Structural and Functional Investigation of the Paraquat Inducible Pathway.

by

Benjamin Francis Cooper

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

College of Life and Environmental Sciences

School of Biosciences

University of Birmingham

October 2022

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Originally proposed to facilitate cell invasion, Mammalian Cell Entry (MCE) domains have subsequently been implicated in the transport of hydrophobic substances. Indeed, multiple phylogenetic and biochemical studies have proposed members of the MCE protein superfamily to traffic lipid species. The *E. coli* Mla pathway, itself containing the MCE protein MlaD, presents the first proteinaceous system facilitating intermembrane GPL trafficking. *E. coli* contains two further MCE proteins, YebT and PqiB, each residing in unique operons which have become of principal interest as putative Gram-negative GPL transporters.

Here we focus upon the Pqi pathway, a tripartite macromolecular complex, encoded by a single PqiA-C operon, anticipated to span the periplasm allowing the transport of lipid species between diderm membranes. We present X-ray structures of the native PqiC lipoprotein and that of a soluble PqiC$^{17-187}$ construct revealing a potentially biologically relevant octameric ring assembly and high-resolution side chain information respectively. We also provide the first structure of native PqiB, via cryo-EM, allowing comparison to the existing soluble PqiB$^{39-546}$ structure and demonstrating flexibility within the pore lining loop regions. Finally, we utilise phenotypic complementation assays to identify residues within PqiA and PqiB essential for Pqi function and provide evidence for a protonation dependent PqiA mechanism.

Acknowledgements

First, I wish to thank my PhD supervisor, Dr Timothy Knowles, for his supervision, guidance and provision of a stimulating project. I've thoroughly enjoyed your friendly and approachable mentoring style, even if it meant dragging you around Cannock chase on a bike…

I also owe great thanks to Prof Andrew Lovering, for offering his vast structural insight into my project, without whom the similarity between PqiA and PfMATE would not have been drawn. I will certainly miss your infectious enthusiasm and the incredible science you disseminate to the 7th floor.

Thank you to the 7th floor of Biosciences, past and present, for providing such an enjoyable working environment. I'm sure I'll never experience anything quite like it again and wish you all the best for the future. Particular thanks go to Dr Gareth Hughes, for his supervision at the start of my PhD and friendship throughout the remainder. Dr Peter Wotherspoon, for providing knowledge and humour, in equal parts, throughout my PhD. Dr Simon Caulton and Rebecca Parr, for crystal fishing and providing the "crystallography help desk" when I required assistance solving my structures. Pooja Sridhar, for assistance in admin tasks and organisation of the 7th floor.

Beyond the 7th floor I wish to thank Dr Jack Bryant for guidance in generating knockouts and performing complementation assays. Dr Doug Browning, for assistance with cloning and Dr Aneika Leney and Dr Jedd Bellamy-Carter, for performing mass spectrometry.

Special thanks go to Dr T J Ragan and particularly Dr Christos Savva of the Midlands Regional cryo-EM facility for educating me in all aspects of cryo-EM and promoting autonomy. Similarly, thanks to Dr Deniz Ugurlar of the ThermoFisher centre for electron microscopy for screening, optimising and ultimately collecting a structure yielding dataset.

I am immeasurably grateful to my family, Mum, Dad and Michael, for their continual support and advice. You have always enabled me to pursue my own aims, despite any sacrifice to yourselves, thus I can say without doubt that I would not have made it this far without you. Thank you for everything you have done for me.

Last, but certainly not least, thank you to Giedrė Ratkevičiūtė. Your appetite for knowledge and desire for perfection have pushed me to achieve more myself. I have aways been able to rely on your love and support both in and outside the lab. I am exceedingly fortunate to have met someone so incredible and talented during my PhD.

The work comprising this thesis was enabled by a Midlands Integrative Biosciences Training Programme studentship (grant no. BB/M01116X/1), jointly funded by the BBSRC and University of Birmingham, as well as a BBSRC Research grant (grant no. BB/S017283/1).

Table of Contents

Abstract i

Acknowledgements ii

Table of Contents iv

List of Figures ix

List of Tables xii

List of Abbreviations xiii

CHAPTER 1: GENERAL INTRODUCTION 1

Structure and organisation of the Gram-negative cell envelope 2

Architecture of the Gram-negative cell envelope 2

Periplasmic Transport of Gram-negative Cell Envelope Components 6

Outer Membrane Proteins 6

Lipopolysaccharide 8

Lipoproteins 9

Glycerophospholipids 11

Maintaining Lipid Asymmetry 20

The Maintenance of Lipid Asymmetry (Mla) pathway 22

MlaFEDB 24

MlaA 28

MlaC 31

Resolving transport directionality 32

Alternative Putative Gram-negative GPL Transport Systems 37

Tol-Pal 37

PbgA 39

YhdP and "AsmA-like" proteins 41

MCE/MlaD Domains 44

The Advent of Mammalian Cell Entry (MCE) *44*

MCE Proteins *46*

MCE Proteins in Lipid Transport? *47*

E. coli MCE proteins 51

YebST (LetAB) *52*

PqiABC *55*

Resolving function from structure *57*

Research Aims 58

CHAPTER 2: MATERIALS AND METHODS 60

Construct design, Expression and Purification 61

Culture Media and Antibiotics *61*

Bacterial Strains *62*

Constructs *63*

Transformation *63*

Expression *64*

Purification – Soluble Constructs *64*

Purification – Membrane Constructs *66*

SDS-PAGE *68*

Western Blotting *68*

Q5 Site Directed Mutagenesis (SDM) *69*

Q5 SDM Primers *71*

Sequencing *72*

Quantification of protein and DNA concentrations *73*

Molecular Biology Techniques 74

Plasmids *74*

Molecular Biology Primers *75*

Polymerase Chain Reaction *75*

Agarose Gel Electrophoresis *77*

Electroporation *78*

Datsenko and Wanner Gene Inactivation *78*

P1 Transduction *81*

Removal of the aph cassette *82*

Phenotypic Complementation Assays *83*

STRUCTURAL DETERMINATION 84

X-ray Crystallography *84*

Cryo-EM *86*

CHAPTER 3: X-RAY CRYSTALLOGRAPHY OF PQIC 88

INTRODUCTION 89

Introduction to X-ray Crystallography *89*

Chapter Aims *96*

RESULTS 98

Expression and Purification of PqiC FL *98*

Expression and Purification of PqiC$^{17-187}$ *98*

Structure Determination of PqiC FL *100*

Structure Determination of PqiC$^{17-187}$ *112*

DISCUSSION 120

PqiC FL *120*

PqiC$^{17-187}$ *125*

Summary *129*

CHAPTER 4: CRYOGENIC ELECTRON MICROSCOPY OF PQIABC 130

INTRODUCTION 131

Introduction to Cryo-EM *131*

Sample Preparation *132*

Image Formation *135*

Data Processing – Single Particle Averaging *139*

Chapter Aims *146*

RESULTS 147
Expression and Purification of PqiABC *147*
Initial cryo-EM Screening of PqiABC *148*
ThermoFisher Collaboration *151*
Processing of the ThermoFisher PqiABC Dataset *154*
PqiB Core Model Building *166*
DISCUSSION 174
C1 Refinement and Processing *174*
C6 Refinement and Processing *179*
PqiB Core Model Building *183*
Summary *187*
CHAPTER 5: *IN-VIVO* PHENOTYPIC COMPLEMENTATION ASSAYS 188
INTRODUCTION 189
Phenotypic Complementation Assays *189*
Chapter Aims *190*
RESULTS 192
PqiABC Knockout Generation *192*
Pqi Mutation Selection *194*
Phenotypic Complementation Assay Mutations and Assays *205*
DISCUSSION 213
PqiA *213*
PqiB *217*
PqiC *219*
Summary *220*
CHAPTER 6: CONCLUSIONS AND FUTURE WORK 221
CONCLUSIONS AND FUTURE WORK 222
Advancement of Pqi Structural Understanding *222*
Advancement of Pqi Functional Understanding *226*

CONCLUDING REMARKS 231

APPENDIX 232

Appendix. 1 – Construct coding insert sequences. *232*

Appendix. 2 – PqiC X-ray Data Collection Parameters. *238*

Appendix. 3 – PqiC FL IMAC Elution Fraction SDS-PAGE. *238*

Appendix. 4 – PqiC$^{17-187}$ IMAC Elution Fraction SDS-PAGE. *239*

Appendix. 5 – PqiABC IMAC Elution Fraction SDS-PAGE. *239*

Appendix. 6 – Cryo-EM FSC curves. *240*

BIBLIOGRAPHY 241

List of Figures

FIG. 1 – ARCHITECTURE OF THE GRAM-NEGATIVE CELL ENVELOPE. 3

FIG. 2 – OUTER MEMBRANE BIOGENESIS TRANSPORT PATHWAYS. 10

FIG. 3 – *E. COLI* GPL BIOSYNTHESIS. 12

FIG. 4 – PROPOSED GRAM-NEGATIVE INTERMEMBRANE GPL TRANSPORT MODELS. 16

FIG. 5 – ENERGY PROFILES FOR INTERMEMBRANE GPL TRANSPORT. 17

FIG. 6 – PUTATIVE INTERMEMBRANE GPL TRANSPORT PATHWAYS. 19

FIG. 7 – REMOVAL OF MIS-LOCALISED OM GPLS VIA PLDA AND PAGP. 21

FIG. 8 – THE MAINTENANCE OF LIPID ASYMMETRY (MLA) PATHWAY. 23

FIG. 9 – GPL BINDING SITES WITHIN THE *E. COLI* MLAFEDB COMPLEX. 26

FIG. 10 – THE STRUCTURE OF MLAA. 30

FIG. 11 – THE TOL-PAL AND PBGA PATHWAYS. 41

FIG. 12 – THE "ASMA-LIKE" PROTEIN FAMILY AND HOMOLOGUES. 43

FIG. 13 – THE MCE1, MCE1$^{105-307}$ AND MCE1$^{163-307}$ PROTEINS AND RESPECTIVE DNA SEGMENTS. 45

FIG. 14 – CORE ARCHITECTURES OF BACTERIAL MCE PROTEINS. 47

FIG. 15 – ACTINOBACTERIAL *MCE* OPERON NOMENCLATURE. 48

FIG. 16 – *E. COLI* MCE PROTEIN CONTAINING OPERONS. 51

FIG. 17 – STRUCTURES OF YEBS AND YEBT. 54

FIG. 18 – STRUCTURES OF PQIA, PQIB AND PQIC. 56

FIG. 19 – DATSENKO AND WANNER GENE INACTIVATION. 79

FIG. 20 – OVERVIEW OF THE CRYSTALLOGRAPHIC METHOD 90

FIG. 21 – THE PROCESS OF VAPOUR DIFFUSION. 92

FIG. 22 – BRAGG'S LAW OF DIFFRACTION. 94

FIG. 23 – PURIFICATION OF PQIC FL AND PQIC$^{17-187}$. 99

FIG. 24 – PQIC FL CRYSTALS. 101

FIG. 25 – OVERLAY OF CFL1 AND CFL2 STRUCTURES. 103

FIG. 26 – ARCHITECTURE OF THE PQIC FL OCTAMER. 104

FIG. 27 – MOLECULAR LIPOPHILICITY POTENTIAL OF PQIC FL OCTAMER SURFACE. 106

Fig. 28 – Lipophilic Residues of Interest Within PqiC FL. 107

Fig. 29 – Coulombic Electrostatic Potential of the PqiC FL Octamer Surface. 109

Fig. 30 – Charged Residues of Interest Within PqiC FL. 110

Fig. 31 – Sequence Conservation of PqiC FL. 111

Fig. 32 – $PqiC^{17-187}$ Crystals. 113

Fig. 33 – $PqiC^{17-187}$ Asymmetric Unit. 115

Fig. 34 – Comparison the CFL2 and Csol Asymmetric Unit Geometry. 116

Fig. 35 – Alignment of CFL2 and Csol Monomers. 117

Fig. 36 – Comparison of CFL2 and Csol Electron Density. 118

Fig. 37 – Structural Alignment of the CFL2 Structure with Dali Server Hits. 121

Fig. 38 – Asymmetric Unit Arrangements of PqiC Homologous 122

Fig. 39 – Crystal Packing of $PqiC^{17-187}$. 126

Fig. 40 – Packing of the b1- b2 Loop in the CFL2 and Csol Structures. 128

Fig. 41– Growth of 3D Electron Microscopy derived structures in the PDB. 131

Fig. 42 – Typical cryo-EM sample preparation. 134

Fig. 43 – Transmission Electron Microscope Anatomy. 136

Fig. 44 – Contrast Transfer Function (CTF) Curves 138

Fig. 45 – The canonical single particle averaging (SPA) pipeline. 140

Fig. 46 – Purification of PqiABC. 148

Fig. 47 – Initial screening of PqiABC. 149

Fig. 48 – Initial screening of PqiABC on graphene oxide coated grids. 150

Fig. 49 – PqiABC ThermoFisher preliminary processing. 153

Fig. 50 – PqiABC final 2D classes. 155

Fig. 51 – Refinement and postprocessing of the ABC1 map. 158

Fig. 52 – Refinement and postprocessing of the ABC2 map. 159

Fig. 53 – Characterisation of PqiABC heterogeneity. 161

Fig. 54 – Refinement and postprocessing of the BFL1 map. 164

Fig. 55 – Refinement and postprocessing of the BFL2 map. 165

Fig. 56 – PqiB core model building. 167

Fig. 57 – Final PqiB BFL structure. 168
Fig. 58 – BFL1 MCE2 alternative pore lining loop density. 170
Fig. 59 – BFL1 MCE1 and MCE3 alternative PLL density. 171
Fig. 60 – Alignment of the BFL structure and PDB: 5UVN. 173
Fig. 61 – Fitting of the PqiC octamer into the ABC1 map. 175
Fig. 62 – Coiled coil flexibility in existing soluble PqiB structures. 177
Fig. 63 – Comparison of C6 PqiB cryo-EM maps. 181
Fig. 64 – Fitting of 5UVN to the EMD-8608 and BFL1 maps. 182
Fig. 65 – PqiB BFL1 MCE loop density. 184
Fig. 66 – PqiB 5UVN additional pore lining strands. 186
Fig. 67 – Phenotypic Complementation Assays. 191
Fig. 68 – *E. coli* K-12 BW25113 Δ*pqiABC* knockout generation. 193
Fig. 69 – Bending of PfMATE TM1. 196
Fig. 70 – Comparison of PfMATE and PqiA. 198
Fig. 71 – ConSurf conservation analysis of PqiA. 199
Fig. 72 – Conserved hydrophobic motifs in *E. coli* MCE protein PLLs. 201
Fig. 73 – PqiC complementation assay residues. 204
Fig. 74 – PqiA phenotypic complementation assays. 207
Fig. 75 – PqiB phenotypic complementation assays. 209
Fig. 76 – PqiC phenotypic complementation assays. 211
Fig. 77 – *pqiAB* phenotypic complementation assay 212
Fig. 78 – Comparison of the PfMATE N-lobe and PqiC C-lobe. 215
Fig. 79 – Location of the 15 cysteines within PqiA. 225
Fig. 80 – Anticipated architecture of the PqiABC complex. 230

List of Tables

TABLE. 1 – STRAINS UTILISED WITHIN THIS STUDY. 62

TABLE. 2 – KEY CONSTRUCT PROPERTIES. 63

TABLE. 3 – Q5 SDM REACTION CONTENTS. 69

TABLE. 4 – Q5 SDM CYCLING CONDITIONS. 70

TABLE. 5 – Q5 KINASE, LIGASE, DPN1 REACTION CONSTITUENTS. 70

TABLE. 6 – PQIABC Q5 SDM PRIMERS. 71

TABLE. 7 – PQI SEQUENCING PRIMERS. 72

TABLE. 8 – THEORETICAL PROTEIN/COMPLEX MOLECULAR MASS AND EXTINCTION COEFFICIENTS. 73

TABLE. 9 – PLASMIDS USED WITHIN THIS STUDY. 74

TABLE. 10 – MOLECULAR BIOLOGY PRIMERS USED WITHIN THIS STUDY. 75

TABLE. 11 – PHUSION® PCR REACTION CONTENTS. 76

TABLE. 12 – PHUSION® PCR CYCLING CONDITIONS. 76

TABLE. 13 – COLONY PCR CYCLING CONDITIONS. 77

TABLE. 14 – PQIC STRUCTURE DATA COLLECTION AND REFINEMENT STATISTICS. 119

TABLE. 15 – OPTIMISED GRID PREPARATION PARAMETERS FOR PQIABC. 151

TABLE. 16 – PQIABC THERMOFISHER DATASET COLLECTION PARAMETERS. 152

List of Abbreviations

ABC	-	ATP Binding Casette
ADP	-	Adenosine Diphosphate
ATP	-	Adenosine Triphosphate
Bam	-	β-barrel Assembly Machinery
CAMP	-	Cationic Antimicrobial Peptides
CDP-DAG	-	Cytidine-diphosphate Diacylglycerol
CL	-	Cardiolipin
Cls	-	Cardiolipin Synthase
Cryo-EM	-	Cryogenic Electron Microscopy
CTF	-	Contrast Transfer Function
DUF	-	Domain of Unknown Function
DDM	-	*n*-dodecyl-β-D-maltoside
ER	-	Endoplasmic Reticulum
FRET	-	Forster Resonance Energy Transfer
FRT	-	FLP Recombinase Target
GPL	-	Glycerophospholipid
IM	-	Inner Membrane
IMP	-	Inner Membrane Protein
IPTG	-	Isopropyl-β-D-1-thiogalactopyranoside
LLG	-	Log Likelihood Gain
Lol	-	Localisation of Lipoprotein
LOS	-	Lipooligosaccharide
LP	-	Lipoprotein
LPS	-	Lipopolysaccharide
Lpt	-	Lipopolysaccharide transport

LSB	-	Lauryl Sulfobetaine
MAM	-	Multivalent Adhesion Molecule
MCE	-	Mammalian Cell Entry
Mla	-	Maintenance of Lipid Asymmetry
MR	-	Molecular Replacement
Mw	-	Molecular Weight
MWCO	-	Molecular Weight Cut Off
OD_{600}	-	Optical Density at 600 nm
OM	-	Outer Membrane
OMP	-	Outer Membrane Protein
OMV	-	Outer Membrane Vesicle
ORF	-	Open Reading Frame
PDB	-	Protein Data Bank`
PE	-	Phosphatidylethanolamine
PG	-	Phosphatidylglycerol
PGP	-	Phosphatidylglycerol Phosphate
Pgs	-	Phosphatidylglycerol Phosphate Synthase
PLL	-	Pore Lining Loop
PMF	-	Proton Motive Force
Pqi	-	Paraquat Inducible
PS	-	Phosphatidylserine
Pss	-	Phosphatidylserine Synthase
Psd	-	Phosphatidylserine Decarboxylase
RCF	-	Relative Centrifugal Force
RMSD	-	Root Mean Square Deviation
SEC	-	Size Exclusion Chromatography
SDM	-	Site Directed Mutagenesis
SDS-EDTA	-	Sodium Dodecyl Sulphate and Ethylenediaminetetraacetic Acid

SDS-PAGE	-	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SNP	-	Single Nucleotide Polymorphism
SPA	-	Single Particle Analysis
STAS	-	Sulphate Transporter and AntiSigma Factor Agonist
TBST	-	Tris Buffered Saline with Tween 20
TCEP	-	(Tris-(2-carboxyethyl)phosphine)
TFZ	-	Translation Function Z Score
TM	-	Transmembrane
UV	-	Ultraviolet
WT	-	Wild Type

Chapter 1: General Introduction

Structure and organisation of the Gram-negative cell envelope

Architecture of the Gram-negative cell envelope

Initially believed to be simple, semipermeable, sac-like structures, subsequent advances in research have revealed biological membranes as complex dynamic systems continually adapting to their surroundings. They are an indispensable component of cellular life, serving as seats for vital biochemical processes, compartmentalising cellular space and defending cells from intrusive toxic substances, whose constituents vary depending upon the organism or organelle within which they reside.

Unicellular organisms lack the ability to modulate their local environment to the same degree as their multicellular counterparts, hence a more sophisticated boundary between themselves and their surroundings is often necessary. Consequently, unicellular organisms typically exhibit more structurally sophisticated membranes and cell wall structures capable of resisting sustained exposure to a wider range of otherwise damaging chemicals. This is particularly true of diderm or Gram-negative bacteria, which possess a cell envelope comprising two distinct membranes separated by an aqueous periplasm, which accommodates a thin peptidoglycan layer (Fig. 1).

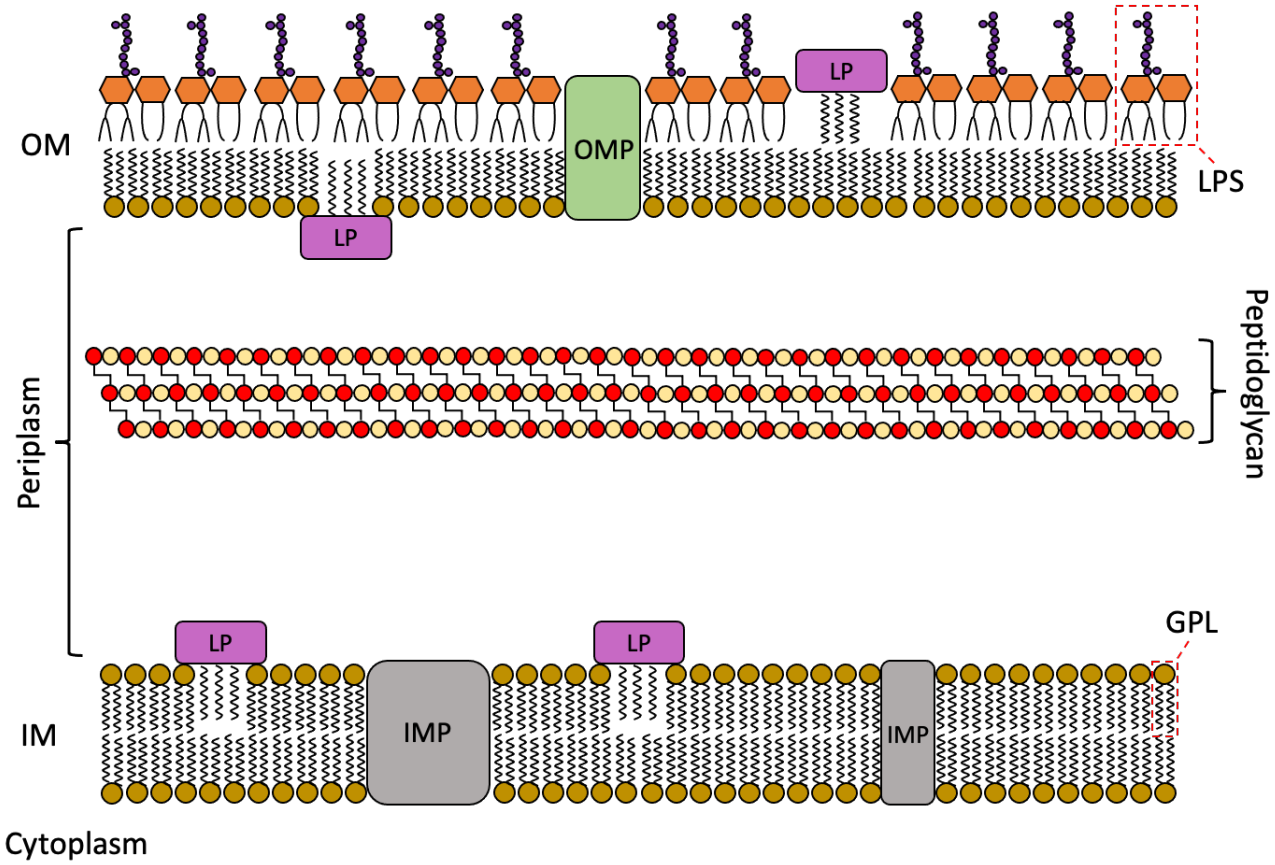


Fig. 1 – Architecture of the Gram-negative cell envelope.

The symmetric inner membrane (IM) bilayer consists exclusively of glycerophospholipids (GPL) and is adorned with both integral inner-membrane proteins (IMPs) and lipoproteins (LPs). The outer membrane (OM) bilayer is asymmetrical, comprising a glycerophospholipid inner leaflet and lipopolysaccharide (LPS) outer leaflet. The OM also encompasses lipoproteins and integral outer membrane proteins (OMPs). The two membranes are separated by the periplasm which harbours a thin peptidoglycan layer. The peptidoglycan mesh comprises alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) (indicated yellow and red circles) cross-linked by a short oligopeptide.

The Gram-negative inner membrane bilayer is of symmetrical nature, with both the inner and outer leaflets composed of glycerophospholipids (GPLs). This bilayer is adorned with both integral and membrane associated proteins. The transmembrane segments of integral inner membrane proteins (IMPs) are typically α-helical, however, these may be interspersed with disordered and/or globular domains, residing within the cytoplasm or periplasm. Lipoproteins may be anchored, via covalently attached acyl chains, to the outer leaflet of the inner membrane allowing the protein to reside at the periphery of the periplasm.

The periplasm is an aqueous environment, accounting for 20 - 40% of the total cell volume, separating the inner and outer membranes that houses a thin layer of peptidoglycan or murein, a polymer comprising linear oligo(*N*-acetylglucosamine - *N*-acetylmuramic acid) chains crosslinked at regular intervals by short peptides resulting in a mesh-like structure (Stock et al., 1977; Vollmer and Bertsche, 2008). This peptidoglycan exoskeleton prevents the risk of osmotic lysis arising from the large cytoplasmic turgor pressure as well as influencing the overall shape of the cell (Vollmer and Bertsche, 2008). Finally, the periplasm provides an oxidising environment allowing the stabilisation of proteins via disulphide bonds (Holmgren and Fagerstedt, 1982; Messens et al., 2007).

Whilst the inner membrane (IM) is composed entirely of GPLs, the canonical outer membrane (OM) is highly asymmetric incorporating the complex glycolipid lipopolysaccharide (LPS) exclusively within the outer leaflet and GPLs within the inner leaflet (Sohlenkamp and Geiger, 2015). Like the IM, the OM also incorporates integral membrane proteins, however, in contrast to their IM counterparts, outer-membrane proteins (OMPs) are almost exclusively of β-barrel architecture. In addition, lipoproteins may be tethered to the inner (more commonly) or outer leaflet of the OM enabling them to protrude into the periplasm or extracellular environment respectively.

The organisation of these three layers into a single structure provides a highly effective, selectively permeable, barrier against exogenous compounds whilst simultaneously permitting the essential flux of both nutrients and waste products. The presence of a fully functional cellular envelope is then essential for the vitality of diderms, however, the assembly and maintenance of such a complex structure is a significant undertaking.

Periplasmic Transport of Gram-negative Cell Envelope Components

Components of the Gram-negative cell envelope are synthesised within the confines of the IM where access to biosynthetic enzymes, precursors and cellular energy supplies are plentiful. Consequently, newly synthesised components must traverse the IM and periplasm, often against a concentration gradient, before OM incorporation all without disrupting the crucial cellular envelope architecture. The amphipathicity of OM components necessitate distinct mechanisms to shield their hydrophobic portions from the aqueous periplasm during trafficking to the OM. These processes must be tightly coordinated to maintain stoichiometric ratios, furthermore, once beyond the IM, cellular energy sources such as ATP and the proton motive force (PMF) become unavailable. Whilst the transport pathways responsible for delivering OMPs (Hagan et al., 2011), LPS (Okuda et al., 2016), and lipoproteins (Konovalova and Silhavy, 2015) to the OM have been extensively delineated the mechanisms facilitating bulk GPL transport remain enigmatic.

Outer Membrane Proteins

Gram-negative bacteria contain two classes of integral membrane proteins. α-helical proteins are located primarily within the IM whilst their β-barrel counterparts reside solely within the OM and consequently are often referred to as Outer Membrane Proteins (OMPs).

OMPs exhibit a canonical topology in which antiparallel β-sheets assemble a cylinder such that hydrophobic residues face the exterior, permitting membrane interactions, and hydrophilic residues line the core providing gateways for exchange of polar solutes. OMPs demonstrate considerable variation in several aspects including size (8-36 β-strands), multimerisation and additional variable domains or loops (Horne et al., 2020).

As with all *E. coli* proteins, OMPs are synthesised in the cytoplasm (Fig. 2). Transport across the IM is achieved via the SecYEG complex which interacts with an inherent amino leader sequence (Kudva et al., 2013). Once in the periplasm, nascent OMPs bind periplasmic chaperones of either the SurA or Skp pathways enabling passage to the OM but also critically preventing premature (mis)folding (Fig. 2). Recent evidence suggests Skp and SurA to act in a sequential manner with Skp initially capturing nascent OMPs as they cross the IM and passing them to SurA which interacts with the Bam complex (Wang et al., 2021).

Despite the considerable structural variations exhibited, folding and insertion of almost all known OMPs is performed by the β-barrel assembly machinery (Bam) complex (Fig. 2). The *E. coli* Bam complex comprises five proteins (Bam A-E) (Wu *et al.* 2005). Of these only BamA, itself an integral β-barrel protein, and BamD, a lipoprotein, are essential (Wu et al., 2005). Accessory lipoproteins BamB, C and E appear neither essential nor highly conserved hence their proposed roles in the modulation of folding, substrate specificity and/or coordination with other cellular processes appear largely dispensable (Charlson et al., 2006; Konovalova et al., n.d.; Sklar et al., 2007; Wu et al., 2005). The exact mechanism by which Bam folds and inserts OMPs to the OM, in the absence of ATP, is unclear and several plausible models have been proposed (Konovalova et al., 2017). A recent cryo-EM study capturing several intermediate states of Bam mediated β-barrel folding has provided the greatest insight into this fundamental process so far (Doyle et al., 2022). The study suggests that, following recognition of their C-terminal β signals, OMPs are inserted into the OM as β sheets which are subsequently “barrelised”. Furthermore, the authors indicate that Bam utilises the inherent tensile forces of the OM to drive the close of OMPs (Doyle et al., 2022).

Lipopolysaccharide

Assembly of LPS occurs in two sequential steps at the IM. Initially, core sugars are attached to a lipid A anchor before the resulting “rough LPS’ molecule is flipped to the periplasmic face of the IM by a dedicated floppase, MsbA, allowing its conversion to “smooth LPS” via the addition of the O-antigen (Doerrler et al., 2004; Ruiz et al., 2004; Zhou et al., 1998) (Fig. 2).

LPS transport to the OM of *E. coli* is facilitated by the Lpt machinery, consisting of 7 essential proteins, depletion of which results in LPS accumulation on the periplasmic face of the IM (Fig 2). LPS is extracted from the IM by the $LptB_2FGC$ complex, an ABC transporter, utilising energy from the LptB catalysed hydrolysis of ATP. At the OM a complex comprising LptD, a β-barrel OMP, and LptE, a lipoprotein, insert LPS directly into the outer leaflet of the OM (Gu et al., 2015). Transfer between these two membrane complexes occurs via a periplasmic bridge structure formed from LptA monomers, each possessing a hydrophobic groove to sequester acyl chains, which oligomerise in a head-to-tail fashion (Suits et al., 2008). Bridge transmission and OM insertion are both active processes, driven by distal ATP hydrolysis by LptB within the $LptB_2FGC$ complex (Okuda et al., 2012). All LPS binding sites along the periplasmic bridge remain occupied at any one time hence extraction from the IM and subsequent transfer to the first bridge position causes a chain of movement resulting in displacement and OM insertion of the molecule residing in the final bridge position. This model has been coined the ‘PEZ’ model due to its resemblance of the mechanics within the popular candy dispenser (Okuda et al., 2016).

Lipoproteins

Lipoproteins are synthesised in the cytoplasm as pre-lipoproteins, possessing an intrinsic N-terminal signal peptide which incorporates the 'Lipobox' consensus sequence of Leu-Ala/Ser-Gly/Ala-Cys, before traversing the IM via the Sec translocon (Okuda and Tokuda, 2011) (Fig. 2). Maturation of bacterial lipoproteins involves processing of the signal peptide, in three distinct phases, occurring at the periplasmic face of the IM. Covalent attachment of diacylglycerol to the sulfhydryl group of the C-terminal lipobox cysteine enables cleavage of the signal sequence up to the aforementioned residue (Inouye et al., 1983; Sankaran and Wu, 1994). Finally, N-acylation of the cysteine amino group yields the mature tri-acylated lipoprotein (Gupta and Wu, 1991).

The identity of the residue succeeding the modified cysteine is critical in determining the ultimate destination of the mature lipoprotein. Whilst most lipoproteins are transported to the OM via the localisation of lipoprotein (Lol) pathway, those possessing an aspartic acid at the +2 position are retained within the IM (Yamaguchi et al., 1988).

The *E. coli* Lol pathway consists of five proteins (Lol A-E). Initially, the mature lipoprotein is extracted from the IM by the LolCDE complex and transferred to the soluble periplasmic chaperone LolA in an ATP dependent manner (Yakushi et al., 2000). LolA interacts with LolB, itself an OM lipoprotein, whose higher affinity for acyl chains strips the lipoprotein from the LolA binding pocket (Taniguchi et al., 2005), however, the ultimate transfer of the lipoprotein from LolB to the OM is poorly characterised (Fig. 2).

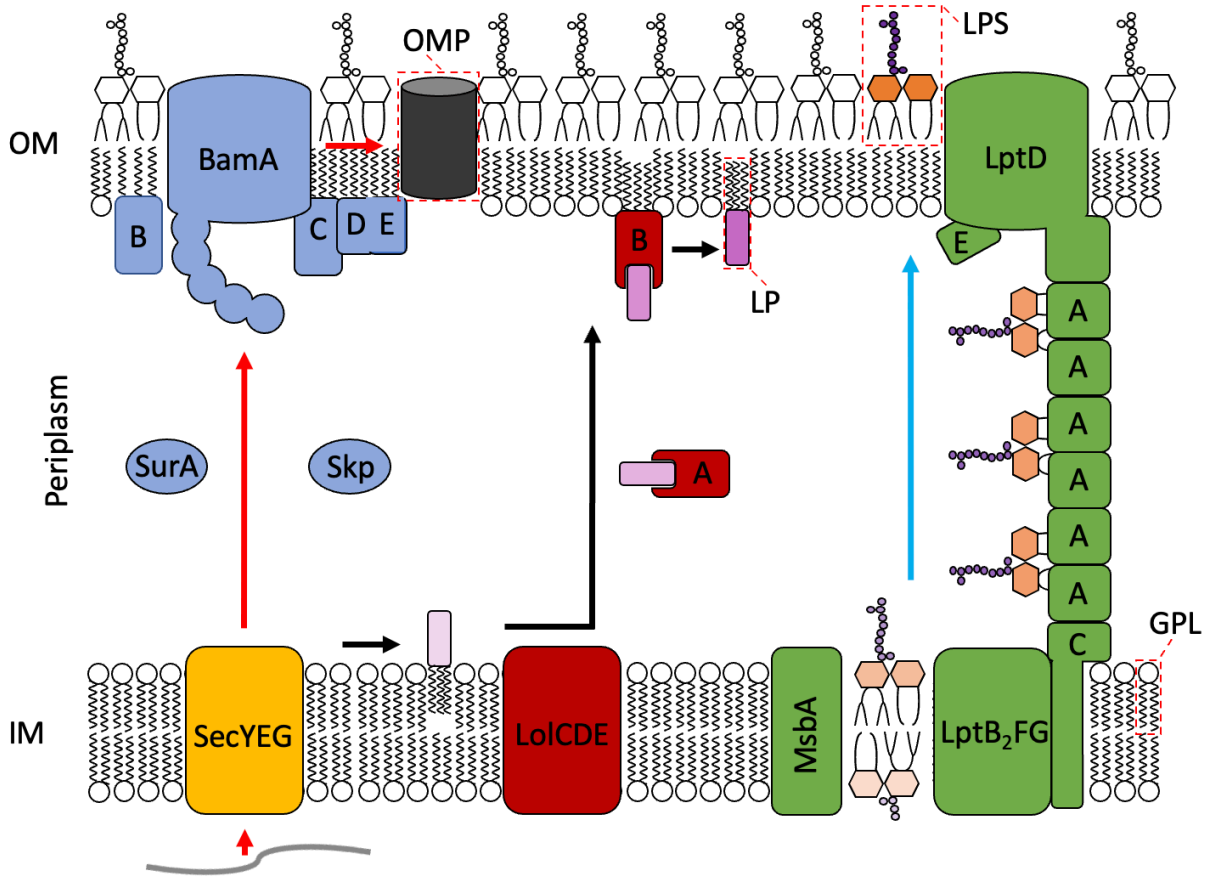


Fig. 2 – Outer Membrane Biogenesis Transport Pathways

Components of the outer membrane are synthesised upon or within the confines of the inner membrane and then transported across the periplasm to the outer membrane. Nascent peptides are transported across the IM by the SecYEG complex. Outer membrane proteins are received by the periplasmic chaperones Skp and SurA and delivered to the β-barrel assembly machinery (BAM) which catalyses their insertion into the OM. Lipoproteins are processed at the cytoplasmic face of the IM following translocation through the Sec machinery. Mature lipoproteins are extracted from the IM by the LolCDE complex and passed to the periplasmic chaperone LolA which in turn delivers them to the OM where LolB facilitates their insertion. Rough LPS is flipped across the IM by MsbA and converted to smooth LPS. Mature LPS is then extracted from the IM by the $LptB_2FGC$ complex and transported directly to the OM via a bridge of LptA monomers where LptD and E facilitate its insertion into the OM.

Glycerophospholipids

Glycerophospholipid Synthesis

Glycerophospholipids (GPLs) are amphipathic molecules comprising a glycerol backbone adjoined to variable hydrophilic head groups and hydrophobic acyl chains. Bacterial membrane GPL composition is highly diverse, varying not only between species but also throughout the bacterial life cycle. For example, *E. coli* exhibits a relatively simple composition comprising approximately 75% phosphatidylethanolamine (PE), 20% phosphatidylglycerol (PG) and 5% cardiolipin (CL) (Raetz and Dowhan, 1990) whereas *Pseudomonas pertucinogena* contains approximately 46% PE, 35% CL, 11% PG and 8% lyso cardiolipin, reviewed in (Sohlenkamp and Geiger, 2015). Beyond lipid head group variability, additional modifications including acyl chain modifications, primarily cis-trans isomerisation, desaturation and cyclopropanation may occur in response to environmental conditions (Cronan, 2002).

Biosynthesis of the primary GPLs in *E. coli* (PE, PG and CL) occurs at the cytoplasmic face of the IM and has been extensively characterised (Sohlenkamp and Geiger, 2015). Both phosphatidylethanolamine and phosphatidylglycerol are synthesised from the core metabolite cytidine-diphosphate diacylglycerol (CDP-DAG) in two steps (Fig. 3). CDP-DAG is first condensed with serine by phosphatidylserine synthase (Pss) forming phosphatidylserine (PS) which is subsequently decarboxylated by phosphatidylserine decarboxylase (Psd) to yield PE. Alternatively, phosphatidylglycerol phosphate synthase (PgsA) may condense CDP-DAG and glycerol-3-phosphate to produce phosphatidylglycerol phosphate (PGP) which is subsequently dephosphorylated by one of three phosphatases (PgpA-C) to yield PG. In *E. coli*, cardiolipin is synthesised in one step by one of three cardiolipin synthase enzymes (ClsA-C). Both ClsA and ClsB catalyse the condensation of two PG molecules whilst ClsC synthesises cardiolipin from one molecule of both PG and PE (Tan et al., 2012).

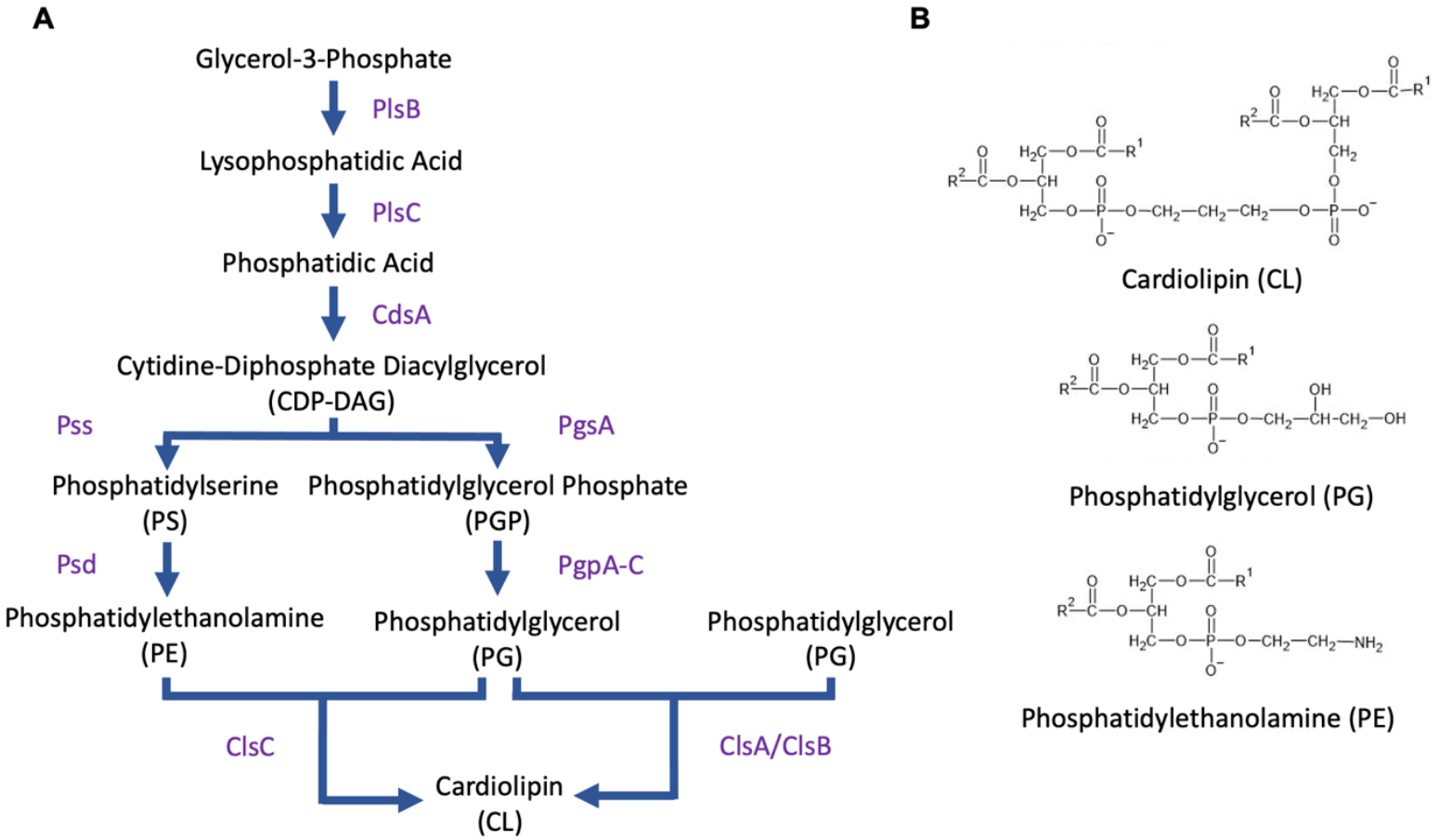


Fig. 3 – *E. coli* GPL biosynthesis.

(A) *E. coli* biosynthetic pathways for PG, PE and CL. Intermediates and enzymes are depicted in black and purple respectively. (B) Structural formulas for CL, PG and PE.

Much of the pioneering work into bacterial phospholipid metabolism was performed by Eugene Kennedy during the 1960s and continued into the late 1990s. The comprehensiveness of his work led to the utilisation of *E. coli* as a model system for studying phospholipid related processes (see (Dowhan, 2013)).

Localising phospholipid synthesis to the cytoplasmic face of the IM ensures biosynthetic enzymes have access to the necessary energy pools and cofactors whilst providing a stable environment for the acyl chains of synthesised molecules to reside. However, newly synthesised GPLs must be rapidly equilibrated across the two IM leaflets to ensure uniform bilayer propagation for sustained growth. This intramembrane translocation requires polar lipid head groups to cross the hydrophobic core of the membrane. In model lipid membranes spontaneous GPL flipping is extremely slow and energetically unfavourable, however, *in vivo* flipping occurs at rates in excess of 30,000 times faster than those observed *in vitro* (Rothman and Kennedy, 1977). Transmembrane XALP*N* peptides (poly Ala, Leu repeats flanked by two “X” residues at either termini with *N* specifying the peptide length), and the N-terminal TM helix of signal peptidase, have been observed to induce GPL translocation when inserted into model *E. coli* phospholipid membranes suggesting the mere presence of TM proteins alone may be sufficient for flipping (Kol et al., 2004). Nonetheless, eukaryotes possess three different classes of proteins to facilitate intramembrane GPL translocation, implying such mechanisms may be present in prokaryotes. Flippases and floppases perform ATP-depended translocation of specific substrates into and out-from the cytoplasmic leaflet respectively whilst scramblases catalyses non-specific, passive mixing between both membrane leaflets (Pomorski and Menon, 2016). To date, no generalised bacterial scramblase or GPL translocator has been identified, however, dedicated translocators of other substrates have been characterised. MsbA is a floppase that utilises ATP to flip newly synthesised rough LPS from the cytoplasmic to the periplasmic face of the IM enabling its conversion to smooth LPS (Ruiz et al., 2009). MsbA has also been implicated in the

transbilayer movement of GPLs as inactivation of a temperature sensitive *E. coli* MsbA mutant caused rapid accumulation of both LPS and GPLs within the inner leaflet of the IM (Doerrler et al., 2004). Conversely, LPS is nonessential in *N. meningitidis* hence the viability of a ΔMsbA mutant indicates its redundancy in GPL translocation thus the aforementioned accumulation may just be a consequence of abnormal LPS flipping (Tefsen et al., 2005). Indeed, one study failed to observe MsbA catalysed flipping of C_6NBD-PG in both the presence and absence of ATP (Kol et al., 2004). Whether, therefore, there is a dedicated GPL translocator in bacteria remains to be seen.

Transport to the Outer Membrane

Intermembrane lipid transport consists of three phases; 1) extraction from the donor membrane, 2) traversion of the aqueous periplasm and 3) insertion into the recipient membrane. The inherent amphipathicity of GPLs makes their transport across the periplasm a considerable undertaking, where exposure of their hydrophobic acyl chains to the aqueous environment is highly energetically unfavourable (Tanford, 1978). Consequently, it seems likely that transport across the periplasm occurs in an assisted manner (Shrivastava and Chng, 2019).

Several models for intermembrane GPL transport have been proposed by extrapolating our current understanding of previously established periplasmic associated transport pathways, namely the Localisation of lipoprotein (Lol) pathway, the Lipopolysaccharide transport (Lpt) pathway and the OMP transport pathway.

The models are divided into two categories, the first of these relies upon GPL binding proteins which sequester GPL acyl chains from the periplasm whilst the second model

pertains to the direct exchange of lipids between membranes without ever having to leave their native environment (Giacometti et al., 2022; Shrivastava and Chng, 2019) (Fig. 4).

GPL binding proteins

Such proteins may be soluble in nature, shuttling between complexes residing in both membranes, operating in a manner analogous to the Lol pathway (Okuda and Tokuda, 2011). Alternatively, these proteins may assemble a physical structure, spanning the periplasm, providing a continuous hydrophobic environment through which GPLs, entire or partial, may pass, emulating the Lpt pathway (Okuda et al., 2016) (Fig. 4). Of these two, an Lpt-like pathway seems more feasible. To shuttle sufficient GPLs to the OM, a Lol like pathway would require large quantities of the periplasmic shuttle and/or a binding stoichiometry sufficient to transmit multiple GPLs simultaneously (Powers and Stephen Trent, 2019).

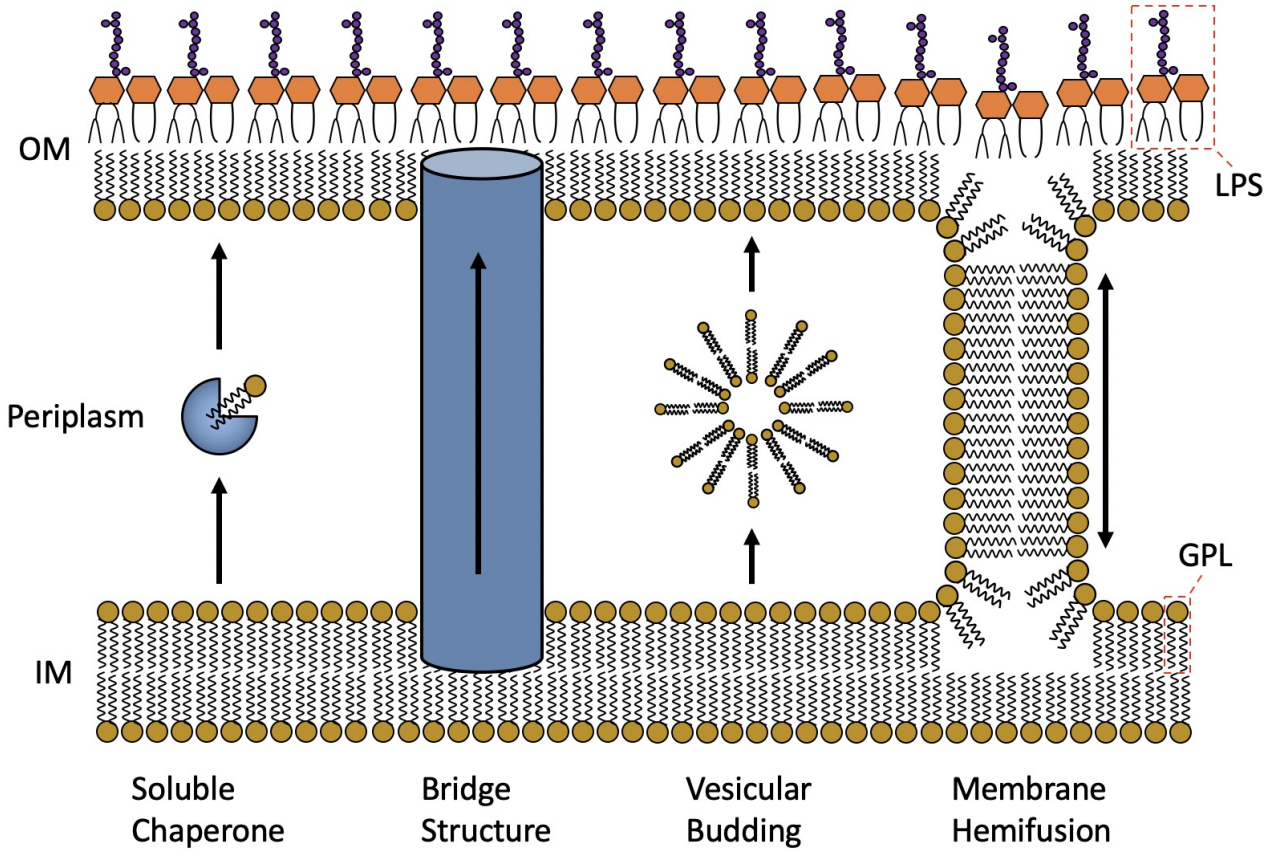


Fig. 4 – Proposed Gram-negative intermembrane GPL transport models.

Left: soluble transport proteins shuttle GPLs across the periplasm. Middle-Left: proteinaceous structures span the periplasm and provide a tunnel through which GPLs may transit. Middle-Right: Direct transport via vesicular budding. Right: Direct transport via the formation of hemifusion structures. Proteins (both soluble and periplasm spanning) are depicted in blue whilst the GPL transport direction is indicated by black arrows. Figure reproduced from (Shrivastava and Chng, 2019).

In both cases, it seems likely that an external energy input would be required. However, the exact utilisation of this input depends on the energetics of the transport system (Fig. 5). Energy may be required to extract lipids from the donor membrane whilst transfer to the recipient membrane occurs spontaneously (Fig. 5A). Alternatively, the inverse scenario may exist, whereby the loading of lipids occurs spontaneously, however, external energy input is required for their subsequent release (Fig. 5B). Finally, energy may be required for both the loading and unloading of lipids from the proteinaceous transporter (Fig. 5C). In all scenarios, a bridge structure appears to provide an obvious advantage as it may physically couple the utilisation of energy sources, freely available at the IM, with distal processes within both the periplasm and OM.

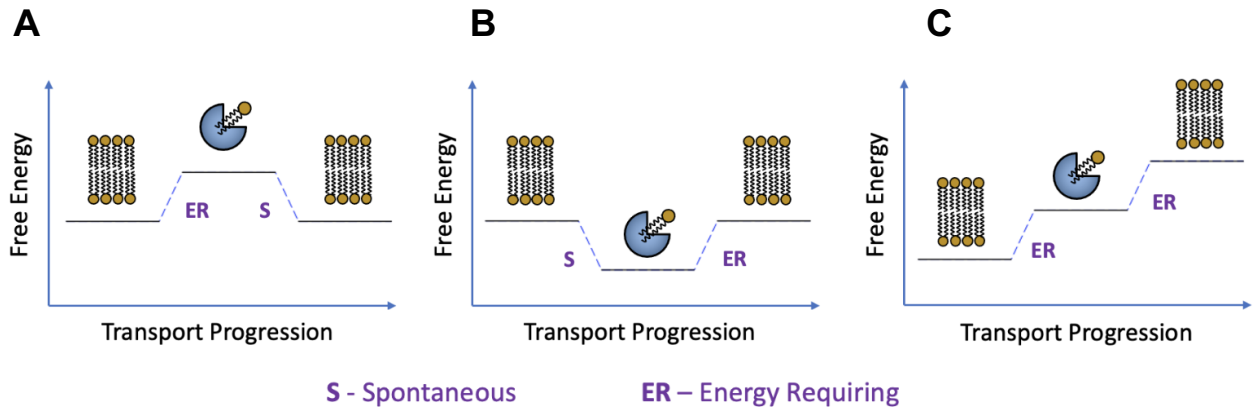


Fig. 5 – Energy profiles for intermembrane GPL transport.

Utilisation of a GPL binding protein to facilitate intermembrane transport likely requires external energy input during (A) GPL loading, (B) GPL offloading or (C) both GPL loading and unloading. Horizontal, solid, black lines represent the free energy state of each phase of the transport process. Spontaneous and energy requiring transitions between these states are indicated by descending and ascending blue, hashed lines respectively. Figure adapted from (Shrivastava and Chng, 2019).

Direct exchange

Alternatively, GPL transfer may occur via vesicular budding or the formation of hemifusion structures enabling direct transfer of GPLs between the two membranes (Fig. 4). It's possible these models would also require an external energy input in addition to proteins ensuring close membrane proximity and/or facilitating membrane curvature. Currently, no evidence for the existence of such transport pathways has been found in bacteria. Furthermore, the lattice-like peptidoglycan layer sandwiched between the two membranes, with its relatively small "pore" size, and the distance between the membranes (15-20 nm) themselves present obvious hurdles for such transport mechanisms to overcome. Zones of adhesion between the IM and OM, so called Bayer's patches, potentially enabling hemifusion events have been postulated though remain highly controversial (Bayer, 1991, 1968; Kellenberger, 1990). However, it must also be noted that machinery for similar fusion events exists in nature, exemplified by the transient fusion of the OMs of adjacent myxococcus cells (Ducret et al., 2013; Pathak et al., 2012).

The following sections explore the systems implicated in intermembrane GPL transport, summarised in Fig. 6.

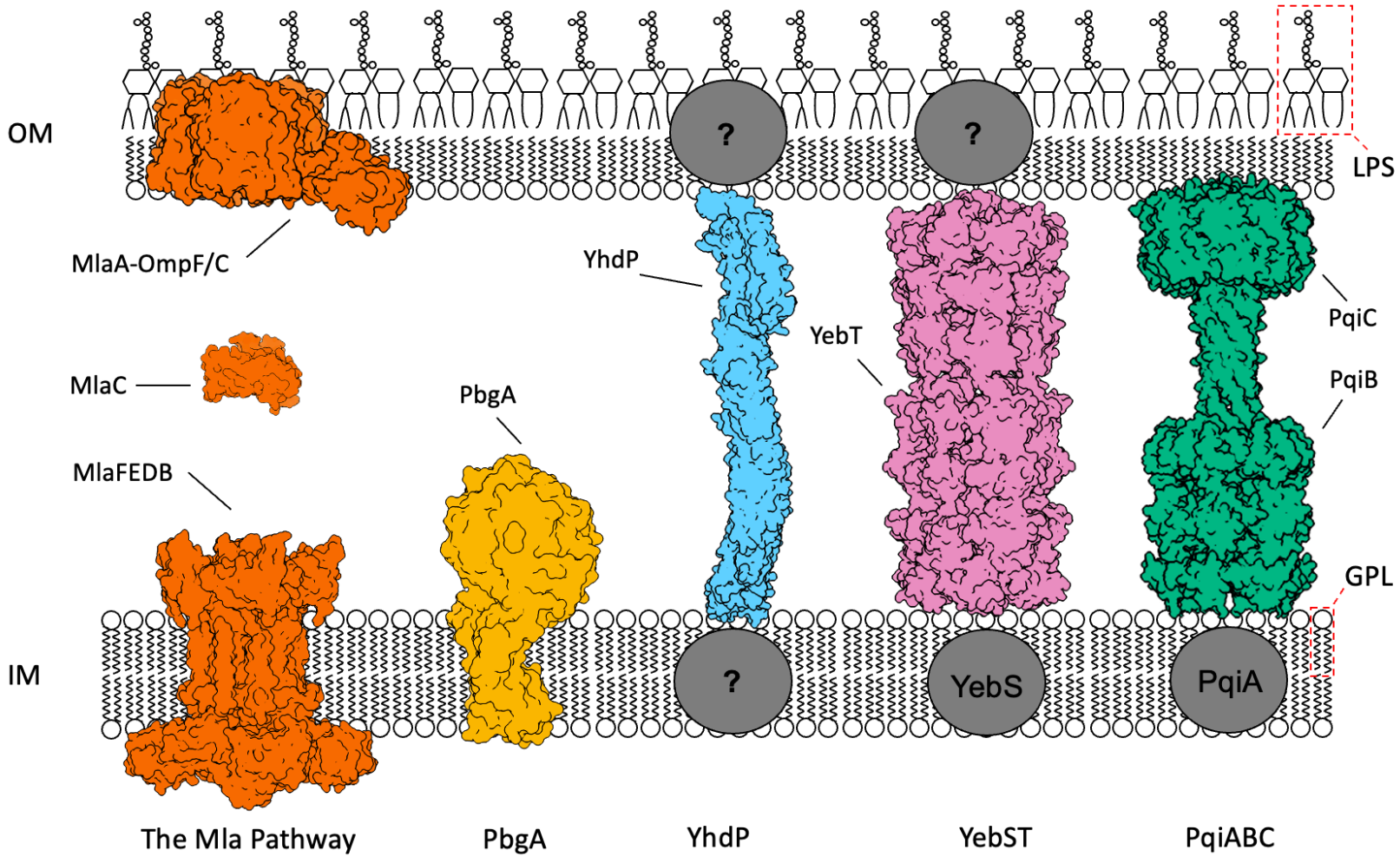


Fig. 6 – Putative intermembrane GPL transport pathways.

Volume representations of the systems implicated in intermembrane GPL transport: The Mla Pathway – Orange (PDB: 7CGE, 5UWA, 5NUQ), PbgA – Yellow (PDB: 6V8Q), YhdP – Blue (Alphafold P46474), YebST – Purple (PDB: 6V0C) and PqiABC (BFL and CFL structures)

Maintaining Lipid Asymmetry

In response to OM damage, LPS may be lost triggering spontaneous GPL flipping to plug the resulting gaps in the membrane (Malinverni and Silhavy, 2009). Additionally, the maintenance of an asymmetric membrane is entropically unfavourable hence spontaneous flipping of OM GPLs, from the inner to the outer leaflet, will occur as cells age. In both aforementioned scenarios, GPLs become surface exposed and substrate for one of two OM β-barrel GPL degrading enzymes, PldA and PagP, which act to restore OM lipid asymmetry (Fig. 7). Structural analysis has revealed the active sites of both proteins are only accessible to GPLs residing within the outer leaflet of the OM (Hwang et al., 2002; Snijder et al., 1999). The phospholipase PldA normally exists as an inactive monomer, however, the presence of outer leaflet GPL initiates its assembly into active dimers which deacylate GPLs and lyso-GPLs at both the sn-1 and sn-2 positions effectively degrading mislocalised molecules (Dekker, 2000). PagP, a palmitoyl transferase induced by the PhoPQ stress response, removes palmitate from the sn-1 position of a GPL donor and attaches it to lipid A producing hepta-acylated LPS and a lyso-GPL by-product which is presumably subsequently degraded by PldA (Bishop, 2005).

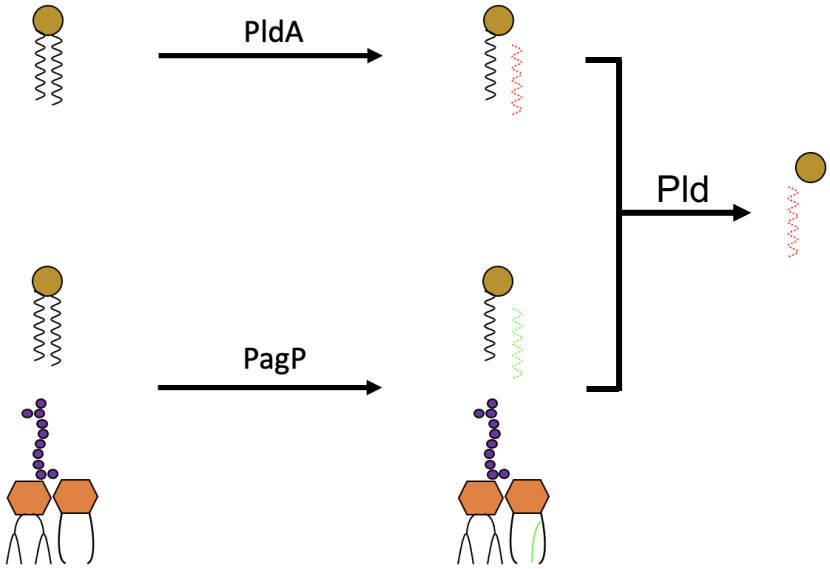


Fig. 7 – Removal of mis-localised OM GPLs via PldA and PagP.

The phospholipase PldA deacylates GPLs and lyso-GPLs at both the sn-1 and sn-2 positions effectively degrading mislocalised molecules. PagP, a palmitoyl transferase, removes palmitate from the sn-1 position of a GPL donor and attaches it to lipid A producing hepta-acylated LPS and a lyso-GPL by-product which is presumably subsequently degraded by PldA.

The Mla pathway was originally proposed by Malinverni and Silhavy as a third system for the removal of mislocalised outer leaflet GPLs (Malinverni and Silhavy, 2009). Loss of function mutations in all Mla components produce a distinct SDS-EDTA sensitive phenotype, a common indicator of OM perturbation. Spontaneous suppressor mutations were mapped to the PldA locus and found to increase expression of the phospholipase suggesting SDS-EDTA sensitivity resulted from outer leaflet GPL accumulation. Furthermore, increased PagP mediated LPS acylation was observed in Δ*mla* and Δ*mla*Δ*pldA* mutants presumably due to an increased availability of outer leaflet GPL substrates. Together, these findings implicated

the Mla pathway in retrograde GPL transport making it the first proteinaceous GPL transport pathway identified.

The Maintenance of Lipid Asymmetry (Mla) pathway

The Mla pathway is unlike PldA and PagP in that it does not rely on enzymatic OMPs to degrade mis-localised GPLs, but instead utilises a transport system of six proteins; an IM complex composed of MlaF, E, D, and B proteins, a periplasmic chaperone, MlaC, and an OM lipoprotein, MlaA (Fig. 8). In *E. coli* the system is expressed from two loci with the *mlaFEDCB* genes all being located within a single operon distinct from that of *mlaA*. In other species, however, all six genes have been observed to be colocalised (Casali and Riley, 2007).

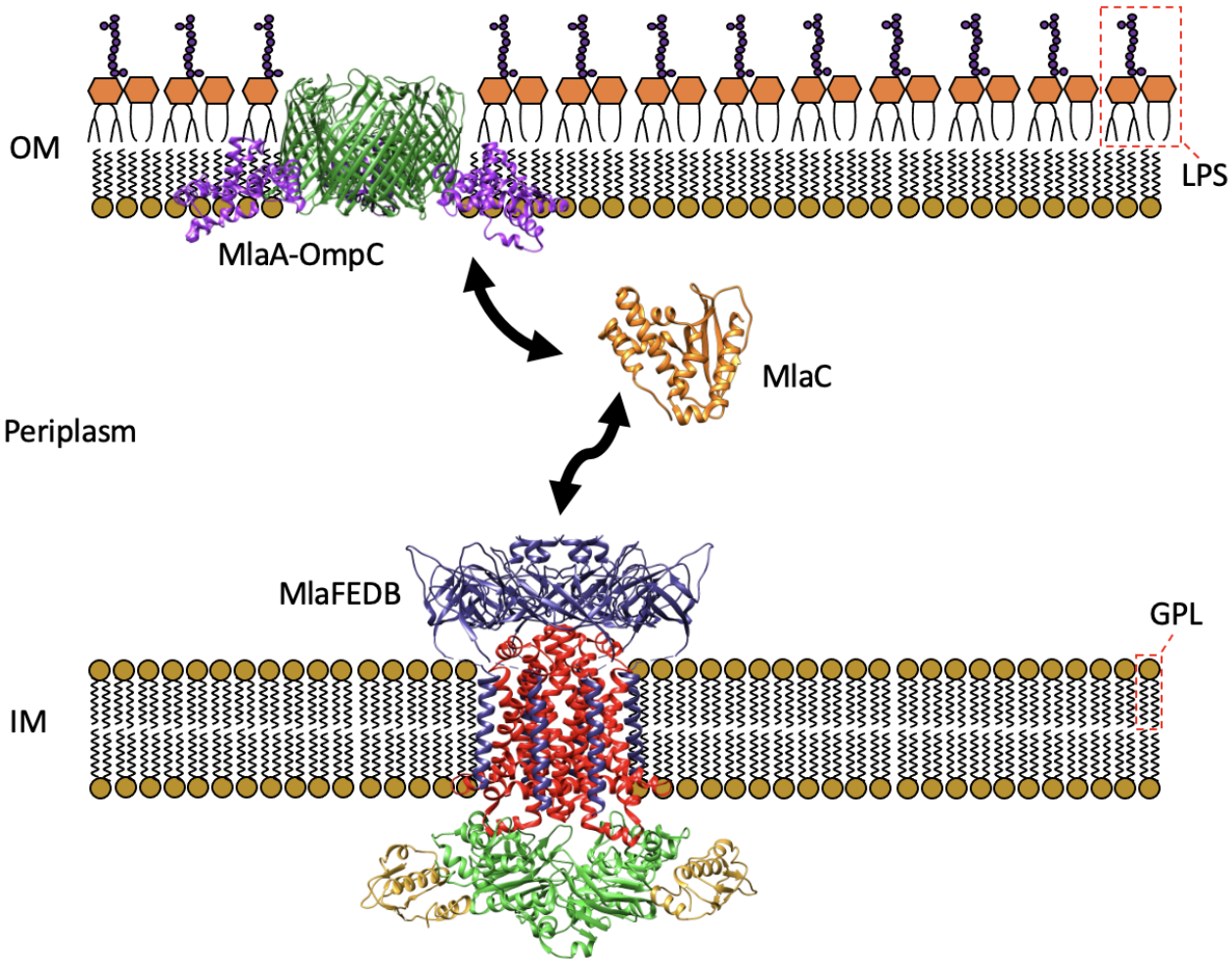


Fig. 8 – The Maintenance of Lipid Asymmetry (Mla) Pathway.

Cartoon representation of the Mla pathway components in their respective cellular locations. Black arrows represent possible GPL exchanges. Colouring: MlaA - Purple, OmpC - Dark Green, MlaC - Orange, MlaD - Dark Blue, MlaE - Red, MlaF - Light Green, MlaB - Gold. MlaFEDB (PDB: 7CGE), MlaC (PDB: 5UWA), MlaA-OmpF (PDB: 5NUP). Please note that the peptidoglycan layer is removed from the periplasm to aid figure clarity.

MlaFEDB

The MlaFEDB complex, existing in a 2:2:6:2 ratio of MlaF, MlaE, MlaD and MlaB (Giacometti et al., 2022; Thong et al., 2016), belongs to the ATP-binding cassette (ABC) superfamily of transporters. ATP hydrolysis is performed by cytosolic facing MlaF which in combination with membrane spanning permease MlaE, forms a classical ABC transporter complex (Thong et al., 2016). MlaF mediated ATP hydrolysis is believed to be intricately linked to the processes of GPL transport performed by MlaE and MlaD. Recently published structural studies indicate ATP induced conformational changes to MlaF are propagated throughout the entire complex initiating rearrangements of the transmembrane helices of both MlaE and MlaD (Chi et al., 2020; Coudray et al., 2020; Mann et al., 2021). This process of ATP hydrolysis is tightly regulated, with MlaB, a cytosolically located STAS domain containing protein believed to regulate both complex assembly and ATPase activity. Indeed, a single amino acid substitution in MlaB, T52A, was shown to abolish all ATPase activity within a purified MlaFEB complex (Thong et al., 2016). Furthermore, preliminary structural investigations have highlighted the ability of MlaB to stabilise the MlaFEDB complex. MlaB interacts at the alpha-helical domains of MlaF, which in turn forms extensive contacts with the transmembrane helices of MlaE (Coudray et al., 2020; Mann et al., 2021; Tang et al., 2020). Two independent publications reported an inability to purify a (sub)complex containing both MlaE and MalF in the absence of MlaB (Kolich et al., 2020; Thong et al., 2016). MlaB, therefore, can be seen to form a stabilising bridge between the nucleotide binding domains of MlaF and the transmembrane domains of MlaE. Recently, a crystal structure of the MlaFB subcomplex was utilised to propose a novel post-translational mechanism for regulating complex assembly. MlaF dimers are stabilised by the binding of MlaB allowing formation of the entire MlaFEDB complex. Reduction in MlaB abundance or its inactivation, potentially via phosphorylation, would result in destabilisation of MlaF dimers and consequently reduce MlaFEDB complex formation (Kolich et al., 2020). Finally, the mammalian cell entry (MCE) domain containing component, MlaD, forms a hexameric structure which sits atop the

complex in the periplasm (Ekiert et al., 2017), forming extensive contacts with the rest of the complex through membrane anchoring helices (Chi et al., 2020). As an MCE domain containing protein, MlaD has been shown to bind all major *E. coli* GPL species (Ekiert et al., 2017). Recent insights highlight the presence of a large hydrophobic cavity running through the centre of the MlaD hexamer (Ekiert et al., 2017), in addition to putative GPL binding sites found at the interface of two MlaD monomers (Ercan et al., 2019). Mutagenesis of several cavity lining residues have proved crucial for correct MlaD function, hence the cavity itself is likely involved in GPL transport (Ekiert et al., 2017). Indeed, recent structural findings suggest at least three phospholipid binding sites are present within the MlaFEDB complex, between the interfaces of MlaD and MlaE, with density observed within three cavities designated central, lower and outer (Fig. 9). Within the central cavity, formed between TM1, TM2 and TM5 of two MlaE subunits and the hexameric ring of MlaD, two phospholipid densities have been observed, with one acyl chain from each monomer extending upwards towards MlaD and the other downwards towards MlaE, suggesting an intermediate transport state (Fig. 9B) (Chi et al., 2020; Coudray et al., 2020). Three phospholipids were resolved in the lower cavity, formed between the TMs of MlaE and MlaD with the authors proposing this as the site of GPL extraction/insertion, implying GPL migration may occur at the cytoplasmic membrane leaflet (Fig. 9C) (Chi et al., 2020). Preliminary liposomal transport assays reinforce this notion (Hughes et al., 2019). Two additional GPLs were also observed in the outer cavity, on the periphery of the MlaE TMs in a V shaped groove, with the lipids oriented such that their acyl chains were directed down towards MlaF suggesting they were derived from the periplasmic leaflet of the inner membrane (Fig. 9D) (Chi et al., 2020). Consequently, the authors proposed this may represent a gateway for GPL entry/exit. A subsequent cryo-EM map of the *A. baumannii* MlaFEDB complex contained density suggestive of seven additional GPLs bound within the MlaD hexamer, however, these remain un modelled in the deposited structure (Mann et al., 2021).

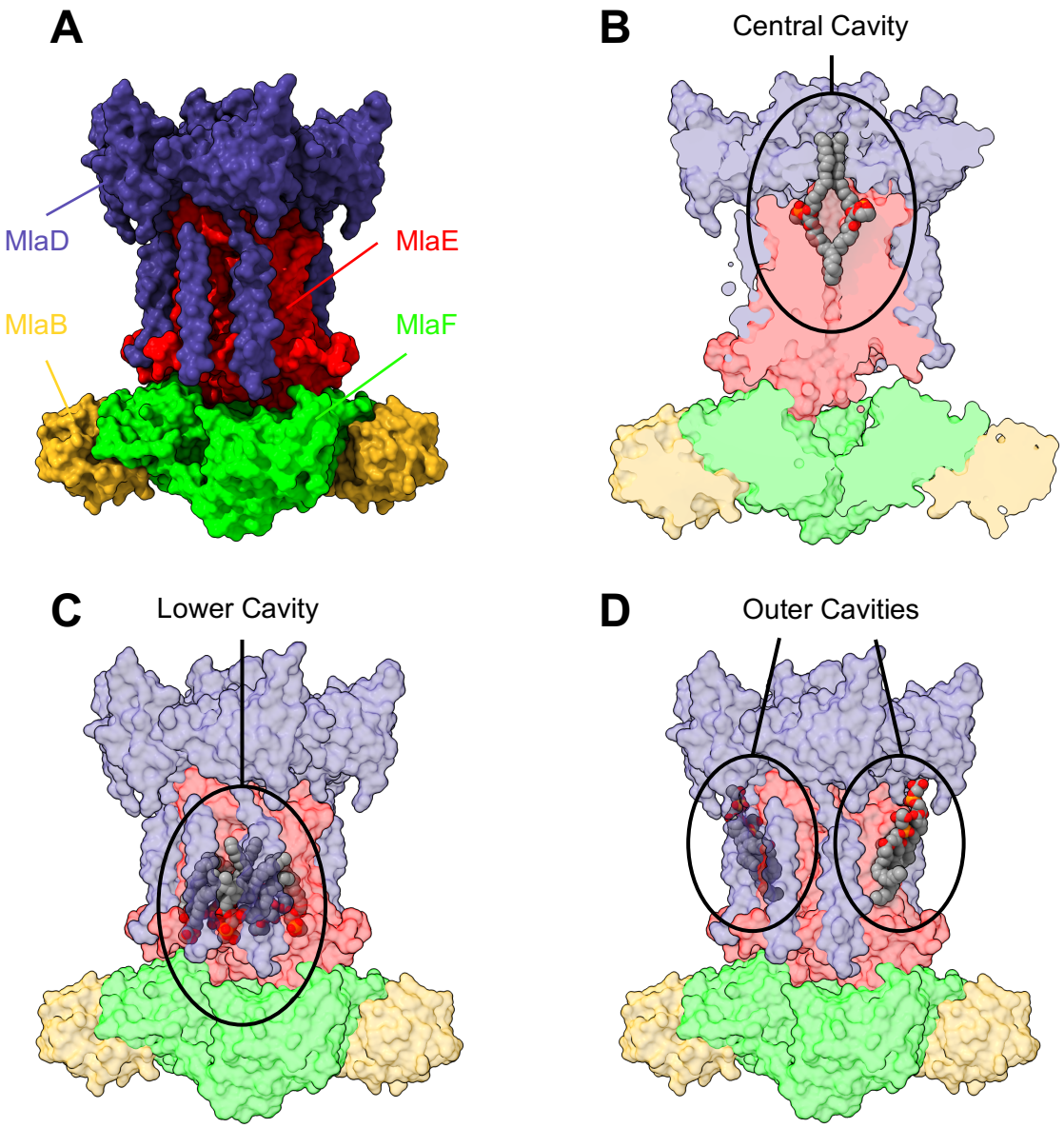


Fig. 9 – GPL binding sites within the *E. coli* MlaFEDB complex.

(A) Molecular surface representation of the *E. coli* MlaFEDB complex solved via cryo-EM. Colouring: MlaA - Purple, MlaD - Purple, MlaE - Red, MlaF - Light Green, MlaB - Gold. (B-D) Transparent Molecular surface representation of the *E. coli* MlaFEDB complex indicating the position of modelled PG molecules within the (B) central, (C) lower and (D) outer cavities. MlaFEDB (PDB:7CGE).

To probe the structural changes occurring during ATP hydrolysis, researchers have also solved the MlaFEDB structure in the ATP bound conformation, using an ATPase attenuated mutation ($MlaF^{E170Q}$) (Chi *et al.* 2020). Two conformations were solved. In the first, a large conformational change in MlaD was noted (twisted by 15^{o} and lifted), resulting in an extension to the central cavity possibly acting to facilitate GPL transit, whilst the arrangement of MlaF, B and E remained unchanged from the apo conformation. In the second structure, tight dimerisation of MlaF was observed, with concomitant changes throughout all TM domains resulting in collapse of the central GPL binding cavity and further movement of MlaD. This configuration may represent an extrusion conformation, however, it remains to be seen whether extrusion proceeds into the membrane or to MlaC within the periplasm. Other studies have solved additional structures of MlaFEDB from both *E. coli* and *A. baumannii*, with representatives of apo, ADP and ATP bound forms (Coudray et al., 2020; Mann et al., 2021; Tang et al., 2020), however, these studies report no major conformational changes between states. Henceforth, functional assays and additional structures, depicting multiple states, remain crucial to entirely resolve the mechanics of GPL translocation.

MlaA

MlaA, perhaps the most poorly understood component of the Mla pathway, exists as a monomeric, doughnut shaped, α-helical lipoprotein embedded entirely within the inner leaflet of the OM (Fig. 10). Purification has revealed a stable association with both OmpC and OmpF (Abellón-Ruiz et al., 2017; Chong et al., 2015), indeed the structure of MlaA was solved in complex with both OmpC and OmpF trimers (Abellón-Ruiz et al., 2017), however, photocross-linking has since implicated OmpC as the predominant binding partner of MlaA (Yeow et al., 2018). Additionally, deletion of OmpF under low osmolarity conditions, where its abundance exceeds that of OmpC, yielded no effect upon lipid asymmetry whist removal of OmpC disrupted OM lipid asymmetry (Chong et al., 2015). Furthermore, mutation of residues involved in the OmpC-MlaA interaction or OmpC trimerisation appeared to disrupt Mla activity suggesting an essential interaction between MlaA and OmpC trimers (Yeow et al., 2018). The structure of MlaA, solved in complex with OmpF, provides one obvious discrepancy (Fig. 10A) (Abellón-Ruiz et al., 2017), however, it must be noted that overexpression was performed within a strain devoid of OmpC. Nonetheless, the authors also solved the structure of the *Klebsiella pneumoniae* MlaA-OmpK36 (OmpC) complex and reported no significant structural differences (Abellón-Ruiz et al., 2017). Interestingly, deletion of OmpC/OmpF had no effect upon MlaA localisation, however, in the same background an MlaA mutant lacking the lipobox motif resided entirely within the IM (Chong et al., 2015). Finally, when grown in the presence of OmpC/F, the deacylated MlaA mutant partitioned between both membranes suggesting the OmpC/F-MlaA interaction alone confers OM localisation of MlaA (Chong et al., 2015). The role of OmpC in the function of MlaA remains unknown but it is speculated that OmpC trimers act purely to stabilise the structure and position MlaA within the inner leaflet of the membrane.

The structure depicts an amphipathic pore at the centre of the MlaA monomer through which GPLs are anticipated to transit (Fig. 10B-C). Atomistic and coarse grained molecular

dynamics simulations suggest outer leaflet GPLs bind MlaA via a ‘ridge’ structure at the pore apex whilst conformational changes may facilitate widening of the MlaA pore permitting the passage of bound GPLs (Fig. 10B-C) (Abellón-Ruiz et al., 2017).

A dominant gain-of-function mutation in the *mlaA* gene (designated *mlaA**), resulting in the deletion of N41 and F42, has been identified which confers a unique cell death pathway associated with disruption of lipid homeostasis. It is speculated that the *mlaA** mutation causes increased transit of GPLs into the outer leaflet where they act as a substrate for PldA. Elevated abundance of PldA products is anticipated to stimulate LPS synthesis (May and Silhavy, 2018), which once incorporated into the OM provokes membrane blebbing (Sutterlin et al., 2016). GPLs are trafficked to the OM, from the IM, to replace those lost in vesiculisation causing the IM to shrink and eventually rupture (Sutterlin et al., 2016). Two independent findings indicate the inner leaflet of the OM as the most likely source of GPLs translocated by MlaA*. First, the sensitivity of MlaA* to MacConkey agar remained constant upon deletion of *mlaC*, *mlaD* or *mlaF* implying the phenotype may not be accounted for by elevated anterograde transport through the entire Mla system (Sutterlin et al., 2016). Second, the structure of MlaA reveals the two residues deleted in MlaA* reside within helix H1, which shield the MlaA pore to inner leaflet GPLs, hence their deletion likely enables the lateral access of inner membrane GPL to the central pore (Abellón-Ruiz et al., 2017).

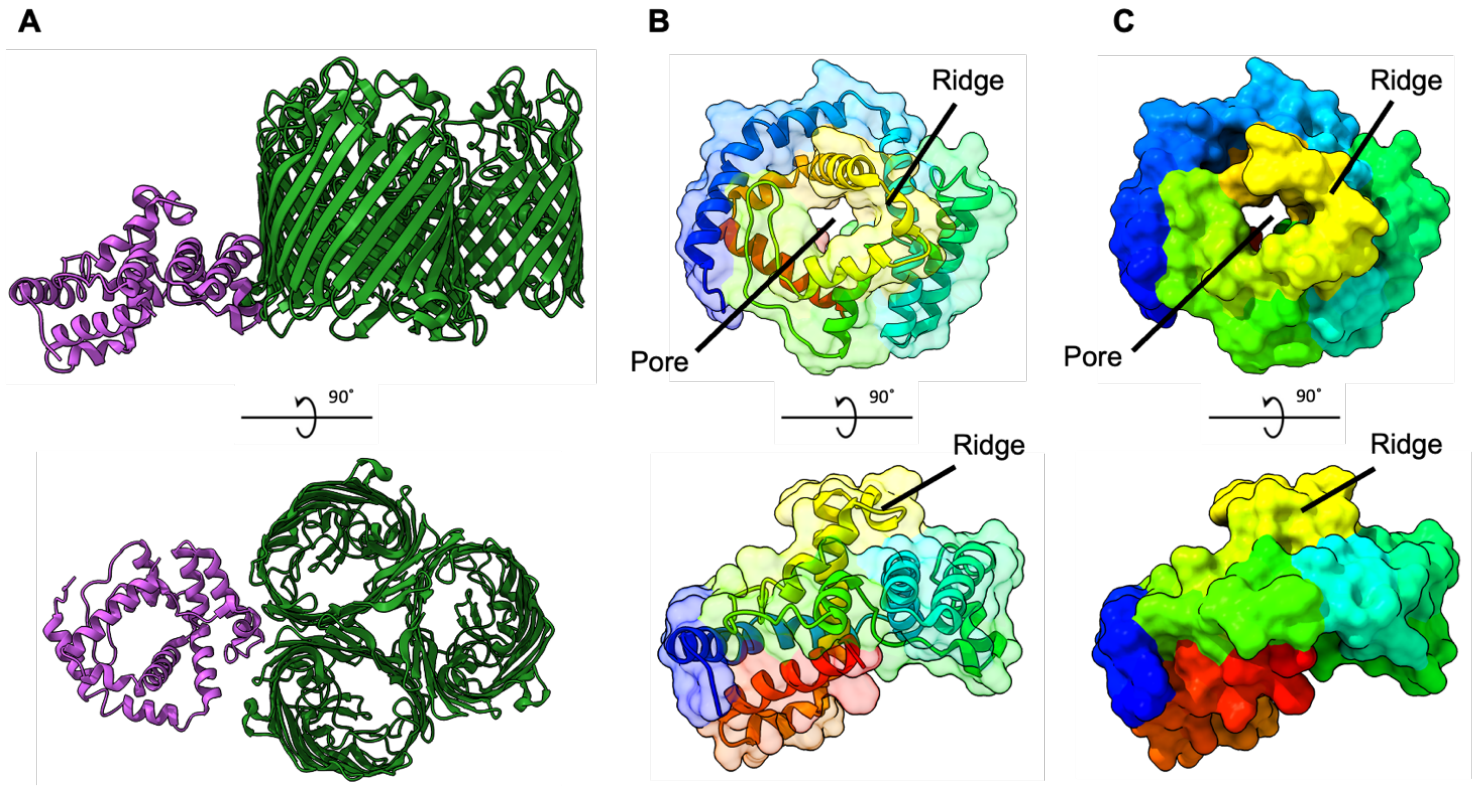


Fig. 10 – The structure of MlaA.

(A) Cartoon representation of the crystal structure of *S. marcescens* MlaA (purple) in complex with OmpF (green) indicating its integration into the membrane. (B) Cartoon representation of MlaA overlaid with a transparent surface both with rainbow colouring. The central pore and ridge structure of MlaA are indicated. (B) Surface representation of MlaA depicted with rainbow colouring. The central pore and ridge structure of MlaA are indicated. Figure adapted from (Abellón-Ruiz et al., 2017). MlaA-OmpF (PDB:5NUQ).

MlaC

The Mla system utilises a soluble carrier protein, MlaC, to bind GPLs and potentially shuttle them between membranes. MlaC represents the best studied protein within the system, perhaps owing to its solubility and relatively low molecular weight, with numerous studies providing both structural and biochemical insights into its sequestering of GPLs. When extracted from whole cells MlaC co-purifies with PE, PG, and potentially CL, with some groups having observed CL binding (Hughes et al., 2019) whilst others not (Ercan et al., 2019).

Central to GPL binding is a large hydrophobic pocket which, thanks to crystallographic studies (Ekiert et al., 2017; Hughes et al., 2019), has been shown to accommodate the acyl chains of all major *E. coli* GPL species. The headgroups of sequestered lipids therefore, appear inconsequential, however, additional research is required to validate this hypothesis. Inherent flexibility present within the hydrophobic cavity suggests MlaC may be able to fine-tune the topography of its binding site to accommodate different GPLs. Indeed, the electron density of a *Pseudomonas putida* MlaC structure (PDB: 4FCZ) shows density for four acyl chains and thus could represent the binding of either cardiolipin or two di-acyl GPLs within the hydrophobic pocket.

The evidence therefore points to a system in which MlaC functions as a generic GPL chaperone. Indeed, MlaC has been shown to interact with the IM MlaFEDB complex as well as the OM MlaA-OmpC complex both *in vitro* (Ekiert et al., 2017; Hughes et al., 2019) and *in vivo* (Ercan et al., 2019). The ability of MlaC to form a lipid-free (apo) conformation, resulting from a closing of the GPL binding cavity in response to a unique β-sheet pivot (Hughes et al., 2019), is suggestive of a molecule that can readily offload its GPLs upon interaction with its binding partners. However, the loading and unloading of GPLs by MlaC is reliant upon

coordination with one of its binding partners, whether that be MlaFEDB or MlaA, and is unable to simply sequester free GPLs from the environment (Hughes et al., 2019).

This use of a soluble carrier protein to facilitate GPL transport may be advantageous owing to its inherent plasticity and potential for rapid response to both external and internal signals, potentially allowing for a more targeted movement of lipids when required. The question therefore must be posed as to whether this system allows for the facilitation of bulk GPL transport or aids a more specific type of lipid transportation.

Resolving transport directionality

Whilst both genetic and biochemical evidence indicate that the Mla pathway functions as a lipid transport pathway, their proposals of transport directionality appear largely conflicting. The Mla pathway was originally characterised as a retrograde pathway (Malinverni and Silhavy 2009). This verdict was dependent upon two key genetic observations; firstly, mapping of Δ*mla* suppressor mutations to the *pldA* promoter region and secondly, indirect observation of outer leaflet GPL accumulation, via LPS hepta-acylation, upon removal of any pathway component (Malinverni and Silhavy, 2009). Following this initial interpretation, additional evidence supporting the retrograde Mla hypothesis has surfaced.

Increased OM vesicle (OMV) shedding has been observed in *H. influenzae and V. cholerae* Δ*mlaA* and Δ*mlaE* mutants, an observation since recorded within other organisms (Baarda et al., 2019; Kulp et al., 2015; Roier et al., 2016), whilst overexpression of the MlaFEDCB operon yielded reduced vesiculation. The exaggerated OMV shedding was interpreted to result from GPL accumulation within the OM implying that, under WT conditions, a retrograde Mla function may prevent such accumulation. Overexpression of the MlaFEDCB operon was also sufficient to partially offset GPL transport defects arising in Δ*tol-pal E. coli* cells

(Shrivastava et al., 2017). When monitored through membrane fractionation and mass spectrometry, *Salmonella enterica* Δ*tol-pal* mutants displayed elevated OM GPL levels, quantified by liquid chromatography - tandem mass spectrometry, implicating the Tol-Pal complex in retrograde GPL transport, however, the direct or indirect nature of this involvement remains unclear (Masilamani et al., 2018).

Evolution experiments in *A. baumannii* revealed both Δ*pldA* and Δ*mla* mutations increase growth and restore the OM barrier function in cells devoid of lipooligosaccharide (LOS) (Powers and Trent, 2018). The assumption being that in the absence of LOS, GPLs would populate both OM leaflets, hence any processes actively removing outer leaflet GPLs would be detrimental to the cell. Therefore, assuming the retrograde model, removal of the Mla pathway would prevent futile removal of GPLs and enable equilibrium to be established between the OM leaflets.

The structure of MlaA also provides evidence potentially supporting a retrograde model (Abellón-Ruiz et al., 2017). The observation that MlaA sits within and spans the entire inner leaflet appears to prevent both lateral diffusion of inner leaflet GPLs and their access to the central amphipathic pore, through which transport is anticipated to occur. Molecular dynamics simulations indicate binding of outer leaflet GPLs to a pronounced ridge structure before migrating into the central pore, suggestive of transport towards the periplasm (Abellón-Ruiz et al., 2017). Furthermore, considering an anterograde model, the ability of MlaA to only deposit GPLs extracted from MlaC into the outer leaflet, where they become accessible to PagP and PldA, seems counterproductive. Such transport would require an additional process relocating transported GPLs to the inner leaflet, however, no such mechanism has been identified. Finally, the high affinity of MlaC for GPLs further argues against anterograde GPL transport as it would be physiologically unrealistic for MlaA to extract GPLs from MlaC.

Despite these findings, there is also a significant body of evidence supporting an anterograde Mla model. *A. baumannii in vivo* pulse-chase assays have been used to monitor phospholipid accumulation within the inner and outer membranes (Kamischke et al., 2019). In *mla* knockouts accumulation of newly synthesised GPLs within the IM, and its relative absence from the OM was observed, indicative of compromised intermembrane GPL flux. However, the possibility of alternative mechanisms of GPL loss from the OM remains a major caveat, particularly loss via OMVs as increased shedding is a canonical Δ*mla* phenotype. Whilst OMV composition was not analysed, the authors stated that OM GPL loss was accounted for within the experimental design by measuring the labelled to unlabelled GPL ratio for each membrane. However, without measuring the total GPL levels, including those in OMVs, some doubt must remain. Indeed, this doubt prompted others to suggest the aforementioned experiments were compromised by the unintentional use of an unreliable Δ*mlaF* mutant strain (Δ$mlaF^{UW}$) (Powers et al., 2020). When compared with their own mutant strain Δ$mlaF^{UGA}$, Δ$mlaF^{UW}$ displayed a significant growth defect, increased antibiotic sensitivity and an altered transcriptomic profile. Whole genome sequencing revealed an unexpected SNP within the coding sequence for ObgE, a protein involved in the regulation and homeostasis of the (p)ppGpp alarmone response**.** The authors proposed the synthetic interaction between *obgE** and Δ*mlaF*, likely compromised lipid biosynthesis, and underpinned the previous observations rather than the disruption of anterograde transport. Finally, utilising their own pulse-chase assay monitoring OM vesiculation as a proxy for anterograde transport, as *Acinetobacter* membranes cannot be reliably fractionated, the authors concluded neither Δ*mlaF* mutant displayed irregular anterograde GPL transport.

To remove the complexity associated with whole cell assays, several studies have been published investigating the directionality of GPL transfer between the IM associated component MlaD and the soluble chaperone MlaC in isolation (Ercan et al., 2019; Hughes et al., 2019). In all cases transfer was observed only between holo-MlaD (lipid bound) and apo-MlaC (lipid unbound), never in the reverse scenario, supporting an anterograde transport

scenario. However, these studies were performed using MlaD in isolation, rather than as part of the MlaFEDB ATPase, hence the potential for a different outcome upon addition of the other complex components remained, particularly those involved in ATPase activity. Indeed, if the Mla pathway does have retrograde activity, MlaC would be required to have very tight affinity for GPLs to extract them from MlaA at the OM in the absence of an energy source. Thus, in isolation it is quite conceivable that MlaC extracts GPLs from MlaD, however, once in the presence of the ATPase, GPLs are pulled away from MlaC. To address this issue, similar transfer experiments have been performed between MlaC and the entire MlaFEDB complex reconstituted into liposomes (Hughes et al., 2019). Once again, transfer was observed towards apo-MlaC, not from holo-MlaC to MlaFEDB, corroborating the previous observations. Interestingly, GPL transfer was found to occur independently of ATP hydrolysis, however, upregulation of ATPase activity was observed only in the presence of MlaC-apo, suggestive of MlaC-apo being the substrate for MlaFEDB rather than MlaC-holo. However, a more recent study utilising a FRET based *in vitro* liposomal transport assay of the whole Mla pathway, reported both ATP-dependent retrograde and ATP-independent anterograde GPL transport (Tang et al., 2020). These findings potentially reconcile both anterograde and retrograde camps, suggesting the Mla pathway is capable of bidirectional transport though, considering the cytoplasmic ATP concentration, retrograde would be favoured physiologically. It must be noted that owing to the orientation of MlaFEDB within the liposome, and the impermeability of ATP to the liposomal lumen, it is unclear how the aforementioned assays were able to concurrently monitor GPL transport and ATP hydrolysis. Although compelling, clearly, additional studies are required to provide clarity in the role the ATPase plays in retrograde and/or anterograde transport. A recent study appears to address this issue utilising a series of *in-vitro* radioactive lipid transport assays (Low et al., 2021). The rate of [^{14}C]-labelled GPL transport from MlaC to nanodisc reconstituted MlaFEDB in the presence of ATP was observed to be approximately 1.5x that in the absence of ATP indicating ATP enhances the transfer of lipids from MlaC to MlaFEDB over a spontaneous baseline level. This notion was supported by assays performed with the MlaF^{K47A}EDB

ATPase dead mutant and ATPase inhibitor vanadate which displayed no increased transfer in the presence of ATP as well as those performed with the non-hydrolysable ATP analogue AMP-PNP (Low et al., 2021). Spontaneous transfer in the inverse scenario, from nanodisc reconstituted MlaFEDB to MlaC, was observed, however, strikingly the presence of ATP completely abolished all MlaFEDB to MlaC transmission suggesting that the physiological presence of ATP promotes unidirectional retrograde transport throughout the Mla pathway.

Alternative Putative Gram-negative GPL Transport Systems

Collectively, Tang et al., and Low et al., offer the most compelling evidence yet for the involvement of the Mla pathway in retrograde transport. Indeed, a consensus verdict that the Mla pathway functions physiologically, at least predominantly, in a retrograde manner is becoming widely adopted throughout the scientific community, however, this raises an obvious question. If not Mla, then what is the identity of the proteinaceous anterograde GPL transport pathway in Gram-negative bacteria, assuming such a system exists. The search for this apparently illusive pathway has revealed several putative GPL transport pathways; Tol-Pal, PbgA, YhdP, Mla, Pqi, Yeb (Let), however, as yet the evidence to categorically implicate any one of these in GPL transport is lacking.

Tol-Pal

The Tol-Pal pathway is highly conserved across Gram-negative bacteria consisting of 5 proteins which assemble into two sub-complexes. The IM houses the TolQRA molecular motor, with estimated stoichiometry of 4-6:2:1, whilst the OM lipoprotein Pal localises periplasmic TolB to the OM periphery (Fig. 11) (Lloubès et al., 2001). Interaction between the two complexes is dependent upon the PMF which drives structural rearrangements in TolA, resulting in its extension across the periplasm, where it forms weak interactions with Pal and potentially TolB (Germon et al., 2001).

The precise function of the Tol-Pal complex remains unclear, however, mutants exhibit elevated sensitivity to antibiotics and detergents, increased leakage of periplasmic contents and increased membrane blebbing, hence its implication in maintaining OM integrity and stability (Lloubès et al., 2001). A role in OM invagination during cell division has also been

proposed due to the active recruitment of complex components to the septum during the cell cycle (Petiti et al., 2019). A recent study utilising steady-state radioactive labelling of cellular lipids, revealed cells lacking the Tol-Pal complex accumulate excess GPLs in the OM, but maintain typical levels of LPS, implying disruption of bulk retrograde GPL transport (Shrivastava et al., 2017). The same study revealed overexpression of the MlaFEDBC operon, but not MlaA, was partially capable of rescuing this phenotype suggesting a degree of cooperativity between the two pathways.

It remains unclear if these phenotypes arise directly from Tol-Pal removal or represent the cumulative effect of disrupted OM processes, including an as yet unidentified retrograde GPL transport system, in the absence of Tol-Pal. Assuming the direct involvement of Tol-Pal in GPL transport, three models have been proposed. Firstly, the Tol-Pal complex may directly bind and transport GPLs, however, current component structures fail to demonstrate GPL binding, placing significant doubt upon their lipid trafficking capabilities. This issue is overcome by the second model in which Tol-Pal components interact with a currently unidentified GPL binding protein. Finally, the two membranes may be brought into close proximity via the activated Tol-Pal complex facilitating direct diffusion of GPLs via hemifusion events. Currently, the experimental evidence necessary to clarify the physiological role of the Tol-Pal complex is lacking, hence the mystery surrounding its involvement in GPL transport, if any, persists.

PbgA

During enteric bacterial infection the PhoPQ system initiates remodelling of cell surface components within the OM, increasing hydrophobicity and decreasing negative lipid charge, preventing binding and insertion of cationic antimicrobial peptides (CAMPs) (Dalebroux et al., 2014; Dalebroux and Miller, 2014). Recent genetic screening suggested PhoPQ mediates insertion of additional OM cardiolipin, boosting OM hydrophobicity, via the IM protein PhoPQ barrier gene A (PbgA) by an unknown mechanism (Dalebroux et al., 2015) (Dalebroux *et al.* 2015).

PbgA (formerly YejM) consists of a 5TM N-terminal domain and a large globular periplasmic domain connected via a dynamic linker. The oligomeric state of PbgA remains unclear with an initial paper suggesting a homo-tetrameric arrangement (Dalebroux et al., 2015), however, later publications observed monomeric species (Clairfeuille et al., 2020; Fan et al., 2020) (Fig. 11). Interestingly, under normal growth conditions the periplasmic domain appears dispensable, whereas the TM region remains essential for cell viability (Dalebroux et al., 2015). PbgA possesses a long, deep cleft running along its surface which, within the TM domain, is lined with hydrophobic residues before becoming increasingly negatively charged within the periplasmic domain (Fan et al., 2020).

PhoPQ signalling elevates transport of cardiolipin and other acidic GPLs to the OM (Dalebroux et al., 2014; Dalebroux and Miller, 2014). Deletion of the globular PbgA periplasmic domain was concurrent with constant OM cardiolipin levels during PhoPQ signalling, hence the protein was implicated in OM cardiolipin trafficking (Dalebroux et al., 2015). An ensuing crystal structure of PbgA in complex with two cardiolipin molecules (PDB: 6V8Q) (Fan et al., 2020) appeared to support this notion, however, the physiological relevance of this structure is questionable considering cardiolipin was supplemented during the crystallisation process. Indeed, the authors themselves acknowledge one of the CL

molecules is likely an artefact of crystal packing (Fan et al., 2020). Several independent publications have cast further doubt on this proposed role, instead presenting PbgA as a periplasmic LPS sensor, including a subsequent crystal structure in complex with a co-purifying LPS molecule (PDB: 6XLP) (Cian et al., 2020; Clairfeuille et al., 2020; Fivenson and Bernhardt, 2020; Guest et al., 2020). This interaction is likely more physiologically relevant considering the co-purification of ligand rather than its supplementation during crystallisation. The same publication presents a comprehensive model in which PbgA regulates the abundance of LpxC in response to periplasmic LPS levels, via interactions with LapB and FtsH, ensuring the synchronisation of LPS synthesis and OM transport (Clairfeuille et al., 2020).

The PbgA TM domain is a member of the DUF3413 family, exhibiting the conserved D-X9-H motif, raising the possibility of a putative catalytic dyad. Missense mutation of these residues only marginally affected OM barrier function fuelling speculation of their involvement in a distinct or regulatory role (Fan et al., 2020). Additional catalytic activity of PbgA remains a prevailing topic with a study suggesting the PbgA periplasmic domain is a Mg^{2+} dependent metallophosphatase whose activity is directly coupled to the transport of CL to the OM (Gabale et al., 2020). Furthermore, the potential binding of a single phosphatidylethanolamine lipid within the TM cleft of PbgA in its LPS complexed crystal structure maintains speculation of PbgA's involvement in lipid trafficking (Clairfeuille et al., 2020).

A prevailing consensus that PbgA functions primarily as an LPS sensor, rather than cardiolipin transporter, appears to have arisen, however, considering its potential catalytic activity, impact upon OM CL levels and speculative lipid binding, the possibility of additional functionalities merits further investigation.

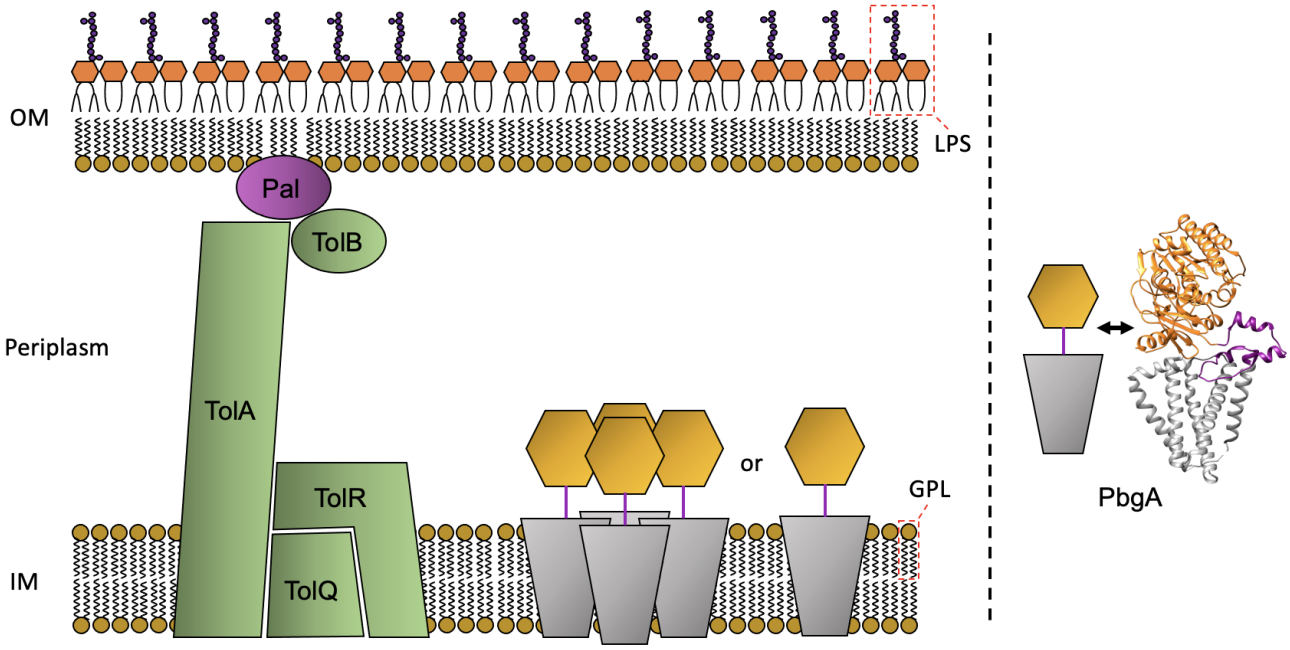


Fig. 11 – The Tol-Pal and PbgA Pathways.

Cartoon representation of the Tol-Pal and PbgA pathways in their respective cellular locations. The transmembrane, linker and periplasmic domains of PbgA are coloured grey, purple and orange respectively. PbgA (PDB: 6V8Q).

YhdP and “AsmA-like” proteins

Recent genetic screens have provided a tantalising clue to yet another protein possibly involved in Gram-negative lipid transport (Grimm et al., 2020). Screening for genetic interactions with the *mlaA** mutant led to the identification of YhdP, a 1266 residue inner membrane associated protein with one clear N-terminal and a potential second C-terminal transmembrane domain, implying it alone may span the periplasm and connect both membranes (Grimm et al., 2020). Deletion of *yhdP* in *mlaA** cells slows the IM-to-OM GPL flux, prolonging cell death, however, the inability to replace GPLs lost from the OM efficaciously instead results in OM rupture (Grimm et al., 2020).

YhdP belongs to the “AsmA-like” protein family which is highly conserved amongst didermic bacteria. *E. coli* itself harbours six “AsmA-like” proteins: YhdP, TamB, YdbH, YicH, YhjG and AsmA (Fig. 12A). Two of these proteins, AsmA and TamB, display homology at their N-termini with the eukaryotic GPL transporter Vps13 (Levine, 2019), which has been shown to contain a channel through which GPLs are transported between membranes (Kumar et al., 2018; Li et al., 2020). “AsmA-like” proteins are typified by their high β-sheet composition, adopting a concave β-taco fold with a hydrophobic interior, somewhat reminiscent of the β-jellyroll fold utilised within the Lpt machinery (Josts et al., 2017). Furthermore, the predicted β-sheet construction of all “AsmA-like” proteins emulates that of Vps13 as well as Atg2, a protein believed to transport endoplasmic reticulum GPLs during autophagosome formation (Fig. 12) (Osawa et al., 2019). Utilising differing genetic approaches, two recent studies revealed the three largest *E. coli* AsmA-like proteins, YhdP, TamB and YdbH, to be redundant in maintaining OM GPL homeostasis further strengthening the association between AsmA-like proteins and GPL transport (Douglass et al., 2022; Ruiz et al., 2021). Furthermore, both studies reported a triple knockout of YhdP, TamB and YdbH to be lethal indicating constitutive function of “AsmA-like” proteins under WT conditions (Douglass et al., 2022; Ruiz et al., 2021).

Considering the aforementioned evidence, there is a strong consensus implicating the “AsmA-like” protein family in GPL transport, however, subsequent structural and biochemicals studies are required to reveal both their transport mechanism and involvement in basal GPL transport.

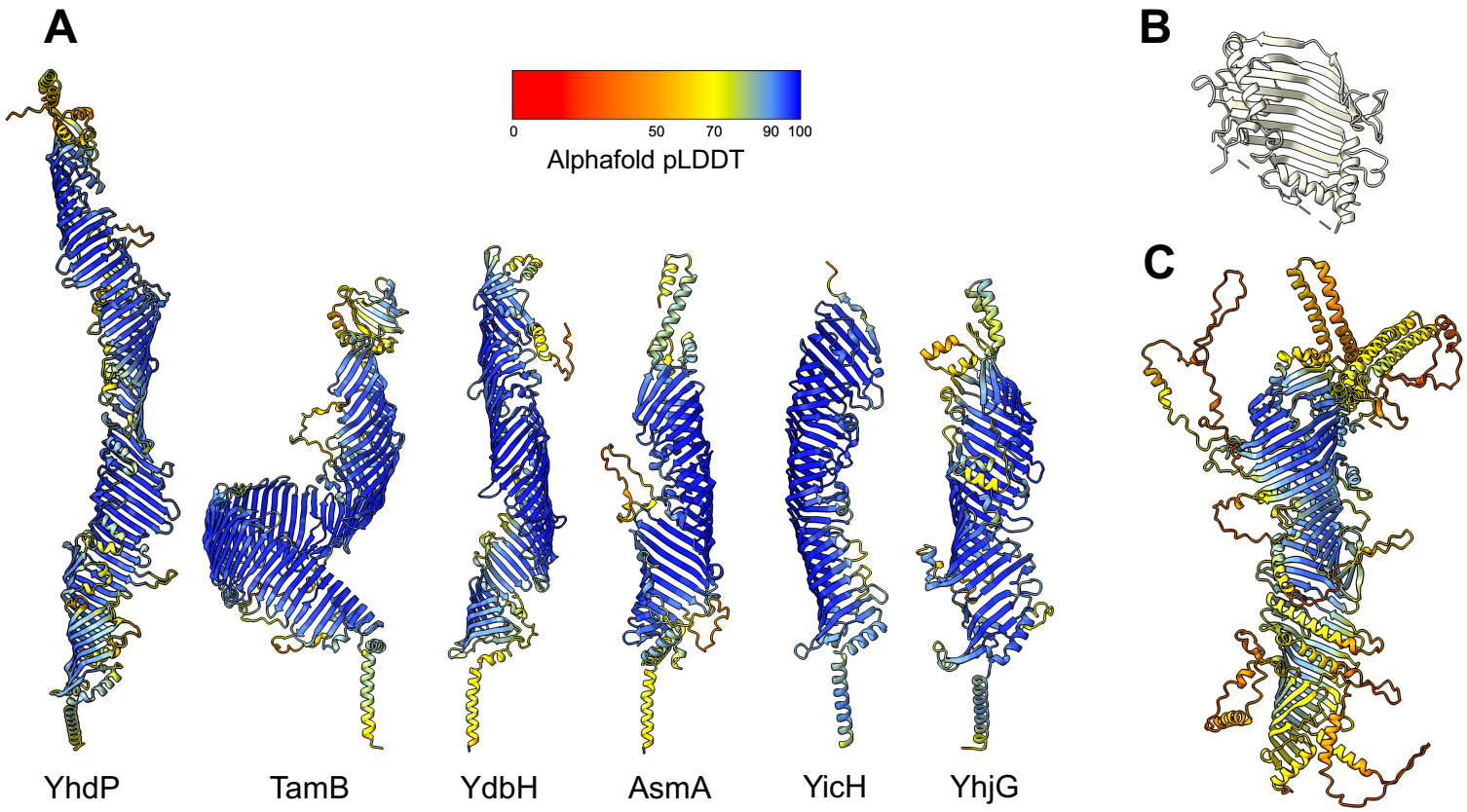


Fig. 12 – The "AsmA-like" Protein Family and Homologues.

(A) AlphaFold predicted structures for the six *E. coli* "AsmA-like" protein family members depicted in cartoon and coloured according to AlphaFold's per-residue confidence measure (pLDDT) (Jumper et al., 2021). (B) Cartoon representation of the crystalised N-terminal region of *C. thermophilum* Vps13 indicating the B-sheet rich nature of the protein (PDB: 6CBC). (C) Cartoon representation of the Alphafold predicted structure for Atg2, indicating its similarity to the "AsmA-like" proteins, coloured according to pLDDT (Jumper et al., 2021).

MCE/MlaD Domains

Despite tentative implications relating to the aforementioned proteins and complexes, the Mla pathway remains the only confirmed proteinaceous Gram-negative phospholipid transport system. The MCE/MlaD domain harboured by the MlaD component of the Mla pathway is a particularly intriguing domain that may indicate a range of novel transport systems potentially including those of GPLs.

The Advent of Mammalian Cell Entry (MCE)

The introduction of a cloned *Mycobacterium tuberculosis* fragment into a non-pathogenic *E. coli* strain (XL1-Blue) conferred the cells the ability to access and inhabit HeLa cells and macrophages (Arruda et al., 1993). The genetic locus enabling cell entry was identified as a 1362 bp ORF, named mammalian cell entry or *mce1*, encoding a ~45 kDa protein (Mce1) (Chitale et al., 2001). Polystyrene latex beads coated in Mce1 were observed to be internalised by non-phagocytic HeLa cells and settled within vacuolar compartments (Chitale et al., 2001). Subsequent experiments assessed the ability of two truncated Mce1 proteins, $Mce1^{105\text{-}307}$ and $Mce1^{163\text{-}307}$ (Fig. 13) to facilitate latex microsphere uptake. Interestingly microspheres coated in $Mce1^{105\text{-}307}$ were internalised at a greater rate than those coated in the full length Mce1 ($Mce1^{1\text{-}454}$) whilst $Mce1^{163\text{-}307}$ coated beads displayed no association with HeLa cells even after 24 hours (Chitale et al., 2001). These findings attribute the ability of Mce1 to enter non-phagocytic cells to a 58 amino acid domain, between residues 105 and 163, of Mce1 which subsequently became known as the **M**ammalian **C**ell **E**ntry (MCE) or more recently MCE/MlaD domain (Chitale et al., 2001).

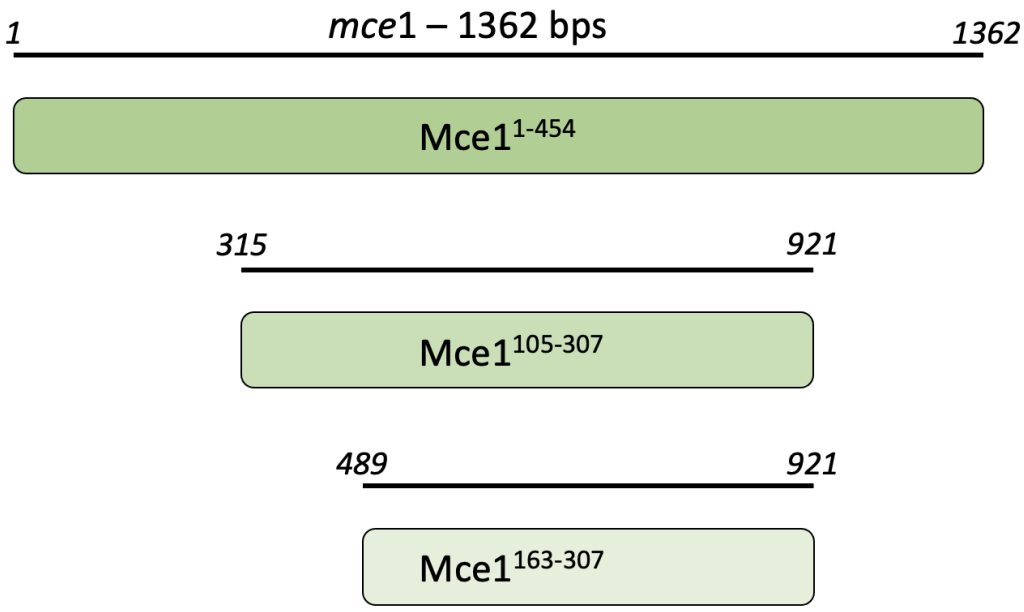


Fig. 13 – The Mce1, Mce1$^{105-307}$ and Mce1$^{163-307}$ proteins and respective DNA segments. Black lines represent the DNA sections of the *M. tuberculosis* H37Rv Mce1 loci (Rv-0169) responsible for the Mce1, Mce1$^{105-307}$ and Mce1$^{163-307}$ proteins. Black italic numbers indicate the positions in the genomic sequence where truncations were made. The Mce1 protein resulting from each DNA section is indicated by the associated green boxes below with the constituent amino acid residues indicated by superscript numbers. Figure adapted from (Chitale et al., 2001).

MCE Proteins

The MCE/MlaD domain is widely conserved, indeed the Pfam hidden Markov model (HMM) for the MCE/MlaD domain (PF02470) includes entries from 4504 species (Finn et al., 2016). Proteins containing MCE/MlaD domain(s) are termed MCE/MlaD proteins, referred to as 'MCE proteins' hereafter. A recent bioinformatic study revealed >97% of identified bacterial MCE proteins conform to four core architectures, termed type I to IV respectively (Fig. 14) (Isom et al., 2017). Type I bacterial MCE proteins consist of a solitary MCE/MlaD domain and represent by far the most abundant architecture whilst Type II MCE proteins are unique to actinobacteria and contain a single MCE/MlaD domain in tandem with a DUF3407 (cholesterol uptake porter) domain. Finally, Type III and IV MCE proteins are found exclusively in proteobacteria and contain 3 and 7 consecutive MCE/MlaD domains respectively (Isom et al., 2017).

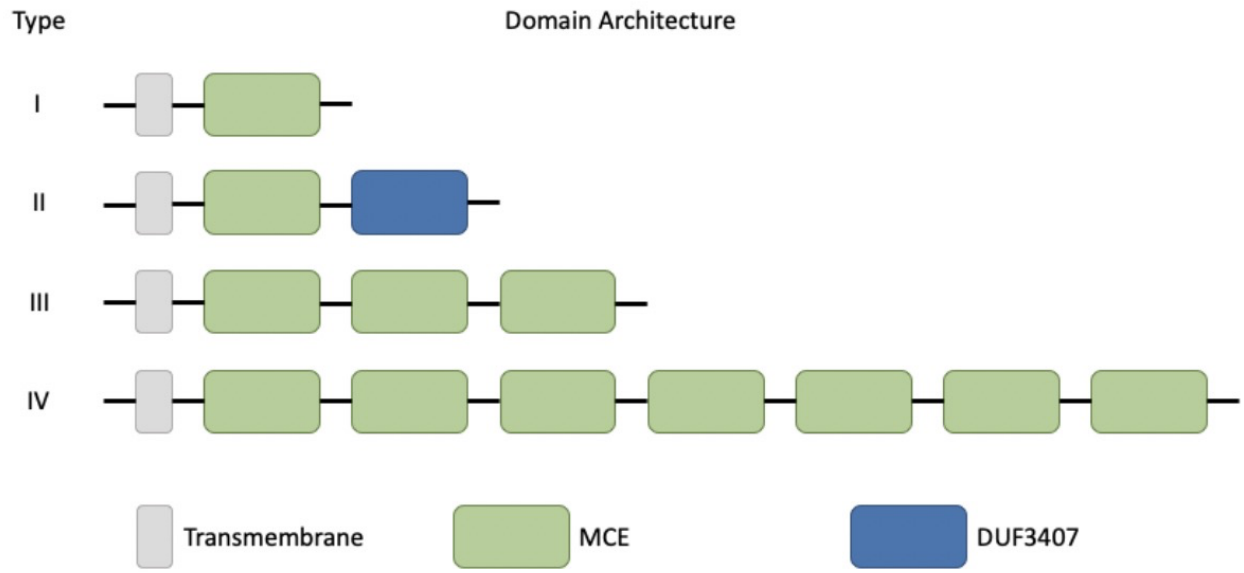


Fig. 14 – Core architectures of bacterial MCE proteins.

Grey, Green and Blue boxes represent transmembrane, MCE and DUF3407 domains respectively whilst black lines indicate connecting sequences. Figure reproduced from (Isom et al., 2017).

MCE Proteins in Lipid Transport?

Given the myriad of bacteria possessing MCE proteins, particularly those pursuing a non-invasive lifestyle, it seems likely that at least a subset of MCE proteins facilitate a physiological function distinct from cell entry.

To date, the majority of research has focussed upon elucidating the role of MCE proteins in Actinobacteria. Actinobacterial species contain differing numbers of *mce* operons, however, all actinobacterial *mce* operons conform to a canonical configuration consisting of 2 *yrbE* permease genes (A and B) followed by 6 *mce* genes (A – F) which encode macromolecular

transport complexes homologous to ATP-binding cassette (ABC) transporters (Casali and Riley, 2007; Kumar et al., 2005). The *mce* operons of Actinobacteria are annotated using a simple nomenclature, first proposed by Tekaia, in which each of the two *yrbE* genes and six *mce* genes are referred to as *yrbE**X**A-B* and *mce**X**A-F* where ***X*** is a numerical integer indicating the operon number (Fig. 15) (Tekaia et al., 1999).

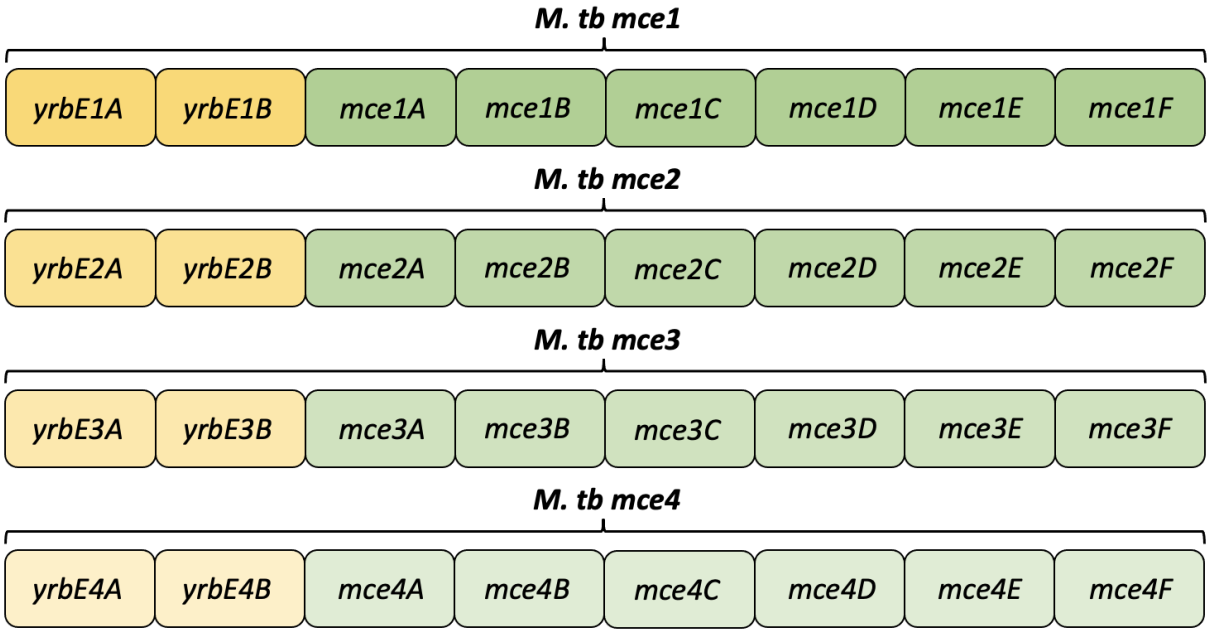


Fig. 15 – Actinobacterial *mce* operon nomenclature.

The four *mce* operons of *Mycobacterium tuberculosis* (*M. tb mce* 1-4) are utilised to illustrate the Actinobacterial *mce* operon nomenclature originally proposed by Tekaia. The numerical character, constant throughout each horizontal, indicates the operon number whilst the alpha character, constant throughout each column, indicates the genes position within the operon.

Expression of mycobacterial *mce* operons appears dependent upon both the cell's growth stage (Singh et al., 2016) and progression through the infection process (Kumar et al., 2005). Thus, MCE proteins may facilitate the adaption of mycobacteria to adverse conditions within the host cell. Indeed, a recent review serves as an excellent summary of the current perspective of mycobacterial *mce* operons as lipid transporters maintaining cell wall homeostasis under stress conditions (Klepp et al., 2022). Furthermore, the expression of each *M. tuberculosis mce* operon is mediated by a specific regulator (Casali et al., 2006; de La Paz et al., 2009; Forrellad et al., 2014; Kendall et al., 2007; Santangelo et al., 2009, 2008; Vindal et al., 2008) presumably enabling induction in response to particular external stressors including but not limited to starvation, hypoxia, surface stress and oxidative stress. Multiple studies, both phylogenetic (Casali and Riley, 2007; Dunphy et al., 2010; Joshi et al., 2006) and biochemical (Forrellad et al., 2014), have reported *mce* operons to encode transport systems that traffic lipids across the cell wall whilst others document the apparent co-expression of *mce* operons and genes related to lipid metabolism, further suggestive of their involvement in lipid trafficking (Casali et al., 2006; de La Paz et al., 2009; Kendall et al., 2007). The *mce4* operon of *M. tuberculosis* has been shown to encode a cholesterol importer (Pandey and Sassetti, 2008) furthermore, MCE4 proteins have been identified as essential for cholesterol uptake in *M. smegmatis* (Klepp et al., 2012; Rank et al., 2021) implicating a conserved involvement of the *mce4* operon, in cholesterol import, in both pathogenic and non-pathogenic mycobacterial species.

Further evidence of the involvement of MCE proteins in lipid transport can be found in the plant kingdom. The TGD2 protein of *Arabidopsis thaliana* has been implicated as the substrate binding component of an ABC-like transporter (Awai et al., 2006). The entire complex consisting of TGD2 along with TGD1 permease and TGD3 ATPase components is anticipated to facilitate the import of ER derived phosphatidic acid into the chloroplast plastid (Awai et al., 2006). Indeed, a soluble TGD2 construct, lacking the N-terminal transmembrane domain, was shown to specifically bind phosphatidic acid within its large C-terminal domain

(Awai et al., 2006). A subsequent study utilised a series of truncations and deletions to isolate a 25 amino acid fragment of TGD2, located within its single MCE domain, which alone was sufficient for phosphatidic acid binding (Lu and Benning, 2009).

The multi-MCE domain containing protein Multivalent Adhesion Molecule 7 (MAM7), is a purported constitutively expressed, outer membrane anchored protein with 7 sequential MCE domains which protrude from the cell surface (Krachler et al., 2011). An array of MAM proteins containing 6 or 7 tandem MCE domains appear widely conserved amongst Gram-negative bacteria, with the obvious omission of *E. coli,* whilst being absent from both Gram-positive and plant pathogenic bacteria (Krachler et al., 2011). Recombinantly expressed MAM7 from a range of Gram-negative pathogens conferred a non-adhesive *E. coli* BL21 strain the ability to attach to 3T3 cells, as such, MAMs have been classified as a distinct group of bacterial adhesins (Krachler et al., 2011). MAM7 mediated host cell attachment reportedly involves the binding of MAM7 to two surface exposed constituents, phosphatidic acid phospholipids and fibronectin (Lim et al., 2014).

The wide conservation of MAMs within pathogenic Gram-negative species identifies them as an attractive target for the development of novel antimicrobial therapeutics. Indeed, MAM7 inhibitors present a promising novel prophylactic treatment, and several works document their encouraging development (Al-Saedi, 2018; Beltran et al., 2016; Hawley et al., 2013; Huebinger et al., 2020; Krachler et al., 2012).

MAMs presents a functionality of MCE proteins, cell attachment, distinct from their alleged canonical role as transport complex components, however, their utilisation of phosphatidic acid phospholipids provides a degree of concordance with the MCE proteins discussed thus far. Notwithstanding these findings, an *E. coli* MAM7 homologue, YebT (LetB) remains firmly presented within scientific literature as a potential Gram-negative trans-envelope phospholipid transport pathway.

E. coli MCE proteins

Indeed, *E. coli* contains three MCE domain containing proteins, MlaD, YebT (also recently referred to as LetB) and PqiB, each residing within unique operons and assembling IM associated, homo-hexamers *in vivo* (Fig. 16) (Ekiert et al., 2017; Nakayama and Zhang-Akiyama, 2017). YebT and PqiB, and the operons within which they reside, have become of principal interest as significant candidates for a novel Gram-negative phospholipid transport system. Presently the literature focuses upon YebT, however, a large degree of overlap is anticipated in the mechanics of these two proteins considering their own striking similarity and that of their additional operonic components.

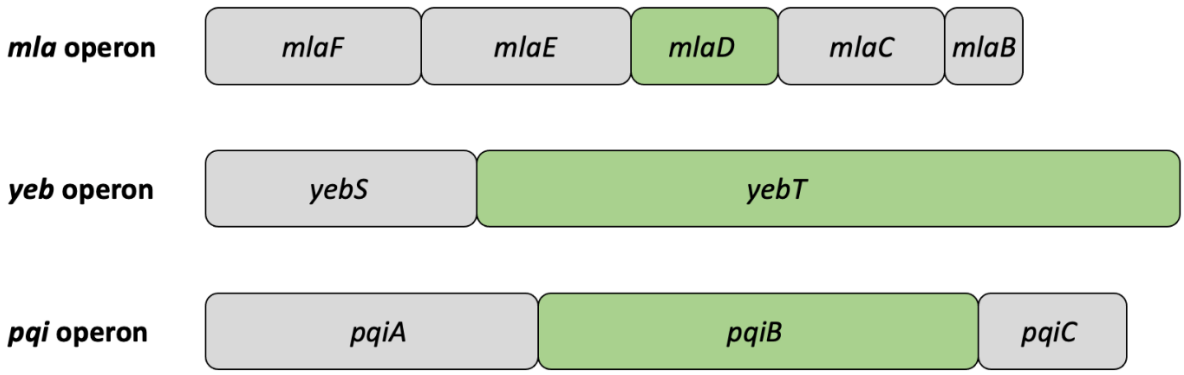


Fig. 16 – *E. coli* MCE protein containing operons.

The three MCE protein containing operons in *E. coli* namely the *mla, yeb* and *pqi* operons which harbour MlaD, YebT and PqiB respectively. MCE containing genes are coloured green whilst the length of each bar corresponds to the size of each gene.

YebST (LetAB)

YebT, encoded by the second of two genes located within the *yebST* operon, is the largest *E. coli* MCE protein comprising seven sequential MCE domains homologous to MAM7 discussed above. Each MCE domain folds to form a seven stranded β-barrel, akin to that of MlaD, resulting in seven vertically stacked β-barrel assemblies per YebT monomer which compose seven hollow ring structures in the final hexameric assembly, a hallmark of MCE proteins (Fig. 17A-B) (Ekiert et al., 2017). The YebT hexamer is of sufficient height to conceivably traverse the periplasm, its seven stacked MCE domains alone being sufficient to achieve the approximately 220 Å required, however, if/how YebT interacts with the OM, directly or otherwise, remains to be determined (Ekiert et al., 2017).

A continuous central hydrophobic channel runs the entire length of the YebT periplasmic domain, with an average diameter of 14 Å, providing a shielded environment large enough for GPLs and other small molecules to conceivably pass (Isom et al., 2020). Apparent pivoting of the MCE domains enables modulation of the central pore aperture size and thus may provide a regulatory mechanism for the transport speculated to occur through YebT (Isom et al., 2020). The same publication assessed the function of YebT mutants possessing fewer MCE domains via a lauryl sulfobetaine (LSB) complementation assay. Whilst all mutants assembled into homo-hexamers with the expected number of ring structures, no mutant with less than five MCE domains was able to restore barrier function during the assays (Isom et al., 2020). These findings led the authors to conclude that the length of YebT is crucial for its function hence five MCE domains represents the minimum required to span the periplasm (Isom et al., 2020).

A follow up publication generated an array of YebT mutants, each lacking one of the seven MCE domains (ΔRing 1-7), to investigate their relative importance and elucidate any specific functionalities. Surprisingly, deletion of the third and sixth MCE domains (ΔRing 3/6)

conferred an even higher level of resistance to LSB than the WT protein (Vieni et al., 2022). In the WT YebT protein, the sixth MCE domain appears to undergo the largest conformational change (Isom et al., 2020) and is thus likely crucial to the regulation of any internal transport, hence its deletion likely de-restricts any transport through its central pore.

Each MCE domain possesses a pore lining loop (PLL), so called as they protrude from the MCE domain towards the centre of the homo-hexameric assembly and form the surface of the central YebT pore (Ekiert et al., 2017). Replacement of the PLL from the sixth MCE domain of YebT with that of the third MCE domain, the only one of similar length, once again conferred a greater resistance to LSB than the WT, comparable to that of the ΔRing 6 mutant (Vieni et al., 2022). The PLL comprises two distinct structural features: 1 – a short horizontal helix which sits perpendicular to the length of the pore and 2 – a hairpin loop which extends vertically along the length of the central pore (Ekiert et al., 2017). One of two hydrophobic motifs has been identified within the PLLs of all *E. coli* MCE domains and appear essential for the function of YebT, PqiB and MlaD (Ekiert et al., 2017; Vieni et al., 2022). Finally, the authors mutated the hairpin and helix sections of the sixth YebT MCE domain, to contain those from the third MCE domain, independently. In both instances, the mutant protein displayed increased resistance to LSB suggesting that interactions between the entire WT PLL and its local environment are essential for correct YebT functionality (Vieni et al., 2022).

YebS (also recently referred to as LetA), is the first product of the *yebST* operon and thus anticipated to form a complex with YebT. Remarkably, YebS displays no sequence homology with any known transporter protein nor ATP binding domain suggesting that, if it does function as a transporter, it does so either passively, via an as yet unidentified ATPase, or potentially the PMF. Topological modelling predicts eight TM helices split into two consecutive PqiA domains (Pfam PF04403), each anticipated to span the IM four times, a notion supported by AlphaFold (Fig. 17C).

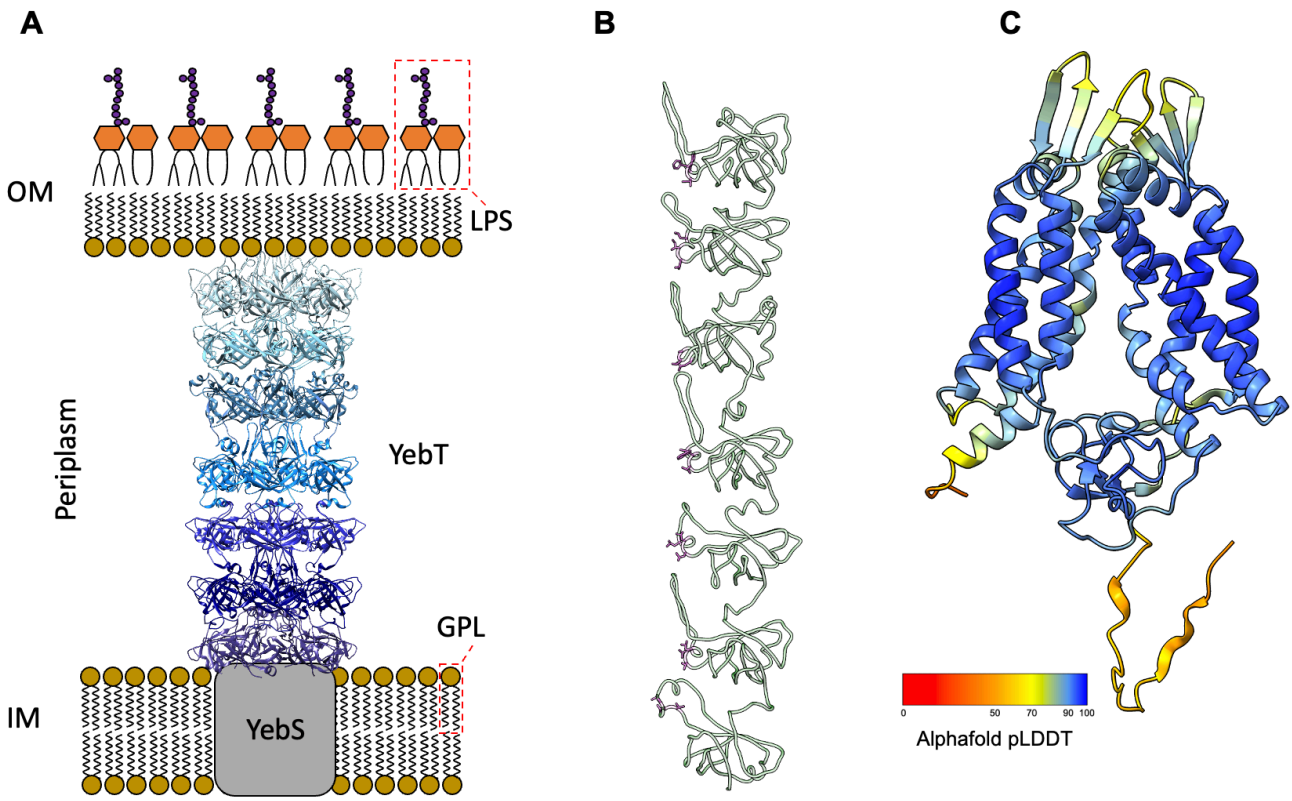


Fig. 17 – Structures of YebS and YebT.

(A) Depiction of YebT in its proposed cellular location, spanning the periplasm, with MCE domains are coloured in an array of blues. YebS is indicated by the grey box within the inner membrane (PDB: 6V0C). (B) Single protomer of YebT highlighting the vertical organisation of the seven MCE domains. The residues constituting the hydrophobic motifs found in all *E. coli* MCE domains are coloured in purple. YebT (PDB: 6V0C). (C) Cartoon representation of the AlphaFold predicted structure for YebS coloured according to Alphafold's per residue confidence measure (pLDDT) (Jumper et al., 2021).

PqiABC

PqiB is encoded by the second of three genes residing within the paraquat-inducible (*pqi*) operon. The functional homo-hexameric oligomer adopts a syringe-like structure, approximately 230 Å in length, also proposed to span the entire periplasmic space (Ekiert et al., 2017). Each PqiB monomer consists of three consecutive MCE domains and a C-terminal helix which, once assembled into the functional hexamer, contribute to the N-terminal, triple ring-barrel assembly and coiled-coil needle projection respectively (Fig. 18A-B) (Ekiert et al., 2017). PqiB also harbours a central hydrophobic channel running its entire length, akin to that of YebT, through which transport is anticipated (Ekiert et al., 2017). Much of the structural information elucidated from studies on YebT have been proven to or are anticipated to be directly relevant to PqiB. Indeed, when presenting the pivoting of the MCE domains in YebT, the authors also re-processed their previous PqiB dataset to reveal the same domain flexion in both proteins (Isom et al., 2020).

The two remaining operonic genes; *pqiA* and *pqiC* encode an integral IM protein and an OM associated lipoprotein respectively (Fig. 18C-D) (Nakayama and Zhang-Akiyama, 2017). PqiA is predicted to adopt a topology analogous to that of YebS, accommodating two PqiA domains, however, experimental structural evidence relating to either *E. coli* protein remains enigmatic. PqiC is predicted to anchor within the periplasmic face of the outer membrane and comprises an ABC auxiliary lipoprotein domain (formerly DUF330). Whilst the structure of the *E. coli* protein remains unsolved, insight can be gained via a 1.45 Å structure of *Enterobacteria cloacae* PqiC (PDB: 6OSX) (Fig. 18D), however, the structure provides little insight into the function of PqiC nor how it interfaces with PqiB. Indeed, interactions between all complex components, including stoichiometry, remain uncharacterised; nevertheless, once assembled into their quaternary complex the trio of Pqi proteins are anticipated to provide a continuous channel between the inner and outer membrane of diderms.

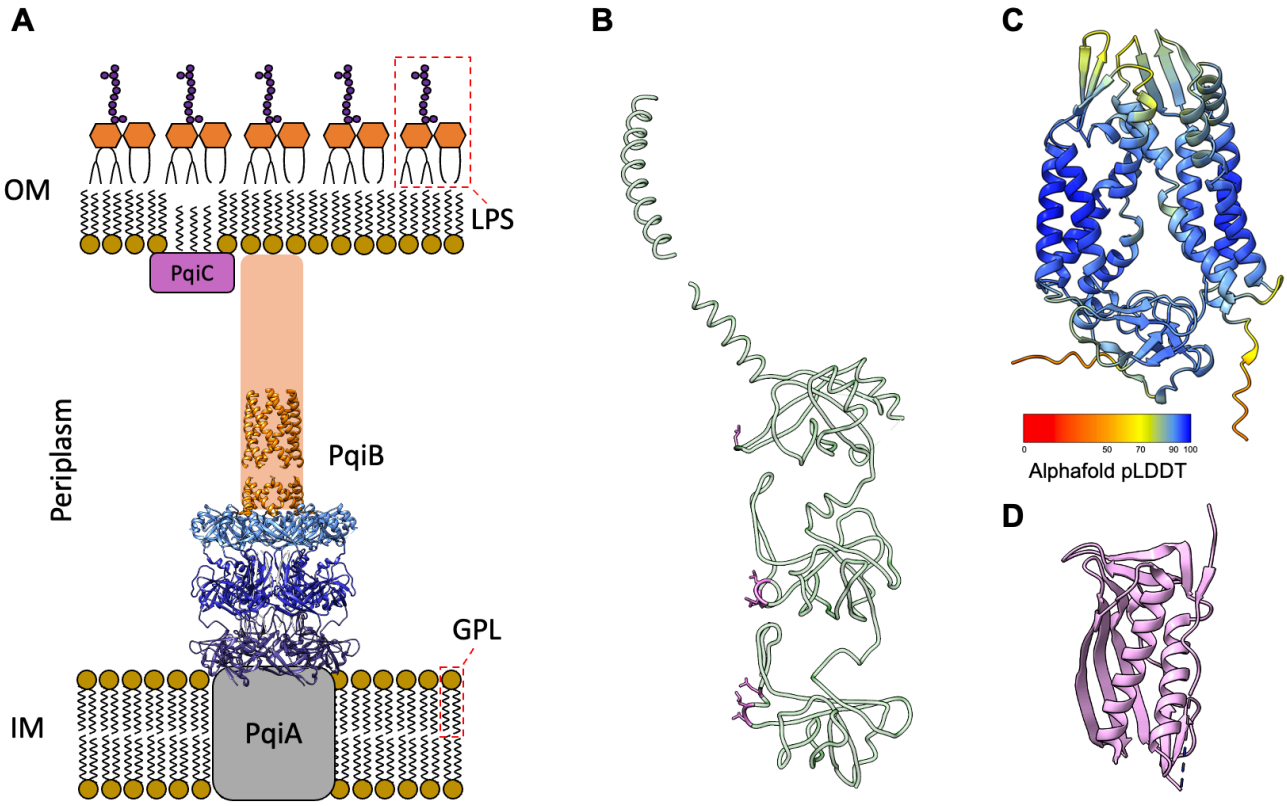


Fig. 18 – Structures of PqiA, PqiB and PqiC.

(A) Depiction of PqiB in its proposed cellular location, spanning the periplasm, with MCE domains are coloured in an array of blues and the helical coiled coil depicted in orange. The anticipated length of the entire coiled coil is indicated by the orange shaded area. PqiA is indicated by the grey box within the inner membrane (PDB: 5UVN). (B) Single protomer of PqiB highlighting the vertical organisation of the seven MCE domains. The residues constituting the hydrophobic motifs found in all *E. coli* MCE domains are coloured in purple. PqiB (PDB: 5UVN). (C) Cartoon representation of the Alphafold predicted structure for PqiA coloured according to Alphafold's per residue confidence measure pLDDT (Jumper et al., 2021). (D) Cartoon representation of the *E. cloaclae* PqiC. (PDB: 6OSX).

Resolving function from structure

Despite these pioneering structural insights, the physiological purpose of either pathway remains cryptic. The *pqi* operon is upregulated in response to paraquat induced oxidative stress, in a soxRS dependent manner, and by starvation via RpoS (Koh and Roe, 1996, 1995). Induction of the *yeb* operon is less clear, however, putative evidence demonstrated induction of YebS under conditions harbouring compromised LPS (Nakayama and Zhang-Akiyama, 2017). Furthermore, soluble constructs of Yeb and PqiB, like MlaD, have been shown to co-purify with GPL species bound, raising speculation of their involvement in lipid trafficking (Ekiert et al., 2017).

Growth experiments revealed 1% LSB was sufficient to inhibit the growth of Δ*pqiAB* and Δ*mlaD* single mutants. Whilst a Δ*yebST* single mutant was not itself sensitive to LSB, a double Δ*pqiAB* Δy*ebST* mutant displayed a more severe phenotype than Δ*pqiAB* alone, and triple Δ*pqiAB* Δ*yebST* Δ*mlaD* mutant failed to grow indicating complete LSB resistance necessitates all three *E. coli* MCE proteins. Critically, the triple mutant exhibited typical growth under normal conditions (LB agar) suggesting the functions of MCE protein containing pathways may be performed, concurrently or as required, by alternative pathways (Isom et al., 2017).

That the Pqi, Yeb and Mla pathways are functionally divergent, and not simply examples of evolutionary redundancy appears to be the current consensus, however, under certain conditions their operations may overlap. Retention of all three operons in genomes that have undergone extensive reductive evolution, such as *Salmonella enterica* serovar Typhi, reinforces this notion (McClelland et al., 2004). The potential for additional components of either pathway to be identified remains, particularly in the case of YebST where the lack of an OM associated element presents an obvious omission. Indeed MlaA, the OM component of the Mla pathway, is expressed from a locus discrete from the *mlaFEDBC* operon in *E. coli*

(Malinverni and Silhavy, 2009). Furthermore, YadH, an IM ABC transporter, was identified in complex with Mla components during a proteomic survey, raising the possibility of more accessory proteins (Babu et al., 2018). If, therefore, the Pqi and Yeb pathways do indeed constitute periplasmic transport pathways, the direction of travel and energy source, if any, driving inter-bilayer movement remain principal uncertainties.

Research Aims

That the *pqi* operon functions as a periplasmic lipid transport system remains a strongly held consensus, however, there is little evidence to support this notion and the Pqi complex itself remains relatively unstudied.

Of the three *pqi* encoded proteins, the only *E. coli* structure resolved is that of a soluble $PqiB^{39\text{-}546}$ construct hence those of PqiA and PqiC remain undetermined. The deposited $PqiB^{39\text{-}546}$ structure represents the biologically anticipated homo-hexameric assembly, however, the oligomeric states of the remaining components, and how these components convene in the periplasmic spanning complex, is of primary interest both structurally and functionally.

There is strikingly little information regarding the cellular function or mechanism of the Pqi pathway. Complementation studies have identified a series of hydrophobic residues within the MlaD and YebT pore lining loops which appear critical for function. Analysis of the PqiB sequence reveals the presence of analogous hydrophobic residues within each of its three MCE domains, thus these are also anticipated to be integral to function. Whilst intriguing, these residues in isolation do little to reveal the mechanism by which PqiB functions let alone that which Pqi mediated transport may occur considering the presence of at least two further complex components.

Considering the significant lack of current information, the scope of this study was broad with the intention to probe the Pqi pathway both structurally and functionally via an array of structural, biophysical, and genetic methodologies. The aims of this study are summarised as follows:

1 – Solve the structure of the outer-membrane associated lipoprotein PqiC.

2 – Determine the architecture of the entire PqiABC complex revealing all interacting partners and their stoichiometries.

3 – Utilise complementation assays to infer critical residues within all three known components of the Pqi pathway.

Chapter 2: Materials and Methods

Construct design, Expression and Purification

Culture Media and Antibiotics

E. coli cell lines were grown in Miller's lysogeny broth (referred to as LB hereafter) comprising 10 g/L Tryptone, 5 g/L Yeast Extract and 10 g/L NaCl. LB was supplemented with the following antibiotics: kanamycin (30 µg/ml), ampicillin (100 µg/ml), carbenicillin (100 µg/ml) or chloramphenicol (35 µg/ml in ethanol) depending on the selection criteria required. For agar plates, LB was combined with 15 g/L Agar prior to sterilisation.

Growth of the PqiABC construct involved further supplementation of LB with: 325 µM Na_2EDTA, 200 µM $ZnSO_4$, 450 µM H_3BO_4, 75 µM $MnCl_2$, 17.5 µM $CoCl_2$, 15 µM $CuSO_4$, 20 µM $(NH_4)_6Mo_7O_{24}$.

Bacterial Strains

The strains utilised within this study are detailed below (Table. 1).

Table. 1 – Strains utilised within this study.

Strain	Usage	Source
E. coli K-12 BW25113	Knockout generation and complementation studies	(Baba et al., 2006)
E. coli K-12 BW25113 *ΔpqiABC::aph*	Complementation studies	This Study
E. coli K-12 BW25113 *ΔpqiABC*	Complementation studies	This Study
E. coli C43 DE3	Recombinant protein expression	(Miroux and Walker, 1996)
E. coli DH5α	Plasmid Replication	New England Biolabs
E. coli NEB5α	Plasmid Replication (SDM)	New England Biolabs

Constructs

The key properties for the constructs utilised within this study are detailed below (Table. 2).

Table. 2 – Key construct properties.

Construct Name	Plasmid	Restriction Sites	His Tag Location	Codon Optimisation	Source
PqiC FL	pET 22b	Nde1 – Xho1	C – Terminal	Y	GenScript
PqiC $^{17\text{-}187}$	pET 26b	Nde1 – Xho1	N - Terminal	Y	GenScript
PqiABC	pET 26b	Nde1 – Xho1	C – Terminal	N	GenScript
PqiAB	pET 26b	Nde1 - Xho1	C - terminal	N	GenScript

Inserts for all the constructs above in Table.2 are provided in Appendix. 1.

Transformation

Competent cells were produced using protocol 1 'Preparation of competent *E. coli*' from the Qiagen Bench Guide (www.gene-quantification.de/Qiagen-BenchGuide.pdf).

Transformation required incubation of 1-5 µl plasmid DNA (for 1-10 ng DNA) with 50 µl competent cells (OD_{600} = 0.5) for 15 minutes on ice. Following incubation, cells were heat shocked at 42 ˚C for 30 seconds and placed back on ice for 10 minutes. Cells were then supplemented with 100 µl SOC outgrowth media (2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose) and incubated at 37 ˚C for 1 hour before plating on an appropriate antibiotic selection plate.

Expression

Single colonies from a fresh transformation were used to inoculate LB starter cultures supplemented with an appropriate antibiotic. Starter cultures were grown overnight at 37 ˚C with continual shaking at 180 rpm.

Expression was performed in larger volumes of antibiotic supplemented LB. Each 1L of expression media was inoculated with 10 ml of starter culture and grown at 37 ˚C until OD_{600} reached 0.4 at which point the temperature was reduced to 18 ˚C. Upon obtaining an OD_{600} of 0.6, expression was induced via the addition of 1 mM IPTG and proceeded overnight at 18 ˚C with continual shaking (180 rpm). Cells were harvested via centrifugation at 6000 RCF for 15 minutes (Beckman Coulter JLA-8.1000).

Purification – Soluble Constructs

Cell lysis

Cell pellets were resuspended in an appropriate volume of 20 mM Tris, 500 mM NaCl; pH 8.5 supplemented with 1 cOmplete™ EDTA-free protease inhibitor cocktail tablet (Roche) per 50 ml of resuspension buffer. Cells were homogenised using a glass Dounce homogeniser prior to lysis via four passes of an Emulsiflex-C3 cell disruptor (Avestin) at an output pressure of 15,000 psi. Cell debris was removed via centrifugation at 50,000 RCF for 1 hour (Beckman Coulter JA-25.50).

IMAC Purification

The resulting supernatant was filtered through a 1.2 µm syringe filter (Sartorius) before clarified lysate was bound overnight to a 5 ml HisTrap HP column (Cytiva) pre-equilibrated in 20 mM Tris, 500 mM NaCl, 50 mM Imidazole; pH 8.5. The column was then washed with 5 column volumes of 20 mM Tris, 500 mM NaCl, 50 mM Imidazole; pH 8.5 before elution in 4 column volumes of 20 mM Tris, 500 mM NaCl, 500 mM Imidazole; pH 8.5 collected in 1.5 ml fractions. The flowthrough, wash and elution fractions were analysed via SDS-PAGE to determine the purification success and identify elution fractions containing the protein of interest.

Size Exclusion Chromatography (SEC)

Elution fractions containing the protein of interest were pooled and concentrated in a 10 kDa MWCO Amicon® Ultra centrifugal filter unit (Merck Millipore) to a suitable volume for size exclusion chromatography (<1 ml, <2 ml and <5 ml for 10/300, 16/600 and 26/60 columns respectively). SEC was performed using Superdex® 200 columns of various sizes, dependent upon the amount of protein loaded, attached to an ÄKTA purifier system (GE Healthcare). Columns were pre-equilibrated with 1.2 column volumes of 20 mM Tris, 150 mM NaCl; pH 8.5 prior to sample injection after which elution, in the same buffer, continued for a further column volume with fractions being collected. The protein contents of each fraction were indicated by continually monitoring the absorbance at 280 nm (A_{280}) during elution. Finally, the protein contents of each fraction were analysed via SDS-PAGE.

Purification – Membrane Constructs

Resuspension and Cell lysis

Cell pellets were resuspended in an appropriate volume of 20 mM Bis-Tris Propane, 500 mM NaCl, 0.5 mM TCEP; pH 8.5 supplemented with 1 cOmplete™ EDTA-free protease inhibitor cocktail tablet (Roche) per 50 ml of resuspension buffer. Cells were homogenised using a glass Dounce homogeniser prior to lysis via four passes of an Emulsiflex-C3 cell disruptor (Avestin) at an output pressure of 15,000 psi. Cell debris was removed via centrifugation at 10,000 RCF for 10 minutes (Beckman Coulter JA-25.50).

Membrane Harvest and Solubilisation

The supernatant was decanted and centrifuged at 165,000 RCF for 1 hour to pellet the membrane fraction (Beckman Coulter Type 70 Ti). The supernatant was decanted, membrane pellets transferred into conical tubes and the wet membrane pellet mass recorded before being frozen at -80 ˚C.

Membrane pellets were resuspended in 20 mM Bis-Tris Propane, 500 mM NaCl, 0.5 mM TCEP, 1% w/v n-dodecyl-β-D-maltoside; pH 8.5 at a ratio of 1 ml per 40 mg of wet membrane mass, homogenised using a Dounce homogeniser, and left rotating at 4 ˚C for 3 hours. Insoluble material was removed via centrifugation at 75,000 RCF for 30 minutes (Beckman Coulter JA-25.50).

IMAC Purification

The resulting supernatant was filtered through a 0.45 µm syringe filter (Sartorius) before clarified lysate was incubated overnight at 4 ˚C with 1 ml of Ni-NTA (nickel-nitrilotriacetic acid) Agarose (Qiagen) pre-equilibrated in 20 mM Bis-Tris Propane, 500 mM NaCl, 50 mM Imidazole, 0.5 mM TCEP, 0.03% w/v n-dodecyl-β-D-maltoside; pH 8.5.

The incubated supernatant was transferred to a gravity column and washed with 5 column volumes of 20 mM Bis-Tris Propane, 500 mM NaCl, 50 mM Imidazole, 0.5 mM TCEP, 0.03% w/v n-dodecyl-β-D-maltoside; pH 8.5 before eluting 12 x 1.5 ml fractions in 20 mM Bis-Tris Propane, 500 mM NaCl, 500 mM Imidazole, 0,03% w/v n-dodecyl-β-D-maltoside; pH 8.5 collected in 1.5 ml fractions. The flowthrough, wash and elution fractions were analysed via SDS-PAGE to determine the purification success and identify elution fractions containing the protein of interest.

Size Exclusion Chromatography (SEC)

Elution fractions containing the protein of interest were pooled and concentrated in a 100,000 MWCO Amicon® Ultra centrifugal filter unit (Merck Millipore) to a suitable volume for size exclusion chromatography (<1 ml or <2 ml for 10/300 or 16/600 columns respectively). SEC was performed using Superdex® 200 (PqiC FL, PqiC $^{17\text{-}187}$) or Superose® 6 (PqiABC) columns of various sizes, dependent upon the amount of protein loaded, attached to an ÄKTA purifier system (GE Healthcare). Columns were pre-equilibrated with 1.2 column volumes of 20 mM Bis-Tris Propane, 150 mM NaCl, 0.5 mM TCEP, 0.03% w/v n-dodecyl-β-D-maltoside; pH 8.5 prior to sample injection after which elution, in the same buffer, continued for a further column volume with fractions being collected. The protein contents of

each fraction were indicated by continually monitoring the absorbance at 280 nm (A_{280}) during elution. Finally, the protein contents of each fraction were analysed via SDS-PAGE.

NB. For crystallography of PqiC FL, the purification protocol above was performed with the substitution of Tris in the place of Bis-Tris Propane and the omission of TCEP.

SDS-PAGE

Protein samples were prepared by combining equal volumes of protein sample and 2x Laemmli sample buffer (Sigma-Aldrich). Samples were then loaded onto SurePAGE™, Bis-Tris, 4-12% polyacrylamide gels (GenScript) alongside a Precision Plus Protein™ all blue prestained protein standard (BioRad). Gels were run in MES SDS running buffer (GenScript) at a constant 165 V until complete before staining with Quick Coomassie Stain (Neo Biotech). Finally, gels were imaged using a GS-800 densitometer (BioRad).

Western Blotting

Samples for Western blotting were first separated via SDS-PAGE before being transferred to a nitrocellulose membrane via a Trans-Blot Turbo transfer system (BioRad) as per manufacturer's instructions. Membranes were initially blocked in TBST supplemented with 5 % (w/v) skimmed milk before overnight incubation at 4 ˚C with a 6xHis monoclonal antibody (Takara Bio) diluted 1:5,000 with TBST supplemented with 5 % (w/v) skimmed milk. Membranes were washed five times with TBST before incubation with anti-mouse HRP conjugated secondary (ThermoFisher Scientific) diluted 1:10,000 in TBST for 1 hour. Membranes were washed a further five times in TBST before incubation with ECL (Cytiva) and visualisation via an Amersham 600 series imager (GE Healthcare).

Q5 Site Directed Mutagenesis (SDM)

The NEB Q5 Site Directed Mutagenesis Kit (New England Biolabs) was used to perform all mutagenesis reactions. Primers were designed using the online NEBaseChanger primer design tool (https://nebasechanger.neb.com), all of which are documented below in Table 6.

Mutagenesis reactions were performed in accordance with the NEB Q5 Site Directed Mutagenesis Kit protocol (E0554).

The SDM PCR reaction contents and cycling conditions are detailed below (Tables 3 and 4 respectively).

Table. 3 – Q5 SDM reaction contents.

Constituent	Volume (µl)
Q5 Hot Start High Fidelity 2x Master Mix	12.5
Forward Primer (10 µM)	1.25
Reverse Primer (10 µM)	1.25
Template DNA (30 ng/µl)	1
Sterile Distilled Water	9

Table. 4 – Q5 SDM cycling conditions.

Step	Temperature (˚C)	Time (Seconds)
Initial Denaturation	98	30
25 cycles	98	10
	Ta* (primer dependent)	30
	72	180
Final Extension	72	120
Hold	4	∞

Following the Q5 SDM reaction, a Kinase, Ligase and Dpn1 (KLD) treatment was performed on the PCR product prior to transformation. The KLD reaction mix is detailed below (Table. 5).

Table. 5 – Q5 Kinase, Ligase, Dpn1 reaction constituents.

Constituent	Volume (µl)
PCR Product	1
2x KLD Reaction Buffer	5
10x KLD Enzyme Mix	1
Sterile Distilled Water	3

The reaction mix was left at room temperature for 30 minutes after which 5 µl of the reaction was transformed into NEB5α competent cells as per the manufacturers guidelines and plated into an appropriate antibiotic LB agar selection plate.

Q5 SDM Primers

Primers used for site directed mutagenesis of the PqiABC construct, are detailed below (Table. 6).

Table. 6 – PqiABC Q5 SDM primers.

Mutation	Fwd Primer (5'-3')	Rev Primer (5'-3')
PqiA G146A	CAAAACCTGGgcgATGGCGGAGA	AGTTGAAAAAGCACCCGTG
PqiA E149A	GGGAATGGCGgcgATTTTCCTCG	CAGGTTTTGAGTTGAAAAAGCAC
PqiA G352A	TGAGTTTGTAgcgCGCTGGTCGATG	ACAACTTCATAAATCAAATGCATTC
PqiA D358A	GTCGATGATTgcgGTTTTCGTTATCGCG	CAGCGGCCTACAAACTCAs
PqiB L119N	GATTAGCGGCaacGGAACGCTGCTG	CCTTCGCGACCAATCTGC
PqiB L122N	CCTGGGAACGaacCTGTCTGGAGTTTATATC	CCGCTAATCCCTTCGCGA
PqiB L123N	GGGAACGCTGaacTCTGG AGTTTATATCGAACTGC	AGGCCGCTAATCCCTTCG
PqiB L240N	GATGGGCTCAaacACAACGCTGCTGAG	TCCACACGCATCCCTGCT
PqiB L243N	ATTGACAACGaacCTGAGTGGCGGTGTC	GAGCCCATCTCCACACGC
PqiB L244N	GACAACGCTGaacAGTGGCGGTG	AATGAGCCCATCTCCACAC
PqiB	AACCGGAAACaacGTCACTGGTGCAC	TTCAGCGATCCGCGTAAA

L377N		
PqiB V377N	CGGAAACCTGaacACTGGTGCAC	GTTTTCAGCGATCCGCGT
PqiC L23A	CGAAATTAATgcgAACTATTAC CAGTTACCTGTGGTGCAGAGCG	CCGGAGCTGCATCCGGCC
PqiC W47A	TCGTCTGTTAgcgGTAG AGCAGGTCACTGTTCCTG	TTGCCCTGGCTGGCGGTA
PqiC D55A	CACTGTTCCTgcgTATCTGGCGG	ACCTGCTCTACCCATAAC
PqiC D84A	CAGCCCGTTGgcgCAACAGTTGC	GCCCACAAGTTGTTGTTG

Sequencing

Sequencing was performed by Source BioScience (www.sourcebioscience.com) utilising their Sanger Sequencing service. Typically, 5 µl of plasmid DNA (50 – 100 ng/µl) was submitted for sequencing with Source BioScience's own T7F or T7R primers dependent upon mutation position. Where necessary custom sequencing primers, detailed below (Table. 7), were submitted concurrently with plasmid DNA.

Table. 7 – Pqi sequencing primers.

Sequencing Primer Name	Sequence (5'-3')
PqiA Seq	ATGACGTTTGACCCGCGTTTG
PqiA Seq 2	ATGCCAAAGGGCATGGCAAG
PqiB Seq Fwd	TTGACGCACGTTGAAATCAAAGC

Quantification of protein and DNA concentrations

Protein and plasmid DNA concentrations were both quantified using a DeNovix DS-11 FX spectrophotometer (DeNovix) utilising the 'dsDNA' and 'Protein A280' programs respectively.

The theoretical molecular mass and extinction coefficient for each protein/complex was determined using the Expasy ProtParam online tool (https://web.expasy.org/protparam) and are detailed below (Table. 8).

Table. 8 – Theoretical protein/complex molecular mass and extinction coefficients.

Protein / Complex	**Theoretical Molecular Mass (Da)**	**Theoretical Extinction Coefficient (M^{-1} cm^{-1})**
PqiC FL	21699.53	48930
PqiC $^{17-187}$	20728.23	39420
PqiABC (Assumed 2:6:8 ratio)	620854.80	837755

Molecular Biology Techniques

Plasmids

The plasmids used within this study are detailed below (Table. 9).

Table. 9 – Plasmids used within this study.

Plasmid	Properties	Source
pKD4	Kanamycin resistance cassette flanked by FLP recombinase target sites.	(Datsenko and Wanner, 2000)
pACBSCE	Self-cleaving plasmid with arabinose inducible expression of the λ Red recombination system.	(Lee et al., 2009)
pCP20	Temperature sensitive plasmid expressing FLP recombinase.	(Datsenko and Wanner, 2000)

Molecular Biology Primers

The primers used for the molecular biology within this study are detailed below (Table. 10).

Table. 10 – Molecular biology primers used within this study.

Primer	Sequence (5'-3')	Description
pqiA_ko_Fwd	AGTCATCTGTAAAATAGCGCATCATTAAGGAGT ACCAATGGTGTAGGCTGGAGCTGCTTC	Forward primer for the amplification of the pKD4 kanamycin cassette for the *pqiABC* knockout.
pqiC_ko_Rev	TATTTACAAATCTTTATAATTAAGGTAGACGCTT TATCTCCATATGAATATCCTCCTTAG	Reverse primer for the amplification of the pKD4 kanamycin cassette for *the pqiABC* knockout.
delABC_comf_Fwd	GAAGCCCTACAGACGCAAGT	Forward primer for colony PCR to confirm disruption off the *pqiABC* genes.
delABC_comf_Rev	CCATTACTACCCTGTCCGCC	Reverse primer for colony PCR to confirm disruption off the *pqiABC* genes.

Polymerase Chain Reaction

PCR of linear fragments for gene disruption were performed using the Phusion® High Fidelity DNA polymerase (NEB). The constituents of each 50 µl reaction and Phusion cycling conditions are detailed below (Tables 11 and 12 respectively).

Table. 11 – Phusion® PCR reaction contents.

Constituent	Volume (µl)
5x Phusion HF Buffer	10
Forward Primer (10 µM)	2.5
Reverse Primer (10 µM)	2.5
Template DNA (25-50 ng/µl)	1
dNTP Mix (10mM)	1
Phusion Polymerase	0.5
Sterile Distilled Water	32.5

Table. 12 – Phusion® PCR cycling conditions.

Step	Temperature (˚C)	Time (Seconds)
Initial Denaturation	98	30
25 cycles	98	10
	60	30
	72	60
Final Extension	72	300
Hold	4	∞

Colony PCR was performed using MyTaqRed Mix (Meridian Bioscience). Colonies were picked using a pipette tip, spotted onto a fresh agar plate and then transferred into a reaction mix consisting of 12.5 µl MyTaqRed Mix, 1 µl forward check primer, 1 µl reverse check primer and 10.5 µl Sterile Distilled Water. Colony PCR cycling conditions are detailed below (Table. 13).

Table. 13 – Colony PCR cycling conditions.

Step	Temperature (˚C)	Time (Seconds)
Initial Denaturation	98	600
25 cycles	98	10
	60	30
	72	90
Final Extension	72	300
Hold	4	∞

Agarose Gel Electrophoresis

DNA fragments were resolved using 1% w/v agarose gels (Bioline) cast and run in Tris-Acetate-EDTA (TAE) buffer with SYBR™ safe DNA gel stain (Invitrogen) as per the manufacturer's guidelines. 1x TAE buffer comprises: 40 mM Tris, 20 mM Acetate and 1 mM EDTA; pH ~8.6. Samples were combined with 6x purple loading dye (NEB), at a sample:buffer ratio of 1:5, prior to loading. A 1 kb DNA ladder (NEB) was run to reference fragment sizes. Gels were run at a constant 100 V until complete and visualised under UV light.

Electroporation

Overnight cultures were diluted 100-fold into 100 ml of fresh LB, supplemented with antibiotic where appropriate, and grown until an OD_{600} of 0.6. Cells were transferred to ice for 10 minutes before centrifugation at 3,220 RCF for 10 minutes at 4 ˚C. Cells were resuspended in 25 ml ice cold 10% v/v glycerol. The centrifugation and resuspension steps were performed a further two times with resuspension volumes of 25 ml and 1 ml respectively.

100 µl aliquots of cells were incubated with 50-300 ng of linear DNA fragments for 30 minutes on ice after which the solution was transferred to a 2 mm electroporation cuvette (Scientific Laboratory Supplies) and pulsed once at 1750 V. 1 ml of SOC media (2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose) (NEB) was immediately added and the mixture transferred to a 15 ml conical tube and grown at 37 ˚C for 1 hour before plating on an appropriate antibiotic selection plate.

Datsenko and Wanner Gene Inactivation

The Datsenko and Wanner methods of gene inactivation involved replacement of a specified gene/operon with a kanamycin cassette followed by its subsequent removal to leave a small scar in place of the disrupted genes (Datsenko and Wanner, 2000). This methodology was deployed to create a knockout of the *pqiABC* genes. The Datsenko and Wanner method is outlined below (Fig. 19).

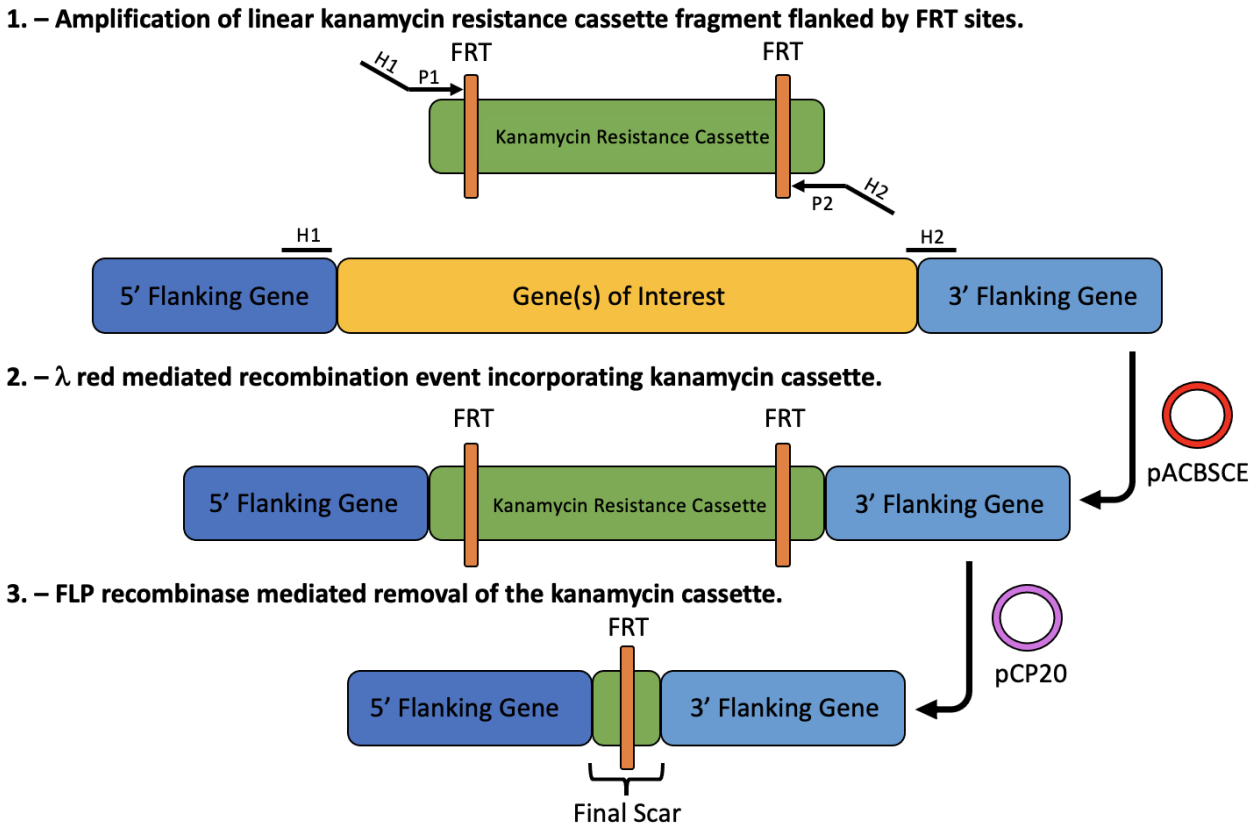


Fig. 19 – Datsenko and Wanner gene inactivation.

1 - Amplification of a linear kanamycin (*aph*) cassette fragment utilising primers with sequences homologous to the 5' and 3' ends of the *aph* cassette (P1 and P2) and the 5' and 3' flanking genes (H1 and H2). 2 - Cells are electro-transformed with the linear fragment which is incorporated into the chromosome in place of the target gene(s) by the λ red recombination machinery (expressed from the PACBSCE plasmid). 3 - FLP recombinase (expressed from the pCP20 plasmid) utilises the FLP recombinase target (FRT) sites located within the incorporated fragment to remove the kanamycin cassette leaving a small scar on the chromosome. Figure adapted from (Datsenko and Wanner, 2000; Isom, 2017).

First, the kanamycin resistance (*aph*) cassette from the pKD4 plasmid was amplified using specific primers to yield a linear fragment comprising the *aph* cassette, flanked immediately

by FLP recombinase target (FRT) sites, and distally by sequences homologous to the regions flanking the genes of interest. The forward primer consisted of 40 bp homologous to the sequence upstream of the genes of interest (H1) followed by 20 bp homologous to the 5' sequence of the *aph* cassette (P1). The reverse primer was of an analogous composition, comprising 40 bp homologous to the end of the gene of interest and its downstream region (H2) followed by 20 bp homologous to the end of the *aph* cassette (P2). The linear PCR fragment was subsequently digested with Dpn1 (NEB) and gel purified using a QIAquick gel extraction kit (Qiagen).

E. coli K-12 BW25113 was made chemically competent, transformed with the pACBSCE plasmid and grown in LB supplemented with chloramphenicol at 37 ˚C until OD_{600} of 0.6. The media was then supplemented with 1% w/v arabinose, to induce expression of the λ red recombination machinery, and transferred to ice for 10 minutes. Cells were electro-transformed with the linear PCR fragment and grown at 37 ˚C for 1 hour before being pelleted (3220 RCF, 10 minutes) and resuspended in 200 µl SOC media (2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose) (NEB).

Half of the mixture was plated on LB agar plates supplemented with kanamycin and incubated overnight at 37 ˚C. The remaining half was left overnight at room temperature before also being plated and incubated in the same fashion. Resultant colonies were checked for corrected insertion of the *aph* cassette via colony PCR utilising primers which bound regions ~300 bp up- and down-stream of the genes of interest (Table 10). A change in size of the PCR product from that obtained from the WT *E. coli* K-12 BW25113 to that of the *aph* cassette indicated successful replacement of the target genes.

Loss of the pACBSCE plasmid was confirmed by replica plating of knockout colonies on chloramphenicol, kanamycin and antibiotic free LB agar plates. Colonies still containing pACBSCE were grown overnight in small LB cultures supplemented with 1% w/v arabinose, streaked onto LB Agar plates and assessed again via replica plating.

P1 Transduction

Preparation of P1 Lysate

P1 transduction was used to transfer freshly generated knockout mutations into fresh recipient strains to mitigate against the possibility of unwanted secondary site mutations introduced during the gene inactivation step. Transduction was performed in accordance with that described by (Thomason et al., 2007).

An overnight culture of the confirmed knockout strain was diluted 1 in 100 into 5 ml of fresh. LB supplemented with 0.2% w/v Glucose and 5 mM $CaCl_2$ and grown at 37 ˚C for 45 minutes. The culture was infected with 100 µl of virulent P1 stock and incubation continued until the culture lysed (~3 hours). Total lysis was ensured by the addition of 200 µl of chloroform followed by a further 5 minute incubation. The mixture was then centrifuged at 10,000 RCF for 10 minutes at 4 ˚C and the supernatant (P1 lysate) stored at 4 ˚C until required.

Transduction

1.5 ml of an overnight culture of the recipient strain grown in LB was pelleted by centrifugation and the pellets resuspended in 750 µl of P1 salts (10 mM $CaCl_2$, 5 mM

$MgSO_4$). 100 µl of resuspended cells were incubated with 50 µl of P1 lysate for 30 minutes at room temperature before addition of 1 ml LB and 200 µl 1M sodium citrate and a further incubation at 37 ˚C of 1 hour. Cells were pelleted in microcentrifuge tubes for 2 minutes at 17,900 RCF and resuspended in 100 µl of LB and plated on LB agar plates supplemented with 30 µg/ml kanamycin and 5 mM sodium citrate and grown overnight at 37 ˚C. The recipient strain and P1 lysate were also plated to validate the selection process and check for contamination. Single colonies from the successful transduction plate were streaked out twice more on LB agar plates supplemented with 30 µg/ml kanamycin and 5 mM sodium citrate to remove non-transduced bacteria and residual phage. Remaining colonies were screened for the knockout of interest via colony PCR.

Removal of the *aph* cassette

Following successful replacement of the target genes, the *aph* cassette may be removed. The desired knockout strain was made chemically competent and transformed with the temperature sensitive plasmid pCP20 and plated on LB agar plates supplemented with carbenicillin and incubated at 30 ˚C for up to 24 hours. Colonies were streaked onto a fresh LB agar plate and incubated overnight at 43 ˚C. Loss of pCP20 was assessed via replica plating on kanamycin, carbenicillin and antibiotic free LB agar plates. Following pCP20 removal, colony PCR (utilising the same ‘check’ primers) was used to determine if the *aph* cassette had been successfully removed leaving only a small scar on the chromosome.

Phenotypic Complementation Assays

Phenotypic complementation assays were performed in accordance with those previously reported for *E. coli* MCE domain proteins (Ekiert et al., 2017; Isom et al., 2020; Vieni et al., 2022).

All strains were grown overnight at 37 °C in LB, supplemented with 30 µg/ml kanamycin where appropriate. The resulting cultures were normalised to OD_{600} = 1 and serial diluted (10^0 – 10^{-7}) in fresh LB in a 96 well plate. 2.5 µl of each dilution was pipetted onto rectangular petri dishes of LB-Millar agar supplemented with 0.25 % lauryl sulfobetaine. Petri dishes were incubated overnight at 37 °C before imaging under white light with a U:GENIUS^3 system (Syngene).

Structural Determination

X-ray Crystallography

Protein Crystallisation

Concentrated purified proteins (5 – 30 mg/ml) were used to set up sitting drop, vapour diffusion crystal trays, maintained at 18 ˚C. Trays were pipetted either by hand in 96-well CombiClover™ 4 chamber crystallisation plates (Molecular Dimensions) or using a Mosquito HTS liquid handler (TTP Labtech) into MRC 3 lens crystallisation plates (SWISSCI). Hand trays drops contained 1 µl protein and 1 µl condition per drop whilst mosquito trays contained 400 nl protein and 400 nl condition per drop. Crystallisation conditions were taken from the JCSG Plus (Mcpherson, 2001), BCS (Chaikuad et al., 2015), PACT Premier (Newman et al., 2005), MIDAS Plus (Grimm et al., 2010), Morpheus (Gorrec, 2009), MemPlus (Newstead et al., 2008b), MemGold (Newstead et al., 2008a) and MemGold 2 (Parker and Newstead, 2012) commercial crystallisation screens (Molecular Dimensions). Formed crystals were picked using an appropriately circular litholoop™ (Molecular Dimensions) and snap frozen in liquid nitrogen for transport. Where necessary, crystals were bathed in a cryoprotecting solution of the mother liquor supplemented with 25% v/v PEG400 prior to freezing.

Data Collection

Diffraction data was collected on an Eiger2 XE 16M detector at the I04 beamline (Diamond Light Source, Oxford, UK). The collection parameters for each of the three datasets can be found in Appendix. 2. Collected datasets were automatically reduced and processed by the ISPyB integrated pipelines (xia2 dials, fast_dp, xia2 3dii, autoPROC and autoPROC + STARANISO). Datasets for structure determination were chosen based on the comparison of

the following statistics; resolution, completeness, multiplicity, I/σI, Rmerge, Rpim, Rmeas, $CC_{(1/2)}$.

Structural Determination and Model Refinement

All structures utilised molecular replacement (MR) to solve the crystallography phase problem. Molecular replacement was performed using the phaser program (McCoy et al., 2007) from within the *PHENIX* software suite. Solutions possessing a translation function Z score (TFZ) exceeding or equal to 8 and a high positive log likelihood gain (LLG) were considered plausible correct solutions and opened in Coot (Emsley and Cowtan, 2004) for manual inspection. Initial molecular replacement models were refined iteratively via repeated cycles of refinement using *phenix.real_space_refine* (Afonine et al., 2012) and REFMAC5 (Murshudov et al., 2011) interspersed with manual model building in Coot (Emsley and Cowtan, 2004) and ISOLDE (Croll, 2018). Model validity was assessed via the MolProbity program distributed with *PHENIX* (Chen et al., 2010).

Cryo-EM

Sample Vitrification

Grids were prepared in a semi-automated fashion using a Vitrobot mark IV (ThermoFisher Scientific) maintained at 4 ˚C, 95% humidity with various blot forces and times. 3 µl of protein sample was applied to a glow discharged (35 mA, 60 Sec) QUANTIFOIL® R1.2/1.3 Cu grid (Quantifoil Micro Tools), blotted and plunged into liquid ethane to form a thin layer of vitreous ice. Frozen grids were maintained in liquid nitrogen hereafter to maintain ice integrity and prevent contamination.

Screening and Data Collection – ThermoFisher Scientific

Screening was performed using a Titan Krios G4 with Selectris imaging filter and Falcon IV detector (ThermoFisher Scientific) at a nominal magnification of 130,000 x (0.96 Å per pixel). The quality of each grid was assessed by inspecting a range of grid areas and considering ice thickness, contamination and particle distribution and orientation. Data was collected from promising grids using the same microscope and detector configuration.

Data Analysis

The potential of the PqiABC dataset was initially assessed via Dr Deniz Ugurlar of the ThermoFisher Centre for Electron Microscopy, who utilised the cryoSPARC software (version 3.2) (cryosparc.com) however the data was subsequently re-processed from raw micrographs to final refined maps as below.

Data was analysed using the RELION software (version 3.1.3) (Scheres, 2012a, 2012b), a brief overview of the major steps are outlined below. Raw micrographs were first motion corrected, using RELION's own implementation of MotionCor2 (Zheng et al., 2017), before CTF estimation using CTFFIND 4.1 (Rohou and Grigorieff, 2015). Manual inspection of the Thon-rings for each micrograph allowed removal of those where the CTF had been estimated poorly. Small scale particle picking was initially performed manually, allowing inspection of a subset of the dataset, however particle picking of the whole dataset was performed using the neural network-based 'Topaz' particle picking framework (Bepler et al., 2019). Picked particles were subjected to rounds of 2D classification allowing further data inspection and removal of unwanted particles prior to initial model generation. Initial models were iteratively improved via successive rounds of 3D classification, with models assessed in UCSF ChimeraX (Pettersen et al., 2021). Finally, individual 3D classes were refined and post-processed to yield the highest possible resolution maps.

Model Building

Maps were first sharpened, via *phenix.auto_sharpen*, to improve clarity and aid the modelling process (Terwilliger et al., 2018). An existing PDB model, segmented and rigid body fit into map density, served as an initial base. Subsequently, segmented regions were re-joined and absent portions built *de-novo* in Coot (Emsley and Cowtan, 2004) before the entire structure was refined utilising *phenix.real_space_refine* (Afonine et al., 2012).

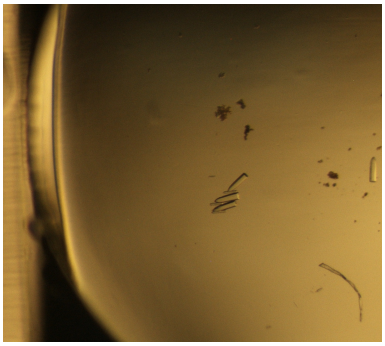
Chapter 3: X-ray Crystallography of PqiC

Introduction

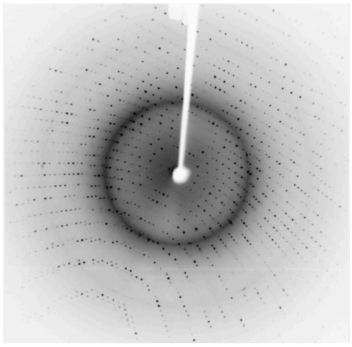
Introduction to X-ray Crystallography

X-ray Crystallography presents a fundamental method of structural determination in which the physical properties of crystals are exploited to determine a 3D molecular structure. X-ray crystallography of proteinaceous crystals remains the most frequently utilised method for the structural elucidation of proteins and macromolecular complexes.

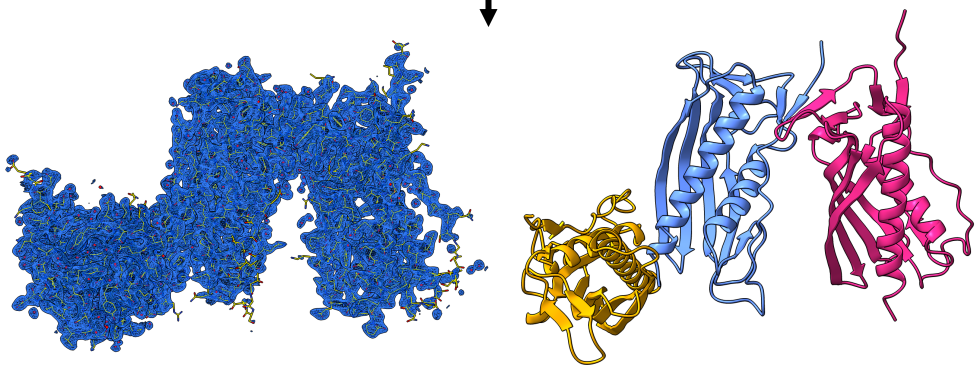
Since the pioneering myoglobin structure unveiled by John Kendrew in 1958, (Kendrew et al., 1958) X-ray crystallography has yielded a myriad of biological structures providing mechanistic insight into fundamental cellular processes and enabled novel applications such as structure-based drug design. The entire crystallographic method, from protein sample to atomic resolution structure, can be summarised into three steps: 1- crystal growth, 2- diffraction and 3- phasing and refinement, discussed in more detail below (Fig. 20).



1. Crystallisation



2. Diffraction



3. Phasing and Refinement

Fig. 20 – Overview of the crystallographic method

Pictorial representation of the three main steps in the crystallographic method: crystallisation, diffraction and phasing and refinement.

Crystal Growth

X-ray crystallography requires the growth of crystals comprising the protein(s) of interest as the scattering caused by a single molecule would fall below the detection threshold. Crystal growth remains the intrinsic bottleneck to structural determination via X-ray crystallography furthermore, the process of crystal formation itself remains poorly understood.

Protein crystallisation is typically performed using the 'vapour diffusion' methodology in which the solubility of a protein in solution is reduced over time hopefully initiating crystal growth and avoiding uncontrolled precipitation however free interface diffusion and batch methods are also frequently utilised. Vapour diffusion experiments may be performed in three different arrangements, sitting drop, hanging drop and sandwich drop, however, the underlying principle remains the same (Fig. 21).

A small drop of sample containing the protein of interest, at high purity (>98%) and high concentration (5-50 mg/ml) (Dessau and Modis, 2011; Newby et al., 2009), is mixed with a volume of crystallisation condition. Crystallisation conditions comprise buffered solutions possessing different pH values and precipitant (salt and organic polymer) concentrations, both factors known to affect the solubility of proteins. This small drop along with a larger reservoir volume of the crystallisation condition are sealed within a closed system allowing vapour diffusion to occur. The buffer and precipitant concentrations within the small drop are lower than those in the reservoir, owed to their dilution upon mixing with the protein sample. Water molecules within both the small drop and larger reservoir are in equilibrium with water molecules within the vapour phase. The concentration gradient between the reservoir and drop causes a net movement of water molecules away from the drop towards the reservoir, further reducing the volume of the drop. This reduction in volume occurs concurrently with increasing protein, buffer and precipitant concentrations within the drop potentially causing

the solution to become supersaturated and form crystals. The movement of water molecules will continue until an equilibrium is established in which the buffer and precipitant concentrations in the drop and reservoir are approximately equal.

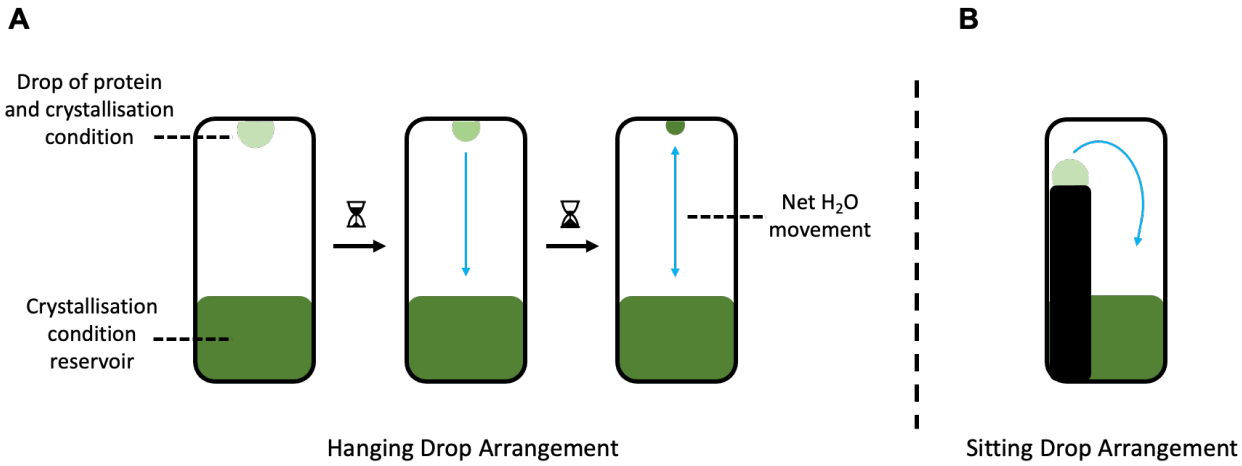


Fig. 21 – The Process of Vapour Diffusion.

(A) The process of vapour diffusion in a hanging drop arrangement. (B) Depiction of the sitting drop vapour diffusion arrangement. The process of vapour diffusion progresses identically to that in the hanging drop arrangement however the small drop is placed on a horizontal shelf rather than being suspended above the reservoir.

Diffraction

The individual positions of each atom comprising a crystal can be inferred from its X-ray diffraction pattern, via an electron density map intermediary, thus crystallography experiments aim to collect these diffraction patterns and solve the structures of crystalised biological molecules. The process by which a diffraction pattern is produced during an X-ray

diffraction experiment, Bragg's law of X-ray diffraction, was first postulated by father and son Lawrence and William Bragg in 1913 (Bragg and Bragg, 1913) and is detailed below.

Bragg's law considers a 2D view of crystal lattice planes – layers of molecules in a crystal lattice (Fig. 22). The law itself states $n \lambda = 2 d \sin \theta$, whereby: n = the order of diffraction, λ = the wavelength of X-rays, d = the distance between diffracting lattice planes and θ = the incident angle. Whilst most of the incident X-ray beam passes straight through the crystal (direct beam), some of the beam is diffracted by atoms within each layer of the lattice. The X-ray scattered by the second, lower lattice plane must travel further than the X-ray diffracted by the first, higher plane. If this extra distance travelled ($2 d \sin\theta$) by the lower wave equates to a whole integer (n) wavelength of the X-ray beam (λ) then the diffracted waves will remain in phase and thus interfere constructively to produce a single spot on the diffraction pattern. Waves travelling an extra distance not equal to an integer number of X-ray wavelengths, do not satisfy Bragg's law and, emerge out of phase and thus interfere destructively and are not detected. An alternative consideration is that the order of diffraction (n) is fixed at a specific integer meaning only lattice planes at a particular spacing (d) will satisfy Bragg's law.

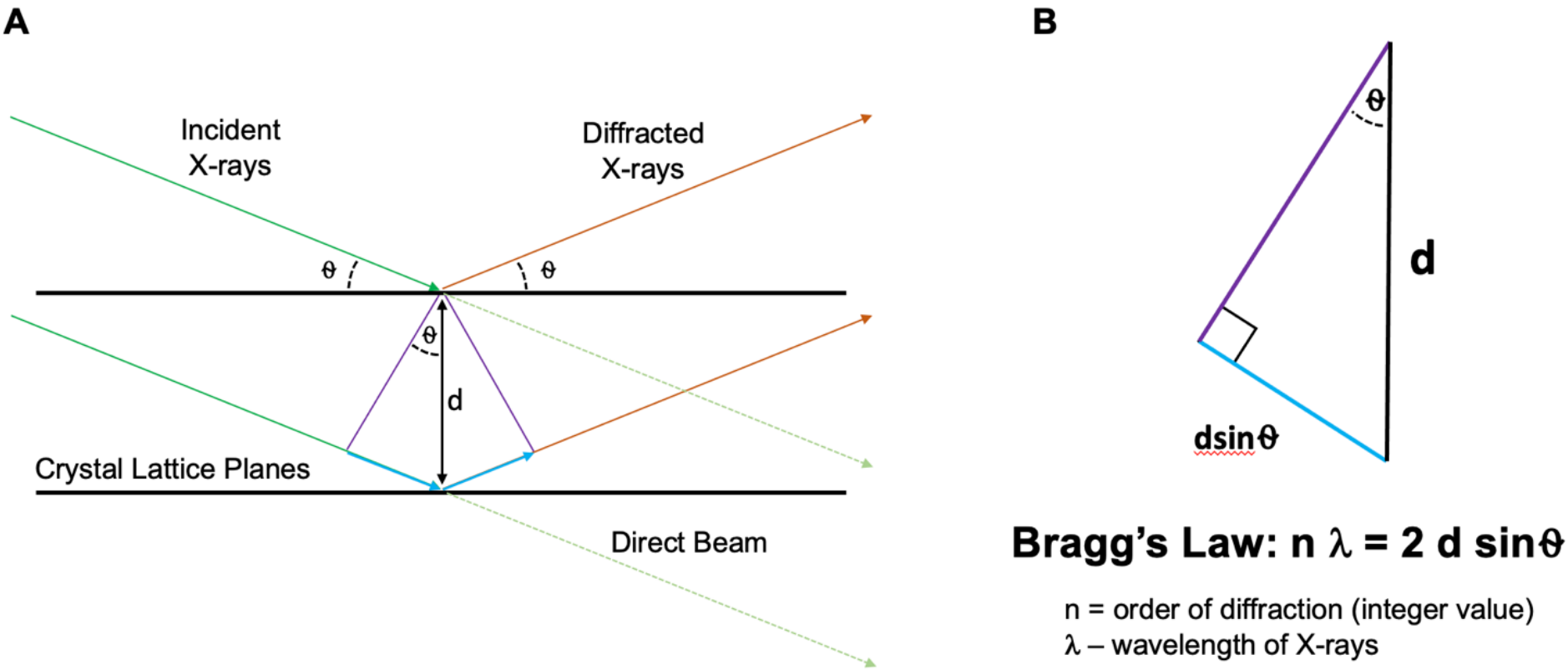


Fig. 22 – Bragg's Law of Diffraction.

Bragg's law of diffraction states that when diffracted X-ray (orange) waves remains in phase they interfere constructively. To remain in phase, the additional distance travelled by X-rays diffracted from lower lattice planes (blue, $2d\sin\theta$) must be equal to an integer number (n) of X-ray wavelengths (λ). Scattered X-rays which emerge out of phase do not satisfy Bragg's law ($n \neq$ integer value) and interfere destructively.

Phasing and Refinement

The phase problem of X-ray crystallography is a fundamental hurdle of the technique which must be overcome, arising due to the inability of phase information to be recorded during the experiment. Diffraction pattern spots provide the amplitude of the diffracted X-rays, however, these must be combined with phase information to reconstruct the electron density, and thus atomic positions, within the crystal. Phase information is typically estimated via one of two alternative methodologies and subsequently combined with the experimental amplitude information to determine an initial electron density map. The first of these methods involves the incorporation of heavy atoms into the unit cell, either via soaking crystals in heavy atom solutions or the insertion of selenomethionine residues to the protein during recombinant expression. Discrepancies in the diffraction patterns of the native and heavy atom containing crystals may be utilised to determine phases associated with the heavy atoms and thus locate them within the unit cell. Finally, the phases and amplitudes for the heavy atom sites are utilised in conjunction with the experimental amplitudes of the native and heavy atom derivative crystals to estimate initial phases for the entire structure.

Alternatively, initial phases may be derived via molecular replacement which utilises a molecule of known structure similar to that being determined. The molecular replacement procedure involves rotation of the known structure in three dimensions to resemble most closely that of the molecule of unknown structure within the unit cell. The rotated model is subsequently translated throughout the unit cell to identify its correct location. Finally, the positioned model may be used to derive an initial set of phases and corresponding electron density map.

Once an initial set of phases have been determined, these may be iteratively refined to achieve a final optimal electron density map. At each refinement step, the model is corrected against the current electron density map generating a new set of phases. These phases are

combined with the experimental amplitude information to generate a map allowing further model correction.

The progression of this cyclic process is monitored via the 'R value', a refinement statistic measuring how well amplitudes calculated from the current model match those observed experimentally. Comparing a completely random set of atoms would yield an R value of approximately 0.6 whilst an R value of 0 would indicate a perfect model. One apparent flaw arises with the use of this statistic, the refinement process itself aims to minimise the difference between experimental and calculated amplitudes thus the R factor itself is biased. This problem is overcome using two R values, R_{work} and R_{free} respectively. Typically, the refinement proceeds with access to only 90-95% of the reflections collected, with the remainder retained purely for validation purposes. Whilst R_{work} assesses the disagreement between model and those reflections utilised during refinement, R_{free} compares the model to reflections not utilised during refinement removing bias. R_{work} values will typically be slightly lower than their R_{free} counterparts (0.25 vs 0.3 for a ~2.5 Å structure), however, both should reduce equally during an unbiased refinement.

Chapter Aims

The structure of *E. coli* PqiC is yet to be determined, hence solving this structure was of principal interest to reveal its architecture whilst also gaining insight into residues both interfacing with PqiB and those potentially facilitating transport. Owed to the relatively small size (~20 kDa) of the PqiC monomer and the desire to achieve high resolution side chain information, X-ray crystallography was identified as the best technique for PqiC structural elucidation. Indeed, the structural determination of several PqiC homologues, referred to later in this chapter, via X-ray crystallography further reinforced the applicability of this

approach. This chapter discusses the crystallisation of PqiC which yielded two unique structures providing high resolution side chain and oligomeric state information respectively.

Results

Expression and Purification of PqiC FL

PqiC FL was expressed from an IPTG inducible pET-22b(+) plasmid transformed into *E. coli* C43 (DE3) cells with induction proceeding overnight at 18 ˚C. The isolated membrane fraction was solubilised in 1% w/v *n*-dodecyl-β-D-maltoside (DDM) before IMAC was utilised to isolate PqiC FL. IMAC elution fractions were analysed via SDS-PAGE (Appendix. 3) before those containing PqiC FL were pooled and further purified via a Superdex S200 16 600 PG size exclusion column (GE Healthcare) yielding a single large single peak after the void peak (Fig. 23A). The composition of elution fractions was determined via SDS-PAGE (Fig. 23B). The presence of PqiC FL was indicated by a major band between 20 and 25 kDa, however, fainter bands at higher molecular masses signify marginal contamination. Fractions containing PqiC FL were pooled and concentrated for crystallisation.

Expression and Purification of $PqiC^{17-187}$

$PqiC^{17-187}$ was expressed from an IPTG inducible pET-26b(+) plasmid transformed into *E. coli* C43 (DE3) cells with induction proceeding overnight at 18 ˚C. IMAC was utilised to isolate $PqiC^{17-187}$ from clarified cell lysate. IMAC elution fractions were analysed via SDS-PAGE (Appendix. 4) before those containing $PqiC^{17-187}$ were pooled and further purified via a Superdex S200 26 600 PG size exclusion column (GE Healthcare) resulting in a single major peak (Fig. 23C). The composition of elution fractions was determined via SDS-PAGE revealing no observable contamination (Fig. 23D). Elution fractions containing $PqiC^{17-187}$, indicated by a band at 20 kDa, were pooled, incubated with TEV-protease (purified historically within the lab) and isolated from the cleaved 6xHis-tag via an IMAC column. The

resulting purified, cleaved protein was concentrated in a 10 kDa MWCO Amicon® Ultra centrifugal filter unit (Merck Millipore) for crystallisation.

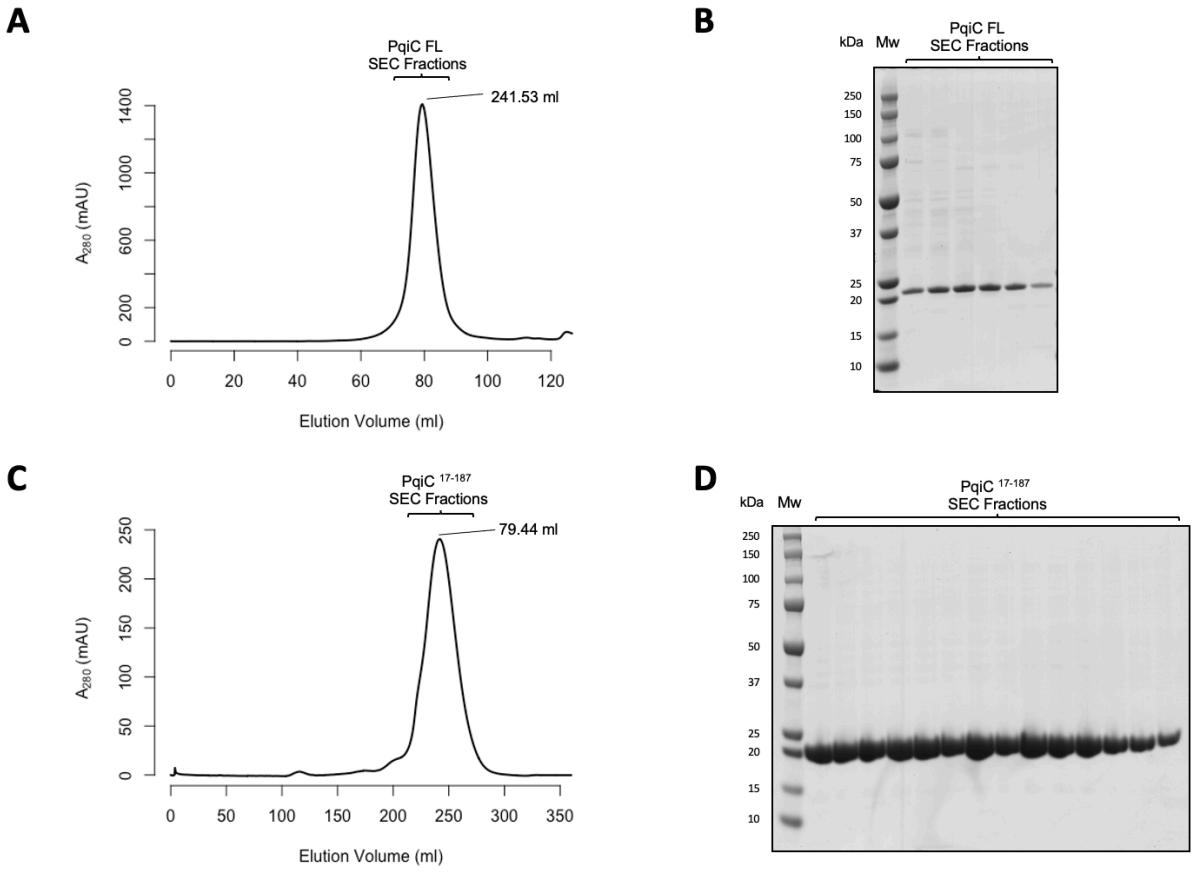


Fig. 23 – Purification of PqiC FL and PqiC$^{17-187}$.

(A) SEC trace of PqiC FL showing a single elution peak from a Superdex S200 16 600 PG column. (B) SDS-PAGE of PqiC FL SEC fractions show a predominant band between 20 and 25 kDa with minor contaminants at higher molecular masses. (C) SEC trace of PqiC$^{17-187}$ showing a single elution peak from a Superdex S200 26 600 PG column. (D) SDS-PAGE of PqiC$^{17-187}$ SEC fractions showing a single band at 20 kDa with no major contaminants.

Structure Determination of PqiC FL

Having determined that we could purify and concentrate large amounts of PqiC FL, we progressed to screening crystallisation conditions. An initial hit was observed in a 400 nl sitting drop well comprising a 1:1 ratio of 6 mg/ml PqiC FL and 0.2 M Potassium citrate tribasic monohydrate, 15 % w/v SOKALAN® CP42, kindly prepared by Giedrė Ratkevičiūtė during a 3-month rotation project under my supervision (Fig. 24A-B). The crystal, belonging to the I 4 2 2 space group, diffracted to 3.4 Å however the dataset suffered a very low signal-to-noise ratio (I/sigma) and required trimming back to 3.8 Å for processing. The data was phased using two copies of the Alphafold PqiC prediction (AF-P0AB10-F1) (Jumper et al., 2021) trimmed to remove the N-terminal signal peptide as the molecular replacement model, producing a solution with a Log Likelihood Gain (LLG) of 692.3 and TF Z-score (TFZ) of 27.8. Refinement proved challenging, due to large areas of ambiguous density, and the final model, hereafter referred to as CFL1, could not be improved beyond an R_{work}/R_{free} of 0.3634/0.3813, Table 14.

This initial model was superseded by a subsequent hit grown in a 400 nl sitting well drop comprising 1:1 ratio of 0.52 mg/ml PqiC WT and 0.3 M lithium sulphate, 0.1 M MES; pH 6.5, 25% v/v PEG 400 (Fig. 24C-D). This crystal, once again belonging to the I 4 2 2 space group, diffracted to 3.2 Å. The data was phased using two copies of chain B from the Csol model (see later in chapter) as the molecular replacement model producing a solution with an LLG and TFZ of 829 and 26.4 respectively. Refinement proved far more successful yielding a final model, hereafter referred to as CFL2, with R_{work}/R_{free} of 0.2469/0.2694, Table 14.

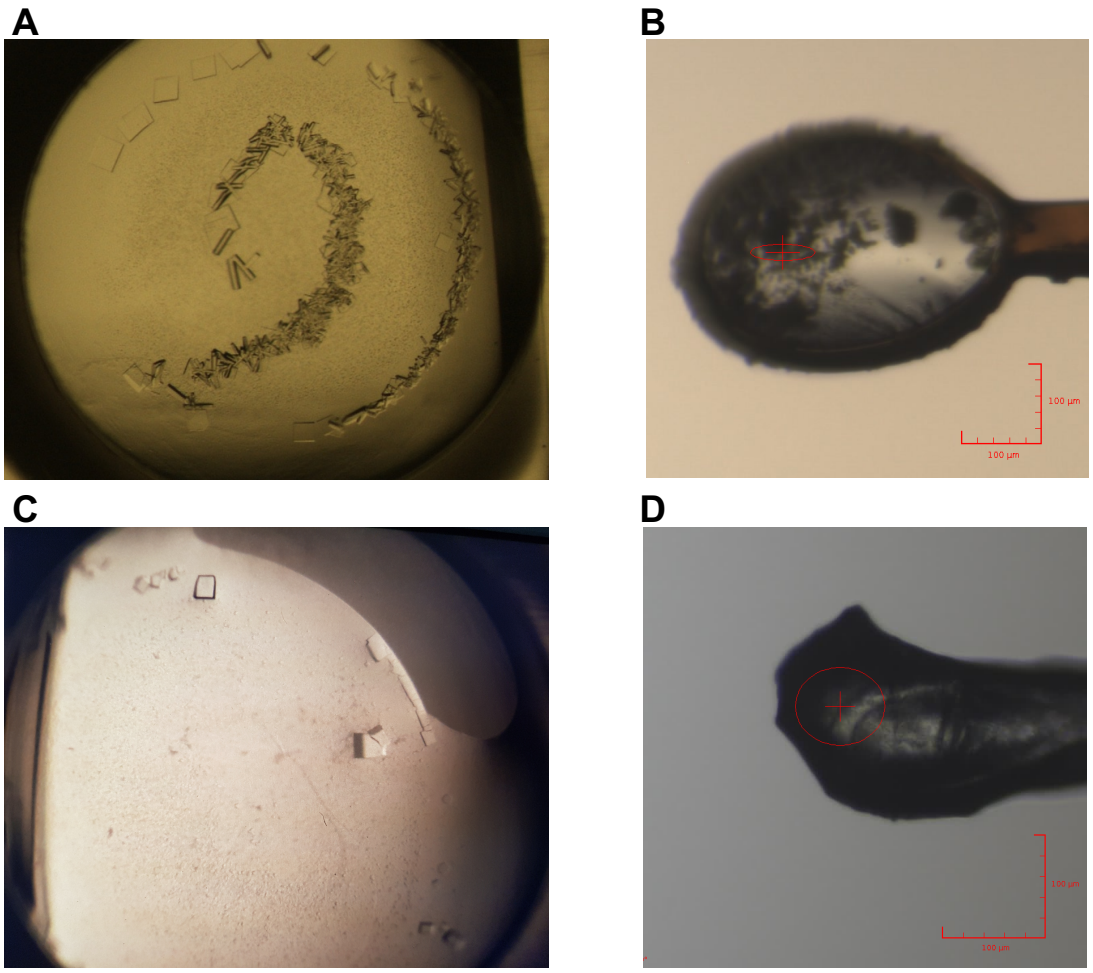


Fig. 24 – PqiC FL Crystals.

(A) and (B) Initial crystals formed in 0.2 M Potassium citrate tribasic monohydrate, 15 % w/v SOKALAN® CP42, from which the CFL1 structure was produced. (C) and (D) Subsequent crystals formed in 0.3 M lithium sulphate, 0.1 M MES; pH 6.5, 25% v/v PEG 400 from which the CFL2 structure was produced.

The CFL1 and CFL2 structures displayed the same overall conformation, thus an alignment of the two asymmetric units yields an RMSD of 0.693 Å across 312 pruned atom pairs and 0.955 Å across all 330 pairs (Fig. 25). The core regions of both structures are nearly identical with instances of deviation located within the loop regions. The greatest deviation is observed in the first β1-β2 loop which we attribute to its presumed flexibility and ability to adopt an array of conformations (Fig. 25). This region was by far the most poorly resolved

within the electron density map and required the loop to be manually built, however, there was insufficient density for the explicit positioning of side chains throughout this region (Fig. 36). Furthermore, it proved impossible to model the entire N-terminus, to the acylation site (Cys16), with the CFL1 and CFL2 structures starting at residues 21 and 23 respectively.

Whilst neither PqiC FL structure yielded-high resolution side chain information, the arrangement of monomers within the unit cell implied quaternary structural information. Observing the symmetry copies for both structures revealed PqiC FL to have crystallised in an octameric state, with the 8 copies forming a “doughnut-like” ring arrangement (Fig. 26), which we believe to be the biological assembly a notion is corroborated by the following evidence. The SEC elution volume of purified PqiC FL is not consistent with that of a monomeric PqiC protomer. Instead, the protein elutes at an earlier elution volume indicative of a larger hydrodynamic radii most likely resulting from multimerisation. Secondly, the homologous structure of *Xanthomonas campestris* XCC0632 (17.3% sequence identity, 30.8% similarity) was solved with an asymmetric unit comprising 8 copies in two semi-circular arrangements (PDB 2IQI). Once again, the symmetry in the unit cell allowed the formation of complete octameric ring structures analogous to those observed in our PqiC FL datasets. Finally, the internal diameter of the pore formed by the octameric PqiC FL arrangement is approximately 40 Å and thus consistent with the width of the PqiB coiled-coil with which PqiC is anticipated to interact (Fig. 26).

The low spherical completeness (31.98 %) of the CFL2 data, following anisotropic truncation via STARANISO, must be noted as a potential cause for concern and raises the possibility that the final model is in fact dominated by the molecular replacement search model. Consequently, the similarity of the CFL2 monomer to the preceding CFL1 monomer, model which was derived from data with a higher completeness (97.83 %), provides some reassurance in its structure. Finally, arguably the most significant feature of both CFL

structures is their octameric arrangement which, considering the monomeric nature of the molecular replacement model, must arise from the diffraction data alone.

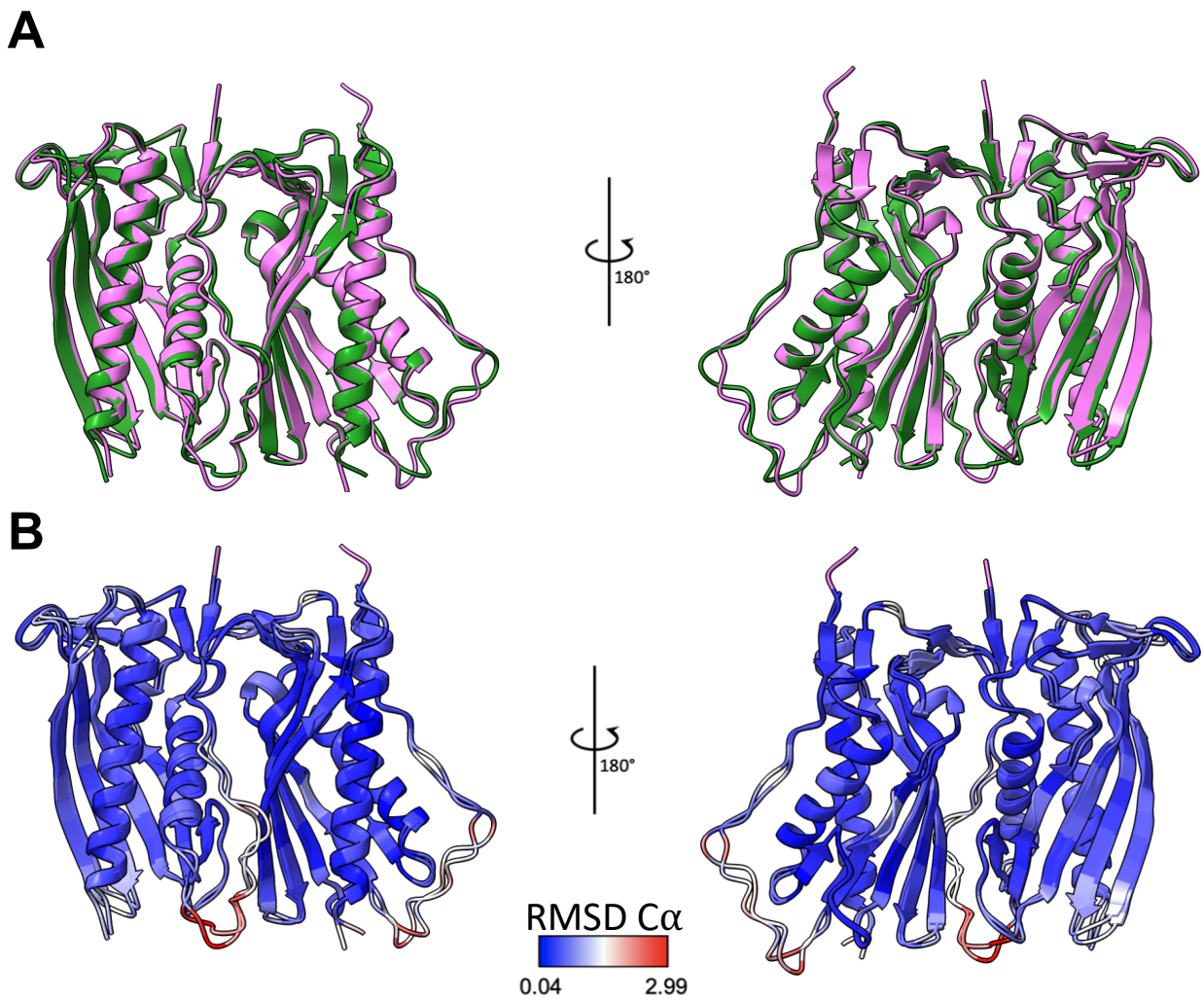


Fig. 25 – Overlay of CFL1 and CFL2 Structures.

(A) Overlay of the aligned CFL1 (orchid) and CFL2 (green) structures highlighting the similarity in their overall architecture. (B) Overlay of the aligned CFL1 and CFL2 structures coloured according to the RMSD of aligned Cα pairs – colour key at bottom centre. The structures deviate very little with discrepancy only appearing in the loop regions, particularly the large β1-β2 loop.

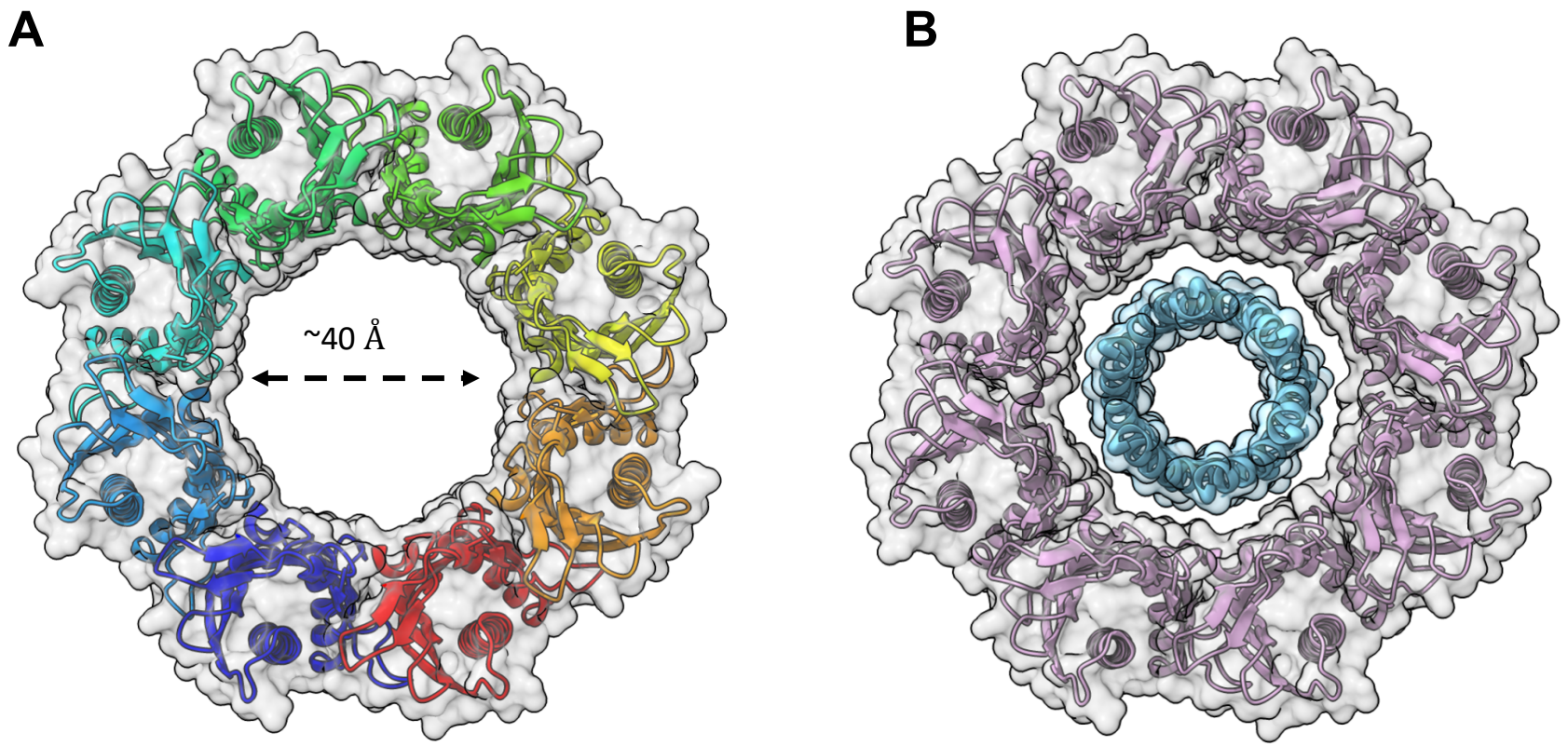


Fig. 26 – Architecture of the PqiC FL Octamer.

(A) Cartoon representation of the CFL2 octamer viewed from the membrane plane. Chains coloured in a rainbow gradient, molecular surface depicted in transparent grey and the approximate diameter of the central pore indicated. (B) Structures of the CFL2 octamer (orchid cartoon, transparent grey surface) and the PqiB coiled-coil (sky blue cartoon and transparent blue surface) viewed from the membrane plane highlighting their potential to interface. (PqiB PDB: 5UVN).

Considering the anticipated biological significance of this octameric arrangement, we examined both the molecular lipophilicity potential (Fig. 27) and coulombic electrostatic potential (Fig. 29) of its surface for indications of structurally and/or functionally critical residues.

First, we examined the lipophilic potential of the structure (Fig. 27). Of primary interest were two rings of lipophilic residues navigating the circumference of the central pore at two distinct heights. The residues comprising these lower (Trp47) and upper rings (Tyr25, Ile73 and Leu78) were the only lipophilic residues located within the primarily hydrophilic central pore (Fig. 28A-B). The apparent distinct and precise location of these residues within the central pore implied a possible relation to function. Indeed, if PqiB were to interface with the central pore of PqiC then these residues may contribute to the stabilisation of such an interaction.

In contrast the lipophilic residues observed on the outer and lower faces of the PqiC FL octamer appeared to be predominantly involved in packing interactions to stabilise the quaternary structure. The top, membrane occluded face was also found to possess lipophilic residues (Tyr64, Val72, Tyr127, Tyr161) which reside predominantly within the centre of this face and flanked by regions of hydrophilicity, however, their role appears unclear (Fig. 28C-D).

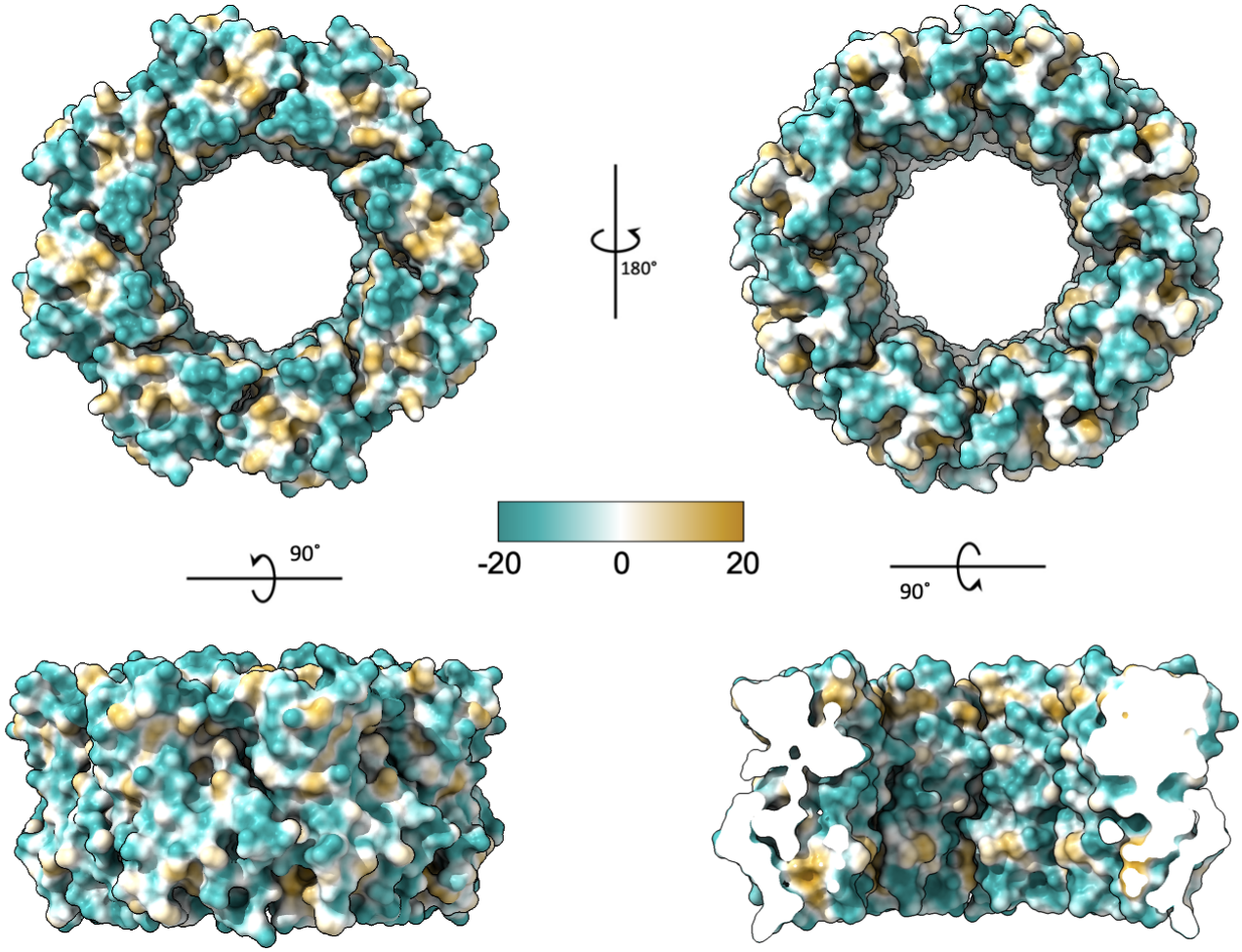


Fig. 27 – Molecular Lipophilicity Potential of PqiC FL Octamer Surface.

Orthogonal views of the PqiC FL octamer surface coloured according to lipophilicity with hydrophilic and lipophilic residues indicated in dark cyan and gold respectively – colour key in centre. Views: (left top) from the membrane plane, (right top) from the periplasmic face, (left lower) outer surface, (right lower) inner surface.

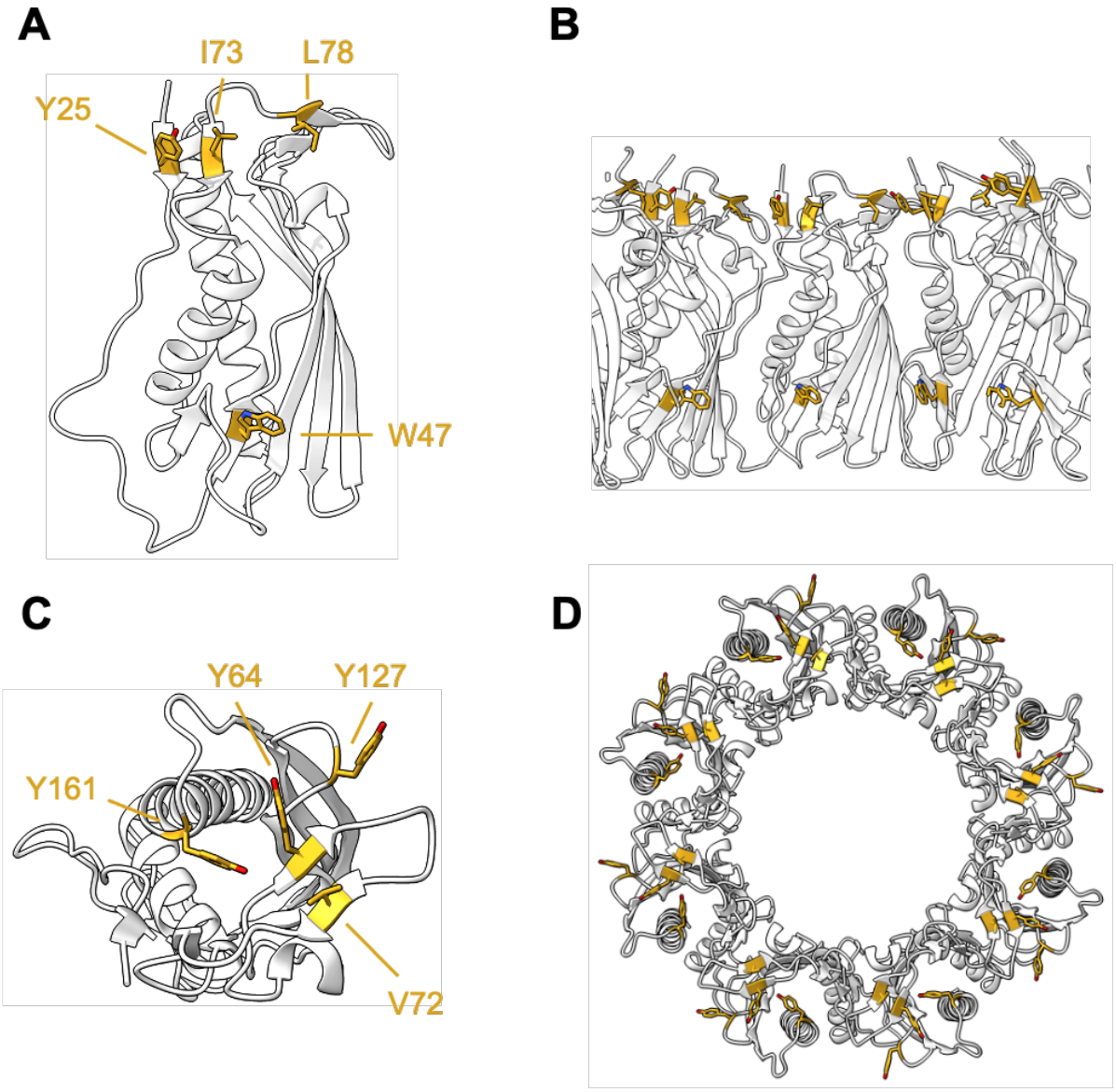


Fig. 28 – Lipophilic Residues of Interest Within PqiC FL.

(A) Cartoon representation of chain A from the CFL2 structure. Lipophilic residues of interest located within the octameric PqiC FL pore are depicted in gold. (B) Cartoon representation of the PqiC FL octamer, viewed from within the central pore, indicating the positioning of the highlighted lipophilic residues. (C) Cartoon representation of chain A from the CFL2 structure. Lipophilic residues of interest located on the membrane occluded face of the PqiC FL octamer are depicted in gold. (D) Cartoon representation of the PqiC FL octamer, viewed from the membrane plane, indicated the positioning of the highlighted lipophilic residues.

We next examined the coulombic electrostatic potential of the structure (Fig. 29). Once again, the area of primary interest was the lining of the central pore where three distinct regions of negative electrostatic potential were observed. In contrast, only a single positively charged residue (Arg88) was found to line the pore. The top negative region, arising from Asp55 and Asp84, perfectly bisects the two aforementioned lipophilic regions. The mid region, consisting of Glu49, sits adjacent to the lower lipophilic region hence at this height the pore circumference encompasses alternating lipophilic and negatively charged regions. Finally, a lower region encompassing Asp114 sits towards the lower lip of the pore. The role of these three negative regions is unclear however, considering their positioning, it's conceivable that these may to form interactions with the coiled-coil of PqiB if the two were to combine as anticipated (Figs. 29,30). Utilising the ConSurf server (Landau et al., 2005) we investigated the sequence conservation of the residues lining the PqiC pore. All the residues involved in the two lipophilic pore lining bands (Tyr25, Trp47, Ile73 and Leu78), as well as those involved in the upper negatively charged band (Asp55 and Asp84) were shown to be highly conserved (Fig. 31). In contrast, the residues belonging to the middle and bottom negatively charged bands (Glu49 and Asp114) showed no significant conservation (Fig. 31). The enriched conservation of these pore lining residues supports their importance in function with Asp55 appearing particularly prominent as a conserved residue.

The membrane occluded face of PqiC displayed distinct regions of electrostatic potential either side of the lipophilic band identified previously. Eight intense regions of positive charge are located around the inner circumference arising from Lys23. The outer circumference houses several aspartate residues, Asp68, Asp128, Adp159 and Asp162, which give rise to regions of negative charge (Fig. 30C-D). Once again, we utilised ConSurf to investigate the sequence conservation of these residues (Lys23, Asp68, Asp128, Asp159 and Asp162) and found them to be highly conserved supporting our notion of their involvement in function (Fig. 31). The distinct regions of positive, negative and lipophilic residues upon this face, arising

from conserved residues, are likely crucial for the correct function of PqiC and require further investigation.

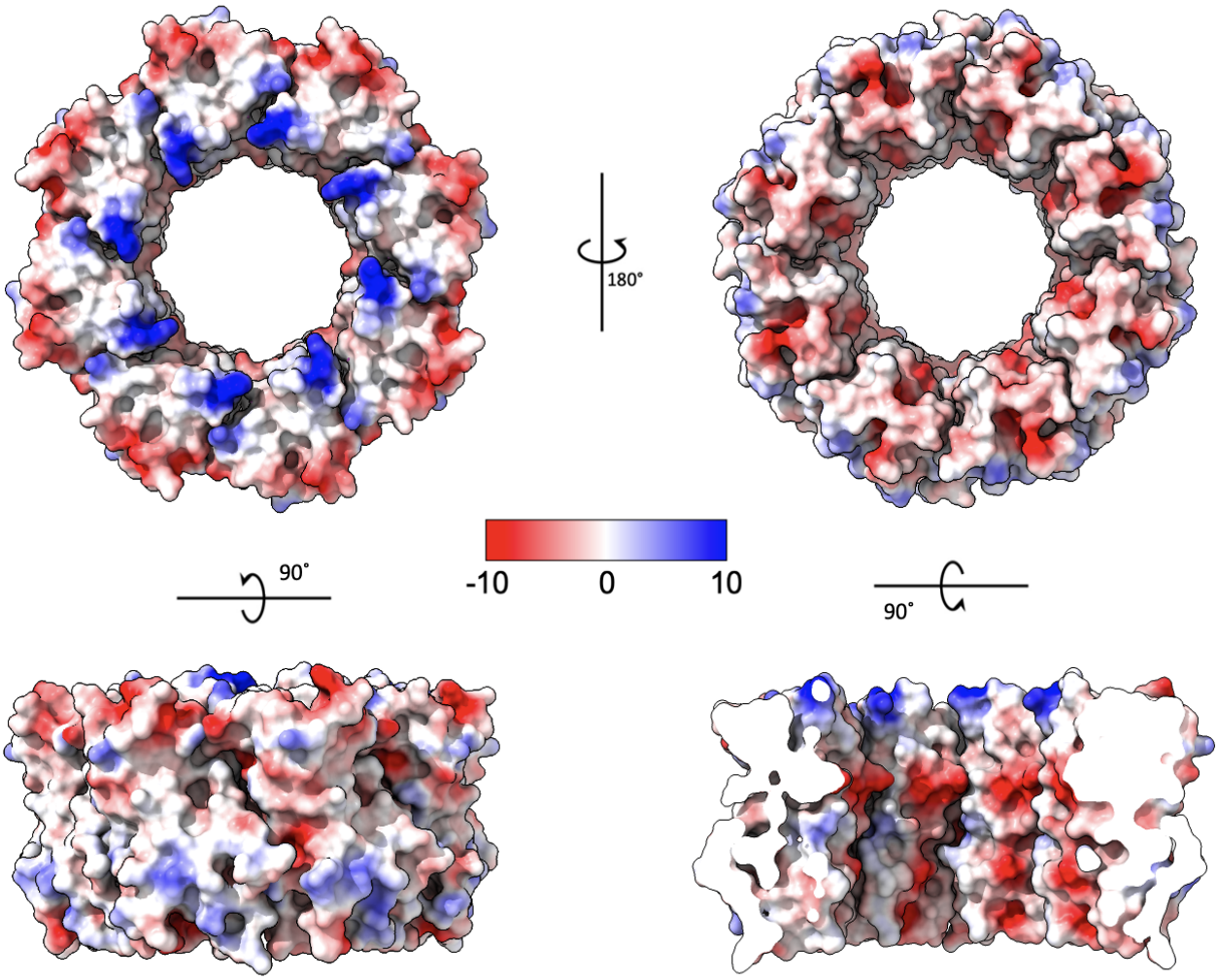


Fig. 29 – Coulombic Electrostatic Potential of the PqiC FL Octamer Surface.

Orthogonal views of the PqiC FL octamer surface coloured according to coulombic electrostatic potential with negative and positive potential indicated in red and blue respectively – colour key in centre. Views: (left top) from the membrane plane, (right top) from the periplasmic face, (left lower) outer surface, (right lower) inner surface.

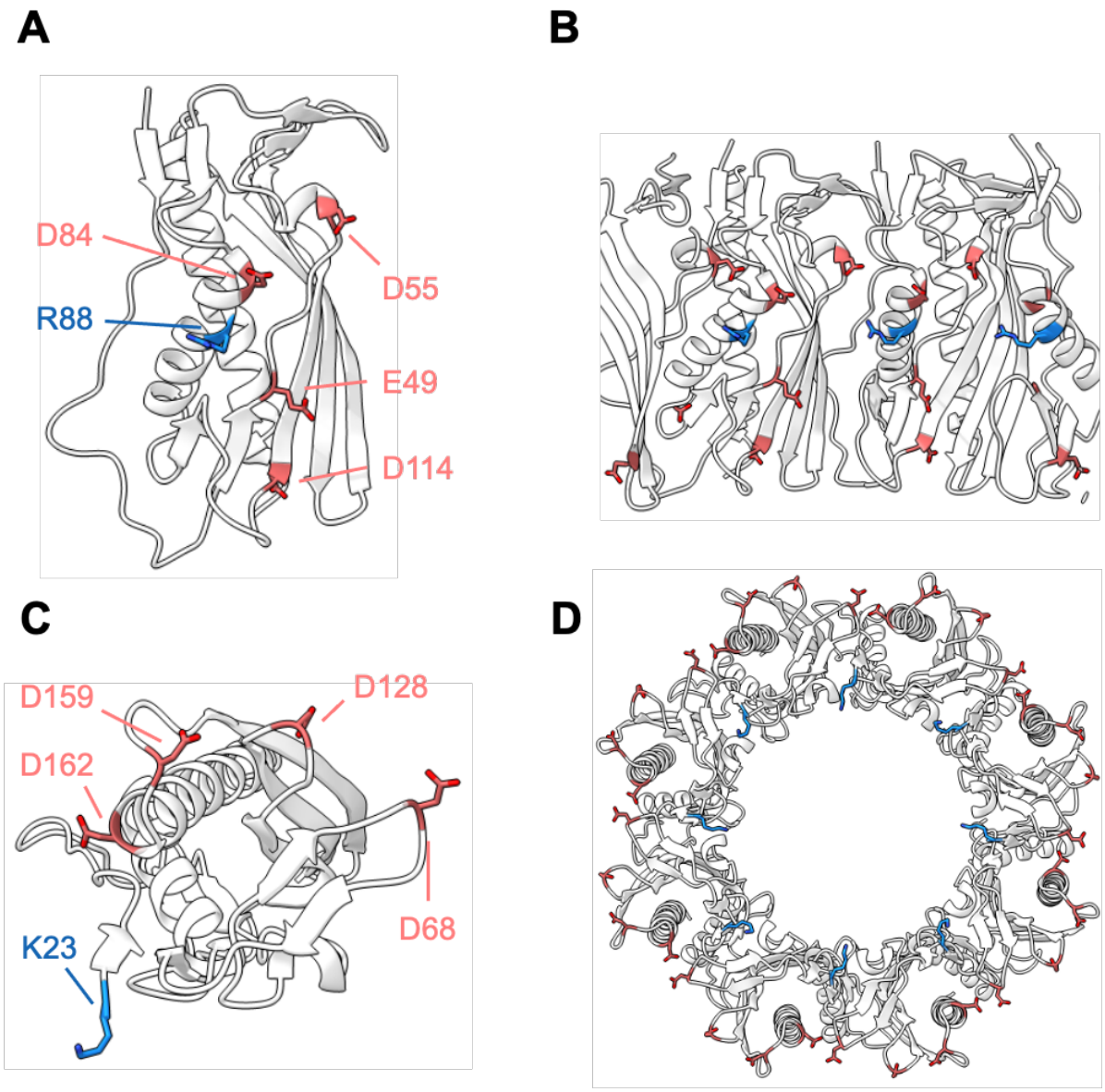


Fig. 30 – Charged Residues of Interest Within PqiC FL.

(A) Cartoon representation of chain A from the CFL2 structure. Positive and negative residues of interest located within the octameric PqiC FL pore are depicted in blue and red respectively. (B) Cartoon representation of the PqiC FL octamer, viewed from within the central pore, indicating the positioning of the highlighted charged residues. (C) Cartoon representation of chain A from the CFL2 structure. Positive and negative residues of interest located within the octameric PqiC FL pore are depicted in blue and red respectively. (D) Cartoon representation of the PqiC FL octamer, viewed from the membrane plane, indicated the positioning of the highlighted charged residues.

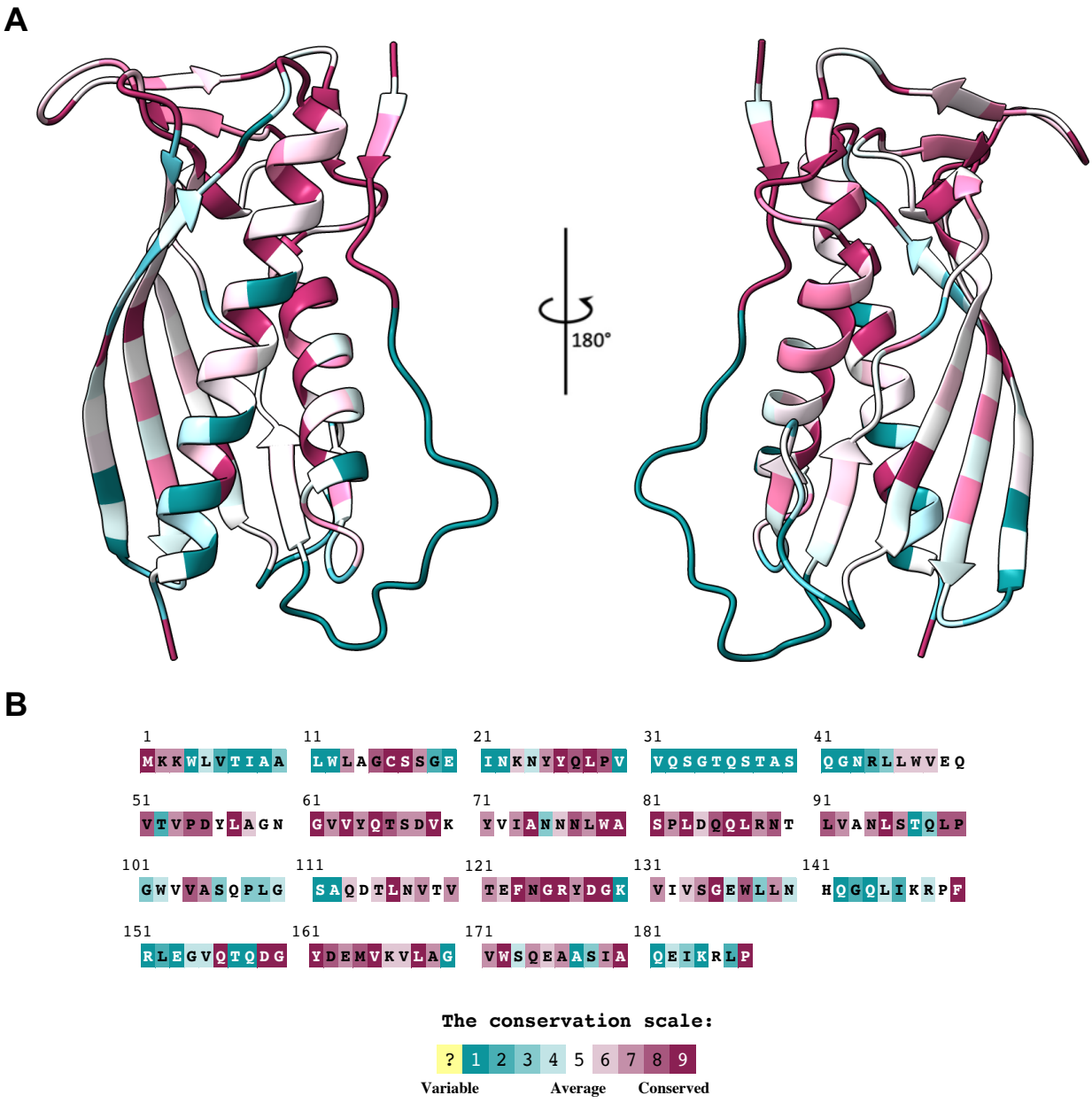


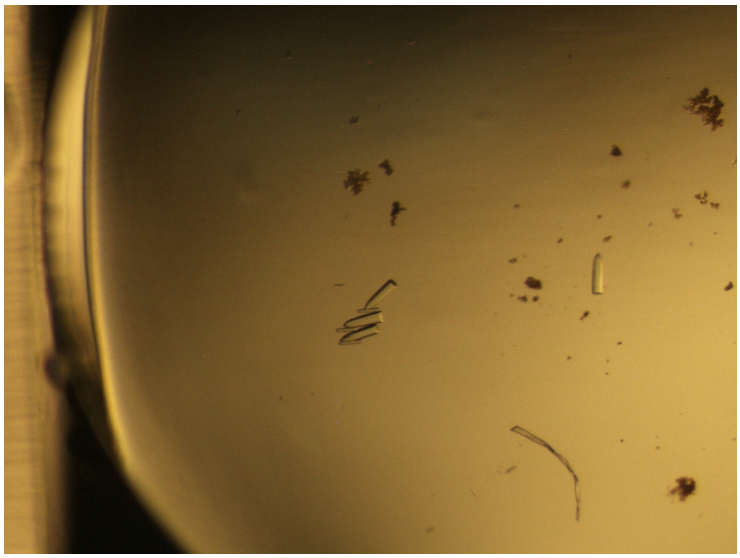
Fig. 31 – Sequence Conservation of PqiC FL.

(A) Orthogonal views of chain B of the CFL2 structure displayed as cartoon representation with residues coloured according to the Consurf sequence conservation. (B) PqiC FL sequence coloured according to the Consurf sequence conservation. Conservation scale is included at the bottom centre of the figure.

Structure Determination of PqiC$^{17-187}$

Whilst the PqiC FL construct appeared to reveal an octameric assembly of PqiC, the resulting maps were not of sufficient resolution to determine accurate side chain information across the entire structure. Consequently, we re-focussed upon crystallisation of the soluble PqiC$^{17-187}$ construct with the aim of resolving the high-resolution geometry currently lacking, with a particular focus upon the β1-β2 loop. The PqiC$^{17-187}$ construct was found to form "needle-like" crystals readily, however, these either diffracted poorly or not at all. A subsequent hit was grown in a 400 nl sitting well drop comprising 1:1 ratio of 33 mg/ml PqiC$^{17-187}$ and 0.1 M sodium chloride, 0.1 M MES; pH 6.5, 33% v/v PEG 400 (Fig. 32). The crystal, belonging to the P 61 space group and diffracted to 2.1 Å. The data was phased using three copies of chain B from the CFL1 model as the molecular replacement model yielding a solution with LLG and TFZ of 2453.6 and 46.2 respectively. Subsequent refinement yielded a model, referred to as Csol hereafter with R_{work}/R_{free} of 0.1988/0.2139, Table. 14.

A



B

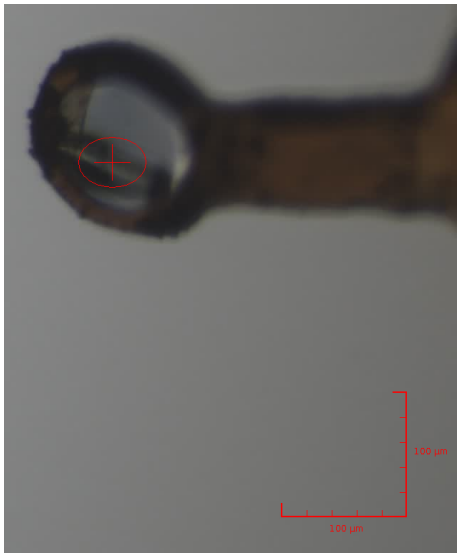


Fig. 32 – PqiC$^{17-187}$ Crystals.

(A) Crystals formed in 0.1 M sodium chloride, 0.1 M MES; pH 6.5, 33% v/v PEG 400 from which the Csol structure was produced. (B) PqiC$^{17-187}$ crystal mounted in the beamline.

The Csol structure displayed a conformation unique to that observed in the CFL structures with three copies present within the asymmetric unit (Fig. 33). Two of these chains, chains A and B, sit in orientations similar to those of the two chains in the CFL asymmetric unit, however, the third chain, chain C, sits perpendicular and slightly below the plane of the other two interfacing with the bottom corner of chain A (Fig. 33). Whilst chains A and B in the CFL structures sit parallel to one another, in the Csol structure chain A sits marginally higher and is rotated by ~15˚ relative to chain B (Fig. 34).

Despite this global geometry change, the monomers from each structure align strongly yielding an RMSD of 0.628 Å across 155 pruned atom pairs and 1.094 Å across all 165 pairs (Fig. 35). The cores of each structure appear almost identical, with only slight variations observed in the loop regions adjoining secondary structure motifs. Once again, the largest

variations in RMSD were localised to the β1-β2 loop, further reinforcing the notion of its inherent flexibility and suggesting its observed position is likely governed by crystal packing interactions (Fig. 35).

The PqiC$^{17-187}$ construct was crystallised with the aim of obtaining a structure of higher resolution than those resulting from the PqiC FL construct to reveal the high-resolution geometry absent from both CFL structures. The Csol structure was solved at 2.1 Å and thus provided the desired overall improvement in resolution over the CFL2 structure, however, not all regions of the final refined map yielded improved electron density. One of the primary aims was to improve the map density for the β1-β2 loop region as this had been the poorest section of the CFL2 map. Whilst the density corresponding to the first residues of this region (residues 21-31) is more clearly resolved than in the CFL2 structure, the density for the extremity of the loop (residues 33-45) is much poorer and mostly absent from the 2mFo – DFc map contoured at 1 sigma (Fig. 36). Despite this, the density corresponding with the core of the protein was a considerable improvement over that of the CFL2 structure with side chain information clearly defined in both β-sheet and α-helical regions (Fig. 36).

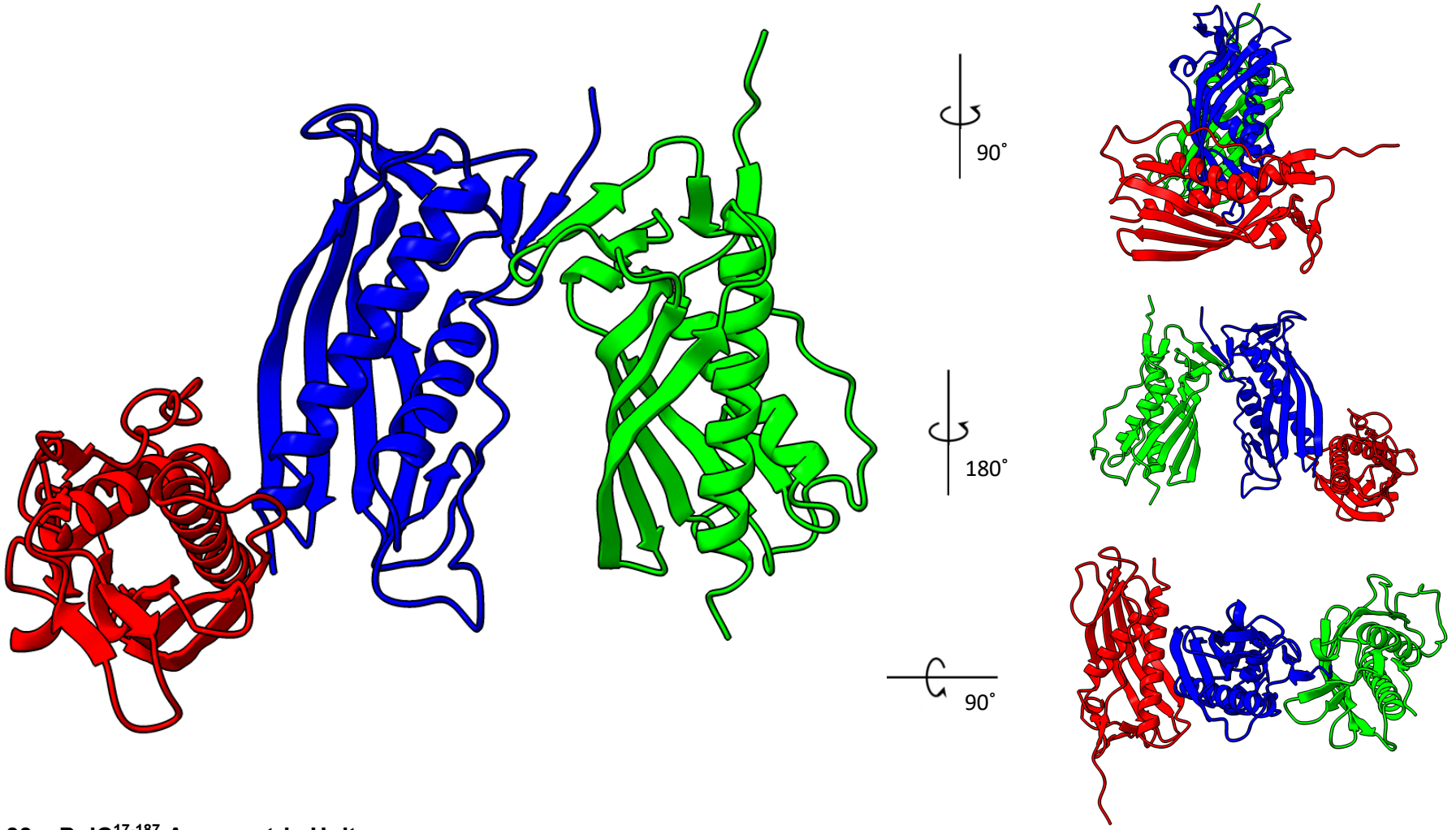


Fig. 33 – PqiC$^{17-187}$ Asymmetric Unit.

Cartoon representation of the PqiC$^{17-187}$ CsoI structure asymmetric unit with chains A, B and C coloured in blue, green and red respectively.

Orthogonal views, with their angle relative to the main image indicated, are depicted to the right of the image.

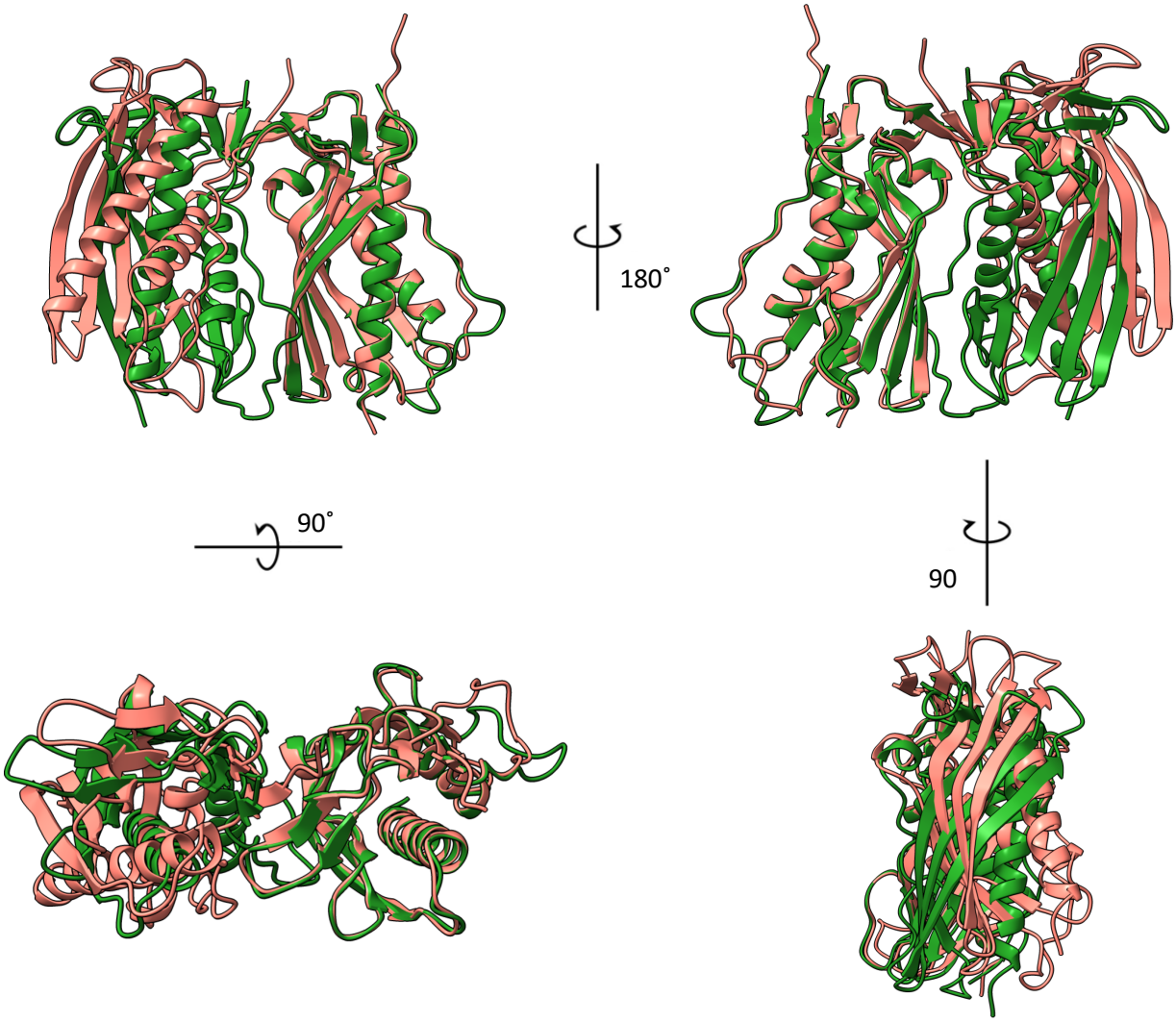


Fig. 34 – Comparison the CFL2 and Csol Asymmetric Unit Geometry.

Overlay of the aligned CFL2 (green) and Csol (salmon) indicating the differing geometries of the A and B chains in both structures. Note Chain C from the Csol structure has been removed for clarity.

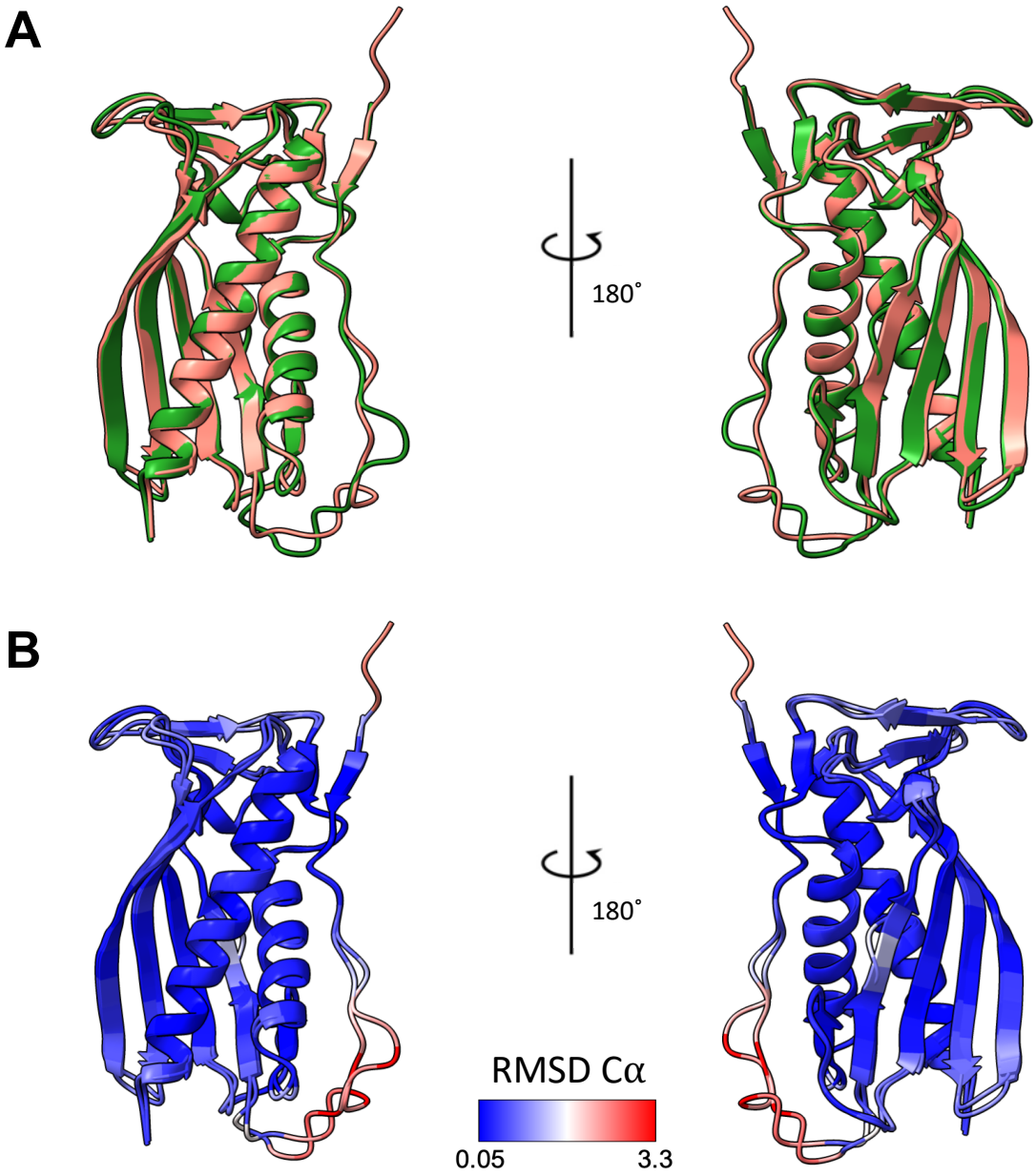


Fig. 35 – Alignment of CFL2 and Csol Monomers.

(A) Overlay of the two aligned chain Bs from the CFL2 (green) and Csol (salmon) structures highlighting the similarity in their overall architecture. (B) Overlay of the two aligned chain Bs from the CFL2 (green) and Csol (salmon) structures coloured according to the RMSD of aligned Cα pairs – colour key at bottom centre. The core of the two structures is near identical with discrepancy only appearing in the loop regions, particularly the large β1-β2 loop.

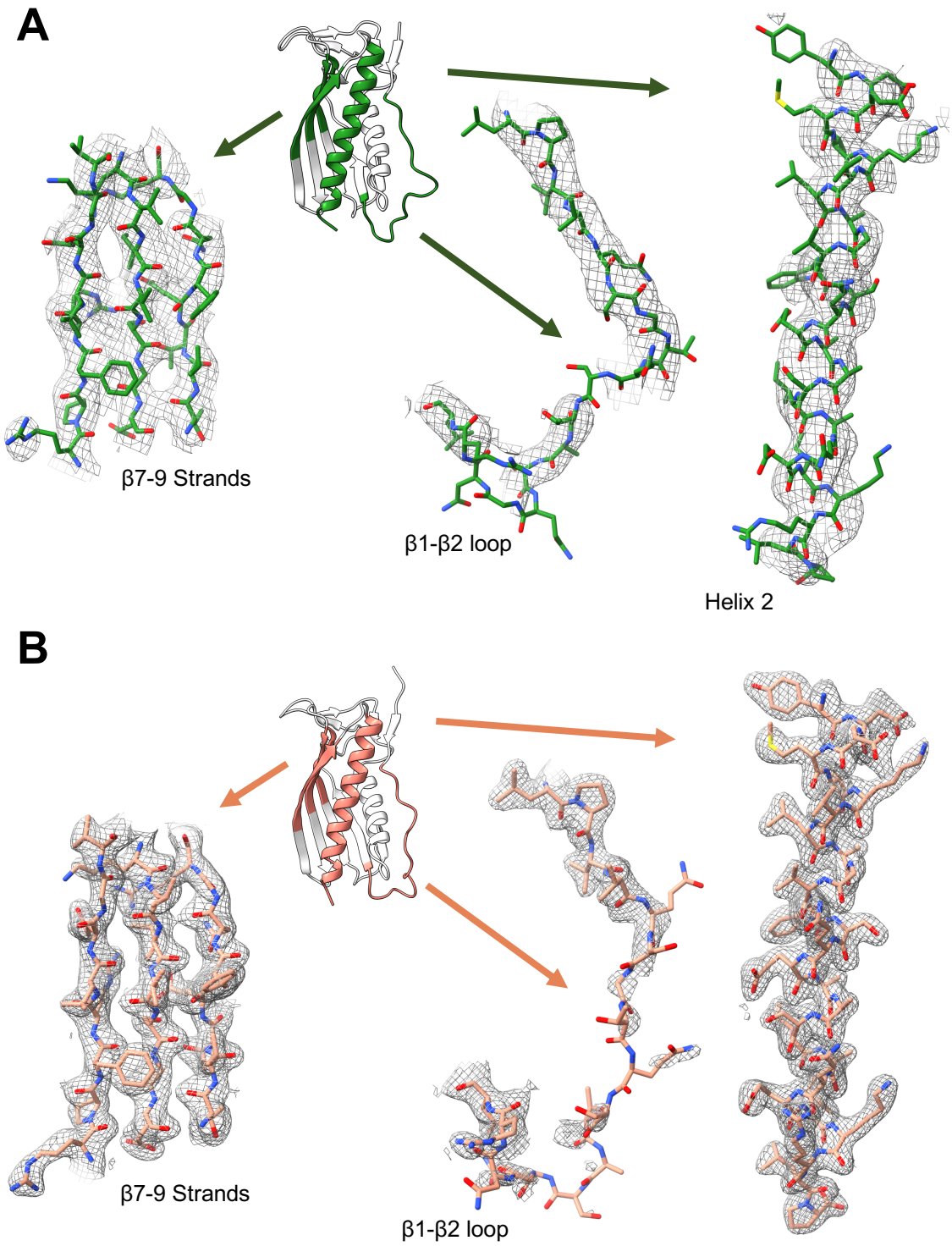


Fig. 36 – Comparison of CFL2 and Csol Electron Density.

The electron density in the final refined map for the (A) CFL2 and (B) Csol structures is displayed for three key regions of the protein - the β7-9 antiparallel strands, helix 2 and the β1-β2 loop. The 2mFo – DFc map (grey) is contoured at 1 sigma and protein displayed as stick representation coloured green (CFL2) and salmon (Csol).

Table. 14 – PqiC structure data collection and refinement statistics.

	PqiC FL – CFL1	PqiC FL - CFL2	PqiC$^{17-187}$ – Csol
Wavelength	0.9763 Å	0.9797 Å	1.0000 Å
Resolution range	43.27 - 3.5 (3.625 - 3.5)	53.77 - 3.211 (3.326 - 3.211)	38.57 - 2.081 (2.155 - 2.081)
Space group	I 4 2 2	I 4 2 2	P 61
Unit cell	132.902 132.902 202.16 90 90 90	132.148 132.148 185.043 90 90 90	133.603 133.603 65.715 90 90 120
Total reflections	23547 (2306)	8788 (132)	67192 (450)
Unique reflections	11775 (1153)	4394 (66)	33852 (225)
Multiplicity	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)
Spherical Completeness (%)	97.83 (99.57)	31.89 (4.96)	84.00 (5.60)
Elliptical Completeness (%)	N/A	90.6 (72.4)	95.7 (56)
Mean I/sigma(I)	8.19 (1.22)	6.86 (1.97)	18.79 (1.24)
Wilson B-factor	142.39	77.90	46.75
R-merge	0.01867 (0.6582)	0.1247 (0.3041)	0.02414 (0.5142)
R-meas	0.0264 (0.9309)	0.1763 (0.4301)	0.03414 (0.7272)
R-pim	0.01867 (0.6582)	0.1247 (0.3041)	0.02414 (0.5142)
CC1/2	1 (0.794)	0.977 (0.787)	0.999 (0.613)
CC*	1 (0.941)	0.994 (0.938)	1 (0.872)
Reflections used in refinement	11545 (1148)	4383 (66)	33847 (225)
Reflections used for R-free	577 (60)	217 (2)	1683 (11)
R-work	0.3634 (0.6265)	0.2469 (0.3901)	0.1988 (0.3199)
R-free	0.3813 (0.6439)	0.2694 (0.2595)	0.2139 (0.3290)
CC(work)	0.877 (0.710)	0.881 (0.758)	0.954 (0.744)
CC(free)	0.960 (0.724)	0.770 (1.000)	0.935 (0.802)
Number of non-hydrogen atoms	2606	2582	4159
Macromolecules	2606	2582	3943
Ligands	0	0	0
Solvent	0	0	216
Protein residues	333	330	504
RMS(bonds)	0.005	0.008	0.010
RMS(angles)	0.86	1.29	1.40
Ramachandran favored (%)	93.31	96.01	96.79
Ramachandran allowed (%)	5.78	3.68	3.21
Ramachandran outliers (%)	0.91	0.31	0.00
Rotamer outliers (%)	0.00	0.71	0.70
Clashscore	10.04	13.06	1.79
Average B-factor	237.93	107.89	57.69
macromolecules	237.93	107.89	58.01

Statistics in parenthesis refer to the highest resolution shell.

Discussion

PqiC FL

In this chapter we have presented the X-ray determined structure of PqiC FL, crystallising as an octameric assembly which we believe to be biologically relevant considering the SEC elution volume of PqiC, the existence of octameric structures for homologous proteins and compatibility of the octamer dimensions with its proposed interaction partner PqiB.. To identify homologous structures deposited to the PDB we utilised the Dali protein structure comparison server (Holm 2020). Dali returned six hits with an RMSD below 2.5 Å; *Enterobacter cloacae* PqiC (PDB: 6OSX), *Xanthomonas campestris* XCC0632 (PDB: 2IQI), *Paraburkholderia phytofirmans* PelC (PDB: 5T10), *Geobacter metallireducens* PelC (PDB: 5T0Z), *Acinetobacter baumannii* LptE (PDB: 5TSE) and *Shewanella oneidensis* LptE (PDB: 2R76) (Fig. 37).

Interestingly, the monomers of the deposited *Paraburkholderia phytofirmans* PelC structure form a dodecameric ring arrangement within the crystal which the authors propose to be the biological assembly (Fig. 38A), raising speculation that PqiC may adopt a similar architecture *in vivo* (Marmont et al., 2017). It must however be noted that the structure of the PelC monomer is distinctly different to that of PqiC, with the largest differences being observed at the N-terminus. Alternatively, the *Xanthomonas campestris* XCC0632 monomer displays a near identical structure to that of PqiC (Fig. 37). Furthermore, deposited structure the contains 8 copies within the asymmetric unit which form two semi-circular arrangements, each containing four monomers (Fig. 38B). Viewing symmetry copies reveals that the crystal is composed of octameric rings analogous to those present within our PqiC FL crystal reinforcing the biological relevance of this arrangement. Curiously the operon to which *Xanthomonas campestris* XCC0632 belongs also houses homologues of the MlaFE ABC-

transport genes in addition to an MCE domain containing protein. Such architectures have been reported previously, in the case of *Legionella pneumophila* (Nakayama and Zhang-Akiyama, 2017), implying ATP powered retrograde transport to proceed through PqiB and PqiC - a notion requiring further investigation.

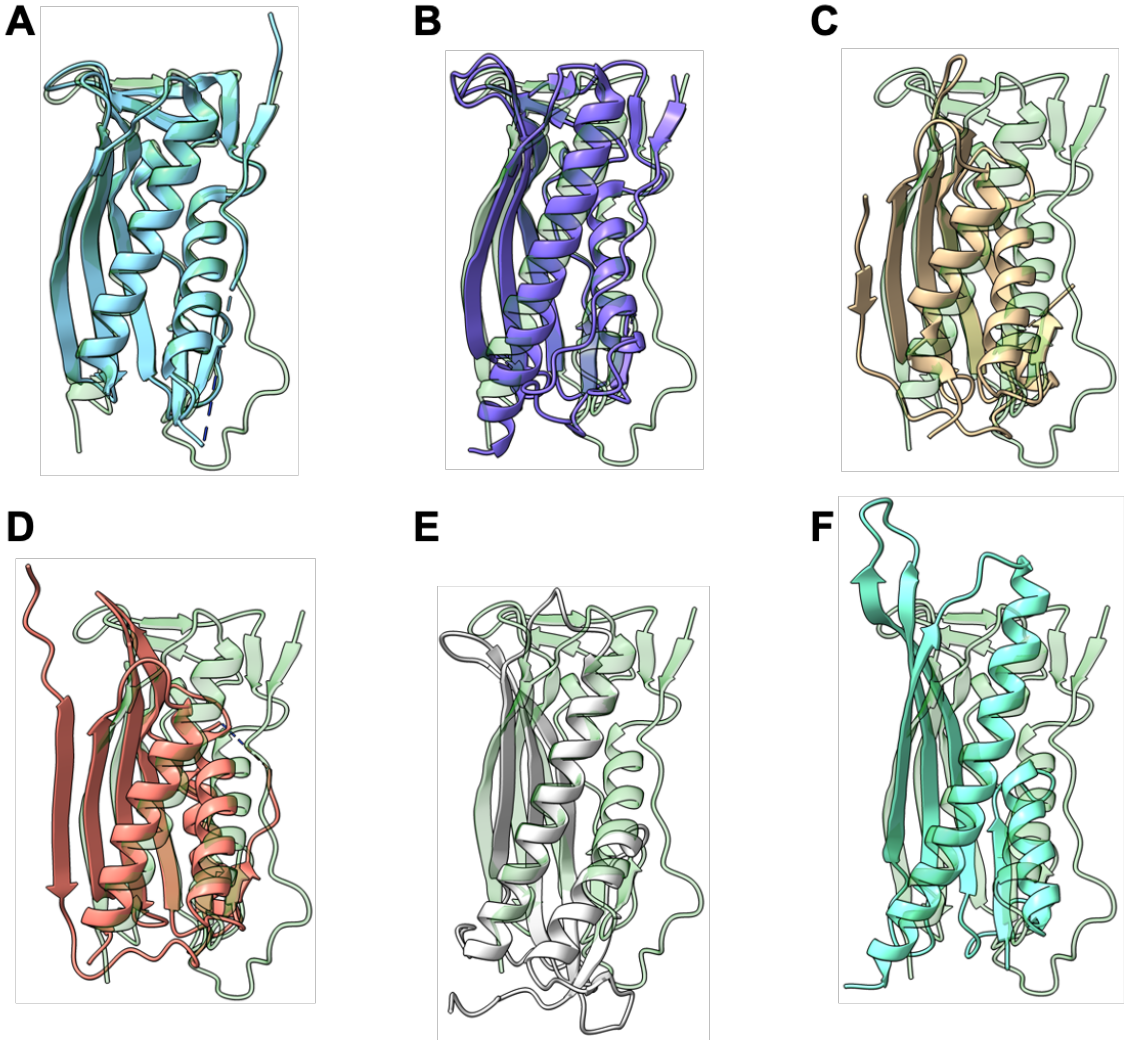


Fig. 37 – Structural Alignment of the CFL2 Structure with Dali Server Hits.

Aligned cartoon representations of CFL2 chain B (transparent green) and single chains from; (A) *Enterobacter cloacae* PqiC (PDB: 6OSX), (B) *Xanthomonas campestris* XCC0632 (PDB: 2IQI), (C) *Paraburkholderia phytofirmans* PelC (PDB: 5T10), (D) *Geobacter metallireducens* PelC (PDB: 5T0Z), (E) *Acinetobacter baumannii* LptE (PDB: 5TSE) and (F) *Shewanella oneidensis* LptE (PDB: 2R76).

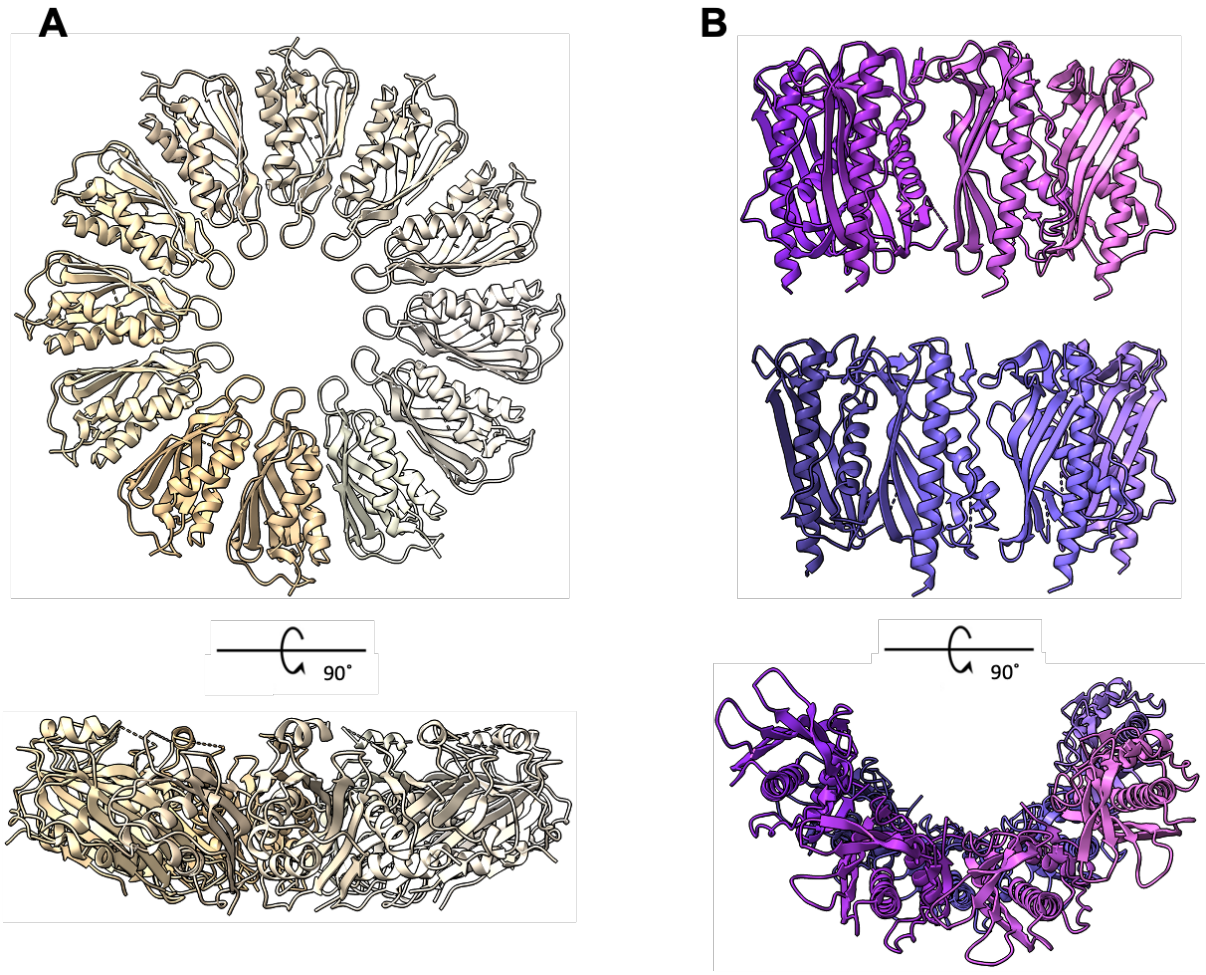


Fig. 38 – Asymmetric Unit Arrangements of PqiC Homologous

(A) Cartoon representation of the anticipated *Paraburkholderia phytofirmans* PelC dodecameric ring biological assembly. PDB: 5T10. (B) Cartoon representation of the *Xanthomonas campestris* XCC0632 asymmetric unit, encompassing two stacked semi-circular arrangements, which form octameric rings throughout the crystal. PDB: 2IQI.

A major area of uncertainty remains the nature of the anticipated interaction between PqiB and PqiC. The corresponding diameters of the PqiB coiled-coil and octameric PqiC pore suggest these to be the likely interfaces, however, the differing symmetries of these components raises questions. PqiB exhibits the canonical hexameric MCE protein architecture, however, we propose PqiC to be octameric in nature. Interactions between MCE proteins and components of differing symmetries have been well documented. In *E. coli* hexameric MlaD interacts predominantly with a dimeric MlaE species within the $MlaF_2E_2D_6B_2$ complex. These interactions are, however, perhaps less intimate than those anticipated between PqiB and PqiC, which must remain a principal area of interest in future studies.

Finally, we probed the lipophilicity and electrostatic potential of the PqiC octamer surface. We identified discrete rings of negatively charged residues interspersed by regions of lipophilic residues within the PqiC octamer pore. We hypothesised these regions to be critical for the interaction of PqiC with the PqiB coiled coil.

In addition to the pore lining face, the membrane occluded face of the PqiC Octamer displayed an intriguing spatial distribution of lipophilic and charged residues, occurring in three concentric rings. The central of these rings housed lipophilic residues which could conceivably interact with the hydrophobic components of the outer membrane. Indeed, molecular dynamics simulations suggest that PqiC may interact more substantially with the membrane, hanging vertically form its surface, rather than simply being anchored in its vicinity via acylation (Rao et al., 2020). The remaining inner and outer concentric rings display positive and negative charge respectively. The presence of such electrostatic potential on an apparently inaccessible protein face is puzzling. Do these residues interact with polar lipid head groups to strengthen the interaction between protein and membrane? Indeed, there remains the possibility of additional, yet undiscovered, components of this pathway that interact with this face of PqiC.

Furthermore, the molecular dynamics simulations revealed a second, less prominent pose in which PqiC monomers lie laterally on the membrane surface (Rao et al., 2020). This arrangement positions PqiC monomers in a similar context to those observed in the dodecameric *Paraburkholderia phytofirmans* PelC structure, raising the possibility of a higher order biological assembly. Indeed, the interface between a hexameric PqiB and dodecameric PqiC may appear more feasible than the hexameric – octameric interaction we currently anticipate.

The structure of PqiC FL provides an insight into the protein's potential biological assembly and, alongside conservation analysis, provides tantalising clues to residues crucial for complex formation and function which must be explored further. The proposed octameric state of PqiC may be validated by the use of native mass spectrometry, SEC-MALS and cryo-EM however these experiments were beyond the scope of the current project.

PqiC$^{17-187}$

We also present the structure of a soluble PqiC$^{17-187}$ construct referred to as Csol. The Csol structure produced a dramatically different asymmetric unit with three copies opposed to the two present within the CFL structures. The orientation of Csol monomers was striking with chains A and B adopting somewhat similar positions to those in the CFL structure but being lifted and rotated respective to each other rather than sitting completely parallel. We propose this difference in chain position results directly from the soluble or membrane/detergent associated nature of the two constructs. As a lipoprotein, native PqiC sits tethered to the membrane with the possibility of interacting directly with the membrane surface itself, a notion supported by molecular dynamics simulations (Rao et al., 2020). This tethered condition is presumably somewhat replicated during purification of the CFL construct, with the membrane plane being replaced by a detergent micelle. The movement of a single PqiC FL monomer is severely restricted as its translation vertically is limited by the presence of the detergent micelle above and the tether length whilst lateral translation is restricted by the additional monomers present within the octameric assembly and the interactions stabilising this quaternary state. Whilst we anticipate the PqiC$^{17-187}$ construct to adopt the same octameric assembly during purification, owing to the similarity of its SEC trace to that of PqiC FL, there are fewer factors enforcing the maintenance of this conformation during the crystallisation process. Indeed, both factors restricting vertical translation in the PqiC FL construct, micelle presence and tether length, are absent in the case of the PqiC$^{17-187}$. We propose that the PqiC$^{17-187}$ monomers adopt a conformation more favourable for crystal packing thus, the molecules within the unit cell no longer assemble an octameric structure but instead a helical arrangement with chains A and B forming a corkscrew with chain C extending radially from this core (Fig. 39).

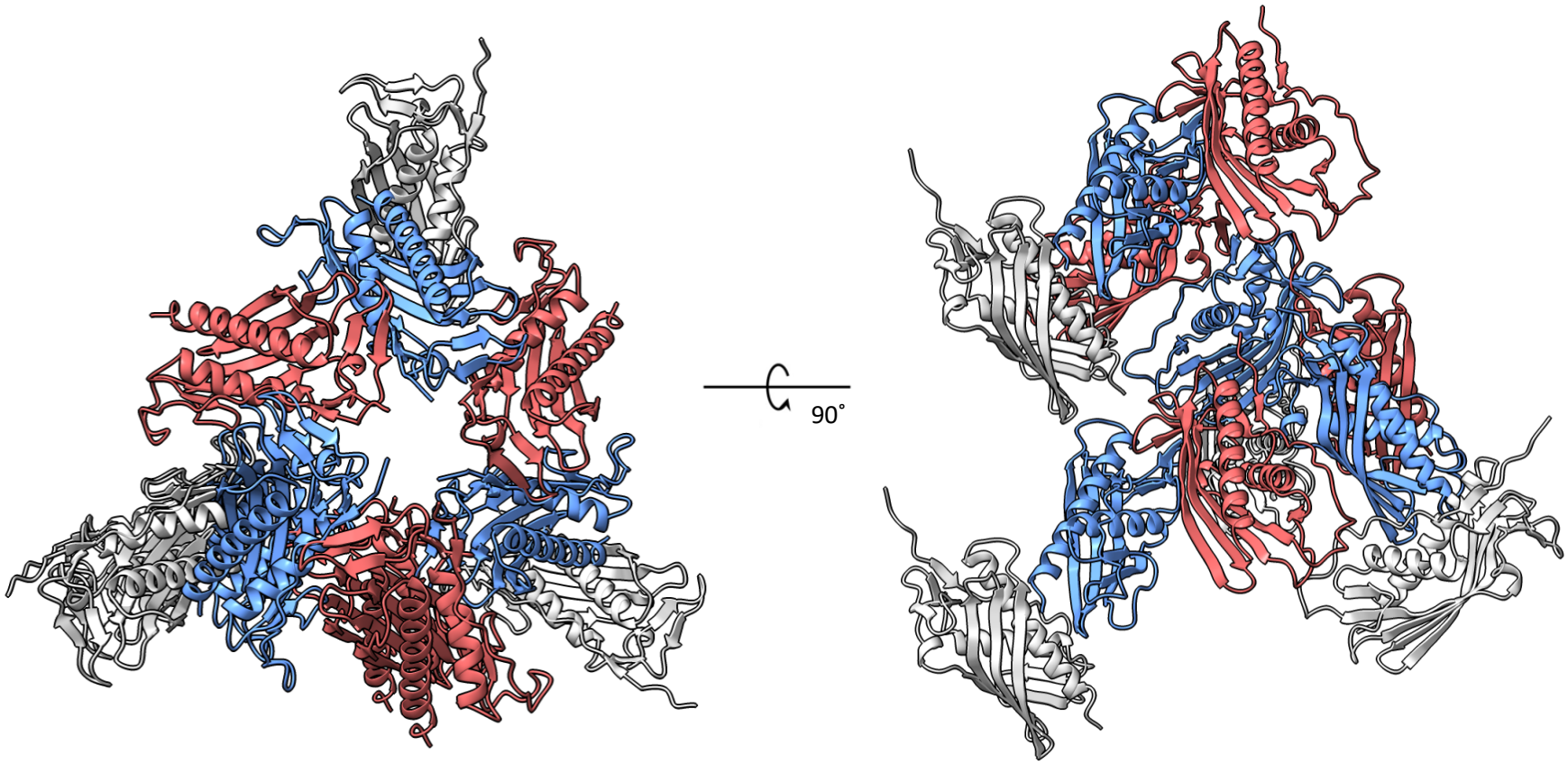


Fig. 39 – Crystal Packing of PqiC$^{17-187}$.

Orthogonal views of the helical PqiC$^{17-187}$ crystal packing arrangement. Copies of the Csol structure are depicted in cartoon representation with chains A, B and C coloured in blue, red and grey respectively.

Finally, it was anticipated that the higher resolution of the CsoI structure may provide improved electron density for the β1-β2 loop, however, this proved not to be the case. Whilst the density corresponding to residues at either end of the loop enabled modelling of side chains, the density within the middle of the loop was worse than that observed within the CFL2 electron density map. We attribute the difference in β1-β2 loop density between the CFL2 and CsoI structures to the relative confinement of the loop in both structures (Fig. 40). In the CFL2 structure, the β1-β2 loop is packed tightly between adjacent PqiC FL monomers within the octameric assembly meaning, that despite its apparent disorder, the number of potential orientations is limited by the physical space available (Fig. 40A). In contrast, the relative angle between $PqiC^{17-187}$ monomers in the CsoI structure mean that the β1-β2 loop is not constrained to the same degree. Indeed, the structure displays a large void between each monomer in which the loop may reside hence there is the potential for many more loop conformations within the CsoI structure which ultimately yield poor or absent electron density (Fig. 40B).

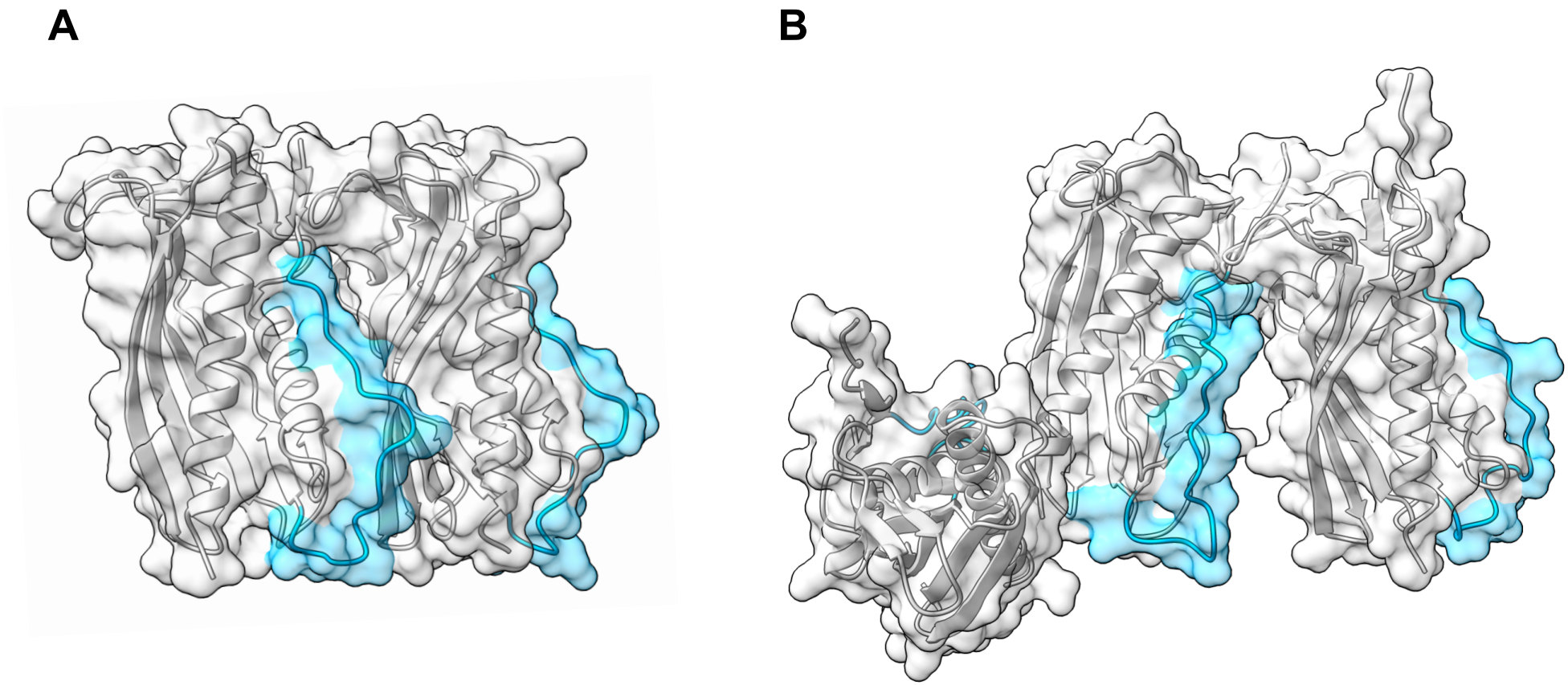


Fig. 40 – Packing of the β1- β2 Loop in the CFL2 and Csol Structures.

Cartoon representations overlayed with transparent surfaces of the (A) CFL2 and (B) Csol structures. The β1-β2 loop is coloured blue indicating its relative confinement within the two structures.

Summary

In this chapter we have presented the structures of PqiC FL and a soluble PqiC$^{17-187}$ construct which reveal high-resolution side chain information as well as a potentially biologically relevant octameric conformation. Probing the surface characteristics of the octameric assembly, in combination with sequence conservation analysis, has identified several highly conserved residues we believe to be involved in Pqi complex formation and/or transport mechanics. Consequently, the PqiC structures discussed within this chapter have informed the design of mutant Pqi operons to be assessed via complementation assays, in Chapter 5, to infer the importance of specific PqiC residues in a biological context.

Chapter 4: Cryogenic Electron Microscopy of PqiABC

Introduction

Introduction to Cryo-EM

Single particle Cryogenic Electron microscopy, hereafter referred to as cryo-EM, is a structural technique only recently widely adopted by the scientific community, however, today its usage is vast, exemplified by the near exponential growth of yearly deposited cryo-EM derived PDB structures (Fig. 41).

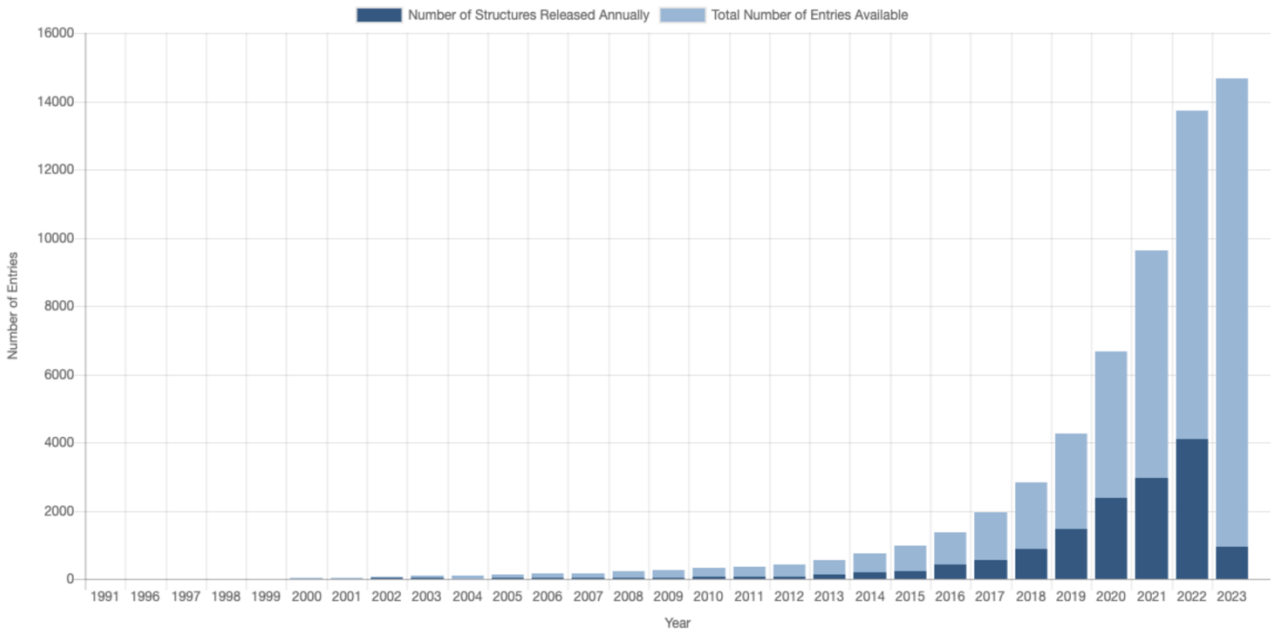


Fig. 41– Growth of 3D Electron Microscopy derived structures in the PDB.

Bar plot indicating the number of 3D Electron Microscopy derived structured deposited to the PDB annually as well as the cumulative total number structures. (Figure: https://www.rcsb.org/stats/growth/growth-em)

Despite clear, early demonstration of the technique's potential with the first 3D reconstruction of the bacteriophage T4 tail (De Rosier and Klug, 1968), the widespread application of cryo-EM was plagued by the low-resolution reconstructions it produced (typically ~ 10 Å) earning it

the rather demeaning title of “blobology”. The breakthrough, coined the “resolution revolution”, came recently in 2012-13, primarily attributed to the availability of direct electron detectors and advanced image processing software (Scheres, 2012a, 2012b) which enabled cryo-EM to emerge as a serious structural determination contender. Indeed, the ability to reconstruct atomic resolution cryo-EM maps has now been demonstrated by several independent groups (Nakane et al., 2020; Yip et al., 2020). Despite conceding in other areas (particularly the requirement of the studied macromolecule to be ~>50 kDa as well as the time and computation requirements of processing), modern cryo-EM displays several advantages over other structural determination techniques including low sample requirement, the ability to capture protein dynamics and structural determination under more native conditions.

The entire cryo-EM method from protein sample to 3D reconstruction can be summarised in three steps: 1- sample preparation, 2- image formation and 3- data processing, discussed in more detail below.

Sample Preparation

As with all structural techniques, the main requirement for cryo-EM is a homogenous sample containing the biomolecule of interest. It should be noted that whilst homogeneity is desirable, the picking and classification phases of single particle processing allow avoidance of contamination to a degree, although these should not be considered as a final purification steps. Cryo-EM sample preparation involves the rapid vitrification of the sample, containing the biomolecule of interest, upon an EM grid within a thin layer of vitreous ice, with the aim of suspending the particles in an array of random orientations. EM grids comprise a 3 mm metal, mesh disk which are overlayed with a perforated carbon foil. These canonical grids may be functionalised with additional support films, including graphene oxide (Pantelic et al.,

2010), amorphous carbon (Williams and Glaeser, 1972), graphene (Liu et al., 2019) and Ni-NTA lipid monolayers (Kelly et al., 2008), to which particles adsorb, effectively concentrating them upon the grid surface.

Typically in semi-automated approaches, 3 µl of sample is applied to the face of a freshly glow discharged EM grid within a humidity-controlled chamber. Excess sample is removed via blotting with filter paper before the grid descends rapidly into the cryogen, usually liquid ethane cooled by liquid nitrogen (Fig. 42). The vitrification process is necessary for the sample to survive the vacuum present within the microscope column whilst also mitigating radiation damage to a lesser degree.

Although the canonical approach, this semi-automated vitrification process suffers several challenges. An inherent consequence of the process is that the majority of sample is wasted, with less than 0.1% of the applied volume remaining on the grid following the blotting step (Rubinstein et al., 2019). Indeed, the use of blotting paper has been shown to result in both protein aggregation and denaturation (Arnold et al., 2017; Glaeser, 2016). Furthermore, the thickness and distribution of the sample layer upon the grid surface is affected by many factors (support, sample properties, environmental conditions etc) meaning that identifying optimal conditions can become an time consuming, trial and error laden process (Passmore and Russo, 2016). Finally, the dexterity required for the handling and transfer of grids necessitates the extensive training of skilled users. To mitigate against the aforementioned sample waste, several alternative application methods have been developed (spraying (Ashtiani et al., 2018; Feng et al., 2017), inkjet dispensing (Dandey et al., 2018; Jain et al., 2012), microcapillary writing (Arnold et al., 2017) and contact pin printing (Ravelli et al., 2020)), all of which reduce the amount of sample applied to the grid and negate the need to remove excess liquid from the grid following application. Where excess liquid is applied, “self-wicking” grids consisting of nanowires maty be utilised to yield a thin film within the grid holes (Wei et al., 2018). Such grids are utilised within the Spotiton system which provides a

completely automated glow discharge, inkjet sample application and rapid vitrification pipeline with the aim of increasing the reproducibility (Dandey et al., 2018; Jain et al., 2012).

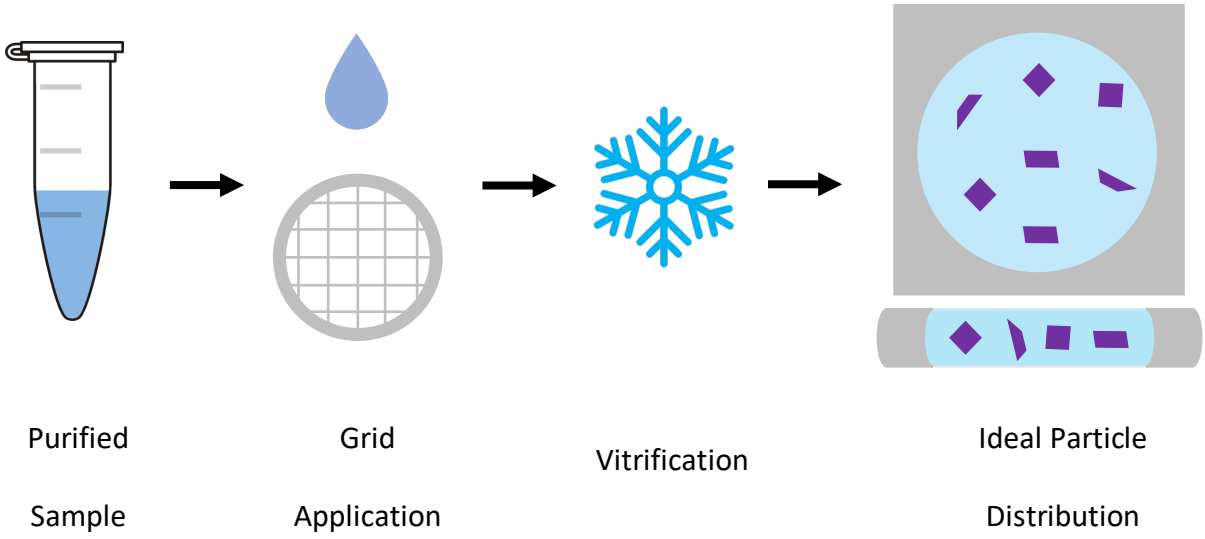


Fig. 42 – Typical cryo-EM sample preparation.

Flow diagram indicating the stages of cryo-EM sample preparation from a purified proteinaceous sample to produce a vitrified grid.

Despite the relative simplicity of this process, particularly considering the development of more automated vitrification systems, sample preparation presents the primary bottleneck within the entire cryo-EM pipeline. Typically, proteins exhibit aggregation, preferential orientation or non-uniform distribution hence screening of various concentrations, buffer constituents, grid types and support films (to name but a few) is required before a condition yielding a uniform distribution of randomly oriented particles is identified.

Image Formation

Microscope Anatomy

Once vitrified samples have been prepared, they are inserted into an electron microscope to allow visualisation of the protein of interest. A typical electron microscope comprises several essential components, an electron source, a series of electromagnetic lenses, apertures and an image detection system, maintained within a high vacuum column (Fig. 43).

Electrons emitted from the electron source are directed down the microscope column with their position guided by a complex series of electromagnetic lenses. Two types of electron course are commonly used: thermionic (tungsten filament, LaB_6 crystal) and field emission (tungsten tip). The later of these is preferred for high-resolution imaging applications owed to the generation of beams with much greater spatial (electrons emanating from the same point) and temporal (electrons emanating with the same wavelength) coherence allowing the formation of superior images. Additional apertures limit the size of the electron beam at various points within the column. Electrons are first focussed onto the sample before being directed to a screen or camera detection system, residing at the bottom of the column, allowing visualisation of the sample. Maintaining the entire microscope column under a high vacuum prevents the unwanted interaction of electrons with air which generates noise in the final images.

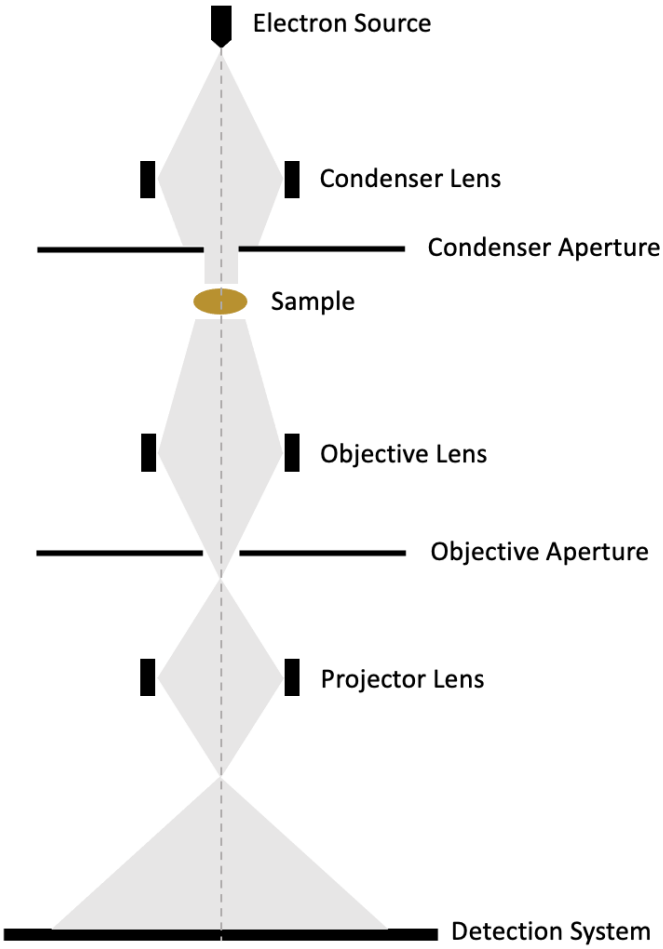


Fig. 43 – Transmission Electron Microscope Anatomy.

Schematic depicting the main components of a transmission electron microscope. The central axis of the microscope column and electron beam are indicated by the dashed line and grey area respectively.

Contrast and Defocus

When incident electrons reach the sample, three primary scenarios occur. 1- most electrons are unscattered and continue directly to the detector. 2 – ~ 5% of electrons are elastically scattered by the sample meaning they retain their incident energy - these are useful for image formation. 3 – ~ 15% of electrons are inelastically scattered resulting in a deposition of energy to the sample which may induce radiation damage. The relative frequency of these useful elastic and harmful inelastic scattering events is one of the primary limitations when imaging biological samples via cryo-EM (Henderson, 1995; Langmore and Smith, 1992).

The majority of contrast in the cryo-EM image is attributed to phase contrast (~90%), arising from the phase discrepancy between unscattered and elastically scattered electrons at the image plane (Bhella, 2019). Amplitude contrast, arising from the removal of electrons scattered inelastically or at high angles by apertures or energy filters, usually provides low resolution image information and is therefore largely discounted.

Proteins themselves are comprised of atoms possessing low scattering potential thus images taken at absolute focus contain very little contrast rendering the proteins themselves almost invisible. To counteract this issue, defocussed images are instead collected, increasing the phase shift of elastically scattered electrons and generating contrast accordingly. At a particular defocus value, electrons scattered to different angles contribute differently to the contrast observed in the final image. While some angles contribute positively, others do so negatively or not at all. The contrast transfer function (CTF) is an oscillating mathematical function, of increasing frequency and decreasing amplitude, which describes the contribution of each spatial frequency to the contrast observed in the final image (Erickson et al., 1997; Wade, 1992) (Fig. 44A).

Higher defocus increases image contrast but inflicts costs in terms of 1 – increased CTF fluctuations (in turn increasing the number of spatial frequencies not contributing to image contrast) and 2 – increased image delocalisation. These effects may be compensated for during computational processing. Finally, the loss of information resulting from the CTF necessitates datasets be collected to contain images possessing a range of defoci values which are subsequently utilised to fully restore all image information (Orlova and Saibil, 2011) (Fig. 44B).

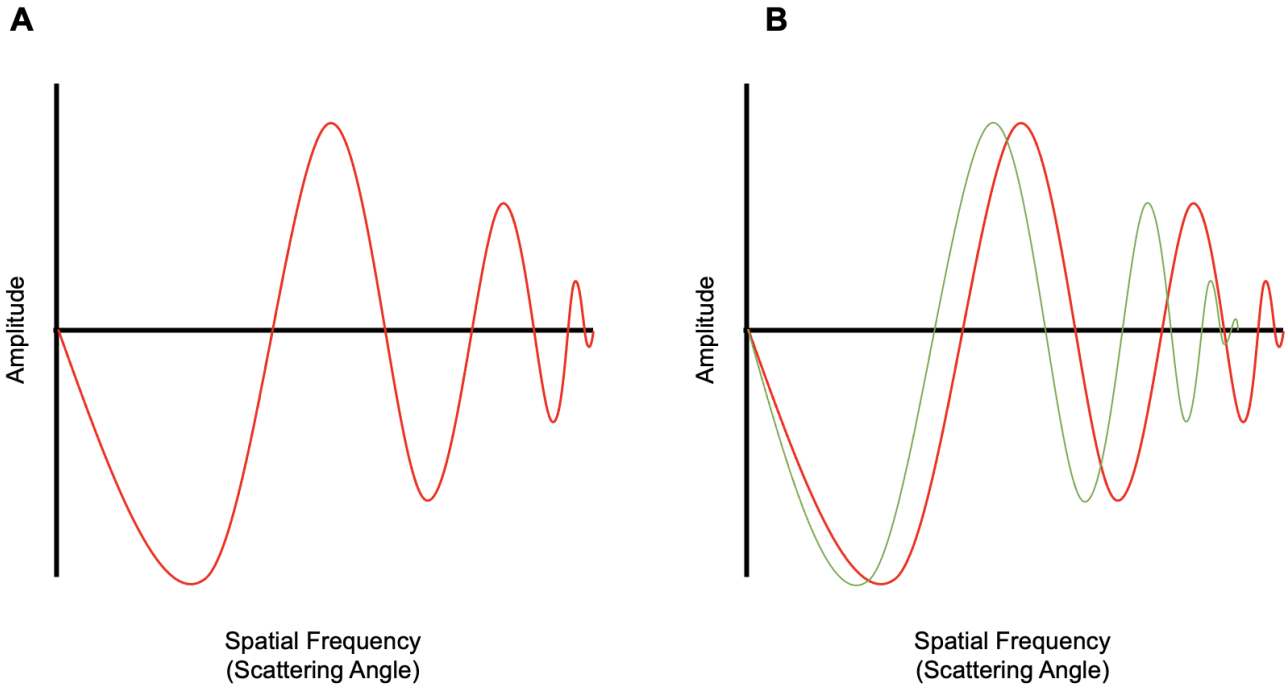


Fig. 44 – Contrast Transfer Function (CTF) Curves.

(A) CTF curve for a single defocus value showing the oscillating function which decreases in amplitude with spatial frequency. Spatial frequencies above and below the 0 contribute positively and negatively to image contrast respectively (B) CTF curves for two differing defoci. The green curve represents a CTF curve at higher defocus, oscillating more frequently and shifted to the left.

Data Processing – Single Particle Averaging

The microscope records a series of movies composed of individual frames which are termed micrographs, each capturing multiple instances of the particle of interest in an array of orientations. The aim of data processing is to convert these micrographs into a 3D structure of the particle of interest, typically performed via a single particle analysis (SPA) pipeline. Whilst specific projects may require additional specialist steps, the canonical SPA pipeline comprises 9 steps to progress from raw micrographs to protein structure, discussed below, (Fig. 45).

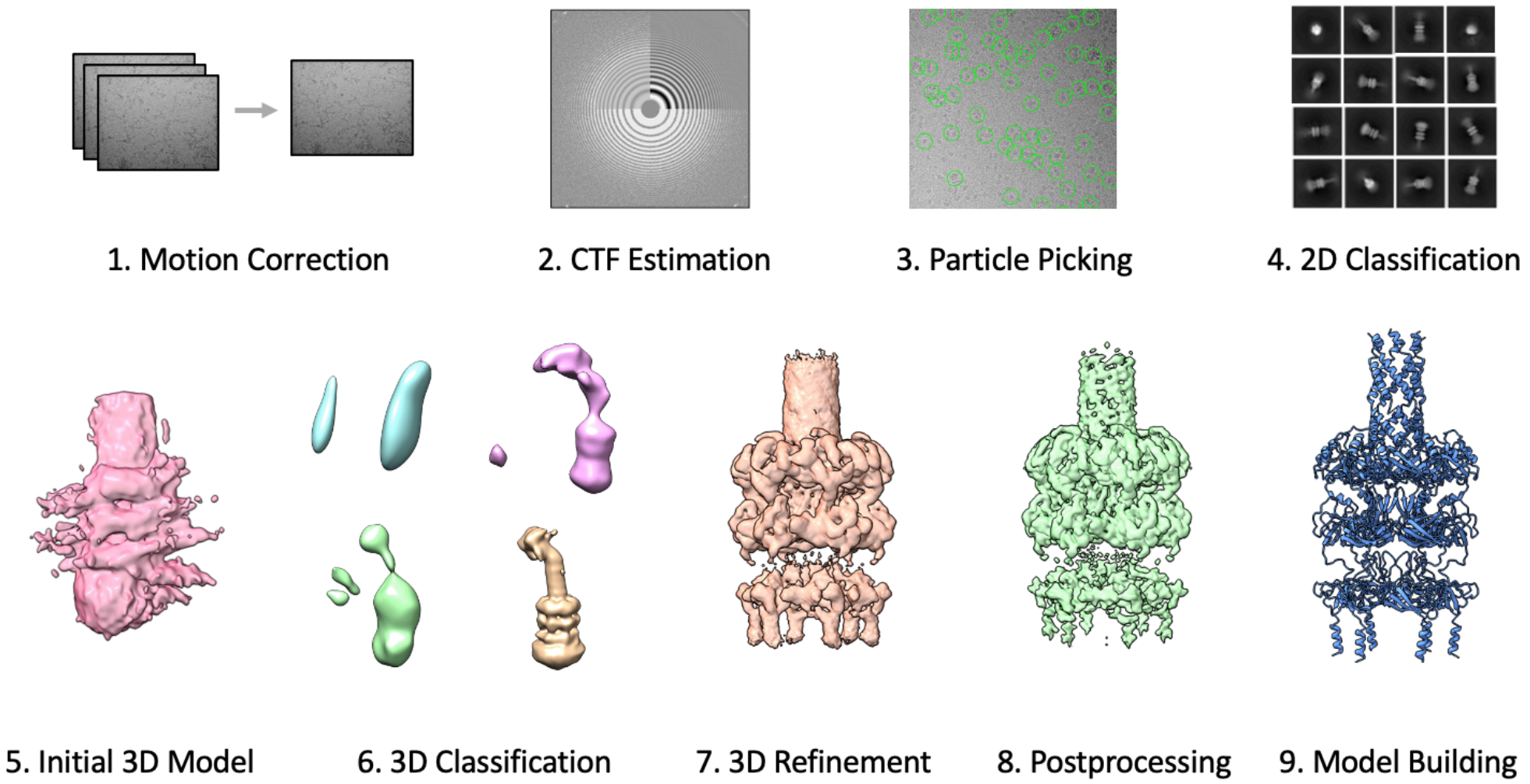


Fig. 45 – The canonical single particle averaging (SPA) pipeline.

Pictorial representation of the 9 steps in the canonical SPA pipeline converting raw micrographs into a structure of the protein of interest.

Motion Correction

During data collection the sample experiences motion for various reasons including drifting of the stage upon which the sample resides, and deformations induced by the electron beam. Several motion correction algorithms exist, executing variations on a fundamental scheme, to align individual micrograph frames to account for this motion. Typically, micrographs are sectioned into a 5x5 grid producing 25 smaller ‘patches’ which are aligned across frames rather than the entire micrograph. This increases accuracy as motion is not consistent across the entire micrograph. The process produces a single image which is the average of all frames for further processing.

CTF Estimation

As mentioned previously, the CTF describes the contribution of each spatial frequency to the final image contrast. The 2D fourier transform of each micrograph allows visualisation of its CTF as a sinusoidal wave plotted through the observed spherical rings (Thon rings). This experimental CTF is compared to a range model CTF curves for differing defocus values allowing accurate prediction of the micrograph defocus.

The oscillating nature of the CTF means that some amplitudes are of a negative nature, thus these phases are flipped resulting in only positive values. Following phase flipping, CTF curves from different images at different defoci are merged thus compensating for the zero values in any particular CTF curve.

Particle Picking

Individual particles must be identified and extracted from the CTF corrected micrographs prior to downstream processing. This process may be performed manually, semi-automatically or via fully autonomous programmes. The simplest manual picking method involves the user identifying particles and selecting them by hand. Whilst accurate, this can be a very time-consuming process and is dependent upon the user identifying all particle orientations. More commonly, a semi-automated process is employed where the user picks a subset of particles which are classified to generate 2D classes (see below). These classes are used as references for faster autopicking programmes, however, these rely on the generation of good reference templates. The development of fully autonomous picking programmes has significantly sped up the task exemplified by the Laplacian of Gaussian (LoG) filter (Scheres, 2015). Recently, several machine learning algorithms have been developed and deployed to tackle particle picking (Bepler et al., 2019; Wagner et al., 2019). These algorithms have the distinct advantage of requiring no prior information, removing the issue of reference bias, furthermore, the algorithm may be trained on a subset of the dataset to improve results.

2D Classification

Autopicking, and to a lesser extent manual picking, results in the selection and extraction of micrograph artefacts such as aggregated protein, ice contamination and carbon surrounding the grid holes. 2D classification presents the first opportunity to remove these ‘junk’ particles whereby particles of the same view are grouped together into classes, thus increasing their signal to noise ratio. Put simply, ‘junk’ particles are distinctly different in appearance than the particle of interest and consequently accumulate in their own classes which can be discarded

from downstream steps. 2D classification also provides a means of assessing the quality of extracted particles prior to 3D reconstruction.

Initial 3D model

3D classification, see below, requires the provision of an initial 3D model to perform projection matching. Whilst a user may provide a model, low pass filtered to remove high resolution information, this process introduces bias into the processing pipeline. Consequently, a preferred method is to generate an *ab-initio* initial model from the experimental particles themselves using a stochastic gradient descent algorithm thus eliminating the introduction of model bias.

3D Classification

The 2D classes obtained from 2D classification are in fact 2D projections of the 3D object itself. Furthermore, projection slice theorem states that the fourier transform of a 2D class, and its comprising particles, is equivalent to a central slice of the 3D objects own fourier transform. Consequently, if the angles relating the 2D projections to the 3D object are known, or can be accurately estimated, the projections may be back projected to reconstruct the 3D object.

To estimate these angles, projections are generated from the provided initial model and compared to each particle. The particle is then assigned the angle from the simulated projection it most closely resembles in a process called projection matching. The process is repeated iteratively as back projection of the particles causes the initial model to evolve,

more closely resembling the 3D object and consequently permitting more accurate assignment of angles.

As the initial model iteratively improves, it becomes possible to identify heterogeneity within the particle population as they are back projected. These particles are assigned to different 3D classes, generating distinct 3D models, allowing the identification of distinct conformations.

3D Refinement

Once 3D classification has accurately assigned the projection angles of all particles and thus generated an optimal 3D model, this model will be refined. Refinement is also an iterative process in which 2D projections are generated via forward projection of the 3D model. These are subsequently projection matched to the extracted particles, which assume the optimised projection angle prior to back projection thus generating an improved model to restart the process.

During refinement, the particles are split into two independently refined subsets allowing accurate determination of the final map resolution via fourier shell correlation (FSC). The FSC describes the correlation between the two independently refined half maps which decreases from 1 to 0 with increasing resolution due to diminishing signal to noise ratio. Resolution is typically reported at the resolution the FSC falls below 0.143, referred to as gold standard FSC, as this is thought to yield resolutions comparable to those obtained from X-ray crystallography.

Following refinement, additional processing steps to further account for microscope aberrations and particle heterogeneity/flexibility may be deployed before repeating the

refinement step to improve map resolution. These steps are not considered part of the canonical SPA pipeline and as such are not discussed here.

Post processing

A refined map will be further postprocessed or sharpened to restore high resolution features which may be absent due to characteristics of the data collection (radiation damage, imaging aberrations) or processing pipeline (incorrect angle assignment). The deterioration of information is first predicted allowing subsequent up-weighting and restoration of high resolution features within the final map.

Model Building

Once a final post processed map has been achieved, it may be utilised to build a model of the protein structure. The model building process itself is performed differently depending upon the resolution of the final map and the availability of prior structural information. For the lowest resolution maps, it may only be possible to fit existing experimental or predicted structures into the density assuming their availability, however, caution must be maintained to prevent over interpretation in these scenarios, particularly when relying on predicted structures. Maps displaying secondary structural elements allow the more definitive positioning and manipulation of existing/predicted structures. Programmes such as ISOLDE utilise molecular dynamics to aid in the fitting of models to low resolution density maps, however, once again over interpretation must be avoided. The highest resolution models may be built *de novo* utilising adapted versions software initially developed for X-ray crystallography including COOT and Phenix.

Chapter Aims

Cryo-EM has been deployed by several lab groups investigating the structure of *E. coli* MCE proteins. Whilst this has yielded numerous structures of the entire MlaFEDB complex, thus far only isolated soluble constructs of YebT and PqiB have been reported, hence the interface between these MCE proteins and additional operon components remains elusive. Furthermore, structural variations between these soluble constructs compared to the native membrane associated proteins remain a possibility. Of the existing MCE protein structures, that of a soluble PqiB$^{39\text{-}546}$ construct remains the least well defined with ambiguous density present within the central pore and diminishing density, likely due to flexibility, preventing accurate modelling of the entire coiled-coil moiety (Ekiert et al., 2017). We hypothesised that the entire PqiABC complex may provide a more stable sample for cryo-EM whilst also enabling us to probe stoichiometry and interactions between all three complex components. The chapter discusses our progress in attempting to solve the entire PqiABC complex via cryo-EM. Whilst a high-resolution structure of the entire complex remains to be determined, we successfully solved and modelled the structure of native PqiB allowing an interesting comparison to the existing soluble structure.

Results

Expression and Purification of PqiABC

The entire PqiABC operon was expressed from the IPTG inducible pET-26b(+) plasmid transformed into *E. coli* C43 (DE3) cells with induction proceeding overnight at 18 ˚C. The isolated membrane fraction was solubilised in *n*-dodecyl-β-D-maltoside (DDM) before IMAC was utilised to isolate PqiABC. IMAC flow-through, wash and elution fractions were analysed via SDS-PAGE (Appendix. 5). Elution fractions containing all three proteins were pooled and further purified by a Superose 6 10 300 Increase size exclusion column (Fig. 46A). The composition of elution fractions was determined via SDS-PAGE. Bands at ~35 kDa, ~60 kDa and ~20 kDa indicating the presence of PqiA, PqiB and PqiC respectively, with PqiA running slightly smaller than anticipated considering its molecular mass of 46 kDa. (Fig. 46B).

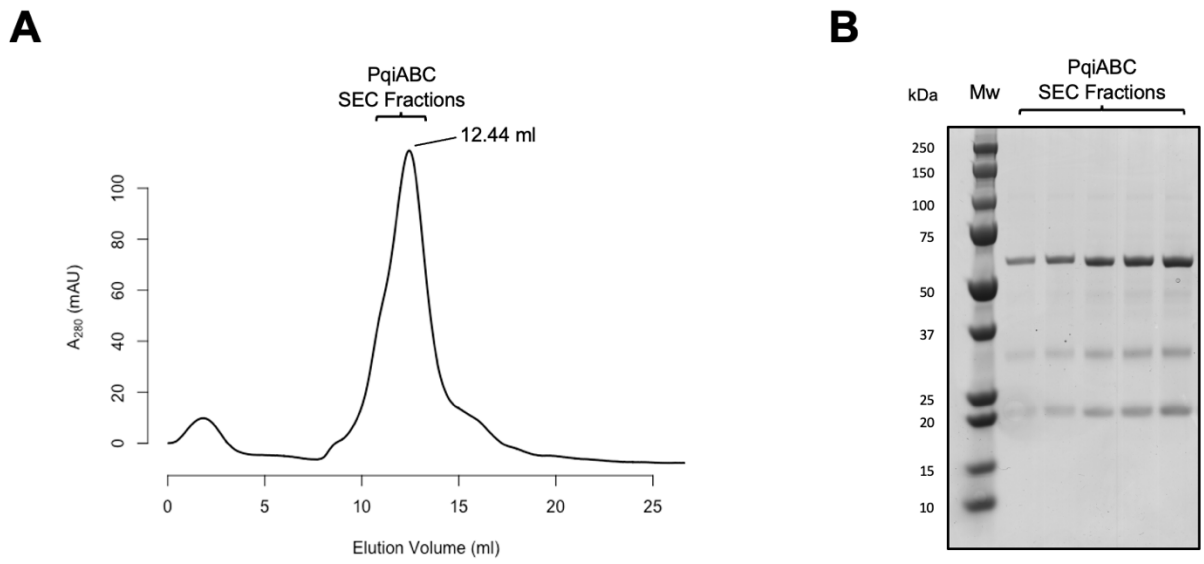


Fig. 46 – Purification of PqiABC.

(A) SEC trace of showing a single elution peak, between 10 and 14 ml, after the void peak from a Superose 6 10 300 Increase column. (B) SDS-PAGE of PqiABC SEC fractions showing three predominant bands at ~35 kDa, ~60 kDa and ~20 kDa indicating the presence of PqiA, PqiB and PqiC respectively.

Initial cryo-EM Screening of PqiABC

Having determined that we could express and purify the entire PqiABC operon as a stable complex we progressed to making and screening cryo-EM grids. Unfortunately, the complex displayed a high affinity for the carbon support film, adhering to its surface and excluding itself almost entirely from the ice containing grid holes (Fig. 47). Attempts to overcome this issue by further increasing sample concentration, to fully saturate the carbon surface and force particles into the grid holes, also proved unsuccessful (data not shown).

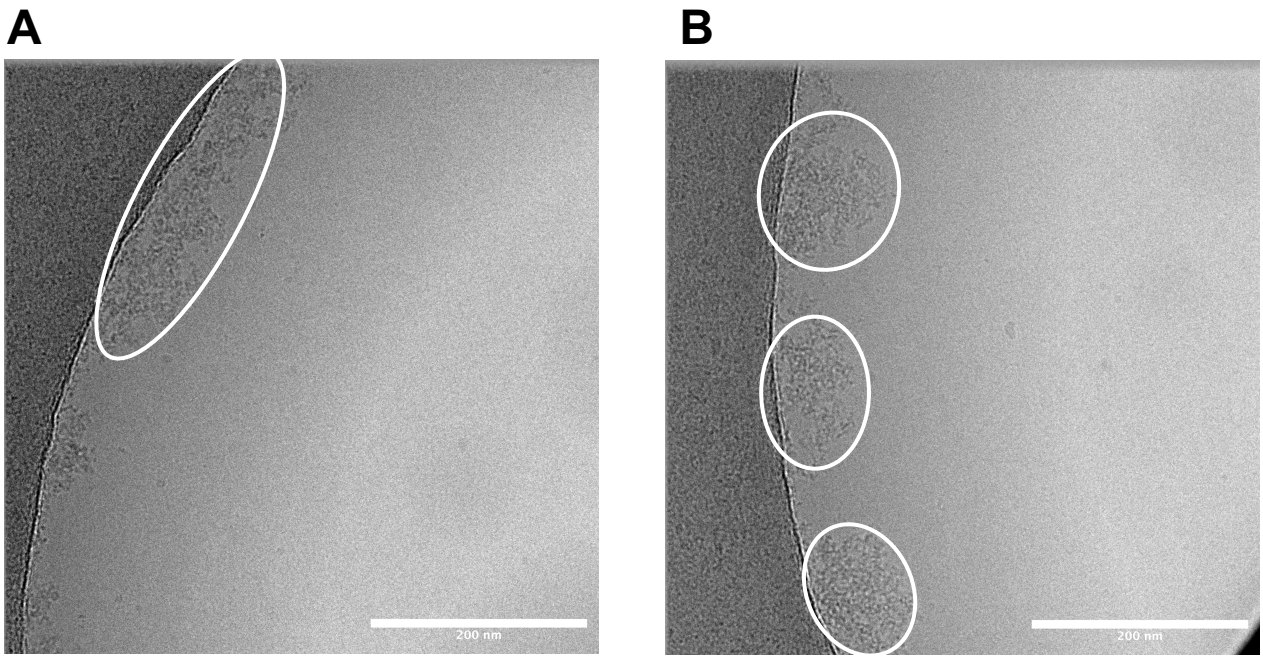


Fig. 47 – Initial screening of PqiABC.

The particles can be seen clustered around periphery of the grid holes adhered to the carbon support film of a Quantifoil R1.2/1.3 Cu grid (darker areas circles in white).

Next, we attempted to exploit the apparent high affinity of PqiABC for carbon by applying the sample to grids prepared with an additional graphene oxide support film across the entire grid surface according to (Bokori-Brown et al., 2016). We hypothesised that the PqiABC complex would adhere to the graphene oxide layer, thus forcing particles into the grid holes where they could be imaged effectively.

The addition of the graphene oxide support film was successful in forcing particles into the grid holes, enabling the collection of a small dataset of micrographs. Unfortunately, during processing it became clear that the micrographs presented two major issues. Firstly, the

micrographs possessed a high degree of background noise and what appeared to be ice contamination making identification of the particle extremely difficult (Fig. 48A). Secondly, we observed extreme preferential orientation of particles, with every PqiABC particle adhered to the graphene oxide layer in the same longitudinal fashion (Fig. 48A). The presence of only a single orientation, despite the presence of symmetry, makes reconstitution of a 3D object an impossible task, thus this dataset yielded no more than a single 2D class (Fig. 48B).

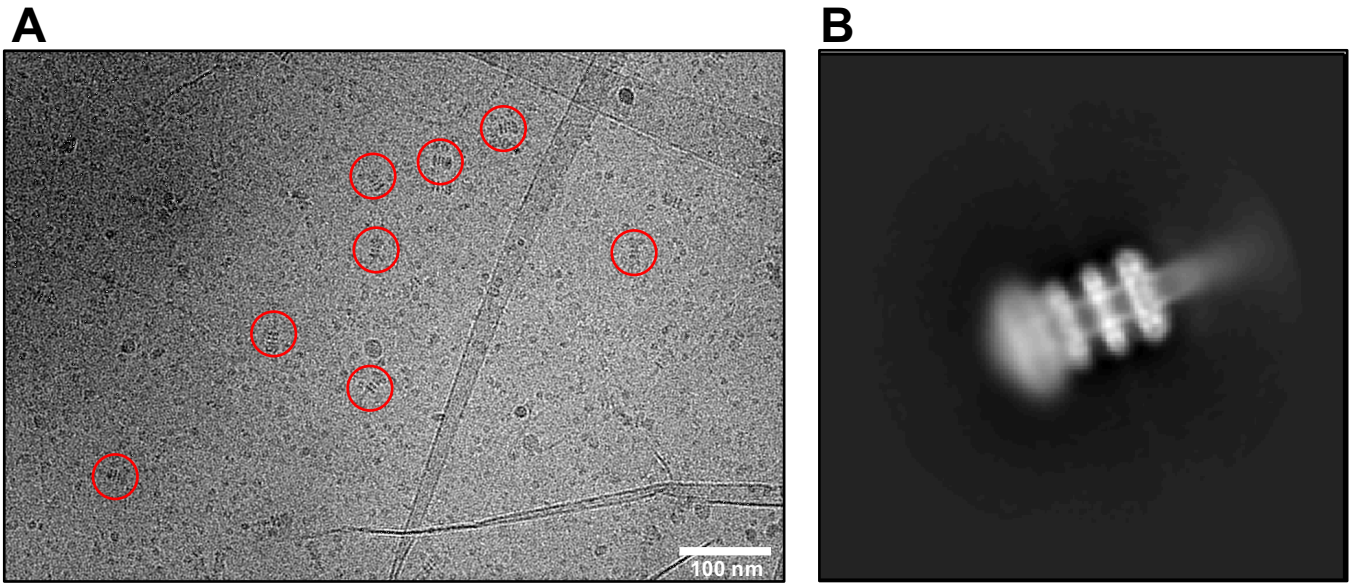


Fig. 48 – Initial screening of PqiABC on graphene oxide coated grids.

(A) Representative micrograph from the PqiABC graphene oxide coated Quantifoil R1.2/1.3 Cu grid dataset demonstrating the high background and preferential orientation issues. Red circles indicate the location of particles. (B) 2D class indicating the single orientation of the PqiABC sample observed in the graphene oxide dataset.

ThermoFisher Collaboration

Grid Optimisation

Following these initial screening sessions, the opportunity to collaborate with the ThermoFisher Centre for Electron Microscopy presented itself as an ideal opportunity to screen a large array of vitrification conditions with the aim of collecting an optimised dataset. Through an extended grid screening session, performed by Dr Deniz Ugurlar at the ThermoFisher Centre for Electron Microscopy, the grid type, sample concentration, blotting parameters and addition of supplementary detergents were optimised resulting in the preparation of a final grid for data collection (Table. 15).

Table. 15 – Optimised grid preparation parameters for PqiABC.

Parameter	Optimised Value
Sample concentration	0.75 mg/ml
Grid Type	Quantifoil R1.2/1.3 Cu
Blot Force	0
Blot Time	5.5 sec
Supplementary detergents	0.07% FOM

Data Collection and Preliminary Processing

Following the screening session, a large dataset of the optimised grid was collected utilising the below parameters, (Table. 16). Collected micrographs were processed on the fly to allow assessment of data quality via inspection of 2D classes, (Fig. 49A). Considering the encouraging 2D classes, a preliminary 3D map was generated and refined by Dr Deniz Ugurlar further indicating the potential of the dataset, however, the PqiC component appeared absent from the complex, (Fig. 49B). All processing at the ThermoFisher Centre for Electron Microscopy was performed utilising the CryoSparc processing pipeline.

Table. 16 – PqiABC ThermoFisher dataset collection parameters.

Collection Parameter	Value
Accelerating Voltage	300 kV
Magnification	130,000 x
Pixel Size	0.96 Å/pixel
Dose Rate	6.15 e/pix/sec
Total Dose	40 e/ $Å^2$
Dose Fractions	204*7 frames (EER)
Exposure Time	5.94 sec
Number of images	8163
Defocus Values	-0.75 to -1.75 with 0.25 increments
Slit Width	10 eV

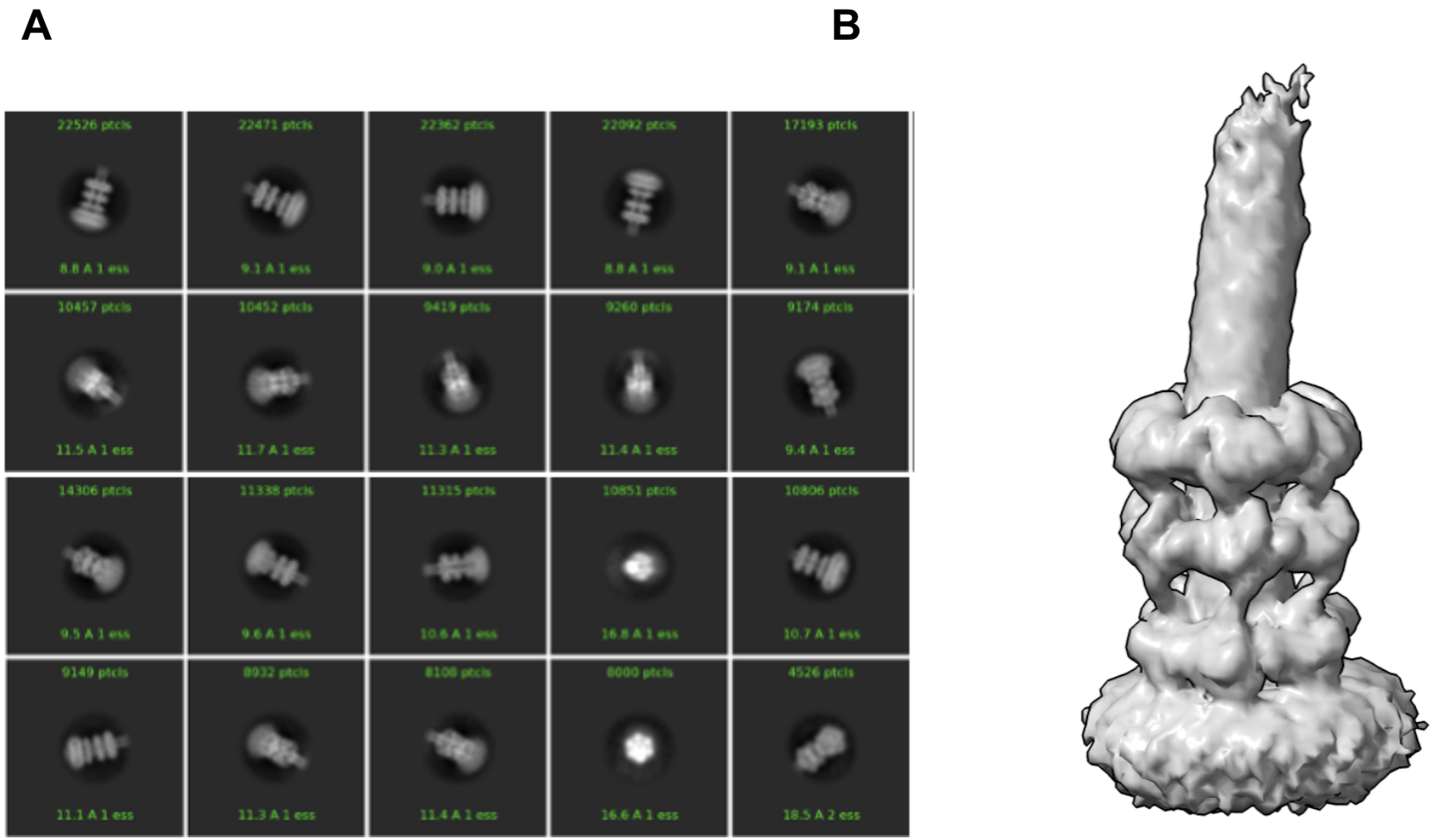


Fig. 49 – PqiABC ThermoFisher preliminary processing.

(A) On the fly 2D classification performed during the PqiABC ThermoFisher data collection indicating the particle orientations present within the dataset. (B) Preliminary refined map produced by Deniz Ugurlar at the ThermoFisher Centre for Electron Microscopy indicating the dataset potential.

Processing of the ThermoFisher PqiABC Dataset

The raw dataset was subsequently transferred to Birmingham for extensive processing utilising the RELION (v3.1.3) processing pipeline. The dataset was motion corrected utilising RELION's own implementation of the UCSF MotionCor2 programme before CTF estimation performed using CTFFIND 4.1. Picking was executed by the AI-based Topaz picking algorithm, utilising the default ResNet16_u64 model and picking threshold of 0, resulting in the selection of 1,138,419 particles. These particles were filtered coarsely via 2D classification, allowing the removal of apparent junk particles, to obtain classes displaying numerous orientations (Fig. 50). These classes were entered into successive rounds of 3D classification, performed using C1 symmetry, the first of which utilised a low pass filtered version of the map produced by Deniz Ugurlar as its initial model. Each iteration of 3D classification saw the evolution of the initial model allowing more accurate assignment of angles and removal of junk and damaged particles resulting in 139,466 particles for refinement.

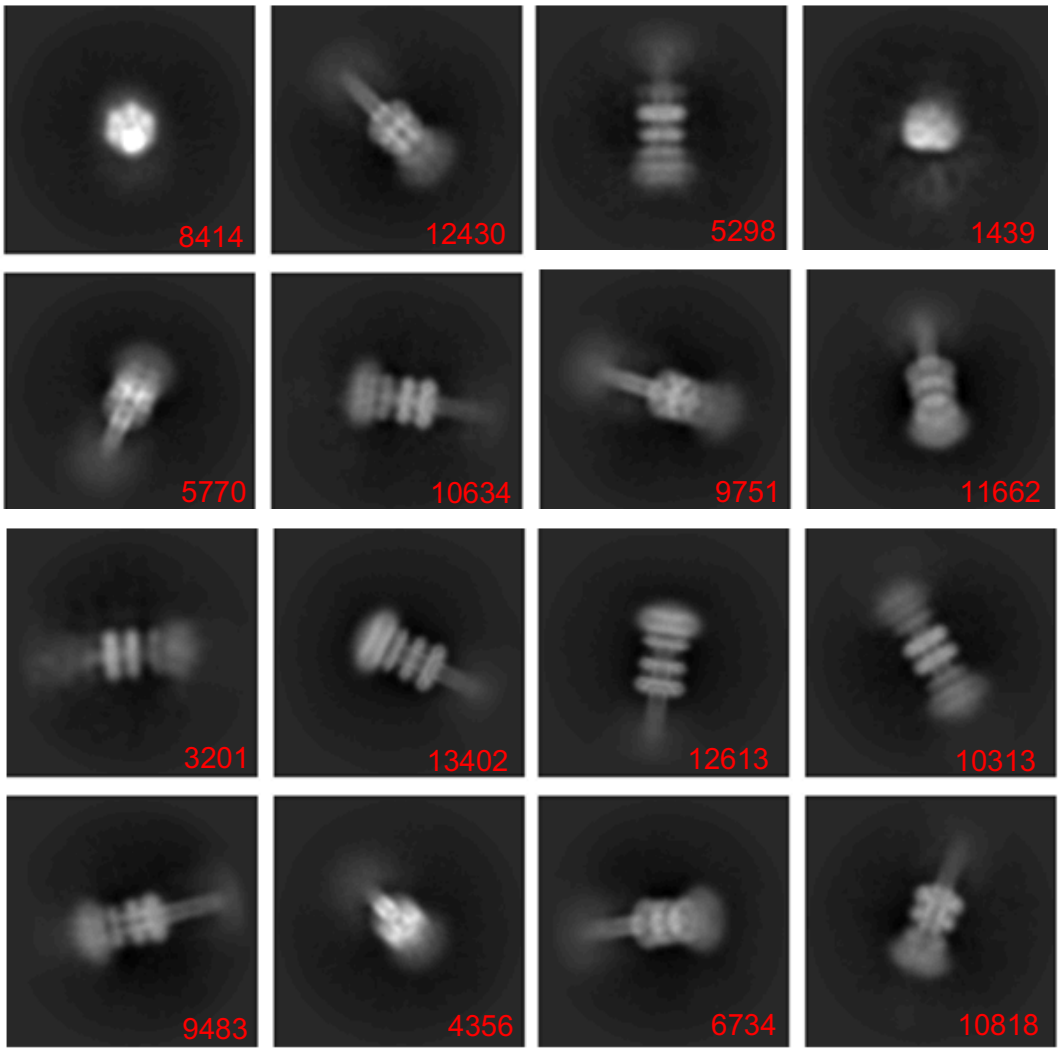


Fig. 50 – PqiABC final 2D classes.

The final 2D classes observed after successive rounds of 2D classification demonstrating numerous orientations of the particle existed within the micrographs. The number of particles within each class is indicated in red.

C1 Refinement and Postprocessing

The refinement and postprocessing steps were initially performed utilising C1 symmetry owed to the anticipated differing oligomeric states, and thus symmetries, of the three Pqi complex constituents. Considering the apparent lack of density corresponding to the PqiA or PqiC moieties of the complex in the map produced by Deniz Ugurlar, our first aim was to discern any densities within our maps which may be attributed to these components.

We first utilised a very broad mask, encompassing all the density depicted in our most promising 3D class (Fig. 51A). This yielded a map with a global resolution of 7.71 Å, hereafter referred to as ABC1, however, local resolution estimates were far more variable (Fig. 51) (Appendix. 6).

All 18 MCE domains and the coiled coil of PqiB are clearly represented at all map contour levels (Fig. 51B-C). The coiled coil itself adopts a tilted conformation, leaning ~12° to one side, indicating a clear propensity for “bent’ particles within the dataset, potentially indicative of inherent flexibility. The large variation in local resolution estimates for the various equivalent MCE domains of PqiB (~6 - 24 Å) is further evidence of fundamental flexibility within the PqiB molecule (Fig. 51B-C).

In addition to the core PqiB density, two additional low-resolution densities are also present at the tip and midpoint of the PqiB coiled coil are present at low map contour levels, (Fig. 51B). Unfortunately, we were unable to further resolve either density, thus it remains difficult to confirm their validity. Finally, closer inspection of the micelle area once again revealed no indication of any PqiA related density with only the transmembrane helices of PqiB present.

Considering the potential for the two additional low-resolution densities to arise from noise, we designed a tighter mask, largely excluding these regions, and repeated the refinement

and postprocessing steps. The resulting map, hereafter referred to as ABC2, possessed a near identical global resolution (7.68 Å), however, a vastly improved range of local resolution estimates (Fig. 52) (Appendix. 6). The PqiB characteristics observed in the ABC1 structure, coil tilting and MCE resolution variation, were represented in the ABC2 map, however, the removal of additional noise aided these observations. Unfortunately, the tighter mask failed to assist in the elucidation of definitive density corresponding to PqiA nor the C-terminus of PqiB (Fig. 52).

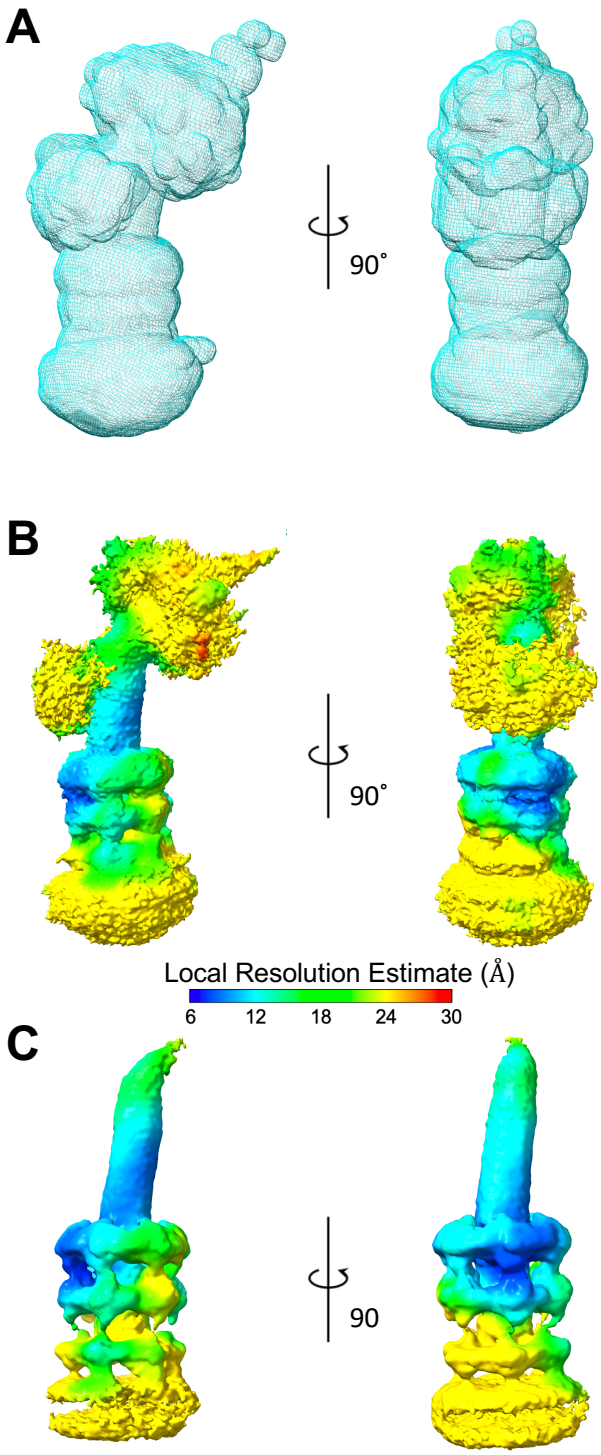


Fig. 51 – Refinement and postprocessing of the ABC1 map.

(A) Broad mask utilised in the refinement and postprocessing steps. (B) ABC1 map depicted at the 0.002 contour level indicating additional densities at the tip and side of the PqiB coil. (C) ABC1 map at the 0.02 contour level indicating the variation in MCE domain resolution. Maps are coloured according to local resolution estimate (Å).

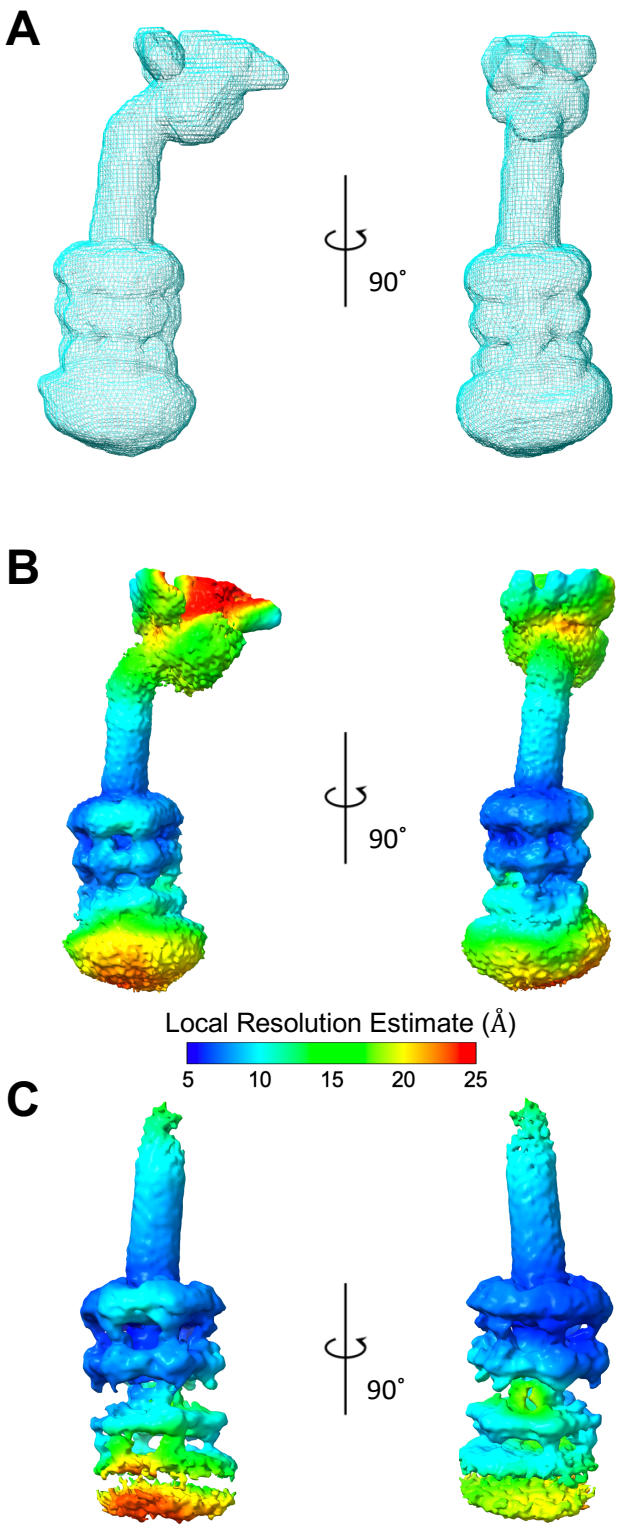


Fig. 52 – Refinement and postprocessing of the ABC2 map.

(A) Narrow mask utilised in the refinement and postprocessing steps. (B) ABC2 map depicted at the 0.01 (B) and 0.025 (C) contour levels. Maps are coloured according to local resolution estimate (Å).

Characterising Flexibility – Multi-body Refinement

Whilst the tighter mask had produced a marginal improvement in map resolution, the apparent inherent flexibility of the core PqiB component ultimately limited resolution. Consequently, we attempted to characterise this flexibility utilising RELION's multi-body refinement tool which characterises relative motion between user defined rigid bodies. Subsequent principal component analysis allows the observed variance to be assigned to distinct eigenvectors (Nakane et al., 2018).

We designed a series of masks isolating the micelle, MCE domains 1-3 and coil regions of the map which we wished to consider as rigid bodies before running the refinement tool, (Fig. 53A). Principal component analysis revealed the first three eigenvectors to account for ~70% of the total observed variance, with the remaining vectors accounting for < 3% each, indicative of three large particle motions (Fig. 53B).

The first and second eigenvectors, accounting for ~30% and ~20% of the variance respectively, describe adjacent radial sweeping motions of the PqiB coiled coil of ~40°, whilst the third eigenvector defines a rotation of the PqiB coiled coil of ~60° (Fig. 53C-E). That all three eigenvectors, accounting for such a large proportion of the total variation, portray motion relating to the PqiB coiled coil reinforce this as a highly flexible moiety and somewhat justify our inability to further refine the structure. Thus, alignment of the molecule is driven primarily by the relatively rigid MCE domain core, however, beyond this, the large variation of conformations hinders robust particle alignment and the accurate assignment of angles.

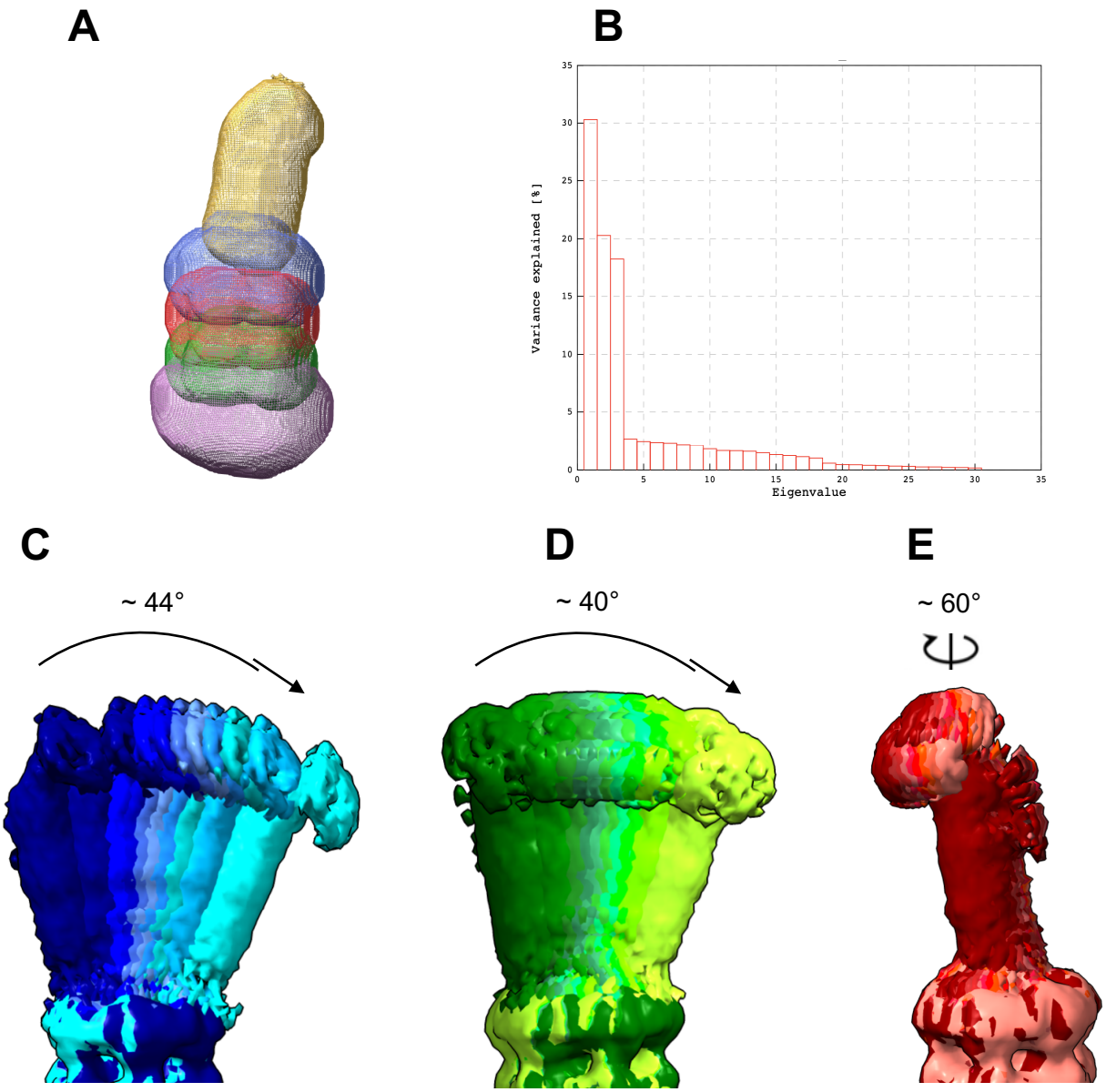


Fig. 53 – Characterisation of PqiABC heterogeneity.

(A) Mesh representation of the masks utilised during multi-body refinement. Masks for the micelle, MCE domains 1-3 and coil are depicted in purple, green, red, blue and yellow respectively. (B) Contribution of each eigenvector to the observed variance. (C) Coil sweeping motion described by the first eigenvector. (D) Adjacent coil sweeping motion described by the second eigenvector. (D) Coil twisting motion described by the third eigenvector. The micelle and bottom two MCE domains were removed to aid clarity.

C6 Refinement and Processing

During our C1 refinement and processing it became clear that our dataset lacked unambiguous density relating to either PqiA or PqiC, furthermore, the heterogeneity exhibited by PqiB prevented us from generating a high-resolution map. Consequently, we decided to refocus our efforts on processing the data utilising sixfold symmetry and tighter masks with the aim of obtaining a higher resolution map of the native membrane associated PqiB component.

Initially we utilised a mask surrounding the entire PqiB density as well as that relating to the micelle (Fig. 54A), yielding a post processed map (BFL1) with a global resolution of 4 Å (Fig. 54B-C) (Appendix. 6). Local resolution estimation revealed the core MCE domains to be the best resolved (~3.5 Å) with resolution decreasing towards the periphery of the map along the coiled coil (~6.5 - 9.5 Å) and micelle (~6.5 - 15.5 Å) moieties respectively (Fig. 54B). Consequently, these portions became imperceivable at higher contour levels, revealing a high-resolution map of the PqiB core comprising density relating to the transmembrane helices, MCE domains and approximately the first third of the coiled coil (Fig. 54C). The resolution of this map permitted the observation of secondary structural elements indicating its potential to facilitate the building of a structural model.

Considering the obvious improvement in explicit PqiB density resulting from the application of symmetry, we designed a mask to isolate the PqiB core (MCE domains and coiled coil base) in the attempt of pushing the resolution of these regions further (Fig. 55A). The resulting post processed map (BFL2) possessed a global resolution of 3.4 Å suggesting a slight improvement over the BFL1 map (Fig. 55B-C) (Appendix. 6). Furthermore, local resolution estimates indicated the resolution of the MCE domains and coiled coil base to be approximately 3.2 Å and 6.2 Å, again indicating marginal improvements over the BFL1 map.

Despite the higher resolution of the BFL2 map, it displayed one key inferiority when compared to the BFL1 map. Whilst all three MCE domains remain well defined, density located between the planes of the first and second MCE domains is significantly more ambiguous, becoming absent at higher map contour levels (Fig. 55C). Therefore, the density correlating to the helices connecting the first and second MCE domains, as well as that of the pore lining loops of the first MCE domain, may not be comprehensively distinguished in the BFL2 map.

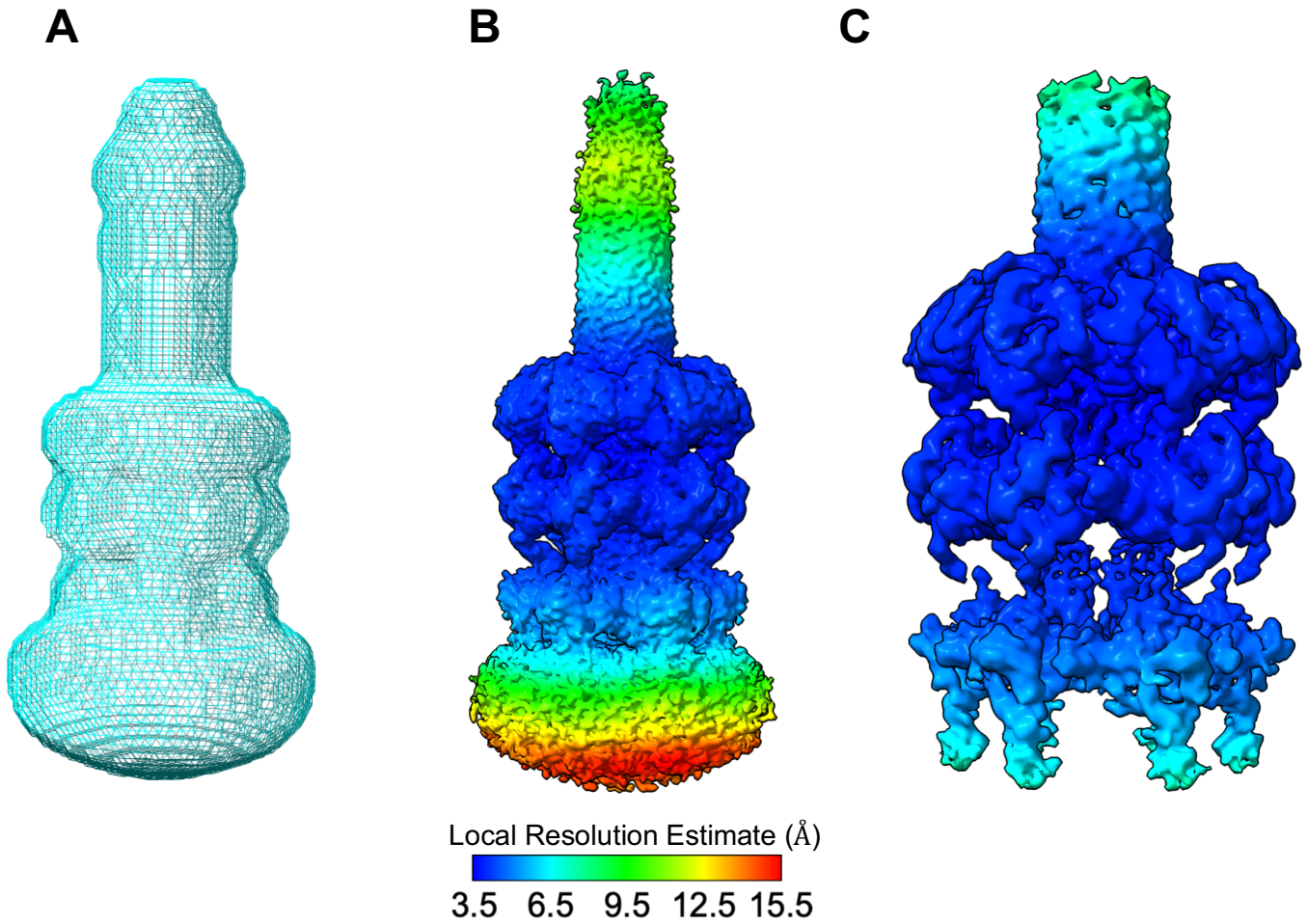


Fig. 54 – Refinement and postprocessing of the BFL1 map.

(A) Mask utilised in the refinement and post processing steps. (B) BFL1 map depicted at the 0.0055 contour level. (C) BFL1 map depicted at the 0.0095 contour level. Maps are coloured according to local resolution estimate (Å).

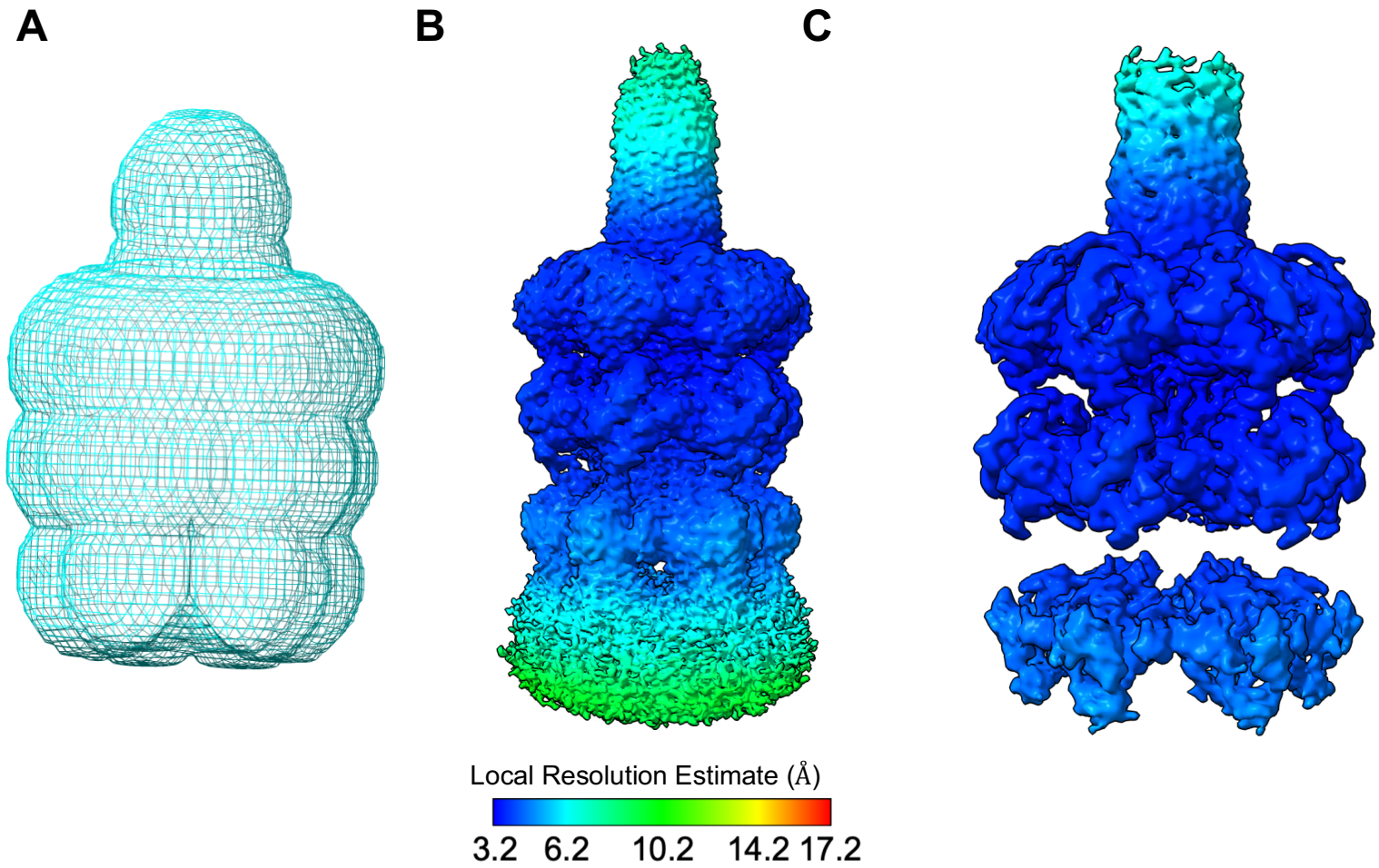


Fig. 55 – Refinement and postprocessing of the BFL2 map.

(A) Mask utilised in the refinement and post processing steps. (B) BFL2 map depicted at the 0.0055 contour level. (C) BFL2 map depicted at the 0.0095 contour level. Maps are coloured according to local resolution estimate (Å).

PqiB Core Model Building

Next, we utilised our six-fold symmetrical maps (BFL1 and BFL2) to build a structure of the native PqiB core. Despite the higher global resolution of the BFL2 map, 3.4 Å vs 4 Å, we built our model from the BFL1 map considering its clearer density for moieties located between the first two MCE domains and that relating to the transmembrane helices. These regions were poorly defined and or absent from the BFL2 map thus the BFL1 map enabled us to generate a more accurate and comprehensive model.

Initially, the map was sharpened to improve its clarity and aid the modelling process (Fig. 56A). Subsequently the existing cryo-EM derived structure of a soluble PqiB construct (PDB: 5UVN) was segmented, into 4 sections comprising each of the 3 MCE domains and the coil moiety, before each portion was positioned into the BFL1 map density via rigid body fitting (Fig. 56B). The positioned segments were re-assembled into continuous protein chains and regions absent from the existing structure built *de-novo* before the entire structure was refined yielding the final BFL structure (Fig. 56C, Fig. 57).

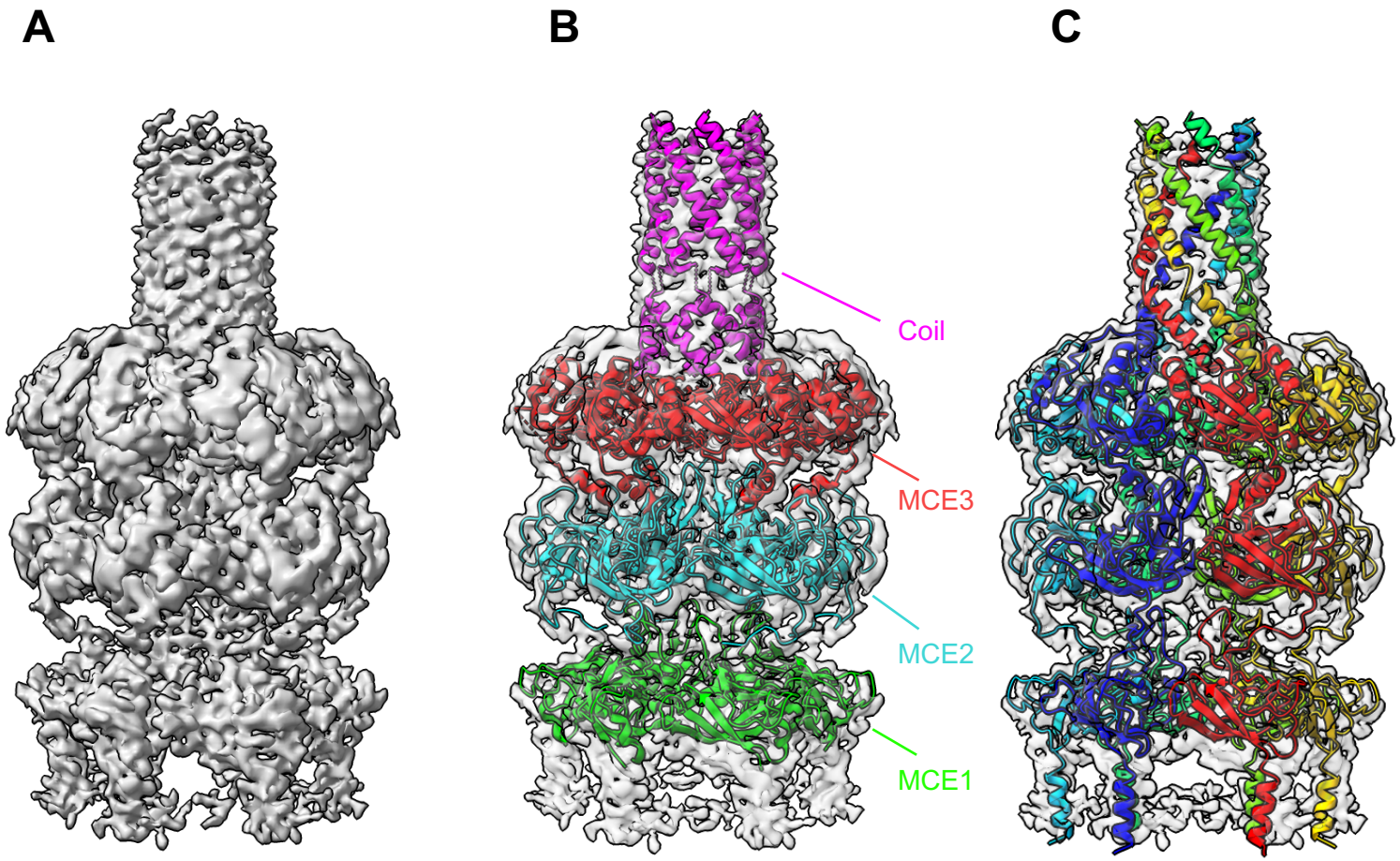


Fig. 56 – PqiB core model building.

(A) Sharpened BFL1 map displayed as a surface at the 8 map contour level. (B) Segments of the existing soluble PqiB structure (PDB: 5UVN) fit into the sharpened BFL1 map. Segments are distinguished by unique colours. (C) Cartoon representation of the BFL model, with rainbow chain colouring, within the sharpened BFL1 map.

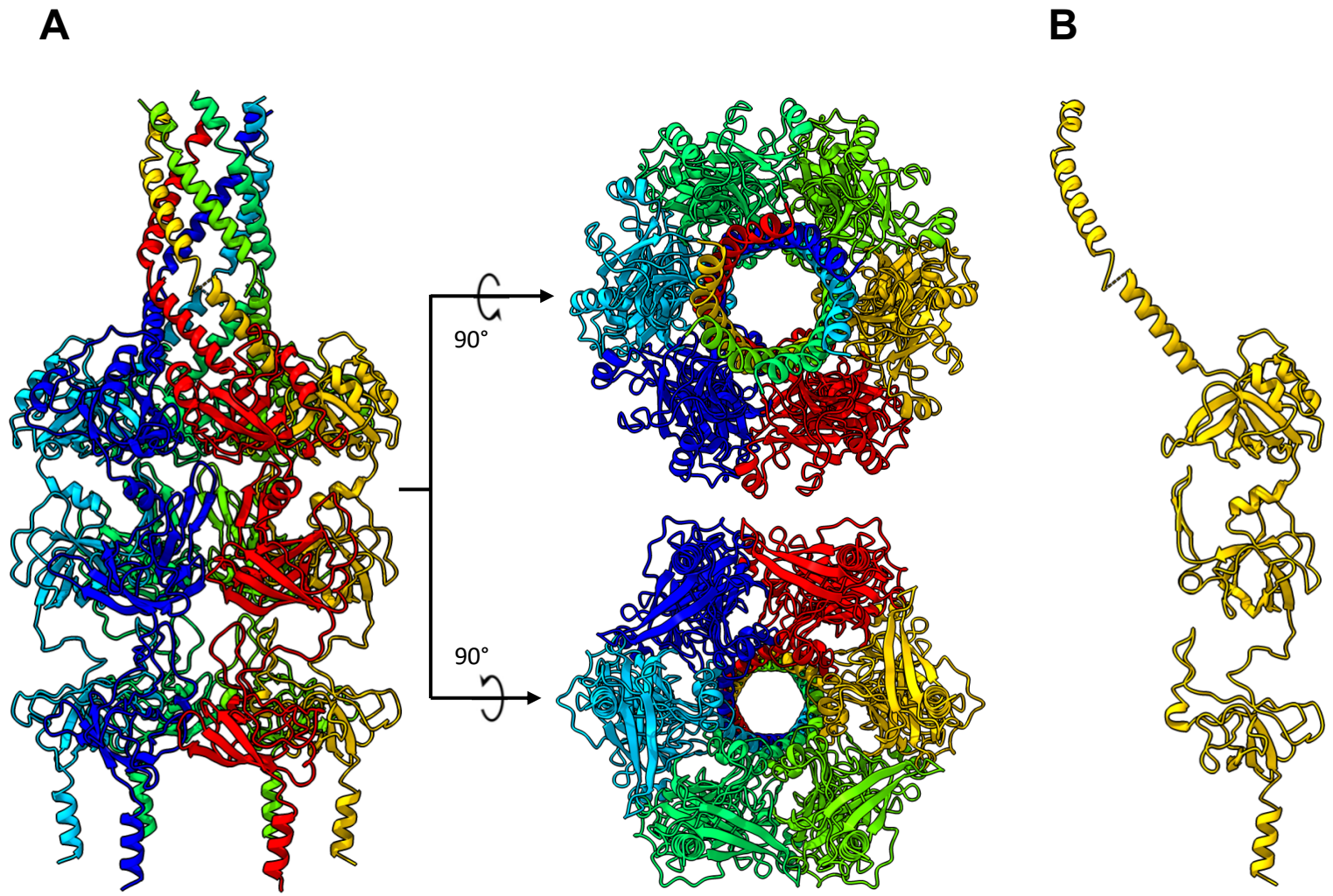


Fig. 57 – Final PqiB BFL structure.

(A) Orthogonal views of the BFL structure, depicted as a cartoon, with rainbow chain colouring. (B) Cartoon representation of a BFL monomer.

The final sharpened BFL1 map displayed density suggestive of differing pore lining loop (PLL) conformations within each of the PqiB MCE domains. This was most evident for the PLL of the second MCE domain where an additional hairpin of density was observed to the right of each of the modelled PLLs, indicating two lateral loop conformations (Fig. 58A). Furthermore, the modelled left density was suggestive of two alternative conformations of the PLL tip, with the second unmodelled position moving the loops further into the central pore (Fig. 58B).

The map also indicated the presence of additional PLL conformations for the first and third MCE domains. Density surrounding the MCE1 PLL was suggestive of the same conformational variability observed for MCE2 however the density itself was less explicit. Once again, an additional hairpin of density to the right, this time less complete, implied the presence of differing lateral conformations whilst an ‘m’ shaped density above the two hairpins alludes to the potential of loop flexion (Fig. 59A). The density surrounding the MCE3 PLL was indicative of much less motion with only the loop tip potentially adopting two conformations, differing by approximately 90°, pointing vertically (modelled) or horizontally towards the pore cavity (Fig. 59B).

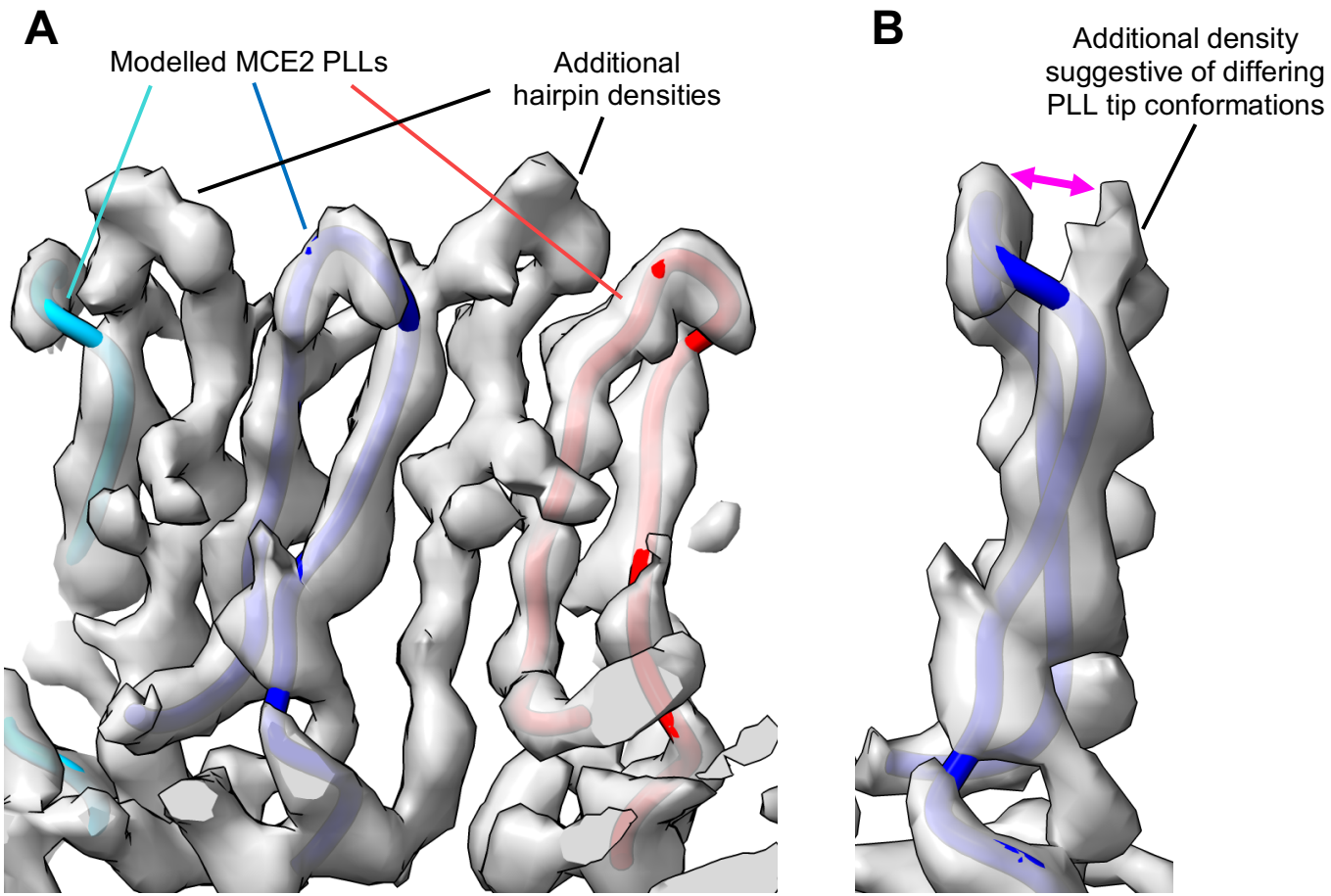


Fig. 58 – BFL1 MCE2 alternative pore lining loop density.

(A) Slab view of the MCE2 PLL region indicating the presence of an additional hairpin of density to the right of each modelled PLL. (B) Slab view of a single MCE2 PLL indicating the densities suggestive of flexion of the loop tip towards the central pore. The Sharpened BFL1 map displayed as a transparent grey surface at the 12.5 map contour level whilst the BFL model is depicted in liquorice form with rainbow chain colouring.

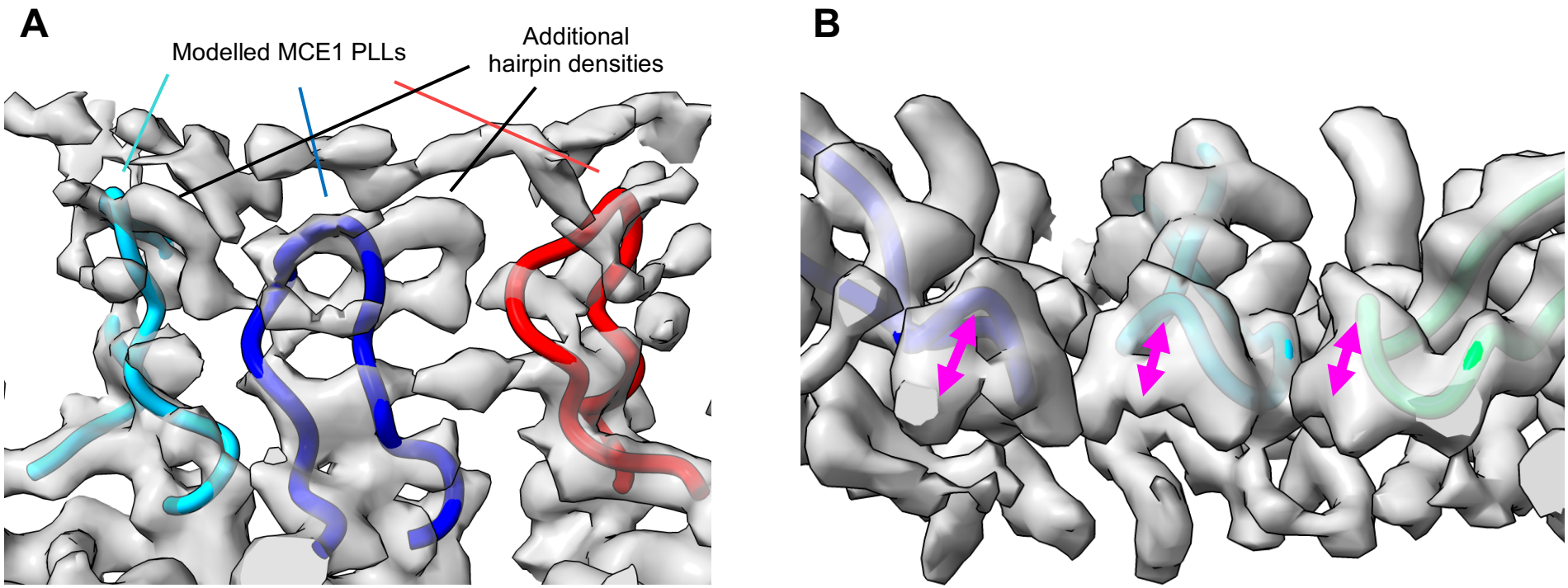


Fig. 59 – BFL1 MCE1 and MCE3 alternative PLL density.

(A) Slab view of the MCE1 PLL region indicating an additional hairpin of density to the right of each modelled PLL and an ‘m’ shaped density above the two hairpins. (B) Slab view of the MCE3 PLL region indicating the potential for the surrounding density to accommodate two conformations of the PLL tip. The Sharpened BFL1 map displayed as a transparent grey surface at the 9.5 map contour level whilst the BFL model is depicted in liquorice form with rainbow chain colouring.

Finally, we compared the BFL structure to the previously deposited structure of a soluble PqiB$^{39\text{-}546}$ construct, PDB: 5UVN. The composition of both structures is nearly identical, with the BFL structure possessing three small additional moieties, thus alignment primarily revealed variations in global architecture (Fig. 60A). The monomers of each structure aligned to yield an RMSD of 0.986 across 229 pruned atom pairs and 2.787 Å across all 400 Å pairs (Fig. 60B). Whilst both structures displayed very similar arrangements of the second and third MCE domains there was a significant discrepancy in the positioning of the first MCE domain. The BFL structure displayed a more elongated arrangement with its first MCE domain residing ~4 A lower than that of 5UVN potentially indicating a distinct conformation from that previously observed (Fig. 60).

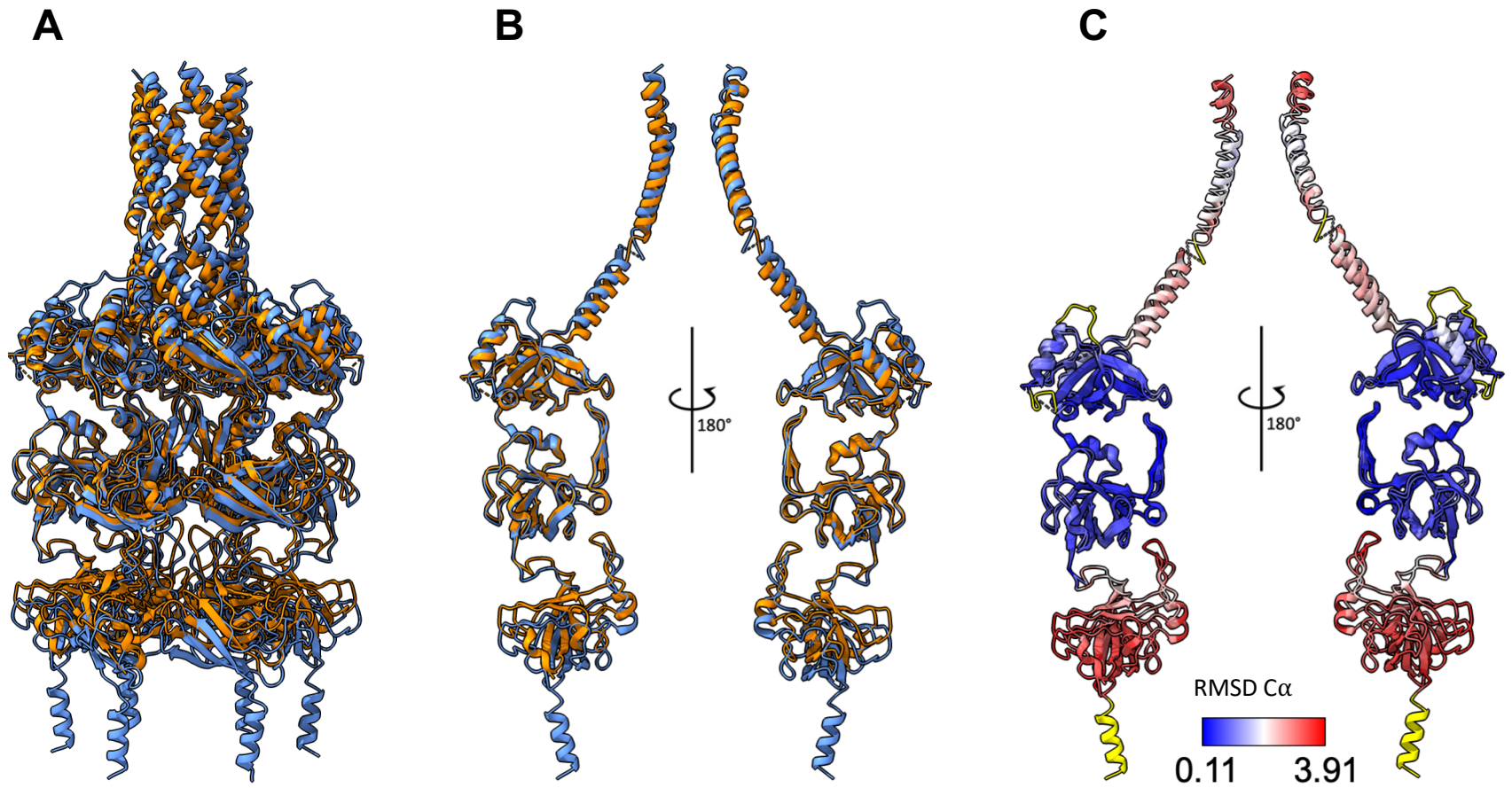


Fig. 60 – Alignment of the BFL structure and PDB: 5UVN.

(A) Alignment of the BFL structure and 5UVN depicted as cartoons coloured blue and orange respectively. (B) Alignment of a single monomer of the BFL structure and 5UVN depicted as cartoons coloured blue and orange respectively. (C) Alignment of a single monomer of the BFL structure and 5UVN depicted as cartoons coloured according to the RMSD of aligned Cα pairs. Regions absent from the 5UVN structure are indicated in yellow.

Discussion

In this chapter we attempted to probe the native PqiABC complex via cryo-EM with the aim of elucidating the overall complex architecture and probing the interactions between complex constituents.

C1 Refinement and Processing

Initially utilising C1 symmetry we produced two maps, ABC1 and ABC2, via the utilisation of two different masks. Both maps contained density for the micelle and core PqiB component, resolved clearer in the ABC2 map, whilst the ABC1 map contained two additional low-resolution densities at the tip and approximately halfway up the coiled coil of PqiB. Both additional densities were of low resolution and extremely noisy, raising concerns about the validity of their existence. The top density appears largely accommodating of the PqiC octamer but only when tilted to approximately 45° (Fig. 61). Such an orientation appears unlikely biologically, however, the observed lean of the PqiB coiled coil may explain its existence within this dataset. The additional density located midway up the PqiB coiled coil is more intriguing, potentially revealing an additional complex component. We attempted to probe the origin of this density via bottom-up proteomics, performed at the University of Birmingham by Jedd Bellamy-Carter, however, no relevant proteinaceous hits were identified. Consequently, we suggest that the density may arise from remnants of the peptidoglycan layer, retained during the purification. The height of the density on the PqiB coiled coil aligns well with the anticipated location of the peptidoglycan layer within the periplasm further reinforcing this notion. Ultimately the differentiation of either density from noise is unambiguous, and their identities cannot be definitively determined, thus the suggestions we have made are purely speculative.

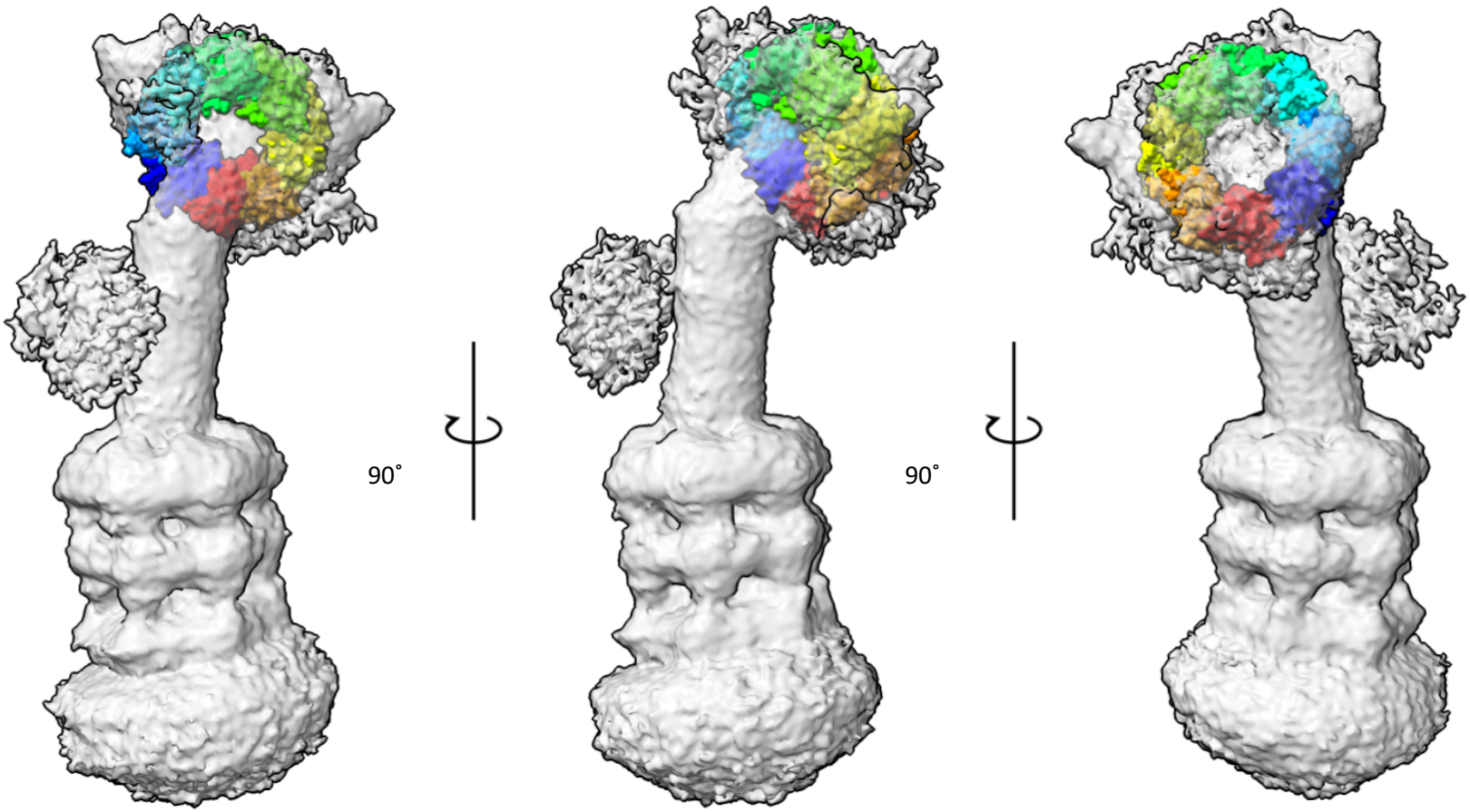


Fig. 61 – Fitting of the PqiC octamer into the ABC1 map.

Orthogonal views of the PqiC octamer fitted into the low-resolution density present at the top of the PqiB coiled coil in the ABC1 map. PqiC is depicted as a surface with protomers coloured according to a rainbow palette.

Unfortunately, inspection of the micelle revealed no evidence of PqiA in either the ABC1 or ABC2 maps thus we were unable to provide insight into the orientation of PqiA nor its interaction with PqiB. The complete absence of PqiA was surprising considering its definite observation on the SDS-PAGE gel of the same sample following size exclusion. A potential explanation arose from the aforementioned bottom-up proteomics analysis which, in addition to PqiA, PqiB and PqiC, revealed the presence of FtsH, an ATP-dependent zinc metalloprotease capable of degrading membrane proteins, within the sample. Thus, the absence of PqiA related density within this dataset maybe due to its degradation, primarily by FtsH, in the period between elution from the size exclusion column and grid vitrification however this seems unlikely. Alternatively, PqiA may denature during the vitrification process due to interactions with the air-water interface, an environment known to be detrimental to proteins (Glaeser, 2018; Glaeser and Han, 2017).

A defining characteristic of both C1 maps is the tilted orientation of the PqiB coiled coil. Indeed, multi-body refinement revealed ~70% of the variation within our dataset to be accounted for by the array of coiled coil positions, with the radial array of poses accounting for ~50% alone. The single existing cryo-EM structure of a soluble PqiB construct (PDB 5UVN) displays a straight coiled coil region, owed to the application of C6 symmetry during the refinement and postprocessing steps, however indications of its flexibility are evident within the associated publication (Ekiert et al., 2017). Initial negative stain EM 2D classes clearly indicate an array of PqiB coiled coil poses, displaying the same radial sweeping motion present in our own cryo-EM 2D classes (Fig. 62A). A subsequent 3D reconstruction of the soluble PqiB construct revealed a tilted coiled coil orientation analogous to that observed in our C1 maps (Fig. 62B).

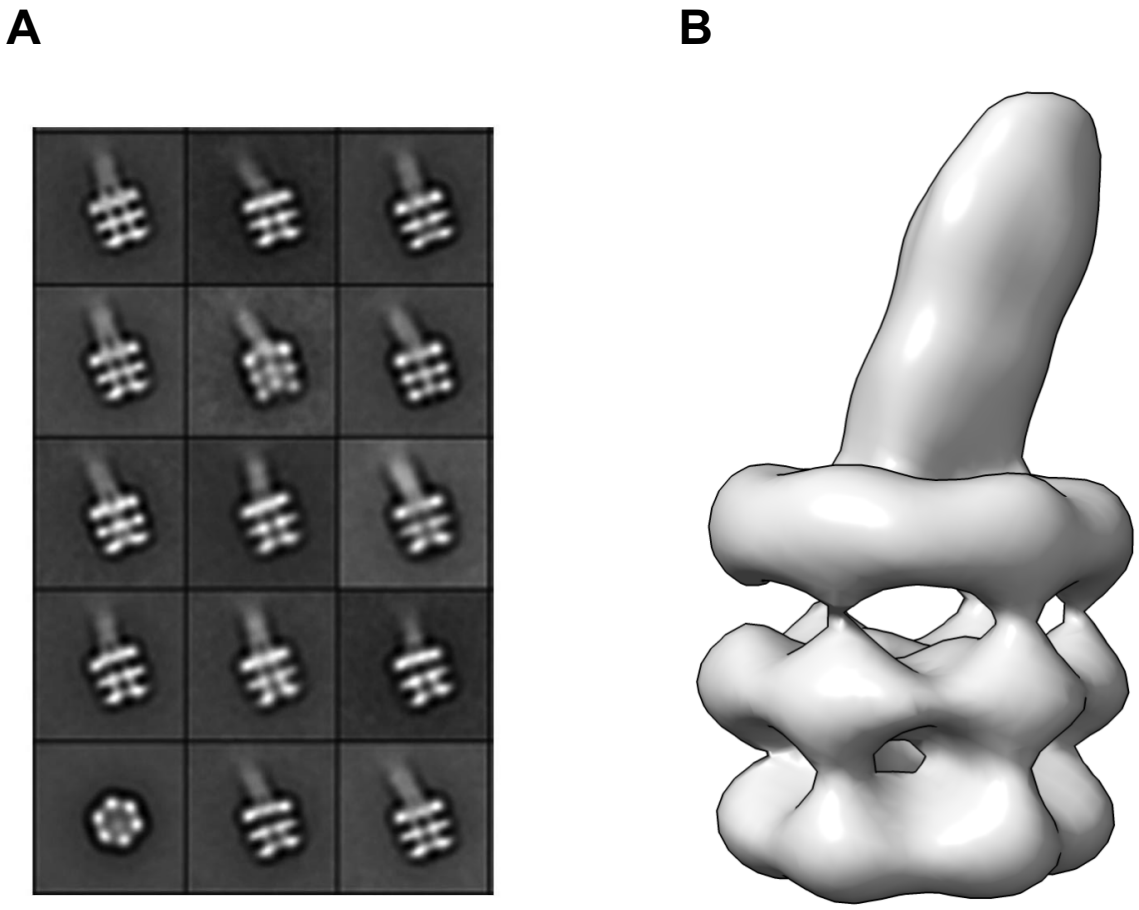


Fig. 62 – Coiled coil flexibility in existing soluble PqiB structures.

(A) Negative stain EM 2D class averages of a soluble PqiB$^{39-546}$ construct indicating a range of coiled coil poses (Ekiert et al., 2017). (B) Negative stain EM 3D reconstruction a soluble PqiB$^{39-546}$ construct displaying a tilted coiled coil as seen in our C1 reconstructions (EMDB EMD-8612).

We anticipate the observed flexibility of the PqiB coiled coil arises due to the removal of PqiB from its native, membrane spanning location. Ordinarily, the expected anchoring of PqiB to both the inner and outer membranes, either directly or via a PqiC intermediate, limits the

flexibility of the coiled coil, thus maintaining its vertical orientation. Removal of these membrane interactions, either via the expression of a soluble construct (Ekiert et al., 2017) or via detergent solubilisation of the entire complex, releases the tension along the PqiB moiety, enabling it to flex and adopt a range of conformations.

Considering the predicted positioning of PqiC at the apex of PqiB, the lack of definitive density for its presence within our maps likely results from its vast array of different spatial locations. Such positional variation prevents the robust alignment of particles, necessary for SPA pipelines, hence the signal is typically interpreted as noise and averaged out. Thus, subsequent attempts to solve the entire PqiABC complex via cryo-EM would likely benefit from attempts to reduce this flexibility either via additional supports or construct modification.

The large variation in PqiB coiled coil orientations prevented the robust alignment of particles, ultimately limiting map resolution. Whilst we utilised RELION's multi-body refinement to characterise heterogeneity across our dataset, it was saturated by the array of coiled coil poses thus masking any smaller, potentially biologically relevant, motions. Indeed, variations in the pore lining loops, modulating the diameter of the central cavity at key locations, have been observed previously in both PqiB and YebT (Isom et al., 2020). Potentially, one of a new generation of heterogeneity algorithms, utilising deep neural networks (Zhong et al., 2021) and deformation fields (Herreros et al., 2021), may better characterise the motion within our dataset however their utilisation was beyond the time scope of the current study.

C6 Refinement and Processing

The lack of definitive density corresponding to either PqiA or PqiC in our C1 processed maps was disappointing, however, the PqiB component was clearly defined. Therefore, we refocussed our efforts upon optimising our maps of the PqiB component via the utilisation of six-fold symmetry and tighter masks.

Our subsequent processing yielded two maps, BFL1 and BFL2 at 4 Å and 3.4 Å global resolution respectively. The mask utilised for the refinement and postprocessing of the BFL1 map was designed with the intent of maintaining the entire length of the PqiB coiled coil, however, the large lateral deflections of the moiety prevented the construction of clear density beyond approximately half of its length. Furthermore, when the map contour level was increased to a level conducive for modelling, the extent of the coiled coil density reduced to ~1/3 of its length. This coverage of the coiled coil density is comparable to that present in the EMD-8608 map from which the existing soluble $PqiB^{39-546}$ structure (PDB: 5UVN) was built (Fig. 63). Indeed, the AlphaFold structure of PqiB (AF-P43671-F1) predicts the interruption of the coiled coil with a small loop region at approximately the height at which these two maps end. Acting as a hinge, this loop region may introduce flexibility to the coiled coil, thus explaining the diminishment of its density beyond this point.

Despite compromising the extent of the coiled coil density, the BFL1 map appeared to encompass the majority of PqiB, with the MCE domains particularly represented at high-resolution (Fig. 63). The second BFL2 map was produced from a smaller mask encompassing the core MCE domains and coiled coil base of the BFL1 map (Fig. 63). Despite the higher reported global resolution of this map, 3.4 vs 4 Å, the map appeared inferior to the BFL1 map owing to ambiguous density of the coiled coil and the loss of density residing between the first and second MCE domains (Fig. 63). The loss of these regions was surprising considering their clear presence within the BFL1 map, however, their absence

would severely hinder any model building attempts thus the BFL2 map was largely discounted.

Comparison of the EMD-8608 map to those generated in this study yields one obvious discrepancy, most evident in the BFL1 map. The three MCE domains are stacked equidistantly in the EMD-8606 map, however, whilst the distance between the second and third MCE domains remains similar in the BFL maps, there is a much larger interval between the first and second domains (Fig. 63). Consequently, the EMD-8086 map displays a more compact domain architecture than that of our own maps. This discrepancy is exemplified by the fitting of the 5UVN model, built from the EMD-8086 map, into our own BFL1 map (Fig. 64). The coiled coil and third (uppermost) MCE domain of the 5UVN model align perfectly with our BFL1 map, however, the discrepancy between map and model increases towards the N-terminus. Whilst the second MCE domain is marginally higher than the BFL1 map would suggest, the first MCE domain resides completely out of register with its corresponding density. Therefore, our BFL1 map appears to indicate distinct differences in the architecture of native and soluble PqiB constructs.

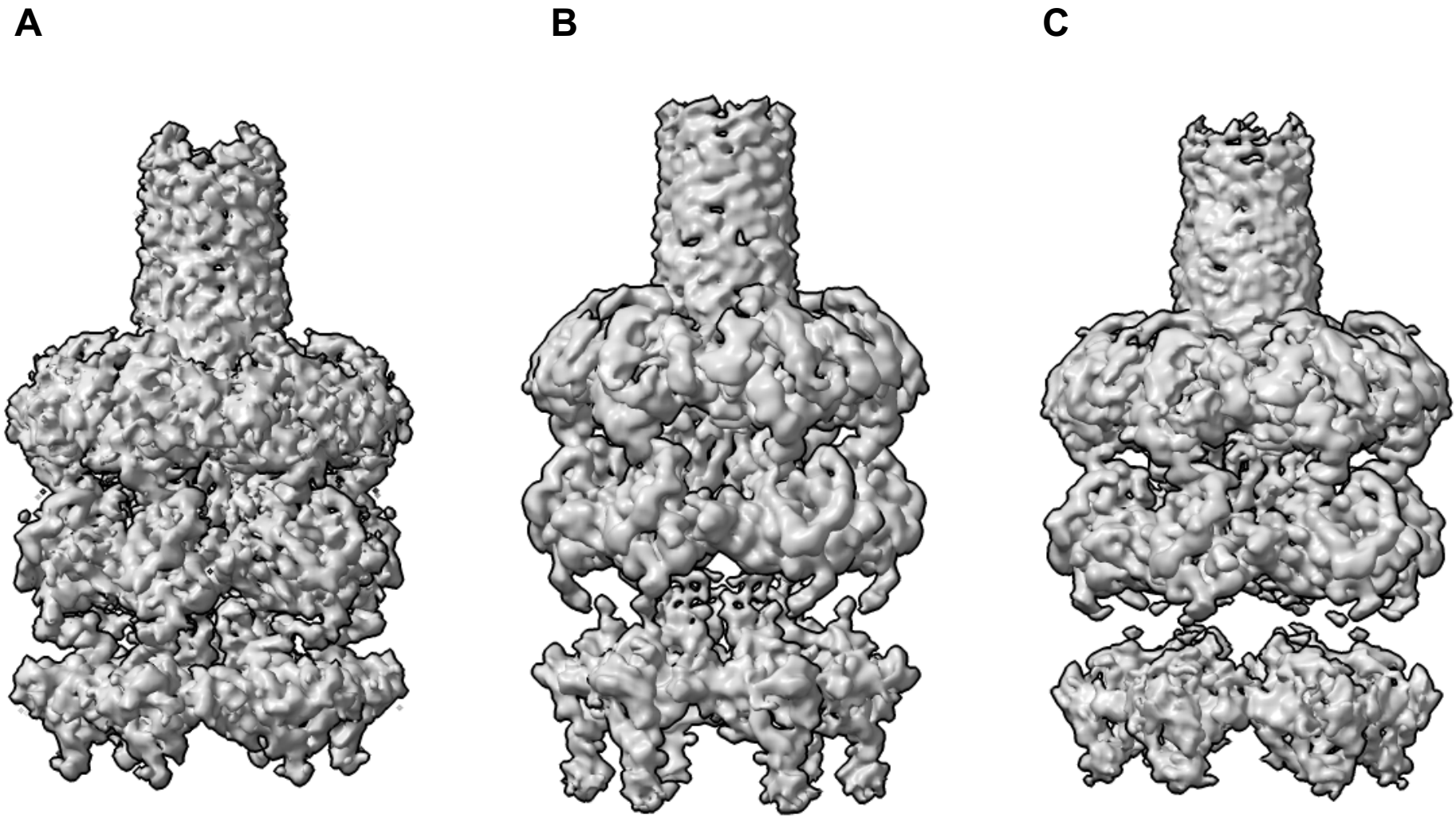


Fig. 63 – Comparison of C6 PqiB cryo-EM maps.

(A) EMD-8608 map from which the soluble PqiB structure PDB: 5UVN was built. (B) BFL1 map. (C) BFL2 map. All maps are depicted as surfaces.

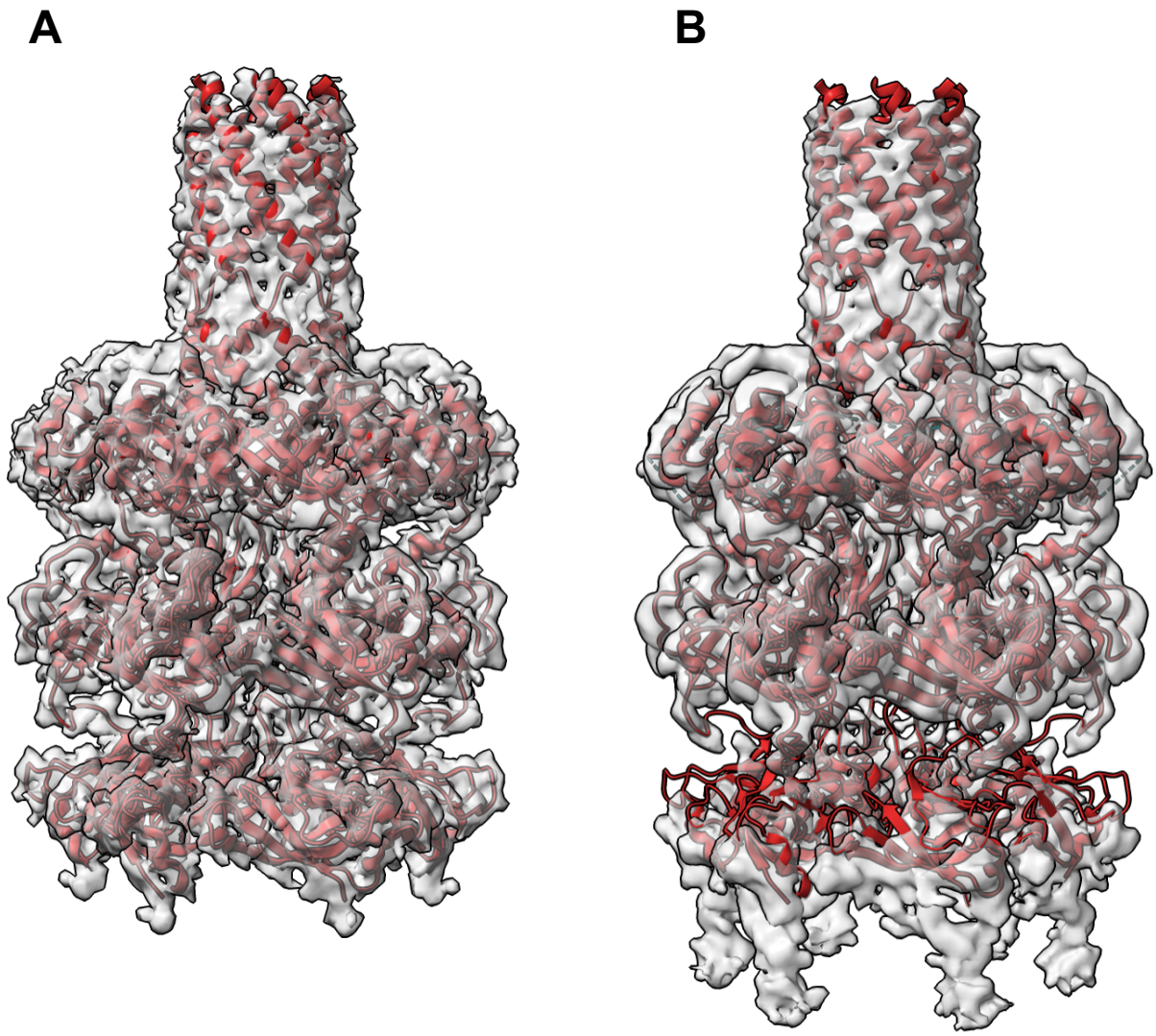


Fig. 64 – Fitting of 5UVN to the EMD-8608 and BFL1 maps.

(A) 5UVN model fit to the EMD-8608 map from which the model was built. (B) 5UVN model fit to our BFL1 map indicating the discrepancy in MCE domain density positioning, 5UVN is displayed as a red cartoon whilst both maps are transparent surfaces.

PqiB Core Model Building

Whilst all three *E. coli* MCE proteins, MlaD, YebT and PqiB, have been probed via cryo-EM, PqiB remains the least characterised with only a single structure of a soluble construct previously deposited. This PqiB$^{39-546}$ structure (PDB: 5UVN) is however incomplete, despite the inherent omission of transmembrane helices, density relating to the coiled coil moiety became increasingly uninterpretable along its length restricting the modelled region to the first ~60 residues. Furthermore, two loop regions in the third MCE domain, residues 348-354 and 388-401, remain unmodelled presumably due to their mobile nature. Utilising our BFL1 map, we built a model of the native PqiB core (BFL), allowing comparison to the existing 5UVN structure, with the aim of clarifying the implied variation between soluble and native PqiB constructs.

The overall domain architecture of our BFL structure is very similar to that of 5UVN, however, their alignment confirms the more elongated nature of the BFL model. We propose that this expanded architecture arises due to tension resulting from the association of PqiB with detergent micelles at either terminus, either directly or indirectly. Under native conditions, PqiB is maintained between the two planes of the inner and outer membranes. Thus, the BFL structure potentially represents a more biologically relevant conformation of the native PqiB protein.

Owing to the quality of our BFL1 map, the resulting BFL model addressed all but one of the aforementioned shortcomings of the 5UVN model, with the entire extent of the coiled coil moiety remaining enigmatic. Most notably, the utilisation of native PqiB allowed the partial modelling of transmembrane helices within the detergent micelle, however, their full extent could not be depicted. The BFL1 map possessed improved density for the third MCE domain enabling modelling of its entirety in the BFL structure, including the two loop regions absent from the 5UVN structure (Fig. 65).

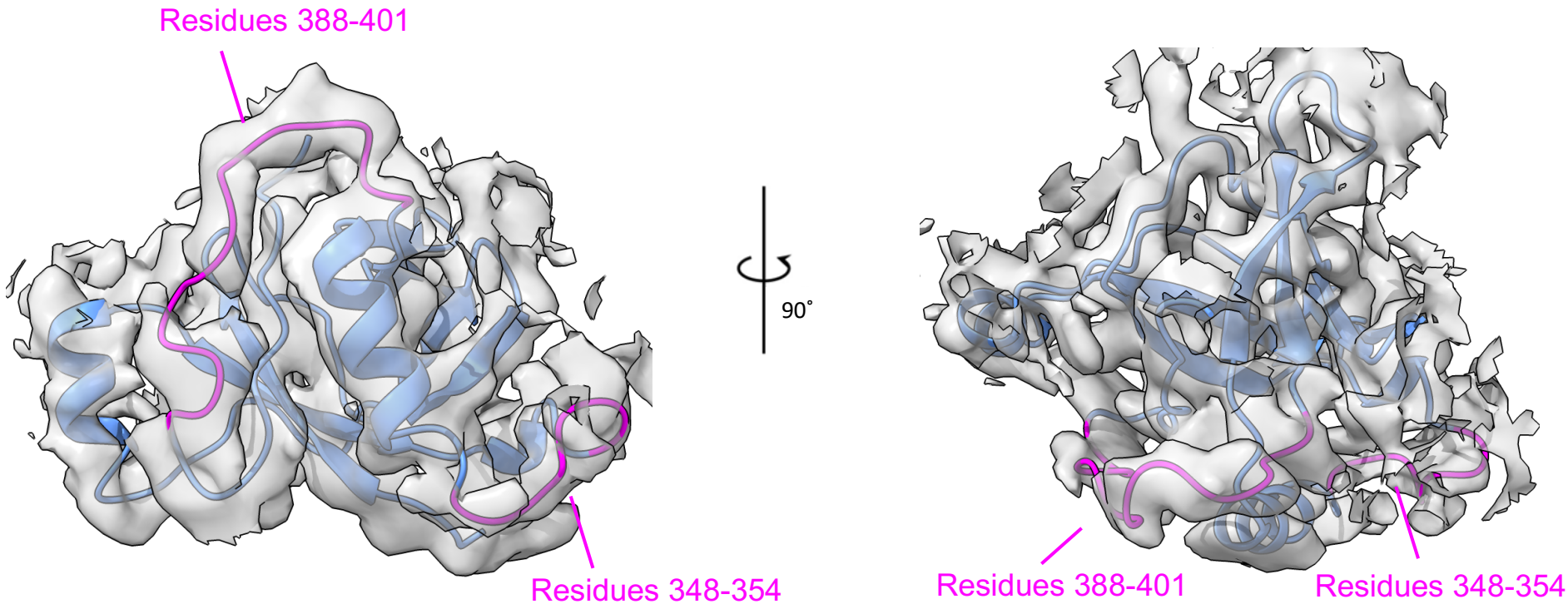


Fig. 65 – PqiB BFL1 MCE loop density.

Cropped BFL1 map around the third MCE domain a single PqiB chain indicating the density allowing modelling of the loop regions absent from the 5UVN structure. Map is displayed as a transparent surface. BFL1 MCE3 is displayed as a blue cartoon with the loop regions coloured magenta.

In addition to the six pore lining loop β-hairpins, each MCE domain of the 5UVN model is lined with six additional strands, positioned in-between the PLL hairpins, of undetermined origin (Fig. 66). The densities from which these ambiguous strands were modelled within 5UVN are resolved more thoroughly in our BFL1 map allowing us to confidently attribute them to differing PLL conformations. Indeed, the BFL1 map displays density suggestive of various PLL conformations particularly within the first and second MCE domains. Differing PLL conformations have been most comprehensively reported for YebT, with multiple discrete orientations reported for MCE domains 1, 5, 6 and 7, where their movement is proposed to modulate the diameter of the central pore potentially regulating transport (Isom et al., 2020; Vieni et al., 2022). Within the same publication, the authors re-analysed their previous $PqiB^{39-546}$ dataset, utilising updated software, and reported the observation of distinct open and closed PLL conformations (Isom et al., 2020). Whilst the authors deposited two YebT structures to the PDB, depicting the extreme open and closed conformations, no updated PqiB maps nor coordinates were made available following the re-analysis, hence direct comparison of their observed loop conformations with the BFL1 map density was not possible.

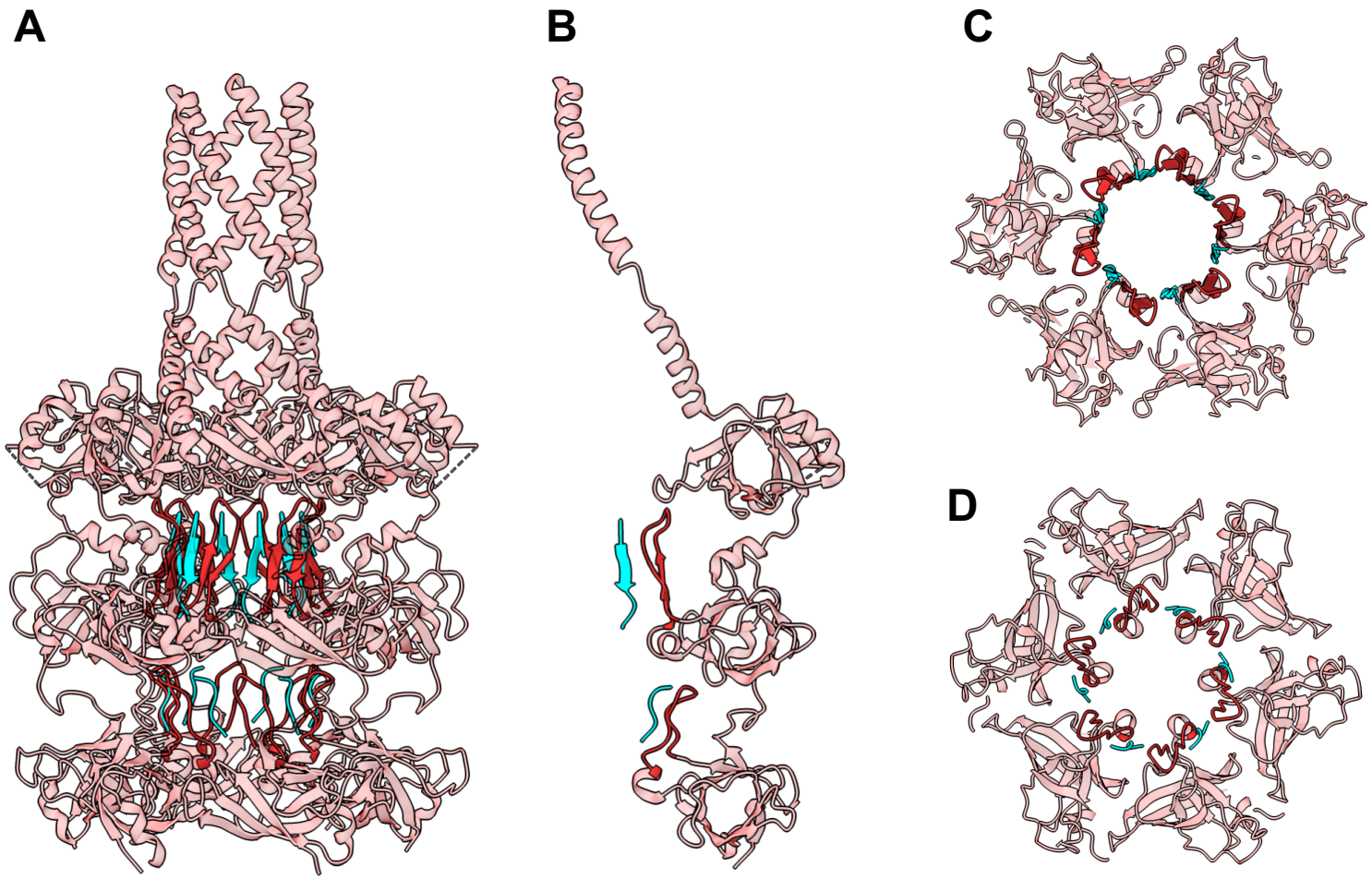


Fig. 66 – PqiB 5UVN additional pore lining strands.

(A) Cartoon of PDB 5UVN coloured pale red with the pore lining loops of the first and second MCE domains coloured red and the additional strands coloured cyan. (B) Cartoon of a single 5UVN chain. (C) Slab ariel view of the second MCE domain. (D) Slab view of the first MCE domain.

Summary

In this chapter we have presented a series of cryo-EM maps resulting from a PqiABC sample. Proteomic analysis of the sample confirmed the presence of all three Pqi components however we were unable to resolve definitive density for PqiA or PqiC, likely due to proteolytic cleavage and flexibility respectively. The production of a masked six-fold symmetrical map enabled us to build a model of the PqiB core which exhibited a different, more elongated, architecture and accounted for the small missing regions of the existing 5UVN structure. Furthermore, the BFL1 map is indicative of several PLL conformations re-affirming their proposed role in modulating the central channel as reported previously (Isom et al., 2020). Consequently, it may be argued that our BFL model presents the most comprehensive representation of the PqiB protein to date, however, its interaction with either PqiA or PqiC still remain enigmatic.

Chapter 5: *In-vivo* Phenotypic Complementation Assays

Introduction

Phenotypic Complementation Assays

The mechanism by which Pqi mediated transport occurs remains enigmatic owing to the absence of published functional studies and the lack of homology to transport systems of known operation. This apparent absence of mechanistic insight may be largely attributed to the current lack of an *in-vitro* assay system, the development of which is ironically hindered by the shortage of functional information creating a somewhat 'chicken and egg like' scenario. Fortunately, we may bypass this issue via the use of phenotypic complementation assays, hereafter referred to as complementation assays, which allow us to probe the relative importance of specific residues within the Pqi complex against a native, whole cell background.

Complementation assays necessitate the generation of a knockout strain, devoid of the gene(s), within which the residues of interest reside (Fig. 67A). This knockout strain is subsequently screened to identify a growth condition upon which the knockout strain displays a distinct phenotype, enabling differentiation of the wild type (WT) and knockout strains (Fig. 67B). Finally, the knockout strain is transformed with plasmids containing the deleted WT region as well as an array of regions harbouring mutations, before growth on the previously identified condition (Fig. 67C). Transformation with the WT region should restore growth to that of the WT strain whilst the relative fitness of those transformed with the various mutated plasmids indicates the severity of the mutation and thus the importance of the mutated residue (Fig. 67D).

E. coli K-12 BW25113 PqiAB and YebST knockout strains have been previously shown to be more sensitive to growth upon media supplemented with the zwitterionic detergent lauryl sulfobetaine (LSB) than their WT parent strains (Isom, 2017). Whilst LSB has been utilised in

several studies for complementation assays probing mutations within YebT, (Isom et al., 2020; Vieni et al., 2022) no complementation assays for specific Pqi mutations have ever been published. Thus, the LSB phenotype of PqiAB presents a currently underutilised tool capable of providing functional insights into the Pqi complex.

Chapter Aims

In this chapter we utilise *in-vivo* complementation assays to probe the essentiality of specific residues within all three components of the Pqi complex. Such assays, combined with structural data from Chapters 3 and 4, and existing literature, allow us to speculate the role of particular residues in complex formation and transport mechanics.

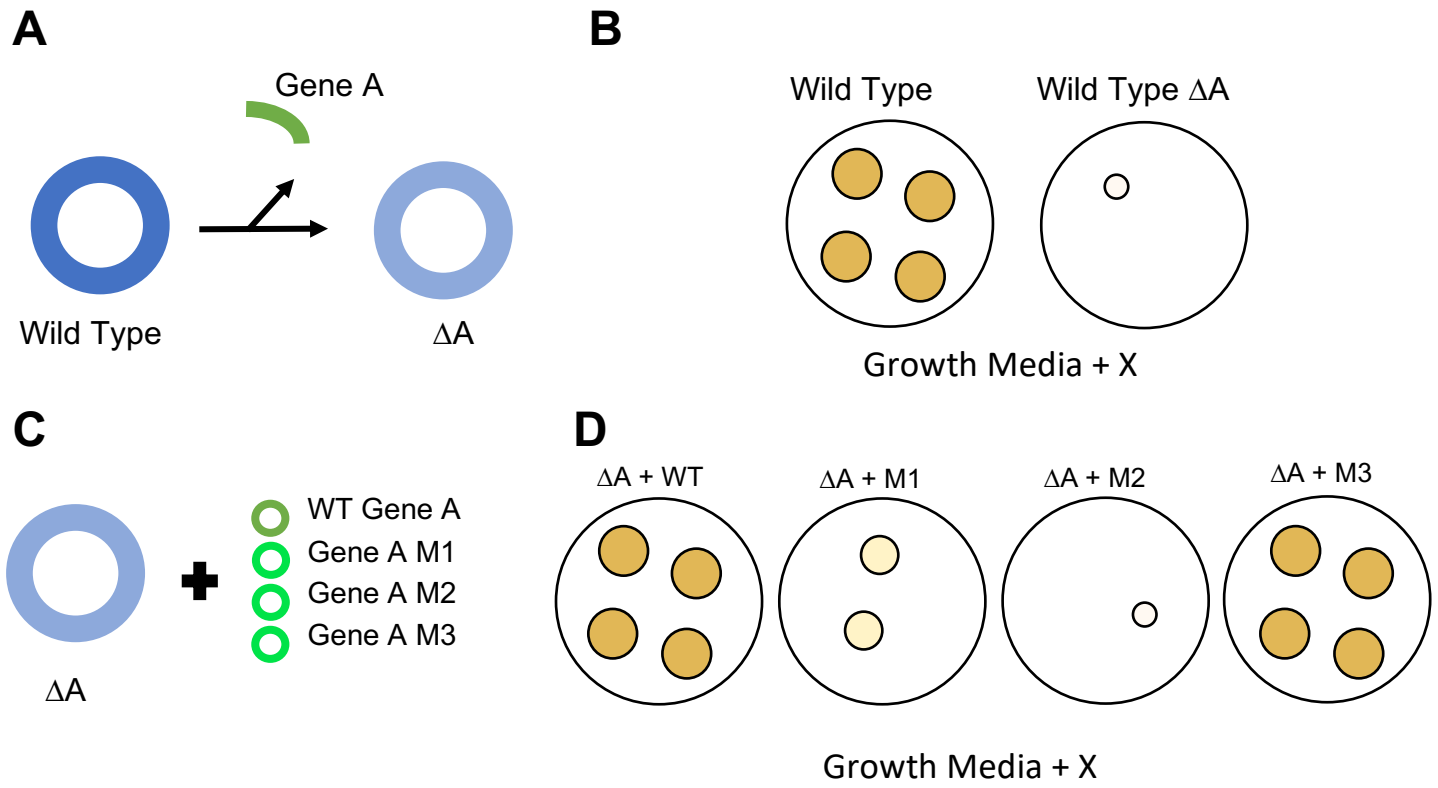


Fig. 67 – Phenotypic Complementation Assays.

(A) Removal of the gene(s) of interest from the WT strain generating a knockout. (B) Identification of a growth condition upon which the WT and knockout strains are phenotypically different. (C) Transformation of the knockout strain with plasmids containing WT and mutant (M) versions of the gene of interest. (D) Growth of the transformed knockout strains on the previously identified growth condition and observation of their resulting phenotypes.

Results

PqiABC Knockout Generation

As stated above, complementation studies require the generation of a knockout strain. Considering we wished to probe mutations in all three Pqi components, and already possessed a plasmid of the entire ABC operon within which to generate mutants, we decided to generate a knockout of the complete *pqi* operon.

The *pqiABC* knockout was generated using the Datsenko and Wanner method for gene inactivation within the *E. coli* K-12 BW25113 background (Datsenko and Wanner, 2000). Briefly, the kanamycin resistance (*aph*) cassette from the pKD4 plasmid was amplified using the pqiA_ko_Fwd and pqiC_ko_Rev primers to yield a linear fragment with flanking regions homologous to the regions up- and down-stream of the *pqi* operon. The WT *E. coli* K-12 BW25113 strain was first transformed with the pACBSCE plasmid, harbouring the λ-red recombination machinery, before introducing the linear *aph* cassette fragment via electroporation enabling recombination to occur. Resulting kanamycin resistant colonies were screened via colony PCR, utilising the delABC_comf_Fwd and delABC_comf_Rev primers, with a reduction in product size from ~4 kb to ~2 kb indicating successful replacement of the operon (Fig. 68).

Finally, the *aph* cassette containing strain (ΔABC::*aph*) was transformed with the pCP20 plasmid, harbouring the FLP recombinase, allowing removal of the *aph* cassette and production of the final *E. coli* K-12 BW25113 Δ*pqi*ABC strain. Removal of the cassette was confirmed via colony PCR, utilising the same check primers, indicated by a further reduction in product size from ~ 2 kb to ~ 700 bp (Fig. 68).

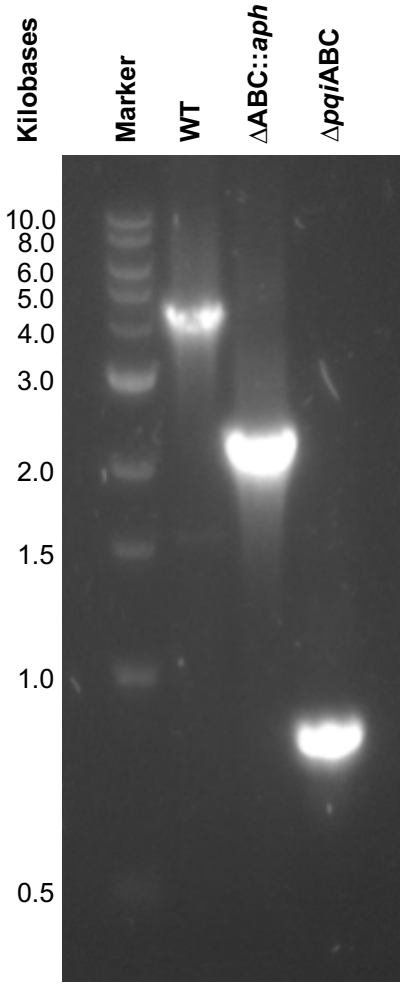


Fig. 68 – *E. coli* K-12 BW25113 Δ*pqiABC* knockout generation.

Gel electrophoresis of colony PCR products confirming the generation of thee Δ*pqiABC* knockout strain. Lane 1 – marker, lane 2 – WT, lane 3 - Δ*ABC::aph*, lane 4 - Δ*pqiABC*.

Pqi Mutation Selection

Phenotypic complementation assays were performed with single mutations in each of the three Pqi components. Residues for mutation were selected based on a range of motives, including structural similarity, existing literature and our own structural models. The residues selected within each *pqi* operon component are introduced below including justification for their mutation.

It should be noted that specific mutations have been established within the literature for probing *E. coli* MCE protein functions (Ekiert et al., 2017; Isom et al., 2020), and thus we utilised these to study PqiB. In contrast, to the best of our knowledge, no mutagenesis studies have been performed in either PqiA or PqiC, nor any of their homologues. Consequently, we decided to mutate our chosen residues within PqiA and PqiC to alanine, reflecting that performed in standard scanning mutagenesis studies (Cunningham and Wells, 1989). Whilst alternative residues may be utilised during scanning mutagenesis if particular hypotheses are to be tested, alanine is typically utilised owed to its uncharged nature, small size and effective imitation of side chain pruning beyond the β carbon (Gray et al., 2017).

PqiA Mutation Selection

Rationale for the mutations within PqiA arose from structural similarities between the predicted AlphaFold structure of PqiA (AF-P0AFL9-F1) and that of the *Pyrococcus furiosus* multidrug and toxic compound extrusion (PfMATE) transporter, discussed in more detail below (Fig. 69).

PfMATE comprises 12 transmembrane helices split into two pseudo-symmetrical lobes. Most MATE transporters harbour a highly conserved, central acidic residue essential for transport. Indeed, PfMATE possesses Asp41 within TM1 of the N-lobe, the protonation of which is proposed to induce bending of the TM1 helix, resulting in the extrusion of bound cargo to the extracellular matrix (Fig. 69). This notion was first proposed following the examination of two PfMATE X-ray structures, displaying the two conformations, which were grown at differing pH (Fig. 69) (Tanaka et al., 2013). Bowing of the TM1 helix was observed to occur about Pro26 and Gly30 with subsequent complementation assays, in addition to the structure of a P26A mutant, confirming the importance of Pro26 for TM1 bending, however, the requirement for Gly30 remains unclear (Tanaka et al., 2013).

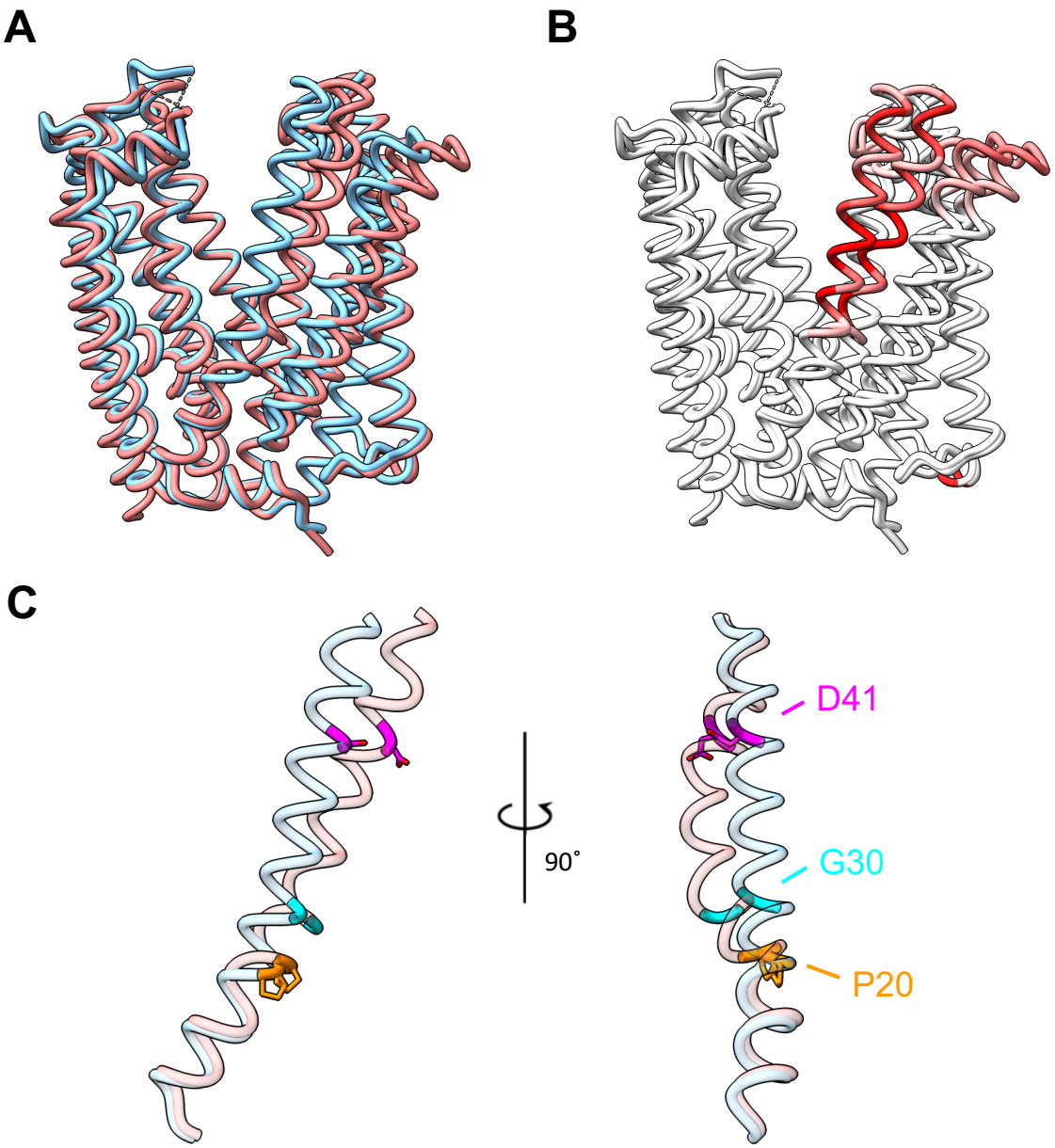


Fig. 69 – Bending of PfMATE TM1.

(A) Aligned cartoon representations of PfMATE transporter in the straight (coral) and bent (blue) conformations. (B) Aligned cartoon representations of the PfMATE transporter in the straight and bent conformations indicating the deflection of the TM1 helix. (C) Cartoon representation of the PfMATE TM1 in the straight (coral) and bent (blue) conformations. Pro26, Gly30 and Asp41 are coloured orange, cyan and magenta respectively. PDB: 3VVN and 3VVO.

Examination of the Pfam hidden Markov model (HMM) logo for the PqiA domain (PF04403) revealed a highly conserved Asp/Glu residue towards the centre of the domain. Indeed, comparison of the PfMATE structures with the AlphaFold predicted structure of PqiA, led to the identification of two acidic residues, Glu149 and Asp358, each residing within separate central helices, at approximately the same height relative to the membrane as PfMATE Asp41. Furthermore, both residues are preceded by glycine, Gly146 and Gly352 respectively, indicating the potential for flexibility within the helix, however, there was a distinct absence of proline from either helix (Fig. 70). Subsequent conservation analysis, via the ConSurf server (Landau et al., 2005), revealed Glu149, Gly352 and Asp358 to be highly conserved whilst Gly146 was moderately conserved itself (Fig. 71). The central positioning and conservation of these four residues raised speculation of their involvement in the function of PqiA, thus we decided to probe these residues via complementation assays.

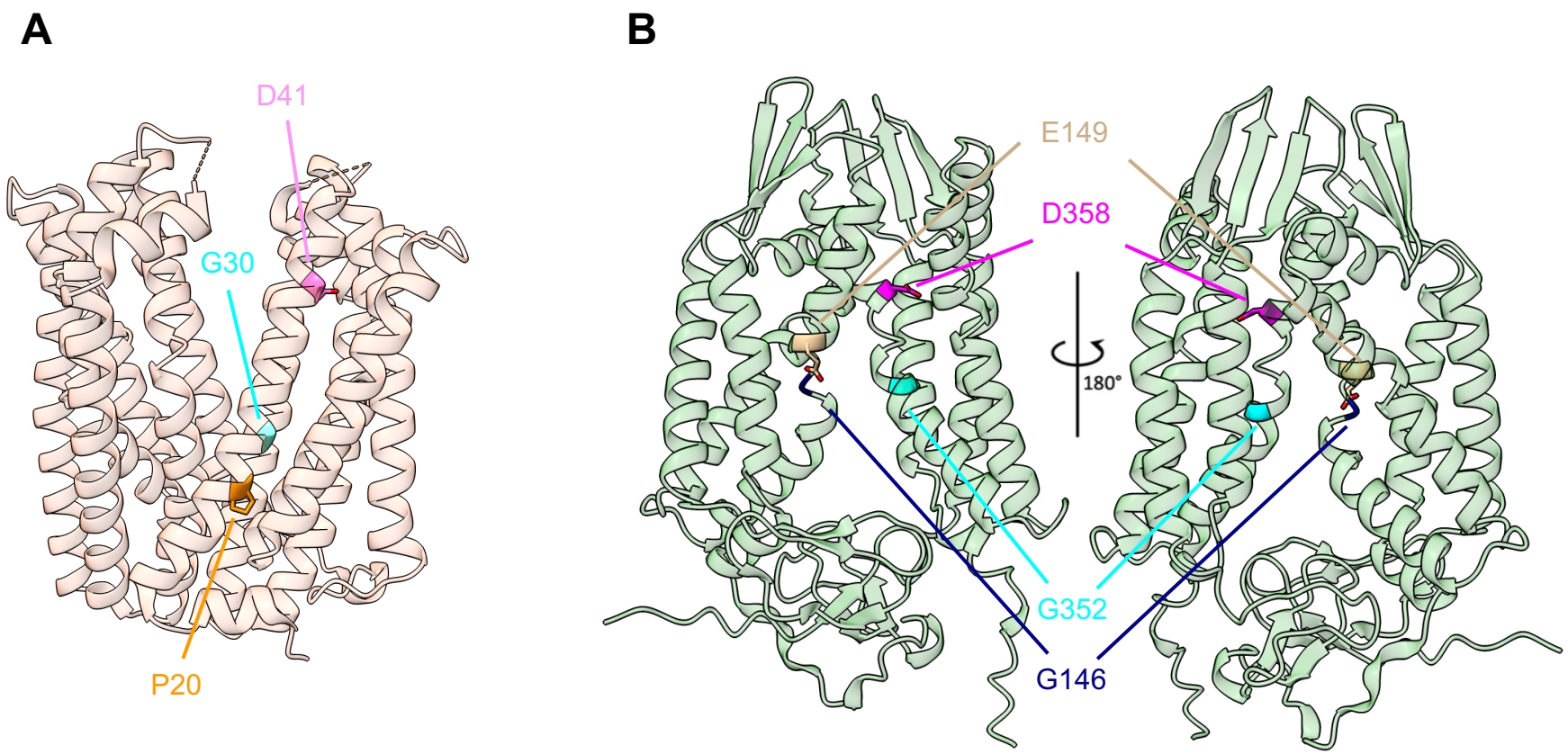


Fig. 70 – Comparison of PfMATE and PqiA.

(A) Cartoon representation of the PfMATE transporter crystallised in the straight conformation (PDB: 3VVN). Pro26, Gly30 and Asp41 are coloured orange, cyan and pink respectively. (B) Cartoon representation of the PqiA AlphaFold prediction (AF-P0AFL9-F1), highlighting the potential similarity to the MATE family of transporters. The four residues selected for mutation are indicated in the following colours: Gly146 – navy, Glu149 – tan, Gly352 – cyan, Asp358 – magenta.

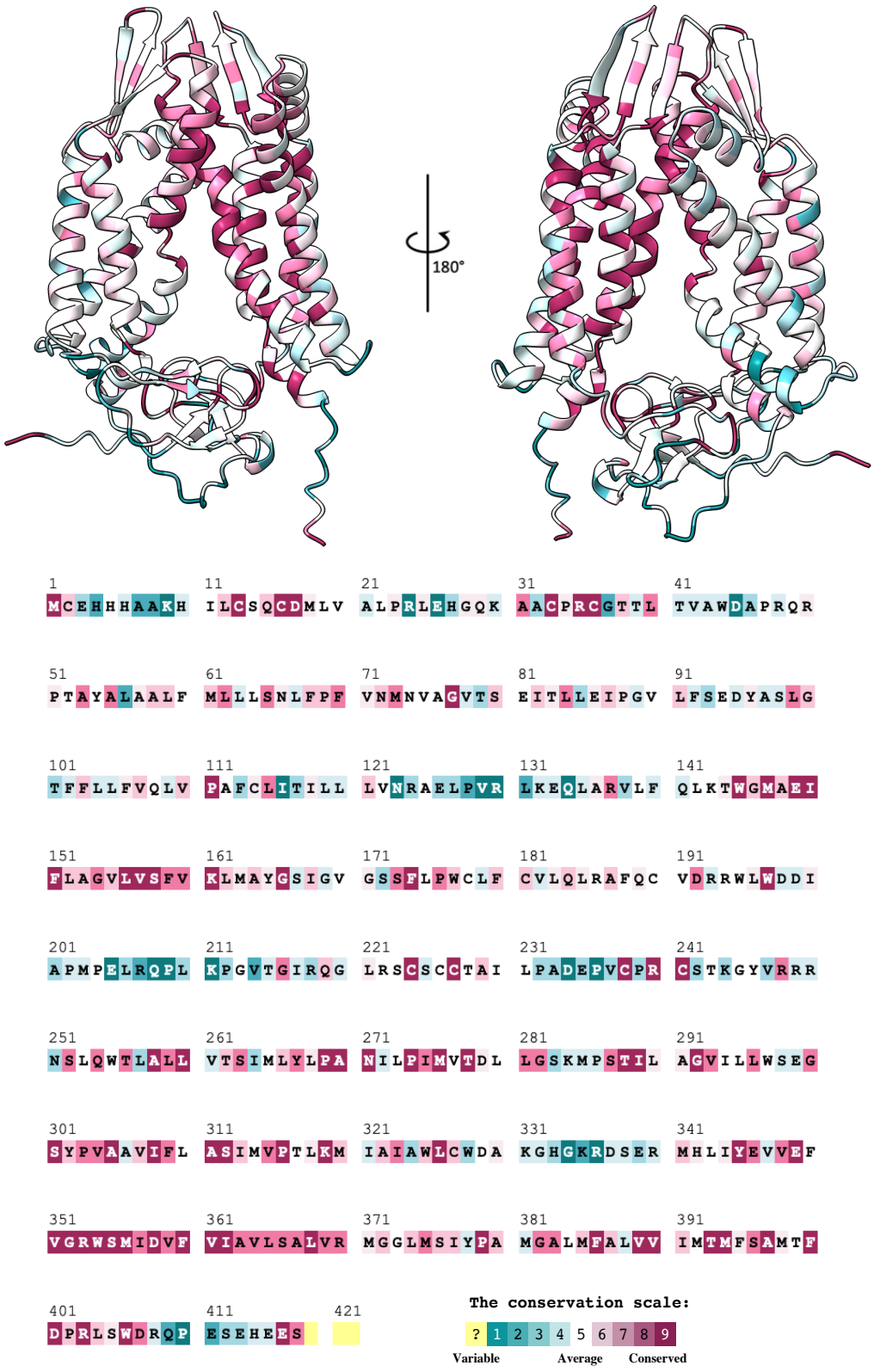


Fig. 71 – ConSurf conservation analysis of PqiA.

(A) Orthogonal views of the AlphaFold PqiA prediction (AF-P0AFL9-F1) displayed as cartoon representation with residues coloured according to ConSurf sequence conservation. (B) PqiA sequence coloured according to Consurf sequence conservation.

PqiB Mutation Selection

Mutations within the PqiB component were focussed upon specific residues within the pore lining loops (PLLs) of each MCE domain. The PLLs of each YebT MCE domain have previously been shown to possess a conserved hydrophobic motif within the short helical region at the base of each loop (Fig. 72) (Isom et al., 2020). The motif comprises three residues arranged in the following pattern ΦxxΦΦ - where Φ indicates a hydrophobic residue and x any residue. Phenotypic complementation assays, involving the substitution of each motif residue to Asn, revealed these residues to be crucial for the function of YebT (Isom et al., 2020; Vieni et al., 2022).

A similar motif was also reported in the PLL of the single MlaD MCE domain comprising two consecutive hydrophobic residues located at the tip of the PLL hairpin (Fig. 72) (Ekiert et al., 2017). Once again phenotypic complementation studies, in which both residues were substituted for Asn simultaneously, proved these residues to be critical for MlaD function (Ekiert et al., 2017).

PqiB displays a combination of both aforementioned motifs with the PLLs of the first and second MCE domains presenting the same ΦxxΦΦ motif observed in YebT (L119, L122, L123, L240, L243, L244), whilst the third MCE harbours the hairpin motif of MlaD (L377, V378) (Fig. 72). Despite this interesting combination of motifs, phenotypic complementation studies of specific PqiB residues remain unpublished thus we decided to perform these assays, probing these conserved residues, within this study. In accordance with published MCE protein PLL complementation assays, all residues were mutated to asparagine.

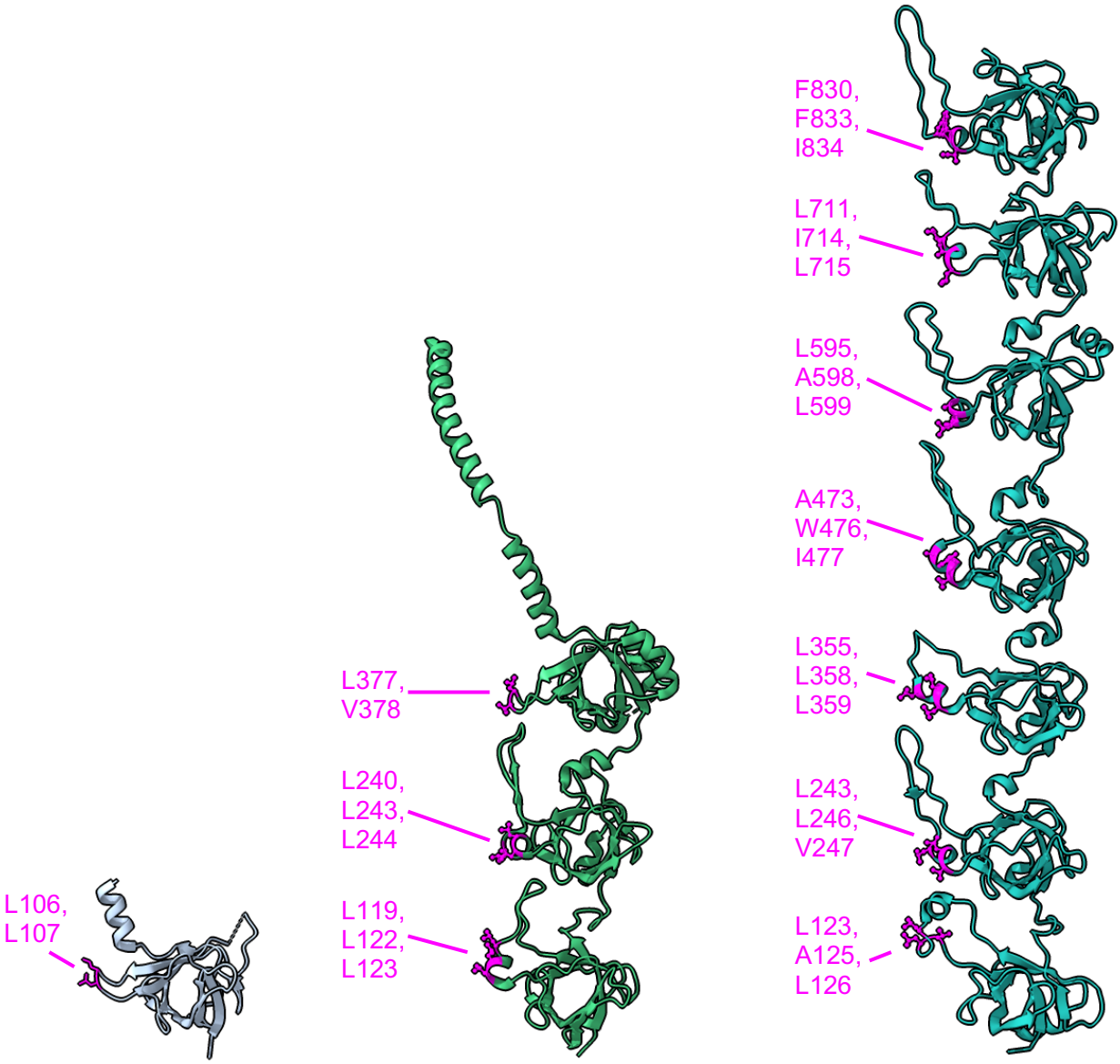


Fig. 72 – Conserved hydrophobic motifs in *E. coli* MCE protein PLLs.

(Left) Cartoon representation of a single MlaD protomer with the two consecutive hydrophobic residues at the tip of the PLL hairpin depicted in magenta. (Middle) Cartoon representation of a single PqiB protomer with the conserved PLL hydrophobic residues depicted in magenta. (Right) Cartoon representation of a single YebT protomer with the PLL hydrophobic residues comprising the ΦxxΦΦ motif depicted in magenta. (PBDs: 5UW2, 5UVN and 6V0C).

PqiC Mutation Selection

PqiC mutations arose directly from residues of interest identified in our CFL2 structure hence the exhaustive rationale for their selection may be found in Chapter 3. A brief justification for each of the four residues selected for mutation (Lys23, Trp47, Asp55, Asp84) may be found below.

Lys23 resides upon the membrane occluded face of the PqiC octamer generating eight regions of positive charge around the inner circumference of this face (Figs. 29, 30). Considering the occlusion, and potential imbedding of this face with the membrane, the role of this charged residue seems unclear. The potential for interactions with polar lipid head groups, or indeed additional proteinaceous components, remain strong possibilities.

Trp47 lines the PqiC octameric pore and gives rise to the lower of two lipophilic band regions which line the PqiC octamer pore at distinct heights (Figs. 27, 28). The precise localisation of these two distinct band regions, and the lack of any other lipophilic residues within the PqiC pore, imply involvement of these residues in either complex function or assembly.

Both Asp55 and Asp84 also line the PqiC octameric pore where they generate the top of three bands of negative electrostatic potential found within the pore surface (Figs. 29, 30). This top region is of particular interest as it perfectly bisects the two bands of lipophilicity previously described, hence, the juxtaposition of these adjacent regions is intriguing. Consurf analysis indicated Asp55 to be amongst the most highly conserved residues within PqiC (conservation score 8). Furthermore, our CFL2 structure indicates the side chain of Asp55 to be particularly prominent, extending towards the centre of the octameric pore. Much like Trp47 above, the precise location, as well as the striking conformation of Asp55, suggest involvement of these residues in complex function or assembly.

The location of all four selected residues in the context of the CFL2 structure and PqiC octamer is displayed below (Fig. 73).

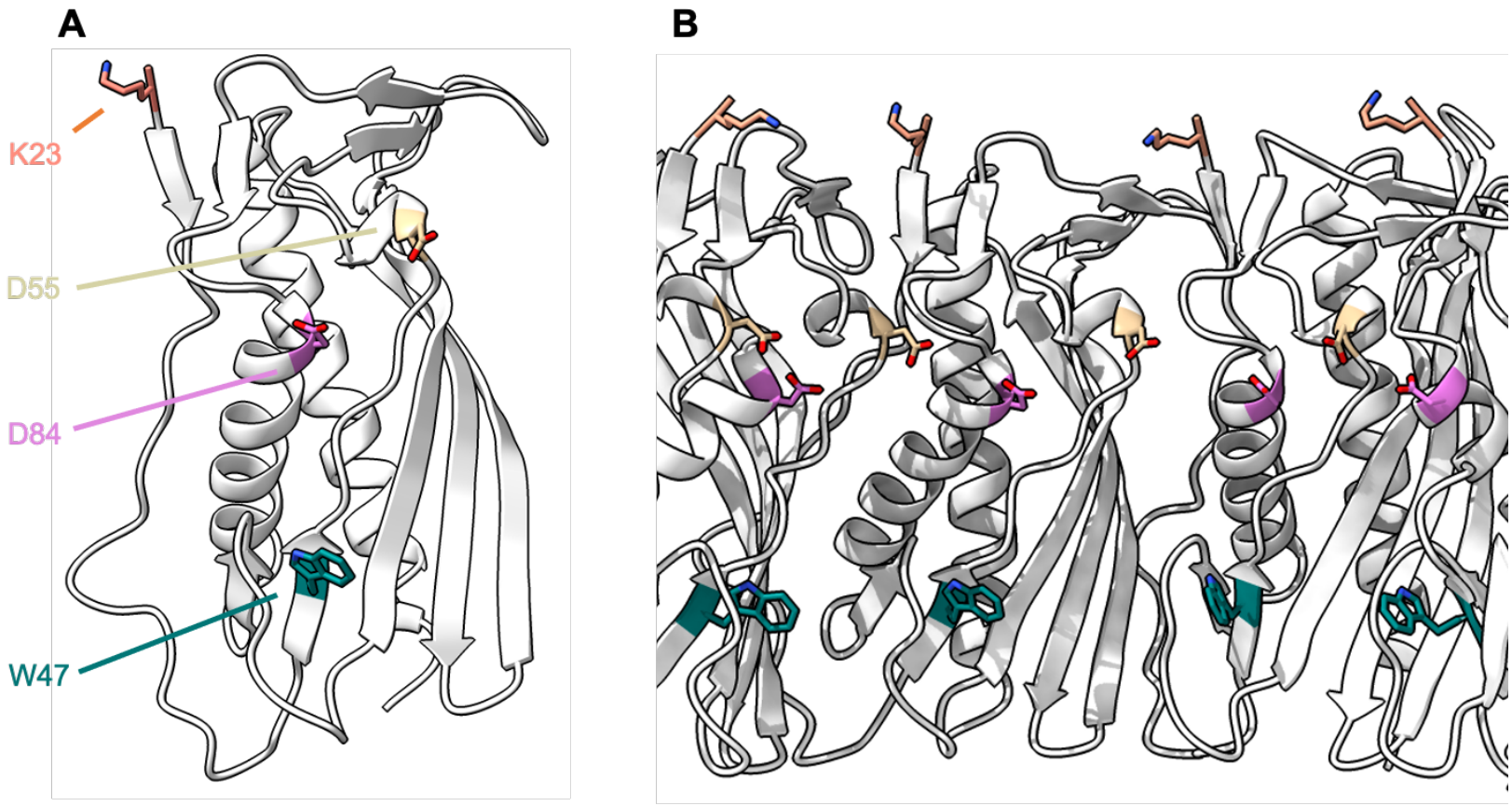


Fig. 73 – PqiC complementation assay residues.

(A) Cartoon representation of the CFL2 structure with Lys23, Trp47, Asp55 and Asp84 depicted in salmon, teal, tan and magenta respectively. (B) Slab cartoon representation of the PqiC Octamer viewed from inside the pore. Lys23, Trp47, Asp55 and Asp84 depicted in salmon, teal, tan and magenta respectively.

Phenotypic Complementation Assay Mutations and Assays

All assays utilised the same three controls: 1 – WT *E. coli* K-12 BW25113 (hereafter referred to as WT), 2 – *E. coli* K-12 BW25113 Δ*pqiABC* knockout (hereafter referred to as ΔABC), and 3 - ΔABC strain transformed with a plasmid harbouring the WT *pqiABC* operon (hereafter referred to as ΔABC + ABC). All investigated mutations constituted single point mutations. Unless otherwise stated, all selected residues were mutated to alanine in accordance with canonical alanine-scanning mutagenesis.

PqiA Phenotypic Complementation Assay

Initially we performed complementation assays to probe mutations within the PqiA component of the *pqi* operon. Growth on LB agar revealed the growth of all strains to be comparable (Fig.74A) whilst differential growth on LB agar supplemented with LSB separated our mutations into three classes (Fig. 74B).

The first of these classes comprised the G146A mutation only, the growth of which was observed to be similar to that of the WT and ΔABC + ABC strains implying the mutation to have minimal or no impact upon Pqi function. The second class encompassed both the E149A and G352A mutations which displayed compromised growth with colonies constrained within the 10^{-2} dilution in comparison to the G146A mutation which displayed growth at the 10^{-3} dilution. Consequently, the E149A and G352A LSB phenotypes are suggestive of a moderate disruption to Pqi function. Finally, the D358A mutation resides within a class of its own owing to its severe LSB phenotype, in which growth is constrained to the 10^{-1} dilution only, indicating a significant disruption to Pqi function.

Western blotting revealed comparable expression of all PqiA mutants confirming any observed differences in viability to arise from compromised Pqi function (Fig. 74C).

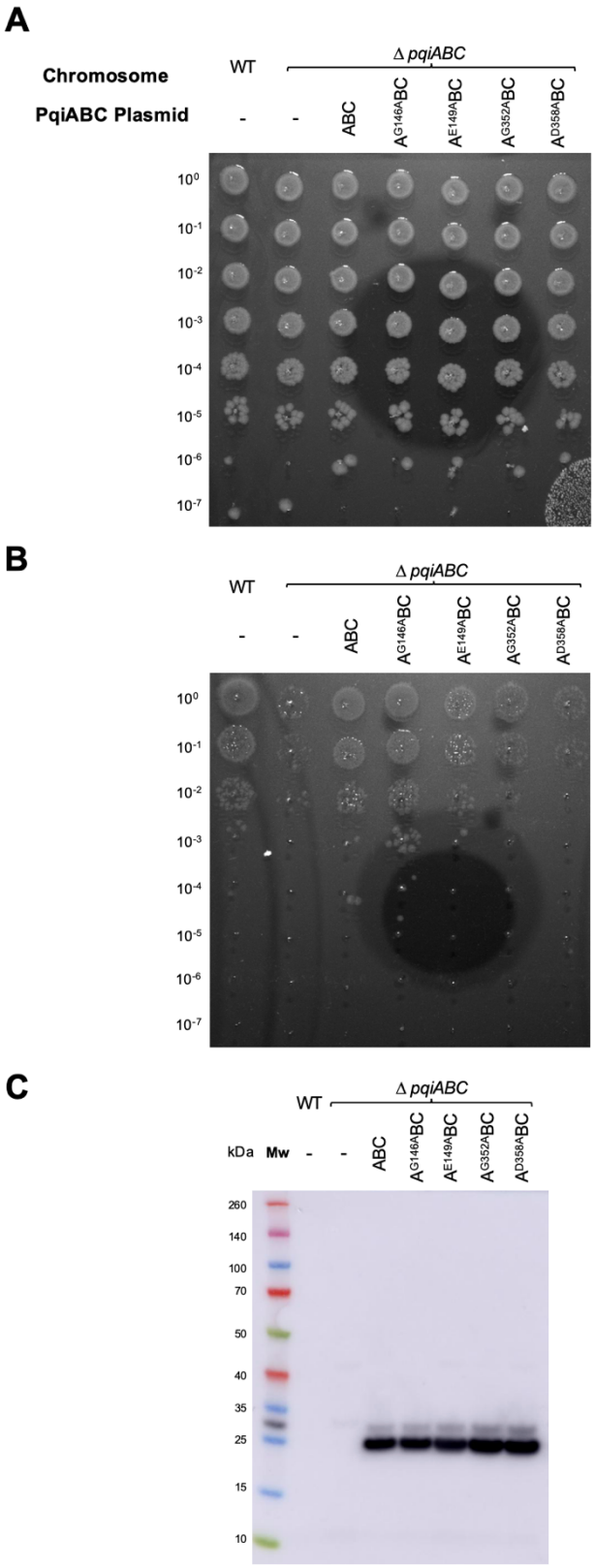


Fig. 74 – PqiA phenotypic complementation assays.

(A) Growth of PqiA strains on LB Agar. (B) Growth of PqiA strains on LB agar supplemented with 0.25% LSB. (C) Anti-His western blot indicating expression of WT and PqiA mutant ABC constructs, detected using ECL reagent.

PqiB Phenotypic Complementation Assay

Next, we probed mutations within the pore lining loops (PLLs) of the PqiB component. Once again, all strains displayed comparable growth upon LB agar (Fig. 75A), however, growth of all PLL residue mutants was diminished when grown on LB supplemented with LSB (Fig. 75B). Indeed, whilst growth of the WT and ΔABC + ABC strains was observed in the 10^{-3} dilution, that of all PqiB mutants was near identical, being constrained to the 10^{-2} dilution or below, resembling that of the ΔABC strain. The reduced growth observed for all PqiB PLL mutations is indicative of significant disruption to Pqi function.

Once again, western blotting revealed comparable expression of all PqiB mutants confirming any observed differences in viability to arise from compromised Pqi function (Fig. 75C).

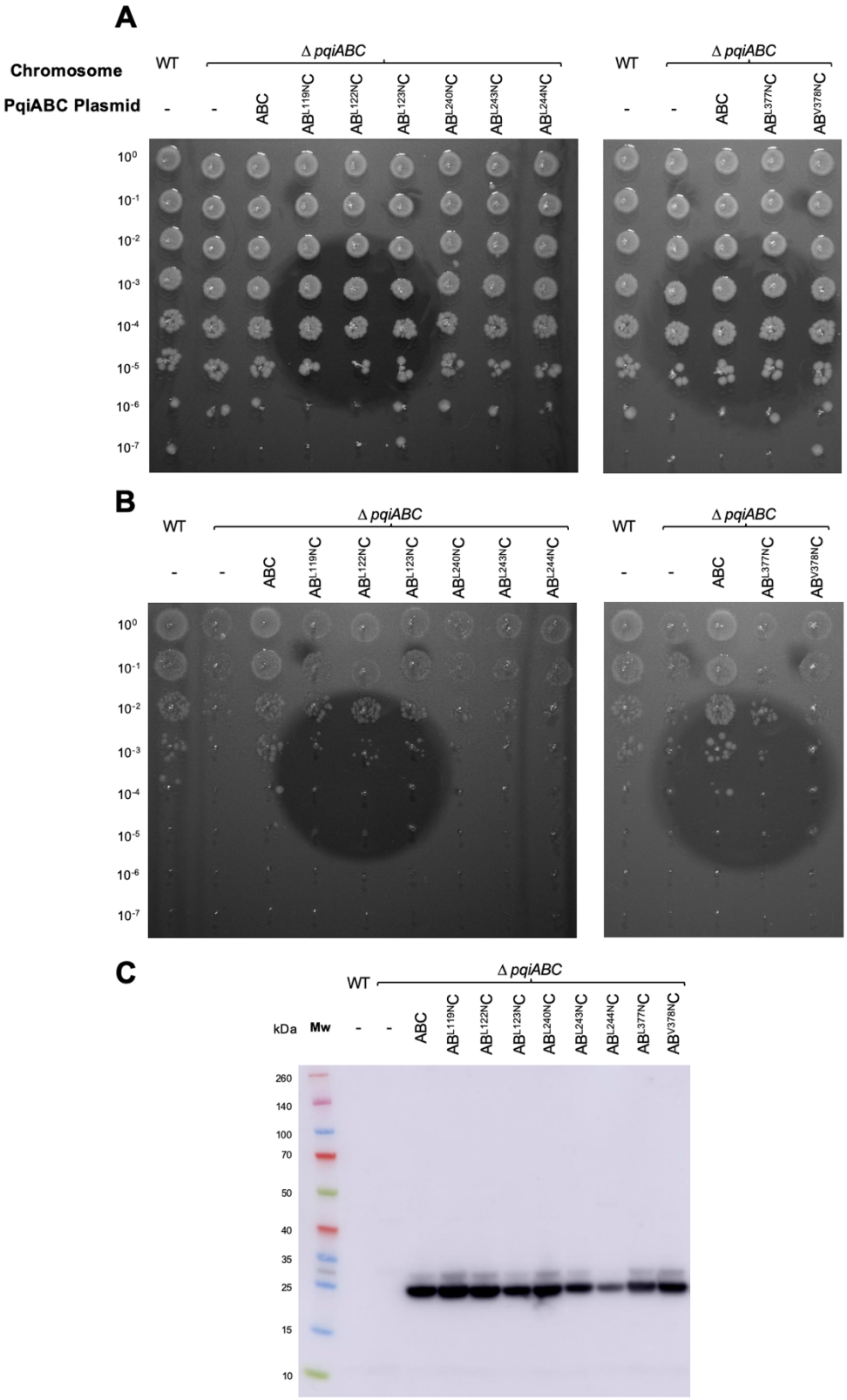


Fig. 75 – PqiB phenotypic complementation assays.

(A) Growth of PqiB strains on LB Agar. (B) Growth of PqiB strains on LB agar supplemented with 0.25% LSB. (C) Anti-His western blot indicating expression of WT and PqiB mutant ABC constructs, detected using ECL reagent.

PqiC Phenotypic Complementation Assay

Finally, we probed mutations within the PqiC component of the *pqi* operon. Once again, all strains displayed comparable growth upon LB agar (Fig. 76A). Contrary to that observed in the PqiA and PqiB assay, when grown on LB agar supplemented with LSB, all PqiC mutant strains displayed growth comparable to that of the WT and ΔABC + ABC strains implying minimal or no impact upon Pqi function (Fig. 76B).

Western blotting revealed comparable expression of all PqiC mutants confirming any observed differences in viability to arise from compromised Pqi function (Fig. 76C).

Considering the apparent ineffectiveness of all our PqiC mutations, we considered the possibility that PqiC itself was superfluous to Pqi function. To test this hypothesis, we attempted to complement the ΔABC strain with a plasmid containing only the *pqiA* and *pqiB* components of the operon (ΔABC + AB strain). Whilst growth of the strain was similar to that of the WT on LB agar (Fig. 77A), when grown on LB agar supplemented with LSB its growth resembled that of the ΔABC strain indicating its inability to complement the knockout and suggested PqiC to be indispensable (Fig. 77B).

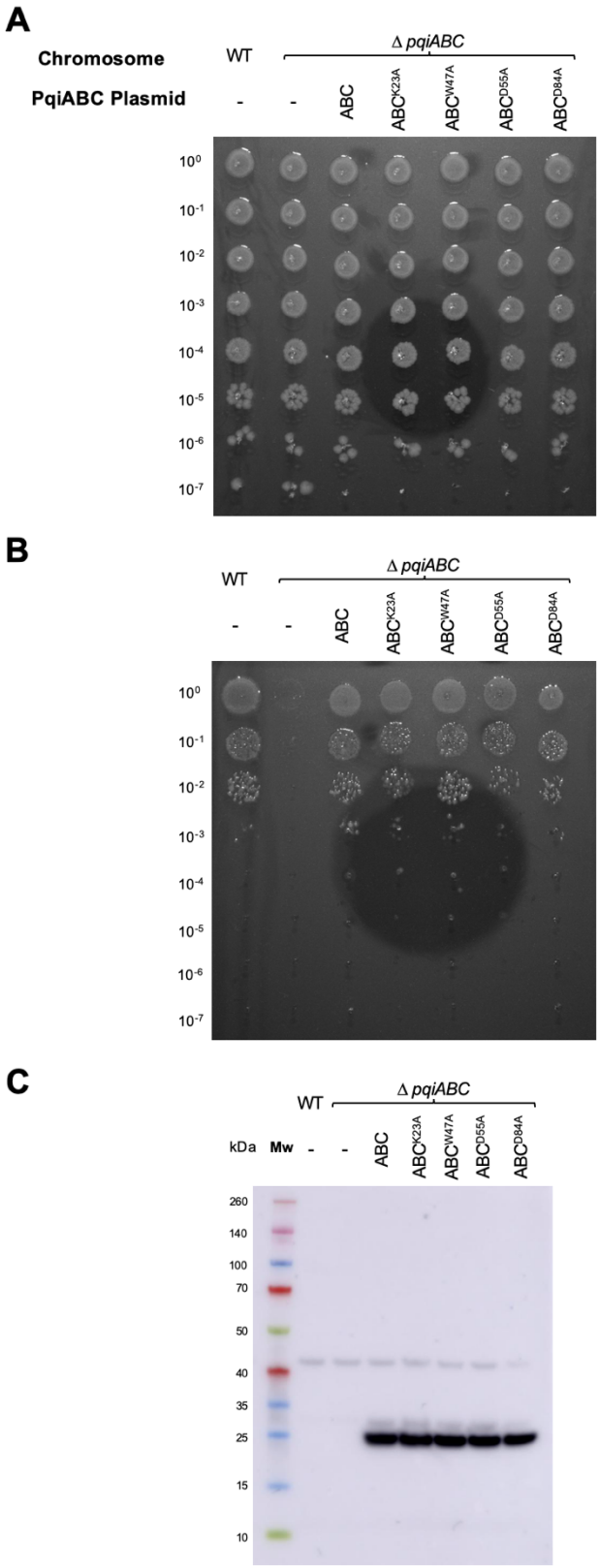


Fig. 76 – PqiC phenotypic complementation assays.

(A) Growth of PqiC strains on LB Agar. (B) Growth of PqiC strains on LB agar supplemented with 0.25% LSB. (C) Anti-His western blot indicating expression of WT and PqiC mutant ABC constructs, detected using ECL reagent.

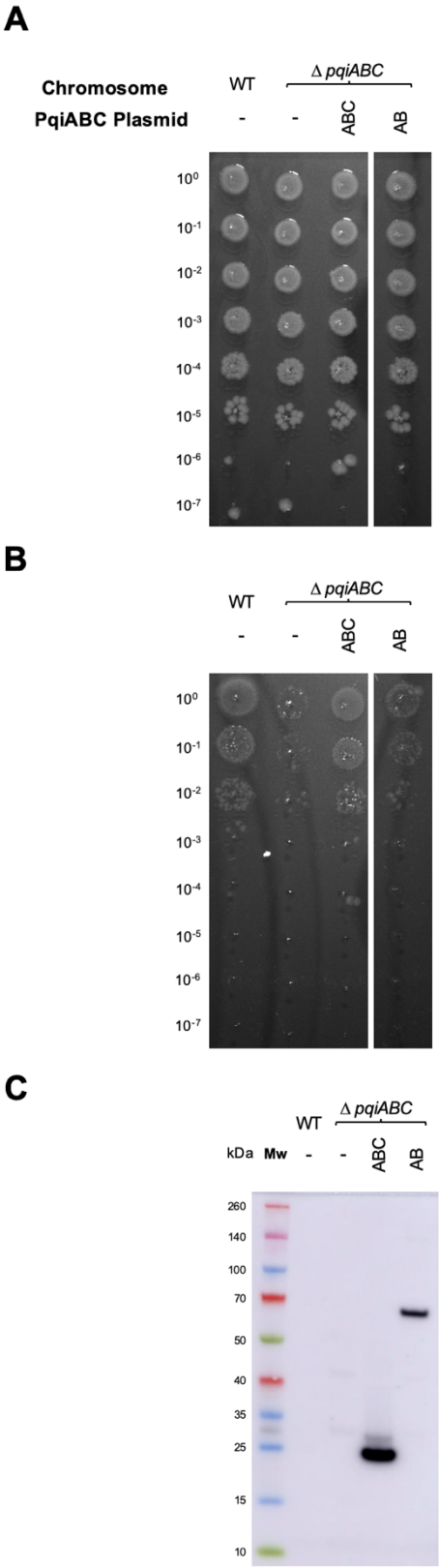


Fig. 77 – *pqiAB* phenotypic complementation assay.

(A) Growth of PqiAB strain on LB Agar. (B) Growth of PqiAB strains on LB agar supplemented with 0.25% LSB. (C) Western blot indicating expression of WT and PqiAB constructs, detected using ECL reagent.

Discussion

PqiA

The multidrug and toxic compound extrusion (MATE) transporter family typically possess a key central acidic residue. Assessment of two *Pyrococcus furiosus* MATE (PfMATE) crystal structures, grown at differing pH's, first revealed the potential role of the conserved central acidic residue (Asp41) (Tanaka et al., 2013). Upon protonation, Asp41 is anticipated to reposition to enter a hydrogen bonding network, comprising residues on neighbouring helices, resulting in the bending of the TM1 helix at Pro26. Deflection of the TM1 helix purportedly collapses a cavity within the N-lobe, resulting in the extrusion of cargo (Tanaka et al., 2013).

Curving of the TM1 helix has proved controversial with some initially doubting its physiological relevance, instead proposing it arose due to interactions with the exogenous lipid monoolein added during the crystallisation process (Radchenko et al., 2015). The subsequent crystallisation of the *Vibrio cholerae* VcmN transporter at differing pH, revealing analogous remodelling of the TM1 in response to protonation, and incorporation of the central Asp35 into a hydrogen bond network, reinforced the initial notion (Kusakizako et al., 2019). Finally, studies utilising fluorescence quenching (Jagessar et al., 2019) and double-electron electron resonance (DEER) spectroscopy (Claxton et al., 2018; Jagessar et al., 2020) have re-iterated the existence of a conserved, polar residue network, within the N-lobe, mediating ion coupled conformational changes.

Glu149 and Asp358 within PqiA were selected based on their potential to emulate this crucial residue, and thus potentially indicate the mechanism by which PqiA mediated transport proceeds. Mutation of both acidic residues to alanine yielded a strong phenotype when grown on media supplemented with LSB, indicating a severe disruption to PqiA function.

Considering the apparent necessity of a polar residue network within close proximity of the central acidic residue in MATE transporters, we decided to probe the residues surrounding Glu149 and Asp358.

Hypothesising that such a network would be conserved, we focussed our search to highly conserved residues. It became immediately obvious that the C-lobe of PqiA, comprising Asp358, contains many more highly conserved residues than its N-lobe counterpart which contained Glu149. Furthermore, the most highly conserved residues within the N-lobe occurred within the region of closest proximity to the C-lobe, potentially enabling their interaction with C-lobe residues and implicating the C-lobe of PqiA to be more influential upon function (Fig. 71). Within the PqiA C-lobe we identified a group of extremely conserved, polar residues (Tyr267, Asn271, Ser312, Ser355) within proximity of Asp358 raising the potential for them to conceivably assemble a hydrogen bonding network analogous to that within the PfMATE N-lobe (Fig. 78). Furthermore, mutation of the preceding Gly352, which potentially forms a hinge point within the Asp358 helix, also produced a severe phenotype indicating its importance for PqiA function. Collectively, the phenotypes of Gly352 and Asp358 and presence of a highly conserved group of polar residues within the PqiA C-lobe, raise the speculative possibility of a ‘MATE-like’ mechanism in which protonation of Asp358 leads to helical rearrangements to eject cargo. Indeed, the polar residues identified present ideal targets for follow up studies.

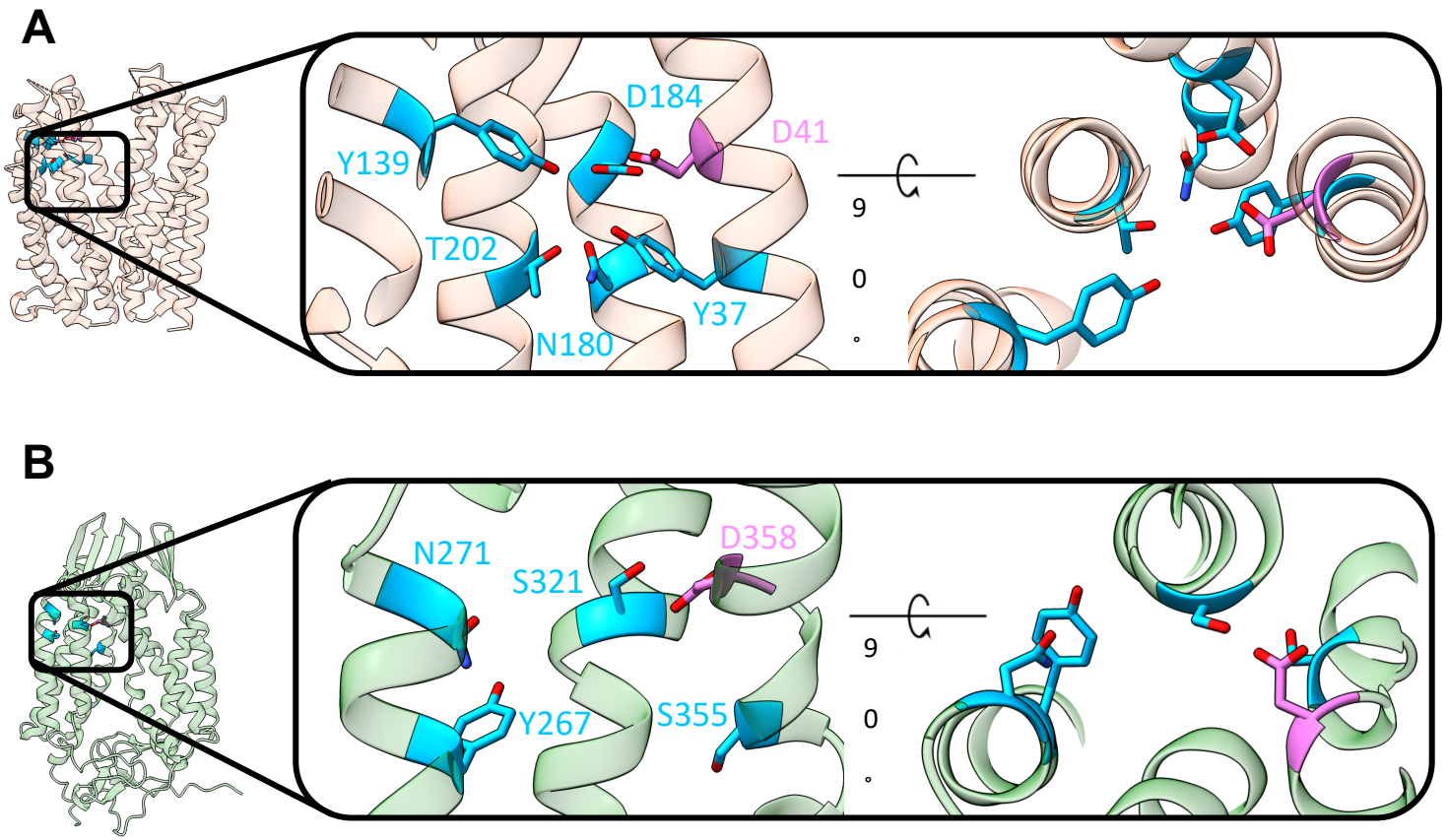


Fig. 78 – Comparison of the PfMATE N-lobe and PqiC C-lobe.

(A) Orthogonal views of the PfMATE N-lobe network of residues believed to be crucial for ion mediated conformational changes of the TM1 helix. The crucial D41 is depicted in magenta whilst Y37, Y139, N180, D184 and T202 are depicted in blue (PDB: 3VVN). (B) Orthogonal views of the AlphaFold PqiA prediction C-lobe demonstrating the potential for an analogous polar network to exist. D358 is depicted in magenta whilst Y267, N271, S312 and S355 are depicted in blue.

Conversely, postulating the involvement of Glu149 in PqiA function is less straightforward. In contrast to Gly352, the preceding Gly146 displayed no phenotype when grown upon agar supplemented with LSB. This apparent indication of redundancy is perhaps not completely surprising given the reduced conservation of Gly146 compared to the other three mutated residues. The severe phenotype and high degree of conservation imply Glu149 as a crucial residue for PqiA function, however, our current lack of knowledge prevents us from speculating its specific involvement at this stage.

It should be noted that both the selection of PqiA residues for mutation and subsequent postulation of their function is largely based upon the AlphaFold predicted PqiA structure (AF-P0AFL9-F1). Hence, whilst AlphaFold provides an invaluable tool for structural biologists, there remains the potential for over interpretation, particularly in the case of PqiA where experimentally derived structures of homologous proteins remain enigmatic. Our concurrent use of conservation analysis provides some mitigation against over interpretation, however, we await the deposition of confirmatory experimental structures with anticipation.

PqiB

Hydrophobic residues within the pore lining loops of *E. coli* MCE proteins have been postulated to be essential for correct function. This notion was first observed for MlaD where mutation of two hydrophobic residues (Leu106 and Leu107) residing at the tip of its single PLL hairpin loop proved critical for function (Ekiert et al., 2017). Subsequently, a distinct hydrophobic motif, ΦxxΦΦ - where Φ indicates a hydrophobic residue and x any residue, was identified within the basal helix of all YebT PLLs (Isom et al., 2020). Complementation assays revealed mutation of these residues to completely abolish YebT function, except for A598 which displayed only diminished activity, once again reinforcing the importance of these residues in MCE protein functionality (Isom et al., 2020; Vieni et al., 2022).

PqiB is unique amongst the MCE proteins of *E. coli* in that its three MCE domains display a combination of the hydrophobic motifs present in MlaD and YebT. Indeed, whilst the first two PqiB MCE domains comprise the hydrophobic motif observed within all YebT MCE domains, the third MCE domain contains two neighbouring hydrophobic residues at the tip of the PLL hairpin, akin to those of MlaD. These two distinct arrangements of hydrophobic residues potentially indicate differences in the function of the three MCE domains hence we chose to probe all these residues via complementation.

Perhaps unsurprisingly given the existing literature, all our PqiB PLL mutants displayed severe growth defects when grown upon LB agar supplemented with LSB indicating a severe disruption to Pqi function. Whilst this confirms the integral involvement of these residues in Pqi function, these results, along with those of all previous complementation assays involving MlaD and YebT, are unable to discern exactly how these mutations affect MCE protein function. Indeed, the observed, impeded growth may arise either due to disruption or alternatively the deregulation of Pqi mediated transport. Which of these two scenarios is correct may be easily identified via an *in vitro* transport assay, however, as yet no such

assay has been developed for either the Pqi or Yeb systems. Whilst liposomal transport assays have been developed for the Mla system (Low et al., 2021; Tang et al., 2020), their utilisation thus far has focussed exclusively upon probing directionality rather than investigating the impact of specific mutations upon transport. The development of transport assays enabling the impact of certain residues upon MCE protein mediated transport to be assessed, particularly PqiB and YebT, presents an obvious means of advancing our understanding of these systems. Thus, such assays must be seen as a priority within the scientific community.

PqiC

We also mutated four PqiC residues which we anticipated to be involved in either complex formation or Pqi mediated transport based upon their location within our CFL2 structure, lipophilic/electrostatic potential and sequence conservation. Curiously, none of our mutations yielded a discernible growth impediment when cultured upon agar supplemented with LSB indicating no disruption of Pqi function. Nevertheless, our inability to complement the Δ*pqi*ABC strain with the *pqiAB* containing plasmid implies PqiC to be indispensable.

Given the apparent essentiality of PqiC, the ineffectiveness of our four selected mutations presumably arises due to one or more of three possibilities: 1 – the mutations performed were not severe enough. 2 – PqiC doesn't form intimate interactions with PqiB. 3 – the octameric structure of PqiC is not biologically relevant.

All the residues selected for mutation were shown to be conserved, with Asp55 being particularly highly conserved implying its importance. Furthermore, three of the four selected residues (W47, D55, D84) were located within the pore of the PqiC octamer, where we anticipate interactions with PqiB. The final residue (K23) generates an intriguing region of positive charge upon the membrane occluded face of the PqiC octamer. Considering the prominent positioning of our selected residues, we anticipated that their simple mutation to alanine would be sufficient to cause significant disruption, however, this was not the case and more severe substitutions may be needed to exert an effect. Indeed, three of our four selected residues exhibit electrostatic potential thus charge reversal mutations may be more appropriate to induce disruption. Substitution of both D55 and D84 for lysine, both independently and concurrently, may elicit greater functional disruption. Furthermore, introducing electrostatic potential in place of the lipophilic W47 may disrupt hydrophobic packing interactions. It must be noted that the selection of residues was largely based upon the assumption of an octameric biological PqiC assembly. Whilst unlikely, the potential for

PqiC to adopt a different architecture when in complex with PqiB remains and thus different residues to those mutated may constitute the interaction interface. Finally, the introduction of cysteine residues with both PqiB and PqiC may provide a method of stabilising their interaction via the formation of disulphide bonds. Whilst this notion presents an attractive proposition, its success is dependent upon overcoming two primary considerations. Firstly, it is unclear how well disulphide bonds between PqiB and PqiC would form considering their differing anticipated symmetries. Secondly, the mystery regarding the structure of the PqiB C-terminus complicates the positioning of a cysteine so that it is able to interact with one placed within PqiC.

Summary

In this chapter we have utilised phenotypic complementation assays to infer the relative importance of specific residues within all three components of the *pqi* operon. Our PqiA mutations suggest the complex may function via a protonation mediated conformational change mechanism analogous to that of MATE transporters. Mutations within PqiB confirm the essentiality of specific PLL hydrophobic motifs for correct function, homologous to those in MalD and YebT, however, fail to reveal the nature of their involvement. Finally, our mutations in PqiC proved to be inconsequential, likely due to their minimal potency, thus more severe charge reversal mutations are required to gain further insight. Consequently, we investigated the potential dispensability of PqiC, however, the inability of a *pqiAB* construct to complement loss of the entire *pqiABC* plasmid implied PqiC to be essential. Indeed, all these results require further investigation and would benefit greatly from an *in vitro* assay where the intricacies of Pqi mediated transport may be revealed.

Chapter 6: Conclusions and Future Work

Conclusions and Future Work

In this study we have provided evidence advancing our understanding of the *E. coli* Pqi pathway. We identified three distinct research aims enabling the elucidation of both structural and functional information with the intent of producing a comprehensive study.

Advancement of Pqi Structural Understanding

Prior to this study, experimentally derived structural knowledge of the *E. coli* Pqi complex amounted to a single, partially complete cryo-EM structure of a soluble PqiB construct. The only insight into *E. coli* PqiC was the crystallographic structure of *Enterobacter cloacae* PqiC (80.3% sequence identity, 88.3% similarity) whilst DeepMind's AlphaFold model (Jumper et al., 2021) provided a potential *E. coli* PqiA structure, itself requiring experimental validation. Considering the lack of existing experimental structural knowledge, the primary contribution of this study has been the determination of structures relating to both the PqiB and PqiC components.

We have presented the first experimental structures of *E. coli* PqiC, encompassing both the native, full-length protein and that of a soluble construct. Collectively, these complementary structures provide both high-resolution side chain geometry and reveal an octameric PqiC arrangement that we anticipate to be biologically relevant. Thus, this study presents a comprehensive structural insight into the *E. coli* PqiC in isolation, however, the nature of its interaction with additional complex components remains to be seen (Fig. 80).

We also presented the first experimental structure of the native *E. coli* PqiB protein, allowing comparison to the previously published soluble $PqiB^{39\text{-}564}$ structure. Our cryo-EM map allowed modelling of several small regions absent from the existing soluble structure, thus

yielding a more comprehensive model. Furthermore, the three MCE domains in our BFL model are much more dispersed than those in the previously published soluble model. We suggest this elongated architecture arises due to the interaction of both PqiB termini with micelles, causing tension along its length. The inherent flexibility of PqiB was evident within our map, revealing the dynamic nature of the PLLs but also, ultimately limiting the extent to which the coiled coil entity could be explicitly modelled. Consequently, the C-terminus of PqiB remains cryptic, an unfortunate reality as this region is anticipated to interface with PqiC and/or the outer membrane and thus remains of considerable interest (Fig. 80).

Indeed, the nature of the PqiB – PqiC interaction is intriguing considering their anticipated differing stoichiometries. One possible resolution involves PqiC acting only to stabilise the interaction between PqiB, thus negating the need for intimate interactions between the two proteins. Structural determination of the PqiB C-terminus likely necessitates reduction in the flexibility of the coiled coil moiety, either via additional stabilisation or construct design, to facilitate its alignment and thus enable accurate modelling. Accordingly, we have devised a PqiAB$^{S467\text{-}TEV}$C construct, comprising a TEV cleavage site midway along the coiled coil, which upon cleavage will yield a sub-complex facilitating the elucidation of the PqiB C-terminus and its interaction with PqiC. Assessment of this concept is ongoing following promising expression, purification, and cleavage steps. Furthermore, the remaining PqiA-PqiB$^{1\text{-}467}$ sub-complex may facilitate examination of the PqiA-PqiB interface; hence this single tool facilitates investigation of interactions at both ends of the Pqi complex. The architecture of the periplasmic spanning Pqi complex, including detailed characterisation of the component interfaces, remains a fundamental objective of future studies.

Finally, a striking feature of the PqiA sequence is the 15 cysteines it possesses, accounting for 3.6% of the 417 residues, hence the necessity for such a large, unpaired number of cysteines is intriguing (Fig. 79). Indeed, the homologous YebS itself possesses 11 cysteines implying their potential involvement in function. Whilst the location of some cysteines varies

between the two proteins, both PqiA and YebS harbour eight homologous cysteines within loop regions at the base of the protein. Furthermore, Consurf sequence conservation analysis revealed these eight cysteines to be amongst the most highly conserved residues in either protein, reinforcing their anticipated importance. The positioning of these eight cysteines within the AlphaFold predicted structures of PqiA and YebS is compelling with two clusters of four cysteines each appearing to adopt a conformation enabling the coordination of a central ion (Fig. 79). Metal coordination by PqiA presents a novel notion not previously considered within the literature, however, the potential role or indeed existence of these coordinated ions remains unclear. Thus, experimental PqiA structural determination, and investigation of its potential metal binding capabilities, present ideal avenues for considerable further investigation.

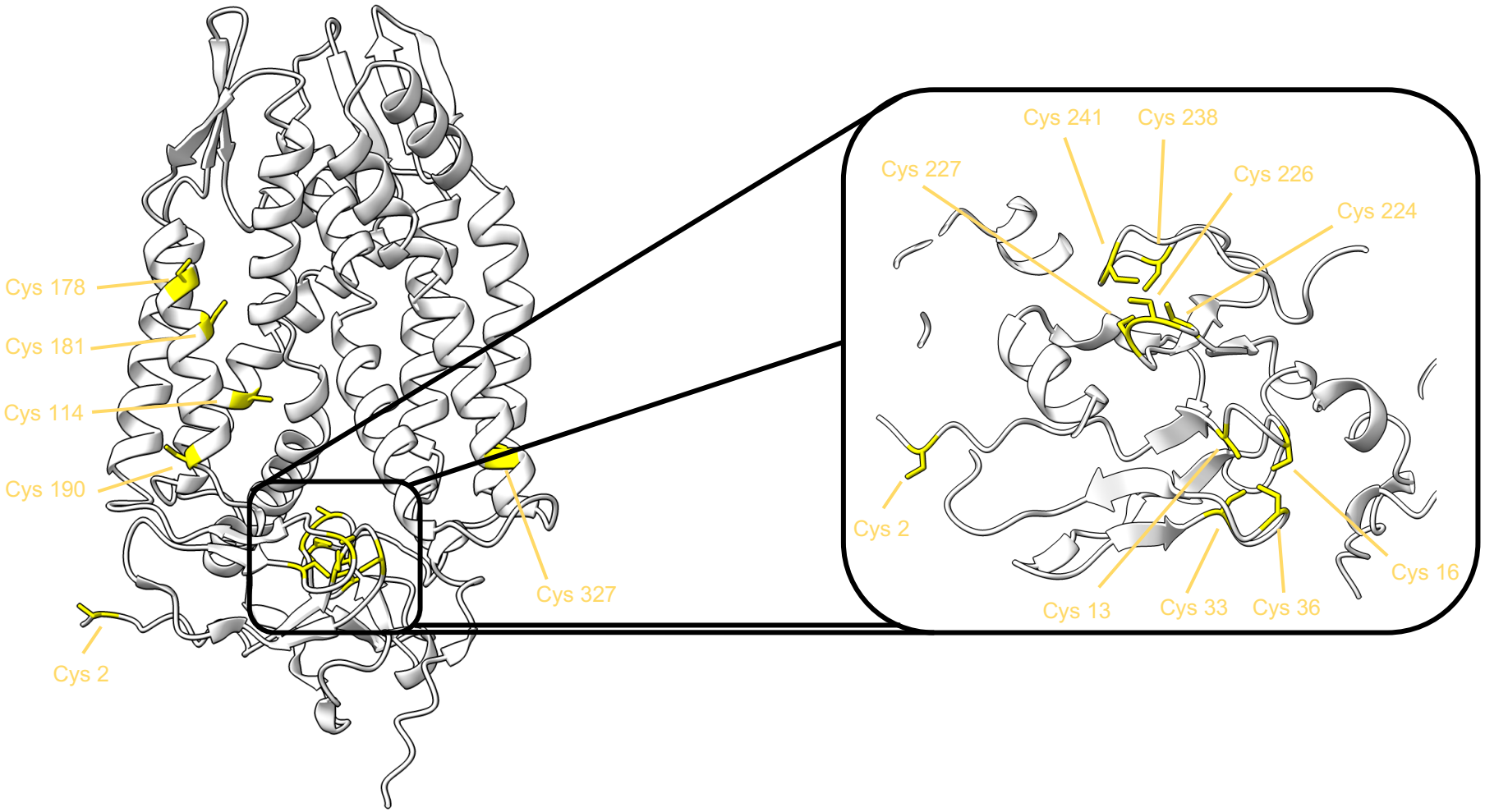


Fig. 79 – Location of the 15 cysteines within PqiA.

Cartoons representation of the AlphaFold predicted PqiA structure (Jumper et al., 2021). The 15 cysteine residues within PqiA are indicated in yellow and labelled. The nine Cys resdiues within the lower loop regions are shown to the right to aid clarity.

Advancement of Pqi Functional Understanding

Prior to this study, information regarding Pqi function had been mostly inferred from studies performed on the more highly scrutinised *E. coli* MCE proteins MlaD and YebT. Indeed, direct evidence implicating Pqi in transmembrane GPL transport involved only the co-purification of PqiB with PE and PG lipids and its membrane spanning, tunnel-like architecture (Ekiert et al., 2017). Phenotypic complementation assays indicated the importance of specific hydrophobic motifs within the PLL components of each MCE domain for the correct function of both MlaD and YebT (Ekiert et al., 2017; Isom et al., 2020). Furthermore, cryo-EM flexibility analysis revealed the diameter of the YebT and PqiB central pores to be modulated, by tilting of the MCE domains as well as flexing of the PLL moieties, providing potential mechanistic insight into these proteins (Isom et al., 2020). Meanwhile, beside their inner and outer membrane locations respectively, the roles of both PqiA and PqiC remained completely obscure. Consequently, we utilised complementation assays to investigate the importance of specific residues within all three Pqi components with the aim of enhancing functional understanding.

Our PqiA complementation assays provide the first insight into specific residues essential for the protein's correct function. Furthermore, we observed similarity in the positioning of these key residues with those involved in the protonation dependent mechanism of MATE transporters. Consequently, we present evidence that PqiA may function via a protonation mediated conformational change providing the first mechanistic insight into the PqiA domain, however, this requires validation.

Given this insight, the obscure nature of PqiA still raises several fundamental questions? If PqiA function is initiated by the PMF then how do lipids traverse through the protein, if at all? Indeed, lipids may be directly loaded or unloaded via PqiB, with PqiA acting purely to harvest

energy from the PMF without the need to handle lipids itself. In this scenario, PqiA may interact with the TM helices of PqiB to propagate its own PMF induced conformational changes throughout the Pqi complex to drive or regulate transport (Fig. 80). Indeed, interactions between the TM helices of MlaD and MalE have been shown to be essential for phospholipid transport indicating MCE proteins can communicate with IM associated complex components (Chi et al., 2020; Coudray et al., 2020). Furthermore, the predominant location of highly conserved PqiA residues within one of its two lobes is intriguing, particularly as PqiA genes comprising only one lobe may be found in species such as *Bdellovibrio bacteriovorus*. The *E. coli* PqiA gene likely arose from a gene duplication and fusion event, however, the evolutionary pressure to maintain both lobes is unclear and requires further investigation. Despite the slight functional insight provided by this study, the magnitude of uncertainty regarding PqiA function is vast and hence provides several avenues for future work.

We also used complementation assays to confirm the essentiality of residues within the PqiB PLL hydrophobic motifs for correct Pqi function. Mutation of any one of these hydrophobic residues completely abolished Pqi function, mirroring results previously obtained with MlaD and YebT (Ekiert et al., 2017; Isom et al., 2020), however, none of these studies, including our own, are able to distinguish between the two potential scenarios underlying this observation. Indeed, the prevention or alternatively the deregulation of Pqi mediated transport are equally likely explanations for the diminished growth resulting from these mutations. Differentiation of these two scenarios requires additional insight either via additional experimentation or the identification of a homologous system.

Whilst complementation provides a platform allowing the indirect inference of residue significance, they are somewhat flawed due to their use of fitness as an indirect measurement of mutation severity. In order to reveal the intricacies of Pqi mediated transport, including directionality, regulation and kinetics, the transport process itself must be

monitored (ideally in isolation). The obvious solution involves the use of an *in vitro* transport assay, analogous to those developed for the Mla system, hence development of such an assay must be viewed as a priority amongst the interested scientific community. Such an assay would not only benefit investigations into the Pqi complex but also homologous pathways, such as YebST, distinguishing their roles and probing species variation.

PqiB is the only Pqi component whose function is clear, to provide a sheltered hydrophobic environment through the periplasm via which GPLs may transit (Fig. 80). Despite this, the mechanics of PqiB mediated transport are still to be resolved. The PLL moieties of MCE proteins appear to be highly mobile entities, likely involved in the regulation of transport, however, the mechanism by which their differing conformations are induced remains unclear. Furthermore, if MCE protein mediated transport is regulated, partially or entirely, via PLL conformations then the differing architectures of PqiB and YebT present an obvious question. YebT comprises MCE domains along its entire length whilst PqiB only contains three MCE domains and thus contains the regulatory PLLs along approximately half its length only. The need for differing numbers of MCE domains, and consequently PLLs, between the two proteins is intriguing. One potential explanation lies in the OM interaction of each complex. Whilst we anticipate these interactions to be distinct, owed to the differing C-termini of YebT and PqiB as well as the presence of PqiC, currently they both remain enigmatic.

The interaction between the Pqi complex and the OM is of significant interest considering its obscurity (Fig. 80). Indeed, our current inability to confirm if PqiB makes OM contact itself presents a fundamental uncertainty, let alone the mechanism of loading/unloading cargo at this interface. Despite being guided by our X-ray structures, none of our PqiC mutations appeared to impede Pqi function hence we were unable to gain further insight into the role of PqiC. This likely results from our introduction of benign mutations, justifying the need for more potent charge reversal mutations, however, a scenario in which PqiC purely stabilises the PqiB - OM interaction without initiating interactions is also possible. The lack of intimate

interactions between PqiB and PqiC would justify the apparent symmetry match between these two components. If PqiC was to act purely as a collar, supporting the PqiB C-terminus at the OM, then it may allow rotation of the PqiB coil moiety within it, reminiscent of a bearing. Furthermore, the interaction between PqiC and the OM may provide insight into its role. Whilst molecular dynamics simulations have suggested the probable nature of their interaction, this may be confirmed via neutron reflectometry. Indeed, a combination of structural and functional studies are likely required to definitively reveal the role of PqiC.

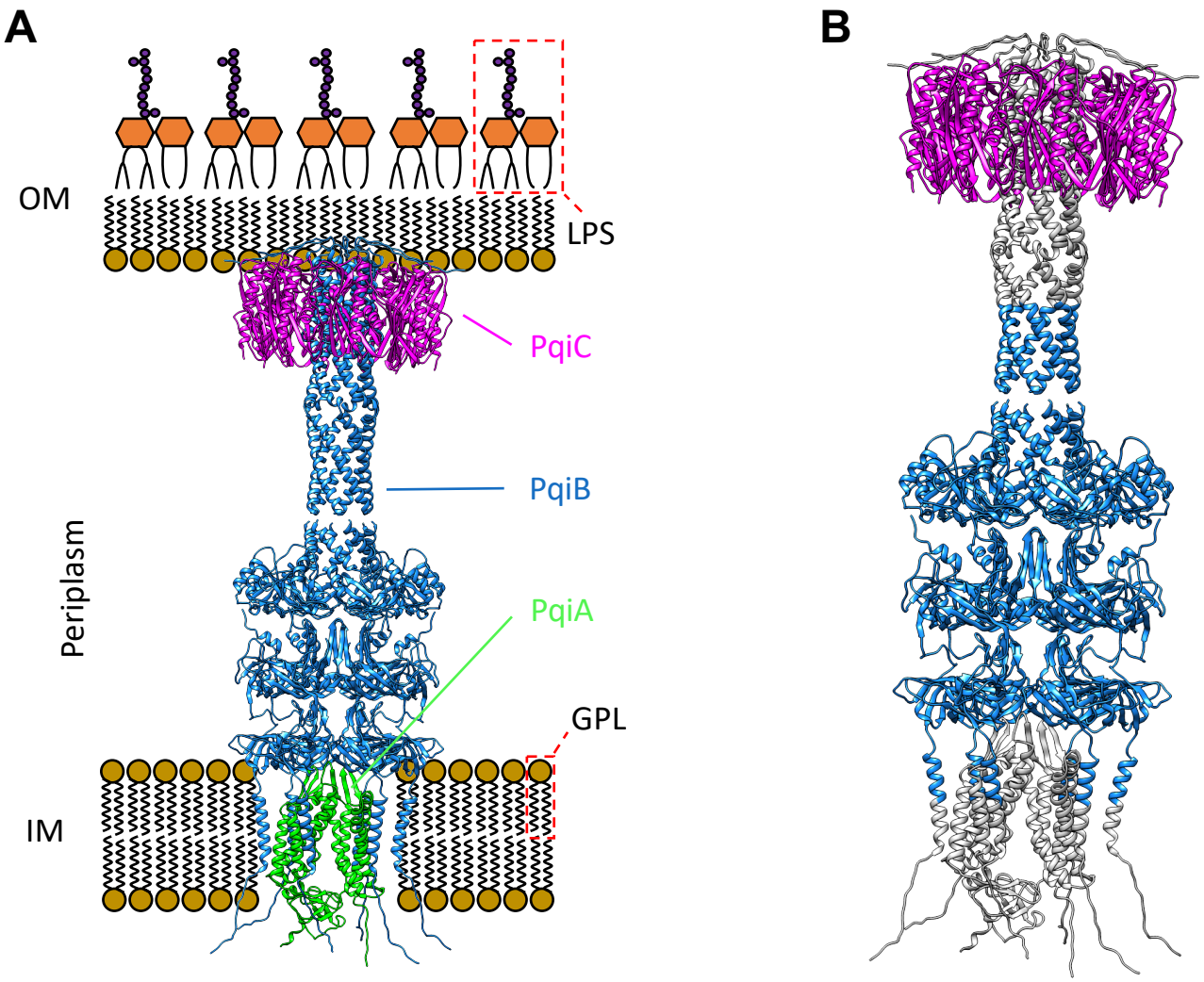


Fig. 80 – Anticipated architecture of the PqiABC complex.

(A) Cartoon representation of the anticipated architecture of the periplasmic spanning PqiABC complex constructed from AlphaFold models. PqiA, PqiB and PqiC are depicted in green, blue and purple respectively. (B) PqiABC complex constructed from AlphaFold models (Jumper et al., 2021). Regions currently determined experimentally are coloured whilst those remaining enigmatic are grey.

Concluding Remarks

Overall, this study has furthered insight into the structure and function of the *E. coli* Pqi complex, however, numerous fundamental uncertainties persist. Despite a strong consensus, evidence implicating Pqi as a lipid transport system is still speculative, furthermore, its transport mechanism, regulation and directionality all remain enigmatic. Finally, the overall architecture of, and interactions maintaining, the periplasmic spanning Pqi complex is unclear. These principal uncertainties provide obvious aims for future studies aiming to characterise the Pqi complex.

Appendix

Appendix. 1 – Construct coding insert sequences.

Below please find the insert coding sequences for each of the constructs utilised in this study. Inserts are located between the Nde1 and Xho1 restriction sites of the pET 26b plasmid (PqiC$^{17\text{-}187}$, PqiABC, PqiAB) or pET 22b (PqiC FL).

PqiC FL

AAGAAATGGCTGGTGACCATTGCGGCGCTGTGGCTGGCGGGTTGCAGCAGCGGCGAG
ATTAACAAGAACTACTATCAGCTGCCGGTGGTTCAGAGCGGTACCCAAAGCACCGCGA
GCCAGGGTAACCGTCTGCTGTGGGTTGAACAAGTGACCGTTCCGGACTACCTGGCGGG
TAACGGCGTGGTTTACCAGACCAGCGATGTGAAATATGTTATCGCGAACAACAACCTGT
GGGCGAGCCCGCTGGACCAGCAACTGCGTAACACCCTGGTGGCGAACCTGAGCACCC
AACTGCCGGGTTGGGTGGTTGCGAGCCAGCCGCTGGGTAGCGCGCAAGACACCCTGA
ACGTGACCGTTACCGAGTTCAACGGTCGTTACGATGGCAAGGTGATCGTTAGCGGTGA
ATGGCTGCTGAACCACCAGGGCCAACTGATTAAACGTCCGTTTCGTCTGGAGGGTGTG
CAGACCCAAGACGGCTATGATGAAATGGTGAAGGTTCTGGCGGGTGTTTGGAGCCAGG
AAGCGGCGAGCATCGCGCAAGAAATTAAACGTCTGCCG

PqiC$^{17\text{-}187}$

(please note the regions highlighted in red and blue refer to the N-terminal hex-histidine tag and TEV cleavage site respectively)

CATCACCACCACCACCACGAGAACCTGTACTTCCAGATGAGCAGCGGTGAAATCAACAA
GAACTACTATCAACTGCCGGTGGTTCAGAGCGGTACCCAAAGCACCGCGAGCCAGGGT

AACCGTCTGCTGTGGGTGGAGCAAGTGACCGTTCCGGACTATCTGGCGGGTAACGGCG
TGGTTTACCAGACCAGCGATGTGAAATATGTTATTGCGAACAACAACCTGTGGGCGAGC
CCGCTGGACCAGCAACTGCGTAACACCCTGGTTGCGAACCTGAGCACCCAACTGCCGG
GTTGGGTGGTTGCGAGCCAGCCGCTGGGTAGCGCGCAAGACACCCTGAACGTGACCG
TTACCGAGTTCAACGGTCGTTACGATGGCAAGGTGATCGTTAGCGGTGAATGGCTGCT
GAACCACCAGGGCCAACTGATTAAACGTCCGTTTCGTCTGGAGGGTGTGCAGACCCAA
GACGGCTATGATGAAATGGTGAAGGTTCTGGCGGGCGTTTGGAGCCAGGAAGCGGCG
AGCATCGCGCAAGAAATTAAACGTCTGCCGTAA

PqiABC

TGCGAACATCATCATGCCGCGAAGCACATCCTGTGCTCGCAGTGTGACATGCTGGTG
GCGTTACCGCGCCTTGAGCATGGTCAGAAAGCGGCATGTCCCCGGTGTGGCACAACGT
TAACCGTGGCGTGGGATGCCCCTCGGCAGCGTCCCACCGCCTATGCGTTGGCTGCACT
GTTCATGCTGTTGCTGTCCAACTTGTTTCCTTTTGTGAATATGAACGTTGCAGGAGTTAC
CAGTGAAATTACATTACTGGAAATTCCCGGCGTGCTTTTTTCTGAGGACTACGCCAGCC
TCGGCACCTTTTTCCTGTTGTTTGTGCAACTGGTTCCCGCGTTTTGTCTGATAACCATTC
TGTTACTGGTGAATCGCGCGGAATTACCGGTCCGTTTAAAAGAGCAACTGGCACGGGT
GCTTTTTCAACTCAAAACCTGGGGAATGGCGGAGATTTTCCTCGCGGGTGTGCTGGTCA
GTTTCGTTAAACTGATGGCTTACGGCAGCATTGGGGTAGGCAGCAGCTTTCTCCCCTGG
TGTTTATTTTGTGTCCTGCAACTGCGCGCTTTTCAGTGCGTTGATCGTCGCTGGTTATGG
GACGACATCGCCCCGATGCCAGAACTGCGCCAGCCGCTAAAACCAGGCGTCACGGGG
ATACGTCAGGGGCTGCGTTCGTGCTCCTGTTGTACGGCAATCCTTCCTGCTGATGAACC
CGTGTGCCCGCGTTGTAGTACCAAAGGGTACGTTCGGCGTAGAAACAGCCTGCAGTGG
ACACTCGCGCTGCTTGTAACGTCCATCATGCTGTATCTTCCGGCTAATATTTTGCCCATC
ATGGTGACGGATTTATTAGGCTCGAAGATGCCGTCGACGATTCTCGCTGGGGTCATTCT
GTTATGGAGCGAAGGCTCTTATCCCGTCGCTGCGGTGATCTTTCTGGCCAGTATTATGG

TGCCAACGTTAAAGATGATCGCCATCGCGTGGCTGTGTTGGGATGCCAAAGGGCATGG

CAAGCGCGACAGTGAAAGAATGCATTTGATTTATGAAGTTGTTGAGTTTGTAGGCCGCT

GGTCGATGATTGACGTTTTCGTTATCGCGGTGCTCTCGGCGCTGGTGCGTATGGGAGG

TTTAATGAGTATTTATCCGGCAATGGGTGCATTAATGTTTGCTTTAGTCGTCATAATGACA

ATGTTTTCTGCTATGACGTTTGACCCGCGTTTGTCGTGGGATCGTCAACCTGAATCAGA

GCATGAGGAGTCCTGACAGTATGGAATCTAATAATGGGGAAGCCAAAATCCAGAAAGTG

AAGAACTGGTCTCCCGTGTGGATATTTCCTATCGTCACGGCGCTCATTGGGGCCTGGGT

TCTTTTTTATCATTACAGCCATCAGGGACCGGAAGTGACCCTGATCACCGCGAATGCGG

AAGGAATTGAAGGTGGCAAAACCACCATTAAAAGCCGTAGCGTTGACGTCGGCGTGGT

TGAAAGCGCCACACTGGCTGATGATTTGACGCACGTTGAAATCAAAGCGCGGCTGAATT

CCGGTATGGAAAAATTGCTGCATAAAGACACCGTCTTTTGGGTGGTGAAACCGCAGATT

GGTCGCGAAGGGATTAGCGGCCTGGGAACGCTGCTGTCTGGAGTTTATATCGAACTGC

AGCCAGGCGCGAAAGGCAGCAAAATGGATAAATACGATTTGCTGGACTCGCCACCGTT

GGCCCCGCCTGATGCGAAAGGTATCCGTGTGATTCTCGATAGCAAAAAAGCCGGGCAG

CTCTCGCCAGGAGATCCGGTGCTGTTCCGTGGCTATCGGGTAGGTTCGGTTGAAACCA

GCACCTTCGATACACAAAAACGCAATATCAGCTATCAACTGTTCATCAATGCACCTTATG

ACCGACTGGTGACCAACAATGTTCGCTTCTGGAAAGATAGTGGCATTGCGGTTGATCTG

ACGTCAGCAGGGATGCGTGTGGAGATGGGCTCATTGACAACGCTGCTGAGTGGCGGT

GTCAGCTTTGATGTGCCGGAAGGTCTGGATTTAGGGCAGCCAGTGGCACCGAAAACAG

CTTTCGTTTTGTATGATGATCAGAAGAGCATTCAGGATTCGTTGTACACCGATCACATTG

ATTATCTGATGTTCTTTAAAGATTCGGTACGCGGTCTGCAACCGGGAGCTCCGGTAGAG

TTCCGGGGTATTCGCCTGGGTACCGTAAGCAAAGTGCCATTCTT

TGCGCCGAATATGCGTCAGACATTTAACGATGATTACCGTATTCCGGTACTGATTCGTAT

CGAGCCAGAGCGGCTGAAAATGCAGCTTGGCGAAAATGCGGATGTTGTTGAGCACCTT

GGCGAATTGTTGAAACGTGGTTTACGCGGATCGCTGAAAACCGGAAACCTGGTCACTG

GTGCACTGTATGTTGATCTCGATTTCTATCCAAATACGCCTGCAATAACCGGTATTCGTG

AATTTAATGGTTATCAGATTATCCCGACCGTTAGCGGCGGCCTGGCGCAAATCCAGCAA

CGACTGATGGAAGCGTTGGATAAGATCAACAAACTGCCATTGAATCCGATGATTGAACA

GGCAACCAGTACGCTTTCTGAAAGTCAGCGCACAATGAAAAACCTGCAAACGACGCTG
GATAGCATGAACAAGATCCTCGCTAGCCAGTCGATGCAGCAGTTGCCGACGGATATGC
AGTCAACGTTGCGTGAATTGAATCGCAGCATGCAGGGCTTCCAGCCTGGCTCCGCAGC
CTACAACAAGATGGTGGCGGATATGCAGCGCCTTGATCAGGTGTTGCGAGAACTGCAA
CCGGTGCTGAAAACGCTCAATGAGAAGAGTAACGCGCTGGTATTTGAAGCGAAGGACA
AAAAAGATCCAGAGCCGAAGAGGGCGAAACAATGAAAAAGTGGCTAGTGACGATTGCA
GCACTGTGGCTGGCCGGATGCAGCTCCGGCGAAATTAATAAAAACTATTACCAGTTACC
TGTGGTGCAGAGCGGTACACAAAGTACCGCCAGCCAGGGCAATCGTCTGTTATGGGTA
GAGCAGGTCACTGTTCCTGACTATCTGGCGGGGAATGGTGTGGTTTATCAAACCAGTGA
TGTGAAGTATGTGATTGCCAACAACAACTTGTGGGCCAGCCCGTTGGATCAACAGTTGC
GCAACACCCTGGTTGCCAACCTGAGTACGCAACTGCCCGGCTGGGTGGTTGCCTCCCA
GCCTCTGGGAAGCGCCCAGGACACGCTCAATGTTACCGTAACGGAGTTTAACGGTCGC
TATGATGGCAAGGTCATTGTCAGTGGTGAGTGGCTGTTGAACCACCAGGGACAACTGAT
CAAACGTCCGTTCCGTCTGGAAGGAGTGCAAACTCAGGATGGTTACGATGAGATGGTTA
AAGTGCTGGCCGGTGTCTGGAGTCAGGAAGCCGCTTCTATTGCACAAGAGATAAAGCG
TCTACCT

PqiAB

TGCGAACATCATCATGCCGCGAAGCACATCCTGTGCTCGCAGTGTGACATGCTGGTG
GCGTTACCGCGCCTTGAGCATGGTCAGAAAGCGGCATGTCCCCGGTGTGGCACAACGT
TAACCGTGGCGTGGGATGCCCCTCGGCAGCGTCCCACCGCCTATGCGTTGGCTGCACT
GTTCATGCTGTTGCTGTCCAACTTGTTTCCTTTTGTGAATATGAACGTTGCAGGAGTTAC
CAGTGAAATTACATTACTGGAAATTCCCGGCGTGCTTTTTTCTGAGGACTACGCCAGCC
TCGGCACCTTTTTCCTGTTGTTTGTGCAACTGGTTCCCGCGTTTTGTCTGATAACCATTC
TGTTACTGGTGAATCGCGCGGAATTACCGGTCCGTTTAAAAGAGCAACTGGCACGGGT
GCTTTTTCAACTCAAAACCTGGGGAATGGCGGAGATTTTCCTCGCGGGTGTGCTGGTCA

GTTTCGTTAAACTGATGGCTTACGGCAGCATTGGGGTAGGCAGCAGCTTTCTCCCCTGG

TGTTTATTTTGTGTCCTGCAACTGCGCGCTTTTCAGTGCGTTGATCGTCGCTGGTTATGG

GACGACATCGCCCCGATGCCAGAACTGCGCCAGCCGCTAAAACCAGGCGTCACGGGG

ATACGTCAGGGGCTGCGTTCGTGCTCCTGTTGTACGGCAATCCTTCCTGCTGATGAACC

CGTGTGCCCGCGTTGTAGTACCAAAGGGTACGTTCGGCGTAGAAACAGCCTGCAGTGG

ACACTCGCGCTGCTTGTAACGTCCATCATGCTGTATCTTCCGGCTAATATTTTGCCCATC

ATGGTGACGGATTTATTAGGCTCGAAGATGCCGTCGACGATTCTCGCTGGGGTCATTCT

GTTATGGAGCGAAGGCTCTTATCCCGTCGCTGCGGTGATCTTTCTGGCCAGTATTATGG

TGCCAACGTTAAAGATGATCGCCATCGCGTGGCTGTGTTGGGATGCCAAAGGGCATGG

CAAGCGCGACAGTGAAAGAATGCATTTGATTTATGAAGTTGTTGAGTTTGTAGGCCGCT

GGTCGATGATTGACGTTTTCGTTATCGCGGTGCTCTCGGCGCTGGTGCGTATGGGAGG

TTTAATGAGTATTTATCCGGCAATGGGTGCATTAATGTTTGCTTTAGTCGTCATAATGACA

ATGTTTTCTGCTATGACGTTTGACCCGCGTTTGTCGTGGGATCGTCAACCTGAATCAGA

GCATGAGGAGTCCTGACAGTATGGAATCTAATAATGGGGAAGCCAAAATCCAGAAAGTG

AAGAACTGGTCTCCCGTGTGGATATTTCCTATCGTCACGGCGCTCATTGGGGCCTGGGT

TCTTTTTTATCATTACAGCCATCAGGGACCGGAAGTGACCCTGATCACCGCGAATGCGG

AAGGAATTGAAGGTGGCAAAACCACCATTAAAAGCCGTAGCGTTGACGTCGGCGTGGT

TGAAAGCGCCACACTGGCTGATGATTTGACGCACGTTGAAATCAAAGCGCGGCTGAATT

CCGGTATGGAAAAATTGCTGCATAAAGACACCGTCTTTTGGGTGGTGAAACCGCAGATT

GGTCGCGAAGGGATTAGCGGCCTGGGAACGCTGCTGTCTGGAGTTTATATCGAACTGC

AGCCAGGCGCGAAAGGCAGCAAAATGGATAAATACGATTTGCTGGACTCGCCACCGTT

GGCCCCGCCTGATGCGAAAGGTATCCGTGTGATTCTCGATAGCAAAAAAGCCGGGCAG

CTCTCGCCAGGAGATCCGGTGCTGTTCCGTGGCTATCGGGTAGGTTCGGTTGAAACCA

GCACCTTCGATACACAAAAACGCAATATCAGCTATCAACTGTTCATCAATGCACCTTATG

ACCGACTGGTGACCAACAATGTTCGCTTCTGGAAAGATAGTGGCATTGCGGTTGATCTG

ACGTCAGCAGGGATGCGTGTGGAGATGGGCTCATTGACAACGCTGCTGAGTGGCGGT

GTCAGCTTTGATGTGCCGGAAGGTCTGGATTTAGGGCAGCCAGTGGCACCGAAAACAG

CTTTCGTTTTGTATGATGATCAGAAGAGCATTCAGGATTCGTTGTACACCGATCACATTG

ATTATCTGATGTTCTTTAAAGATTCGGTACGCGGTCTGCAACCGGGAGCTCCGGTAGAG

TTCCGGGGTATTCGCCTGGGTACCGTAAGCAAAGTGCCATTCTT

TGCGCCGAATATGCGTCAGACATTTAACGATGATTACCGTATTCCGGTACTGATTCGTAT

CGAGCCAGAGCGGCTGAAAATGCAGCTTGGCGAAAATGCGGATGTTGTTGAGCACCTT

GGCGAATTGTTGAAACGTGGTTTACGCGGATCGCTGAAAACCGGAAACCTGGTCACTG

GTGCACTGTATGTTGATCTCGATTTCTATCCAAATACGCCTGCAATAACCGGTATTCGTG

AATTTAATGGTTATCAGATTATCCCGACCGTTAGCGGCGGCCTGGCGCAAATCCAGCAA

CGACTGATGGAAGCGTTGGATAAGATCAACAAACTGCCATTGAATCCGATGATTGAACA

GGCAACCAGTACGCTTTCTGAAAGTCAGCGCACAATGAAAAACCTGCAAACGACGCTG

GATAGCATGAACAAGATCCTCGCTAGCCAGTCGATGCAGCAGTTGCCGACGGATATGC

AGTCAACGTTGCGTGAATTGAATCGCAGCATGCAGGGCTTCCAGCCTGGCTCCGCAGC

CTACAACAAGATGGTGGCGGATATGCAGCGCCTTGATCAGGTGTTGCGAGAACTGCAA

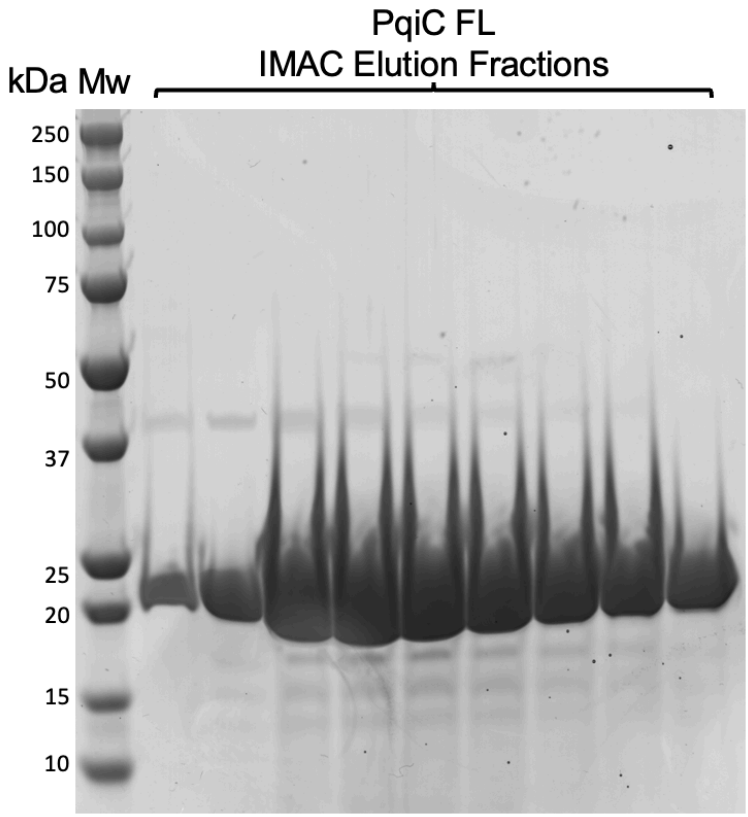
CCGGTGCTGAAAACGCTCAATGAGAAGAGTAACGCGCTGGTATTTGAAGCGAAGGACA

AAAAAGATCCAGAGCCGAAGAGGGCGAAACAA

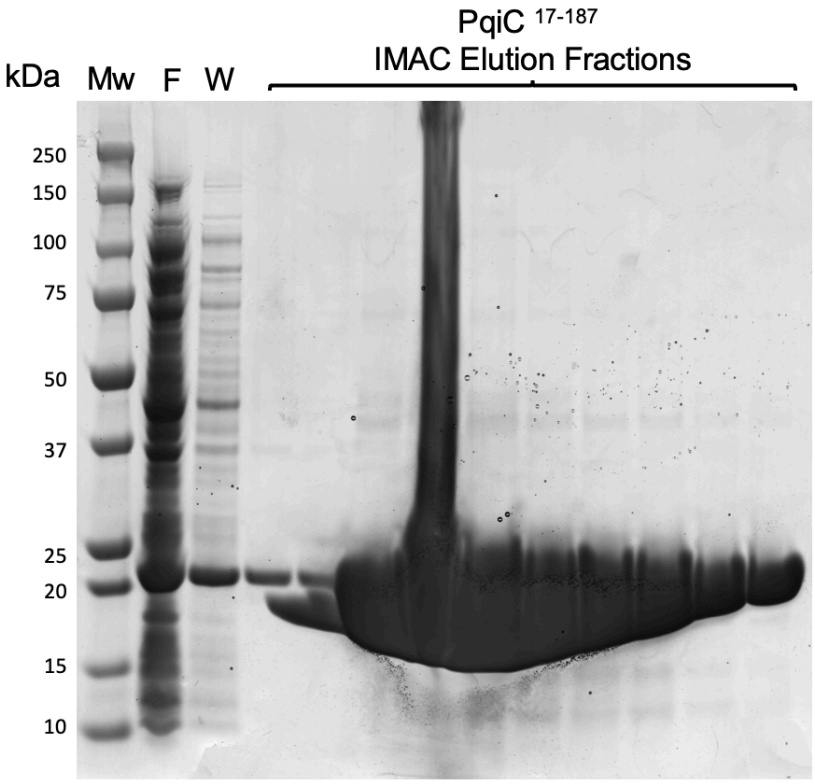
Appendix. 2 – PqiC X-ray Data Collection Parameters

Parameter	CFL1 Collection	CFL2 Collection	CsoI Collection
Wavelength (Å)	0.9763	0.9797	1.0000
Oscillations (˚)	0.1	0.1	0.1
Images	3600	3600	1800
Transmission (%)	100	100	100
Exposure (s)	0.01	0.005	0.04
Beam Size (µm)	80 x 20	86 x 75	43 x 30

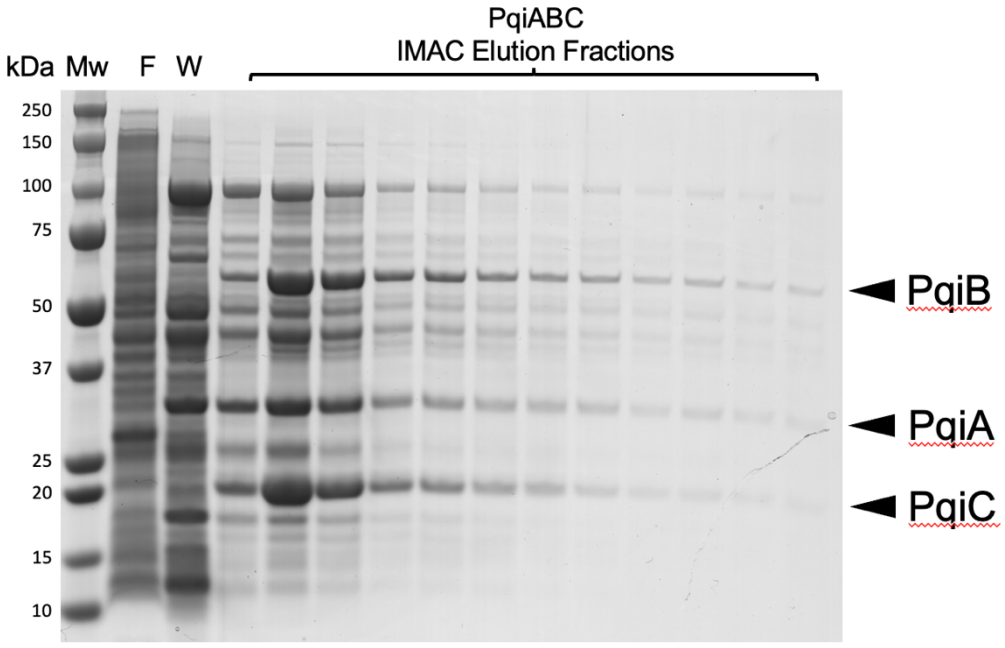
Appendix. 3 – PqiC FL IMAC Elution Fraction SDS-PAGE.



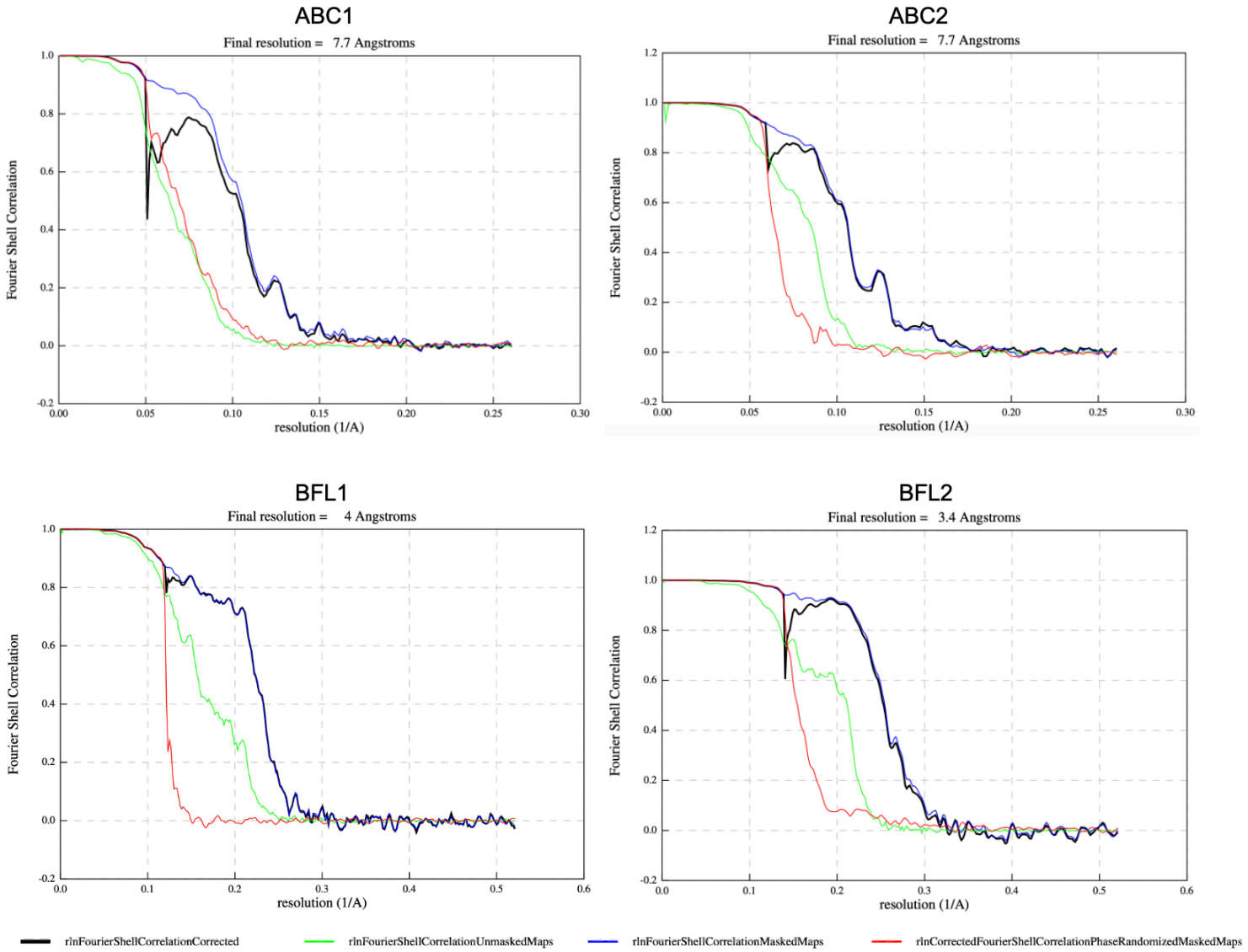
Appendix. 4 – PqiC$^{17-187}$ IMAC Elution Fraction SDS-PAGE.



Appendix. 5 – PqiABC IMAC Elution Fraction SDS-PAGE.



Appendix. 6 – Cryo-EM FSC curves.



Bibliography

Abellón-Ruiz, J., Kaptan, S.S., Baslé, A., Claudi, B., Bumann, D., Kleinekathöfer, U., Van Den Berg, B., 2017. Structural basis for maintenance of bacterial outer membrane lipid asymmetry. Nature Microbiology 2, 1616–1623. https://doi.org/10.1038/s41564-017-0046-x

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. urn:issn:0907-4449 68, 352–367. https://doi.org/10.1107/S0907444912001308

Al-Saedi, F.M.A., 2018. Development of a bacterial adhesin into a next generation antimicrobial agent. University of Birmingham.

Arnold, S.A., Albiez, S., Bieri, A., Syntychaki, A., Adaixo, R., McLeod, R.A., Goldie, K.N., Stahlberg, H., Braun, T., 2017. Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts. Journal of Structural Biology 197, 220–226. https://doi.org/10.1016/j.jsb.2016.11.002

Arruda, S., Bomfim, G., Knights, R., Huima-Byron, T., Riley, L.W., 1993. Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. Science 261, 1454–1457. https://doi.org/10.1126/science.8367727

Ashtiani, D., Venugopal, H., Belousoff, M., Spicer, B., Mak, J., Neild, A., de Marco, A., 2018. Delivery of femtolitre droplets using surface acoustic wave based atomisation for cryo-EM grid preparation. Journal of Structural Biology 203, 94–101. https://doi.org/10.1016/j.jsb.2018.03.012

Awai, K., Xu, C., Tamot, B., Benning, C., 2006. A phosphatidic acid-binding protein of the chloroplast inner envelope membrane involved in lipid trafficking.

Baarda, B.I., Zielke, R.A., Van, A.L., Jerse, A.E., Sikora, A.E., 2019. Neisseria gonorrhoeae MlaA influences gonococcal virulence and membrane vesicle production. PLoS Pathogens 15, e1007385. https://doi.org/10.1371/journal.ppat.1007385

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology 2, 2006.0008. https://doi.org/10.1038/MSB4100050

Babu, M., Bundalovic-Torma, C., Calmettes, C., Phanse, S., Zhang, Q., Jiang, Y., Minic, Z., Kim, S., Mehla, J., Gagarinova, A., Rodionova, I., Kumar, A., Guo, H., Kagan, O., Pogoutse, O., Aoki, H., Deineko, V., Caufield, J.H., Holtzapple, E., Zhang, Z., Vastermark, A., Pandya, Y., Lai, C.C.L., El Bakkouri, M., Hooda, Y., Shah, M., Burnside, D., Hooshyar, M., Vlasblom, J., Rajagopala, S.V., Golshani, A., Wuchty, S., F Greenblatt, J., Saier, M., Uetz, P., F Moraes, T., Parkinson, J., Emili, A., 2018. Global landscape of cell envelope protein complexes in Escherichia coli. Nature Biotechnology 36, 103–112. https://doi.org/10.1038/nbt.4024

Bayer, M.E., 1991. Zones of membrane adhesion in the cryofixed envelope of Escherichia coli. Journal of Structural Biology 107, 268–280. https://doi.org/10.1016/1047-8477(91)90052-X

Bayer, M.E., 1968. Areas of Adhesion between Wall and Membrane of Escherichia coli The envelope of Escherichia coli, J. gen. Microbiol.

Beltran, S., Munoz-Bergmann, C.A., Elola-Lopez, A., Quintana, J., Segovia, C., Trombert, A.N., 2016. The expression of heterologous MAM-7 in Lactobacillus rhamnosus reduces its intrinsic capacity to inhibit colonization of pathogen Vibrio parahaemolyticus in vitro. Biological Research 49, 1–10. https://doi.org/10.1186/S40659-015-0064-1/FIGURES/7

Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., Berger, B., 2019. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nature Methods 2019 16:11 16, 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

Bhella, D., 2019. Cryo-electron microscopy: an introduction to the technique, and considerations when working to establish a national facility. Biophysical Reviews 11, 515. https://doi.org/10.1007/S12551-019-00571-W

Bishop, R.E., 2005. The lipid A palmitoyltransferase PagP: Molecular mechanisms and role in bacterial pathogenesis. Molecular Microbiology 57, 900–912. https://doi.org/10.1111/j.1365-2958.2005.04711.x

Bokori-Brown, M., Martin, T.G., Naylor, C.E., Basak, A.K., Titball, R.W., Savva, C.G., 2016. Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat Commun 7, 11293. https://doi.org/10.1038/ncomms11293

Bragg, W.H., Bragg, W.L., 1913. The reflection of X-rays by crystals. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 88, 428–438. https://doi.org/10.1098/RSPA.1913.0040

Casali, N., Riley, L.W., 2007. A phylogenomic analysis of the Actinomycetales mce operons. BMC Genomics 8, 60. https://doi.org/10.1186/1471-2164-8-60

Casali, N., White, A.M., Riley, L.W., 2006. Regulation of the Mycobacterium tuberculosis mce1 operon. Journal of Bacteriology 188, 441–449. https://doi.org/10.1128/JB.188.2.441-449.2006/ASSET/8FFD3E5A-551B-4FE0-8D28-0A42B1B0203A/ASSETS/GRAPHIC/ZJB0020653920007.JPEG

Chaikuad, A., Knapp, S., von Delft, F., 2015. Defined PEG smears as an alternative approach to enhance the search for crystallization conditions and crystal-quality improvement in reduced screens. Acta Cryst D 71, 1627–1639. https://doi.org/10.1107/S1399004715007968

Charlson, E.S., Werner, J.N., Misra, R., 2006. Differential Effects of yfgL Mutation on Escherichia coli Outer Membrane Proteins and Lipopolysaccharide. J Bacteriol 188, 7186–7194. https://doi.org/10.1128/JB.00571-06

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D: Biological Crystallography 66, 12. https://doi.org/10.1107/S0907444909042073

Chi, X., Fan, Q., Zhang, Y., Liang, K., Wan, L., Zhou, Q., Li, Y., 2020. Structural mechanism of phospholipids translocation by MlaFEDB complex. Cell Research 1–9. https://doi.org/10.1038/s41422-020-00404-6

Chitale, S., Ehrt, S., Kawamura, I., Fujimura, T., Shimono, N., Anand, N., Lu, S., Cohen-Gould, L., Riley, L.W., 2001. Recombinant Mycobacterium tuberculosis protein associated with mammalian cell entry. Cellular Microbiology 3, 247–254. https://doi.org/10.1046/j.1462-5822.2001.00110.x

Chong, Z.-S., Woo, W.-F., Chng, S.-S., 2015. Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in Escherichia coli. Molecular Microbiology 98, 1133–1146. https://doi.org/10.1111/mmi.13202

Cian, M.B., Giordano, N.P., Masilamani, R., Minor, K.E., Dalebroux, Z.D., 2020. Salmonella enterica serovar typhimurium uses PbgA/YejM to regulate lipopolysaccharide assembly during bacteremia. Infection and Immunity 88. https://doi.org/10.1128/IAI.00758-19

Clairfeuille, T., Buchholz, K.R., Li, Q., Verschueren, E., Liu, P., Sangaraju, D., Park, S., Noland, C.L., Storek, K.M., Nickerson, N.N., Martin, L., Dela Vega, T., Miu, A., Reeder, J., Ruiz-Gonzalez, M., Swem, D., Han, G., DePonte, D.P., Hunter, M.S., Gati, C., Shahidi-Latham, S., Xu, M., Skelton, N., Sellers, B.D., Skippington, E., Sandoval, W., Hanan, E.J., Payandeh, J., Rutherford, S.T., 2020. Structure of the essential inner membrane lipopolysaccharide–PbgA complex. Nature 584, 479–483. https://doi.org/10.1038/s41586-020-2597-x

Claxton, D.P., Jagessar, K.L., Ryan Steed, P., Stein, R.A., Mchaourab, H.S., 2018. Sodium and proton coupling in the conformational cycle of a MATE antiporter from Vibrio cholerae. Proceedings of the National Academy of Sciences of the United States of America 115, E6182–E6190. https://doi.org/10.1073/PNAS.1802417115/SUPPL_FILE/PNAS.1802417115.SAPP.PDF

Coudray, N., Isom, G.L., Macrae, M.R., Saiduddin, M.N., Bhabha, G., Ekiert, D.C., 2020. Structure of bacterial phospholipid transporter MlaFEDB with substrate bound. eLife 9, 1–73. https://doi.org/10.7554/eLife.62518

Croll, T.I., 2018. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallographica Section D: Structural Biology 74, 519–530. https://doi.org/10.1107/S2059798318002425/IC5101SUP2.MP4

Cronan, J., 2002. Phospholipid modifications in bacteria. Current Opinion in Microbiology 5, 202–205. https://doi.org/10.1016/S1369-5274(02)00297-7

Cunningham, B.C., Wells, J.A., 1989. High-Resolution Epitope Mapping of hGH-Receptor Interactions by Alanine-Scanning Mutagenesis. Science 244, 1081–1085. https://doi.org/10.1126/science.2471267

Dalebroux, Z.D., Edrozo, M.B., Pfuetzner, R.A., Ressl, S., Kulasekara, B.R., Blanc, M.P., Miller, S.I., 2015. Delivery of cardiolipins to the salmonella outer membrane is necessary for survival within host tissues and virulence. Cell Host and Microbe 17, 441–451. https://doi.org/10.1016/j.chom.2015.03.003

Dalebroux, Z.D., Matamouros, S., Whittington, D., Bishop, R.E., Miller, S.I., 2014. PhoPQ regulates acidic glycerophospholipid content of the Salmonella Typhimurium outer membrane. Proceedings of the National Academy of Sciences of the United States of America 111, 1963–1968. https://doi.org/10.1073/pnas.1316901111

Dalebroux, Z.D., Miller, S.I., 2014. Salmonellae PhoPQ regulation of the outer membrane to resist innate immunity. Current Opinion in Microbiology 17, 106–113. https://doi.org/10.1016/j.mib.2013.12.005

Dandey, V.P., Wei, H., Zhang, Z., Tan, Y.Z., Acharya, P., Eng, E.T., Rice, W.J., Kahn, P.A., Potter, C.S., Carragher, B., 2018. Spotiton: New Features and Applications. J Struct Biol 202, 161–169. https://doi.org/10.1016/j.jsb.2018.01.002

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97, 6640–6645.

de La Paz, M., Klepp, L., Nuñez-García, J., Blanco, F.C., Soria, M., del Carmen García-Polayo, M., Bianco, M.V., Cataldi, A.A., Golby, P., Jackson, M., Gordon, S.V., Bigi, F., 2009. Mce3R, a TetR-type transcriptional repressor, controls the expression of a regulon involved in lipid metabolism in mycobacterium tuberculosis. Microbiology 155, 2245–2255. https://doi.org/10.1099/MIC.0.027086-0/CITE/REFWORKS

De Rosier, D.J., Klug, A., 1968. Reconstruction of Three Dimensional Structures from Electron Micrographs. Nature 1968 217:5124 217, 130–134. https://doi.org/10.1038/217130a0

Dekker, N., 2000. Outer-membrane phospholipase A: known structure, unknown biological function. Molecular Microbiology 35, 711–717. https://doi.org/10.1046/j.1365-2958.2000.01775.x

Dessau, M.A., Modis, Y., 2011. Protein Crystallization for X-ray Crystallography. J Vis Exp 2285. https://doi.org/10.3791/2285

Doerrler, W.T., Gibbons, H.S., Christian, R., Raetz, H., 2004. MsbA-dependent translocation of lipids across the inner membrane of Escherichia coli. Journal of Biological Chemistry 279, 45102–45109. https://doi.org/10.1074/jbc.M408106200

Douglass, M.V., McLean, A.B., Trent, M.S., 2022. Absence of YhdP, TamB, and YdbH leads to defects in glycerophospholipid transport and cell morphology in Gram-negative bacteria. PLOS Genetics 18, e1010096. https://doi.org/10.1371/JOURNAL.PGEN.1010096

Dowhan, W., 2013. A retrospective: Use of Escherichia coli as a vehicle to study phospholipid synthesis and function. Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids 1831, 471–494. https://doi.org/10.1016/j.bbalip.2012.08.007
Doyle, M.T., Jimah, J.R., Dowdy, T., Ohlemacher, S.I., Larion, M., Hinshaw, J.E., Bernstein, H.D., 2022. Cryo-EM structures reveal multiple stages of bacterial outer membrane protein folding. Cell 185, 1143-1156.e13. https://doi.org/10.1016/J.CELL.2022.02.016
Ducret, A., Fleuchot, B., Bergam, P., Mignot, T., 2013. Direct live imaging of cell-cell protein transfer by transient outer membrane fusion in Myxococcus xanthus. eLife 2013. https://doi.org/10.7554/eLife.00868
Dunphy, K.Y., Senaratne, R.H., Masuzawa, M., Kendall, L.V., Riley, L.W., 2010. Attenuation of Mycobacterium tuberculosis Functionally Disrupted in a Fatty Acyl-Coenzyme A Synthetase Gene fadD5. The Journal of Infectious Diseases 201, 1232–1239. https://doi.org/10.1086/651452
Ekiert, D.C., Bhabha, G., Isom, G.L., Greenan, G., Ovchinnikov, S., Henderson, I.R., Cox, J.S., Vale, R.D., 2017. Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. Cell 169, 273-285.e17. https://doi.org/10.1016/j.cell.2017.03.019
Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. urn:issn:0907-4449 60, 2126–2132. https://doi.org/10.1107/S0907444904019158
Ercan, B., Low, W.Y., Liu, X., Chng, S.S., 2019. Characterization of Interactions and Phospholipid Transfer between Substrate Binding Proteins of the OmpC-Mla System. Biochemistry 58, 114–119. https://doi.org/10.1021/acs.biochem.8b00897
Erickson, H.P., Klug, A., Huxley, H.E., Klug, A., 1997. Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 261, 105–118. https://doi.org/10.1098/rstb.1971.0040
Fan, J., Petersen, E.M., Hinds, T.R., Zheng, N., Miller, S.I., 2020. Structure of an inner membrane protein required for phopq-regulated increases in outer membrane cardiolipin. mBio 11. https://doi.org/10.1128/mBio.03277-19
Feng, X., Fu, Z., Kaledhonkar, S., Jia, Y., Shah, B., Jin, A., Liu, Z., Sun, M., Chen, B., Grassucci, R.A., Ren, Y., Jiang, H., Frank, J., Lin, Q., 2017. A Fast and Effective Microfluidic Spraying-plunging Method for High-Resolution Single-Particle Cryo-EM. Structure 25, 663-670.e3. https://doi.org/10.1016/j.str.2017.02.005
Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research 44, D279–D285. https://doi.org/10.1093/nar/gkv1344
Fivenson, E.M., Bernhardt, T.G., 2020. An essential membrane protein modulates the proteolysis of LpxC to control lipopolysaccharide synthesis in escherichia coli. mBio 11. https://doi.org/10.1128/mBio.00939-20
Forrellad, M.A., McNeil, M., Santangelo, M.D.L.P., Blanco, F.C., García, E., Klepp, L.I., Huff, J., Niederweis, M., Jackson, M., Bigi, F., 2014. Role of the Mce1 transporter in the lipid homeostasis of Mycobacterium tuberculosis. Tuberculosis 94, 170–177. https://doi.org/10.1016/J.TUBE.2013.12.005
Gabale, U., Peña Palomino, P.A., Kim, H.A., Chen, W., Ressl, S., 2020. The essential inner membrane protein YejM is a metalloenzyme. Scientific Reports 10, 1–14. https://doi.org/10.1038/s41598-020-73660-6
Germon, P., Ray, M.C., Vianney, A., Lazzaroni, J.C., 2001. Energy-dependent conformational change in the TolA protein of Escherichia coli involves its N-terminal domain, TolQ, and TolR. Journal of Bacteriology 183, 4110–4114. https://doi.org/10.1128/JB.183.14.4110-4114.2001
Giacometti, S.I., MacRae, M.R., Dancel-Manning, K., Bhabha, G., Ekiert, D.C., 2022. Lipid Transport Across Bacterial Membranes. Annual Review of Cell and Developmental Biology 38, 125–153. https://doi.org/10.1146/annurev-cellbio-120420-022914
Glaeser, R.M., 2018. PROTEINS, INTERFACES, AND CRYO-EM GRIDS. Curr Opin Colloid Interface Sci 34, 1–8. https://doi.org/10.1016/j.cocis.2017.12.009

Glaeser, R.M., 2016. How good can cryo-EM become? Nat Methods 13, 28–32. https://doi.org/10.1038/nmeth.3695
Glaeser, R.M., Han, B.-G., 2017. Opinion: hazards faced by macromolecules when confined to thin aqueous films. Biophys Rep 3, 1–7. https://doi.org/10.1007/s41048-016-0026-3
Gorrec, F., 2009. The MORPHEUS protein crystallization screen. J Appl Cryst 42, 1035–1042. https://doi.org/10.1107/S0021889809042022
Gray, V.E., Hause, R.J., Fowler, D.M., 2017. Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. Genetics 207, 53–61. https://doi.org/10.1534/genetics.117.300064
Grimm, C., Chari, A., Reuter, K., Fischer, U., 2010. A crystallization screen based on alternative polymeric precipitants. Acta Crystallogr D Biol Crystallogr 66, 685–697. https://doi.org/10.1107/S0907444910009005
Grimm, J., Shi, H., Wang, W., Mitchell, A.M., Wingreen, N.S., Huang, K.C., Silhavy, T.J., 2020. The inner membrane protein YhdP modulates the rate of anterograde phospholipid flow in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 117, 26907–26914. https://doi.org/10.1073/pnas.2015556117
Gu, Y., Stansfeld, P.J., Zeng, Y., Dong, H., Wang, W., Dong, C., 2015. Lipopolysaccharide is inserted into the outer membrane through an intramembrane hole, a lumen gate, and the lateral opening of lptd. Structure 23, 496–504. https://doi.org/10.1016/j.str.2015.01.001
Guest, R.L., Guerra, D.S., Wissler, M., Grimm, J., Silhavy, T.J., 2020. YejM modulates activity of the YciM/FtsH protease complex to prevent lethal accumulation of lipopolysaccharide. mBio 11. https://doi.org/10.1128/mBio.00598-20
Gupta, S., Wu, H.C., 1991. Identification and subcellular localization of apolipoprotein N-acyltransferase in Escherichia coli. FEMS Microbiology Letters 78, 37–41. https://doi.org/10.1016/0378-1097(91)90251-5
Hagan, C.L., Silhavy, T.J., Kahne, D., 2011. β-Barrel Membrane Protein Assembly by the Bam Complex. Annual Review of Biochemistry 80, 189–210. https://doi.org/10.1146/annurev-biochem-061408-144611
Hawley, C.A., Watson, C.A., Orth, K., Krachler, A.M., 2013. A MAM7 Peptide-Based Inhibitor of Staphylococcus aureus Adhesion Does Not Interfere with In Vitro Host Cell Function. PLoS ONE 8, 81216. https://doi.org/10.1371/JOURNAL.PONE.0081216
Henderson, R., 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Quarterly Reviews of Biophysics 28, 171–193. https://doi.org/10.1017/S003358350000305X
Herreros, D., Lederman, R.R., Krieger, J., Jiménez-Moreno, A., Martínez, M., Myška, D., Strelak, D., Filipovic, J., Bahar, I., Carazo, J.M., Sanchez, C.O.S., 2021. Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. IUCrJ 8, 992–1005. https://doi.org/10.1107/S2052252521008903
Holmgren, A., Fagerstedt, M., 1982. The in vivo distribution of oxidized and reduced thioredoxin in Escherichia coli. Journal of Biological Chemistry 257, 6926–6930. https://doi.org/10.1016/S0021-9258(18)34518-6
Horne, J.E., Brockwell, D.J., Radford, S.E., 2020. Role of the lipid bilayer in outer membrane protein folding in Gram-negative bacteria. Journal of Biological Chemistry 295, 10340–10367. https://doi.org/10.1074/jbc.REV120.011473
Huebinger, R.M., Do, D.H., Carlson, D.L., Yao, X., Stones, D.H., De Souza Santos, M., Vaz, D.P., Keen, E., Wolf, S.E., Minei, J.P., Francis, K.P., Orth, K., Krachler, A.M., 2020. Bacterial adhesion inhibitor prevents infection in a rodent surgical incision model. https://doi.org/10.1080/21505594.2020.1772652 11, 695–706. https://doi.org/10.1080/21505594.2020.1772652
Hughes, G.W., Hall, S.C.L., Laxton, C.S., Sridhar, P., Mahadi, A.H., Hatton, C., Piggot, T.J., Wotherspoon, P.J., Leney, A.C., Ward, D.G., Jamshad, M., Spana, V., Cadby, I.T.,

Harding, C., Isom, G.L., Bryant, J.A., Parr, R.J., Yakub, Y., Jeeves, M., Huber, D., Henderson, I.R., Clifton, L.A., Lovering, A.L., Knowles, T.J., 2019. Evidence for phospholipid export from the bacterial inner membrane by the Mla ABC transport system. Nature Microbiology 1. https://doi.org/10.1038/s41564-019-0481-y

Hwang, P.M., Choy, W.Y., Lo, E.I., Chen, L., Forman-Kay, J.D., Raetz, C.R.H., Privé, G.G., Bishop, R.E., Kay, L.E., 2002. Solution structure and dynamics of the outer membrane enzyme PagP by NMR. Proceedings of the National Academy of Sciences of the United States of America 99, 13560–13565. https://doi.org/10.1073/pnas.212344499

Inouye, S., Franceschini, T., Sato, M., Itakura, K., Inouye, M., 1983. Prolipoprotein signal peptidase of Escherichia coli requires a cysteine residue at the cleavage site. The EMBO journal 2, 87–91. https://doi.org/10.1002/j.1460-2075.1983.tb01386.x

Isom, G.L., 2017. A role for proteobacterial mammalian cell entry domains in phospholipid trafficking and infection. University of Birmingham.

Isom, G.L., Coudray, N., MacRae, M.R., McManus, C.T., Ekiert, D.C., Bhabha, G., 2020. LetB Structure Reveals a Tunnel for Lipid Transport across the Bacterial Envelope. Cell 181, 653-664.e19. https://doi.org/10.1016/j.cell.2020.03.030

Isom, G.L., Davies, N.J., Chong, Z.-S., Bryant, J.A., Jamshad, M., Sharif, M., Cunningham, A.F., Knowles, T.J., Chng, S.-S., Cole, J.A., Henderson, I.R., 2017. MCE domain proteins: conserved inner membrane lipid-binding proteins required for outer membrane homeostasis. Scientific Reports 7, 8608. https://doi.org/10.1038/s41598-017-09111-6

Jagessar, K.L., Claxton, D.P., Stein, R.A., McHaourab, H.S., 2020. Sequence and structural determinants of ligand-dependent alternating access of a MATE transporter. Proceedings of the National Academy of Sciences of the United States of America 117, 4732–4740. https://doi.org/10.1073/PNAS.1917139117/-/DCSUPPLEMENTAL

Jagessar, K.L., McHaourab, H.S., Claxton, D.P., 2019. The N-terminal domain of an archaeal multidrug and toxin extrusion (MATE) transporter mediates proton coupling required for prokaryotic drug resistance. The Journal of Biological Chemistry 294, 12807. https://doi.org/10.1074/JBC.RA119.009195

Jain, T., Sheehan, P., Crum, J., Carragher, B., Potter, C.S., 2012. Spotiton: A prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. J Struct Biol 179, 68–75. https://doi.org/10.1016/j.jsb.2012.04.020

Joshi, S.M., Pandey, A.K., Capite, N., Fortune, S.M., Rubin, E.J., Sassetti, C.M., 2006. Characterization of mycobacterial virulence genes through genetic interaction mapping.

Josts, I., Stubenrauch, C.J., Vadlamani, G., Mosbahi, K., Walker, D., Lithgow, T., Grinter, R., 2017. The Structure of a Conserved Domain of TamB Reveals a Hydrophobic β Taco Fold. Structure(London, England:1993) 25, 1898. https://doi.org/10.1016/J.STR.2017.10.002

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kamischke, C., Fan, J., Bergeron, J., Kulasekara, H.D., Dalebroux, Z.D., Burrell, A., Kollman, J.M., Miller, S.I., 2019. The Acinetobacter baumannii Mla system and glycerophospholipid transport to the outer membrane. eLife 8. https://doi.org/10.7554/eLife.40171

Kellenberger, E., 1990. The “Bayer bridges” confronted with results from improved electron microscopy methods. Molecular Microbiology 4, 697–705. https://doi.org/10.1111/j.1365-2958.1990.tb00640.x

Kelly, D.F., Dukovski, D., Walz, T., 2008. Monolayer purification: A rapid method for isolating protein complexes for single-particle electron microscopy. Proceedings of the National Academy of Sciences of the United States of America 105, 4703. https://doi.org/10.1073/PNAS.0800867105

Kendall, S.L., Withers, M., Soffair, C.N., Moreland, N.J., Gurcha, S., Sidders, B., Frita, R., Ten Bokum, A., Besra, G.S., Lott, J.S., Stoker, N.G., 2007. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in Mycobacterium smegmatis and Mycobacterium tuberculosis. Molecular Microbiology 65, 684–699. https://doi.org/10.1111/J.1365-2958.2007.05827.X

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., Phillips, D.C., 1958. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. Nature 1958 181:4610 181, 662–666. https://doi.org/10.1038/181662a0

Klepp, L.I., Forrellad, M.A., Osella, A.V., Blanco, F.C., Stella, E.J., Bianco, M.V., Santangelo, M. de la P., Sassetti, C., Jackson, M., Cataldi, A.A., Bigi, F., Morbidoni, H.R., 2012. Impact of the deletion of the six mce operons in Mycobacterium smegmatis. Microbes and Infection 14, 590–599. https://doi.org/10.1016/J.MICINF.2012.01.007

Klepp, L.I., Sabio y Garcia, J., FabianaBigi, 2022. Mycobacterial MCE proteins as transporters that control lipid homeostasis of the cell wall. Tuberculosis 132, 102162. https://doi.org/10.1016/J.TUBE.2021.102162

Koh, Y.-S., Roe, J.-H., 1996. Dual regulation of the paraquat-inducible gene pqi-5 by SoxS and RpoS in Escherichia coli. Molecular Microbiology 22, 53–61. https://doi.org/10.1111/j.1365-2958.1996.tb02655.x

Koh, Y.S., Roe, J.H., 1995. Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus in Escherichia coli. Journal of bacteriology 177, 2673–8. https://doi.org/10.1128/jb.177.10.2673-2678.1995

Kol, M.A., De Kroon, A.I.P.M., Killian, J.A., De Kruijff, B., 2004. Transbilayer Movement of Phospholipids in Biogenic Membranes. Biochemistry 43, 2673–2681. https://doi.org/10.1021/bi036200f

Kolich, L.R., Chang, Y.-T., Coudray, N., Giacometti, S.I., MacRae, M.R., Isom, G.L., Teran, E.M., Bhabha, G., Ekiert, D.C., 2020. Structure of MlaFB uncovers novel mechanisms of ABC transporter regulation. bioRxiv 2020.04.27.064196. https://doi.org/10.1101/2020.04.27.064196

Konovalova, A., Kahne, D.E., Silhavy, T.J., 2017. Outer Membrane Biogenesis. Annual Review of Microbiology 71, 539–556. https://doi.org/10.1146/annurev-micro-090816-093754

Konovalova, A., Mitchell, A.M., Silhavy, T.J., n.d. A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. eLife 5, e15276. https://doi.org/10.7554/eLife.15276

Konovalova, A., Silhavy, T.J., 2015. Outer membrane lipoprotein biogenesis: Lol is not the end. Philosophical Transactions of the Royal Society B: Biological Sciences 370. https://doi.org/10.1098/rstb.2015.0030

Krachler, A.M., Ham, H., Orth, K., 2011. Outer membrane adhesion factor multivalent adhesion molecule 7 initiates host cell binding during infection by Gram-negative pathogens. Proceedings of the National Academy of Sciences 108, 11614–11619. https://doi.org/10.1073/pnas.1102360108

Krachler, A.M., Mende, K., Murray, C., Orth, K., 2012. In vitro characterization of multivalent adhesion molecule 7-based inhibition of multidrug-resistant bacteria isolated from wounded military personnel. Virulence 3, 389. https://doi.org/10.4161/VIRU.20816

Kudva, R., Denks, K., Kuhn, P., Vogt, A., Müller, M., Koch, H.G., 2013. Protein translocation across the inner membrane of Gram-negative bacteria: The Sec and Tat dependent protein transport pathways. Research in Microbiology 164, 505–534. https://doi.org/10.1016/j.resmic.2013.03.016

Kulp, A.J., Sun, B., Ai, T., Manning, A.J., Orench-Rivera, N., Schmid, A.K., Kuehn, M.J., 2015. Genome-Wide Assessment of Outer Membrane Vesicle Production in

Escherichia coli. PLOS ONE 10, e0139200. https://doi.org/10.1371/journal.pone.0139200
Kumar, A., Chandolia, A., Chaudhry, U., Brahmachari, V., Bose, M., 2005. Comparison of mammalian cell entry operons of mycobacteria: In silico analysis and expression profiling. FEMS Immunology and Medical Microbiology 43, 185–195. https://doi.org/10.1016/j.femsim.2004.08.013
Kumar, N., Leonzino, M., Hancock-Cerutti, W., Horenkamp, F.A., Li, P.Q., Lees, J.A., Wheeler, H., Reinisch, K.M., De Camilli, P., 2018. VPS13A and VPS13C are lipid transport proteins differentially localized at ER contact sites. Journal of Cell Biology 217, 3625–3639. https://doi.org/10.1083/JCB.201807019
Kusakizako, T., Claxton, D.P., Tanaka, Y., Ishitani, R., Mchaourab, H.S., Correspondence, O.N., 2019. Structural Basis of H +-Dependent Conformational Change in a Bacterial MATE Transporter. Structure/Folding and Design 27, 293-301.e3. https://doi.org/10.1016/j.str.2018.10.004
Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., Ben-Tal, N., 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic acids research 33. https://doi.org/10.1093/NAR/GKI370
Langmore, J.P., Smith, M.F., 1992. Quantitative energy-filtered electron microscopy of biological molecules in ice. Ultramicroscopy 46, 349–373. https://doi.org/10.1016/0304-3991(92)90024-E
Lee, D.J., Bingle, L.E., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J., Hobman, J.L., 2009. Gene doctoring: a method for recombineering in laboratory and pathogenic Escherichia coli strains. BMC Microbiol 9, 252. https://doi.org/10.1186/1471-2180-9-252
Levine, T.P., 2019. Remote homology searches identify bacterial homologues of eukaryotic lipid transfer proteins, including Chorein-N domains in TamB and AsmA and Mdm31p. BMC Molecular and Cell Biology 20, 43. https://doi.org/10.1186/s12860-019-0226-z
Li, P.Q., Lees, J.A., Patrick Lusk, C., Reinisch, K.M., 2020. Cryo-EM reconstruction of a VPS13 fragment reveals a long groove to channel lipids between membranes. Journal of Cell Biology 219. https://doi.org/10.1083/jcb.202001161
Lim, J., Stones, D.H., Hawley, C.A., Watson, C.A., Krachler, A.M., 2014. Multivalent Adhesion Molecule 7 Clusters Act as Signaling Platform for Host Cellular GTPase Activation and Facilitate Epithelial Barrier Dysfunction. PLOS Pathogens 10, e1004421. https://doi.org/10.1371/JOURNAL.PPAT.1004421
Liu, N., Zhang, J., Chen, Y., Liu, C., Zhang, X., Xu, K., Wen, J., Luo, Z., Chen, S., Gao, P., Jia, K., Liu, Z., Peng, H., Wang, H.W., 2019. Bioactive Functionalized Monolayer Graphene for High-Resolution Cryo-Electron Microscopy. Journal of the American Chemical Society 141, 4016–4025. https://doi.org/10.1021/JACS.8B13038/ASSET/IMAGES/LARGE/JA-2018-13038S_0005.JPEG
Lloubès, R., Cascales, E., Walburger, A., Bouveret, E., Lazdunski, C., Bernadac, A., Journet, L., 2001. The Tol-Pal proteins of the Escherichia coli cell envelope: An energized system required for outer membrane integrity? Research in Microbiology 152, 523–529. https://doi.org/10.1016/S0923-2508(01)01226-8
Low, W.Y., Thong, S., Chng, S.S., 2021. ATP disrupts lipid-binding equilibrium to drive retrograde transport critical for bacterial outer membrane asymmetry. Proceedings of the National Academy of Sciences of the United States of America 118, e2110055118. https://doi.org/10.1073/PNAS.2110055118/SUPPL_FILE/PNAS.2110055118.SAPP.PDF
Lu, B., Benning, C., 2009. A 25-Amino Acid Sequence of the Arabidopsis TGD2 Protein Is Sufficient for Specific Binding of Phosphatidic Acid. The Journal of Biological Chemistry 284, 17420. https://doi.org/10.1074/JBC.M109.016014

Malinverni, J.C., Silhavy, T.J., 2009. An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. Proceedings of the National Academy of Sciences of the United States of America 106, 8009–8014. https://doi.org/10.1073/pnas.0903229106
Mann, D., Fan, J., Somboon, K., Farrell, D.P., Muenks, A., Tzokov, S.B., DiMaio, F., Khalid, S., Miller, S.I., Bergeron, J.R.C., 2021. Structure and lipid dynamics in the maintenance of lipid asymmetry inner membrane complex of A. baumannii. Commun Biol 4, 1–9. https://doi.org/10.1038/s42003-021-02318-4
Marmont, L.S., Rich, J.D., Whitney, J.C., Whitfield, G.B., Almblad, H., Robinson, H., Parsek, M.R., Harrison, J.J., Lynne Howell, P., 2017. Oligomeric lipoprotein PelC guides Pel polysaccharide export across the outer membrane of Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences of the United States of America 114, 2892–2897. https://doi.org/10.1073/PNAS.1613606114/-/DCSUPPLEMENTAL
Masilamani, R., Cian, M.B., Dalebroux, Z.D., 2018. Salmonella Tol-Pal reduces outer membrane glycerophospholipid levels for envelope homeostasis and survival during bacteremia. Infection and Immunity 86. https://doi.org/10.1128/IAI.00173-18
May, K.L., Silhavy, T.J., 2018. The Escherichia coli phospholipase PldA regulates outer membrane homeostasis via lipid signaling. mBio 9. https://doi.org/10.1128/mBio.00379-18
McClelland, M., Sanderson, K.E., Clifton, S.W., Latreille, P., Porwollik, S., Sabo, A., Meyer, R., Bieri, T., Ozersky, P., McLellan, M., Harkins, C.R., Wang, C., Nguyen, C., Berghoff, A., Elliott, G., Kohlberg, S., Strong, C., Du, F., Carter, J., Kremizki, C., Layman, D., Leonard, S., Sun, H., Fulton, L., Nash, W., Miner, T., Minx, P., Delehaunty, K., Fronick, C., Magrini, V., Nhan, M., Warren, W., Florea, L., Spieth, J., Wilson, R.K., 2004. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of Salmonella enterica that cause typhoid. Nature Genetics 36, 1268–1274. https://doi.org/10.1038/ng1470
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. urn:issn:0021-8898 40, 658–674. https://doi.org/10.1107/S0021889807021206
Mcpherson, A., 2001. A comparison of salts for the crystallization of macromolecules. Protein Sci 10, 418–422.
Messens, J., Collet, J.-F., Belle, K.V., Brosens, E., Loris, R., Wyns, L., 2007. The Oxidase DsbA Folds a Protein with a Nonconsecutive Disulfide *. Journal of Biological Chemistry 282, 31302–31307. https://doi.org/10.1074/jbc.M705236200
Miroux, B., Walker, J.E., 1996. Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. Journal of Molecular Biology 260, 289–298. https://doi.org/10.1006/JMBI.1996.0399
Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. urn:issn:0907-4449 67, 355–367. https://doi.org/10.1107/S0907444911001314
Nakane, T., Kimanius, D., Lindahl, E., Scheres, S.H., 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. eLife 7, e36861. https://doi.org/10.7554/eLife.36861
Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P.M.G.E., Grigoras, I.T., Malinauskaite, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E.V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., Hardwick, S.W., Chirgadze, D.Y., Murshudov, G., Aricescu, A.R., Scheres, S.H.W., 2020. Single-particle cryo-EM at atomic resolution. Nature 2020 587:7832 587, 152–156. https://doi.org/10.1038/s41586-020-2829-0
Nakayama, T., Zhang-Akiyama, Q.-M., 2017. pqiABC and yebST, Putative mce Operons of Escherichia coli, Encode Transport Pathways and Contribute to Membrane Integrity. Journal of bacteriology 199, e00606-16. https://doi.org/10.1128/JB.00606-16

Newby, Z.E.R., O'Connel, J.D., Gruswitz, F., Hays, F.A., Harries, W.E.C., Harwood, I.M., Ho, J.D., Lee, J.K., Savage, D.F., Miercke, L.J.W., Stroud, R., 2009. A general protocol for the crystallization of membrane proteins for X-ray structural investigation. Nat Protoc 4, 619–637. https://doi.org/10.1038/nprot.2009.27

Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J.L., Stuart, D.I., Perrakis, A., 2005. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. Acta Crystallogr D Biol Crystallogr 61, 1426–1431. https://doi.org/10.1107/S0907444905024984

Newstead, S., Ferrandon, S.B., Iwata, S.O., 2008a. Rationalizing α-helical membrane protein crystallization. Protein Science 17, 466–472. https://doi.org/10.1110/PS.073263108

Newstead, S., Hobbs, J., Jordan, D., Carpenter, E., Iwata, S., 2008b. Insights into outer membrane protein crystallisation. Mol Membr Biol 25, 631–638. https://doi.org/10.1080/09687680802526574

Okuda, S., Freinkman, E., Kahne, D., 2012. Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. Science 338, 1214–1217. https://doi.org/10.1126/SCIENCE.1228984/SUPPL_FILE/OKUDA.SM.PDF

Okuda, S., Sherman, D.J., Silhavy, T.J., Ruiz, N., Kahne, D., 2016. Lipopolysaccharide transport and assembly at the outer membrane: The PEZ model. Nature Reviews Microbiology 14, 337–345. https://doi.org/10.1038/nrmicro.2016.25

Okuda, S., Tokuda, H., 2011. Lipoprotein Sorting in Bacteria. Annual Review of Microbiology 65, 239–259. https://doi.org/10.1146/annurev-micro-090110-102859

Orlova, E.V., Saibil, H.R., 2011. Structural Analysis of Macromolecular Assemblies by Electron Microscopy. Chem Rev 111, 7710–7748. https://doi.org/10.1021/cr100353t

Osawa, T., Kotani, T., Kawaoka, T., Hirata, E., Suzuki, K., Nakatogawa, H., Ohsumi, Y., Noda, N.N., 2019. Atg2 mediates direct lipid transfer between membranes for autophagosome formation. Nature Structural & Molecular Biology 2019 26:4 26, 281–288. https://doi.org/10.1038/s41594-019-0203-4

Pandey, A.K., Sassetti, C.M., 2008. Mycobacterial persistence requires the utilization of host cholesterol. Proc Natl Acad Sci U S A 105, 4376–4380. https://doi.org/10.1073/pnas.0711159105

Pantelic, R.S., Meyer, J.C., Kaiser, U., Baumeister, W., Plitzko, J.M., 2010. Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. Journal of structural biology 170, 152–156. https://doi.org/10.1016/J.JSB.2009.12.020

Parker, J.L., Newstead, S., 2012. Current trends in α-helical membrane protein crystallization: An update. Protein Science 21, 1358–1365. https://doi.org/10.1002/PRO.2122

Passmore, L.A., Russo, C.J., 2016. Chapter Three - Specimen Preparation for High-Resolution Cryo-EM, in: Crowther, R.A. (Ed.), Methods in Enzymology, The Resolution Revolution: Recent Advances In CryoEM. Academic Press, pp. 51–86. https://doi.org/10.1016/bs.mie.2016.04.011

Pathak, D.T., Wei, X., Bucuvalas, A., Haft, D.H., Gerloff, D.L., Wall, D., 2012. Cell Contact–Dependent Outer Membrane Exchange in Myxobacteria: Genetic Determinants and Mechanism. PLoS Genetics 8, e1002626. https://doi.org/10.1371/journal.pgen.1002626

Petiti, M., Serrano, B., Faure, L., Lloubes, R., Mignot, T., Duché, D., 2019. Tol Energy-Driven Localization of Pal and Anchoring to the Peptidoglycan Promote Outer-Membrane Constriction. Journal of Molecular Biology 431, 3275–3288. https://doi.org/10.1016/j.jmb.2019.05.039

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci 30, 70–82. https://doi.org/10.1002/pro.3943

Pomorski, T.G., Menon, A.K., 2016. Lipid somersaults: Uncovering the mechanisms of protein-mediated lipid flipping. Progress in Lipid Research 64, 69–84. https://doi.org/10.1016/j.plipres.2016.08.003

Powers, M.J., Simpson, B.W., Trent, M.S., 2020. The Mla pathway in Acinetobacter baumannii has no demonstrable role in anterograde lipid transport. eLife 9. https://doi.org/10.7554/eLife.56571
Powers, M.J., Stephen Trent, M., 2019. Intermembrane transport: Glycerophospholipid homeostasis of the Gram-negative cell envelope. Proceedings of the National Academy of Sciences of the United States of America 116, 17147–17155. https://doi.org/10.1073/pnas.1902026116
Powers, M.J., Trent, M.S., 2018. Phospholipid retention in the absence of asymmetry strengthens the outer membrane permeability barrier to last-resort antibiotics. Proceedings of the National Academy of Sciences of the United States of America 115, E8518–E8527. https://doi.org/10.1073/pnas.1806714115
Radchenko, M., Symersky, J., Nie, R., Lu, M., 2015. Structural basis for the blockade of MATE multidrug efflux pumps. Nature Communications 2015 6:1 6, 1–11. https://doi.org/10.1038/ncomms8995
Raetz, C.R., Dowhan, W., 1990. Biosynthesis and function of phospholipids in Escherichia coli. The Journal of biological chemistry 265, 1235–1238.
Rank, L., Herring, L.E., Braunstein, M., 2021. Evidence for the mycobacterial mce4 transporter being a multiprotein complex. Journal of Bacteriology 203. https://doi.org/10.1128/JB.00685-20/SUPPL_FILE/JB.00685-20-S0002.XLSX
Rao, S., Bates, G.T., Matthews, C.R., Newport, T.D., Vickery, O.N., Stansfeld, P.J., 2020. Characterizing Membrane Association and Periplasmic Transfer of Bacterial Lipoproteins through Molecular Dynamics Simulations. Structure 28, 475-487.e3. https://doi.org/10.1016/j.str.2020.01.012
Ravelli, R.B.G., Nijpels, F.J.T., Henderikx, R.J.M., Weissenberger, G., Thewessem, S., Gijsbers, A., Beulen, B.W.A.M.M., López-Iglesias, C., Peters, P.J., 2020. Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification. Nat Commun 11, 2563. https://doi.org/10.1038/s41467-020-16392-5
Rohou, A., Grigorieff, N., 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology 192, 216–221. https://doi.org/10.1016/J.JSB.2015.08.008
Roier, S., Zingl, F.G., Cakar, F., Durakovic, S., Kohl, P., Eichmann, T.O., Klug, L., Gadermaier, B., Weinzerl, K., Prassl, R., Lass, A., Daum, G., Reidl, J., Feldman, M.F., Schild, S., 2016. A novel mechanism for the biogenesis of outer membrane vesicles in Gram-negative bacteria. Nature Communications 7. https://doi.org/10.1038/ncomms10515
Rothman, J.E., Kennedy, E.P., 1977. Rapid transmembrane movement of newly synthesized phospholipids during membrane assembly. Proceedings of the National Academy of Sciences of the United States of America 74, 1821–1825. https://doi.org/10.1073/pnas.74.5.1821
Rubinstein, J.L., Guo, H., Ripstein, Z.A., Haydaroglu, A., Au, A., Yip, C.M., Di Trani, J.M., Benlekbir, S., Kwok, T., 2019. Shake-it-off: a simple ultrasonic cryo-EM specimen-preparation device. Acta Crystallogr D Struct Biol 75, 1063–1070. https://doi.org/10.1107/S2059798319014372
Ruiz, F., Hazemann, I., Mitschler, A., Joachimiak, A., Schneider, T., Karplus, M., Podjarny, A., IUCr, 2004. The crystallographic structure of the aldose reductase–IDD552 complex shows direct proton donation from tyrosine 48. Acta Crystallographica Section D Biological Crystallography 60, 1347–1354. https://doi.org/10.1107/S0907444904011370
Ruiz, N., Davis, R.M., Kumar, S., 2021. YhdP, TamB, and YdbH Are Redundant but Essential for Growth and Lipid Homeostasis of the Gram-Negative Outer Membrane. mBio 12. https://doi.org/10.1128/MBIO.02714-21/FORMAT/EPUB
Ruiz, N., Kahne, D., Silhavy, T.J., 2009. Transport of lipopolysaccharide across the cell envelope: The long road of discovery. Nature Reviews Microbiology 7, 677–683. https://doi.org/10.1038/nrmicro2184

Sankaran, K., Wu, H.C., 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. Journal of Biological Chemistry 269, 19701–19706.
Santangelo, M. de la P., Blanco, F., Campos, E., Soria, M., Bianco, M.V., Klepp, L., Alito, A., Zabal, O., Cataldi, A., Bigi, F., 2009. Mce2R from Mycobacterium tuberculosis represses the expression of the mce2 operon. Tuberculosis 89, 22–28. https://doi.org/10.1016/J.TUBE.2008.09.002
Santangelo, M.P., Blanco, F.C., Bianco, M.V., Klepp, L.I., Zabal, O., Cataldi, A.A., Bigi, F., 2008. Study of the role of Mce3R on the transcription of mce genes of Mycobacterium tuberculosis. BMC Microbiology 8, 1–8. https://doi.org/10.1186/1471-2180-8-38/TABLES/2
Scheres, S.H.W., 2015. Semi-automated selection of cryo-EM particles in RELION-1.3. Journal of Structural Biology 189, 114–122. https://doi.org/10.1016/j.jsb.2014.11.010
Scheres, S.H.W., 2012a. A Bayesian View on Cryo-EM Structure Determination. Journal of Molecular Biology 415, 406–418. https://doi.org/10.1016/J.JMB.2011.11.010
Scheres, S.H.W., 2012b. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology 180, 519–530. https://doi.org/10.1016/J.JSB.2012.09.006
Shrivastava, R., Chng, S.S., 2019. Lipid trafficking across the Gram-negative cell envelope. Journal of Biological Chemistry 294, 14175–14184. https://doi.org/10.1074/jbc.AW119.008139
Shrivastava, R., Jiang, X., Chng, S.-S., 2017. Outer membrane lipid homeostasis via retrograde phospholipid transport in Escherichia coli. Molecular Microbiology 106, 395–408. https://doi.org/10.1111/mmi.13772
Singh, P., Katoch, V.M., Mohanty, K.K., Chauhan, D.S., 2016. Analysis of expression profile of mce operon genes (mce1, mce2, mce3 operon) in different Mycobacterium tuberculosis isolates at different growth phases. The Indian Journal of Medical Research 143, 487. https://doi.org/10.4103/0971-5916.184305
Sklar, J.G., Wu, T., Gronenberg, L.S., Malinverni, J.C., Kahne, D., Silhavy, T.J., 2007. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. Proc Natl Acad Sci U S A 104, 6400–6405. https://doi.org/10.1073/pnas.0701579104
Snijder, H.J., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K.H., Varheij, H.M., Egmond, M.R., Dekker, N., Dijkstra, B.W., 1999. Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. Nature 401, 717–721. https://doi.org/10.1038/401717a0
Sohlenkamp, C., Geiger, O., 2015. Bacterial membrane lipids: diversity in structures and pathways. FEMS Microbiology Reviews 40, 133–159. https://doi.org/10.1093/femsre/fuv008
Stock, J.B., Rauch, B., Roseman, S., 1977. Periplasmic space in Salmonella typhimurium and Escherichia coli. J Biol Chem 252, 7850–7861.
Suits, M.D.L., Sperandeo, P., Dehò, G., Polissi, A., Jia, Z., 2008. Novel Structure of the Conserved Gram-Negative Lipopolysaccharide Transport Protein A and Mutagenesis Analysis. Journal of Molecular Biology 380, 476–488. https://doi.org/10.1016/J.JMB.2008.04.045
Sutterlin, H.A., Shi, H., May, K.L., Miguel, A., Khare, S., Huang, K.C., Silhavy, T.J., 2016. Disruption of lipid homeostasis in the Gram-negative cell envelope activates a novel cell death pathway. Proceedings of the National Academy of Sciences of the United States of America 113, E1565-74. https://doi.org/10.1073/pnas.1601375113
Tan, B.K., Bogdanov, M., Zhao, J., Dowhan, W., Raetz, C.R.H., Guan, Z., 2012. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. Proceedings of the National Academy of Sciences of the United States of America 109, 16504–16509. https://doi.org/10.1073/pnas.1212797109
Tanaka, Y., Hipolito, C.J., Maturana, A.D., Ito, K., Kuroda, T., Higuchi, T., Katoh, T., Kato, H.E., Hattori, M., Kumazaki, K., Tsukazaki, T., Ishitani, R., Suga, H., Nureki, O., 2013.

Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. Nature 2013 496:7444 496, 247–251. https://doi.org/10.1038/nature12014
Tanford, C., 1978. The Hydrophobic Effect and the Organization of Living Matter. Science 200, 1012–1018. https://doi.org/10.1126/science.653353
Tang, X., Chang, S., Qiao, W., Luo, Q., Chen, Y., Jia, Z., Coleman, J., Zhang, K., Wang, T., Zhang, Z., Zhang, C., Zhu, X., Wei, X., Dong, C., Zhang, X., Dong, H., 2020. Structural insights into outer membrane asymmetry maintenance in Gram-negative bacteria by MlaFEDB. Nature Structural and Molecular Biology. https://doi.org/10.1038/s41594-020-00532-y
Taniguchi, N., Matsuyama, S.I., Tokuda, H., 2005. Mechanisms underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed β-barrel structures. Journal of Biological Chemistry 280, 34481–34488. https://doi.org/10.1074/jbc.M507388200
Tefsen, B., Bos, M.P., Beckers, F., Tommassen, J., De Cock, H., 2005. MsbA is not required for phospholipid transport in Neisseria meningitidis. Journal of Biological Chemistry 280, 35961–35966. https://doi.org/10.1074/jbc.M509026200
Tekaia, F., Gordon, S.V., Garnier, T., Brosch, R., Barrell, B.G., Cole, S.T., 1999. Analysis of the proteome of Mycobacterium tuberculosis in silico. Tubercle and Lung Disease 79, 329–342. https://doi.org/10.1054/TULD.1999.0220
Terwilliger, T.C., Sobolev, O.V., Afonine, P.V., Adams, P.D., 2018. Automated map sharpening by maximization of detail and connectivity. urn:issn:2059-7983 74, 545–559. https://doi.org/10.1107/S2059798318004655
Thomason, L.C., Costantino, N., Court, D.L., 2007. E. coli Genome Manipulation by P1 Transduction. Current Protocols in Molecular Biology 79, 1.17.1-1.17.8. https://doi.org/10.1002/0471142727.mb0117s79
Thong, S., Ercan, B., Torta, F., Fong, Z.Y., Wong, H.Y.A., Wenk, M.R., Chng, S.-S., 2016. Defining key roles for auxiliary proteins in an ABC transporter that maintains bacterial outer membrane lipid asymmetry. eLife 5. https://doi.org/10.7554/eLife.19042
Vieni, C., Coudray, N., Isom, G.L., Bhabha, G., Ekiert, D.C., 2022. Role of Ring6 in the Function of the E. coli MCE Protein LetB. Journal of Molecular Biology 434, 167463. https://doi.org/10.1016/j.jmb.2022.167463
Vindal, V., Ashwantha Kumar, E., Ranjan, A., 2008. Identification of operator sites within the upstream region of the putative mce2R gene from mycobacteria. FEBS Letters 582, 1117–1122. https://doi.org/10.1016/J.FEBSLET.2008.02.074
Vollmer, W., Bertsche, U., 2008. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. Biochimica et Biophysica Acta (BBA) - Biomembranes, Structural proteomics of the cell envelope of Gram-negative bacteria 1778, 1714–1734. https://doi.org/10.1016/j.bbamem.2007.06.007
Wade, R.H., 1992. A brief look at imaging and contrast transfer. Ultramicroscopy 46, 145–156. https://doi.org/10.1016/0304-3991(92)90011-8
Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T.R., Lill, P., Gatsogiannis, C., Raunser, S., 2019. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Communications Biology 2019 2:1 2, 1–13. https://doi.org/10.1038/s42003-019-0437-z
Wang, X., Peterson, J.H., Bernstein, H.D., 2021. Bacterial Outer Membrane Proteins Are Targeted to the Bam Complex by Two Parallel Mechanisms. mBio 12. https://doi.org/10.1128/MBIO.00597-21
Wei, H., Dandey, V.P., Zhang, Z., Raczkowski, A., Rice, W.J., Carragher, B., Potter, C.S., 2018. Optimizing “Self-Wicking” Nanowire Grids. J Struct Biol 202, 170–174. https://doi.org/10.1016/j.jsb.2018.01.001
Williams, R.C., Glaeser, R.M., 1972. Ultrathin carbon support films for electron microscopy. Science (New York, N.Y.) 175, 1000–1001. https://doi.org/10.1126/SCIENCE.175.4025.1000

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J., Kahne, D., 2005. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell 121, 235–245. https://doi.org/10.1016/j.cell.2005.02.015

Yakushi, T., Masuda, K., Narita, S.I., Matsuyama, S.I., Tokuda, H., 2000. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. Nature Cell Biology 2000 2:4 2, 212–218. https://doi.org/10.1038/35008635

Yamaguchi, K., Yu, F., Inouye, M., 1988. A single amino acid determinant of the membrane localization of lipoproteins in E. coli. Cell 53, 423–432. https://doi.org/10.1016/0092-8674(88)90162-6

Yeow, J., Tan, K.W., Holdbrook, D.A., Chong, Z.S., Marzinek, J.K., Bond, P.J., Chng, S.S., 2018. The architecture of the OmpC–MlaA complex sheds light on the maintenance of outer membrane lipid asymmetry in Escherichia coli. Journal of Biological Chemistry 293, 11325–11340. https://doi.org/10.1074/jbc.RA118.002441

Yip, K.M., Fischer, N., Paknia, E., Chari, A., Stark, H., 2020. Atomic-resolution protein structure determination by cryo-EM. Nature 2020 587:7832 587, 157–161. https://doi.org/10.1038/s41586-020-2833-4

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., Agard, D.A., 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature Methods 2017 14:4 14, 331–332. https://doi.org/10.1038/nmeth.4193

Zhong, E.D., Bepler, T., Berger, B., Davis, J.H., 2021. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. Nature Methods 2021 18:2 18, 176–185. https://doi.org/10.1038/s41592-020-01049-4

Zhou, Z., White, K.A., Polissi, A., Georgopoulos, C., Raetz, C.R.H., 1998. Function of Escherichia coli MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. Journal of Biological Chemistry 273, 12466–12475. https://doi.org/10.1074/jbc.273.20.12466