale cantilevers Resolves different structural conformations of the DNA double Helix. *Nano Letters* **12**:3846–3850. doi: [10.1021/nl301857p](https://doi.org/10.1021/nl301857p).
- Lindahl E**, Azuara C, Koehl P, Delarue M. 2006. NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Research* **34**:W52–W56. doi: [10.1093/nar/gkl082](https://doi.org/10.1093/nar/gkl082).
- Liu H**, Naismith JH. 2009. A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expression and Purification* **63**:102–111. doi: [10.1016/j.pep.2008.09.008](https://doi.org/10.1016/j.pep.2008.09.008).
- Ludtke SJ**, Baldwin PR, Chiu W. 1999. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology* **128**:82–97. doi: [10.1006/jsbi.1999.4174](https://doi.org/10.1006/jsbi.1999.4174).
- Lukyanova N**, Kondos SC, Farabella I, Law RHP, Reboul CF, Caradoc-Davies TT, Spicer BA, Kleinfeld O, Traore D, Ekkel SM, Voskoboinik I, Trapani JA, Hatfaludi T, Oliver K, Hotze EM, Tweten RK, Whisstock JC, Topf M, Saibil HR, Dunstone MA. Conformational changes during pore formation by the perforin-related protein pleurotolysin. *PLOS Biology*. doi: [10.1371/journal.pbio.1002049](https://doi.org/10.1371/journal.pbio.1002049).
- Malhotra S**, Sankar K, Sowdhamini R. 2014. Structural interface parameters are discriminatory in recognising near-native poses of protein-protein interactions. *PLOS ONE* **9**:e80255. doi: [10.1371/journal.pone.0080255](https://doi.org/10.1371/journal.pone.0080255).



- Marchioretto M**, Podobnik M, Dalla Serra M, Anderluh G. 2013. What planar lipid membranes tell us about the pore-forming activity of cholesterol-dependent cytolysins. *Biophysical Chemistry* **182**:64–70. doi: [10.1016/j.bpc.2013.06.015](https://doi.org/10.1016/j.bpc.2013.06.015).
- Mikhaylov A**, Sekatskii SK, Dietler G. 2013. DNA trace: a comprehensive software for polymer image processing. *Journal of Advanced Microscopy Research* **8**:241–245. doi: [10.1166/jamr.2013.1164](https://doi.org/10.1166/jamr.2013.1164).
- Mindell JA**, Grigorieff N. 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology* **142**:334–347. doi: [10.1016/S1047-8477\(03\)00069-8](https://doi.org/10.1016/S1047-8477(03)00069-8).
- Murzin AG**, Lesk AM, Chothia C. 1994. Principles determining the structure of beta-sheet barrels in proteins. I. A theoretical analysis. *Journal of Molecular Biology* **236**:1369–1381. doi: [10.1016/0022-2836\(94\)90064-7](https://doi.org/10.1016/0022-2836(94)90064-7).
- Nakamura M**, Sekino N, Iwamoto M, Ohno-Iwashita Y. 1995. Interaction of theta-toxin (perfringolysin-O), a cholesterol-binding cytolysin, with liposomal membranes - change in the aromatic side-chains upon binding and insertion. *Biochemistry* **34**:6513–6520. doi: [10.1021/bi00019a032](https://doi.org/10.1021/bi00019a032).
- Owen RH**, Boulnois GJ, Andrew PW, Mitchell TJ. 1994. A role in cell-binding for the C-terminus of pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae*. *FEMS Microbiology Letters* **121**:217–222. doi: [10.1111/j.1574-6968.1994.tb07101.x](https://doi.org/10.1111/j.1574-6968.1994.tb07101.x).
- Pettersen EF**, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**:1605–1612. doi: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084).
- Ramachandran R**, Heuck AP, Tweten RK, Johnson AE. 2002. Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin. *Nature Structural Biology* **9**:823–827.
- Ramachandran R**, Tweten RK, Johnson AE. 2004. Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment. *Nature Structural & Molecular Biology* **11**:697–705. doi: [10.1038/nsmb793](https://doi.org/10.1038/nsmb793).
- Reboul CF**, Mahmood K, Whisstock JC, Dunstone MA. 2012. Predicting giant transmembrane beta-barrel architecture. *Bioinformatics* **28**:1299–1302. doi: [10.1093/bioinformatics/bts152](https://doi.org/10.1093/bioinformatics/bts152).
- Reboul CF**, Whisstock JC, Dunstone MA. 2014. A new model for pore formation by cholesterol-dependent cytolysins. *PLOS Computational Biology* **10**:e1003791. doi: [10.1371/journal.pcbi.1003791](https://doi.org/10.1371/journal.pcbi.1003791).
- Rosado CJ**, Kondos S, Bull TE, Kuiper MJ, Law RH, Buckle AM, Voskoboinik I, Bird PI, Trapani JA, Whisstock JC, Dunstone MA. 2008. The MACPF/CDC family of pore-forming toxins. *Cellular Microbiology* **10**:1765–1774. doi: [10.1111/j.1462-5822.2008.01191.x](https://doi.org/10.1111/j.1462-5822.2008.01191.x).
- Rossjohn J**, Feil SC, McKinstry WJ, Tweten RK, Parker MW. 1997. Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form. *Cell* **89**:685–692. doi: [10.1016/S0092-8674\(00\)80251-2](https://doi.org/10.1016/S0092-8674(00)80251-2).
- Sali A**, Blundell TL. 1993. Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**:779–815. doi: [10.1006/jmbi.1993.1626](https://doi.org/10.1006/jmbi.1993.1626).
- Sato TK**, Tweten RK, Johnson AE. 2013. Disulfide-bond scanning reveals assembly state and  $\beta$ -strand tilt angle of the PFO  $\beta$ -barrel. *Nature Chemical Biology* **9**:383–389. doi: [10.1038/nchembio.1228](https://doi.org/10.1038/nchembio.1228).
- Shatursky O**, Heuck AP, Shepard LA, Rossjohn J, Parker MW, Johnson AE, Tweten RK. 1999. The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* **99**:293–299. doi: [10.1016/S0092-8674\(00\)81660-8](https://doi.org/10.1016/S0092-8674(00)81660-8).
- Shen MY**, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Science* **15**:2507–2524. doi: [10.1110/ps.062416606](https://doi.org/10.1110/ps.062416606).
- Shepard LA**, Heuck AP, Hamman BD, Rossjohn J, Parker MW, Ryan KR, Johnson AE, Tweten RK. 1998. Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy. *Biochemistry* **37**:14563–14574. doi: [10.1021/bi981452f](https://doi.org/10.1021/bi981452f).
- Sonnen AF**, Plitzko JM, Gilbert RJ. 2014. Incomplete pneumolysin oligomers form membrane pores. *Open Biology* **4**:140044. doi: [10.1098/rsob.140044](https://doi.org/10.1098/rsob.140044).
- Tilley SJ**, Orlova EV, Gilbert RJ, Andrew PW, Saibil HR. 2005. Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* **121**:247–256. doi: [10.1016/j.cell.2005.02.033](https://doi.org/10.1016/j.cell.2005.02.033).
- van Heel M**. 1984. Multivariate statistical analysis of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy* **13**:165–183. doi: [10.1016/0304-3991\(84\)90066-4](https://doi.org/10.1016/0304-3991(84)90066-4).
- Xu L**, Huang B, Du H, Zhang XC, Xu J, Li X, Rao Z. 2010. Crystal structure of cytotoxin protein suilysin from *Streptococcus suis*. *Protein & Cell* **1**:96–105. doi: [10.1007/s13238-010-0012-3](https://doi.org/10.1007/s13238-010-0012-3).



## Appendix B

The following publication makes a significant contribution to the MACPF field of research. It is reproduced in the article format.

*Conformational changes during pore formation by the perforin-related protein pleurotolysin*

PLoS Biology (2015), 13(2):e1002049

N Lukoyanova\*, S Kondos\*, I Farabella, R Law, **CF Reboul**, T Caradoc-Davies, B Spicer, O Kleifed, D Traore, S Ekkel, I Voskoboinik, T Hatfaludi, Oliver K, E Hotze, R Tweten, JC Whisstock, M Topf, H Saibil, MA Dunstone

RESEARCH ARTICLE

# Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin

Natalya Lukoyanova<sup>1‡</sup>, Stephanie C. Kondos<sup>2,3‡</sup>, Irene Farabella<sup>1</sup>, Ruby H. P. Law<sup>2,3</sup>, Cyril F. Reboul<sup>2,3</sup>, Tom T. Caradoc-Davies<sup>3,4</sup>, Bradley A. Spicer<sup>2</sup>, Oded Kleinfeld<sup>2</sup>, Daouda A. K. Traore<sup>2</sup>, Susan M. Ekkel<sup>2</sup>, Ilia Voskoboinik<sup>5</sup>, Joseph A. Trapani<sup>5</sup>, Tamas Hatfaludi<sup>2</sup>, Katherine Oliver<sup>6</sup>, Eileen M. Hotze<sup>7</sup>, Rodney K. Tweten<sup>7</sup>, James C. Whisstock<sup>2,3</sup>, Maya Topf<sup>1</sup>, Helen R. Saibil<sup>1\*</sup>, Michelle A. Dunstone<sup>2,3,8\*</sup>

**1** Department of Crystallography/Biological Sciences, Institute of Structural and Molecular Biology, Birkbeck College, London, United Kingdom, **2** Department of Biochemistry and Molecular Biology, Monash University, Clayton, Melbourne, Australia, **3** The ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Melbourne, Australia, **4** Australian Synchrotron, Clayton, Melbourne, Australia, **5** Peter MacCallum Cancer Centre, East Melbourne, Australia, **6** Institute of Structural and Molecular Biology, University College London, London, United Kingdom, **7** Department of Microbiology and Immunology, The University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America, **8** Department of Microbiology, Monash University, Melbourne, Australia

‡ These authors contributed equally to this work and are joint first authors.

\* [REDACTED] (HRS) [REDACTED] (MAD)



**OPEN ACCESS**

**Citation:** Lukoyanova N, Kondos SC, Farabella I, Law RHP, Reboul CF, Caradoc-Davies TT, et al. (2015) Conformational Changes during Pore Formation by the Perforin-Related Protein Pleurotolysin. *PLoS Biol* 13(2): e1002049. doi:10.1371/journal.pbio.1002049

**Academic Editor:** Raimund Dutzler, University of Zurich, SWITZERLAND

**Received:** June 3, 2014

**Accepted:** December 10, 2014

**Published:** February 5, 2015

**Copyright:** © 2015 Lukoyanova et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All structure factor and coordinate files for the crystal structures are available in the Protein Data Bank database (accession numbers 4OEB (PlyA), 4OEB (PlyB), 4OV8 (TMH1 lock)). Cryo-EM maps of the Ply pore and prepore structures are available in the Electron Microscopy Data Bank (accession numbers EMD-2793 (pore), EMD-2794 (TMH1 lock), EMD-2795 (TMH2 helix lock) and EMD-2796 (TMH2 strand lock)). The structure coordinate files for the fitted pore model is available in the Protein Data Bank database (accession number 4V2T). The structure coordinate

## Abstract

Membrane attack complex/perforin-like (MACPF) proteins comprise the largest superfamily of pore-forming proteins, playing crucial roles in immunity and pathogenesis. Soluble monomers assemble into large transmembrane pores via conformational transitions that remain to be structurally and mechanistically characterised. Here we present an 11 Å resolution cryo-electron microscopy (cryo-EM) structure of the two-part, fungal toxin Pleurotolysin (Ply), together with crystal structures of both components (the lipid binding PlyA protein and the pore-forming MACPF component PlyB). These data reveal a 13-fold pore 80 Å in diameter and 100 Å in height, with each subunit comprised of a PlyB molecule atop a membrane bound dimer of PlyA. The resolution of the EM map, together with biophysical and computational experiments, allowed confident assignment of subdomains in a MACPF pore assembly. The major conformational changes in PlyB are a ~70° opening of the bent and distorted central β-sheet of the MACPF domain, accompanied by extrusion and refolding of two α-helical regions into transmembrane β-hairpins (TMH1 and TMH2). We determined the structures of three different disulphide bond-trapped prepore intermediates. Analysis of these data by molecular modelling and flexible fitting allows us to generate a potential trajectory of β-sheet unbending. The results suggest that MACPF conformational change is triggered through disruption of the interface between a conserved helix-turn-helix motif and the top of TMH2. Following their release we propose that the transmembrane regions assemble into β-hairpins via top down zippering of backbone hydrogen bonds to form the membrane-inserted β-barrel. The intermediate structures of the MACPF domain during refolding into the β-barrel pore establish a structural paradigm for the transition from soluble

files for the sets of prepore models are available in the Protein Data Bank database (accession numbers 4V3A (TMH1 lock), 4V3M (TMH2 helix lock) and 4V3N (TMH2 strand lock)).

**Funding:** HRS acknowledges support from the Wellcome Trust (grant 079605/2/06/2) for EM facilities and doctoral training support to KO (MRC G1001602), the BBSRC (BB/D00873/1) and the ERC (294408). MT acknowledges the BBSRC (BB/K01692X/1) and Leverhulme Trust (RPG-2012-519). We thank D. Houldershaw and R. Westlake for computing support, L. Wang for EM support, and E.V. Orlova, D.K. Clare and A.P. Pandurangan for discussion. MAD is a National Health and Medical Research Council of Australia (NHMRC) Career Development Fellow. MAD acknowledges support from the Australian Research Council (ARC) [DP120104058, DP0986811, CE140100011] and the National Health and Medical Research Council [606471]. JCW is an NHMRC Senior Principal Research Fellow. JCW also acknowledges the support of an Australian Research Council Federation Fellowship. CFR is supported by the Australian Postgraduate Award. DT acknowledges the support of an ARC Super Science Fellowship. RKT acknowledges the National Institutes of Health, NIAID [AI037657]. The authors acknowledge the support of the Victorian Life Sciences Computation Initiative (VLSCI, Melbourne, Australia) and the Multi-modal Australian ScienceS Imaging and Visualisation Environment (MASSIVE) ([www.massive.org.au](http://www.massive.org.au)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** CDC, cholesterol dependent cytotoxicity; EM, electron microscopy; HTH, helix-turn-helix; MACPF, membrane attack complex/perforin-like family; SCCC, segment-based cross-correlation score; TMH, transmembrane hairpin.

monomer to pore, which may be conserved across the whole superfamily. The TMH2 region is critical for the release of both TMH clusters, suggesting why this region is targeted by endogenous inhibitors of MACPF function.

## Author Summary

Animals, plants, fungi, and bacteria all use pore-forming proteins of the membrane attack complex-perforin (MACPF) family as lethal, cell-killing weapons. These proteins are able to insert into the plasma membranes of target cells, creating large pores that short circuit the natural separation between the intracellular and extracellular milieu, with catastrophic results. However, the pore-forming proteins must undergo a substantial transformation from soluble precursors to a large barrel-shaped transmembrane complex as they punch their way into cells. Using a combination of X-ray crystallography and cryo electron microscopy, we have visualized, for the first time, the mechanism of action of one of these pore-forming proteins—pleurotolysin, a MACPF protein from the edible oyster mushroom. This enabled us to propose a model of the pleurotolysin pore by fitting the crystallographic structures of the pore proteins into a three-dimensional map of the pore obtained by cryo electron microscopy. We then designed a set of double mutants that allowed us to chemically trap intermediate states along the trajectory of the pore formation process, and to determine their structures too. By combining these data we proposed a detailed molecular mechanism for pore formation. The pleurotolysin first assembles into rings of 13 subunits, each of which then opens up by about 70° during pore formation. This process is accompanied by refolding and extrusion of two compact regions from each subunit into long hairpins that then zipper together to form an 80-Å wide barrel-shaped channel through the membrane.

## Introduction

Membrane pore-forming proteins have the unique property of being expressed as metastable, water-soluble monomers that convert into a membrane inserted form. These proteins typically assemble into prepore oligomers on the target membrane surface. A dramatic conformational change then permits membrane insertion and formation of transmembrane pores [1–4].

The membrane attack complex/perforin-like family (MACPF) proteins form the largest superfamily of pore-forming proteins identified to date. They include perforin and complement component-9 (C9), mammalian pore-forming proteins that function as weapons of the humoral and cellular immune system, respectively [5]. The superfamily also includes a wide range of molecules implicated in defense or attack [6–8]. For example, invasion by the protozoan parasites *Plasmodium* spp. and egress by *Toxoplasma gondii* requires MACPF proteins, plants utilize the MACPF fold to combat bacterial infection [9], and MACPF-related proteins can be identified in numerous Gram negative and Gram positive bacteria. Finally, a significant group of MACPF proteins play important, but poorly understood, roles in embryonic development and neurobiology [10–12].

Despite the absence of detectable sequence identity, the first crystal structures of MACPF proteins revealed that the pore-forming domain unexpectedly shared homology with the pore-

forming bacterial cholesterol dependent cytolysins (CDCs) family [13–15]. This structural similarity extended across the key elements involved in pore formation (originally annotated as three non-contiguous domains 1–3 in CDCs). The central, common feature of the MACPF/CDC fold is a four stranded, highly twisted  $\beta$ -sheet decorated with three small clusters of  $\alpha$ -helices. Two of these helical bundles contain the regions destined to insert into the membrane (transmembrane hairpins TMH1 and TMH2). The third  $\alpha$ -helical region comprises a short helix-turn-helix (HTH) motif formed *via* a sequence insertion at the bend of the central  $\beta$ -sheet. The HTH motif packs on top of TMH2. These structural similarities, together with commonality of a pore-forming function, suggested that MACPF proteins share a common ancestor with CDCs and assemble into giant pores via a CDC-like mechanism [13,14,16–19].

Previous studies have provided important insight into pore formation by CDCs. Electron microscopy (EM), biochemical, and biophysical studies of CDCs showed that monomers assemble into prepore oligomers on the membrane surface without major conformational changes in the subunits [17,19–22]. However, conversion to the pore form involves dramatic secondary and tertiary conformational changes in which the highly twisted  $\beta$ -sheet opens up and the assembly collapses  $\sim 40$  Å towards the membrane surface, allowing unfurling of TMH1 and TMH2 and their insertion into the membrane as amphipathic  $\beta$ -hairpins [19–22].

The CDCs form initial interactions with the membrane through a C-terminal lipid binding immunoglobulin-like (Ig) domain. In the MACPF branch of the superfamily a wide variety of domains are found both N- and C-terminal to the pore-forming MACPF domain. For example, perforin includes a C-terminal lipid and calcium binding C2 domain (a variation of the Ig fold). Similar to the CDC Ig domain, this region mediates initial interaction of perforin with the target membrane. The MACPF domains in the complement membrane attack complex proteins are flanked by arrays of small disulphide constrained domains (e.g., thrombospondin, epidermal growth factor, and complement control protein domains). Rather than interacting directly with membranes, the role of these regions includes mediation of key protein-protein interactions that recruit the MACPF domain to the target cell surface [23–25].

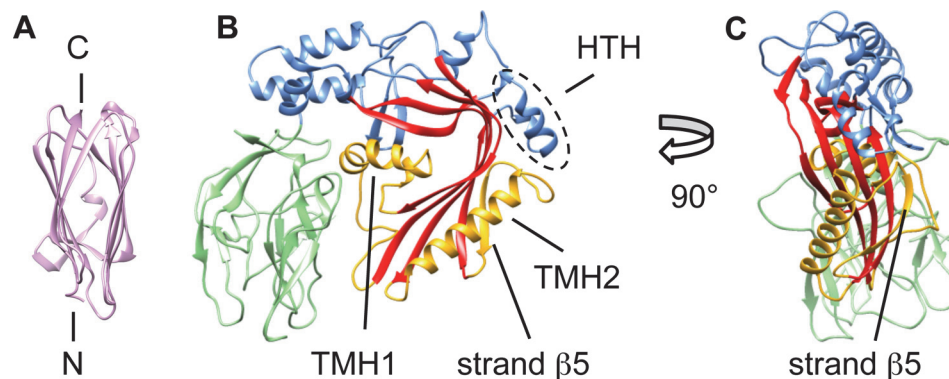
The molecular structures of key intermediates in the assembly of MACPF and CDC pore complexes remain obscure, but are necessary to understand the transition from a monomeric form into oligomeric membrane prepores and then into pores. Here we have analysed this transition, using a variety of structural and biophysical approaches. Structures of MACPF and CDC oligomeric assemblies by EM have been very limited in resolution, owing to their heterogeneity and flexibility. To gain further insight into the structural conversions in pore formation, we chose pleurotolysin (Ply), a MACPF protein consisting of two components, PlyA and PlyB, from *Pleurotus ostreatus* [26,27]. Previous studies have shown that PlyA binds membranes and is required to recruit the pore-forming MACPF protein PlyB to the membrane surface. PlyA and PlyB together form relatively small and regular pores in liposomes [27,28]. As well as determining the structure of the pleurotolysin pore, we used protein-engineering approaches to trap and structurally characterise three distinct prepore intermediates. Together these approaches allowed us to visualise a potential molecular trajectory of a MACPF protein during pore formation.

## Results

### Crystal Structures of the Pleurotolysin Components

The 1.85 Å X-ray crystal structure of PlyA (Fig. 1A; S1 Table) revealed a  $\beta$ -sandwich fold, unexpectedly related to the actinoporin-like family of pore-forming toxins [29]. Previous studies





**Figure 1. Crystal structures of the two pleurotolysin components: PlyA and PlyB.** (A) The structure of PlyA showing a  $\beta$ -sandwich fold similar to that seen in actinoporins [29]. (B) The structure of PlyB, with the bent, central  $\beta$ -sheet characteristic of the MACPF/CDC superfamily (red). The transmembrane hairpin regions are labelled as TMH1 and TMH2 (yellow) and the helix-turn-helix motif is labelled HTH (outlined by the dashed oval). The trefoil of C-terminal  $\beta$ -rich domains is shown in green. The upper part of the central sheet is flanked mainly by helical regions (blue). The conserved pore-forming core consists of the bent  $\beta$ -sheet and the TMH domains. (C) PlyB seen edge-on, clearly showing strand  $\beta$ 5.

doi:10.1371/journal.pbio.1002049.g001

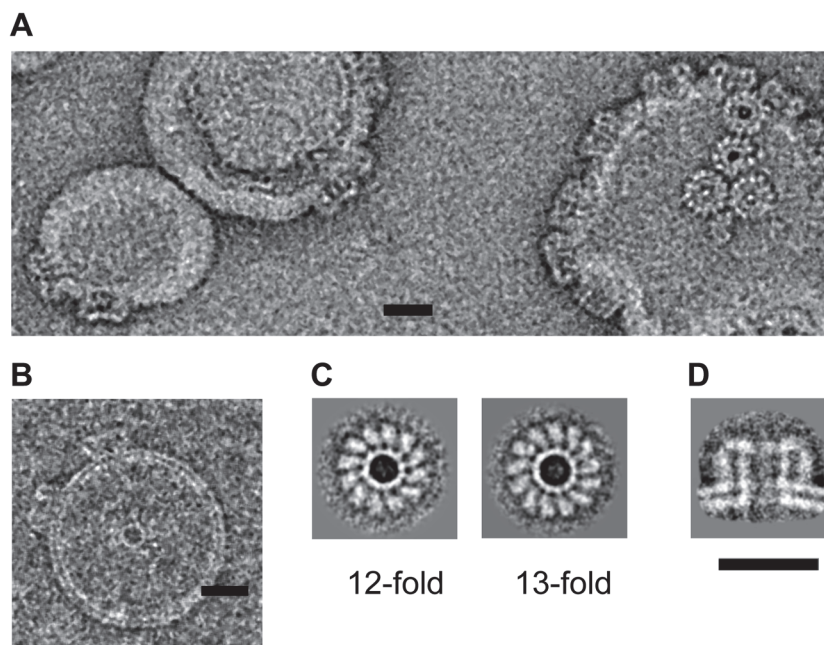
suggest that actinoporin-like proteins interact with membranes *via* one end of the  $\beta$ -sandwich, with the N-terminal sequence responsible for forming the pore [29]. However, PlyA lacks the proposed actinoporin N-terminal transmembrane region consistent with the observation that PlyA binds membranes, but is unable to form pores on its own [27].

The 2.2 Å structure of PlyB (Fig. 1B and 1C; S2 Table) reveals an N-terminal MACPF domain (blue/red/yellow) followed by three small  $\beta$ -rich domains clustered in a globular trefoil-like arrangement (green). The MACPF domain of PlyB contains a central, four-stranded bent and twisted  $\beta$ -sheet characteristic of the MACPF/CDC superfamily (red). The TMH1 cluster of helices (yellow) is located on the inside of PlyB, next to the concave face of the central  $\beta$ -sheet. TMH2 (yellow) comprises a single large  $\alpha$ -helix and an additional  $\beta$ -strand (termed “strand  $\beta$ 5”), located on the edge of the central  $\beta$ -sheet. Together, the central  $\beta$ -sheet and the TMH regions constitute the topologically conserved MACPF/CDC pore-forming fold.

## Cryo-EM Structure of the Pleurotolysin Pore

EM images of liposomes with added PlyAB showed distinctive, ring shaped pore structures (Fig. 2A and 2B). Analysis of negative stain EM images of oligomeric rings of Ply on membranes showed that the majority of the oligomers had 13-fold symmetry (75%), but 12- (15%), 11- (5%), and 14-fold (5%) rings were also present (Fig. 2C). For 3-D reconstruction, we extracted 14,700 individual cryo-EM images of pore side views in liposomes (Fig. 2D). The images were analysed by the single particle approach, following the method developed for the CDC pneumolysin [17]. This allowed us to sort the pore views by symmetry, enabling determination of an 11 Å resolution cryo-EM map of a liposome-embedded 13-fold pleurotolysin pore from 8,770 views (Fig. 3A and 3B). We used the crystal structures of PlyA and PlyB together with biophysical data (S1 Fig.) to interpret the map. A single PlyB moiety was fitted into the upper part of the pore structure (Fig. 3C). The C-terminal trefoil (green) and the  $\alpha$ -helices at the top of the MACPF domain (blue) unambiguously fit the EM density with only minor structural rearrangement. The core of the MACPF domain undergoes a massive opening but does not collapse as in CDCs (Fig. 3C).

The structure was modeled by flexible fitting in a multistep procedure [30]. In the pore map, the position of PlyB is clearly recognizable in the upper part of each subunit, while the V-



**Figure 2. Electron microscopy of pleurotolysin pores.** Representative views of negatively stained (A) and vitrified (B) Ply pores on liposomes. (C) Averaged views of 12-fold and 13-fold symmetric pores on lipid monolayers (negative stain). (D) Averaged side view of Ply pores on liposomes (cryo-EM). Scale bar, 20 nm.

doi:10.1371/journal.pbio.1002049.g002

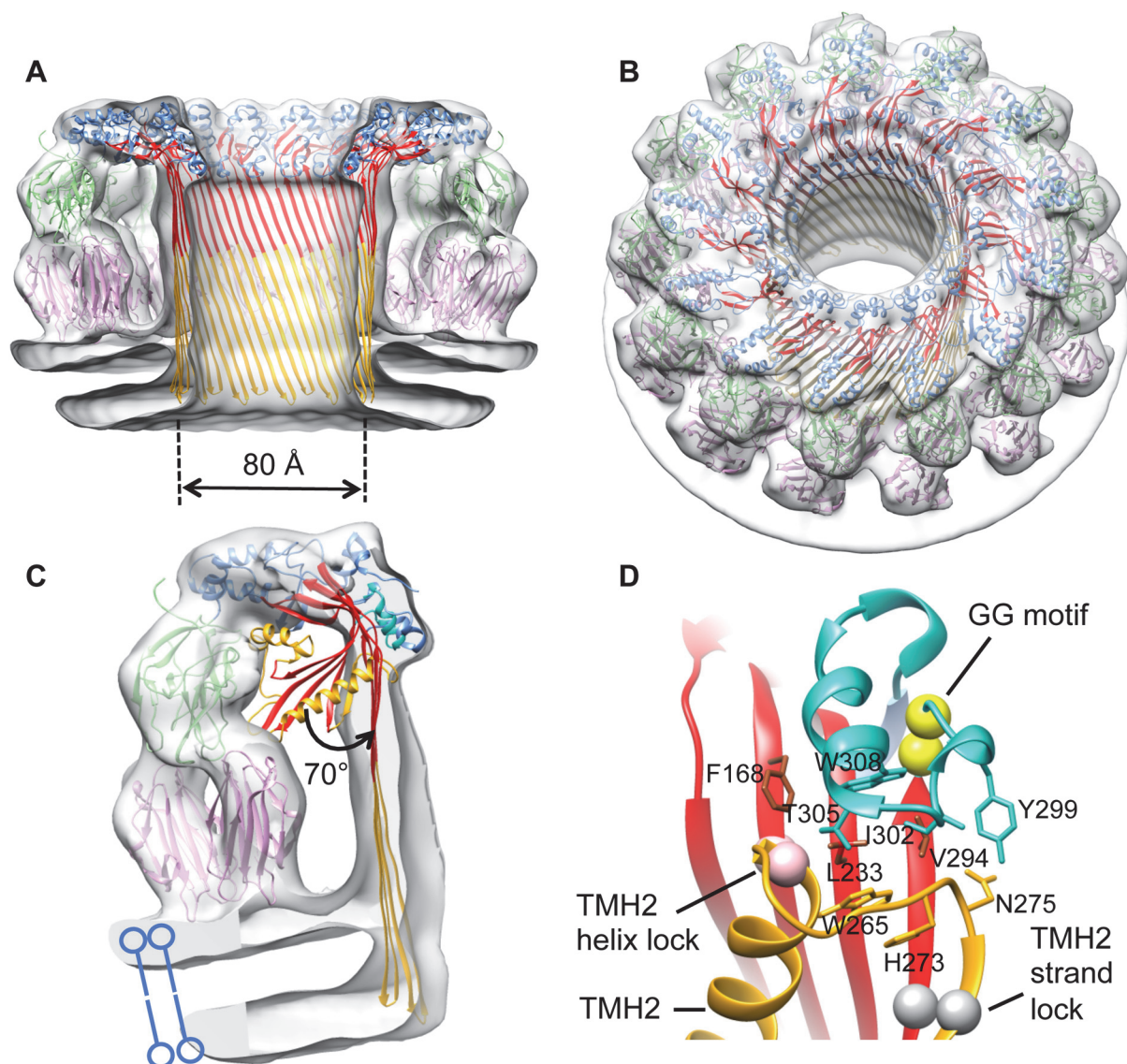
shaped density at the base of each asymmetric unit accommodates two PlyA molecules. The positions of PlyB subdomains were refined without TMH1 and TMH2, because these trans-membrane regions are expected to refold to form the  $\beta$ -barrel of the pore. The best fits were further refined with Flex-EM [30] via simulated annealing rigid-body dynamics.

To identify the sequence forming the transmembrane  $\beta$ -hairpins we carried out fluorescence spectroscopy studies using single cysteine mutants in TMH1, as previously performed on CDCs [20]. This approach revealed an alternating pattern of emission between residues 128–147 consistent with a  $\sim 30$  Å membrane-spanning amphipathic  $\beta$ -hairpin structure (S1 Fig). This information provided a useful restraint for the fitting. In the resulting pore model, each MACPF domain forms a four-stranded  $\beta$ -sheet (Fig. 3A–3C).

$\beta$ -barrels are limited to discrete architectures, each with a characteristic strand tilt relative to the barrel axis [31]. For a barrel composed of  $n$  strands, the shear number  $S$  describes the register of hydrogen bonding between residues in adjacent  $\beta$ -strands and defines the strand tilt and the dimensions of the formed barrel: the greater the strand tilt, the wider and shorter the barrel [32]. Only three Ply barrel models, with  $S = 0$  ( $0^\circ$  tilt),  $S = n/2$  ( $20^\circ$  tilt), and  $S = n$  ( $36^\circ$  tilt) have dimensions comparable with the Ply pore cryo-EM map (S2 Fig). The  $S = n/2$  model gave the best fit in diameter and height (CC = 0.90 versus 0.73 for  $S = 0$  barrel and 0.74 for  $S = n$ ).

This 52-stranded  $\beta$ -barrel was combined with a 13-mer ring of fitted PlyB molecules. Because of steric clashes with the barrel, further refinement using Flex-EM was performed on the HTH motif (residues 298–313) (Figs. 1B, and 3C, 3D). After refinement of the central asymmetric unit, the pore was rebuilt with C13 symmetry in Chimera [33] to give the final pore model. In this pore, the central  $\beta$ -sheet has straightened and opened by  $\sim 70^\circ$ , as measured from the fitting, and TMH1 and TMH2 are fully unwound into  $\beta$ -hairpins to form a  $\beta$ -barrel spanning the membrane bilayer (Fig. 3A–3C). The pore channel is thus formed by a 52-strand  $\beta$ -barrel that is 80 Å in inner diameter and over 100 Å in height.

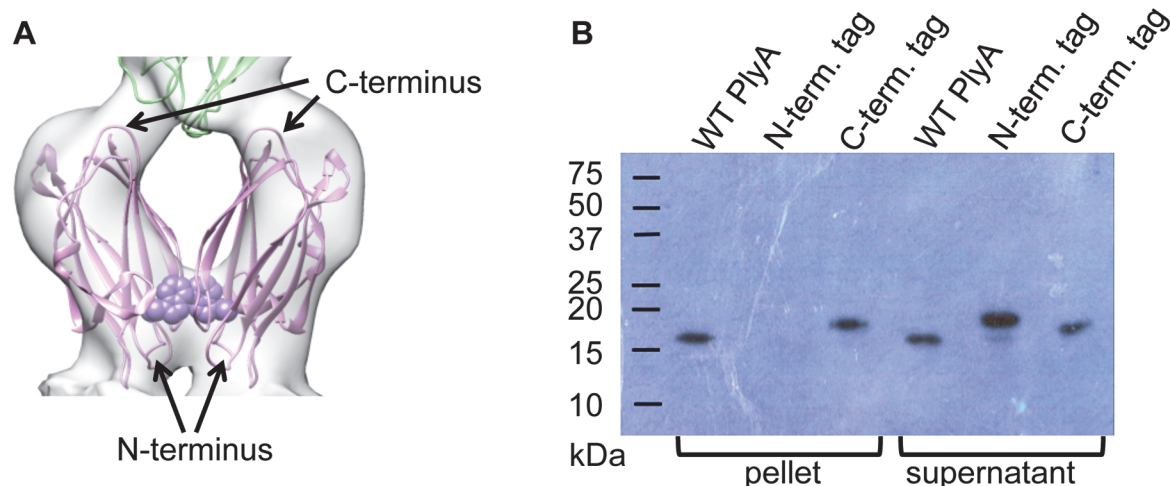




**Figure 3. Structure of the pleurotolysin pore.** (A) Cut away side and (B) tilted surface views of the cryo-EM reconstruction of a pleurotolysin pore with the fitted atomic structures. (C) Segment of the pore map corresponding to a single subunit with pore model fitted into the density. The PlyB crystal structure is superposed to show a 70° opening of the MACPF  $\beta$ -sheet (red) and movement of the HTH motif (cyan). TMH regions (yellow) are refolded into transmembrane  $\beta$ -hairpins. The PlyB C-terminal trefoil (green) sits on top of the PlyA dimer (pink). (D) Interface between TMH2, the HTH region, and the underlying sheet in the PlyB crystal structure. The position of the TMH2 helix lock (pink spheres) and TMH2 strand lock (grey spheres) are shown. The highly conserved “GG” motif (296–297) in the HTH region is represented as yellow spheres.

doi:10.1371/journal.pbio.1002049.g003

The PlyB C-terminal trefoil sits in the cavity formed by a V-shaped wedge of density contacting the membrane (Figs. 3C and 4A). This density can be accounted for by two PlyA molecules, revealing a tridecameric PlyB/2xPlyA pore assembly. The symmetrical shape of PlyA precludes discrimination of up/down orientation in the density. However, in the crystal structure of PlyA, we noted two different V-shaped dimers (termed N-dimer and C-dimer) in the asymmetric unit (S3A and S3D Fig.). Both forms fitted adequately into EM density, placing either the PlyA N-terminus (N-dimer) or C-terminus (C-dimer) in proximity to the membrane surface. We tested the orientation of PlyA by adding a hexahistidine tag to the N-terminus



**Figure 4. Validation of the orientation of PlyA.** (A) Proposed orientation of PlyA dimer (pink) and interface with PlyB C-terminal trefoil (green). Trp 6 is shown as purple spheres. (B) Western blot showing PlyA binding to red blood cells when untagged or C-terminally tagged but not when N-terminally tagged.

doi:10.1371/journal.pbio.1002049.g004

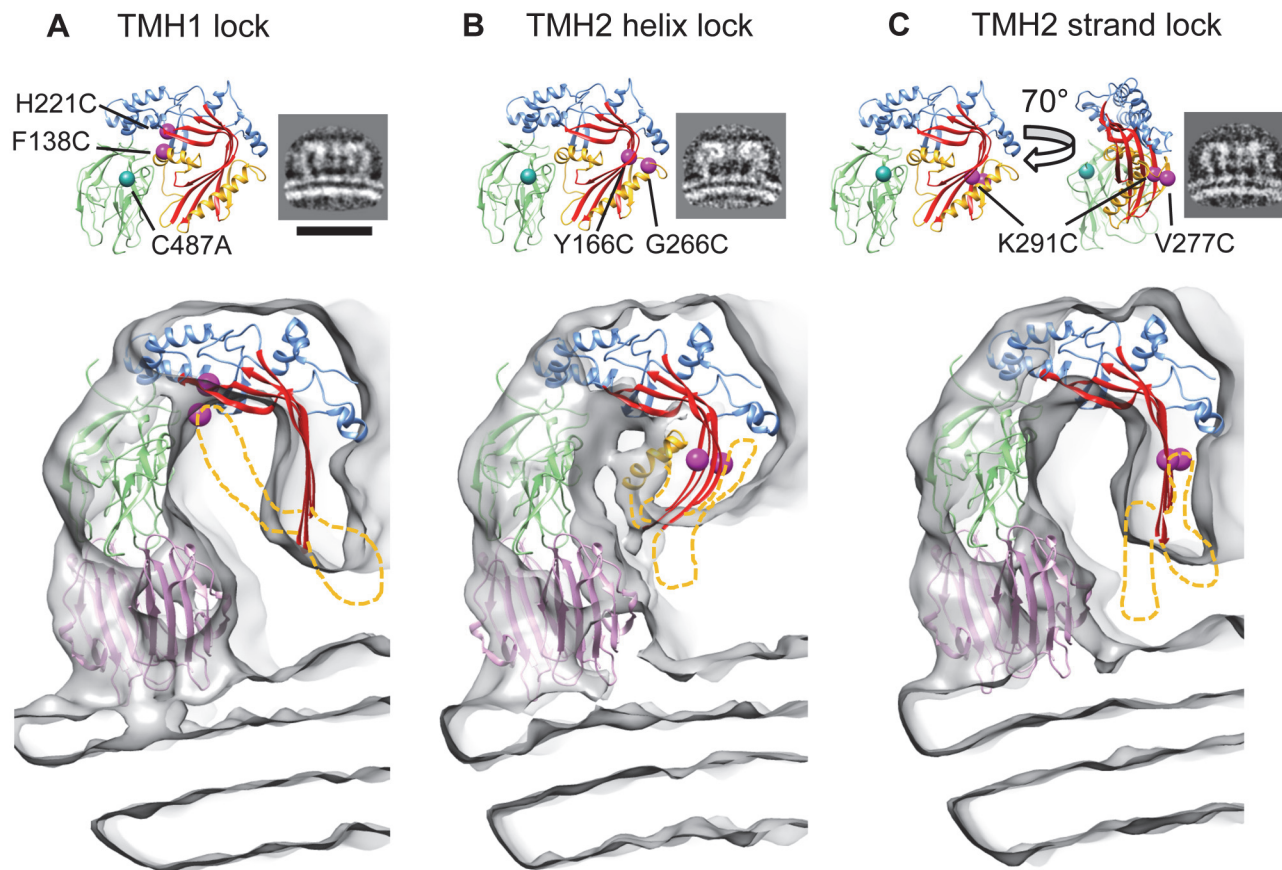
(Fig. 4A and 4B), which abrogated membrane binding of PlyA to red blood cells whereas a C-terminal tag had no effect on binding (Fig. 4B). Also, mutation of Trp 6 (W6E), located in the PlyA N-dimer interface, reduced membrane binding and led to 100-fold lower pore-forming activity (Fig. 4A, denoted as purple spheres; S4A and S4B Fig.). These data support an N-dimer-like arrangement of PlyA molecules (Fig. 4A), consistent with the known orientation of actinoporins on the membrane surface [29].

The resulting fit of 26 PlyA and 13 PlyB subunits had a cross-correlation coefficient of 0.8 with the map which includes part of the membrane as measured in Chimera [33]. To evaluate the local quality of fit, the segment-based cross-correlation coefficient (SCCC) [34] was determined and plotted on the pore subunit structure (S5 Fig.). This analysis shows that the fit is more reliable for PlyB than for PlyA, because the map resolution is better in the region occupied by PlyB.

To probe the mechanism of pore assembly, we engineered a series of disulphide bonds to limit movement in either TMH1 or TMH2. As performed for perfringolysin O and other CDCs [35], the TMH regions were trapped by introducing cross-links to the central sheet or other adjacent regions in the monomer structure. This trapping allows oligomer assembly but prevents the TMH region from unfolding enough to insert into the membrane. The disulphide trap mutants were engineered on a background PlyB variant that lacks the wild type cysteine (C487A) in order to avoid incorrect disulphide bond formation. PlyB<sup>C487A</sup> retains wild type activity according to haemolysis assay (S6A Fig.). We then determined the cryo-EM structures of three different prepore-locked variants.

## Structure of a TMH1 Trapped Intermediate

Oxidised TMH1 variant PlyB<sup>F138C,H221C</sup> (TMH1 lock) (Fig. 5A) possessed no detectable lytic activity (S6B Fig.), but reduction of the disulphide restored wild type lytic activity. Furthermore, the oxidised form could assemble into oligomeric prepores on erythrocytes or liposomes, and these prepores could be converted into lytic pores by disulphide reduction (S6C Fig.). These data suggest that the TMH1 lock prepore structure is an intermediate in the formation of the pore. The crystal structure of the TMH1 trapped variant was determined and is otherwise indistinguishable from the wild type (S7 Fig.; S3 Table).



**Figure 5. Three dimensional (3-D) reconstructions of disulphide locked pleurotolysin prepores.** (A) PlyB crystal structure with positions of TMH1 disulphide lock indicated by magenta spheres and corresponding side view average of the liposome-bound prepores (cryo-EM). Scale bar, 20 nm. Main panel, cut-away view of the prepores cryo-EM map with the model obtained by flexible fitting. No TMH density is seen in the TMH1 lock prepores map. (B) The equivalent panels are shown for the TMH2 helix lock map and model. Partial density is seen for the TMH1 region. (C) Equivalent views of the TMH2 strand lock map and model. No density is seen for the TMH1 regions. These regions must therefore be disordered and they were omitted from the fits. The disordered regions are shown schematically as yellow dashed lines; the long TMH1 helix is illustrated in (B) but was not part of the fitting. Mutated residues are shown: TMH1 lock; F138C (located in the TMH1 region, yellow), H221C. TMH2 helix lock; Y166C, G266C (located in the TMH2 helix region, yellow). TMH2 strand lock; V277C (located in the fifth  $\beta$ -strand, TMH2 region, yellow), K291C, all on the C487A background.

doi:10.1371/journal.pbio.1002049.g005

In order to analyse the degree of  $\beta$ -sheet opening we created a library of thousands of molecular models and then performed constrained fitting into the prepores map. This procedure is described in [S8 Fig](#). Briefly, using the monomer and pore structures as end points, we generated two series of angular sweeps of the beta sheet opening or closing, using the TEMPy software [36]. Each of these 2,300 generated  $\beta$ -sheet conformations was combined twice with the rest of the PlyB structure, either with the monomer or with the pore conformation, using MODELLER/Flex-EM [30]. The resulting 4,600 models were assessed by their fit to the prepores map and by statistical potentials, and the best ranking fits were used to estimate the angle of  $\beta$ -sheet opening in the prepores conformation ([S4 Table](#)).

Remarkably, the cryo-EM 3-D structure of the TMH1 lock prepores showed that the central sheet of the MACPF domain in the prepores assembly had opened substantially ( $53^\circ \pm 9^\circ$ ) ([Fig. 5A](#)). In these prepores, the top half of the barrel has formed, but not the lower, transmembrane part. Indeed, no density could be observed for most of TMH1 and TMH2 and it appears that these regions are largely unstructured. Thus, these data reveal that locking TMH1 has little effect in limiting the MACPF  $\beta$ -sheet opening.



## Structures of TMH2 Trapped Intermediates Identify a Key Region for Controlling Sheet Opening

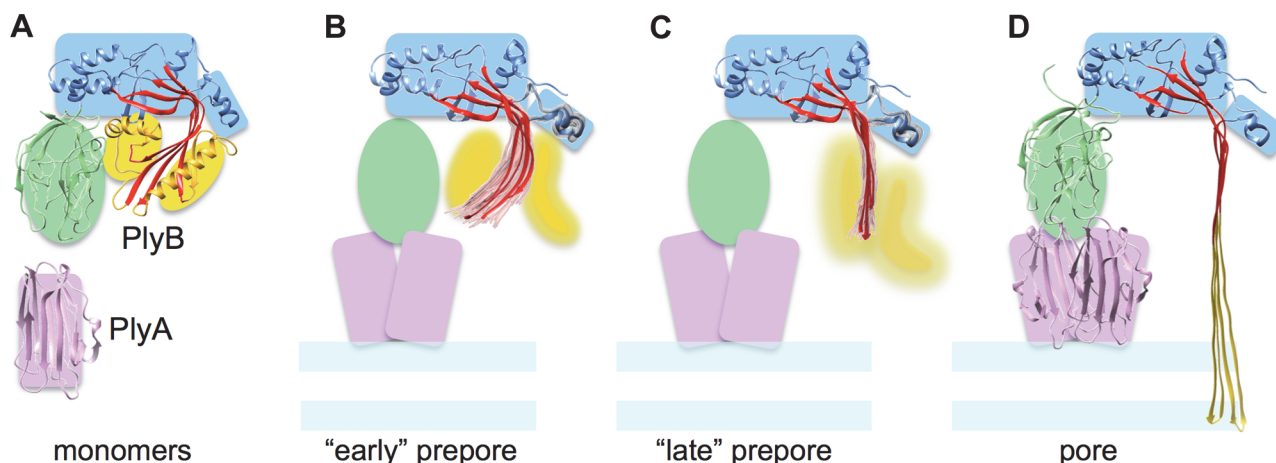
To lock TMH2 to the central core  $\beta$ -sheet of the MACPF domain we identified a second variant PlyB<sup>Y166C,G266C</sup> (TMH2 helix lock) that could be trapped in the prepore state (Fig. 5B). As with TMH1 lock, addition of reducing agent allows it to continue its trajectory to the pore form (S6D and S6E Fig.). The cryo-EM structure of the TMH2 helix lock prepore revealed a very different picture from the TMH1 lock. The core sheet opening is only  $37^\circ \pm 11^\circ$  with some density remaining for TMH1, suggesting that this region remains partly ordered (Fig. 5B). These data show that the MACPF fold oligomerises without substantial relief of the twist in the core sheet or TMH1 release. As in the TMH1 lock prepore, no density can be seen extending to the membrane surface, consistent with the absence of lytic function. The  $\beta$ -sheet opening was analysed as above (S8 Fig.). Remarkably, the results reveal that restricting the movement of the top of the TMH2  $\alpha$ -helix prevents unbending of the MACPF core  $\beta$ -sheet. This finding is opposite to expectation, since TMH1 forms an extensive buried interface, whereas TMH2 makes fewer contacts with the domain core sheet. However, our results are consistent with observations that TMH2 is important for controlling pore formation in other superfamily members, including interactions with the MAC inhibitor CD59, and the pH trigger of the CDC listeriolysin O [25,37,38].

To further probe TMH2 function, a third PlyB disulphide lock was created to join strands  $\beta$ 4 of the central  $\beta$ -sheet to  $\beta$ 5 of TMH2, PlyB<sup>V277C,K291C</sup> (TMH2 strand lock; Figs. 5C, S6F, and S6G). Cryo-EM and modelling analysis showed that the central  $\beta$ -sheet was open to the same extent as in the TMH1 lock (Fig. 4C,  $49^\circ \pm 8^\circ$ ). This result provides an interesting contrast to the consequences of the TMH2 helix lock and highlights that simply restricting TMH2 movement through locking strands  $\beta$ 4 and  $\beta$ 5 does not prevent opening of the core sheet. The restrictions enforced by the TMH2 strand lock are, however, sufficient to prevent extension of the TMH2 hairpin since no density is seen extending into the membrane. These data collectively imply that neither TMH region can enter the membrane without the other, suggesting that the TMH sequences have evolved for cooperative folding and assembly.

## Discussion

Here, we present a series of structures that identify the major conformational changes during MACPF pore formation. The final pore structure reveals that individual PlyB monomers in the pore have the orientation seen for those in the distantly related CDCs [17]. Although sequence-related to perforin, their pores differ in several respects. Like CDCs, perforin is a thin, key-shaped molecule, but it does not open up in the pore state [18]. This difference likely arises from the divergent structures surrounding the conserved MACPF core, as well as from its longer TMH regions. In addition, C-terminal labelling indicated the opposite  $\beta$ -sheet orientation in the perforin pore [18]. A model based on a more recently determined C8 structure [39] suggests that the closely related terminal complement proteins would have the CDC orientation, but there are currently no other data available for a more definitive conclusion on perforin.

Our findings highlight a critical role of the interface between the top of TMH2 and the surrounding region in controlling sheet opening. The results of the constrained fitting suggest that a key trigger for the conformational change includes displacement of the HTH motif away from the bend in the sheet. Highly conserved glycine residues [14] adjacent to the HTH motif may provide the hinge point for this motion (Fig. 3D). Consistent with this model, mutation of the equivalent glycine residues in a CDC prevents oligomerisation [40]. It is notable that the HTH packs against the top of TMH2, suggesting that interactions between these two regions may govern unlocking of the bent conformation (Fig. 3D). After sheet unbending, we propose



**Figure 6. Schematic diagram of sheet opening in pleurotolysin pore formation.** (A) PlyA and B monomer structures with coloured shapes indicating the major elements used in flexible fitting. (B) and (C) show the two types of prepore structures found, which can be considered as early and later stages of sheet opening. The top 20 models (pink for the MACPF  $\beta$ -sheet and grey for HTH region) and best scoring model (red) in the pleurotolysin prepore maps are shown. (B) TMH2 helix lock, in which the best scoring model has a rotation angle ( $\beta$ -sheet opening) of  $37^\circ$  relative to the monomer structure. (C) TMH2 strand lock, in which the best scoring model has a rotation angle of  $49^\circ$  relative to the monomer. (D) Pore model with  $70^\circ$  open sheet and membrane inserted TMH regions.

doi:10.1371/journal.pbio.1002049.g006

that membrane insertion and pore formation follow a top down, zippering mechanism with the barrel assembling towards the membrane surface, energetically driven by refolding of the TMH regions. This mechanism would also minimize the free energy cost of inserting naked hairpins with unsatisfied hydrogen bond potential into the membrane.

Analysis of intermediate prepore structures provides the basis for a molecular movie (S1 Movie) that illustrates a possible trajectory of the core  $\beta$ -sheet in a MACPF protein unbending from the soluble monomer conformation to the transmembrane pore (Fig. 6). The pore structure shows that Ply shares some features with CDCs, in particular the orientation of monomers and opening of the molecule to release the TMH regions. On the other hand it resembles perforin regarding its longer TMH regions that refold into a  $\sim 100$ -Å-long  $\beta$ -barrel that reaches down through the membrane without any collapse of the molecule.

This work provides new insights into the assembly of a two-component MACPF/CDC family member, suggesting a basis for the study of more complex assembly systems such as the complement MAC. Furthermore, the intermediate structures of the MACPF/CDC domain during its refolding into the  $\beta$ -barrel pore establish a structural paradigm for the transition of the prepore to pore, which is likely to be conserved across the MACPF/CDC protein family.

## Materials and Methods

### Protein Production and Crystallography

PlyA and PlyB were expressed in *Escherichia coli*. PlyA was expressed as a soluble protein; PlyB required refolding. Crystals of selenomethionine-labelled PlyA were grown in 50 mM Na citrate (pH 5.6), 12% (w/v) PEG3350, 0.2 M  $\text{MgSO}_4$ , and the structure determined using single-wavelength anomalous dispersion (S1 Table). Crystals of PlyB were grown in 0.2 M  $\text{NH}_4\text{Ac}$ , 0.1 M Na citrate (pH 5.0) and 30% (w/v) PEG8000. In addition to two native datasets, diffraction data were also collected from four different heavy atom derivatives (S2 Table). The PlyB structure was determined using multiple isomorphous replacement with anomalous scattering. The PlyB TMH1 lock structure was determined by molecular replacement using the structure of wild type PlyB.

## Electron Microscopy and Atomic Structure Fitting

Pleurotolysin WT pores and engineered disulphide bond oligomers assembled on sphingomyelin/cholesterol liposomes were imaged by negative stain and cryo-EM, and sorted according to diameter and symmetry by multivariate statistical analysis and multireference alignment. Three-dimensional reconstructions were obtained by a combination of angular reconstitution and projection matching for tridecameric pores and prepores. The PlyA and PlyB crystal structures and a model for a 52-stranded  $\beta$ -barrel [31] were fitted into the electron microscopy maps using a combination of manual fitting, rigid body refinement, and flexible fitting. The degree of unbending of the MACPF  $\beta$ -sheet in the prepore intermediates was estimated by using a multistep procedure (S8 Fig.) to generate a library of sheet conformations in angular sweeps, and selecting the top 20 best fitting ones for each prepore map. The goodness-of-fit of the central asymmetric unit was assessed locally using SCCC (S5 Fig.) [33,34].

Detailed protocols are available in [S1 Methods](#).

## Supporting Information

**S1 Fig. TMH1 unravels to form a membrane penetrating amphipathic  $\beta$ -hairpin.** (A) The fluorescence emission of NBD covalently attached to the individual cysteine substituted residues along the putative TMH1 hairpin is shown in the absence and presence of cholesterol-rich liposomes. Dark grey, residues in a hydrophobic environment; light grey, residues facing the barrel lumen; black bars, proposed  $\beta$ -hairpin residues; and the asterisks denote mutant proteins that could not be produced in functional form. (B) Collisional quenching of the NBD probe by a C12 doxyl group positioned on the acyl chain of phosphatidylcholine. The fluorescence data from TMH1 show an identical pattern to that observed previously for both membrane-inserted TMHs of perfringolysin O where two residues immediately prior to the  $\beta$ -turn are membrane associated and then the alternating pattern of membrane inserted residues is offset by one residue in the second strand [41]. The position of proline residue 137 in TMH1 is consistent with the suggested position of the  $\beta$ -turn. (C) Schematic of the TMH1  $\beta$ -hairpin with strand 1, predicted hairpin, then strand 2. Orange boxes are residues that are in contact with the membrane, blue boxes are residues facing the lumen of the barrel, black boxes are the  $\beta$ -turn residues, and white boxes represent mutant proteins that could not be produced in functional form. (TIF)

**S2 Fig. Fitting of different  $\beta$ -barrel models into the pleurotolysin pore reconstruction.** The diameter and the height of  $\beta$ -barrels depend on the number of strands ( $n$ ) and shear ( $S$ ) of the  $\beta$ -hairpins relative to the barrel axis. (A)  $S = 0$  ( $0^\circ$  tilt) corresponds to a  $\beta$ -barrel with strands parallel to the barrel axis. (B)  $S = n/2$  ( $20^\circ$  tilt) model. For a pleurotolysin pore with 13-fold symmetry and each subunit contributing two hairpins the shear  $S$  is 26 [31]. (C)  $S = n$  ( $36^\circ$  tilt) and higher correspond to greater tilts of the  $\beta$ -strands and are typically associated with small transmembrane  $\beta$ -barrels. (D) Fitting of pleurotolysin  $\beta$ -barrel models into the cryo-EM map of the pleurotolysin pore revealed that the barrel height in models with  $S = n$  (purple) and higher shear is not sufficient to cross the membrane. Of the three possible models,  $S = n/2$  (blue) fits significantly better (cross-correlation 0.90) than  $S = 0$  (green; cross-correlation 0.73) and  $S = n$  (purple; cross-correlation 0.74). Coloured bars indicate the bottom of each corresponding barrel fitted in the pore map. (TIF)

**S3 Fig. Comparison of PlyA orientation with actinoporins and domain 4 in CDCs.** Two crystal forms of PlyA revealed two different V-shaped dimers. (A) An N-terminus-N-terminus dimer (N-dimer) of PlyA, which when fitted into EM density is positioned on the membrane

similarly to actinoporins [29]; (B) Sticholysin II [29], phosphocholine is shown in space filling format (grey spheres). (C) Superposition of PlyA and sticholysin II structures. (D) C-terminus-C-terminus dimer (C-dimer) of PlyA, which fits into EM density in the orientation corresponding to that of the C2 domain of perforin [18]. (E) Domain 4 of perfringolysin O [42] with the cholesterol binding residues shown as cyan spheres. Both PlyA dimers (A and D) could be fitted into the pleurotolysin pore density, with a slightly better fit for the N-dimer (cross-correlation 0.74 versus 0.71 for C-dimer).

(TIF)

**S4 Fig. Effect of His tags on the binding of PlyA to membranes and haemolysis.** To test the subunit orientation in the fit, we mutated a residue (W6E) located in the centre of the PlyA N-N dimer interface (Fig. 4A). (A) Western blot analysis of a red blood cell (RBC) pull-down assay by PlyA. This assay shows a reduced affinity of W6E PlyA compared to WT PlyA. (B) The W6E PlyA variant has significantly reduced (90–120-fold, duplicate experiment) pleurotolysin lytic activity.

(TIF)

**S5 Fig. Local assessment of the quality of fit for a pore subunit.** The local quality of fit in each region (defined as in Fig. 1) was assessed using SCCC [34]. The pseudo-atomic pore subunit model is shown colour-coded according to the SCCC score from blue to red (see colour key).

(TIF)

**S6 Fig. Concentration and time dependence of haemolysis activity of wild type and mutant pleurotolysin.** (A) Titration based haemolysis assay of wild type pleurotolysin and PlyB<sup>C487A</sup> mutant. Titration (B) and time course haemolysis assays (C) of TMH1 lock mutant, TMH2 helix lock mutant (D, E) and TMH2 strand lock mutant (F, G). In (B), (D), (F) all disulphide locked mutants are compared to the PlyB<sup>C487A</sup> background in reducing and non-reducing conditions. All the locked mutants lack haemolytic activity in non-reducing conditions (empty squares) compared to reducing conditions (black squares). Titration assay shows that reduced TMH1 lock (B) and TMH2 strand lock (F) mutants recovered the same activity as the background mutant, whereas the TMH2 helix lock (D) recovered 50% lysis activity at 3-fold higher concentration than the background mutant. In the time-course haemolysis assay of prepores (C, E, G), red blood cells were incubated with disulphide locked mutants to form prepores on the surface of the cells. Unbound PlyA and PlyB were then washed off. Conversion of prepores to pores was monitored (at 620 nm) in real time by decrease of light scattering resulting from lysis of the cells (black circles), compared to either pre-lysed cells (empty triangles) or prepore loaded RBC in non-reducing conditions (empty circles). Note that due to the long time required to lyse the cells there is settling of the RBC in the negative control sample. These data show that all assembled pleurotolysin prepores can convert to pores under reducing conditions.

(TIF)

**S7 Fig. Overlay of PlyB crystal structure with PlyB<sup>F138C,H221C</sup> (TMH1 lock, purple).** RMSD 0.47 Å (463 Cα). The disulphide bond is shown as sticks, with the  $2F_o - F_c$  electron density contoured at  $1\sigma$  level shown in blue (inset).

(TIF)

**S8 Fig. Flowchart of PlyB conformational analysis for the three prepore maps.** The input models (A) correspond to the PlyB crystal structure and the pore subunit model (based on density fitting of PlyA and PlyB crystal structures and the  $\beta$ -barrel into the pore density map). These were used as initial conformations for two series of angular sweeps of the MACPF  $\beta$ -sheet (B), generated in steps of 0.5 Å translation and 1° rotation around the centre of mass. Each sampled conformation of the MACPF  $\beta$ -sheet was then combined with the PlyB

monomer structure or with the pore model, resulting in a library of  $\sim 4,600$  models (C). Each model was then rigidly fitted into each of the prepore maps (D) and the goodness-of-fit of the MACPF  $\beta$ -sheet was assessed using SCCC [34]. The top 20 models were then re-ranked (E) based on the DOPE score [43].

(TIF)

**S9 Fig. Resolution curves for cryo-EM reconstructions.**

(TIF)

**S1 Movie. Movie of the trajectory from soluble monomers (crystal structures) through to the final assembled pore shows the opening of the  $\beta$ -sheet.** The cryo-EM reconstruction of the first intermediate (TMH2 helix lock) is shown with the model. This is followed by the cryo-EM reconstruction of the TMH1 lock with the best-fit model. The final images show the 11 Å resolution cryo-EM reconstruction of the assembled MACPF pore.

(MOV)

**S1 Methods. Detailed protocols not described in the Materials and Methods section.**

(DOCX)

**S1 Table. Data collection and statistics for PlyA crystals.**

(DOCX)

**S2 Table. Data collection and statistics for PlyB crystals.**

(DOCX)

**S3 Table. Data collection and statistics for PlyB<sup>C487A,F138C,H221C</sup> crystals.**

(DOCX)

**S4 Table. Domain-orientation score [30] (DOS, translation, and rotation parameters) for 20 best fitting sheet conformations for TMH2 helix lock, TMH1 lock, and TMH2 strand lock pleurotolysin prepores (Fig. 6).**

(DOCX)

## Acknowledgments

We thank D. Houldershaw and R. Westlake for computing support, L. Wang for EM support, and E.V. Orlova, D.K. Clare, and A.P. Pandurangan for discussion. We thank M.J. de Veer for the gift of sheep red blood cells. We acknowledge C. Rosado and W-T. Kan for assistance with vector constructions and protein production, technical support from Australian Synchrotron staff (beamline MX-1 and MX-2), and the Monash Platforms (Protein Production, Biological Electron Microscopy, Proteomics and Crystallography).

## Author Contributions

Conceived and designed the experiments: SCK BAS RKT JCW HRS MAD. Performed the experiments: NL SCK IF RHPL TTC BAS OK DT SME TH KO EMH MT. Analyzed the data: NL SCK IF CFR KO RKT JCW MT HRS MAD. Contributed reagents/materials/analysis tools: IV JAT. Wrote the paper: NL SCK IF RKT JCW MT HRS MAD.

## References

1. Degiacomi MT, Iacovache I, Pernot L, Chami M, Kudryashev M, et al. (2013) Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nat Chem Biol* 9: 623–629. doi: [10.1038/nchembio.1312](https://doi.org/10.1038/nchembio.1312) PMID: [23912165](https://pubmed.ncbi.nlm.nih.gov/23912165/)



2. Iacovache I, Bischoffberger M, van der Goot FG (2010) Structure and assembly of pore-forming proteins. *Curr Opin Struct Biol* 20: 241–246. doi: [10.1016/j.sbi.2010.01.013](https://doi.org/10.1016/j.sbi.2010.01.013) PMID: [20172710](https://pubmed.ncbi.nlm.nih.gov/20172710/)
3. Meusch D, Gatsogiannis C, Efremov RG, Lang AE, Hofnagel O, et al. (2014) Mechanism of Tc toxin action revealed in molecular detail. *Nature* 508: 61–65. doi: [10.1038/nature13015](https://doi.org/10.1038/nature13015) PMID: [24572368](https://pubmed.ncbi.nlm.nih.gov/24572368/)
4. Unno H, Goda S, Hatakeyama T (2014) Hemolytic lectin CEL-III heptamerizes via a large structural transition from alpha-helices to a beta-barrel during the transmembrane pore formation process. *J Biol Chem* 289: 12805–12812. doi: [10.1074/jbc.M113.541896](https://doi.org/10.1074/jbc.M113.541896) PMID: [24652284](https://pubmed.ncbi.nlm.nih.gov/24652284/)
5. Tschopp J, Masson D, Stanley KK (1986) Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytolysis. *Nature* 322: 831–834. PMID: [2427956](https://pubmed.ncbi.nlm.nih.gov/2427956/)
6. Kadota K, Ishino T, Matsuyama T, Chinzei Y, Yuda M (2004) Essential role of membrane-attack protein in malarial transmission to mosquito host. *Proc Natl Acad Sci U S A* 101: 16310–16315. PMID: [15520375](https://pubmed.ncbi.nlm.nih.gov/15520375/)
7. Kafsack BF, Pena JD, Coppens I, Ravindran S, Boothroyd JC, et al. (2009) Rapid membrane disruption by a perforin-like protein facilitates parasite exit from host cells. *Science* 323: 530–533. doi: [10.1126/science.1165740](https://doi.org/10.1126/science.1165740) PMID: [19095897](https://pubmed.ncbi.nlm.nih.gov/19095897/)
8. Birmingham CL, Canadien V, Kaniuk NA, Steinberg BE, Higgins DE, et al. (2008) Listeriolysin O allows *Listeria monocytogenes* replication in macrophage vacuoles. *Nature* 451: 350–354. doi: [10.1038/nature06479](https://doi.org/10.1038/nature06479) PMID: [18202661](https://pubmed.ncbi.nlm.nih.gov/18202661/)
9. Morita-Yamamuro C, Tsutsui T, Sato M, Yoshioka H, Tamaoki M, et al. (2005) The Arabidopsis gene CAD1 controls programmed cell death in the plant immune system and encodes a protein containing a MACPF domain. *Plant Cell Physiol* 46: 902–912. PMID: [15799997](https://pubmed.ncbi.nlm.nih.gov/15799997/)
10. Henstridge MA, Johnson TK, Warr CG, Whisstock JC (2014) Trunk cleavage is essential for *Drosophila* terminal patterning and can occur independently of Torso-like. *Nat Commun* 5: 3419. doi: [10.1038/ncomms4419](https://doi.org/10.1038/ncomms4419) PMID: [24584029](https://pubmed.ncbi.nlm.nih.gov/24584029/)
11. Stevens LM, Frohnhofer HG, Klingler M, Nusslein-Volhard C (1990) Localized requirement for torso-like expression in follicle cells for development of terminal anlagen of the *Drosophila* embryo. *Nature* 346: 660–663. PMID: [2385293](https://pubmed.ncbi.nlm.nih.gov/2385293/)
12. Zheng C, Heintz N, Hatten ME (1996) CNS gene encoding astrotactin, which supports neuronal migration along glial fibers. *Science* 272: 417–419. PMID: [8602532](https://pubmed.ncbi.nlm.nih.gov/8602532/)
13. Hadders MA, Beringer DX, Gros P (2007) Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* 317: 1552–1554. PMID: [17872444](https://pubmed.ncbi.nlm.nih.gov/17872444/)
14. Rosado CJ, Buckle AM, Law RH, Butcher RE, Kan WT, et al. (2007) A common fold mediates vertebrate defense and bacterial attack. *Science* 317: 1548–1551. PMID: [17717151](https://pubmed.ncbi.nlm.nih.gov/17717151/)
15. Slade DJ, Lovelace LL, Chruszcz M, Minor W, Lebioda L, et al. (2008) Crystal structure of the MACPF domain of human complement protein C8 alpha in complex with the C8 gamma subunit. *J Mol Biol* 379: 331–342. doi: [10.1016/j.jmb.2008.03.061](https://doi.org/10.1016/j.jmb.2008.03.061) PMID: [18440555](https://pubmed.ncbi.nlm.nih.gov/18440555/)
16. Rossjohn J, Feil SC, McKinsty WJ, Tweten RK, Parker MW (1997) Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form. *Cell* 89: 685–692. PMID: [9182756](https://pubmed.ncbi.nlm.nih.gov/9182756/)
17. Tilley SJ, Orlova EV, Gilbert RJ, Andrew PW, Saibil HR (2005) Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* 121: 247–256. PMID: [15851031](https://pubmed.ncbi.nlm.nih.gov/15851031/)
18. Law RH, Lukyanova N, Voskoboinik I, Caradoc-Davies TT, Baran K, et al. (2010) The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* 468: 447–451. doi: [10.1038/nature09518](https://doi.org/10.1038/nature09518) PMID: [21037563](https://pubmed.ncbi.nlm.nih.gov/21037563/)
19. Hotze EM, Tweten RK (2012) Membrane assembly of the cholesterol-dependent cytolysin pore complex. *Biochimica et biophysica acta* 1818: 1028–1038. doi: [10.1016/j.bbamem.2011.07.036](https://doi.org/10.1016/j.bbamem.2011.07.036) PMID: [21835159](https://pubmed.ncbi.nlm.nih.gov/21835159/)
20. Ramachandran R, Tweten RK, Johnson AE (2005) The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation. *Proc Natl Acad Sci U S A* 102: 7139–7144. PMID: [15878993](https://pubmed.ncbi.nlm.nih.gov/15878993/)
21. Shatursky O, Heuck AP, Shepard LA, Rossjohn J, Parker MW, et al. (1999) The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* 99: 293–299. PMID: [10555145](https://pubmed.ncbi.nlm.nih.gov/10555145/)
22. Czajkowsky DM, Hotze EM, Shao Z, Tweten RK (2004) Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. *Embo J* 23: 3206–3215. PMID: [15297878](https://pubmed.ncbi.nlm.nih.gov/15297878/)
23. Aleshin AE, DiScipio RG, Stec B, Liddington RC (2012) Crystal structure of C5b-6 suggests structural basis for priming assembly of the membrane attack complex. *J Biol Chem* 287: 19642–19652. doi: [10.1074/jbc.M112.361121](https://doi.org/10.1074/jbc.M112.361121) PMID: [22500023](https://pubmed.ncbi.nlm.nih.gov/22500023/)

24. Aleshin AE, Schraufstatter IU, Stec B, Bankston LA, Liddington RC, et al. (2012) Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC). *J Biol Chem* 287: 10210–10222. doi: [10.1074/jbc.M111.327809](https://doi.org/10.1074/jbc.M111.327809) PMID: [22267737](https://pubmed.ncbi.nlm.nih.gov/22267737/)
25. Hadders MA, Bubeck D, Roversi P, Hakobyan S, Forneris F, et al. (2012) Assembly and regulation of the membrane attack complex based on structures of C5b6 and sC5b9. *Cell Rep* 1: 200–207. doi: [10.1016/j.celrep.2012.02.003](https://doi.org/10.1016/j.celrep.2012.02.003) PMID: [22832194](https://pubmed.ncbi.nlm.nih.gov/22832194/)
26. Sakurai N, Kaneko J, Kamio Y, Tomita T (2004) Cloning, expression, and pore-forming properties of mature and precursor forms of pleurotolysin, a sphingomyelin-specific two-component cytolysin from the edible mushroom *Pleurotus ostreatus*. *Biochim Biophys Acta* 1679: 65–73. PMID: [15245918](https://pubmed.ncbi.nlm.nih.gov/15245918/)
27. Tomita T, Noguchi K, Mimuro H, Ukaji F, Ito K, et al. (2004) Pleurotolysin, a novel sphingomyelin-specific two-component cytolysin from the edible mushroom *Pleurotus ostreatus*, assembles into a trans-membrane pore complex. *J Biol Chem* 279: 26975–26982. PMID: [15084605](https://pubmed.ncbi.nlm.nih.gov/15084605/)
28. Ota K, Leonardi A, Mikelj M, Skočaj M, Wohlschlager T, et al. (2013) Membrane cholesterol and sphingomyelin, and osteolysin A are obligatory for pore-formation by a MACPF/CDC-like pore-forming protein, pleurotolysin B. *Biochimie* 95: 1855–1864. doi: [10.1016/j.biochi.2013.06.012](https://doi.org/10.1016/j.biochi.2013.06.012) PMID: [23806422](https://pubmed.ncbi.nlm.nih.gov/23806422/)
29. Mancheno JM, Martin-Benito J, Martinez-Ripoll M, Gavilanes JG, Hermoso JA (2003) Crystal and electron microscopy structures of sticholysin II actinoporin reveal insights into the mechanism of membrane pore formation. *Structure* 11: 1319–1328. PMID: [14604522](https://pubmed.ncbi.nlm.nih.gov/14604522/)
30. Topf M, Lasker K, Webb B, Wolfson H, Chiu W, et al. (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* 16: 295–307. doi: [10.1016/j.str.2007.11.016](https://doi.org/10.1016/j.str.2007.11.016) PMID: [18275820](https://pubmed.ncbi.nlm.nih.gov/18275820/)
31. Reboul CF, Mahmood K, Whisstock JC, Dunstone MA (2012) Predicting giant transmembrane beta-barrel architecture. *Bioinformatics* 28: 1299–1302. doi: [10.1093/bioinformatics/bts152](https://doi.org/10.1093/bioinformatics/bts152) PMID: [22467914](https://pubmed.ncbi.nlm.nih.gov/22467914/)
32. Murzin AG, Lesk AM, Chothia C (1994) Principles determining the structure of beta-sheet barrels in proteins. I. A theoretical analysis. *J Mol Biol* 236: 1369–1381. PMID: [8126726](https://pubmed.ncbi.nlm.nih.gov/8126726/)
33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612. PMID: [15264254](https://pubmed.ncbi.nlm.nih.gov/15264254/)
34. Pandurangan AP, Shakeel S, Butcher SJ, Topf M (2014) Combined approaches to flexible fitting and assessment in virus capsids undergoing conformational change. *J Struct Biol* 185: 427–439. doi: [10.1016/j.jsb.2013.12.003](https://doi.org/10.1016/j.jsb.2013.12.003) PMID: [24333899](https://pubmed.ncbi.nlm.nih.gov/24333899/)
35. Hotze EM, Wilson-Kubalek EM, Rossjohn J, Parker MW, Johnson AE, et al. (2001) Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane beta-sheet from a prepore intermediate. *J Biol Chem* 276: 8261–8268. PMID: [11102453](https://pubmed.ncbi.nlm.nih.gov/11102453/)
36. Vasishtan D, Topf M (2011) Scoring functions for cryoEM density fitting. *J Struct Biol* 174: 333–343. doi: [10.1016/j.jsb.2011.01.012](https://doi.org/10.1016/j.jsb.2011.01.012) PMID: [21296161](https://pubmed.ncbi.nlm.nih.gov/21296161/)
37. Lockert DH, Kaufman KM, Chang C-P, Hüsler T, Sodetz JM, et al. (1995) Identity of the segment of human complement C8 recognized by complement regulatory protein CD59. *J Biol Chem* 270: 19723–19728. PMID: [7544344](https://pubmed.ncbi.nlm.nih.gov/7544344/)
38. Schuerch DW, Wilson-Kubalek EM, Tweten RK (2005) Molecular basis of listeriolysin O pH dependence. *Proc Natl Acad Sci U S A* 102: 12537–12542. PMID: [16105950](https://pubmed.ncbi.nlm.nih.gov/16105950/)
39. Lovelace LL, Cooper CL, Sodetz JM, Lebiada L (2011) Structure of human C8 protein provides mechanistic insight into membrane pore formation by complement. *J Biol Chem* 286: 17585–17592. doi: [10.1074/jbc.M111.219766](https://doi.org/10.1074/jbc.M111.219766) PMID: [21454577](https://pubmed.ncbi.nlm.nih.gov/21454577/)
40. Ramachandran R, Tweten RK, Johnson AE (2004) Membrane-dependent conformational changes initiate cholesterol-dependent cytolysin oligomerization and intersubunit beta-strand alignment. *Nat Struct Mol Biol* 11: 697–705. PMID: [15235590](https://pubmed.ncbi.nlm.nih.gov/15235590/)
41. Shepard LA, Heuck AP, Hamman BD, Rossjohn J, Parker MW, et al. (1998) Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an alpha-helical to beta-sheet transition identified by fluorescence spectroscopy. *Biochemistry* 37: 14563–14574. PMID: [9772185](https://pubmed.ncbi.nlm.nih.gov/9772185/)
42. Farrand AJ, LaChapelle S, Hotze EM, Johnson AE, Tweten RK (2010) Only two amino acids are essential for cytolytic toxin recognition of cholesterol at the membrane surface. *Proc Natl Acad Sci U S A* 107: 4341–4346. doi: [10.1073/pnas.0911581107](https://doi.org/10.1073/pnas.0911581107) PMID: [20145114](https://pubmed.ncbi.nlm.nih.gov/20145114/)
43. Shen MY, Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15: 2507–2524. PMID: [17075131](https://pubmed.ncbi.nlm.nih.gov/17075131/)