AN INVESTIGATION INTO LIPID EXCHANGE IN THE MAINTENANCE OF LIPID ASYMMETRY PATHWAY

by

PETER J. WOTHERSPOON

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

February 2022

I.Abstract

The Maintenance of Lipid Asymmetry (Mla) pathway is a 6 component system implicated in the transport of phospholipids between the Gram-negative bacterial membranes, specifically in the transport of ectopic phospholipids from the outer leaflet of the outer membrane to the inner membrane. However, papers from 2019 proposed a potential role in phospholipid export. As phospholipid export to the outer membrane currently represents the last unexplored frontier of outer membrane biogenesis, the prospect of identifying a pathway responsible presents an attractive topic for further investigation. Moreover, deficiencies in Mla activity have been shown to reduce virulence and increase susceptibility to membrane destabilising agents, cementing Mla as a topic of significant interest to the field of antimicrobial research.

As such, the original objective of this project was to investigate lipid transport between the inner membrane component of the Mla pathway and the periplasmic chaperone with the intent of understanding the discrepancies between the findings in the 2019 papers and the original categorisation of Mla function. However, subsequent publications have provided clear evidence that Mla plays no role in lipid export.

Thus we chose to shift the main objective of the project to understanding the mechanism by which ATPase activity in MlaFE results in lipid exchange between MlaC and MlaD. We present structural data of a stalled intermediate complex of MlaC and MlaD, showing the binding interface, from which we have identified several residues with significant effects on lipid exchange between the two proteins. Based on these results, and existing understanding of MlaFEDB conformational changes we propose a mechanism by which ATPase induced conformational changes in MlaD might result in the expulsion of lipid from the MlaC binding pocket.

II.Acknowledgements

I would like to acknowledge and thank the BBSRC MIBTP for providing both the opportunity and funding for this research project.

Pooja Sridhar, thank you for teaching me everything I know about lab work and putting up with my antics since I was a rotation student.

Tim Knowles, thank you for providing support and insight, in addition to being an amazing supervisor over the past 3 years.

Gareth Hughes, thanks mate.

Thank you to Andy Lovering, Ian Cadby, Simon Caulton, Richard Meek, Hannah Walters-Morgan, Amber Wilson and Rebecca Parr for assistance with crystal fishing and collecting diffraction data as well as general advice regarding crystallography, ITC and MST.

David Hardy, thanks for being a constant source of good advice both in and out of the lab.

Steph Nestorow, thank you for always being helpful.

Giedre Ratkeviciute, thank you for general day to day help and running the follow-up *in vivo* mutation experiments we discussed, sorry for increasing your workload.

Ben Cooper, thank you for advice regarding MST and for maintaining the order of the Cooper-Knowles Lab.

Thank you to Savva Christos and TJ Ragan from the Midlands Regional Cryo-EM Facility for advice and assistance regarding cryo-EM.

Thank you to Aneika Leney and Jeddidiah Bellamy-Carter for assistance and advice regarding native mass spectrometry.

Thank you to Zoe Stroud, Rachael Grime, Steve Hall and Charles Moore-Kelly for a lot of help during my first lab year.

Rachel Holyfield, thank you for being a very good rotation student.

Thank you Mr. Chalk for the generous donation of 1 whisker.

III. Contents Listings

Table of Contents

List of Figures

List of Tables

List of Abbreviations

β-OG	-	n-octyl β-D-glucopyranoside
AAA+	-	ATPases associated with diverse cellular activities
ABC	-	ATP-binding cassette
ADP	-	Adenosine diphosphate
ATP	-	Adenosine triphosphate
AUC	-	Analytical ultra-centrifugation
BDO-CL	-	1',3'-bis[1,2-dioleoyl-sn-glycero-3-phospho]-glycerol
CFU	-	Colony forming units
CL	-	Cardiolipin
dansyl	-	5-(dimethylamino)naphthalene-1-sulfonyl
dansyl-PE	-	1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(5-dimethylamino-1-naphthalenesulfonyl)
DDM	-	*n*-dodecyl-β-D-maltoside
E. coli	-	*Escherichia coli*
EDTA	-	Ethylenediaminetetraacetic acid
EM	-	Electron microscopy
FASII	-	Type II fatty acid synthesis
FRET	-	Förster resonance energy transfer
GDP	-	Guanosine diphosphate
GlcNAc	-	N-acetylglucosamine
GTP	-	Guanosine triphosphate
IPTG	-	Isopropyl-β-D-thiogalactopyranoside
K_D	-	Equilibrium dissociation constant
Kdo	-	Keto-deoxyoctulosonate
LB	-	Lysogeny broth

LLG	-	Log-likelihood gain
LoG	-	Laplacian of gaussian
Lol	-	Localization of lipoproteins
MCE	-	Mammalian cell entry
MST	-	Microscale thermophoresis
Mw	-	Molecular weight
NBD	-	Nucleotide binding domain
NTF2	-	Nuclear Transport Factor 2
PAGE	-	Polyacrylamide gel electrophoresis
*p*Bpa	-	*para*-benzoyl-L-phenylalanine
PE	-	Phosphatidylethanolamine
PG	-	Phosphatidylglycerol
POPE	-	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POPG	-	1-palmitoyl-2-oleoylglycero-3-phosphoglycerol
POTRA	-	Polypeptide-transport associated
QCM-D	-	Quartz crystal microbalance with dissipation monitoring
SDS	-	Sodium dodecyl sulphate
SEC	-	Size exclusion chromatography
SPR	-	Surface plasmon resonance
SRP	-	Signal recognition particle
TCEP	-	Tris(2-carboxyethyl)phosphine
TEM	-	Transmission electron microscopy
TFZ	-	Translation function Z-score
TLC	-	Thin-layer chromatography
TMD	-	Transmembrane domain
Tris	-	Tris(hydroxymethyl)aminomethane
UDP	-	Uridine diphosphate
UMP	-	Uridine monophosphate

1.Introduction

The Maintenance of Lipid Asymmetry (Mla) pathway is a multicomponent pathway spanning the Gram-negative bacterial periplasm, with inner and outer membrane components as well as a periplasmic chaperone. It has been observed to have roles in virulence, vesicle blebbing and maintenance of the outer membrane barrier function[6, 50, 106]. Initial characterisation of Mla function by Malinverni and Silhavy suggested that its primary role and mode of function was the transport of ectopic phospholipids from the outer leaflet of the outer membrane to the inner membrane[106]. Subsequent reports have also identified a possible role in transport of phospholipids in the direction of the outer membrane and suggested an involvement in outer membrane biogenesis[71, 77]. As phospholipid export remains the last unexplained component of outer membrane biogenesis investigation of this role is of critical importance.

Moreover, while MlaFEDB is classified as an ATP-binding cassette (ABC) protein, it is structurally distinct from the seven ABC protein superfamilies[182]. As such, the mechanisms of MlaFEDB ATPase activity are not explained within the paradigms of existing ABC transporters and determination of the ATP induced conformational changes that drive lipid exchange between MlaC and MlaD are of great interest.

In this thesis, we aim to investigate the mechanisms by which lipids are exchanged between MlaFEDB and MlaC with emphasis on the interaction between MlaC and MlaD, which is currently not well understood, to the point that not even the binding orientation or stoichiometry is known. To that end, this section will serve to introduce the Gram-negative outer membrane, including its role, composition and biogenesis; the protein pathways involved in outer membrane assembly, with reference to ABC proteins involved in interaction with acylated molecules; the synthesis of

phospholipids as the substrate of the Mla pathway; the structure of MlaC and MlaD; and a brief note on the role of specific residues in protein – protein interactions.

1.1.Gram-Negative Bacteria

Bacteria are prokaryotes, unicellular organisms with no nucleus and limited compartmentalisation when compared to eukaryotic cells, delineated from their external environment by a lipidic membrane. Bacterial cells have long been further subcategorised by their response to Gram-staining. The distinction being a result of the thick external peptidoglycan layer present in some bacteria, which is capable of retaining a hexamethyl pararosaniline chloride primary stain. As such the distinction between a Gram-positive and Gram-negative bacteria is reliant on the properties of the bacterial membrane[9] [Figure 1.1]. Generally Gram-positive bacteria have a single membrane supplemented with the aforementioned thick external peptidoglycan layer and Gram-negatives have a comparatively thinner layer of peptidoglycan and two lipid membranes.

Gram-negative bacteria are becoming an emergent global health risk with the advent of widespread multidrug resistant bacteria. Antibiotics function by targeting essential processes for bacterial viability. They can inhibit peptidoglycan synthesis like beta-lactams or vancomycin, inhibit protein synthesis by targeting the bacterial ribosome like aminoglycosides, disrupt the cell membrane like polymyxins or interact with a variety of other processes, however, they must be able to access their target to be efficacious[15]. The presence of the outer membrane in Gram-negative bacteria confers an intrinsic resistance to many broad spectrum antibiotics that are effectively employed clinically to treat Gram-positive infections. Modifications to the protein or lipid composition of the outer membrane can confer further resistance to a wider

range of antibiotics through an increase in drug efflux pumps or a reduction in membrane permeability[39].

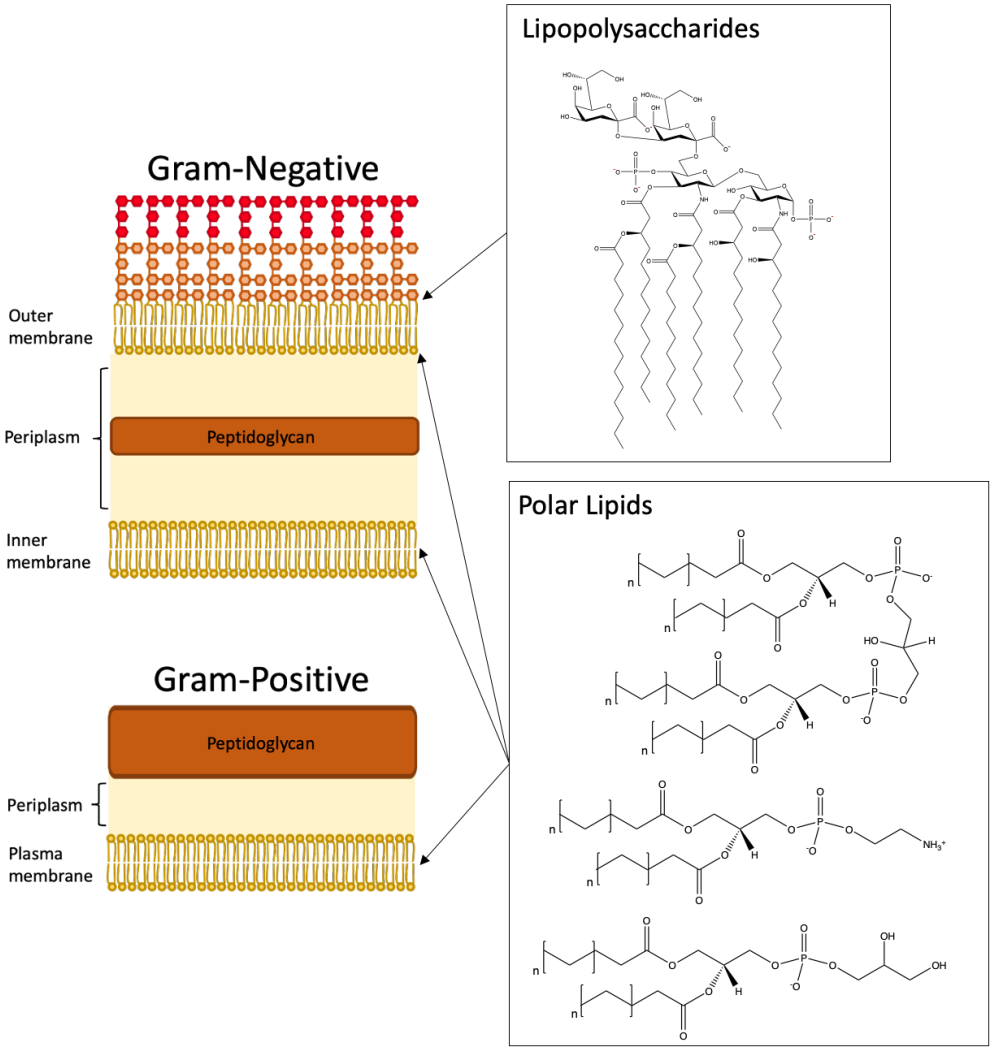

Figure 1.1 - Diagram depicting the structure of the Gram-negative and Gram-positive bacterial membranes. The outer leaflet of the Gram-negative outer membrane is composed primarily of lipopolysaccharides, while the inner leaflet and inner membrane are composed primarily of polar lipids. Gram-Positive bacteria on the other hand lack an outer membrane.

Beyond conferring increased resistance to antimicrobial agents the outer membrane is the main interface though which Gram-negative bacteria interact with their external environment. As such it is inundated with porins, exporter and importer proteins to

allow the uptake of small ions and molecules as well as a whole host of other proteins involved in secretion, chemotaxis, motility, pathogenesis and a multitude of other processes more specific to individual bacterial species[83]. Thus, the formation and maintenance of this membrane is a key physiological process for Gram-negative bacteria. It is necessary for them to synthesise, transport, assemble and organise the structural and functional components of this membrane in manner conducive to maintaining effective barrier function, as dysregulation of any of these processes may compromise its role as a barrier and interface. Many proteins and protein pathways involved in these roles have been shown to be ubiquitous across Gram-negative bacteria and their role often necessitates that some protein components are surface exposed. It is for these reasons that investigation into the pathways responsible for outer membrane maintenance and biogenesis is of great interest to the field of antimicrobial research[15]. Surface exposed drug targets that may compromise a major intrinsic resistance mechanism of Gram-negative bacteria could themselves be used as antimicrobial agents or allow for an increase in the scope of existing antimicrobial agents.

1.2.Components of the Bacterial Cell Envelope

The bacterial plasma membrane is based on a bi-layered sheet of amphipathic phospholipid molecules [Figure 1.2]. The phospholipids are arranged with inward pointing hydrophobic acyl chains and outward facing polar phosphate head groups[61]. This is a stable arrangement that avoids exposing the hydrophobic acyl chains to the surrounding aqueous environment resulting is a lipid bilayer with an internal hydrophobic environment and a polar surface, which makes diffusion of most charged molecules through the membrane energetically unfavourable. Although this arrangement is permeable to small uncharged molecules and can be easily disrupted by amphipathic molecules, such as detergents. This phenomenon underpins the

barrier function of the plasma membrane. The lipid bi-layer makes up one of the two major components of the bacterial cell envelope, the other being the peptidoglycan layer. The peptidoglycan layer is assembled outside of the inner membrane from a precursor known as Lipid II, which itself is assembled from N-acetylglucosamine and N-acetylmuramic acid pentapeptide precursors in the cytoplasm [Figure 1.3]. The peptidoglycan layer provides rigidity to resist osmotic stress as well as acting as a mesh that excludes particles larger than 2 nm[189].

Gram-positive bacteria have a 'thick', 30-100 nm, peptidoglycan layer outside their inner membrane, which is in some cases additionally decorated with complex sugars and fatty acids as with bacteria of the *Mycobacterium* genus. Gram-negative bacteria on the other hand have a 'thin' peptidoglycan layer outside their inner membrane, and a second lipid membrane outside of the peptidoglycan layer known as the outer membrane, tethered to the peptidoglycan by Braun's lipoprotein[167].

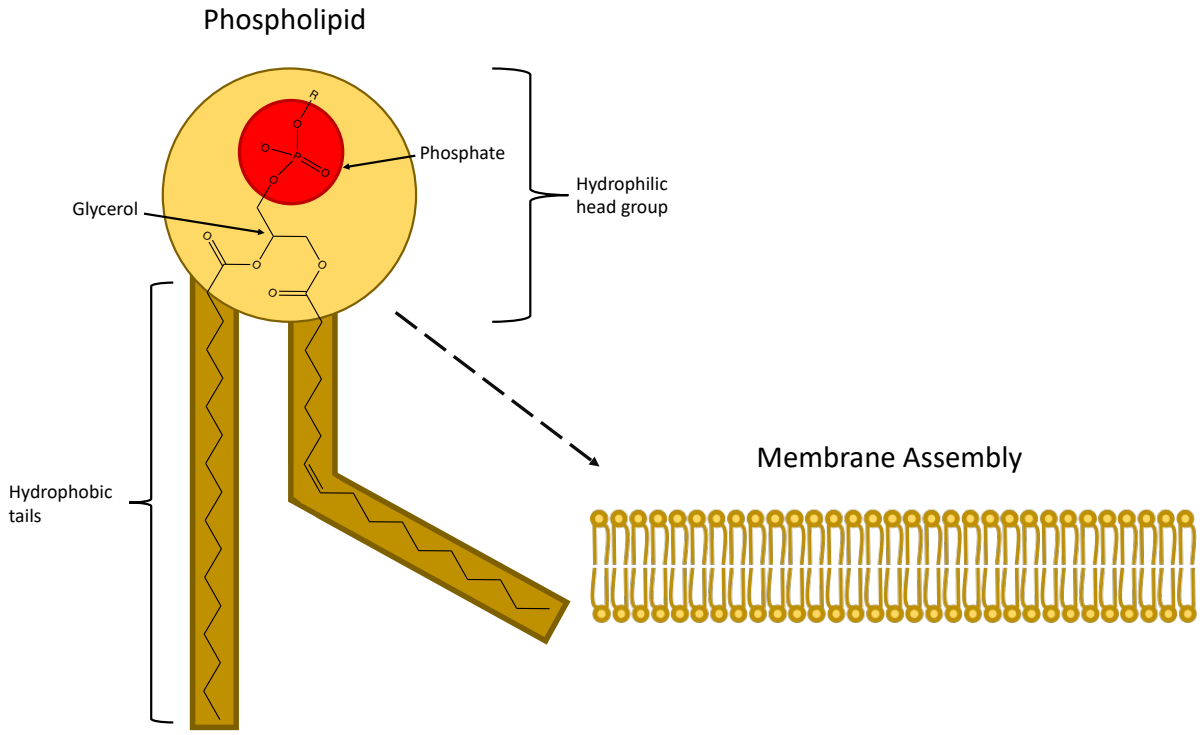

Figure 1.2 - Diagram depicting the structure of a standard diacyl phospholipid as well as its assembly into a biological membrane.

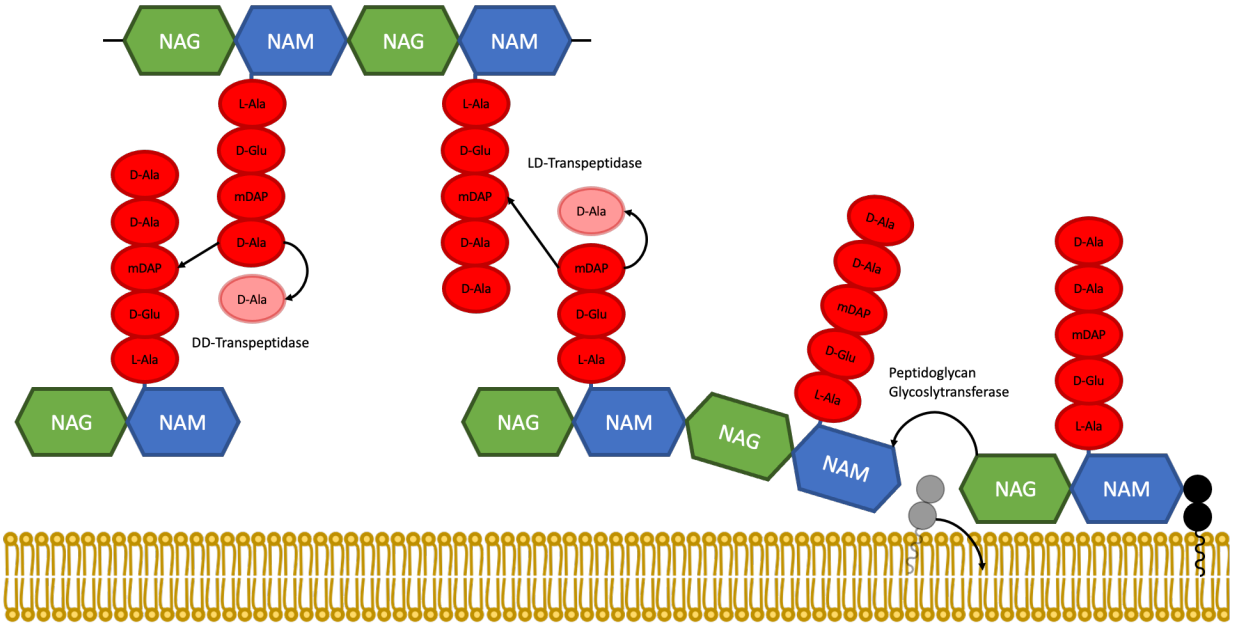

Figure 1.3 - A simplified schematic of peptidoglycan assembly from Lipid II. Transglycosylation extends the NAG-NAM chain linearly and then crosslinking occurs between the pentapeptides.

The outer membrane consists of an asymmetric bilayer of lipopolysaccharides and phospholipids[117] organised in a similar manner to the inner plasma membrane, however, under ideal conditions the outer layer of this lipid sheet contains only lipopolysaccharides and the inner only phospholipids[33]. Lipopolysaccharides are large molecules based around a core, usually hexa-acylated, lipid referred to as lipid A decorated with a long chain glycan polymer that can be substantially varied in its monomer components[66, 157] [Figure 1.4]. The long glycan polymers of lipopolysaccharides recruit doubly charged cations to stabilise a quasi-crystalline structure that significantly slows the diffusion of hydrophobic solutes[122, 123]. Thus, this bilayer acts as a supplementary permeability barrier to the inner membrane and is markedly less permeable to many common antibiotics, detergents, bile salts and protein defence factors that would otherwise be capable of disrupting the inner membrane[39, 143]. Moreover, this two membrane system is host to multidrug efflux transporters, such as the AcrAB-TolC complex (TolC-like complexes) which spans the

inner and outer membranes, and is capable of opposing the influx of agents that challenge cell viability by actively removing them from the periplasmic space before they are able to penetrate the inner membrane. This effectively makes the periplasmic space an “air-lock” for chemicals that threaten cell viability[181]. AcrAB-TolC and other IM efflux transporters are also capable of drug efflux from the cytoplasm in the event that they penetrate the IM.

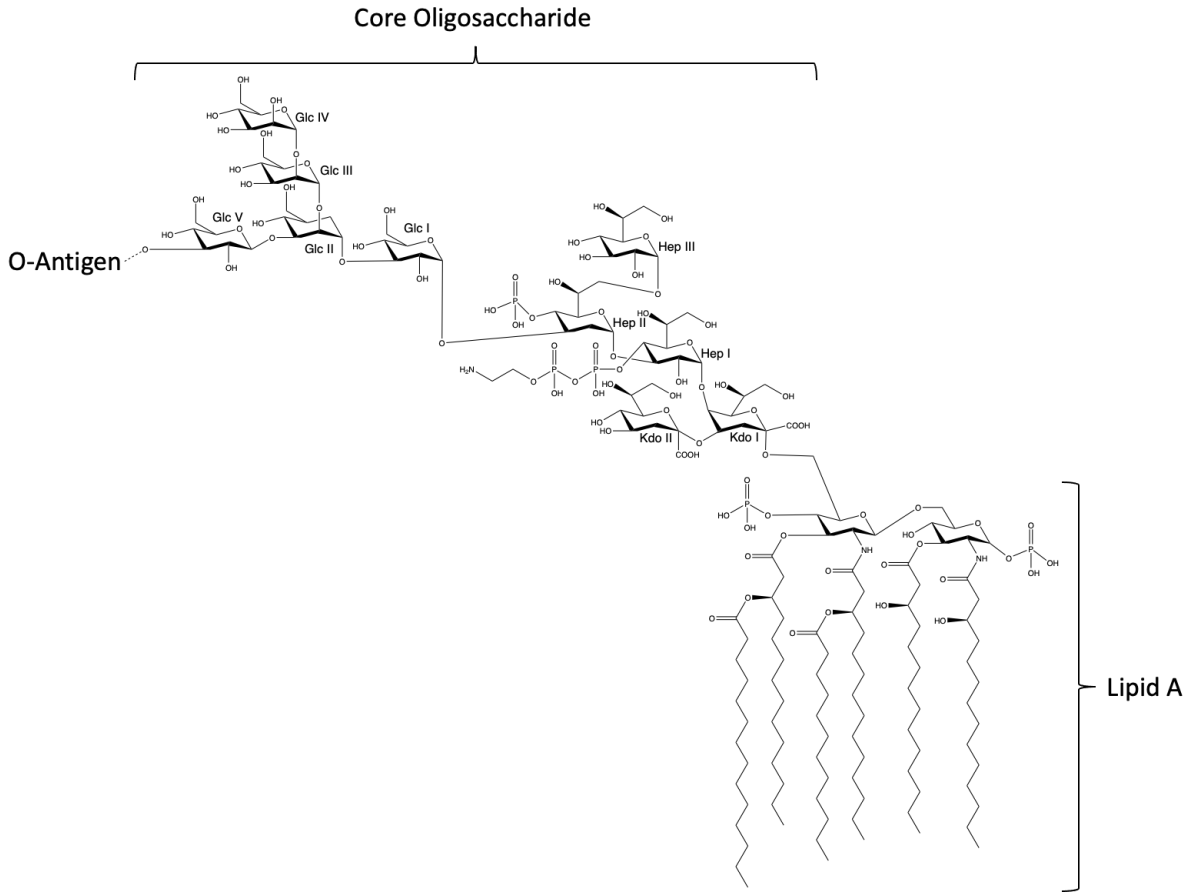

Figure 1.4 - The general structure of an LPS molecule. The O-antigen can be highly variable even among different strains of the same species and this is simply shown as “O-antigen”. The molecule depicted here has an *E. coli* R1-type core oligosaccharide attached to Lipid A.

It is worth considering that the barrier properties of the OM that make it impermeable to exogenous toxins also exclude many key ions and nutrients that would otherwise be passively absorbed. As such the OM is also host to many porins to facilitate the

passive transport of various molecules required for cellular function[11], such as the major general porins, OmpC/OmpF in *E. coli*. These porins are expressed at high copy number and are possibly the most abundant integral outer membrane proteins. These porins, however, also represent a point of access for several classes of hydrophobic anti-biotics such as β-lactams and fluoroquinolones[67]. Gram-negative bacteria can subvert the issue of compromising their own OM with passive import proteins through deficiencies or mutations in general porins that have associations with multidrug resistance. Overall, the combination of an outer membrane with reduced permeability, efflux transporters and general porins result in a barrier capable of excluding toxins without impeding the key cellular function of acquiring metabolites, and the major components of the outer membrane that contribute to this are the lipid and lipopolysaccharide bilayer and the outer membrane proteins.

1.3.Biogenesis of the Gram-negative Outer Membrane

These components of the outer membrane are synthesised in the cytoplasm and at the inner membrane. As such they, must then be transported and integrated into the outer membrane. This means proteins, lipids and lipopolysaccharides which begin their synthesis in the cytoplasm must be transported through the inner membrane and peptidoglycan layer. They must then be assembled at their target destination in their active conformations without disrupting existing structural components of the OM[200]. This process comes with added complexity in that the periplasmic space is devoid of any standard cellular energy source, be it ATP or the proton motive force[126]. Investigations into the mechanisms that permit the assembly of OM components under these conditions have been conducted and the processes and pathways involved in the synthesis and export of lipopolysaccharides and membrane proteins are generally well established. Likewise, the processes of phospholipid synthesis are well understood, however, the mechanisms by which lipids are transported from the

inner membrane to the outer membrane are still in contention[164, 167]. The following sub-sections will discuss current models for the synthesis and export of OM components.

1.3.1.Biogenesis of LPS

Lipopolysaccharide (LPS) synthesis in *E. coli* can be thought to occur in two phases, synthesis of the core Lipid A molecule and synthesis and attachment of the O-antigen. Export of LPS occurs somewhat concurrently with synthesis, however, the final stage of export to the outer membrane occurs subsequently to synthesis. Synthesis begins with the Lpx machinery in the cytoplasm and a uridine diphosphate (UDP)-linked N-acetylglucosamine (GlcNAc) precursor molecule. This precursor molecule is built on by several phases of acylation carried out by soluble Lpx proteins as shown in[24, 73] [Figure 1.5]. Later phases of this synthesis are tied to transport of the nascent LPS molecule, as they are carried out by the peripheral membrane proteins LpxH and LpxB, the former of which incorporates the precursor molecule into the inner leaflet of the inner membrane[16, 125]. The remaining components of the Lpx pathway are integral membrane proteins and conduct further acylation and modifications at the inner leaflet of the inner membrane[26, 51, 158].

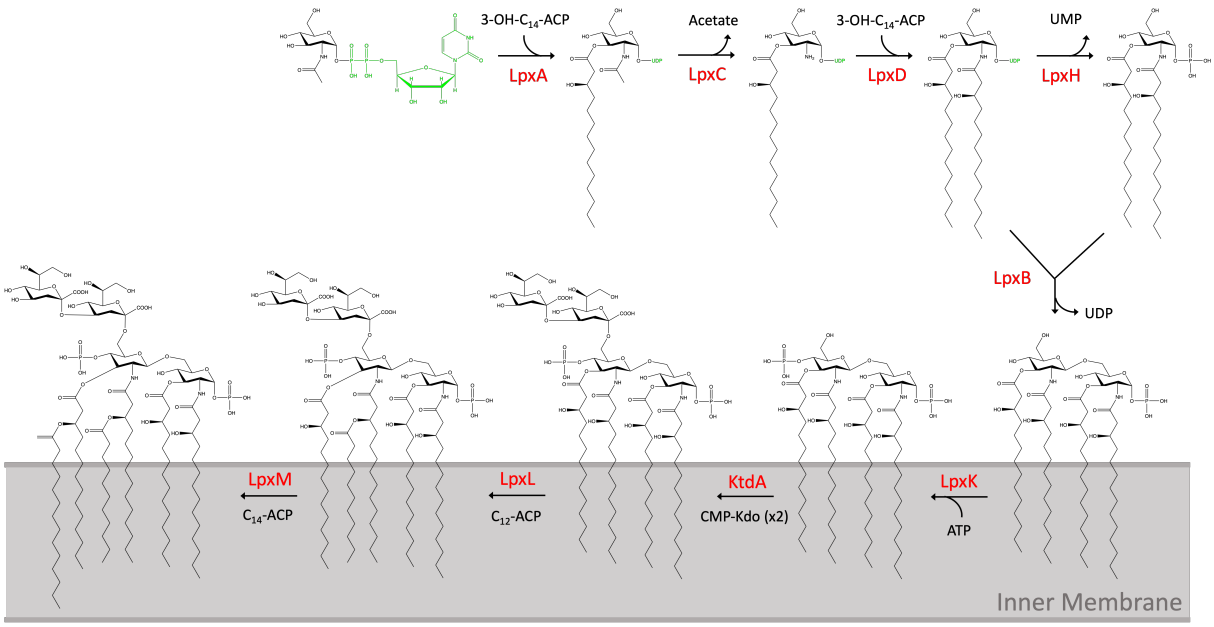

Figure 1.5 - Depiction of the LPS biosynthesis pathway in the cytoplasm and at the IM. UDP-GlcNAc is sequentially modified by Lpx proteins in a synthesis process tied to membrane integration.

At this point attachment of the core oligosaccharides occurs. Generally the core oligosaccharide is separated into the inner core and outer core, with the inner core being more conserved and the outer core being moderately variable across even different strains of the same bacteria. *E. coli* has 5 different core oligosaccharides referred to as R1, R2, R3, R4 and K12. Due to the variability of the outer core there is an array of membrane-associated enzymes involved in the addition of each unit of the oligosaccharide. Formation of the outer core is the final step that occurs on the cytoplasmic facing side of the inner membrane[193]. At this stage the ATP-binding flippase MsbA is responsible for flipping the LPS precursor to the periplasmic leaflet of the inner membrane[210]. This role was also identified by Raetz and his associates by investigating the subcellular localization of LPS precursors in MsbA knockouts. Further extensions to the oligosaccharide by way of addition of the O-antigen is then conducted on the periplasmic facing side of the inner membrane before the translocation of the completed LPS to its final destination at the outer membrane. This

comes with the caveat that the O-antigen must also be synthesised and transported to the periplasmic facing side of the inner membrane.

Assembly of the O-antigen has been shown to occur through three distinct pathways one of which is specific to a single strain of S.enterica. These pathways are generally referred to by the proteins involved, those being the Wzx/Wzy dependent pathway, the Wzm/Wzt dependent pathway and the synthase dependent pathway. The first pathway to be discovered, the Wzx/Wzy dependent pathway was first identified by Subbaiah & Stocker in 1964 when they noticed mutations which resulted in *S.Typhimurium* with LPS deficient in O-antigen[176]. Lui et *al*. later determined that the *wzx* gene encoded a transporter of UDP-linked O-subunits responsible for flipping them to the periplasmic leaflet of the inner membrane[101]. Polymerisation of the O-antigen is then conducted by the Wzy glycosyltransferase, a fact not confirmed until Woodward et *al.* reconstructed the pathway *in vitro* in 2010[199]. Discovery of the Wzm/Wzt dependent pathway came when Kopmann & Jann demonstrated that some *E. coli* serotypes synthesised their O-antigen by a different mechanism[87]. The Wzm/Wzt dependent pathway is characterised by cytosolic polymerisation of the O-antigen followed by export of the completed O-antigen by the Wzm/Wzt ABC transporter. The synthase dependent pathway was identified in 1996 by Keenleyside and Whitfield and is specific to a strain of *S.enterica*, wherein the O-antigen is simultaneously synthesised and exported[80]. Common among all these pathways is that the polysaccharide precursors of the O-antigen are assembled attached to a polyisoprenoid carrier. The O-antigen is then ligated to a terminal sugar of the core oligosaccharide by the ligase WaaL[188].

From this point the completed LPS must be transported from the inner membrane to the outer membrane. This is performed by the Lpt machinery consisting of seven core proteins, LptA-G [Figure 1.6]. LptBCFG forms a subassembly at the inner membrane

and LptDE at the outer membrane with LptA bridging the two sub-assemblies[18, 172]. At the inner membrane, an LptB dimer acts as the nucleotide binding domain for the transmembrane LptFG components of the Lpt machinery. These components form an ATPase that ties ATP hydrolysis to the transport of LPS from the inner membrane to a periplasmic bridge formed by LptA and domains of LptC and LptD[104, 172]. The bridge domains all fold in a β-jellyroll motif, with a hydrophobic groove capable of binding the acyl-chains of LPS and shuttling it to the C-terminal, transmembrane domain, of LptD. More recently Li et *al.* have also provided structural evidence for the mechanism that ties ATPase activity to LPS exchange between LptBFG and LptC using vanadate stalled complexes[97]. They showed that the uncharacteristically outward open conformation of LptBFG allows lateral flow of LPS into the binding cavity. A transmembrane helix of LptC supplements the LptFG binding pocket and separates the transmembrane helices of LptF from LptG in the open configuration, maintaining the wide outward open configuration. ATP binding results in the closing of that cavity, effectively "squeezing" the LPS out and up towards the β-jellyroll of LptF and LptC, where it is passed by LptA to LptD [Figure 1.7]. The C-terminal domain of LptD forms a 26 stranded transmembrane β-barrel plugged with LptE which allows for the final incorporation of LPS into the outer membrane through an intermembrane hole in LptD, which allows lateral flow of LPS from LptD into the membrane[172]. A somewhat complete picture of the molecular mechanism of the Lpt pathway has been provided by structures of LptDE presented by Botos et *al.* and $LptB_2FG$ presented by Luo et *al.* in combination with evidence derived from crosslinking LPS to LptA provided by Okuda et *al.*[18, 104, 126].

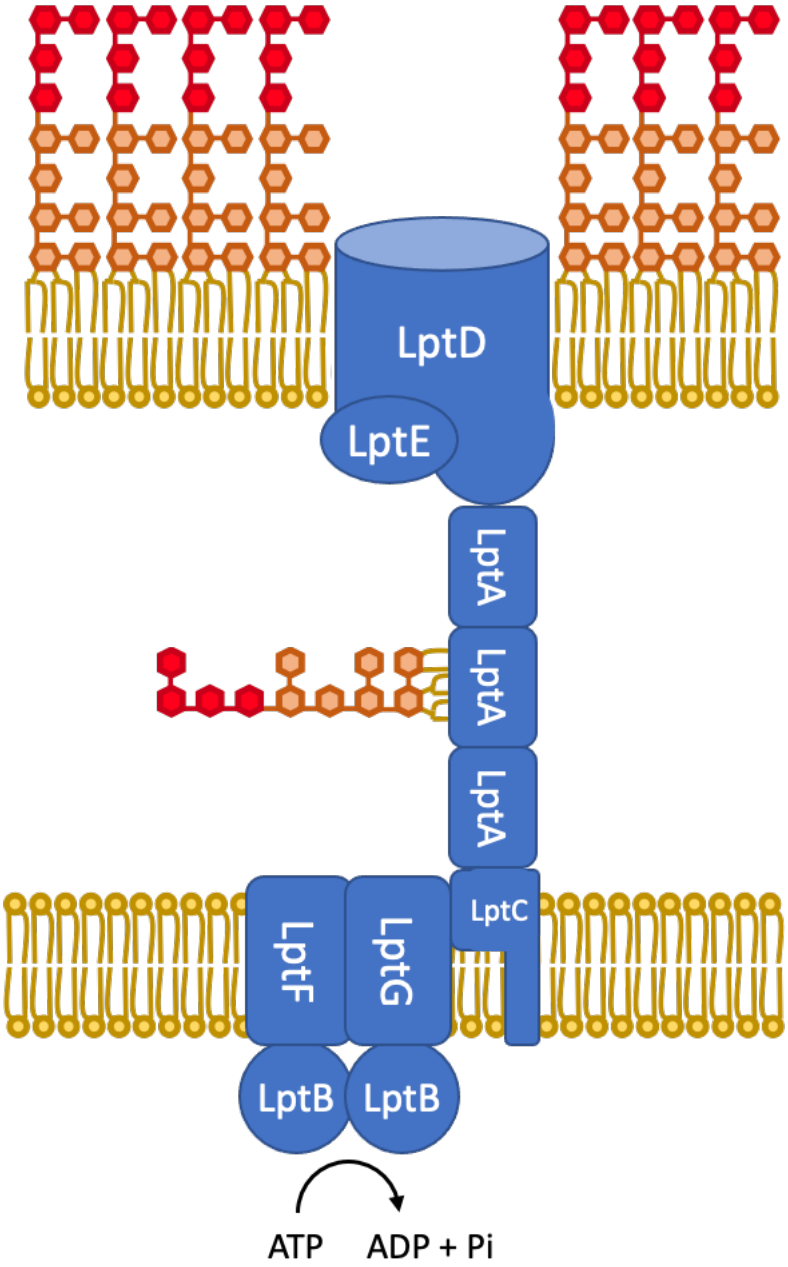

Figure 1.6 - Simplified diagram of the Lpt system involved in the export of LPS to the OM. LPS transport is tied to ATP hydrolysis by $LptB_2FG$ at the IM.

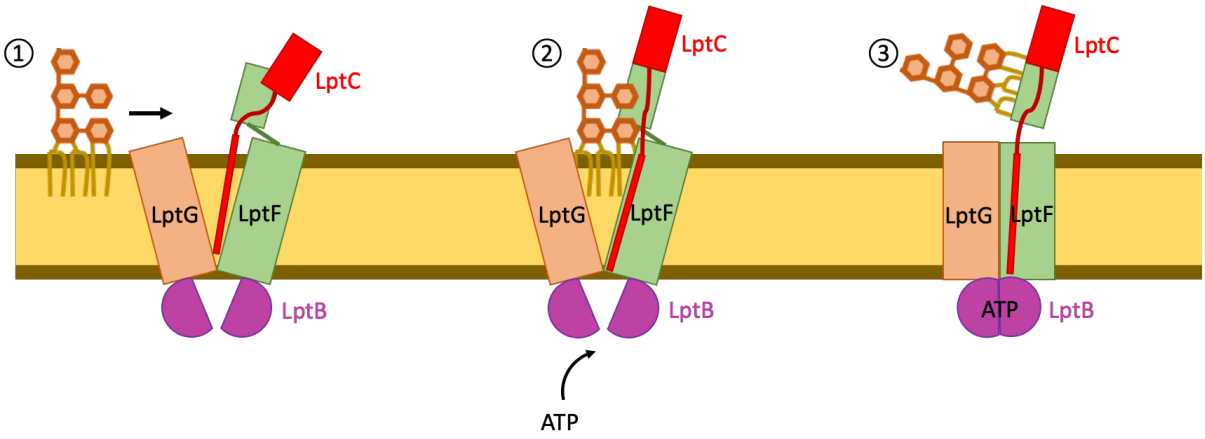

Figure 1.7 - Proposed mechanism for Lpt activity. (1) A transmembrane helix of LptC maintains the wide outward open configuration of LptFG allowing lateral flow of LPS into the binding site. (2) LPS enters the binding cavity and shifts the transmembrane helix of LptC aligning the β-jellyroll domains of LptC and LptF. (3) Binding of ATP results in the collapse of the LPS binding pocket and it is "squeezed" out to the β-jellyroll domain of LptF. Hydrolysis of bound ATP reopens the LPS binding pocket and the cycle repeats. LPS molecules are pushed along the β-jellyroll "rail" towards LptC and eventually into LptA by the next LPS molecule.

1.3.2.Biogenesis of OMPs & Lipoproteins

Synthesis and export of membrane proteins to the inner and outer membranes in bacteria begins with the standard processes of protein synthesis. Coding regions of DNA are transcribed to mRNA by RNA polymerase and associated ancillary proteins[209]. Proteins are then synthesised in accordance with the sequence encoded by their respective mRNA transcripts by the bacterial ribosome[173, 194]. The bacterial ribosome recognises the Shine-Dalgarno sequence on mRNA and initiates protein synthesis by assembling with initiator tRNA and fMet complementary to the start codon[93]. During elongation, the nascent protein is often sequestered by a chaperone such as the trigger factor protein or signal recognition particle (SRP), which prevents premature folding and directs the nascent protein to further export or folding machinery[156].

Many of the components of the Mla pathway find themselves incorporated into or exported beyond the inner membrane. The SEC translocon is heavily implicated in both of these roles as it is thought to be responsible for the vast majority of inner membrane integration and translocation[25, 166]. However, some proteins are integrated exclusively or with the assistance of YidC, and in more rare cases fully folded proteins may be transported across the inner membrane by the Tat translocon[47, 151, 165] [Figure 1.8].

With regards to integral inner membrane proteins, the vast majority are primarily composed of hydrophobic α-helixes due the mechanisms of their integration into the membrane by SEC. SEC is a multiprotein complex consisting of the ATPase SecA, which provides energy for the movement of poly-peptides through the SecYEG translocon, and the SecYEG translocon itself which forms a transmembrane channel through which an unfolded poly-peptide chain can be passed. The poly-peptide is

passed to SEC by the N-terminal signal sequence which stays bound in the translocation pore of SEC until translocation is complete at which point it is released into the membrane by a lateral gate[47]. This signal sequence generally consists of a short positively charged N-terminal region (referred to as the n-region) followed by a longer hydrophobic region (referred to as the h-region) and folds into an α-helix. Translocation through SEC may continue until the C-terminal of the poly-peptide or until a stop transfer sequence is reached. In this way, SEC can act as both an insertase for integral membrane proteins at the inner membrane and a translocase for proteins bound for the periplasm, outer membrane or secretion[47].

YidC on the other hand is responsible for the SEC-independent insertion of small membrane spanning proteins with short periplasmic domains, as well as assisting in the SEC-dependent insertion of larger proteins likely at the stage of lateral transfer of the transmembrane domains into the membrane from SEC as well as in helix packing[118, 151]. In SEC-independent insertion, YidC is believed to allow the formation of an intermediately inserted membrane protein by providing an accessible hydrophobic cavity for the transmembrane region of a protein to bind. Dissociation from YidC then results in the full insertion of the membrane spanning region of the protein.

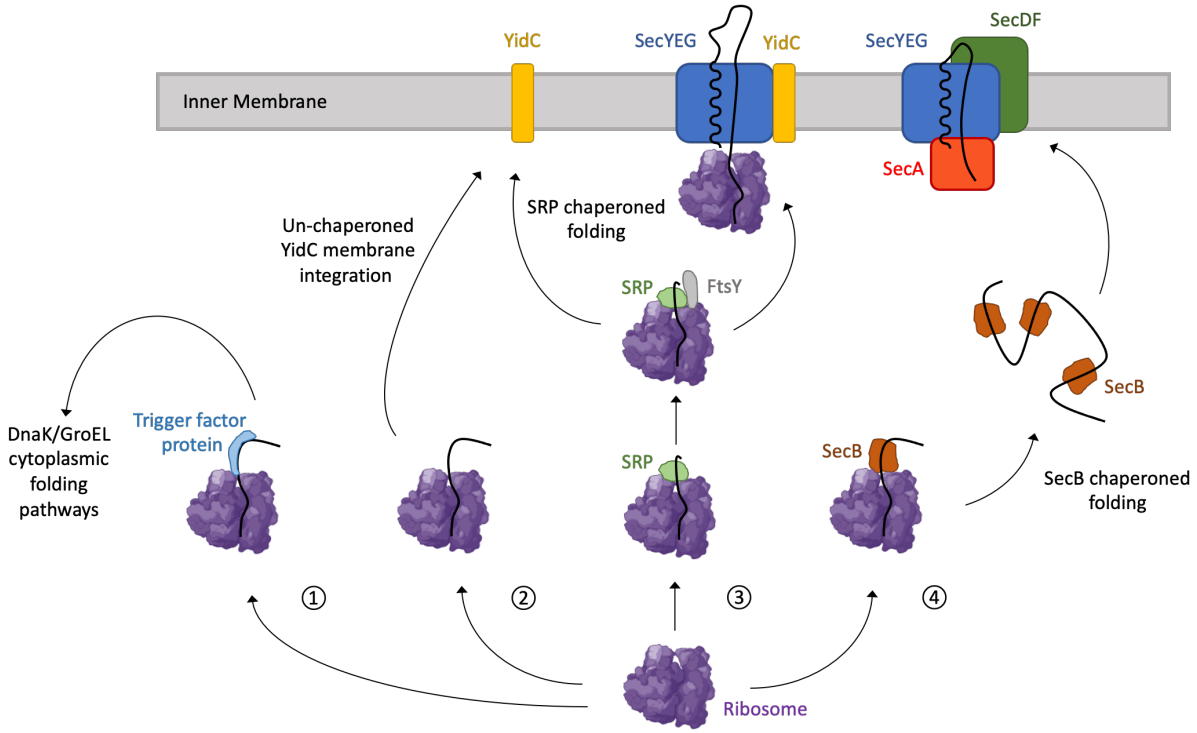

Figure 1.8 - Schematic diagram depicting protein synthesis and transport. During synthesis of proteins by the bacterial ribosome (purple) the nascent chain can be sequestered by chaperone proteins and targeted to different transport/folding apparatus. (1) The nascent chain is sequestered by the Trigger factor protein (light blue) and targeted to DnaK or GroEL for folding in the cytoplasm. (2) The nascent chain encounters membrane insertase YidC (yellow) and is integrated into the IM. (3) The nascent chain is sequestered by the SRP (light green) and with the assistance of FtsY (grey) is targeted to YidC or SecYEG (blue) (and YidC) for membrane integration or export to the periplasm. Sec translocation occurs concurrently with synthesis.(4) The nascent chain is sequestered by SecB. SecB ensures that premature folding of the polypeptide chain does not occur and chaperones it to the SecYEG translocon. The soluble ATPase SecA (red) couples ATP hydrolysis to protein translocation. SecDF uses the PMF to complete translocation in combination with SecA. It is possible for the nascent chain to instead be sequestered by SecA and subsequently chaperoned to SecYEG.

The outer membrane component of the Mla pathway is a lipoprotein. Lipoproteins are associated with the membrane in a different manner to transmembrane proteins and are fully transported into the periplasmic space by SEC, initially tethered to the membrane only by the signal peptide region which is released laterally into the bilayer by SEC[21]. The lipoprotein is then modified by the diacylglycerol transferase Lgt, which recognises a conserved lipobox signal sequence and adds diacylglycerol to conserved cysteine sulfhydryl groups[152] [Figure 1.9]. The lipobox is a 4 amino acid

sequence consisting of the conserved cysteine and 3 preceding residues. Most often these residues are leucine at position -3, alanine (or serine) at position -2 and alanine (or glycine) at position -1, resulting in a sequence similar to -LAAC-[105]. Once the protein is tethered to the membrane by diacylglycerol, the lipoprotein specific peptidase LspA can cleave the membrane tethered N-terminal SEC signal peptide and the new N-terminal can be acylated by the phospholipid transacylase, Lnt. As previously stated in bacterial species other than *E. coli* the Tat pathway can sometime fill in the role of SEC in translocating lipoproteins into the periplasm[120].

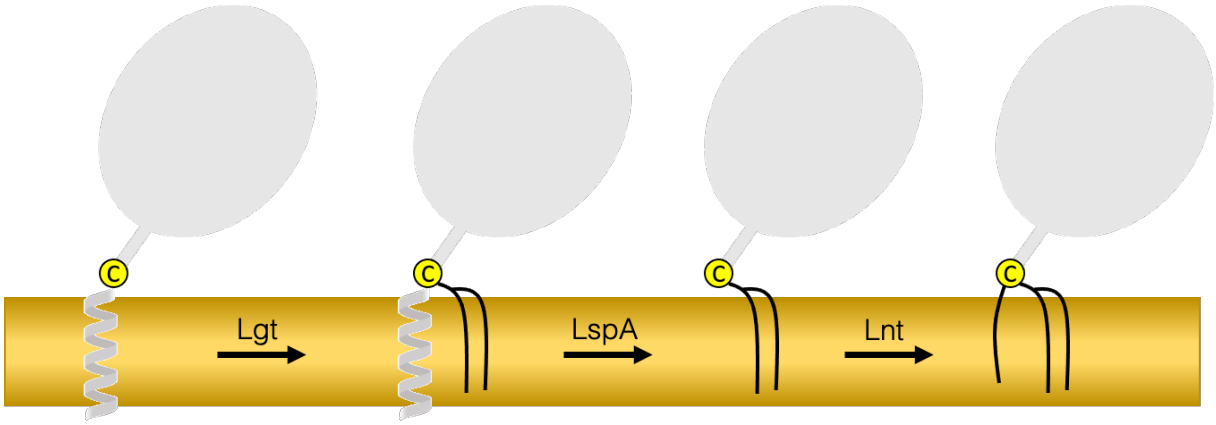

Figure 1.9 - Processing pathway of lipobox proteins. Lgt modifies the lipobox cysteine with diacylglycerol. LpsA cleaves the signal peptide leaving the conserved cysteine as the new N-terminal. Lnt then catalyses the phospholipid dependent acylation of the α-amino group of the N-terminal cysteine.

Lipoproteins at the outer membrane must be transferred there from the inner membrane. This process is handled by the localization of lipoproteins (Lol) export pathway, which acts on lipoproteins lacking an inner membrane retention signal. It has been shown that an Asp residue at position 2 (adjacent to the acylated cysteine) is the major determining signal, however, the position 3 residue can contribute to Lol avoidance[78]. LolCDE at the inner membrane is an ATPase complex responsible for releasing lipoproteins from the inner membrane and passing them to the periplasmic chaperone LolA [Figure 1.10]. An interaction between LolA and LolC is believed to be

required for this exchange[78]. A 2018 paper by Kaplan et *al.* provided crystallographic evidence for the mechanisms by which LolA interacts with LolC. A hairpin loop of LolC, which inserts into the binding pocket of LolA acts as a “hook” to recruit and prime the binding pocket. They suggested that the lipoproteins may then compete with the hairpin loop for the LolA binding pocket. The chaperone LolA contains a hydrophobic cavity for binding the lipoprotein acyl modifications allowing for energetically favourable transport across the periplasm[206]. Once a lipoprotein has been passed to LolA in the periplasm, LolA chaperones the lipoprotein to LolB at the outer membrane. LolB itself is an outer membrane lipoprotein. Narita et *al*. demonstrated that the exchange of lipoproteins between LolA and LolB is mediated by a greater affinity of LolB for the acyl modifications. Once exchanged to LolB, lipoproteins are released into the inner leaflet of the OM[120].

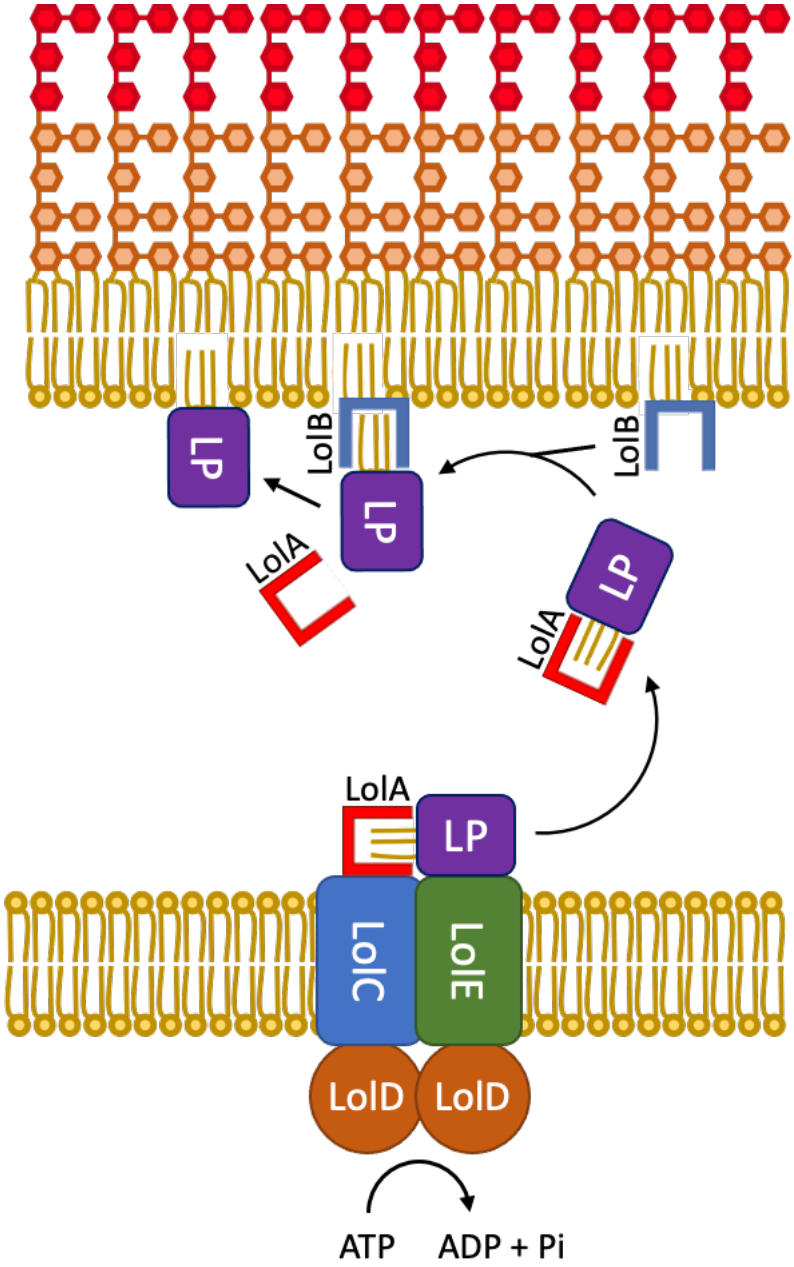

Figure 1.10 - Diagram of lipoprotein export to the OM by the Lol pathway. The LolCDE ATPase couples ATP hydrolysis to the removal of lipoproteins (purple) from the IM. The lipoprotein is then exchanged to the periplasmic chaperone LolA. LolA transports the lipoprotein to LolB at the OM and the lipoprotein is subsequently released into the membrane.

While the are no integral outer membrane proteins among those in the Mla pathway, Mla activity is reliant on expression of an outer membrane porin. Integral outer membrane proteins of this type also use the SEC translocon for transportation into the periplasm. However, once in the periplasm they are separated from their membrane tethered signal sequence by a signal peptidase and chaperoned to the outer membrane still unfolded by one of several outer membrane protein chaperones such as SurA, Skp or DegP[82]. Struyve et *al.* noticed that C-terminal residues were responsible for OM targeting and integration and Robert et *al.* later determined the generalised consensus motif to be -Φ-X-Φ-X-Φ-X-Y-Φ-F/W (where X represents any amino acid and Φ is hydrophobic)[144, 174]. It was determined by the investigation of Sklar et *al.* (2007) into the effects of chaperone mutants on OMP folding, that SurA is the predominant chaperone pathway in *E. coli* and that Skp and DegP may function more to rescue OMPs not properly sequestered by SurA or that are in some way aberrantly interacting with other components of the pathway[170, 218]. These chaperones assist in targeting unfolded OMPs to the Bam complex at the OM and ensuring that they do not fold or aggregate prematurely. The Bam complex is responsible for the folding most of the β-barrel proteins at the outer membrane and consists of the integral outer membrane protein BamA and lipoproteins BamB-E [Figure 1.11]. BamA is a β-barrel protein with 16 anti-parallel β-strands and five N-terminal polypeptide-transport associated (POTRA) domains that extend into the periplasm[7]. BamA was the first discovered component of the Bam complex, by Voulhoux et *al.* in 2003, when they observed a highly conserved OMP, Omp85, with homologues in mitochondria and chloroplast[190]. They showed that knockouts of this protein resulted in OMP folding defects. The remaining lipoprotein components were discovered by collaborators of Kahne and Silhavy over the next 4 years by identifying co-purified proteins[107, 169, 200]. Due to its essential role in protein folding, a mechanistic understanding of the complex has been long sought after. It has been shown that OMPs can spontaneously fold into

some synthetic liposomes but are incapable of spontaneously folding into a membrane of *E. coli* lipids, thus it is generally assumed that Bam merely assists an energetically favourable process. It was proposed that unfolded polypeptides are passed up the POTRA domains and integrated into the BamA barrel between strands 1 and 16, creating an expanded compound barrel of BamA and a nascent partially folded protein which can then dissociate from BamA into the membrane[7]. Recent structural evidence presented in 2020 shows the Bam complex folding a nascent BamA, into the membrane and validates prior models of Bam activity, while also improving understanding of how the release of nascent proteins from Bam can be energetically favourable[185]. While Bam is responsible for folding the vast majority of outer membrane proteins, there are some that fold by other mechanisms and there are several proteins chaperoned by the Tam complex, however, these pathways are unrelated to Mla expression or activity[217].

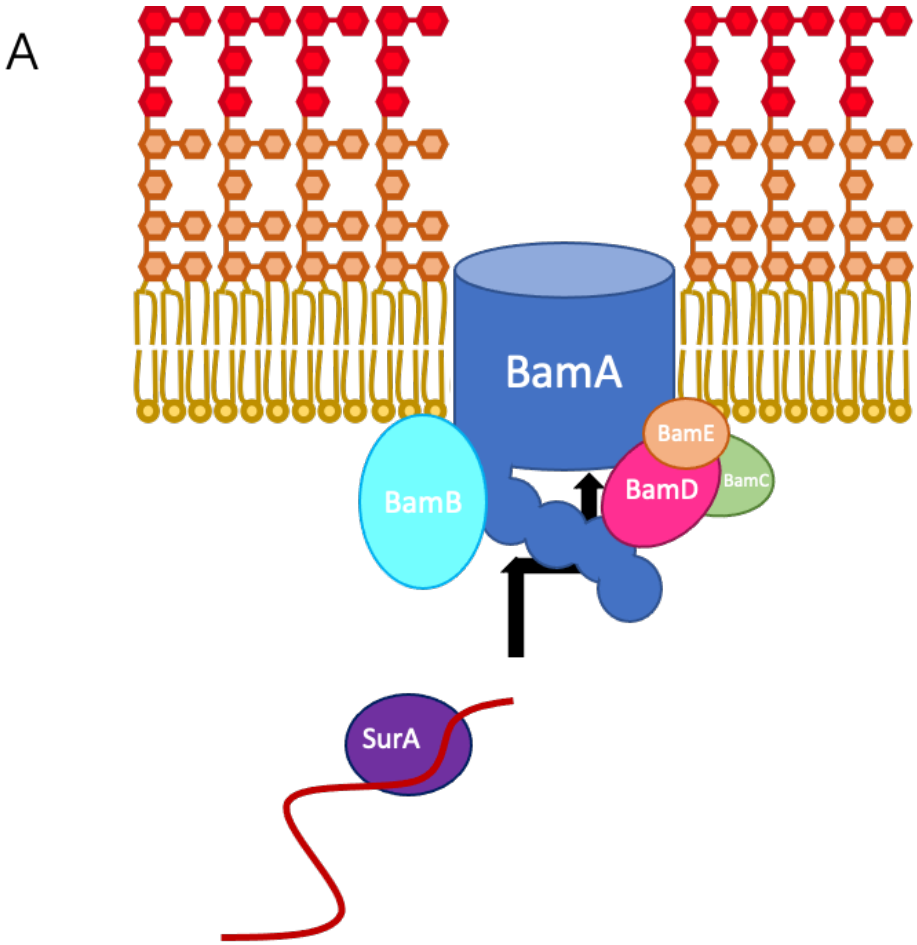

Figure 1.11 - (A) Diagram of periplasmic chaperone SurA delivering an unfolded b-barrel protein to the Bam complex. (B) Top view demonstrating how the nascent protein is incorporated into the BamA barrel and subsequently released due to preferential hydrogen bonding arrangements in the separated barrel structures.

1.3.3.Biogenesis of Phospholipids

Polar lipid synthesis can fundamentally be reduced to the formation of phosphatidic acid via two successive acylation reactions of glycerol-3-phosphate. Phosphatidic acid is the precursor of all bacterial phospholipids, and is as a result the key branching point in the phospholipid synthesis pathway. As such, it would be fastidious to explain the source of the fatty acids used in glycerol-3-phosphate acylation. While many bacteria can prepare recycled or use exogenous fatty acids for glycerol-3-phosphate acylation through the activity of the two component fatty acid kinase system of FakA and FakB in Gram-positive bacteria or acetyl-CoA/acetyl-ACP synthetases in Gram-negative bacteria [Figure 1.12], they are also capable of synthesising their own fatty acids through the bacterial type II fatty acid synthesis pathway (FASII)[130]. Components of the FASII pathway and the associated reactions have long since been identified, through the work of Wakil *et al.*, who determined mechanisms and identified components of the FASII pathway as early as the 1960s as part of a series of papers on the mechanism of fatty acid synthesis[191].

With regards to the incorporation of exogenous fatty acids, the known mechanism are as follows. The acetyl-CoA synthetase of *E. coli*, FadD, is a partially membrane associated ATP dependent synthetase, which concomitantly transports exogenous fatty acids across the inner membrane and catalyses the formation of acyl-CoAs. It was initially identified in 1969 by Overath et *al.* who proposed its role in converting fatty acids to a metabolically active state[129]. The OMP FadL complements FadD activity by transporting fatty acids across the outer membrane. The acetyl-ACP synthetase, Aas, is an integral inner membrane protein, which was discovered in 1976 by Ray and Cronan and similarly catalyses the formation of acyl-ACPs in an ATP dependent manner[141]. Aas has been shown to be a bifunctional protein by Hsu et *al.* and is also capable of 2-acyl-glycerophosphoethanolamine acyltransferase

activity[69]. The two component FakA/FakB system on the other hand was only recently discovered by Parsons et *al.* in 2014, when investigating the incorporation of exogenous fatty acids in *S.aureus*[130]. The Fak system is cytosolic and acts on exogenous fatty acids proposed to spontaneously flip to the inner leaflet of the membrane. FakB removes exogenous fatty acids from the membrane by exchanging them for acyl-phosphates. The fatty acid kinase, FakA, then phosphorylates the fatty acid to an acyl-phosphate which can be fed into phospholipid synthesis.

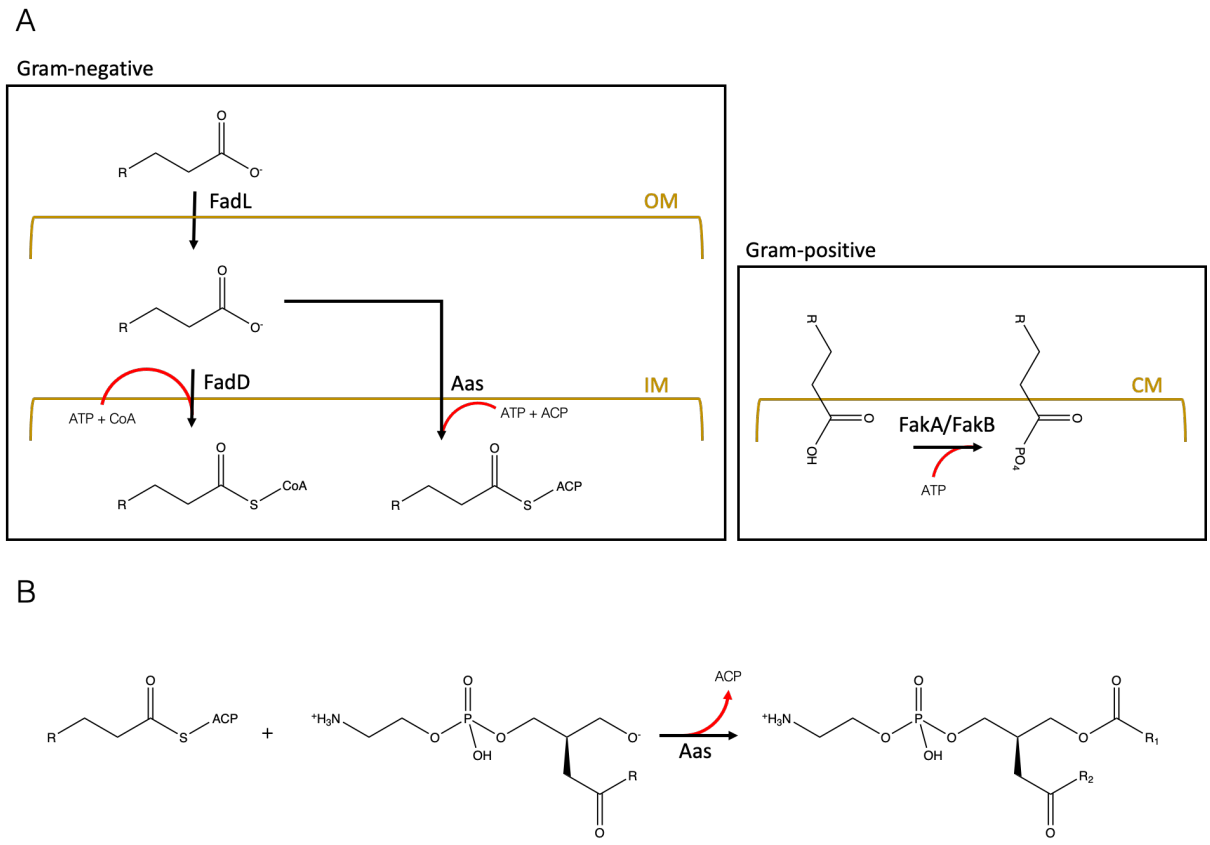

Figure 1.12 - (A) Pathways for exogenous fatty acid incorporation in Gram-negative bacteria (through FadL/FadD or Aas) and Gram-positive bacteria (through FakA/FakB). The acyl-ACP product of Aas can be processed by PlsB, the acyl-CoA product of FadD can be processed by PlsC and the acyl-phosphate product of FakA/FakB can be processed by PlsY. These represent the entry points for exogenous fatty acids into PL synthesis. (B) Aas is also capable of generating phosphatidylethanolamine from lyso-phosphatidylethanolamine and an acyl-ACP.

FASII on the other hand relies of the synthesis of fatty acids from the core bacterial metabolite acetyl-CoA. The FASII pathway has two distinct stages, initiation and

elongation [Figure 1.13]. Initiation begins with the precursor molecule acetyl-CoA generated in other metabolic pathways. Through the activity of the multicomponent acetyl-CoA carboxylase complex, AccABCD, acetyl-CoA is converted into malonyl-CoA in two half reactions. The first involves the ATP dependent carboxylation of a biotin molecule using one molecule of acetyl-CoA and the second involves the transfer of the carboxyl group from biotin to another molecule of acetyl-CoA to form malonyl-CoA[36]. From this point, malonyl-CoA is loaded onto an acyl-carrier protein by the acyl carrier protein transacylase, FabD. The acyl carrier protein acts as a platform for lipid elongation and helps direct the elongating acyl chain to the required enzyme[146]. Different phases of elongation are then carried out by the Fab pathway proteins, starting with the ketoacyl-ACP synthase, FabH, which catalyses the condensation reaction of malonyl-ACP with acetyl-CoA to form 3-oxobutanoyl-ACP[65]. In some bacteria FabH is also capable of using propionyl-CoA, a product of amino acid catabolism, instead of acetyl-CoA resulting in the formation of odd chain length fatty acids, although generally substrate specificity of acetyl-CoA is higher[31]. The other ketoacyl-ACP synthases, FabB and FabF, both elongate acyl-ACPs of length C6 – C16 through a condensation reaction of the acyl-ACP and malonyl-ACP. As FabB and FabF require an acyl-ACP substrate and, as is also the case with FabH, produce a ketoacyl-ACP product, reduction of the ketoacyl-ACP condensation product to an acyl-ACP for further elongation by FabB and FabF constitute the majority of reactions in the elongation phase. The first stage of ketoacyl reduction is catalysed by FabG and involves the NADPH dependent reduction of a 3-ketoacyl-ACP to a 3-hydroxyacyl-ACP. Then the hydroxyacyl-ACP is dehydrated to a trans-2-enoyl-ACP by either FabA or FabZ. An enoyl-ACP reductase, FabI in *E. coli*, then performs the NADH dependent reduction of trans-2-enoyl-ACPs to produce an acyl-ACP for further elongation or for lipid synthesis once sufficiently long fatty acids have been produced. Unsaturated fatty acids are produced by a hydroxyacyl-ACP dehydration step by FabA, which is capable of producing both trans-2-enoyl-ACPs and cis-3-enoyl-ACPs

through a subsequent isomerization of the trans-2-enoyl-ACP product. The resulting cis-3-enoyl-ACP is then passed back to FabB/F for elongation[23, 195, 203].

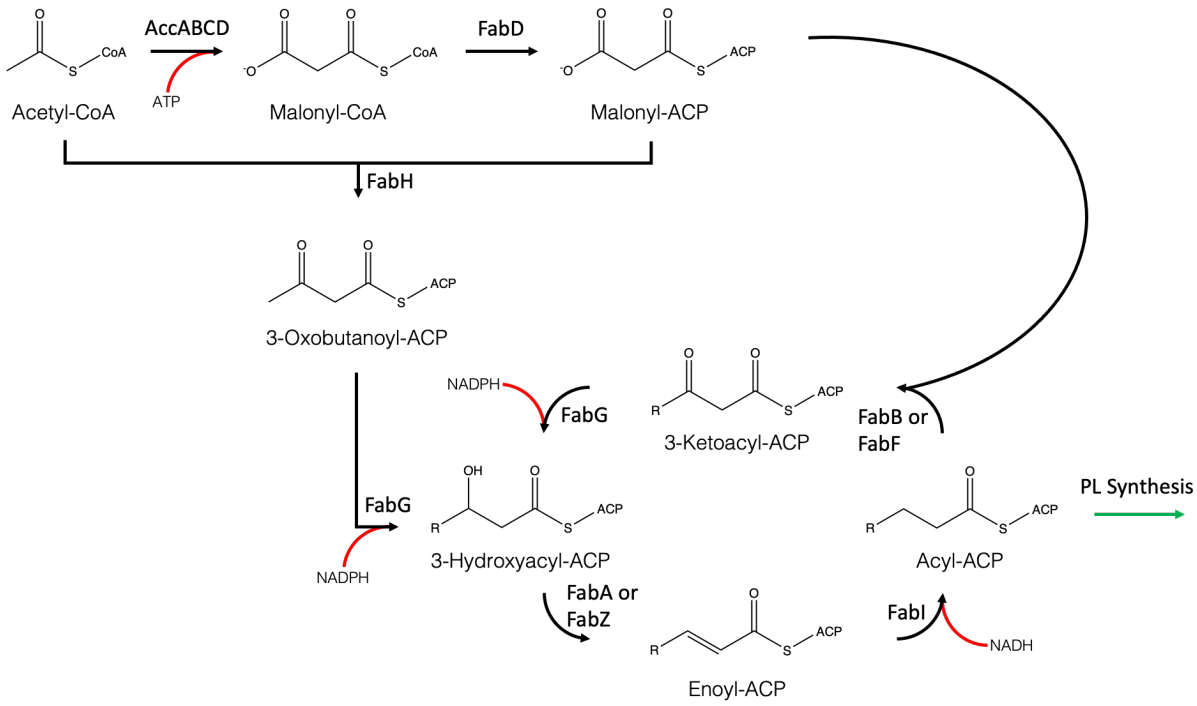

Figure 1.13 - Initiation and elongation of fatty acids by the FASII pathway. The process starts with the synthesis of malonyl-ACP from two molecules of acetyl-CoA and then proceeds with the successive addition of malonyl-ACP to a growing ketoacyl-ACP chain.

Once acyl-ACPs of sufficient length have been produced there are several ways they can enter the phospholipid biosynthesis pathway [Figure 1.14]. They can be directly used for the acylation of glycerol-3-phosphate via the activity of the peripheral membrane protein, PlsB, or they can be released from their acyl-carrier proteins via the activity of PlsX resulting in an acyl-phosphate which the transmembrane protein, PlsY, can subsequently use to acylate glycerol-3-phosphate. This is also the stage at which the acyl-phosphate products of the fatty acid kinase system are incorporated into the phospholipid synthesis pathway. The product of glycerol-3-phosphate acylation is a lyso-phosphatidic acid, which is released into the inner membrane and then further acylated to phosphatidic acid through the activity of the peripheral

membrane protein, PlsC, and an acyl-ACP/CoA substrate[202]. Phosphatidic acid can then be modified into an array of other phospholipids through the common precursor CDP-diacylglycerol which is formed by the integral membrane protein CDP-DAG synthase, CdsA, from CTP and phosphatidic acid at the cytosolic face of the inner membrane[131]. The major *E. coli* phospholipid PE is generated through formation of phosphatidylserine by PssA, which substitutes the CMP component of CDP-DAG with L-serine[149]. Phosphatidylserine is then decarboxylated by Psd into PE[115]. The other major *E. coli* phospholipids PG and cardiolipin are generated through the formation of phosphatidylglycerol phosphate by the substitution of the CMP component of CDP-DAG with glycerol-3-phosphate by PgsA. Phosphatidylglycerol phosphate is then dephosphorylated by one of several multipurpose phosphatases to produce PG. Cardiolipin is formed from PG and another molecule of either PG or PE in a condensation reaction that releases the head-group of the secondary lipid catalysed by a cardiolipin synthase[45]. Large components of the work identifying these Pls proteins was conducted by Charles Rock and his collaborators.

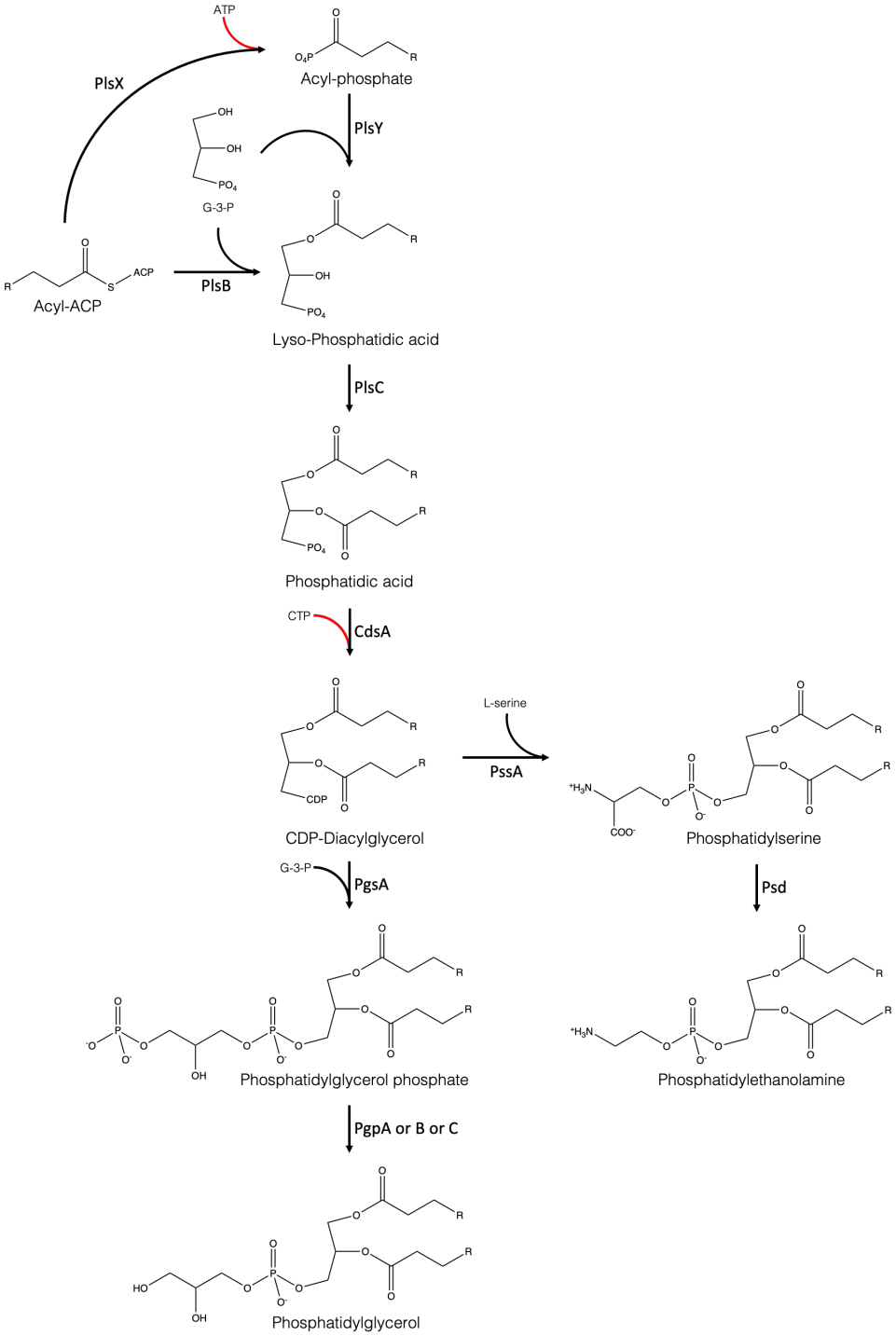

Figure 1.14 - The bacterial synthesis pathway of the common phospholipid precursor CDP-diacylglycerol showing the down-stream synthesis of phosphatidylglycerol and phosphatidylethanolamine.

All proteins that modify phosphatidic acid are either peripheral or integral membrane proteins and all modifications occur at the cytosolic side of the inner membrane. This system of synthesis describes how newly formed phospholipids are incorporated at the inner membrane, but it does not explain how phospholipids are trafficked to the outer membrane. While the exact molecular mechanisms of this have still not been elucidated, it has long been known that lipid export is PMF dependent. Donohue-Rolfe and Schaechter showed in 1980 that depleting the PMF with chemicals that conduct protons across the membrane, such as sodium azide or 2,4-dinitrophenol prevent the export of newly synthesised lipids to the OM[43]. While protein pathways that interact and transport lipids have since been identified, a clear picture of lipid export has yet to be established. Moreover, lipid transport has an additional level of complexity as it has been shown that exogenous lipids can be integrated into the outer membrane and then further metabolised by the synthesis pathway components at the inner membrane. An example of which is with phosphatidylserine liposomes, which can be fused with the outer membrane, and then transported back to the inner membrane for modification into PE by Psd[92]. This discovery came as a further investigation into the seminal work of Jones and Osborn in 1977, who showed that exogenous phospholipids distributed themselves evenly between the inner and outer membrane by determining the distribution of radiolabelled lipids between isolated inner and outer membrane fractions[76]. They also showed in the same investigation that Lipid A derivatives did not appear in the inner membrane fraction. This suggests an additional system for retrograde lipid transport or a system capable of anterograde and retrograde lipid transport.

Currently, it has not been definitively established what mechanisms are responsible for the trafficking of phospholipids from the location of their synthesis at the inner membrane to the outer membrane. Considering known trafficking pathways for lipopolysaccharides and proteins, it stands to reason that a pathway for phospholipid

transport would follow the paradigms observed in those pathways as they are similar molecules. As such, a lipid trafficking protein complex would likely either span the periplasmic gap and form a transport channel as with the LPT complex or involve a periplasmic chaperone, like the Lol pathway. As the energy for outer membrane integration cannot be stored in the lipid, like it can be with an unfolded protein which is not in its lowest energy conformation, a lipid trafficking protein complex would have to generate the energy required for transport at the inner membrane.

It has also been suggested that phospholipids may be transported by cellular lipid structures such as periplasmic vesicles or transmembrane lipid bridges, however, convincing evidence for the formation of these structures has yet to be provided[164]. Cryo-electron microscopy of chemically fixed plasmolysed *E. coli* seemed to reveal points of membrane adhesion, termed as Bayer bridges, however, more contemporary investigations seemed to suggest that these adhesion sites were simply artifacts of the chemical fixation process[10, 81]. The concept of periplasmic membrane vesicles is also believed to be rather unlikely due to the size constraints of the periplasm[164]. Although, bacteria have been observed to release so called 'outer-inner' membrane vesicles containing cytoplasmic material and inner membrane lipids suggesting the potential ability to combine their inner and outer membranes under certain conditions[134]. However, this may only occur in situations where vesicles are released from the bacteria and likely does not suggest an alternative role of this phenomenon in outer membrane biogenesis.

1.4.Introducing The Mla Pathway

1.4.1.Discovery and Characterisation

Interest in the Mla pathway stems from investigations into the role of mammalian cell entry (MCE) proteins conducted by Arruda et *al*., which showed that expression of certain virulence genes from *Mycobacterium tuberculosis* could enable an otherwise non-invasive strain of *E. coli* to invade HeLa cells and survive inside human macrophages for more than 24 hours[5]. Thus the proposed role for the protein coded by this DNA fragment, mammalian cell entry protein 1A, was to facilitate host cell invasion. This activity was further shown to be conferred by a 58 amino acid segment of the protein. Further genes were identified to be in the same operon as *mce1A* and three other similar operons with MceA1 homologues were identified in *M.tuberculosis*[114]. Casali and Riley later followed up with a report identifying *mce* operons across various species of bacteria[28]. They also identified homology with proteins expressed in plant chloroplasts associated with the uptake of phosphatidic acid, which led them to tentatively propose a role in exogenous lipid import or membrane biogenesis. The ubiquity of MCE domain proteins as well as their proposed role as lipid binding proteins drew much research attention to MCE proteins in the field of antimicrobial research. It was not until Malinverni and Silhavy in 2009, that a more concrete description of the role of an *mce* domain protein was proposed[106]. Malinverni and Silhavy investigated the six component Mla protein system, presenting the operon structure of the system in *E. coli*. They determined that the MlaF, -E, -D and -B components likely constituted an IM ABC transport component and based on signal peptides determined that MlaC was likely secreted. An earlier 1994 paper had already determined that MlaA, then referred to as VacJ, was an outer membrane lipoprotein[178]. There has since been a large compendium of structural investigation into all components of the Mla pathway and the roles of pathway

components are widely accepted to be as originally predicted. MlaA is an outer membrane lipoprotein, MlaC is a periplasmic substrate binding protein and MlaFEDB form an inner membrane ATPase complex in a similar paradigm to the Lol pathway[50].

Malinverni and Silhavy also presented a phenotype associated with Mla mutants, suggesting an impeded OM barrier function and increased sensitivity to amphipathic molecules[106]. They showed that growth of Mla mutants was severely impeded on media containing SDS/EDTA compared to WT strains and subsequently presented convincing evidence for the role of the Mla pathway in *E. coli*.

The function of Mla as initially proposed by Malinverni and Silhavy was in the maintenance of outer membrane asymmetry by removal of polar lipids from the predominantly lipopolysaccharide based outer leaflet of the OM[106] [Figure 1.15]. It is established that PL may compensate for LPS loss in the outer leaflet of the OM in the event of LPS shedding due to extracellular stressors or OM assembly defects, however, accumulation of outer leaflet PL increases the susceptibility of the OM to further destabilisation by surfactants like detergents and bile salts and thus there exists a need for retrograde PL transport[106, 121].

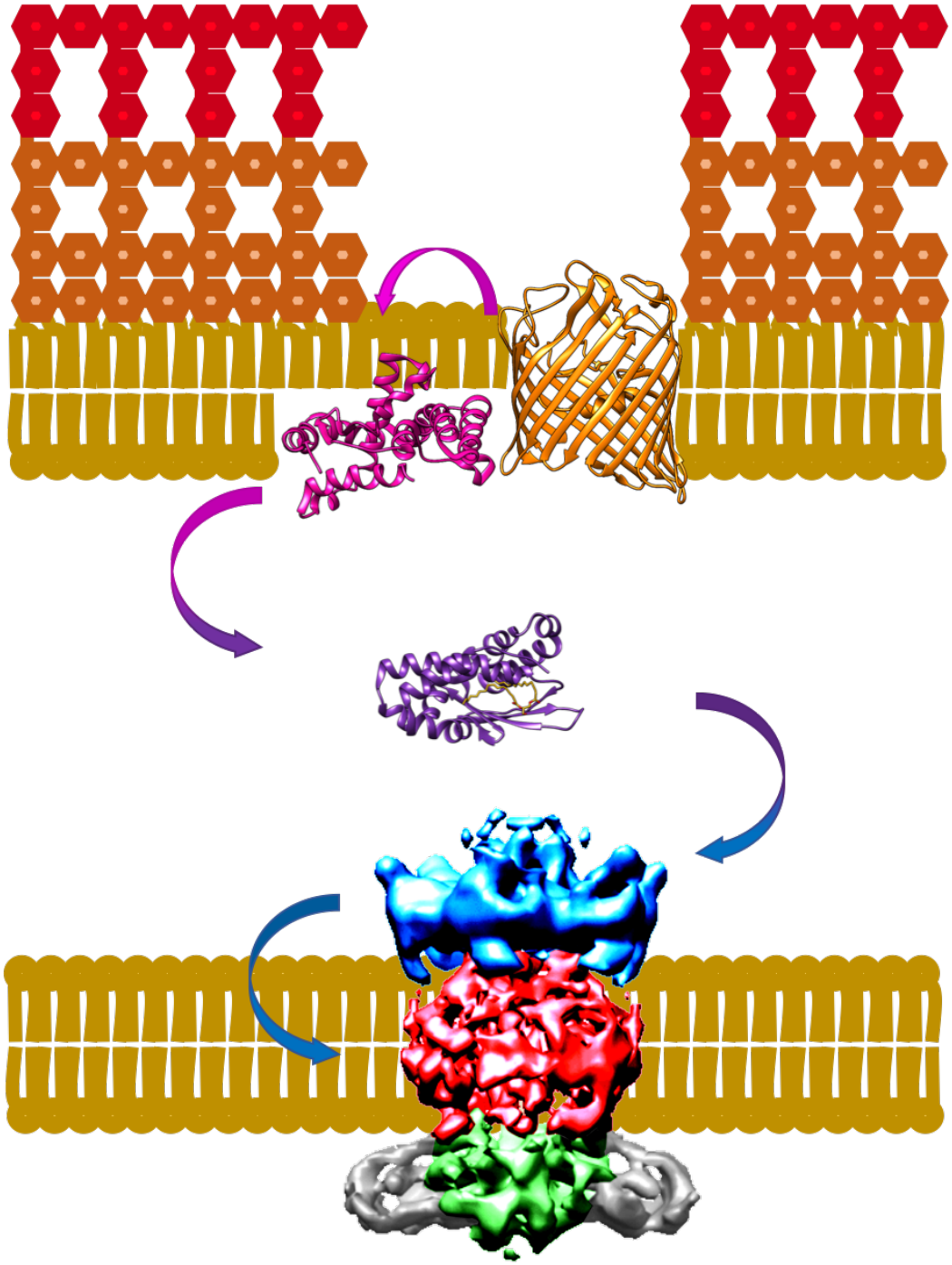

Figure 1.15 - The Mla pathway showing the directionality of lipid transport as originally proposed by Malinverni and Silhavy. MlaA (pink) forms a complex with a β-barrel OMP (orange) in the OM. The periplasmic chaperone MlaC (purple) transport lipids from MlaA to the MlaFEDB (green, red, blue and grey respectively) ATPase at the IM.

Their conclusion was drawn from a series of experiments involving deletion mutations of mla genes, specifically mlaC and mlaD and the resultant effect on SDS-EDTA sensitivity as well as the synergistic effects of Δmla mutations on cells with and without ΔpldA mutations reported through the activity of the protein PagP[106]. PldA is an OM protein that forms a catalytically active dimer in the OM in the presence of PL and serves to remove the fatty acid lipid tails from the phosphodiester head group effectively degrading PL localised to the outer leaflet[38]. Similarly, PagP is also an OM protein that removes the fatty acid tails from lipids, however, it acts only on palmitoyl moieties in the sn-1 position and transfers them to lipid A forming a hepta-acylated LPS. It was found that SDS-EDTA sensitivity induced by Δmla mutations could be rescued by increased PldA expression lending to the conclusion that Δmla mutations result in sensitivity as a result of OM PL accumulation[106]. Furthermore single Δmla mutations and double ΔmlaΔpldA mutations resulted in increased PagP activity owing to increased substrate availability reported through increased presence of hepta-acyl LPS further linking loss of Mla pathway activity to OM PL accumulation and suggesting a retrograde direction for lipid transportation[106].

Conversely, recent papers provide evidence which suggests that deactivation of the Mla pathway homologue in *Acinetobacter baumannii* results in decreased abundance of OM PL and accumulation of newly synthesised PL in the inner membrane[77]. This was determined via direct analysis of membrane lipids using TLC and 2-C^{13} acetate 'pulse' grows to track newly synthesised lipids combined with a technique of isopycnic centrifugation to separate bacterial IM and OM[77, 127]. This new evidence determined through more direct quantification of PL implies a function in populating the OM with newly synthesised PL from the IM, which supports the idea that the Mla pathway in *A. baumannii* is involved in net anterograde lipid transfer. Further investigation into the interactions between purified Mla pathway proteins in *E. coli*

supports the idea of anterograde directionality[71]. It was shown that the direction of lipid transport between a soluble mutant of the IM MCE protein MlaD and the periplasmic chaperone MlaC occurred spontaneously in the direction of MlaC[71]. This directionality was maintained even in the presence of the full IM complex, MlaFEDB, and a surface tethered IM complex was observed to enzymatically shuttle surface deposited PL to MlaC[71]. While more recent papers have resolved most issues of contention regarding the directionality of lipid transport by the Mla system associated with these observations, this investigation began at a stage where the directionality of lipid transport was contentious, A more detailed introduction to the controversy of lipid transport directionality will be presented following a more detailed introduction to the Mla components.

1.4.2.The Mla Operon

The mla operon is differentially organised across different species of Gram-negative bacteria. In some species, notably *Neisseria gonorrhoeae* and other members of the Neisseria genus, all components of the Mla system are expressed on one polycistronic operon at a single gene locus[6]. Comparatively, the Mla system in *E. coli* is spread across two gene loci, with the majority of components, coding for the inner membrane ATPase complex and the periplasmic chaperone, being expressed together on a single polycistronic operon[59]. The OM lipoprotein MlaA is encoded at a substantially different genomic loci, separated by approximately 1 million base pairs. Speculations for the reason of this genomic divergence in the mla operon organisation are unknown but owing to its original identification as a virulence associated chromosome locus in *Shigella flexneri* and its proposed role in outer membrane vesicle blebbing, differential regulation of MlaA may be necessary[6].

1.4.3.The Mla Proteins

1.4.3.1.MlaA

The OM lipoprotein MlaA is the component of the Mla pathway that exchanges lipids with MlaC at the OM. The structure of MlaA from *Klebsiella pneumoniae* (PDB: 5NUP) and *Serratia marcescens* (PDB: 5NUQ) were determined by Abellón-Ruiz et *al.* who managed to crystallise the full length proteins[1] [Figure 1.16]. MlaA was shown to form a complex with trimers of the OM porin OmpF with either a 3:3 or 1:3 ratio, binding at the interface between two porins and causing minimal conformational changes to the porins[1]. While both structures presented by Abellón-Ruiz et *al.* were in complex with OmpF, prior research by Chong et *al.* demonstrated that it forms complexes with both OmpC and OmpF through affinity co-purification experiments[204]. They showed that the affinity of MlaA for these OMPs was sufficient to localise it to the OM, even if it lacks a lipid anchor. Using a construct of MlaA targeted to the periplasm through the SEC translocon rather than to the OM via the Lol pathway by replacing its native signal sequence with a pelB leader sequence they demonstrated that, MlaA still localised to the OM only in cells expressing OmpC. They also showed that both ΔmlaA and ΔompC mutants result in cells with ectopic OM lipids using the same PagP assay as Malinverni and Silhavy.

Baarda et *al.* has since identified MlaA from various species of bacteria, including *N. gonorrhoeae* that lack the conserved N-terminal cysteine residue that is normally functionalised with the lipoprotein lipid anchor[6]. It is reasonably clear then that the presence of OmpC or OmpF is somewhat involved in the folding of MlaA into its functional conformation, especially because MlaA is atypically membrane embedded for a lipoprotein. Structurally, MlaA is periplasmic facing, helical protein spanning the

inner leaflet of the OM. It folds into 6 α-helices, 5 of which are parallel to the membrane and 1 of which, the C-terminal helix, is perpendicular. MlaA forms a ring, with a central channel, which molecular dynamics simulations suggests fills with water molecules[1]. The channel spans the inner leaflet of the OM and the integrity of this channel has been shown to be required for proper activity. Further molecular dynamics simulations suggest that in a PE bilayer MlaA is capable of tilting the head groups of outer leaflet PE towards it[1]. It is suggested that the thickness of MlaA prevents inner leaflet PE from accessing the channel[1]. The channel is predicted to enlarge and allow translocation of PL across it. Interestingly, the activity of the pore is resilient to single mutations of involved residues suggesting the overall conditions of the pore are more important for function than any single residue, which is supported by a lack of conserved pore residues determined by ConSurf[1]. However, a notable exception to this are the D32 and R177 residues which are not robust to mutation, suggesting the salt bridge formed between them is likely important for function[1]. Modification of residues on the periplasmic face have also been shown to reduce activity possibly because they are involved in interactions with the periplasmic chaperone MlaC[1]. Aside from the proposed role of Mla in maintaining OM asymmetry, Mla has also been identified as having roles in vesicle blebbing and biofilm formation in investigations into the effects of MlaA mutants in invasive bacterial species. MlaA was originally identified as a factor that modulated virulence in some bacteria[6, 201]. Baarda et *al.* showed that the virulence of some bacterial strains are differentially affected by *mlaA* mutants[6]. Certain bacterial species lose virulence in response to MlaA deficiency while some exhibit improved virulence. Baarda et *al.* suggested that differential regulation of MlaA may play a role in host colonisation for some species, wherein MlaA downregulation contributes to the production of membrane vesicles conducive to colonisation.

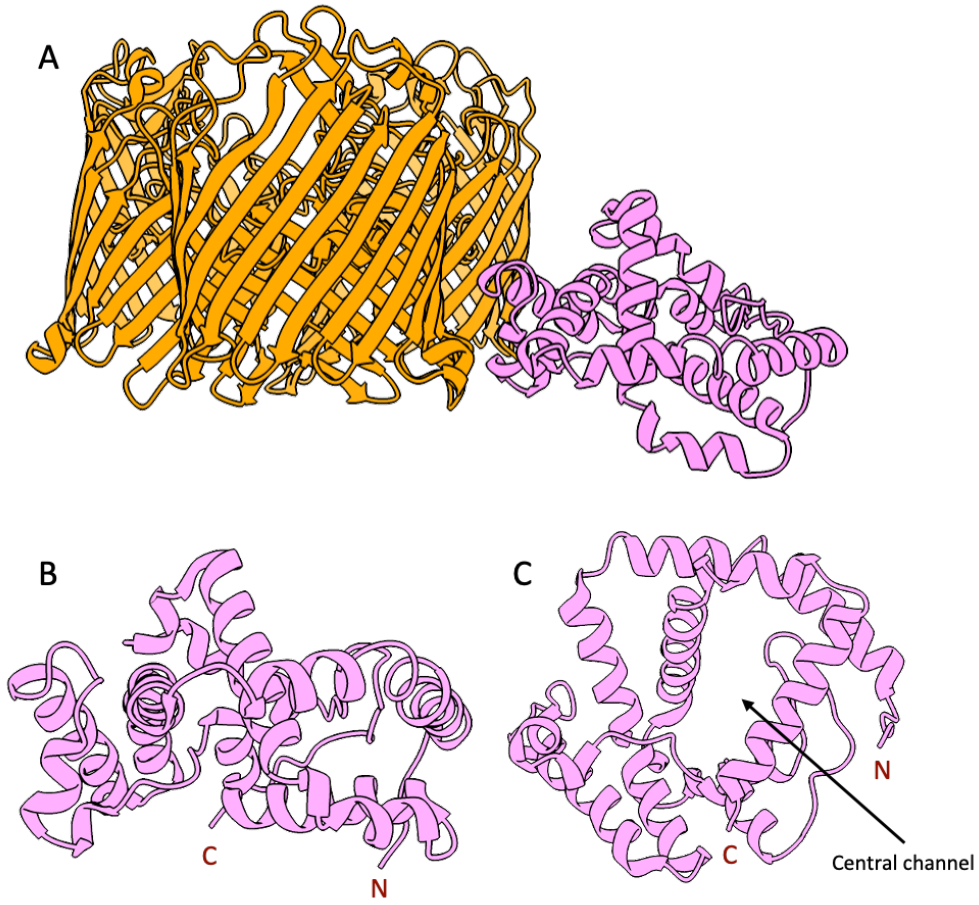

Figure 1.16 – (A) Ribbon diagram of the 5NUQ structure of MlaA (pink) in complex with an OmpF (orange) trimer. The relative position of MlaA suggests the protein is substantially membrane integrated. (B) Side and (C) Bottom views of MlaA showing the N and C terminals as well as the central channel.

1.4.3.2.MlaC

The periplasmic lipid chaperone MlaC [Figure 1.17] (PDB: 5UWA, 2QGU, 6GKI) is natively found to be in complex with PLs based on TLC of chloroform:methanol extracts of purified MlaC protein[50, 71, 90]. Structures presented by Ekiert and Kuzin contained density in the lipid binding pocket which they modelled to be either diacylglycerol (PDB: 5UWA) or PE (PDB: 2QGU)[71, 90]. The PL substrate sits in the binding pocket with the acyl chains deeply embedded into the core of the protein. Molecular dynamics simulation show lipid tails to be mobile in the binding pocket[70].

The lipid binding pocket is composed mainly of hydrophobic residues, though, among homologues of MlaC there are variations in the specific residues. It has been observed that the sequence identity between homologs can be low, with a 29% sequence identity between the protein in *Ralstonia solanacearum* and *A. baumanii.* The conserved and similar residues are almost all non-polar residues associated with the binding pocket. In the absence of PL in the binding pocket it has been shown that the β4 sheet may reorient to cover the pocket opening[71]. Reopening of the pocket may be the step that prevents rapid unregulated uptake of membrane PL by MlaC, as it has been shown by NMR that MlaC does not uptake lipids spontaneously upon exposure. As the structure of MlaC is significant to the research conducted as part of this investigation it will be covered in full in a dedicated section (Section 1.7.1).

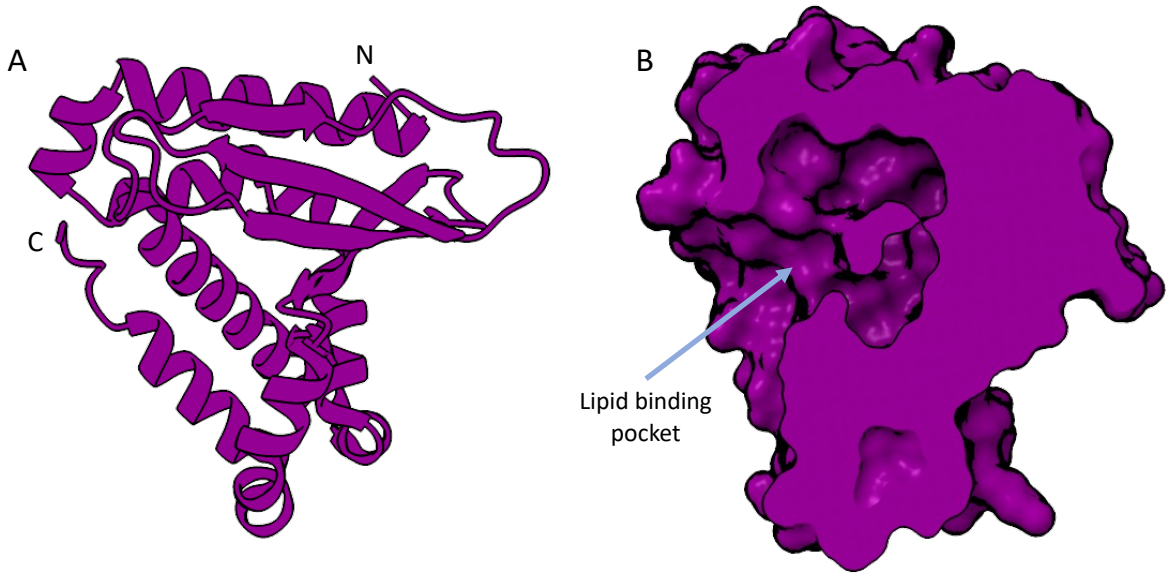

Figure 1.17 – MlaC is a globular protein with a frontal lipid binding pocket. (A) Ribbon diagram of the 5UWA structure of MlaC. (B) Molecular surface diagram cut to show the lipid binding pocket.

1.4.3.3.MlaFEDB

The IM transmembrane complex MlaFEDB is composed of the membrane tethered homo-hexameric MCE protein complex MlaD on the periplasmic face, the transmembrane dimeric permease protein MlaE, the dimeric cytoplasmic facing ATP-binding protein MlaF and two copies of the STAS-domain protein MlaB interacting with MlaF[50]. The MlaD hexamer forms a ring exhibiting 6-fold symmetry with a central helical bundle. Soluble constructs of the MlaD protein have been shown to stably bind lipids and are purified in the lipid bound state, suggesting lipid uptake is independent of MlaFEB activity and is not tied to the ATPase activity of MlaF[71]. MlaFE forms the core of a conventional ATP-binding cassette (ABC) with the primarily helical transmembrane MlaE and the ATPase MlaF forming a complex with two-fold symmetry[50]. MlaB is a small ~10kDa protein that interacts with MlaF and has been shown to be required for the assembly and activity of the MlaFEDB complex[184]. Structural evidence for this arrangement of protein subunits was provided by Ekiert et *al.* in 2018. Higher resolution structures of MlaFEDB have since become available, provided by multiple research groups[30, 34, 182] [Figure 1.18].

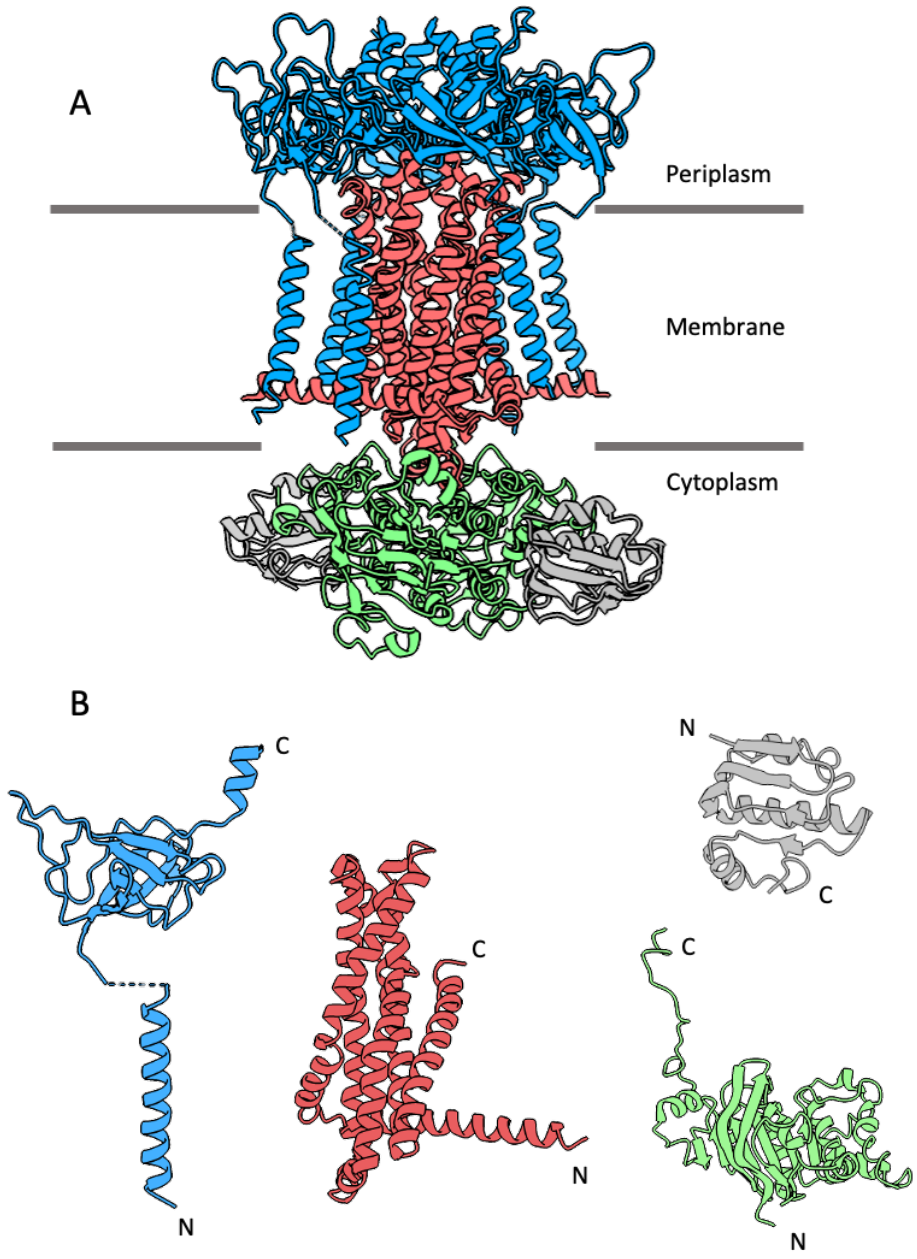

Figure 1.18 – (A) Ribbon diagram of the 7CGE MlaFEDB complex showing the approximate position of the complex in the membrane. The MlaD (blue) hexamer is exposed to the periplasmic side of the IM, and tethered to the IM by transmembrane helices, which also interact with MlaE (red). MlaF (green) and MlaE assemble into a classic ABC transporter. MlaB (grey) binds to MlaF on the cytoplasmic side of the IM. (B) Each monomeric unit of the complex separately showing N and C terminal.

The ATPase role of MlaFE was determined in the original phylogenomic analysis of *mce* operons carried out by Casali and Riley based on conserved motifs associated

with ABC proteins[28]. ABC proteins are a widely conserved family of proteins found across all kingdoms of life and generally responsible for energy generation in multicomponent transmembrane transporters. In bacteria they mainly play roles in uptake and export systems and account for a non-trivial percentage of the bacterial genome[100]. To carry out ATPase activity, the ABC protein assembles into a dimer with two-fold symmetry as can be seen with MlaFB in [Figure 1.18]. The key domains of an ABC protein can be broken down into the nucleotide binding domain (NBD) and the transmembrane domain (TMD). The NBD and TMD can be two domains of a single protein or in some cases they can be two separate proteins that form into a complex. The NBD generally contains several key motifs for function: The Walker A motif responsible for phosphate binding with consensus sequence -G-XXXX-GK-[TS]- (with X denoting any amino acid), the Walker B motif responsible for magnesium ion coordination and ATP hydrolysis with consensus sequence Φ-Φ-Φ-Φ-DE (with Φ denoting a hydrophobic residue), the switch motif (or H-loop) providing a histidine residue and the ABC signature motif with consensus sequence -LSGGQ- responsible for coordinating the γ-phosphate[37]. The ATP binding pocket of the NBD dimer is organised from these key motifs, with the Walker A, Walker B and H-loop motif of one NBD and the signature motif of the other NBD coordinating and catalysing the ATP hydrolysis. Substrate binding tends to occur in a pocket between the two TMD monomers, and ATPase activity results in conformational changes that translocate the substrate.

A general mechanism has been suggested for ABC exporter proteins, wherein nucleotide-free state corresponds to an inward open substrate binding pocket between the TMDs, allowing for substrate binding. Binding of the substrate causes conformational change that enhances the affinity of the NBD for ATP, resulting in ATP binding and tight dimerization of the NBD. The resultant conformational change leads

to an outward open configuration and release of the substrate. ATP hydrolysis results in a transition back to the inward open state[100]. Likewise Type I ABC importers, which tend to have associated substrate binding proteins, generally adopt an inward open configuration in the nucleotide-free state. Conformational changes induced upon binding of a substrate binding protein cause an increased affinity of the NBD for ATP, resulting in ATP binding and tight dimerization of the NBD. The resultant conformational change leads to an outward open configuration to accept the substrate. ATP hydrolysis upon substrate binding results in transition back to the inward open state and release of both the substrate and the substrate binding protein. This mechanism was demonstrated by Jue Chen with the *E. coli* maltose transporter complex[29]. MlaE on the other hand adopts an outward open relaxed state, more characteristic of Type II ABC importers. Lewinson et *al.* proposed a distinct mechanism ABC proteins with a relaxed outward open conformation based on the dynamics of the the *E. coli* vitamin B_{12} importer complex BtuCD. It was suggested that binding of the substrate binding accessory protein to the outward open configuration results in the opening of a channel through the TMD. The substrate is able to translocate through the channel. ATPase activity serves to cause the dissociation of the substrate binding protein from the ABC protein, recovering the outward open configuration. Currently it is ambiguous if MlaFE conforms to any of these mechanisms, however, MlaF has been determined to unambiguously be an NBD protein[95].

For the most part MlaF conforms to the conserved motifs of an NBD protein and as such was easily classifiable as an ABC protein component even when it was originally identified as YrbF, before its function had been investigated. Residues 41-47 form the Walker A motif of MlaF with sequence G-PSGI-GKT, residues 165-170 from the Walker B motif with sequence LIMFDE, His-203 is the H-loop histidine and residues 145-149 form the signature sequence. However, the signature sequence of MlaF does not

exactly match the consensus sequence for this highly conserved motif with a methionine at position 149 instead of a glutamine[85]. It is proposed that this difference in the signature sequence of MlaF may affect the specificity for ATP, allowing for the use of dATP although the significance of this is unknown. While this D/M substitution is very uncommon among ABC proteins it seems to be conserved in Mla-like and TGD ABC proteins, another lipid transport protein in plant chloroplasts. Further structural details regarding primarily MlaD will be addressed in a dedicated section (Section 1.7.2) as the structure of MlaD is fundamentally important to the content of this report.

1.4.4.Controversy of Lipid Transport Directionality

As has been previously alluded, there has recently been some debate about the directionality of transport and the specific cellular role of the Mla system stemming initially from two papers both published in 2019 by Kamischke et *al.* and Hughes et *al.*, which both provided evidence for lipid transport towards the outer membrane[71, 77]. Kamischke et *al.* observed inner membrane lipid accumulation in Mla mutants and Hughes et *al.* observed that the directionality of lipid transport in solubilized components of the Mla pathway appeared to transfer unilaterally in the direction associated with anterograde lipid transport. This was diametrically opposed to the role of the pathway as originally suggested by Malinverni and Silhavy as well as several papers providing genetic and functional evidence in support of the original retrograde transfer of lipids[106, 138, 145].

Roier et *al.*, while investigating the mechanisms of outer membrane vesicle formation and hypervesiculating strains of *Haemophilus influenzae*, noticed that mutants lacking Mla produced significantly higher levels of OMVs[145]. Roier et *al.* demonstrated that investigation into the Mla pathway from a perspective of OMV production and related virulence phenotypes could provide further insight into the role

of the pathway. Their investigation, as well as a similar investigation into virulence in *N. gonorrhoeae* by Baarda et *al*. suggested that Mla may have auxiliary roles outside of the initially proposed role or removing ectopic lipids[6]. They noted increased serum resistance and a significantly improved capacity to colonise host tissue in strains that lacked Mla expression. Considering this phenotype in environments representative of host invasion and that Mla was downregulated during iron starvation under control of the ferric uptake regulator they proposed that the hypervesiculation phenotype was physiologically relevant in combatting host immune response. During invasion OMVs deplete host response factors such as antibodies and contribute to colonization[135]. Both groups came to the conclusion that Mla regulation was a mechanism of tuning virulence during initial colonization and that downregulation of Mla was a physiologically relevant mechanism of OMV biogenesis, permitting the build-up of outer-leaflet PLs resulting in membrane blebbing and eventual vesiculation. This branch of work presents Mla as a PL import pathway in pathogenic bacteria wherein its down-regulation plays a role in invasion.

Powers and Trent in 2018 took a genetic and evolutionary approach to investigate the role of the Mla system in a lipid A deficient strain of *A. baumannii*[138]. *A. baumannii* is rather unique in that it can survive even in the absence of OM LPS, although, LPS deficiency does cause growth defects. By encouraging growth under selection pressures that prevented the reversion of LPS deficient strains to those capable of producing LPS, they generationally cycled their *A.baumannii* strain until they recovered growth rates to near that of WT. Whole genome sequencing of these strains revealed that to recover WT growth rates, mutations that inactivated Mla and PldA had occurred. Increased resistance of these generationally mutated strains to Gram-positive antibiotics suggested that the integrity of the OM had been somewhat recovered. Powers and Trent determined that mechanisms that maintain OM asymmetry were an inherent detriment to cell growth in cases where asymmetry

cannot be maintained. Genetic evidence as presented here suggests that downregulation of Mla results in recovery of OM integrity due to repression of retrograde lipid transport, further supplementing evidence for the retrograde transport directionality of the Mla pathway.

Sutterlin et *al.* discovered a mutant variant of MlaA that they found to be associated with a novel cell death pathway[177]. This mutant, denoted *mlaA**, was a result of the deletion of the two adjacent residues Asn41 and Phe42. The phenotypes they observed to result from this mutation were reduced OM barrier function when compared to a *Δmla* strain, increased production of LPS, hypervesiculation and cell death in stationary phase associated with depletion of cellular fatty acids. They determined that the *mlaA** phenotypes were not recoverable by knockouts of other Mla components and thus determined that the *mlaA** performed a function independent of the remaining components of the system. They proposed that this gain function was the ability to flip OM lipids from the inner leaflet to the outer leaflet. Yeow et *al*. would latter comment that this gain function was likely due to removal of residues key to preventing inner leaflet PLs from accessing the MlaA pore[204[. Sutterlin et *al*. also observed that attenuating PldA function suppressed cell death by preventing over-production of LPS[177]. While they did not provide a conclusive mechanism for this suppression of cell death, they did suggest that increased PldA reaction products may be a signal for LPS production. Most importantly to the topic of anterograde lipid transport though, was their observation of a shrinking inner membrane during cell death that was unaffected by ATP synthase mutations, that they attributed to a diffusive flow of lipids supplementing those lost from the OM due to vesiculation. They suggested that this diffusive lipid flow strongly supported the theory of lipid transport through zones of membrane adhesion. Later investigations into *mlaA** mutants by Grimm et *al*. revealed a gene, *yhdP*, that when downregulated, reduced the depletion of the inner membrane[62]. YhdP is now thought to primarily play

a role in high flux transport of lipids between membranes from comparison of its function to known eukaryotic proteins that share homology. Due to the size of its periplasmic domain Grimm et *al*. suggested it likely spans the periplasm and proposed that the simplest explanation for its function is a direct role in lipid transport. As *ΔyhdP* strains are viable Grimm et *al*. proposed that YhdP either plays a redundant role in standard PL transport to the OM or only functions as a high flux transporter. They believe that this high flux transport function through YhdP induced by OM dysregulation by the *mlaA** was the underlying cause of the novel mechanism of cell death observed by Sutterlin et *al*., rather than any further involvement of Mla pathway proteins.

Supporting evidence for anterograde transport directionality came mainly in the later paper by Kamischke et *al*.[77]. Kamischke et *al*. presented a ^{13}C-acetate 'pulse' chase assay that suggested *mla* mutants built up newly synthesised lipids at the inner membrane. They observed that transport of newly synthesised C^{13}-tagged lipids to the OM was attenuated and that *mla* mutants had decreased abundance of OM lipids. Their observations led them to suggest that the Mla system may play a dual role in both the maintenance of OM asymmetry and the export of nascent PLs. However, this has since been contradicted by Powers et *al*., who highlighted some flaws in the Kamischke paper and demonstrated that Mla mutants in fact do not present the phenotypes observed by Kamischke et *al*.[137]. They identified that a mutation to *obgE* compounding with further mutations to Mla components was responsible for the phenotypes observed. ObgE is a GTPase involved in a pathway of bacterial stress response and Powers et *al.* suggested the mutant generated by Kamischke et *al.* may cause downstream inhibition of fatty acid biosynthesis, leading to an inability to properly incorporate isotopically labelled ^{13}C-acetate into exportable phospholipids.

Hughes et *al*. on the other hand observed *in vitro* exchange of lipids, preferentially in the direction of MlaC between purified samples of Mla proteins[71]. They noted that net lipid exchange over a 30 min incubation occurred from MlaD to MlaC and that holo-MlaC was incapable of delivering lipids to an apo-MlaD. This directionality was observed in soluble constructs of the two proteins, measured at the endpoints after exchange, as well as between the MlaFEDB complex and MlaC, measured continuously by quartz crystal microbalance with dissipation monitoring (QCM-D).

With regards to the directionality of lipid transport observed by Hughes et *al.*, several papers have recently been published that observe ATP-dependent transport of lipids from the MlaFEDB complex to MlaC. Both Tang et *al*. and Low et *al*. performed *in vitro* lipid exchange assays that suggested anterograde transport is almost entirely inhibited by ATP[102, 182]. Tang et *al*. developed an assay that incorporated solubilised MlaFEDB and solubilised MlaA/OmpF in liposomes with an apo preparation of MlaC. They observed lipid transport between the liposomes using fluorescently labelled PE in a rhodamine/NBD FRET pair. They observed high rate retrograde transport in the presence of 1 mM ATP and 2 mM $MgCl_2$[182]. Low et *al*. on the other hand investigated only transport between MlaFEDB in nano-discs and MlaC using C^{14}-labelled lipids. Their assay showed spontaneous transfer of C^{14}-labelled lipids from MlaC to MlaFEDB, which was enhanced by ATP hydrolysis[102]. In both these assays, the absence of ATP leads to clear evidence of anterograde transport suggesting that cellular ATP concentrations may need to be maintained above a threshold to prevent anterograde lipid transport. Regardless, it is currently generally established that the Mla pathway is primarily a retrograde lipid transport pathway.

1.5.Other MCE Proteins

Two other *mce* domain proteins have also been identified in *E. coli*, however, structural evidence suggests that they function through a substantially different mechanism to the Mla pathway. These two proteins are PqiB and YebT. The PqiABC complex, identified as a paraquat inducible gene in *E. coli* by Koh and Roe in 1995 comprises a periplasm spanning channel, as shown by structural evidence provided by Ekiert et *al.*[50, 84]. PqiB consists of a transmembrane domain, then 3 consecutive *mce* domains followed by a long helical domain. Much like MlaD, PqiB assembles into a hexamer with a central pore. The long helical domain appears to create a helix bundle proposed to span the periplasm. PqiB is proposed to interface with PqiA at the IM and PqiC at the OM. Genetic evidence provided by Nakayama et *al.* shows that Pqi also has a role in maintaining membrane integrity[119]. It is proposed that they are also involved in the transport of lipids between the IM and OM. YebST similarly forms a long transmembrane channel and has been shown to be involved in maintaining OM integrity. YebT consists of a transmembrane domain, then 7 consecutive *mce* domains followed by a short helical region and assemble into a hexameric arrangement[50]. Unlike with MlaD and PqiB, plasmid expression of YebT has shown to be acutely toxic. Nakayama et *al.* proposed this toxicity may in some way be caused by the same mechanism as toxicity associated with *mlaA** observed by Sutterlin et *al.*[119].

1.6.Detailed Structural Analysis of MlaC and MlaD

1.6.1.The Structure of MlaC

MlaC is the soluble periplasmic lipid chaperone of the Mla system, but it does not share any sequence similarity with known chaperones of other ATP transporters. It has been described as being composed of two hybridised structural domains. These consist of what has been called an “extended” NTF2 (Nuclear Transport Factor 2), or NTF2-like domain and an AAA+ (ATPases Associated with diverse cellular Activities) helical bundle domain[205]. Although, the structural elements homologous to these domains are formed from non-contiguous sequences in this protein. Interestingly, AAA+ domains are known to assemble into homo-hexameric rings and while MlaC is definitively monomeric, MlaD assembles into a homo-hexameric ring. The AAA+ helical bundle that shares homology with MlaC generally performs the function of transmitting ATP hydrolysis induced conformational change to adjacent monomers in the homo-hexameric assembly of true AAA+ domain proteins, however, that does not appear to be the function in MlaC.

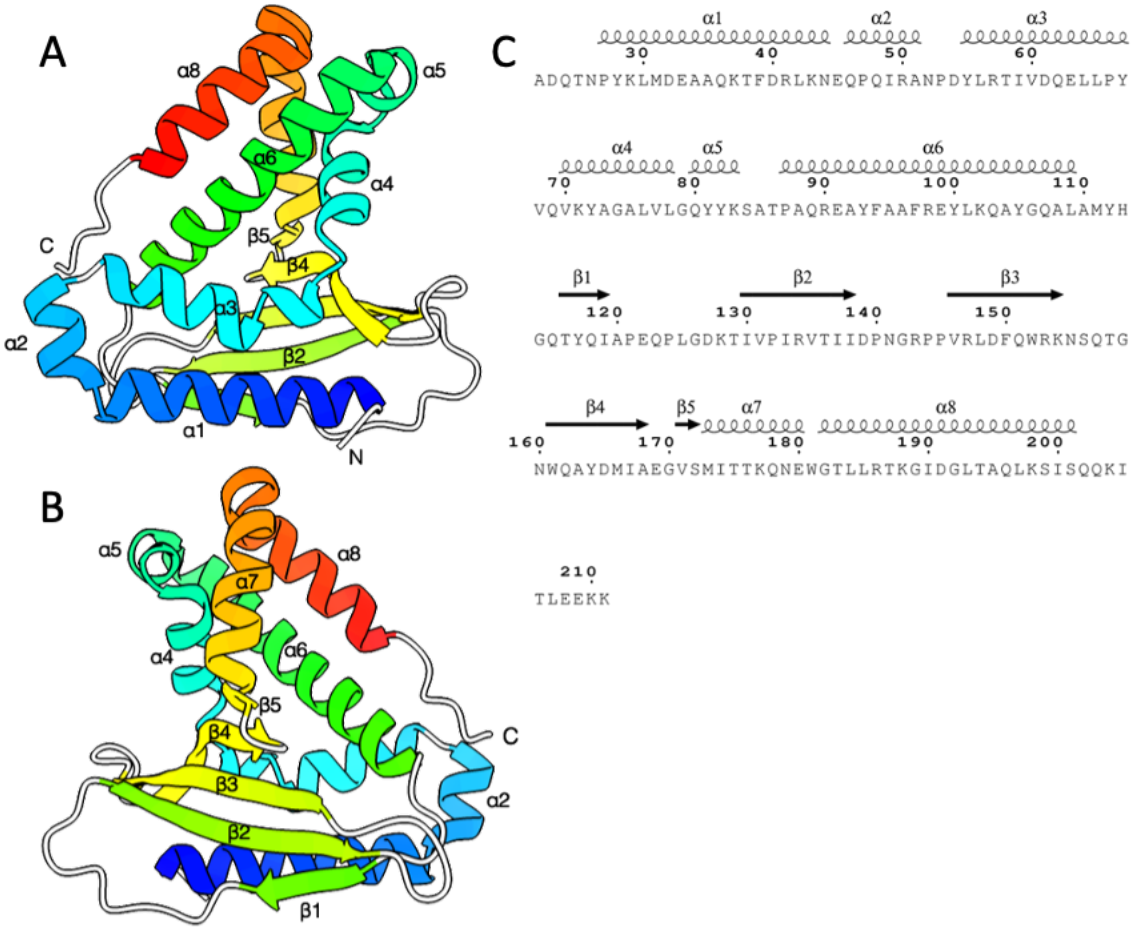

Figure 1.19 - Ribbon diagram of the 5UWA structure of *E. coli* holo-MlaC (A) alongside sequence information (C). Secondary structure elements are aligned to the protein sequence. Front (A) and back (B) views are displayed for clarity. Secondary structure elements are coloured blue to red successively from the N- to C-terminal.

Instead MlaC forms a hydrophobic pocket between α1, α6, β3, and β4, which the acyl tails of glycerophospholipids make direct contact with. This binding pocket appears to be bifurcated at the front by Tyr105, a residue that is highly conserved across MlaC homologues, although the pocket is connected behind the frontal Tyr105 residue. MlaC natively purifies in complex with its glycerophospholipid substrate in what is often referred to as holo-MlaC[70] [Figure 1.19]. Structural evidence presented by Ekiert et *al.* suggests that both chains of the phospholipid substrate bind into only one

of the frontal pockets [Figure 1.20]. The role of the second binding pocket is currently unknown[50]. It was shown that extended incubation of the holo-protein with detergent is capable of delipidating it to generate an Apo-protein[71]. The Apo-protein exhibits some conformational changes when compared to the holo-protein [Figure 1.21]. The most significant of which occurs in the region around β1 – 5 and α7. The entire β-sheet pivots substantially, with a deflection of approximately 15°. The loop between β4 and β5 shifts to block the opening of the lipid binding pocket and α7 becomes disordered to accommodate this change. The overall result is a substantial reduction in the volume of the lipid binding pocket[71] [Figure 1.22]. Molecular dynamics simulations identify a correlation in the conformational changes of the central α6 helix with β2, β3 and α2, possibly suggesting that the central α6 helix plays a role in the transition between holo- and apo- states[70]. As the Tyr105 residue, which is positioned on the central α6 helix, is repositioned in the holo-state by the lipid, this may be one of the components that translates lipid binding to conformational change elsewhere in the protein [Figure 1.23]. MlaC has been shown to accept lipids from soluble MlaD as well as the FEDB complex. The idea that other Mla proteins are necessary for this conformational transition from the apo- state to an open state capable of accepting lipids is further supported by the inability for MlaC to spontaneously bind free lipids[71].

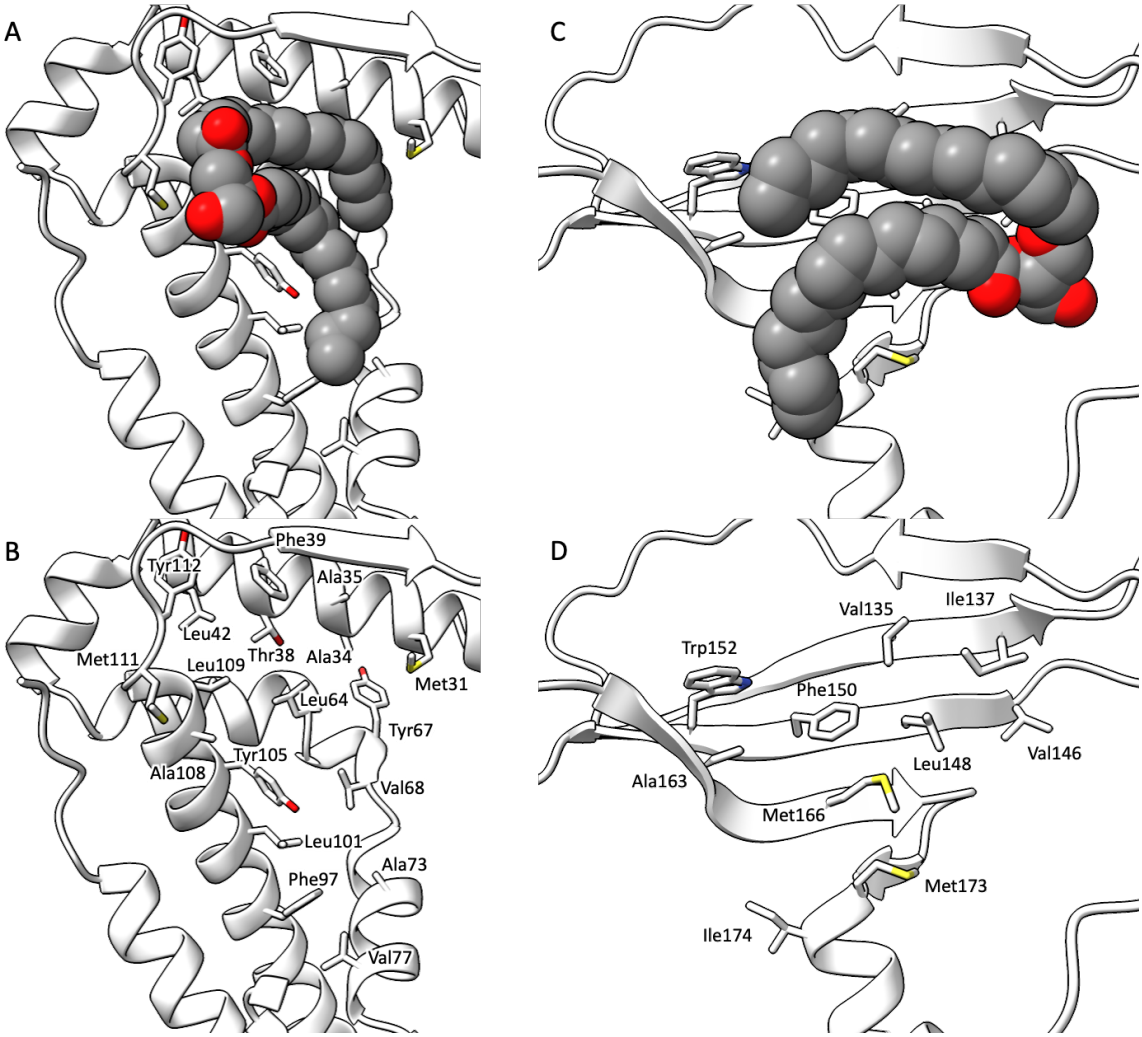

Figure 1.20 - MlaC ribbon diagram showing relevant residues in the binding pocket involved in interacting with the phospholipid substrate. The protein is shown from a front and rear view with the lipid present and absent. (A) Front view with residues 133-179 hidden to allow clear view of lipid. (B) Front view with lipid removed to show residues involved in coordinating the acyl tails. (C) Rear view with residues 24-113 hidden to allow clear view of lipid. (D) Rear view with lipid removed to show residues involved in coordinating the acyl tails.

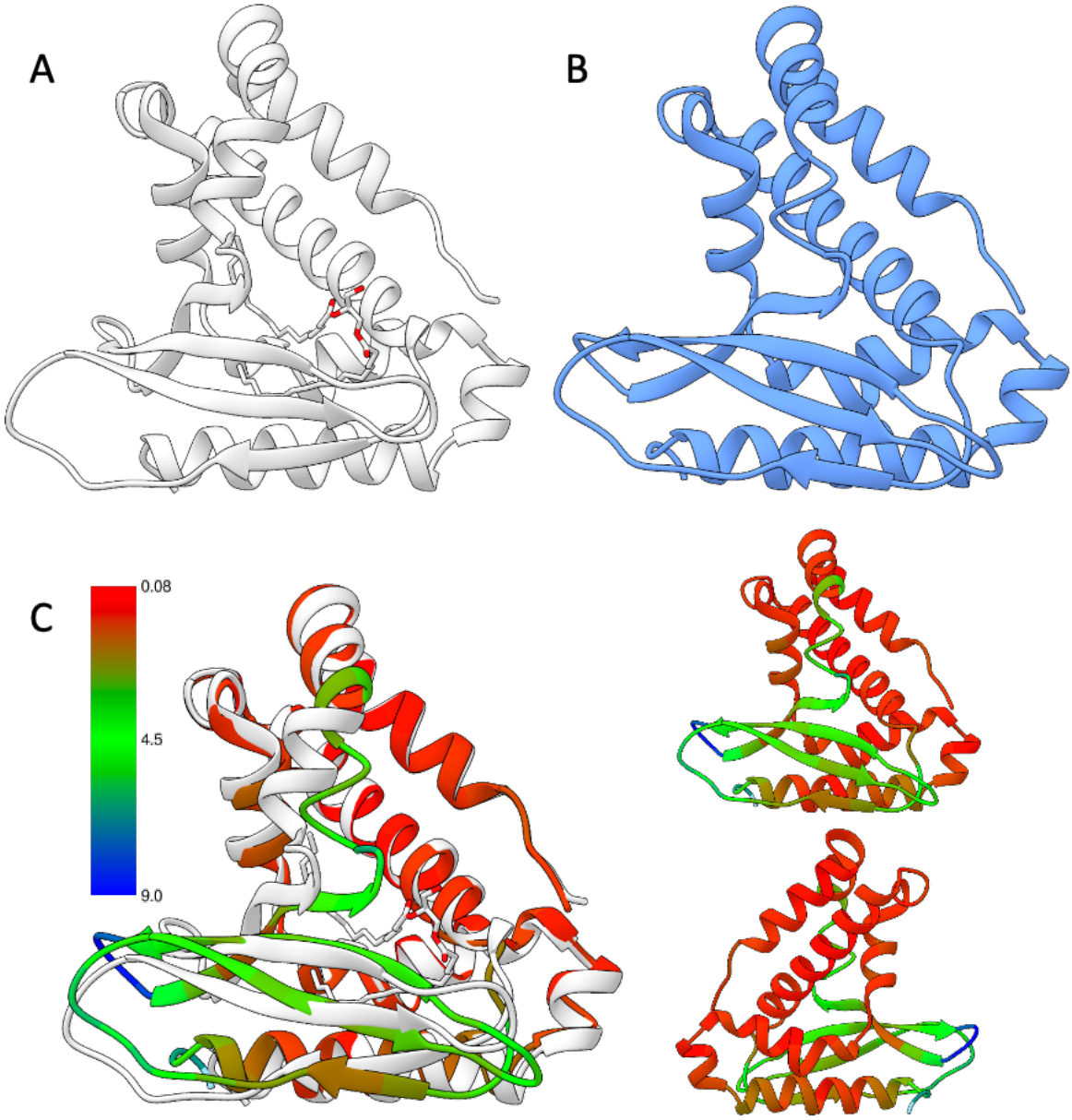

Figure 1.21 - Ribbon diagram of MlaC-holo (A) and MlaC-apo (B). Structures are overlayed (C) and the -apo structure is coloured by C-alpha RMSD from the -holo structure. The colour key is shown on the far left. The greatest deviation of the -apo structure from the -holo structure is in the β3 – β4 loop, which deviates by up to 9 Å, however, the entire β-sheet assembly as well as the beginning of the α7 helix deviates by up to 4.5 Å. Front and back views of the -apo structure have been included in (C) for clarity.

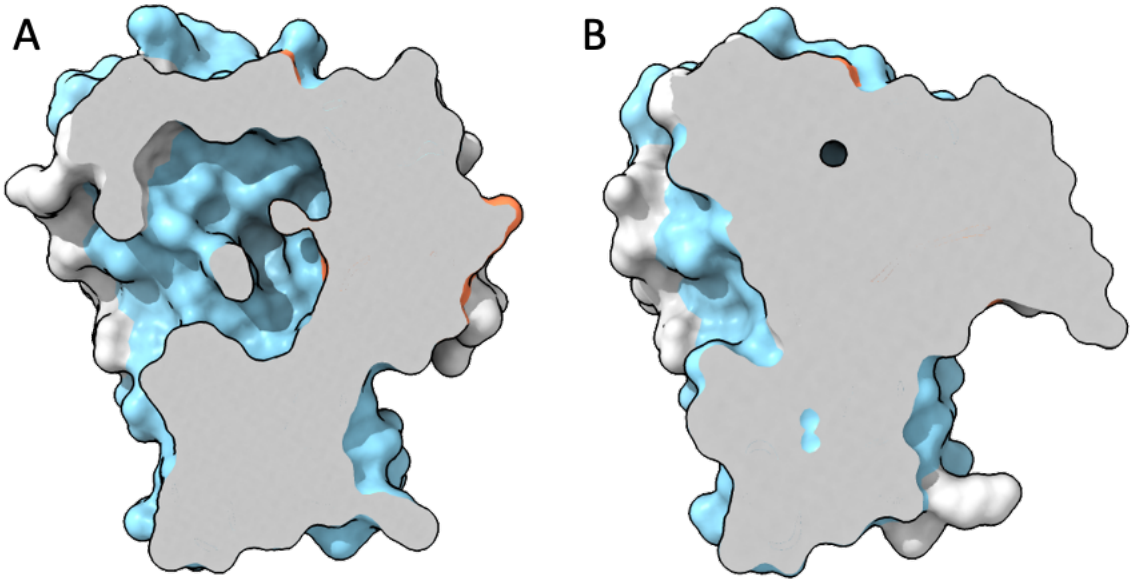

Figure 1.22 - Molecular surface diagram of MlaC-holo (A) and MlaC-apo (B) cut to show the binding pocket. The binding pocket is entirely collapsed in the -apo structure.

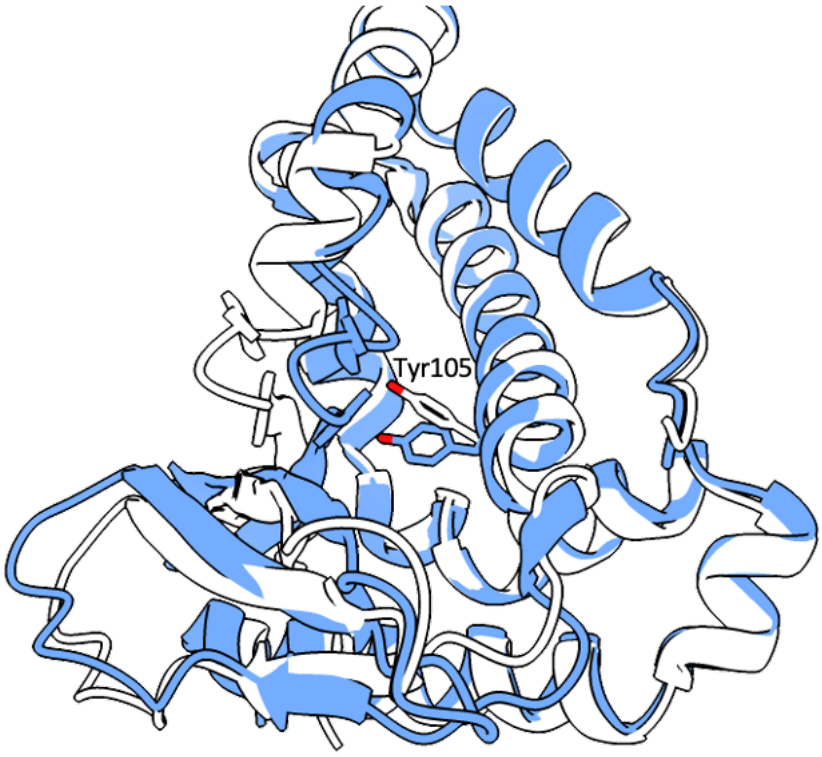

Figure 1.23 - Ribbon diagram showing deflection of the Tyr105 residue in the holo-protein (white) compared to the apo-protein (blue). The only displayed side chain residue is that of Tyr105.

At the time of writing, there are currently 6 structures of MlaC and MlaC homologues deposited in the PDB according to the MlaC PF05494 Pfam classification. The PDB IDs for these structures are 5UWA[50] (*E. coli*), 6GKI[71] (*E. coli*), 4FCZ[91] (*Pseudomonas putida*), 5UWB[50] (*P. putida*), 2QGU[90] (*R. solanacearum*) and 6HSY[205] (*Pseudomonas aeruginosa*). Of these structures, all but the 6GKI structure have been crystalised from a natively purified protein, while the 6GKI protein was made apo prior to crystallization. It is worth noting that the sequence similarity between some of these homologues is as low as 16%, however, the RMSD of structural alignments is cited as being between 1.6 and 3.1 Å for a Cα alignment. While very few residues are exactly conserved among these proteins the relative positions of residues with similar properties is maintained within the core fold. This sequence variability across homologues is likely to some degree permitted by the loss of substrate specificity that allows MlaC and its homologues to bind a large variety of cellular lipids. The sequence variability does, however, manifest as a significant functional difference between some MlaC homologues, notably Ttg2D from *P. aeruginosa*, which is shown to be capable of binding two glycerophospholipids in its active site owing to a larger hydrophobic cavity. This expansion of the binding cavity is attributed to differences in the α8 helix[205]. The purpose of this expanded binding site may be to better accommodate cardiolipin.

1.6.1.1.A Proposed Inhibitor of MlaC

While the Mla pathway is not required for cellular viability it has been suggested in the literature that targeting the Mla system may potentiate the use a wider range of existing antibiotics on Gram-negative bacteria. Bernier et *al*. in 2018 suggested that the increased permeability of the OM in cells exhibiting Mla defects presents an opportunity for combining existing antibiotics ineffective on Gram-negative bacteria with chemicals targeting the Mla system[13]. They showed clear evidence that Mla

defects in *Burkholderia cenocepacia* sensitized them to azithromycin, rifampicin and other antimicrobial agents that are generally minimally effective on Gram-negative strains. They concluded that chemical inhibition of Mla could be an effective adjuvant for many existing antibiotics outside of their current functional spectrum. As such, the ability to screen for chemical inhibitors of the Mla system and validate the effectiveness of proposed Mla inhibitors is a valuable topic in the field of antimicrobial resistance.

In 2020 Huang et *al*. identified clorobiocin as a potential inhibitor of MlaC through simulated docking[219]. They suggested that clorobiocin has high affinity for the MlaC lipid binding pocket and showed that it presents synergistic effects when used with the antimicrobial peptide LL-37 in reducing *A.baumannii* colony forming units. As LL-37 is believed to disrupt the bacterial membrane, they attributed the observed synergistic effect to inhibition of Mla activity. Clorobiocin is structurally similar to novobiocin, an aminocoumarin antibiotic that they also investigated but did not observe synergistic effects with LL-37. Novobiocin is known to target the ATPase activity of DNA gyrase by competing for the ATP binding pocket. A 2017 paper by May et *al*. suggests that novobiocin may also interact with the ABC complex of the Lpt system[109]. Interestingly, they present both structural and functional evidence that it does not interact with the ATP binding pocket of LptB but instead binds to LptB at the interface between the NBD and TMD. Moreover, they showed that rather than a negative effect on Lpt activity, novobiocin was capable of increasing the rate of LPS transport in an *in vitro* assay and recovering Lpt activity in a mutant with a defective LptB – LptFG interface. The effects of novobiocin on LPS transport were not addressed in the Huang paper and may have been a potential oversight, explaining why they did not observe a synergistic killing effect of a known DNA gyrase inhibitor and an OM destabilising agent. We believe that an assay for Mla lipid exchange may play a role in determining if this is the case. That aside, the potential to screen for

inhibitors *in silico* given an understanding of a proteins binding site is a powerful tool, and we believe that validating speculative residues involved in binding between MlaD and MlaC may among other things open *in silico* screening for molecules that perturb the interaction as a potential avenue for further investigation.

1.6.2.The Structure of MlaD

MlaD is a component of the IM MlaFEDB complex and interfaces with the transmembrane protein MlaE. It is periplasmic facing and tethered to the membrane by a single transmembrane helical domain. As part of the FEDB complex it forms a homo-hexameric ring but it has been shown to stably assemble into an SDS resistant hexameric state even in the absence of the remainder of the FEDB complex and without its transmembrane domain[50]. Ekiert et al. presented a crystal structure of the hexameric assembly at 2.85 Å resolution [Figure 1.24]. MlaD is an mce domain protein and adopts the characteristic 7 stranded β-barrel fold [Figure 1.25]. The hexameric ring has a hydrophobic channel through its centre, lined mainly by the residues from the β5 – β6 loop and the inward facing residues of the central α-helical bundle [Figure 1.24C]. The channel has a minimum diameter of approximately 6 – 8 Å at constriction points formed by the tip of the β5 – β6 and by Phe150 of the C-terminal helix. This channel diameter is suitable to allow the transport of small hydrophobic molecules and phospholipids. It was determined that the central α-helical bundle, which accounts for slightly less than half of the multimerization interface, plays a role in the overall oligomerisation state of the protein, as a Δ143 – 153 truncated C-terminal helix resulted in the formation of heptameric complexes and mutations in this region have been shown to severely impair Mla activity *in vivo*[50].

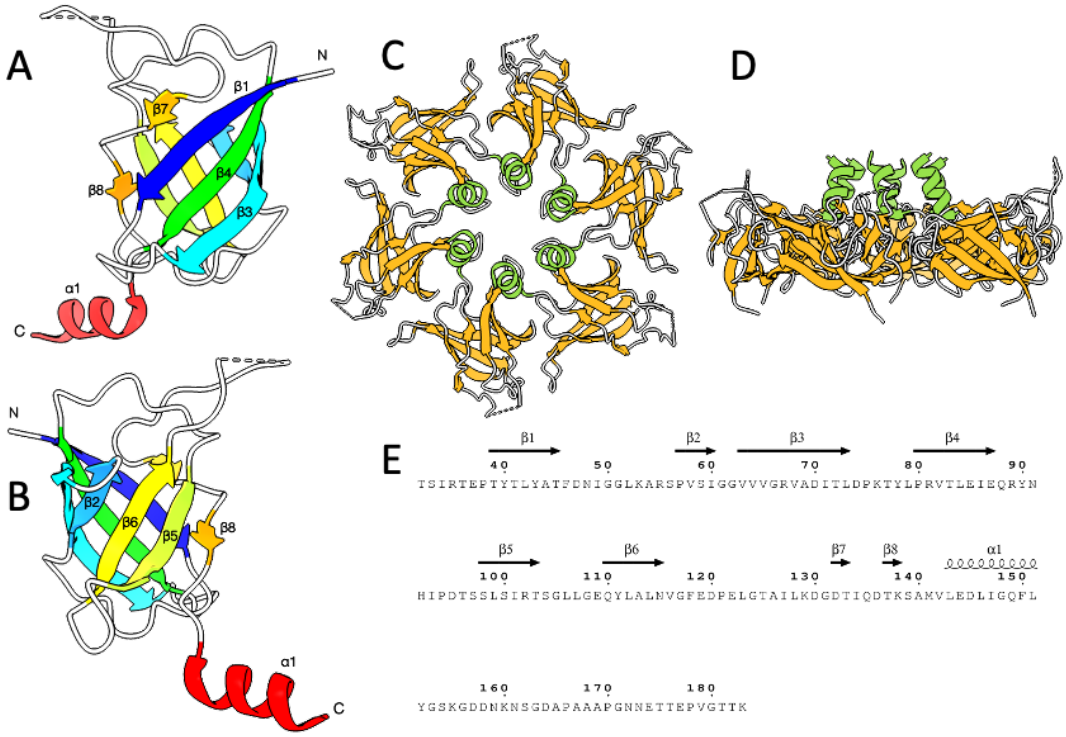

Figure 1.24 - Ribbon diagram of the 5UW2 structure of the *E. coli* holo-MlaD monomeric unit (A) as well as a top (C) and side (D) view of the hexameric assembly. Sequence information Is provided alongside (E) with secondary structure elements aligned to the protein sequence. Front (A) and back (B) views of the monomeric unit are displayed for clarity. Secondary structure elements of the monomeric unit are coloured blue to red successively from the N- to C-terminal. The hexameric assembly is coloured by secondary structure, with β-strands coloured orange and α-helices coloured green. Residues beyond the α1 helix are unmodelled.

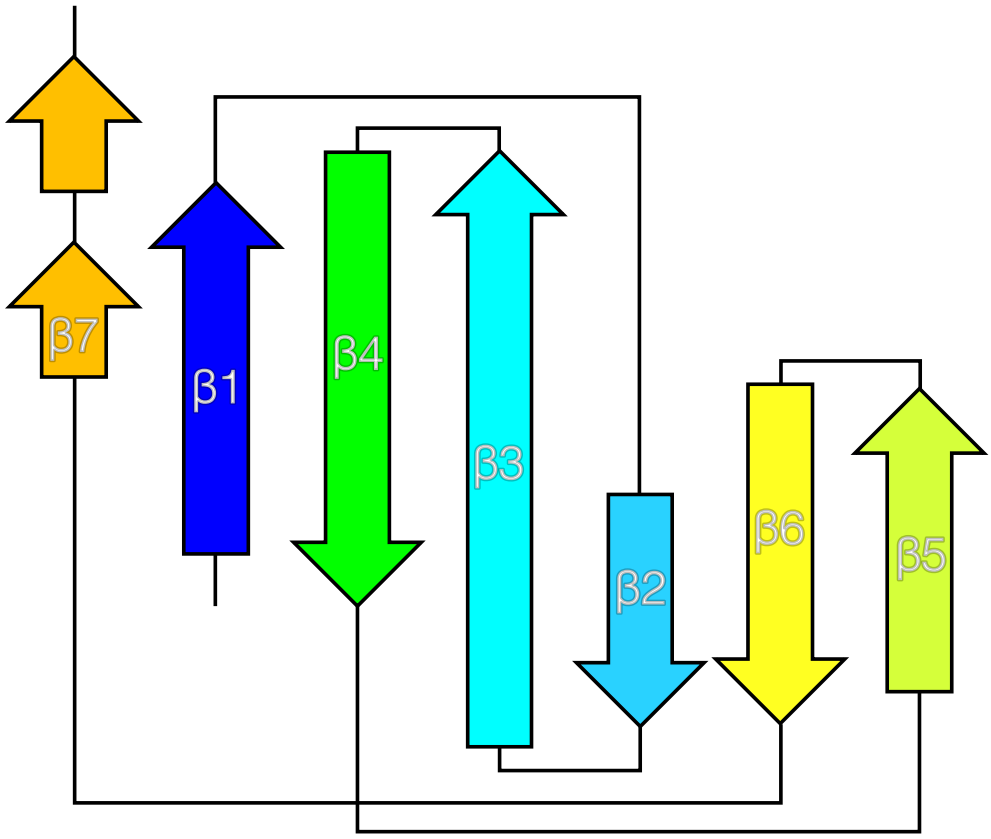

Figure 1.25 - Topology diagram of the MCE structural motif. Colouring is consistent with colouring of the MlaD monomeric unit in [Figure 1.24].

MlaD, like MlaC, is known to natively purify with lipids bound. This has been shown by both by TLC and native mass spectrometry[184]. Native mass spectrometry even suggests that there are multiple lipid binding sites, and that each MlaD hexamer purifies with multiple lipids. It was determined, using an estimated molecular weight of ~750 Da for the lipid substrate that the MlaD hexamer has at least 4 lipid binding sites. However, crystal structures of soluble MlaD constructs lack defined density for the lipid substrate. This may be attributed to differences in lipid position averaging over the crystal lattice because of large variability in occupancies and binding orientations.

1.6.2.1.The Structure of MlaFEDB Reveals Lipid Binding Locations

More recent cryo-EM structures of the FEDB complex show density for the lipid substrates bound in several different positions. However, these substrates are predominantly bound at interfaces between MlaD and MlaE in pockets referred to as the central, lower and outer cavities by Ximin Chi et *al.*[30] [Figure 1.26]. It is not entirely surprising that the central hydrophobic pore of MlaD forms a substrate binding pocket with the TMD of the FEDB ABC complex, however, if this is representative of the lipid binding pocket in the soluble construct of MlaD it would suggest that entire hydrophobic acyl tails of the bound GPLs are exposed to the bulk solvent. Cryo-EM models show single acyl chains from each substrate lipid interacting with MlaD only at the β5 – β6 loop region of the MlaD central pore. Chi et *al.* suggest that the phospholipid head group interacts with the positively charged Lys53 and Arg55 on the bottom face of MlaD[30] [Figure 1.28]. This charge dependent interaction with the lipid head group lends some support to prior evidence suggesting MlaD may have a preference for PG lipids. It has also been shown by Chi et *al.* that residues of the β5 – β6 loop of MlaD, specifically Leu106 and Leu107, are core to the function of MlaD supporting evidence for their direct role lipid interaction [Figure 1.27]. What was identified as the outer lipid binding cavity also sees an interaction between the phospholipid head group and two positively charged residues of MlaD, Arg55 and Arg67 [Figure 1.28]. However, the interaction between these outer phospholipids and the complex seems to primarily be mediated by interactions of the acyl tails with hydrophobic residues on the TM helices of MlaE and MlaD. The remaining lipid binding sites, identified as the lower binding cavity, are not likely to provide any insight into the binding of lipids in the soluble construct as they rely entirely on the N-terminal TM helices of MlaD and TM helices of MlaE which are not present in soluble MlaD constructs.

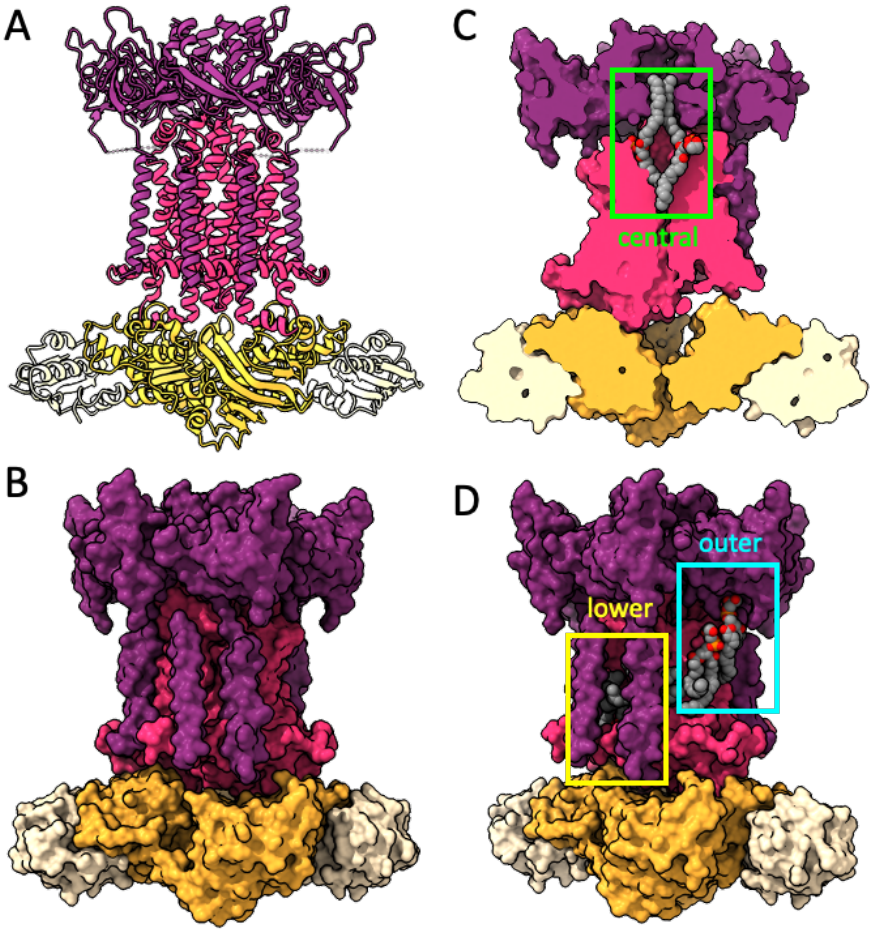

Figure 1.26 - Ribbon diagram of the 7CGE structure of the *E. coli* MlaFEDB complex as determined by Chi et *al.* with lipids removed (A). The MlaD component (purple), MlaE component (magenta), MlaF component (yellow) and MlaB component (tan) are differentiated by colour. A molecular surface diagram of the complex is presented alongside (B). The molecular surface diagram is cut to show the orientation of lipids in the central binding pocket, outlined in green (C). The lower and outer binding locations are highlighted in yellow and cyan respectively (D). One MlaD TM helix has been made transparent to better show lipids in the outer binding location. Both the lower and outer binding locations are primarily formed between the TM helices of MlaD and MlaE, while only the central binding location involves residues present in soluble constructs of MlaD.

The ATP-dependant conformational changes that enable phospholipid transport through the FEDB complex have been identified using an MlaF mutant with attenuated ATPase activity through an E170Q mutation in the Walker B motif. Two new conformations dubbed EQ_{tall} and EQ_{close} were identified. EQ_{close} depicted the ATP-bound conformation and shows substantial reorganisation of the TM helices of MlaE

resulting in the collapse of the central binding cavity and inducing a 30° rotation of the MlaD hexamer[30]. The EQ_{tall} conformation was thought to be some form of intermediate lipid transit state and exhibited an expansion of the central binding cavity owing to large conformational changes in the MlaD hexamer, wherein the entire hexamer lifts and tilts outward. In both of these conformations very marginal changes occur to the core fold of each MlaD monomer. The major structural differences occur in the central α-helix bundle and the TM domain as well as in the relative positions of the monomers in the hexamer. These conformational changes are thought to be induced primarily in response to conformational changes in MlaE, which in turn are caused by ATP-dependent conformational changes in MlaF[30]. Despite strong evidence that conformational changes in MlaD are driven primarily by conformational changes of other components of the complex, there is a lack of evidence to suggest that substrate binding has no effect on the conformation of MlaD as no structure exists of a delipidated MlaD.

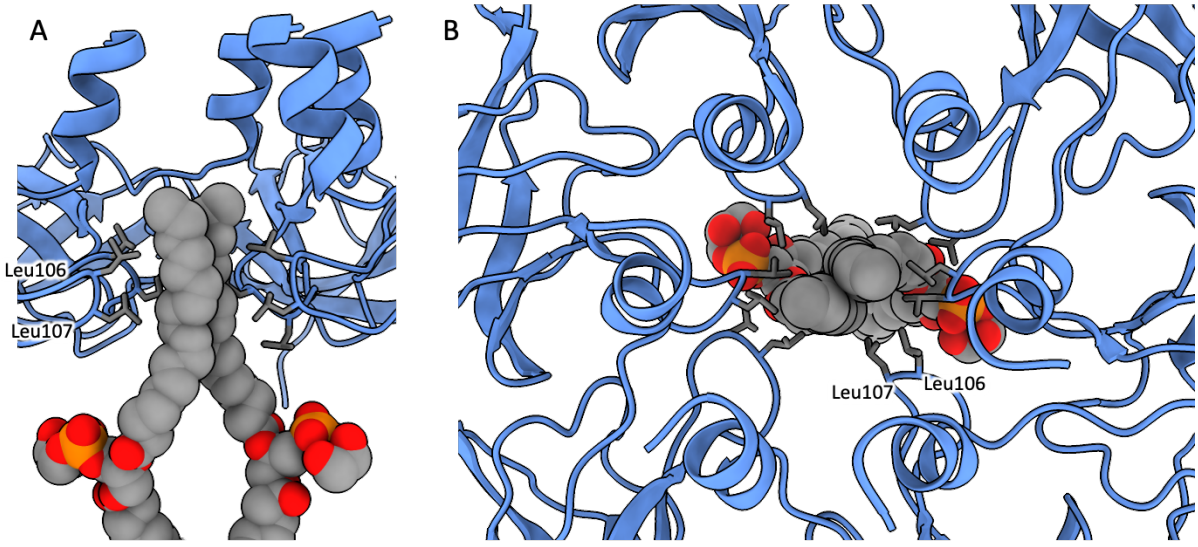

Figure 1.27 - Ribbon diagram showing the interaction of the β5 – β6 loops of MlaD with phospholipids. The side view (A) has 2 MlaD monomer units removed to enhance visibility. The two residues on the β5 – β6 loop, Leu106 and 107, are shown in grey. The top view (B) shows how the Leu residues from each monomer surround and interact with the phospholipids.

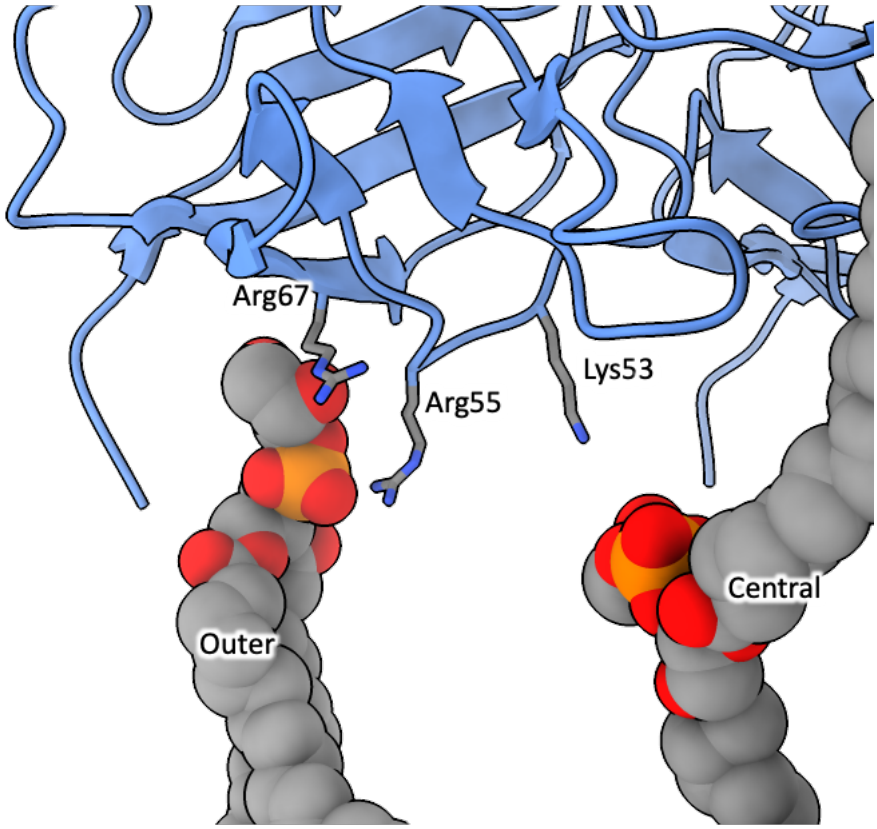

Figure 1.28 - Ribbon diagram demonstrating the proximity of Lys53, Arg55 and Arg67 to phospholipid head groups. It is suggested that Lys53 and Arg55 interact with the head group of the phospholipid in the central binding location during transport. Arg55 is also believed to interact with the lipid in the outer binding location along with Arg67.

1.6.3.Current Understanding of the MlaC – MlaD Interaction

Compared to the emerging literature on the mechanism of action of the MlaFEDB complex, there have been few investigations into the interaction between MlaFEDB and MlaC, and there is little understanding of the mechanism by which the exchange of phospholipids between these proteins occurs. Ercan et *al*. (2019) investigated potential residues of interest in the interaction using *para*-benzoyl-L-phenylalanine (*p*Bpa) mutations[53]. The photo-crosslinking amino acid *p*Bpa has been shown to react with C–H bonds when activated by light between 350 – 360 nm, and this property allows it to be used as a probe for local protein-protein interactions[44]. The

incorporation of *p*Bpa into proteins *in vivo* relies on a system developed by Chin et *al*. which used alternating positive and negative selection to encourage the mutation of an aminoacyl-tRNA synthetase capable of selectively attaching *p*Bpa to the CUA tRNA anticodon[212]. Using this mechanism, Ercan et *al*. identified residues from MlaC and MlaD that are within range to photo-crosslink during interaction[53]. They determined that the Val171 residue of MlaC is a key residue in its interaction with both MlaD and MlaA. They also found that MlaC with mutations to Pro124, Glu169 and Ser172 crosslinked to MlaD but with a reduced rate compared to Val171 [Figure 1.29]. Mutations to MlaD suggested that Phe118, Tyr152, Gln149 and Met141 are the key residues of MlaD involved in interacting with MlaC [Figure 1.30]. Regarding the significant residues of MlaC, all but the Pro124 residue are positioned around the opening of the lipid binding site. The residues of MlaD are primarily positioned around the central helix bundle with the exception of Phe118. Ercan et *al*. concluded that MlaC likely delivers lipids to the central hydrophobic pore. Of significant note is the no MlaC *p*Bpa mutations reduced SDS/EDTA sensitivity of the cells to the level of a knockout, however, the Glu169 mutation did result in a 2-fold reduction in colony forming units (CFUs), suggesting that conservation of this residue may be significant to lipid exchange or folding. Comparatively, all Δ*p*Bpa MlaD mutants significantly reduced cell viability on SDS/EDTA plates.

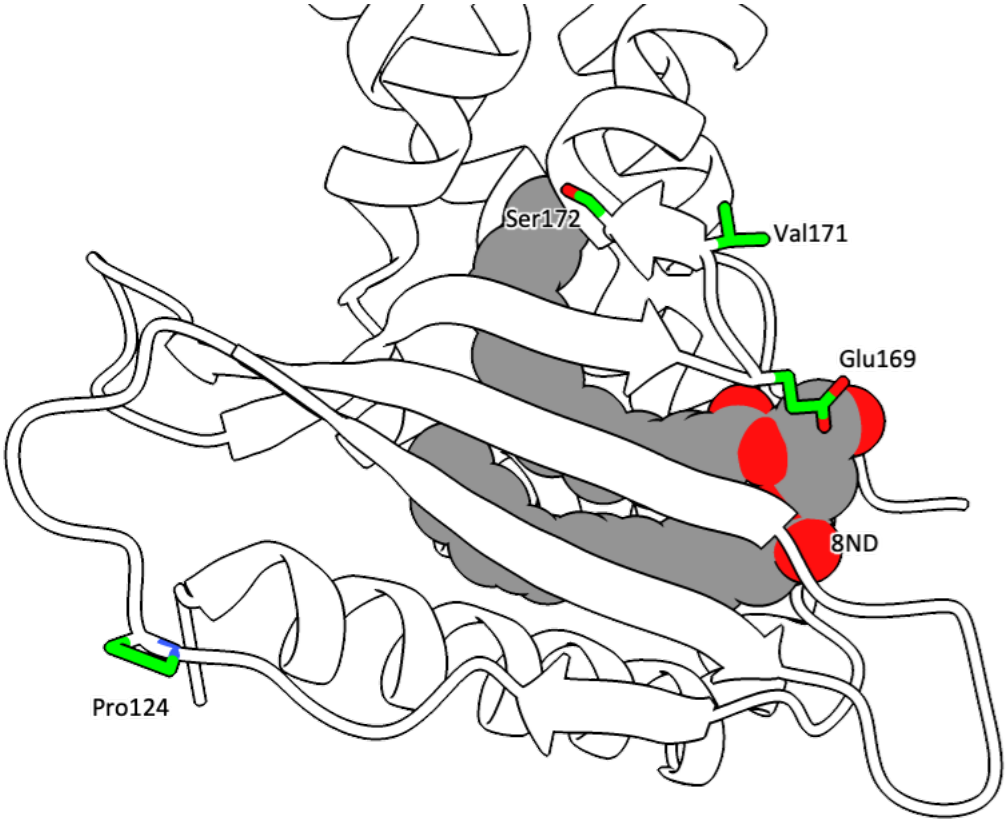

Figure 1.29 - Ribbon diagram of the 5UWA structure of natively purified MlaC. Residues identified by Ercan et al. to be involved in interactions with MlaD are highlighted in green. Most identified residues localise to the β4 - β5 region, which sees significant conformational change between the -apo and -holo states. The exception to this is Pro124, which localises to a loop region far from the other identified residues.

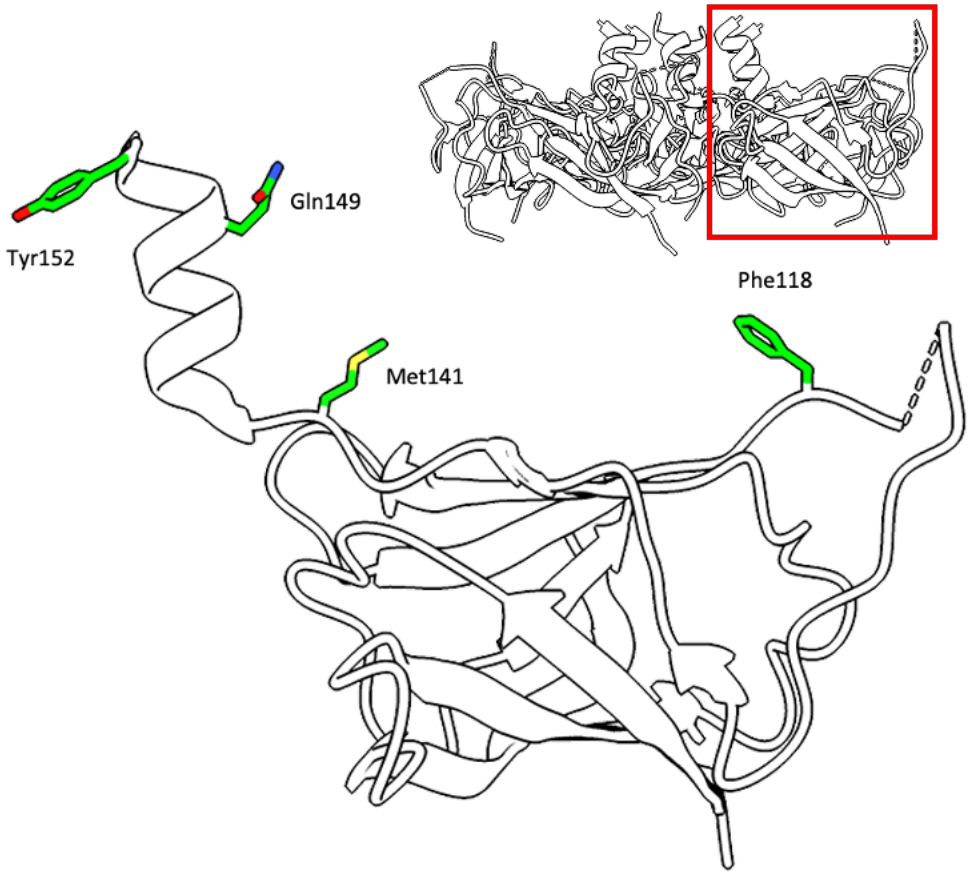

Figure 1.30 - Ribbon diagram of the 5UW2 structure of natively purified MlaD. Residues identified by Ercan et al. to be involved in interactions with MlaC are highlighted in green. The hexameric assembly is included (top right) to show the relative orientation of the monomer.

1.7.A Brief Overview of Protein – Protein Interactions

A principal objective of molecular biology is to develop a mechanistic understanding of cellular processes at a molecular level. Protein-protein interactions are central to a vast array of key biological processes, and with that comes a need for a high degree of specificity with regards to their interaction interfaces. These features of protein surface interactions and the specific structural surface elements that underpin them have sometimes been referred to as the quinary structure, the fifth level of protein structural complexity. In comparison to the four preceding levels of complexity, that

pertain only to isolated proteins and protein complexes, the quinary structure addresses how proteins engage with each other in a crowded cellular environment where transient interactions between macromolecules occur[110].

It has long since been determined that shape complementarity can be a major predictor of protein-protein interaction partners and this methodology has formed the basis of many computational approaches to molecular docking. This alone, however, is a rather simplistic approach to understanding the dynamics of interaction partners. As early as 1995 Schreiber and Fersht used high resolution crystal structures of proteins to conduct a comprehensive assessment of the interfacing residues via mutations and kinetics and determined that electrostatic interactions have a guiding role in protein association[159]. They determined from observed co-operative effects of spatially localised charged residues and simulated Brownian dynamics that association of two proteins can be broken down into two phases. A translation of two molecules in space until they collide wherein charged residues contribute to the rate of association in a manner proportional to their electrostatic field potential and a further rotational step where oppositely charged residues guide the docking of the two proteins. They determined that the effects of uncharged residues in an interacting interface come into play only after the initial association stage. Li et *al*. later showed that a few key residues they referred to as 'hot spots' can account for substantial portions of the binding free energies in transient interactions[96]. This was later supported by evidence suggesting that conformationally invariant surface residues act as anchor points that stabilise intermediate binding states[140]. It has generally been observed that these anchoring residues as well as conserved surface residues tend to locate to pre-organised binding groove regions. Molecular dynamics simulations identified conformationally rigid surface residues play a key role in minimising the enthalpy change of complexation and this contributes significantly to favourable binding free energies[180.

Subsequent to initial association protein-protein interactions are generally classified into several distinct groups in literature, typified primarily by the binding affinity of the proteins. There are the permanent interactions, which have high affinity interactions in the micromolar to femtomolar range and are sometimes unstable in an unbound state making their interaction obligate[124]. Then there are the strong transient interactions which have tight, nanomolar binding affinities and remain bound until a triggering event changes their interaction affinity. G-proteins are an example of this class of transient protein interaction. The GDP bound α-subunit remains tightly bound to the βγ-subunits until a ligand bound G-protein coupled receptor interacts with the G-proteins complex, causing conformational changes that result in the exchange of GDP for GTP in the α-subunit and subsequent release of the α-subunit from the βγ-subunits due to reduced binding affinities. The weak transient interactions on the other hand have affinities in the millimolar to micromolar range and interactions occur and break continuously. Current evidence suggests that the binding interaction between MlaC and MlaD is transient and independent of substrate binding[71].

Contrary to the dynamics of protein-protein association, investigations into the mechanics of dissociation are much more sparse. A recent paper has proposed multiple mechanisms applicable to different protein-protein pairs, although, proposes that the underlying mechanisms can often involve changes in the electrostatic potential of either molecule or mechanically induced conformational changes due to events like ligand binding.

1.8.Aims

The Mla pathway is for many reasons a protein pathway of significant interest and currently the most major gaps in knowledge regarding the pathway pertain to the interaction between MlaC and MlaD. It is our intention to resolve this. As such, the aims of this thesis are as follows:

1) To determine if MlaC forms any significant interaction with the phosphate head group of its phospholipid substrate

2) To determine if any significant conformational changes in MlaD are associated with phospholipid binding

3) To determine the interface of binding between MlaC and MlaD, as well as binding orientations and stoichiometries

4) To identify specific residues involved in the interaction between MlaC and MlaD

5) To provide enough evidence to propose a mechanism for lipid exchange between MlaC and MlaD in response to ATPase activity

2.Materials and Methods

2.1.Transformation of Bacterial Cell Lines

DNA corresponding to *E. coli* MlaC was previously cloned into a custom IPTG inducible plasmid (pBE1203) (Amp^R) by Ekiert *et al.*, with the leading pelB signal sequence (residues 1 – 21) replaced with an N-terminal hexa-histidine tag and TEV cleavage site[50].

DNA corresponding to the periplasmic domain of *E. coli* MlaD (residues 32 – 183) with an N-terminal hexa-histidine tag (MlaDΔTM) was synthesised (Genscript) and cloned into an IPTG inducible pET26b (Kan^R) vector between the XhoI and NdeI restriction sites in a manner to introduce a stop codon before the intrinsic pET26b C-terminal hexa-histidine tag and to remove the pET26b pelB signal sequence.

The *E. coli mlaFEDCB* gene was previously cloned into an IPTG inducible pET22/42 (Amp^R) plasmid by Thong *et al.*, with site-directed mutagenesis of *mlaE* to introduce an N-terminal hexa-histidine tag[184].

Constructs were transformed into BL21(DE3) strain *E. coli* cells (Novagen) using heat shock. Between 1 – 5 μL of DNA was incubated with 20 – 50 μL of bacterial cells for 30 mins on ice. Bacteria were heat shocked at 42°C for 30 sec and then put on ice for a further 2 mins. SOC solution was then added at 5 times cell volume and cells were grown at 37°C for 45 mins. Once grown, cells were plated on lysogeny broth (LB) agar with resistance selection and grown at 37°C O/N.

2.2.Growth and Purification of Soluble Mla Constructs

Transformed cells were grown at 37°C in LB to an OD_{600} = 0.6, before induction with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and overnight expression at 18°C. Cells were harvested by centrifugation at 4500 x *g* for 10 min and resuspended in a buffer of 50 mM Tris(hydroxymethyl)aminomethane (Tris) pH 8.0, 500 mM NaCl, 10 mM imidazole supplemented with EDTA-free protease inhibitors. Cells were lysed at 17000psi in a C3 Emulsiflex cell disruptor (Avestin) before centrifugation at 45000 x *g* for 45 min. The lysate was then filtered through a 0.45 micron filter (MerckMillipore) before being passed through a HisTrap™ Ni-NTA column (Cytiva). The column was washed with 5 column volumes of 50 mM Tris pH 8.0, 500 mM NaCl, 10mM imidazole (MlaC) or 50 mM Tris pH 8.0, 500 mM NaCl, 50 mM imidazole (MlaDΔTM) before the remaining protein was eluted with 50 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazole. Protein fractions were pooled and gel filtered through a Superdex S75 (GE Healthcare) size exclusion column (MlaC) or a Superdex S200 (GE Healthcare) column (MlaDΔTM) in a buffer of 20 mM Tris pH 8.0, 150 mM NaCl. MlaC was incubated with TEV protease (NEB) overnight at a 100:1 molar ratio to cleave the poly-His tag before being flowed through a Ni-NTA column to remove the TEV protease.

2.3.SDS-PAGE

Protein samples were mixed at a 1:1 ratio with 2x Laemmmli sample buffer (Sigma-Aldrich) containing 4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125M Tris-HCl pH 6.8. Samples that were boiled were heated at 95°C for 10 mins in sample buffer. Samples were loaded into Tris-glycine 4-12% gels (Bio-Rad) and were electrophoresed in a Tris-MES buffer (Genscript) containing 50 mM Tris, 50 mM MES, 0.1% SDS, 1 mM EDTA at 160V for 45 mins. Gels

were stained with Instant*Blue*™ Ultrafast Protein Stain (Sigma-Aldrich). Samples electrophoresed in non-reducing conditions were instead mixed at a 1:1 ratio with a sample buffer containing 4% SDS (w/v), 20% glycerol (v/v), 0.004% bromophenol blue (w/v) and 0.125M Tris-HCl pH 6.8.

2.4.Determination of Protein Concentration

Protein concentrations were determined through absorbance at 280 nm. This was conducted using a DS-11 nanodrop spectrophotometer (DeNovix). Extinction coefficients were estimated from the sequence of the respective protein by summing the contributions of Trp, Tyr and Cys to the extinction coefficient (ε = 5500{nTrp} + 1490{nTyr} + 125{nCys}). For the major proteins/complexes purified in this investigation the extinction coefficients evaluated as follows: MlaC – 34380 $M^{-1}cm^{-1}$, MlaD (monomer) – 8940 $M^{-1}cm^{-1}$, MlaD (hexamer) – 53640 $M^{-1}cm^{-1}$ and MlaFEDB – 216315 $M^{-1}cm^{-1}$.

2.5.Formation of Apo Proteins

Native protein was bound to a Ni-NTA column and washed with 5 column volumes of 50mM Tris pH 8.0, 500mM NaCl, 10mM imidazole, 50mM n-octyl β-D-glucopyranoside (β-OG) followed by a 1 hour incubation. This was repeated three times before reflowing overnight. The column was then washed with 10 column volumes of 50mM Tris pH 8.0, 500mM NaCl, 10mM imidazole to remove detergent before the protein was eluted using 50mM Tris pH 8.0, 500mM NaCl, 500mM imidazole and buffer exchanged into 20mM Tris pH 8.0, 150mM NaCl by dialysis through 3.5K molecular weight cut off Snakeskin dialysis tubing (ThermoFisher)[71].

2.6.Thin-Layer Chromatography

To determine the lipid binding state of the natively purified proteins, as well as to validate the effectiveness of detergent washes or lipid loading we used thin-layer chromatography. Lipids were extracted from either MlaC or MlaDΔTM by chloroform methanol extraction. The protein was diluted to 0.5mg/ml and 2ml was mixed with 1ml of chloroform and 2ml of methanol. The samples were vortexed for 5 mins, incubated at 50°C for 30 mins then vortexed again for 5 mins. The mixture was then centrifuged at 2250 x *g* for 10 mins and the lower chloroform phase was extracted. The chloroform phase was evaporated and the dried lipids were resuspended in 100µl of chloroform. Silica TLC plates (Sigma-Aldrich) were loaded with 5µl of sample and lipid standards before being run with a 6.5:2.5:1 (chloroform : methanol : acetic acid) solvent until the solvent front was 1cm for the top of the plate. The plate was allowed to dry before staining with 10% phosphomolybdic acid (PMA) in ethanol. The plate was developed by heating until the solvent front was clearly visible.

To investigate lipid transport between MlaC and MlaDΔTM we used TLC combined with size exclusion chromatography (SEC) separation. Equimolar concentrations of MlaC and MlaDΔTM were incubated together on ice for 30 min. The mixed sample was separated by gel filtration through a Superdex S75 column in a buffer of 20 mM Tris pH 8.0, 150 mM NaCl. Elution fractions were analysed by SDS-PAGE and fractions containing MlaC and MlaDΔTM were separated. Lipids from the separated samples of MlaC and MlaDΔTM were then extracted and run on TLC plates as described above.

2.7.Phospholipid Loading of Mla Components

Liposomes were prepared from chloroform solubilised 1-palmitoyl-2-oleoylglycero-3-phosphoglycerol (POPG) (Avanti), 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) (Avanti) or 1',3'-bis[1,2-dioleoyl-sn-glycero-3-phospho]-glycerol (BDO-CL) (Avanti) by evaporation under a stream of nitrogen to deposit multilamellar lipid stacks. The dried lipid was resuspended to 2mg/ml (POPG), 0.5mg/ml (POPE) or 0.45mg/ml (BDO-CL) in 20mM Tris pH 8.0, 150mM NaCl by bath sonication for 1 hour to form small unilamellar vesicles[150]. Loading of apo-MlaDΔTM was conducted by overnight incubation of the protein with prepared liposomes at a 5-fold molar excess of lipid. MlaDΔTM was then separated from remaining liposomes by His-trap purification and buffer exchanged into 20mM glycine pH 9.0, 150mM NaCl by dialysis through 3.5K molecular weight cut off Snakeskin dialysis tubing (ThermoFisher). Loading of apo-MlaC was conducted by on column reflow of MlaC over a loaded sample of MlaDΔTM and subsequent separation by Ni-affinity purification. Lipid loaded MlaC was kept in a buffer of 20mM Tris pH 8.0, 150mM NaCl.

2.8.Crystallisation and Data Collection

Crystallisation of MlaC-POPG – MlaC-POPG was centrifugally concentrated to 45 mg/ml using a 3 kDa Amicon centrifugal filter unit (MerckMillipore) and set up in sitting drop vapour diffusion crystal trays (Molecular Dimensions) using the JBScreen Classic (Jena Biosciences) crystal screen with 2μl of protein in 20mM Tris pH 8.0, 150mM NaCl and 2μl of reservoir solution per well. Crystallisation was observed in a well containing 2.2M ammonium sulphate, 20% w/v glycerol. Crystals produced from initial screening were used for streak seeding in hanging drop plates (Molecular

Dimensions). A cat whisker was used to streak the drops. Crystallisation occurred in a drop containing 2.0M ammonium sulphate, 20% w/v glycerol. The crystals were fished and plunge frozen in liquid nitrogen. No additional cryoprotectant was used. X-ray diffraction data was collected at Diamond Light Source beamline IO4, indexed to a H3 space group and processed using XIA2[195]. The structure was solved by molecular replacement with 5UWA as the search model. Phasing, building/adjustment and refinement was conducted using the CCP4[196], Phenix[99] and Coot[52] program suites.

Crystallization of MlaDΔTM -apo – MlaDΔTM -apo was spin-concentrated to 15 mg/ml using a 10 kDa Amicon centrifugal filter unit and set up in sitting drop vapour diffusion crystal trays using the MIDAS*plus*™ (Molecular Dimensions) crystal screen with 2µl of protein and 2µl of reservoir solution per well. Crystallisation was observed in a well containing 0.1 M MES pH 5.5, 12% w/v polyvinylpyrrolidone. Further additive screening resulted in the formation of diffraction quality crystals in wells containing 0.1 M MES pH 5.5, 12% w/v polyvinylpyrrolidone, 10 mM betaine monohydrate and 0.1 M MES pH 5.5, 12% w/v polyvinylpyrrolidone, 5% DMSO. The crystals were additionally cryoprotected with 25% glycerol. The crystals were fished and plunge frozen in liquid nitrogen. X-ray diffraction data was collected at Diamond Light Source beamline IO4, indexed to a $P2_12_12$ space group and processed using XIA2. Ellipsoidal truncation of the data was conducted using the UCLA-DOE Institute Diffraction Anisotropy Server[186]. The structure was solved by molecular replacement with 5UW2 as the search model. Phasing, building/adjustment and refinement was conducted using the CCP4[196], Phenix[99] and Coot[52] program suites.

2.9.Surface Plasmon Resonance

Determination of the binding affinity, reported as the equilibrium dissociation constant (K_D), of MlaC-apo/-holo for MlaDΔTM-apo/-holo was conducted by surface plasmon resonance (SPR). The proteins were first exchanged into a buffer of 0.01 M HEPES pH 7.4, 150 mM NaCl (HBS). SPR was conducted using a Biacore 3000 (GE Healthcare) onto a Sensor Chip NTA (GE Healthcare). The chip was washed with a buffer of HBS with 0.005% v/v surfactant P20 (HBS-P). The chip was then regenerated with 5 μL of 0.1 M HEPES pH 8.3, 150 mM NaCl, 350 mM EDTA, 0.005% surfactant P20 before being functionalised with 20 μL of HBS-P with 500 μM $NiCl_2$. MlaDΔTM-apo/-holo was then immobilised on the chip by its hexa-histidine tag. MlaDΔTM-apo/-holo was bound to produce a response of ~2000 RU. 5 μL injections of MlaC were then conducted at various concentrations. To account for potential lipid exchange between MlaC and MlaDΔTM the chip was regenerated and re-loaded with MlaD after each injection of MlaC. The experiments were conducted at a flow rate of 5 $\mu L min^{-1}$. The data was processed using the BIAevaluation software package (GE Healthcare) and PRISM (Graphpad).

2.10.Site-Directed Mutagenesis

Forward and reverse DNA primers for site-directed mutagenesis were chemically synthesised (ThermoFisher). Primers were designed to anneal to plasmid DNA back-to-back with modifications incorporated into the forward primer. Primers were adjusted to a concentration of 10 μM in nuclease-free water. Primers and template DNA were mixed and adjusted to a concentration of 1 μM forward/reverse primer DNA, 20 ng template DNA in 12.5 μL of nuclease-free water. To this, 12.5 μL of Q5®

High-Fidelity 2x Master Mix (NEB) containing dNTPs and a high-fidelity DNA polymerase was added for a final volume of 25 μL. PCR amplification was conducted using a standard thermocycling protocol, with a 30 sec initial denaturation at 98°C followed by 25 cycles of a 10 sec 98°C denaturation, 30 sec annealing at a calculated primer annealing temperature and 30 sec/kilo-base elongation at 72°C. A final 2 min extension at 72°C was conducted before the sample was cooled to 4°C. Thermocycling was conducted in a Veriti™ Thermal Cycler (Applied Biosciences). The template DNA was then broken down while the PCR product was phosphorylated and ligated in a combined kinase-ligase-DpnI reaction by treated with a KLD Enzyme Mix (NEB) for 5 min at room temperature.

DH5-α strain *E. coli* cells (NEB) were transformed with the ligated PCR product using heat shock. 5 μL of the ligated PCR product was incubated with 20 μL of bacterial cells for 30 mins on ice. Bacteria were heat shocked at 42°C for 30 sec and then put on ice for a further 5 mins. SOC was then added at 5 times cell volume and cells were grown at 37°C for 45 mins. Once grown, cells were plated on LB agar with resistance selection and grown at 37°C O/N. Colonies of transformed cells were then selected and grown at 37°C in LB O/N. Plasmid DNA was then extracted and isolated through alkaline lysis.

Bacterial cells from a 5 mL culture were pelleted by centrifugation at 2250 x *g* for 10 min at room temperature. Cells were resuspended in 250 μL of a resuspension buffer (Qiagen) containing 50 mM Tris pH 8.0, 10 mM EDTA, 100 μg/mL RNase A. Cell lysis was induced by the addition of 250 μL of an alkaline lysis buffer (Qiagen) containing 200 mM NaOH, 1% SDS and allowed to continue for 5 min. Neutralisation was

conducted by the addition of 350 μL neutralisation buffer (Qiagen) containing 4.2 M guanidinium hydrochloride (Gu-HCl) pH 4.8, 0.9 M potassium acetate. Precipitated material was removed by centrifugation at 13000rpm for 10 min. The supernatant was applied to a QIAprep spin column (Qiagen) and the flowthrough discarded. The column was first washed with 0.5 mL of a buffer (Qiagen) containing 5 M Gu-HCl, 30% isopropanol. The column was then washed with 0.75 mL of a buffer (Qiagen) containing 10 mM Tris pH 7.5, 80% ethanol. DNA was eluted from the column with 50μL of 10 mM Tris pH 8.5.

2.11.Generation of Disulphide-Bonded Complex

MlaC Val171Cys and MlaDΔTM Met141Cys mutants were generated by site-directed mutagenesis as described in section 2.9. Transformation of BL21(DE3) with these mutants was conducted as described in section 2.1. Growth and purification of mutants was conducted as described in section 2.2 with the addition of 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) to all buffers. Purified mutants were co-incubated at 50 μM for 30 min before subsequent dialysis into a buffer of 20 mM Tris pH 8.0, 150 mM NaCl at 4°C O/N. Assessment of disulphide bonding was conducted through non-reducing SDS-PAGE as described in 2.3.

2.12.Glutaraldehyde Crosslinking of MlaC to MlaDΔTM

Proteins to be crosslinked were first dialysed into a buffer of 20 mM HEPES pH 7.4, 150 mM NaCl. Both MlaC and MlaDΔTM were co-incubated at room temperature for 30 mins at a concentration of 20 μM prior to crosslinking. Crosslinking was initiated by the addition of glutaraldehyde at 0.115% v/v and allowed to continue for 2 mins

before the reaction was terminated by the addition of 100μL of 1 M Tris pH 8.0 per 1 mL of reaction volume.

2.13.Preparation of BDO-CL Stabilized Complex

MlaDΔTM loaded with BDO-CL at a minimum concentration of 100 μM was incubated with a 5 times excess of MlaC-apo at 4°C for 1 hour. Excess MlaC as well as larger aggregates were removed by SEC purification on an S200 size exclusion column in a buffer of 20 mM Tris pH 8.0, 150 mM NaCl.

Samples for cryo-EM that were further stabilised by glutaraldehyde crosslinking were instead purified by size exclusion in a buffer of 20 mM HEPES pH 7.4, 150 mM NaCl. The sample was then incubated with 0.115% glutaraldehyde for various times from 30 sec to 30 mins before the reaction was terminated by the addition of 100μL of 1 M Tris pH 8.0 per 1 mL of reaction volume. The crosslinked sample was then run on an S200 column in a buffer of 20 mM Tris pH 8.0, 150 mM NaCl.

2.14.Cryo-Electron Microscopy

Initial Screening – An un-crosslinked sample of the MlaC-MlaDΔTM complex was provided for initial screening. Quantifoil 300 mesh gold R1.2/1.3 holey carbon grids were glow discharged for 120 sec at 40 mA and treated with graphene oxide. The protein sample was adjusted to 0.02 mg/ml and 3 μL of the sample was pipetted onto the grid. Blotting and sample vitrification was conducted with the assistance of a Vitrobot System (ThermoFisher). Samples were blotted for 4 sec and then plunge frozen in a liquid ethane cryogen. Data was collected on a FEI Titan Krios G3

microscope (ThermoFisher) using a K3 detector (Gatan) in super-resolution mode with a physical pixel size of 0.87 Å. Processing of the data was conducted using the RELION software package[154].

Final Data Collection – A crosslinked sample of the MlaC-MlaDΔTM complex was provided for data collection. Quantifoil 300 mesh gold R2/2 holey carbon grids were glow discharged for 120 sec at 40 mA. The protein sample was adjusted to 0.8 mg/ml and 3 μL of the sample was pipetted onto the grid. Blotting and sample vitrification was conducted with the assistance of a Vitrobot System. Samples were blotted for 4 sec and then plunge frozen in a liquid ethane cryogen. Data was collected on a FEI Titan Krios G3 microscope using a K3 detector in super-resolution mode with a physical pixel size of 1.086 Å. Processing of the data was conducted using the RELION[154] and CCP-EM software package[198].

2.15.Fluorescence Assay for Lipid Exchange

Liposomes were prepared from a mixture of chloroform solubilised POPG supplemented with 3% 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(5-dimethylamino-1-naphthalenesulfonyl) (dansyl-PE) (Avanti). The chloroform was evaporated under a stream of nitrogen to deposit multilamellar stacks. The dried lipids were resuspended at 1 mg/ml in 20 mM Tris pH 8.0, 150 mM NaCl. Prior to beginning the assay, MlaDΔTM (or an MlaDΔTM mutant) and the fluorescent liposomes were preincubated for a period of 3 hours. The final assay contained MlaDΔTM at 50 μM and fluorescent lipids at approximately 500 μM (calculated from the Mw of POPG). The assay was initiated by the addition of MlaC (or an MlaC mutant) to a final concentration of 200 μM. The assay was conducted in a cuvette with 3 mm path length. Assay progress was tracked by emission at 518 nm using an excitation

wavelength of 280 nm. Data was collected for up to 3600 seconds using a PTI fluorimeter (Horiba Scientific) set with excitation and emission slits at 2nm. Excitation/emission spectra were collected prior to the assay for individual assay components at 50 μM.

2.16.Binding Affinity Determination by Fluorescence Assay

The determination of binding affinity was carried out using the fluorescence assay as described in 2.15, with the concentration of MlaDΔTM reduced to 5 μM. The assay was repeated with a dilution series of MlaC starting at 400 μM. PRISM software was used to fit kinetic models to initial rates of activity. For comparison to MST data, the kinetic models were selected from those used by the MO.Affinity Analysis software (NanoTemper) used in 2.17.

2.17.MicroScale Thermophoresis

Proteins of interest were first dialysed into a buffer of 10 mM PO_4^{3-}, 140 mM NaCl, 2.7 mM KCl (PBS). The His-tagged protein was then adjusted to a concentration of 200 nM by dilution in PBS with 0.05% Tween™ 20 (PBS-T). This was mixed with RED-tris-NTA dye (NanoTemper) at 100 nM in PBS-T with a 1:1 volume ratio. This mixture was incubated for 30 mins at room temperature before centrifugation at 10000rpm for 10 mins to remove aggregates. A dilution series was set up of the non-His-tagged protein by sequential 2-fold dilution in PBS. To each dilution an equal volume of the dye labelled protein was added. The mixtures were allowed to reach a stable equilibrium over a period of 1 hour. Once the dilution series was prepared and equilibrated each sample was loaded into a standard MST capillary tube (NanoTemper). Thermophoretic measurements were performed on a Monolith™ NT.115 system

(NanoTemper). LED power was calibrated to produce a fluorescence intensity of between 200 – 300 counts during preliminary capillary scan and samples were measured with an MST power of 60%. MO.Affinity Analysis software (NanoTemper) was used to analyse and fit the data.

2.18.Growth and Purification of MlaFEDB

Transformed cells were grown at 37°C in LB to an OD_{600} = 0.6, before induction with 1mM IPTG. The culture was grown for a further 3 hours at 37°C. Cells were then pelleted by centrifugation at 4500 x *g* for 20 mins and then resuspended in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl supplemented with EDTA-free protease inhibitors. Cells were lysed at 17000psi in a C3 Emulsiflex cell disruptor. Cell debris was removed by centrifugation at 4500 x *g* for 10 mins at 4°C. The membrane fraction was separated from the supernatant by centrifugation at 145000 x *g* for 1 hour at 4°C in an Optima L-70 ultra-centrifuge (Beckman Coulter). The membrane pellet was resuspended in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, 1% *n*-dodecyl-β-D-maltoside (DDM) (ThermoFisher) and allowed to solubilise for 1 hour at 4°C. Un-solubilised material was removed by a second centrifugation step at 145000 x *g* for 1 hour at 4°C. The supernatant was incubated with 1 mL Ni-NTA resin (Qiagen) for 1 hour at 4°C and then loaded onto gravity flow column (Bio-Rad). The column was washed with 10 mL of wash buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.05% DDM, 20 mM imidazole 8 times. The remaining protein was eluted with 8 mL of elution buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.05% DDM, 50 mM imidazole. The eluted sample was spin-concentrated in a 10 kDa Amicon centrifugal filter unit to a volume of approximately 500 μL. The sample was further

filtered through a Superdex S200 column into a buffer of 20 mM Tris pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.05% DDM.

2.19.Fluorescence Assay with MlaFEDB

MlaFEDB was spin-concentrated and reconstituted into POPG liposomes supplemented with dansyl-PE by rapid dilution. Liposomes were prepared from a mixture of chloroform solubilised POPG supplemented with 3% dansyl-PE. The chloroform was evaporated under a stream of nitrogen to deposit multilamellar stacks. The dried lipids were resuspended at 10 mg/ml in 20 mM Tris pH 8.0, 150 mM NaCl, 0.05% DDM. DDM solubilised MlaFEDB at 0.5 mg/ml and fluorescent liposomes at 5 mg/ml were incubated together for 4 hours at 4°C. The sample was then injected into a Superdex S75 size exclusion column and eluted in a buffer of 20 mM Tris pH 8.0, 150 mM NaCl. The void peak was collected and concentrated to slightly above 5 μM.

Samples of MlaFEDB reconstituted into fluorescent liposomes were mixed with MlaC (or buffer) to a final concentration of 5 μM and 50 μM respectively. The assay was initiated by the addition of MlaC and assay progress was tracked by emission at 518 nm using an excitation wavelength of 280 nm. Data was collected for up to 3600 seconds using a PTI fluorimeter.

3.Results: Determination of the Effects of Substrate Binding on Mla Components

3.1.Introduction

In this chapter we aim to investigate, through structural methods the reason for observed substrate specificity in MlaC and the effects of substrate binding on the conformational state MlaD. In 2016, Thong et *al.* determined by mass spectrometry that a soluble construct of natively purified MlaD bound an equal amount of PE and PG lipids[184]. This is unexpected in that the distribution of membrane lipids they observed in their bacterial membranes was 72% PE, 17% PG and 11% CL. Moreover, in the pivotal work on membrane lipids carried out by Donohue-Rolfe and Schaechter they observed an even further reduced relative percentage of PG compared to PE in the outer membrane[43]. It is clear from TLCs presented in Hughes et *al.* that MlaC has an even greater preference for PG than MlaD[71]. We propose that structural insight into the binding orientation of a single species, physiologically representative substrate of MlaC may prove useful in determining the methods by which lipids are passed between the two proteins. Current models of the lipid binding components of the Mla system and its homologues have been generated from crystallography of the natively purified proteins. It is likely that the models suffer from an issue of crystallographic averaging as the lipids present in a natively purified sample of MlaC are not homologous over a crystal lattice, owing to the abundance of different lipid species that these proteins bind. This likely explains the lack of defined density for the lipid phosphate head group in the 5UWA model. The 5UWA model docks 1,2-dipalmitoyl-sn-glycero-3-phosphate, a species of phosphatidic acid, which is a precursor of major *E. coli* phospholipids as a placeholder. While phosphatidic acid is present in the inner membrane and it fits well to the density observed, it is not likely to be abundant, and is most likely not the major substrate of MlaC but it fits appropriately to the density as a placeholder. The 2QGU *R.solanacearum* homologue

is docked with dipalmitoyl-3-sn-phosphatidylethanolamine and while the modelled lipid fits well to the map a preference for PG has been observed in *E. coli*. We believed that a more accurate model of lipid binding with what has been observed as the preferential substrate in our organism of interest may provide useful information regarding the residues of MlaC involved in lipid exchange. To this end we have selected what we believe to be an appropriate lipid substrate for co-crystallization with MlaC. Lipidomics suggests that acyl chains in the IM and OM of *E. coli* K1062 and K-12 strain LM3118 consist primarily of $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{16:1}$, $C_{18:0}$ and $C_{18:1}$ with minor contributions from odd length and modified acyl chains such as cylco-$C_{19:0}$[128]. However, the most abundant acyl chain in *E. coli* K1062 and K-12 strain LM3118 was determined to be $C_{16:0}$, with an abundance of nearly 3 times that of other acyl chains.

With this in mind, we attempted to crystallize MlaC in complex with a POPG lipid, which we believe to be more representative of its preferred physiological substrates, in an attempt to determine interactions between MlaC and its substrate that may be relevant in lipid exchange.

With regards to MlaD, we aimed to produce a structure of the delipidated soluble construct. It has previously been observed that MlaC exhibits conformational change upon the removal of its natively purified lipid, and that this conformational change is likely relevant to its mechanism of action. We expect that a soluble construct of MlaD may present conformational changes between its native and apo state that would otherwise be induced by conformational changes in the MlaFEB complex. Thus, our aim with MlaD was to determine conformational changes between the apo and native state that may be relevant to lipid exchange.

3.2.Introduction to Methods

3.2.1.Crystallography

Crystallography is a technique used in chemistry and structural biology to determine the bonding arrangements of atoms in molecules. The technique leverages the ordered repetitive nature of crystalline lattices along with the properties of electromagnetic waves, allowing them to diffract around obstacles or openings. Crystallography remains one of the most used techniques to determine the structure of biological macromolecules[113].

A crystal is, in all but very rare cases involving quasiperiodic tiling, a regular arrangement of particles in 3-dimensional space into an ordered lattice[162]. It can essentially be thought of as a 3-dimensional extrapolation of a tessellation of patterned tiles wherein the shape of the tile describes the lattice system. The 3-dimensional tile, translated through space to fill the plane is referred to as the unit cell. The specific orientations of molecules within the unit cell can be thought of as the pattern on the 3-dimensional tile and the smallest repeating unit in that pattern is referred to as the asymmetric unit. It has been shown mathematically that there are only 6 different tile shapes, or lattice systems that can regularly tile a 3-dimensional plane. These lattice systems are further subdivided into 14 Bravais lattice groups, which define the location of lattice points in the system and further describe the symmetry of the pattern in the 3-dimensional tile[22] [Figure 3.1]. These can then be further sub-classified into 230 space groups, which give detailed descriptions of the symmetry of the pattern in the tile, or in other words, how the asymmetric unit is positioned in the unit cell. Any regular organisation of non-standard shapes in 3-dimensional space can be described as one of these space groups[63]. With specific

regards to naturally derived protein crystals there are further limitations to only 65 non-enantiomeric space groups due to the chiral nature of biological molecules.

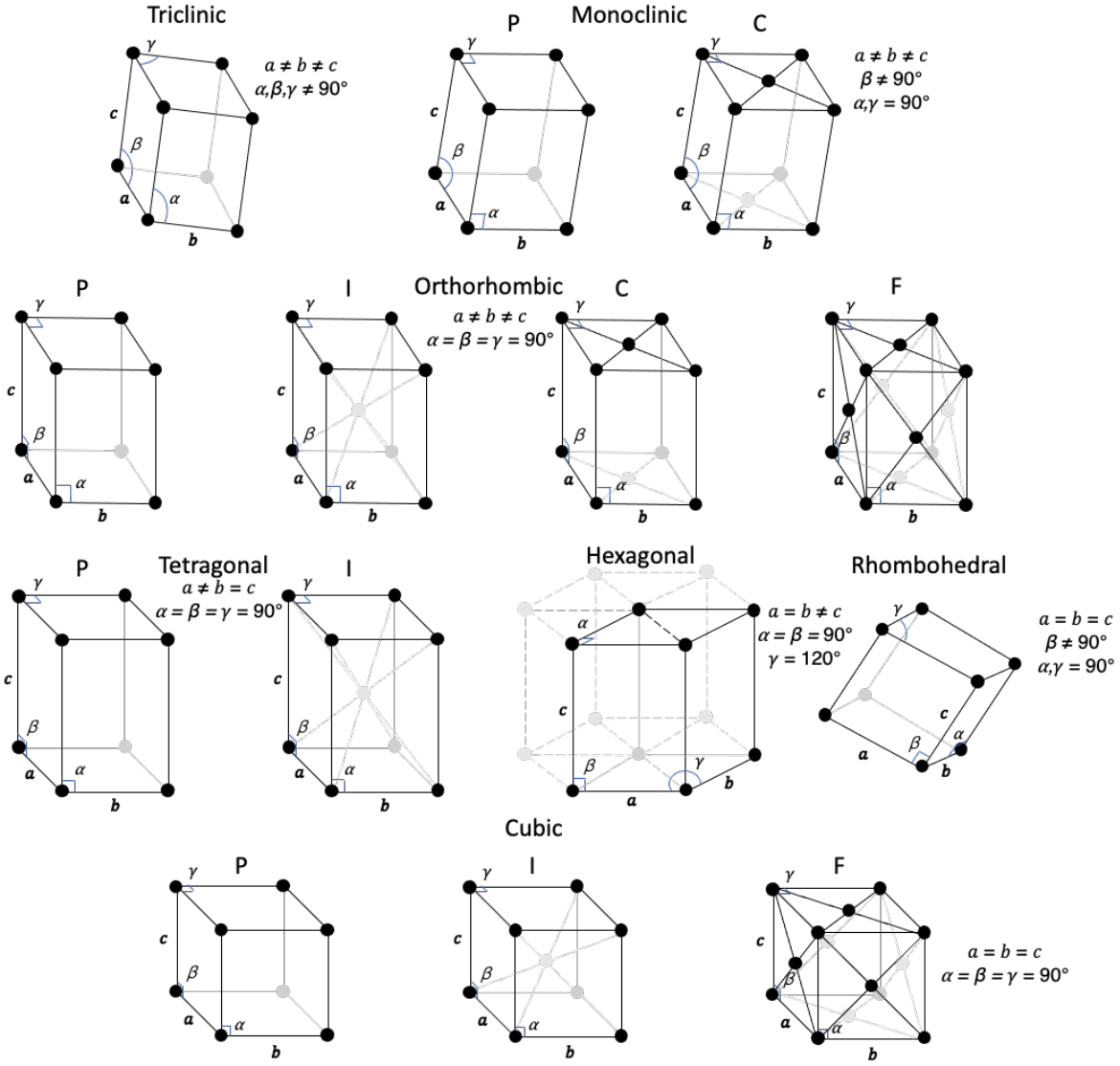

Figure 3.1 - The 14 Bravais lattice groups along with the angles and edge lengths that define them. The monoclinic, orthorhombic, tetragonal and cubic groups are further sub-divided into P (primitive), I (body-centred), C (base-centred) and F (face-centred).

Formation of protein crystals occurs when the concentration of protein in a solution exceeds the limit of solubility. In most cases the protein will form an amorphous precipitate, however, under the right conditions the protein may remain in solution. In this state the solution will be super-saturated, in a similar state to a saturated hot salt solution. Much like a hot saturated salt solution this state is unstable and the

solubilised material will eventually precipitate in either an ordered or disordered manner. If the protein molecules arrange into an ordered manner as they fall out of solution they form micro-crystals, which can act as nucleation sites for the formation of larger macroscopic crystals[111].

Screening for the conditions required to crystallise a protein is usually done using the vapor diffusion method, wherein a solution of protein is mixed with a buffered precipitant salt mix in a small drop of usually μl volume. The drop is equilibrated against a larger reservoir of the precipitant solution. This allows for slow evaporation of the drop and a controlled increase in protein concentration to ideally create the super-saturated conditions to allow nucleation. Usually many different conditions are screened in hopes of finding the optimal precipitant mix to allow for crystallisation[111].

Once a protein has been crystallised, it is in a state of ordered regularity and can be used in a diffraction experiment. A protein crystal can be approximated to a diffraction grating due to this regularity and, as was originally shown in Thomas Young's classic double slit experiment [Figure 3.2], constructive wave interference from a coherent beam diffracted through multiple equally spaced apertures will occur at points where the extra distance travelled by each subsequent diffracted beam is an integer multiple of the beam's wavelength, or at points that satisfy[207]:

$$n\lambda = d \sin \theta$$

Which is to say that two coherent waves originating and in phase at points a and b, respectively, will interfere constructively at point P if the extra distance travelled by the wave from point a is a multiple of its wavelength. As the extra distance travelled by the wave from point a is a function of the distance between points a and b as well

as the angle between the two waves, we arrive at the situation described in [Figure 3.3].

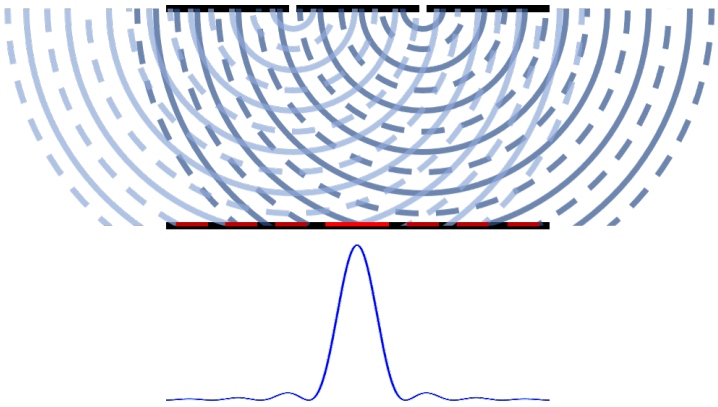

Figure 3.2 - Diagram representing Young's double slit experiment, showing how two diffracting apertures result in bright and dark bands on a detector surface due to constructive and destructive interference. The band intensities are represented graphically below.

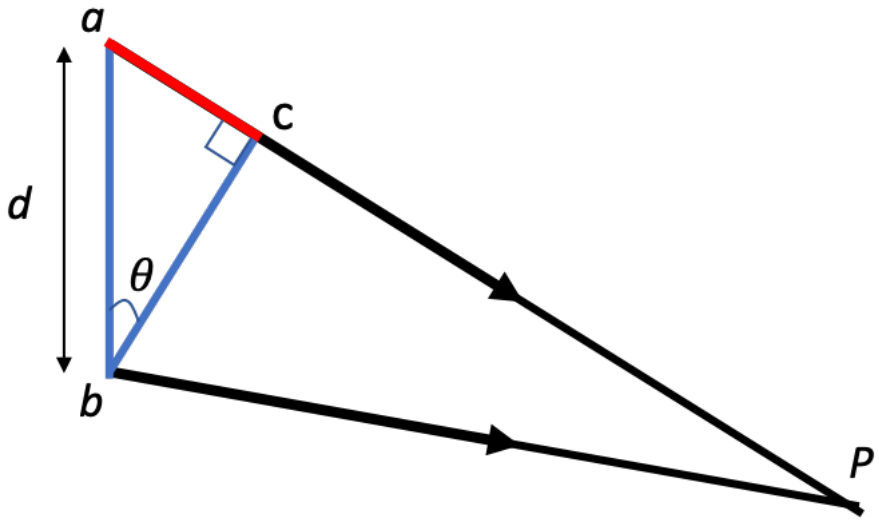

Figure 3.3 - The diagram shows the derivation for the equation $n\lambda = d \sin\theta$. The variable d represents the distance between the two diffracting apertures, P is a point on the detector and θ is the angle between aP and the normal to ab. This equation only holds true when θ is sufficiently small, such that bP and cP are approximately parallel.

Since the distance between diffraction apertures in a crystallography experiment is negligible compared to the distance from the apertures to the collector it can be

assumed that the waves from adjacent diffraction points are very nearly parallel. Due to this the angle between the waves can be expressed as the angle between the normal at a and the vector between a and P, which leads to the more conventional representation of multiple slit diffraction [Figure 3.4]. While bands produced by two aperture diffraction are diffuse around points that satisfy the condition for constructive interference owing to the fact that two waves that are out of phase by only several degrees do not interfere completely destructively as the number of apertures tends towards infinity the diffraction bands become more discreet at a single position. This is because for every wave that is out of phase by 1° there is a wave that is out of phase by (n+1) °. This means that at any point where the waves are not perfectly in phase in multi-aperture diffraction there will be a series of waves out of phase from each other by some amount. In that series will be waves that are perfectly out of phase and completely cancel resulting in no net signal at points that do not satisfy the condition for constructive interference.

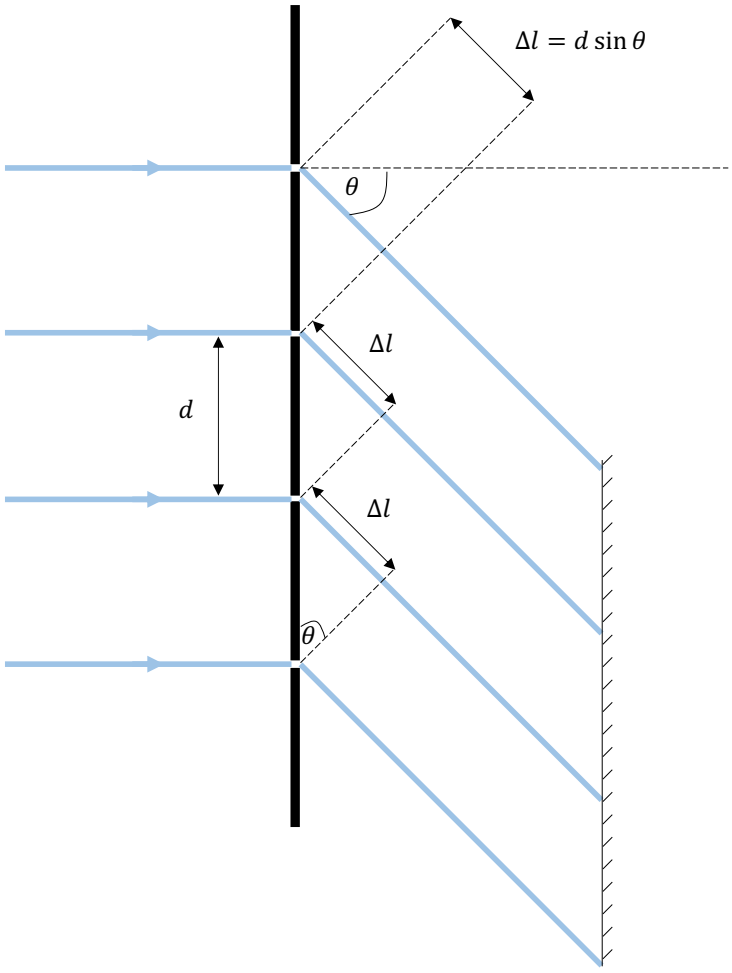

Figure 3.4 - Diagram of multiple slit diffraction describing the additional path length from each subsequent diffraction aperture relative to the previous aperture.

Diffraction in a crystallography experiment differs slightly from diffraction through apertures in that the incident beam is not propagated through a diffraction grating but instead impinges the surface of the crystal at a small glancing angle. The extra distance travelled by each subsequent diffracted beam must therefore also account for the extra distance travelled by the incident beam [Figure 3.5]. As such we arrive at the condition for constructive interference across multiple unit cells in a crystalline lattice, commonly referred to as the Bragg equation[20]:

$$n\lambda = 2d \sin\theta$$

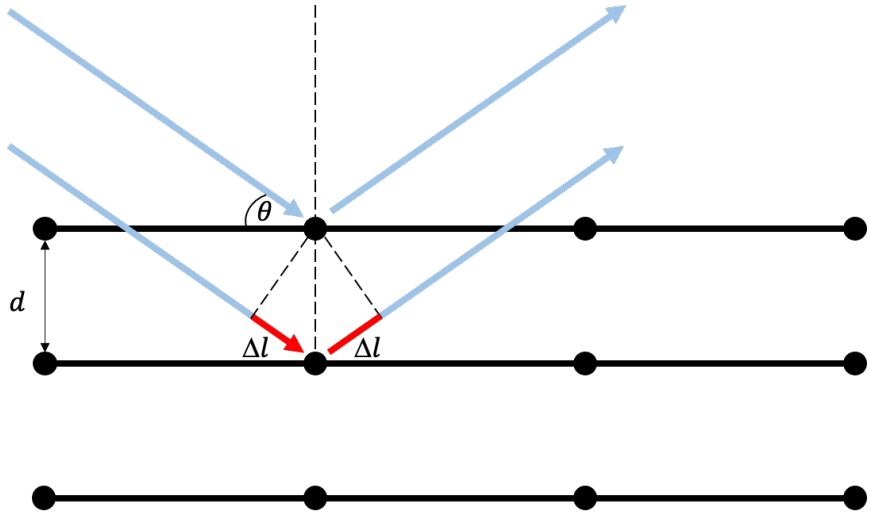

Figure 3.5 - Schematic diagram of Bragg diffraction demonstrating why the difference in path length is 2d sinθ in a crystallography experiment. Reflections will appear at diffraction angles where the difference in path length is an integer multiple of the wavelength.

For all diffraction spacings and incident angles that satisfy the Bragg equation there will be a reflection at the detector. This reflection, quite contradictorily, is a superposition of many diffracted waves that have travelled different distances to the detector. This superposition is not created by lattice regularity and thus is not completely destructive. It is instead a result of the structure of diffracting density within the unit cell. This brings us to the structure factor equation and the reverse Fourier transform of crystallography. The structure factor equation describes how density within the unit cell contributes to final superposition of the wave at each reflection. The equation states[19]:

$$F_{hkl} = \sum_{j=1}^{N} f_j e^{-2\pi i(hx_j + ky_j + lz_j)}$$

This describes the wave diffracted from each scattering entity, j, as the polar form of a complex number proportional to its scattering factor, f_j. This means that the sum,

F_{hkl}, is a polar complex number with a modulus,$|F_{hkl}|$, and argument or phase, φ. A wave propagating through its phase angle as it travels through space can also be modelled as a polar complex number with a modulus equal to its amplitude and argument equal to its phase angle [Figure 3.6]. The phase component of the complex number in this equation is represented as a function of the xyz coordinates of the scattering entity in the unit cell, $(hx_j + ky_j + lz_j)$. Thus the coordinates of a scattering entity at xyz compared to unit cell coordinates $(0, 0, 0)$ is related to the phase angle of a theoretical wave diffracting from it. The superposition formed by waves scattering off all entities in the unit cell can thus be expressed by this polar complex number, F_{hkl}, as summation of polar complex numbers is representative of the formation of the superposition of the waves they represent. This is what it means to say that the positional information of scattering atoms is stored in the phases, which leads to the phase problem.

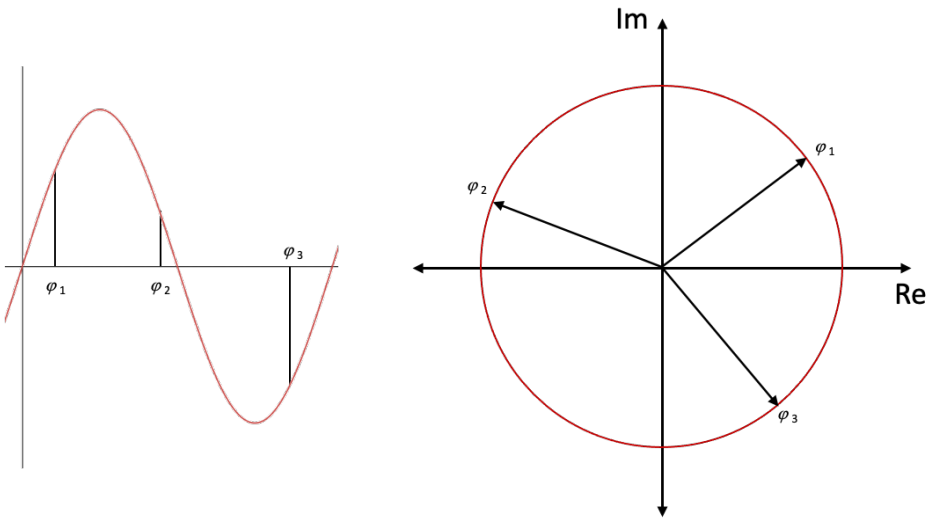

Figure 3.6 - Diagram demonstrating how a sinusoidal wave propagating through space can be represented on a polar graph.

In a crystallography experiment we cannot measure the phases of the waves contributing to the superposition of any given reflection. In fact, we do not even collect

the amplitude or phase of the superposition. What is collected is the intensity of the superposition of multiple waves. As intensity is proportional to amplitude squared, and the absolute square of a complex number is real, $|z|^2 \equiv z\overline{z}$, the intensity derived from the squared modulus of an amplitude represented by a complex number is completely real. This presents an issue, as the electron density equation requires phase information to perform a reverse Fourier transform:

$$\rho(xyz) = \frac{1}{V}\sum_{hkl}|F_{hkl}|e^{2\pi i\varphi_{hkl}-2\pi i(hx+ky+lz)}$$

The initial calculative need for phases is subverted by using the Patterson function, a Fourier transform of the intensities which is the calculative equivalent to a convolution of the density function with its inverse[132]. The output of the Patterson function, however, is interatomic distances rather than atomic density. For simple unit cells, this is sufficient to solve the phase problem and determine the distribution of atoms. However, for unit cells containing macromolecules in a crystal with lattice imperfections it is not possible to reconstruct the unit cell from interatomic distance vectors. A Patterson map of N scattering entities will have at least $N(N-1)$ peaks meaning back-calculating from the Patterson map to phases for a macromolecule is often too computationally intensive to be practical.

There are two main methods to solve a complex unit cell. The first is to introduce atoms with large scattering factors to the unit cell, which contribute more to the intensity of all collected reflections. The difference between the structure factors of a crystal with heavy atoms present and one without will allow for the determination of the phases associated with the heavily scattering atoms and thus they can be positioned in the unit cell. The positions of all other atoms can be determined in Patterson space using interatomic distance vectors from these heavy atoms[99].

The other method of solving the phase problem is through molecular replacement, which essentially uses a molecule of similar structure and known symmetry information about the unit cell derived from the reflection pattern to model Patterson maps for that homologue in different rotational and translational orientations. Correlation between the Patterson map of a correctly positioned homologue and the Patterson map derived from the experimental intensities allows for the solution of the phase problem in Patterson space[99].

3.3.Results

3.3.1.Construct Design of MlaC

We received a construct of MlaC designed for cytoplasmic expression and purification by Ni-affinity chromatography from Ekiert et *al.*[50] [Figure 3.7]. The construct lacks the leading 21 residue signal sequence natively present in the *mlaC* gene and instead replaces it with a hexa-histidine tag and TEV cleavage site. This TEV cleavable His-tag allowed for the separation of MlaC from Ni-affinity purified soluble constructs of MlaD, which is required for the purification of preparative quantities of lipid loaded MlaC.

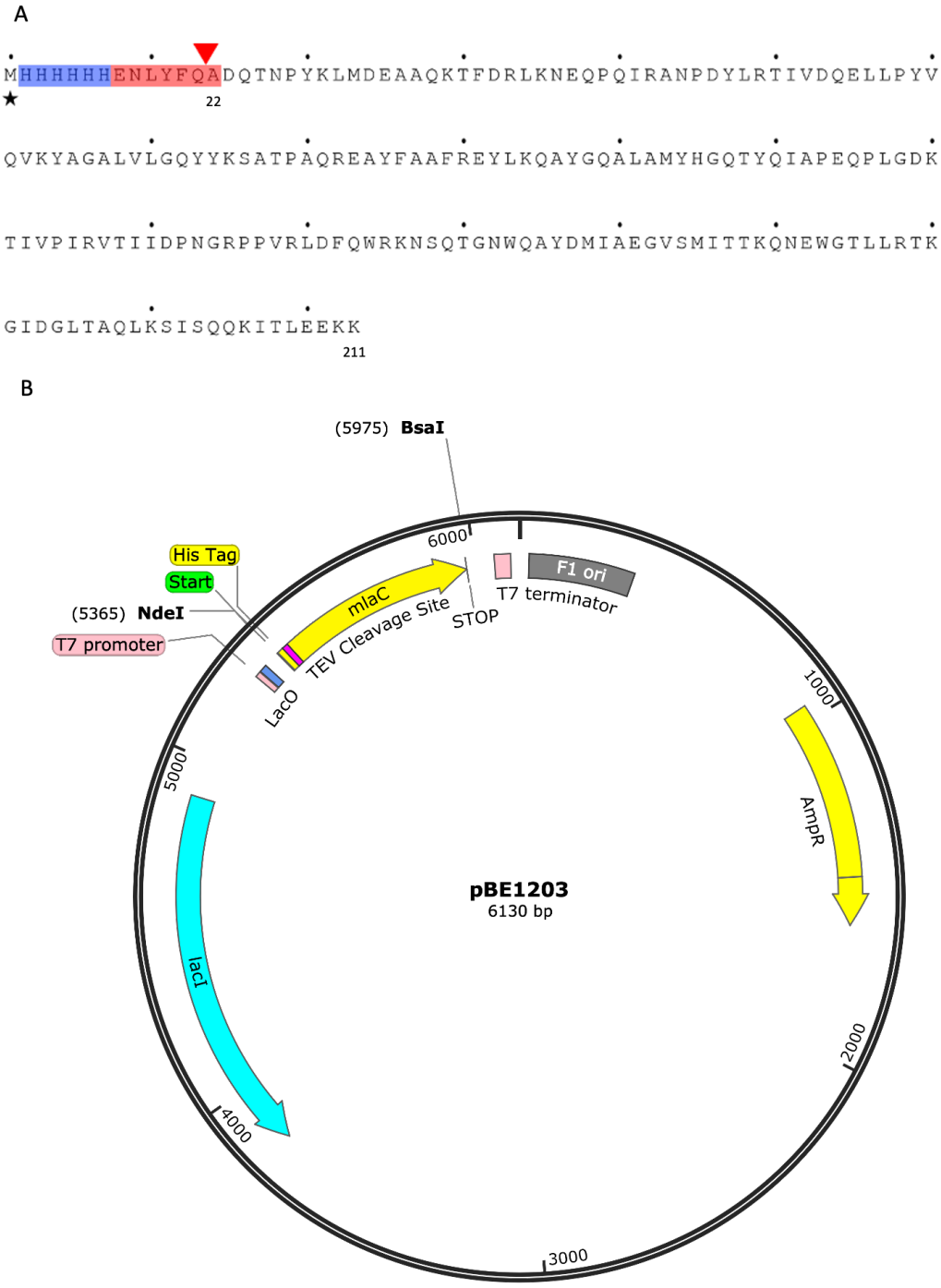

Figure 3.7 – (A) Sequence of the MlaC construct used in this investigation. The leading Met is marked with a star, the hexa-histidine tag is highlighted blue and the TEV recognition sequence is highlighted in red. The TEV cleavage site is marked with a red arrow. The first and last MlaC residue is numbered in accordance with [Figure 1.19]. (B) Plasmid map of pBE1203, with major features labelled and relevant restriction enzyme sites marked.

3.3.2.Construct Design of MlaD

We designed a construct of MlaD for cytoplasmic expression as a soluble protein and purification by Ni-affinity chromatography [Figure 3.8]. It lacks the 31 residue α-helical transmembrane domain that would ordinarily tether it to the membrane and instead replaces it with a hexa-histidine tag. Removal of the transmembrane domain as well as cytoplasmic purification has been shown in previous investigations to have no significant effect on the folding and assembly of the MlaD hexamer.

A

```
.         .         .         .         .         .         .
MHHHHHHTSIRTEPTYTLYATFDNIGGLKARSPVSIGGVVVGRVADITLDPKTYLPRVTL
★       32

.         .         .         .         .         .         .
EIEQRYNHIPDTSSLSIRTSGLLGEQYLALNVGFEDPELGTAILKDGDTIQDTKSAMVLE

.         .         .         .
DLIGQFLYGSKGDDNKNSGDAPAAAPGNNETTEPVGTTK
                                        183
```

B

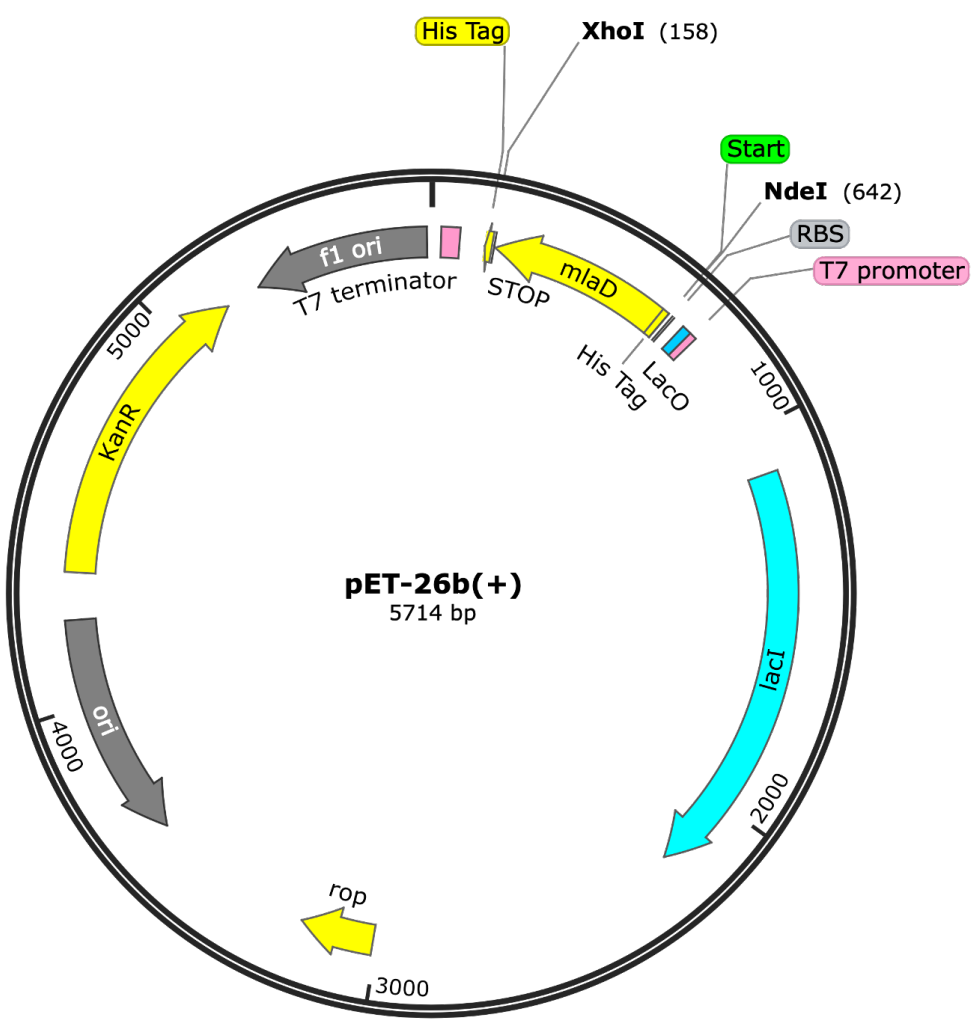

Figure 3.8 – (A) Sequence of the MlaD construct used in this investigation. The leading Met is marked with a star and the hexa-histidine tag is highlighted blue. The first and last MlaD residue is numbered in accordance with [Figure 1.24]. (B) Plasmid map of pET-26b with MlaD inserted between the XhoI and NdeI. Major features are labelled.

3.3.3.Expression and Purification of MlaC

MlaC was purified by nickel affinity and subsequent size exclusion chromatography. Figure 3.9 shows SDS-PAGE of the Ni-affinity purification [Figure 3.9A] showing the presence of a major band at approximately 23 kDa with relatively minor contaminants, corresponding to the expected molecular weight of MlaC with a 6xHis-tag and a TEV cleavage site. Size exclusion chromatography of the pooled Ni-affinity fractions resulted in a single major peak [Figure 3.9D]. The peak elution volume was consistent with a monomeric protein based on calibration of the column. SDS-PAGE of the major SEC peak showed no major contaminants [Figure 3.9E].

TEV cleavage and subsequent re-purification resulted in a decrease in Mw of approximately 1 kDa, consistent with the loss of the 6xHis-tag [Figure 3.9B]. To identify bound phospholipid species, TLC of chloroform-methanol extracted lipids was performed [Figure 3.9C]. MlaC was bound with lipids with preference for PG, which is consistent with the lipid profile presented in the literature and suggests a correctly folded protein. Preparations including an extra detergent wash phase during the Ni-affinity purification stage resulted in a delipidated preparation, henceforth referred to as MlaC-apo.

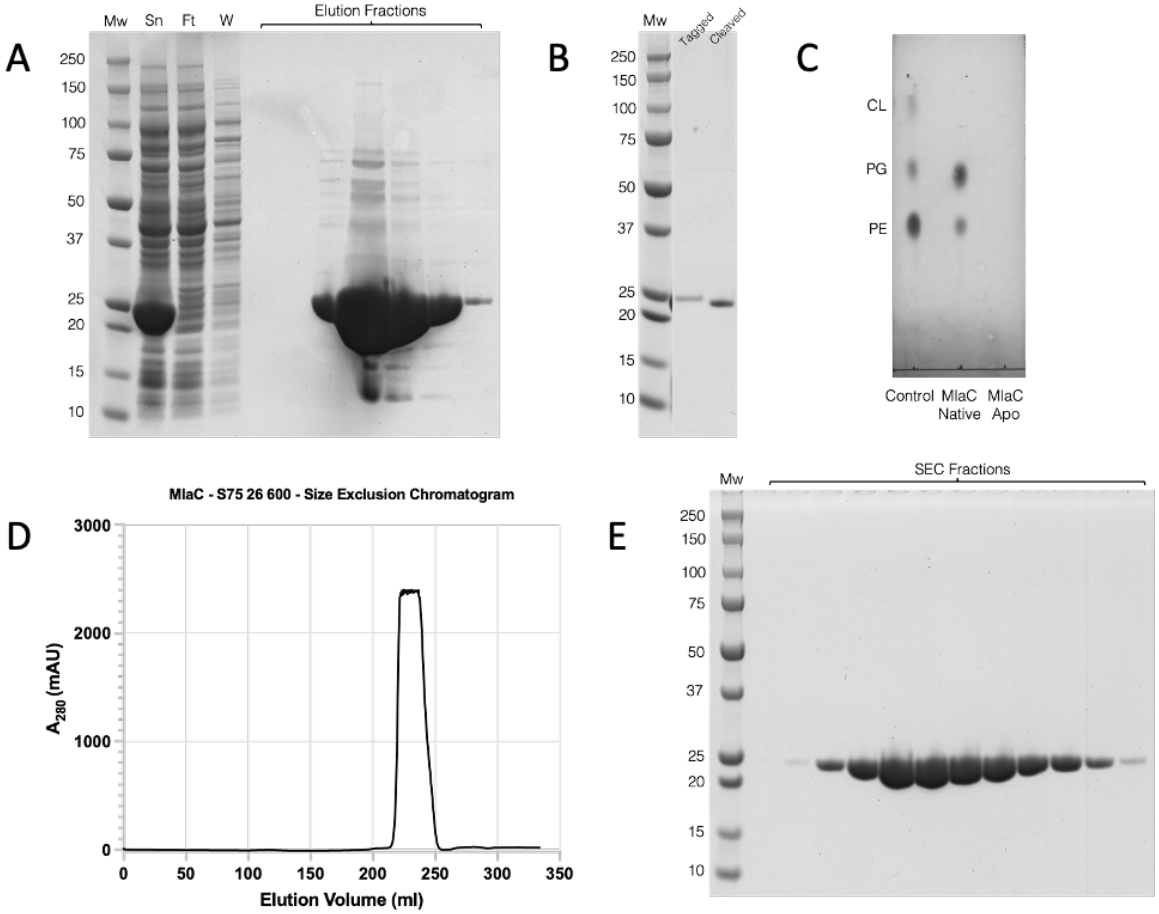

Figure 3.9 - (A) SDS-PAGE of MlaC Ni-affinity purification fractions. Samples from the pre-column supernatant (Sn), the column flowthrough (Ft) and a 20 mM imidazole wash (W) were run alongside fractions eluted with 500mM imidazole. A major band with molecular weight between 20 and 25kDa is present in the Sn but absent in the Ft and W. This band is also the major represented band in the elution fractions, however, multiple contaminant bands are visible. (B) SDS-PAGE of MlaC before (Tagged) and after (Cleaved) treatment with TEV protease. The cleaved band runs slightly below the tagged band suggesting successful cleavage of the hexa-histidine tag. (C) TLC of chloroform-methanol extracts from both a natively purified MlaC as well as MlaC purified with an additional detergent wash step to generate an -apo protein. An *E. coli* polar lipid control was run alongside, with spots associated with cardiolipin (CL), phosphatidylglycerol (PG) and phosphatidylethanolamine (PE) marked. PG and PE are clearly present in the native sample but not in the -apo sample. (D) SEC trace of pooled MlaC elution fractions showing a single elution peak. (E) SDS-PAGE of SEC elution fractions show a single band between 20 and 25kDa with no major contaminants.

3.3.4.Native Mass Spectrometry of MlaC

To ensure that the -apo samples of MlaC were sufficiently delipidated and did not retain any detergent from the lipid washing procedure we conducted native mass spectrometry of the apo and natively purified samples [Figure 3.10]. The major m/z peaks correspond to the 8+, 9+ and 10+ ions in both the native and apo samples. The major peaks in the apo sample correspond to a protein of 21731 Da with minor species of less than 3% abundance corresponding to a bound lipid of approximately 750 Da. Comparatively, the native sample shows a broad distribution of peaks corresponding to MlaC in complex with multiple lipids. There is evidence of a minor population of apo MlaC in the native sample. Mass spectrometry as well as data processing was conducted by Dr. Aneika Leney. This data was submitted for publication as part of the supplementary material of Hughes et *al.* (2019)[71].

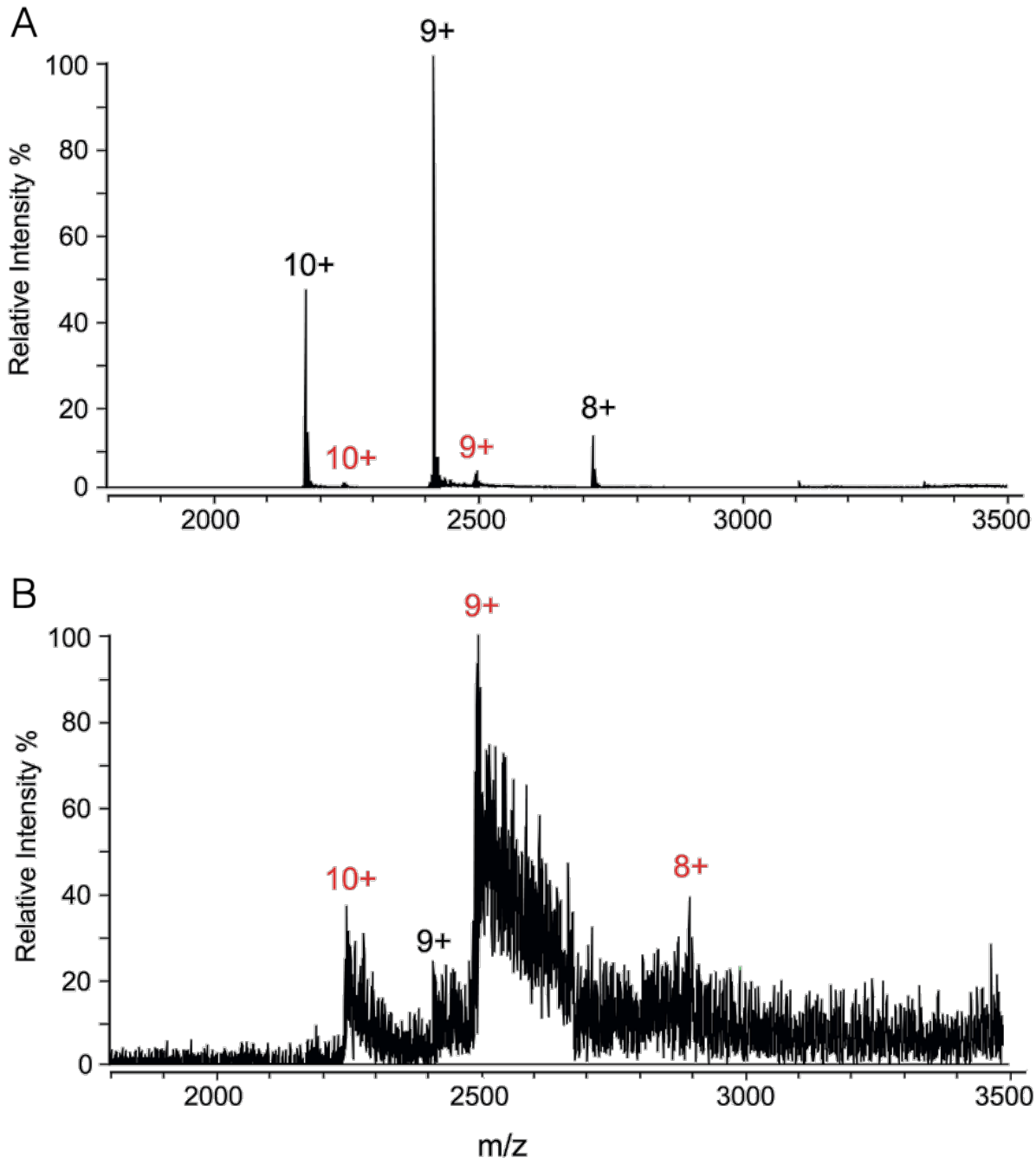

Figure 3.10 - Native mass spectrometry data of MlaC-apo (a) and MlaC-native (b). The spectra of MlaC-apo shows m/z peaks associated with major charge state distributions of a 21731 Da protein (black) as well as a protein approximately 750 Da larger (red), which we attribute to the binding of a lipid. The spectra of MlaC-native shows a large distribution of peaks corresponding to a protein bound to various lipids. The peaks associated with a potentially -apo population (black) as well as those of the protein bound to the 750 Da lipid (red) are labelled.

3.3.5.Expression and Purification of MlaD-Native and MlaD-Apo

Initial Ni-NTA purification of a soluble construct of MlaD, referred to henceforth as MlaDΔTM, resulted in a defined single band that runs at approximately 19 kDa when boiled at 95 °C for 10 mins [Figure 3.11A]. Minor contaminant bands at approximately 15 KDa are present in the Ni-affinity sample and persist through size exclusion. The SDS resistance of the protein is exemplified in samples of the SEC purified protein that have not been boiled prior to SDS-PAGE, which run as though they were a 90 kDa protein [Figure 3.11C]. Although this is not exactly consistent with the expected Mw of approximately 102 kDa for a hexameric complex, it is generally accepted that partial folding of a protein sample can affect its migration during electrophoresis. Regardless, it is clear that the protein purifies in a multimeric state. This is further supported by the elution volume of the sample on an S200 16 60 column, wherein the sample elutes in a single peak at a volume consistent with a protein of Mw between 100 and 200 kDa. Taking this into consideration, it is quite possible that the minor contaminant bands at Mws above the predominant 19 kDa band that persist through size exclusion are partially folded multimeric states rather than other contaminating proteins. Contaminant bands below the predominant 19 kDa band may be attributed to hexameric complexes with degraded monomers being below the resolving capacity of the S200 column to separate. To determine if this was the case, we conducted trypsin digest mass spectrometry on extracted gel bands. Trypsin digest mass spectrometry was carried out by the University of Birmingham Advanced Mass Spectrometry Facility. Both the smaller [Figure 3.12] and larger bands [Figure 3.13] returned a protein with accession number P64604 corresponding to MlaD from *E. coli* [3].

Purification of native MlaDΔTM resulted in lipids bound suggesting a correctly folded and assembled hexamer capable of coordinating lipid substrates [Figure 3.11D].

While there is not a clear preference for PG lipids as with MlaC, the lipids bound to MlaDΔTM are not entirely reflective of the *E. coli* polar lipid control as would be expected if MlaDΔTM had no preference. Likewise, with MlaC, preparations including an extra detergent wash phase during the Ni-affinity purification stage resulted in a delipidated preparation, henceforth called MlaDΔTM-apo.

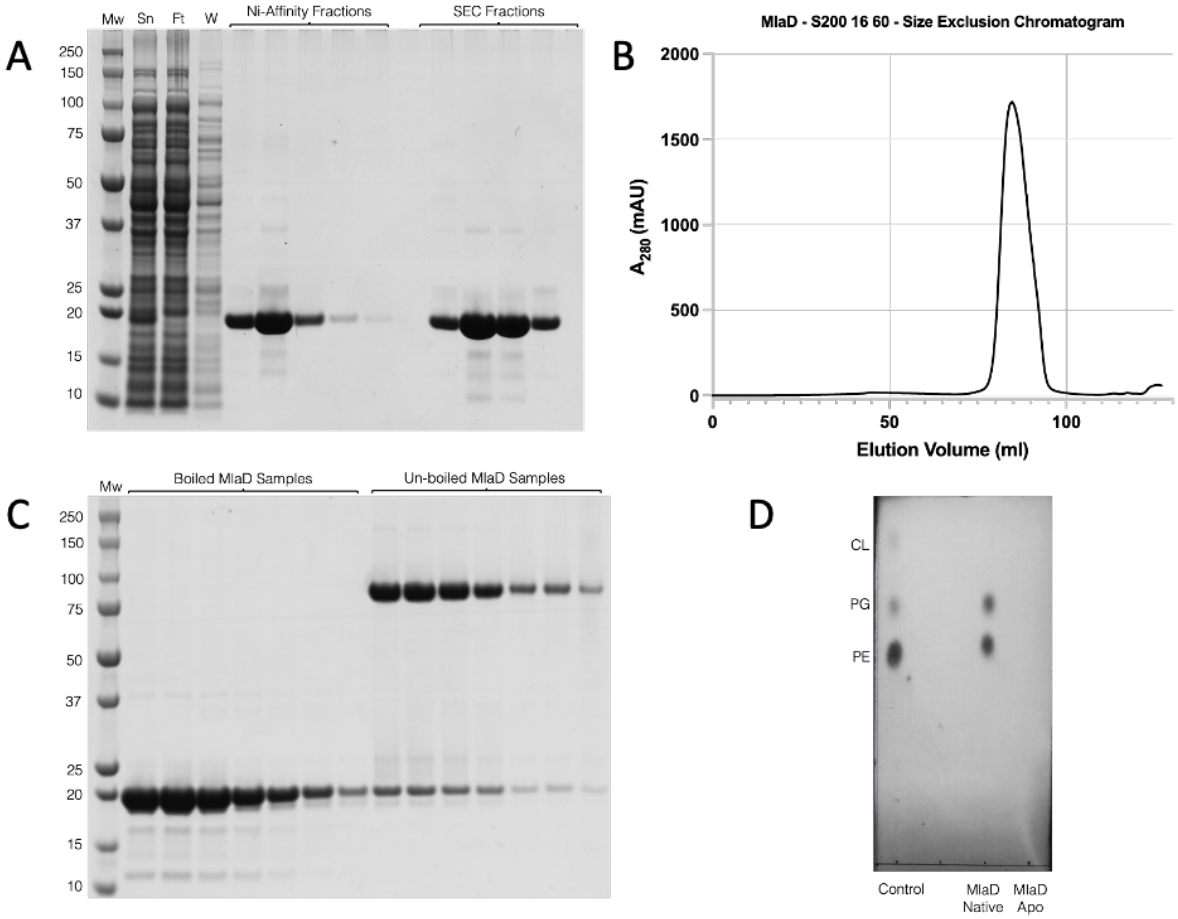

Figure 3.11 - (A) SDS-PAGE of MlaD Ni-affinity purification fractions and SEC elution fractions. Samples from the pre-column supernatant (Sn), the column flowthrough (Ft) and a 50 mM imidazole wash (W) were run alongside fractions eluted with 500 mM imidazole. A band with molecular weight of approximately 20 kDa is present in the Sn but absent in the Ft and W. This band is also the major represented band in the elution fractions. The same band is present in the SEC elution fractions. There are very faint contaminant bands present in the SEC fractions. (B) SEC trace of pooled MlaD elution fractions showing a single elution peak. (C) SDS-PAGE of MlaD samples boiled at 95°C compared to those run without boiling. Un-boiled, the protein runs with a molecular weight of between 75 and 100 KDa. (D) TLC of chloroform-methanol extracts from both a natively purified MlaD as well as MlaD purified with an additional detergent wash step to generate an -apo protein. An *E. coli* polar lipid control was run alongside, with spots associated with cardiolipin (CL), phosphatidylglycerol (PG) and phosphatidylethanolamine (PE) marked. PG and PE are clearly present in the native sample but not in the -apo sample.

Protein Accessions	Sequence	Contaminant	Modifications	Theo. MH+ [Da]
P64604	TSGLLGEQYLALNVGFEDPELGTAILKDGDTIQDTK	FALSE		3821.93312
P64604	TSGLLGEQYLALNVGFEDPELGTAILK	FALSE		2848.49787
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE	1xDeamidated [N]	2729.21326
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE		2728.22925
P64604	NSGDAPAAAPGNNETTEPVGTTK	FALSE	1xDeamidated [N]	2200.00002
P64604	SAMVLEDLIGQFLYGSK	FALSE	1xOxidation [M3]	1886.95683
P64604	TSGLLGEQYLALNVGFEDPELGTAILKDGDTIQDTK	FALSE	1xDeamidated [Q/N]	3822.91714
P64604	SAMVLEDLIGQFLYGSK	FALSE		1870.96191
P64604	NSGDAPAAAPGNNETTEPVGTTK	FALSE	2xDeamidated [N1; N12]	2200.98404
P64604	TEPTYTLYATFDNIGGLKAR	FALSE		2231.13427
P64604	TEPTYTLYATFDNIGGLK	FALSE		2003.99605
P64604	NSGDAPAAAPGNNETTEPVGTTK	FALSE		2199.01601
P64604	YNHIPDTSSLSIR	FALSE		1502.75978
P64604	SAMVLEDLIGQFLYGSK	FALSE	1xDeamidated [Q11]	1871.94593
P64604	SPVSIGGVVVGR	FALSE		1126.65788
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE	3xDeamidated [N6; N17; N]	2731.1813
P64604	VTLEIEQR	FALSE		987.54694

Figure 3.12 - Trypsin digest mass spectrometry sequence results for the 15 KDa band extracted from an SEC fraction lane from the gel in [Figure 3.11A]. The sequences correspond to *E. coli* MlaD.

Protein Accessions	Sequence	Contaminant	Modifications	Theo. MH+ [Da]
P64604	TSGLLGEQYLALNVGFEDPELGTAILKDGDTIQDTK	FALSE		3821.93312
P64604	TSGLLGEQYLALNVGFEDPELGTAILK	FALSE		2848.49787
P64604	NSGDAPAAAPGNNETTEPVGTTK	FALSE	1xDeamidated [N]	2200.00002
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE	2xDeamidated [N6; N17]	2730.19728
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE		2728.22925
P64604	GDDNKNSGDAPAAAPGNNETTEPVGTTK	FALSE	1xDeamidated [N]	2729.21326
P64604	SAMVLEDLIGQFLYGSK	FALSE	1xOxidation [M3]	1886.95683
P64604	NSGDAPAAAPGNNETTEPVGTTK	FALSE		2199.01601
P64604	SAMVLEDLIGQFLYGSK	FALSE		1870.96191
P64604	TEPTYTLYATFDNIGGLK	FALSE		2003.99605
P64604	TEPTYTLYATFDNIGGLKAR	FALSE		2231.13427
P64604	SAMVLEDLIGQFLYGSK	FALSE	1xDeamidated [Q11]	1871.94593
P64604	YNHIPDTSSLSIR	FALSE		1502.75978
P64604	ARSPVSIGGVVVGR	FALSE		1353.79611
P64604	SAMVLEDLIGQFLYGSK	FALSE	1xDeamidated [Q11]; 1xOxidation [M3]	1887.94084
P64604	SPVSIGGVVVGR	FALSE		1126.65788
P64604	DGDTIQDTK	FALSE		992.45309
P64604	YNHIPDTSSLSIR	FALSE	1xDeamidated [N2]	1503.7438
P64604	VADITLDPK	FALSE		971.54079
P64604	VTLEIEQR	FALSE		987.54694

Figure 3.13 - Trypsin digest mass spectrometry sequence results for the faint 37 KDa band extracted from an SEC fraction lane from the gel in [Figure 3.11A]. The sequences correspond to *E. coli* MlaD.

3.3.6.Characterisation of MlaC and MlaD Functionality

To determine if the functionality of purified MlaDΔTM and MlaC was consistent with previous publications, exchange of lipids between MlaDΔTM and MlaC was analysed. Natively purified MlaDΔTM was incubated with MlaC-apo for 30 mins on ice, before subsequent separation by SEC [Figure 3.14A]. Lipids were extracted from elution fractions containing exclusively either MlaDΔTM or MlaC as determined by SDS-PAGE [Figure 3.14B]. TLC results suggest that nearly all lipids initially coordinated by the MlaDΔTM sample were exchanged to the sample of MlaC-apo and this is consistent with the expected direction of lipids exchange for these soluble constructs *in vitro* [Figure 3.14C]. It is worth noting that the strong preference of MlaC for PG lipids is not observed in the sample receiving lipids from MlaDΔTM, but instead reflect the lipid profile of MlaDΔTM as was also observed by Hughes et *al.*[71].

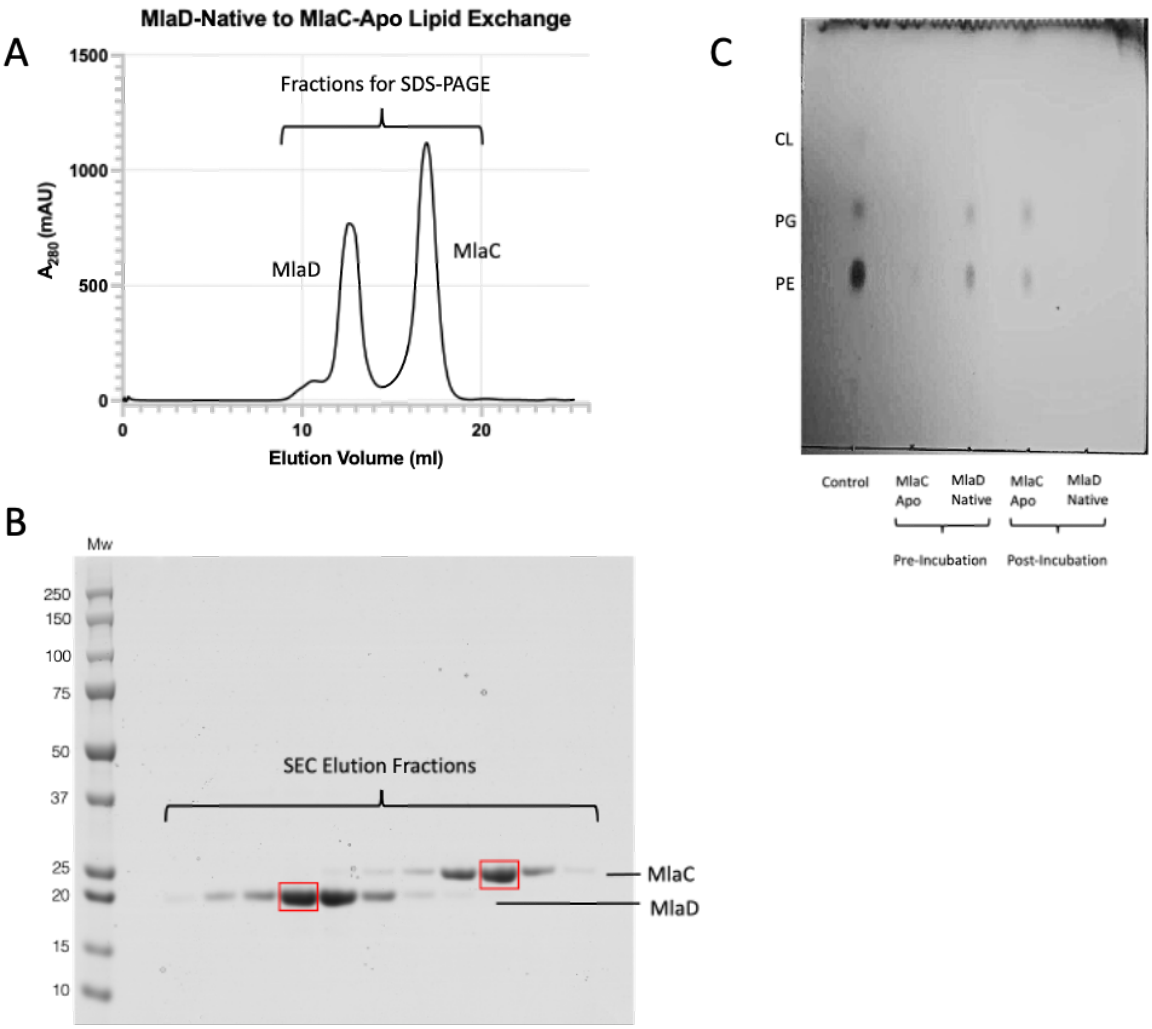

Figure 3.14 - (A) SEC trace showing the separation of MlaC from MlaD post incubation of MlaC-apo with MlaD-native. The SDS-PAGE run of the elution fractions (B) show sufficient separation of the two proteins. (C) shows the TLC of lipids extracted from the fractions highlighted red in (B). Pre-incubation MlaC-apo shows no evidence of bound lipids, however, after incubation with MlaD-native, all lipids appear to have been removed from MlaD-native and transferred to MlaC-apo.

3.3.7.Lipid Loading of Solubilized Samples

It has previously been established by NMR that spontaneous lipid uptake by MlaC-apo occurs at a negligible rate, if at all. To determine if this was also the case for MlaDΔTM we performed on column incubation of MlaDΔTM with various lipids and subsequently assessed lipid binding with chloroform:methanol extraction and TLC. We determined that incubation of MlaDΔTM-apo with single species lipids resulted in effective lipid uptake [Figure 3.15]. MlaDΔTM was capable of accepting both liposomal POPG and POPE. The validation of POPG and POPE uptake by MlaDΔTM has significance to the aims of this investigation as it has previously been shown that MlaC will readily accept lipids from MlaDΔTM. To ensure that MlaC was capable of accepting single species of lipids from MlaDΔTM, lipid exchange experiments as previously described with natively purified MlaDΔTM were conducted. We thus attempted to show that reflowing MlaC-apo over a large excess of column bound, lipid loaded MlaDΔTM would prove to be an effective method of generating preparative amounts of lipid loaded MlaC. Validation of POPG and POPE exchange between MlaDΔTM and MlaC-apo was shown by choloform:methanol extraction and TLC. MlaC-apo was capable of accepting both POPG and POPE from MlaDΔTM [Figure 3.15]. While TLC is not a technique appropriate for quantitative assessment it would appear that POPG is much more readily accepted by MlaDΔTM than POPE.

We also made attempts to load preparative amounts of MlaC with BDO-CL at this stage, however, we encountered issues during our attempts to separate MlaC from MlaDΔTM after incubation in the presence of BDO-CL. This issue will be addressed in Chapter 4 as it is significant to the methodologies discussed in that chapter.

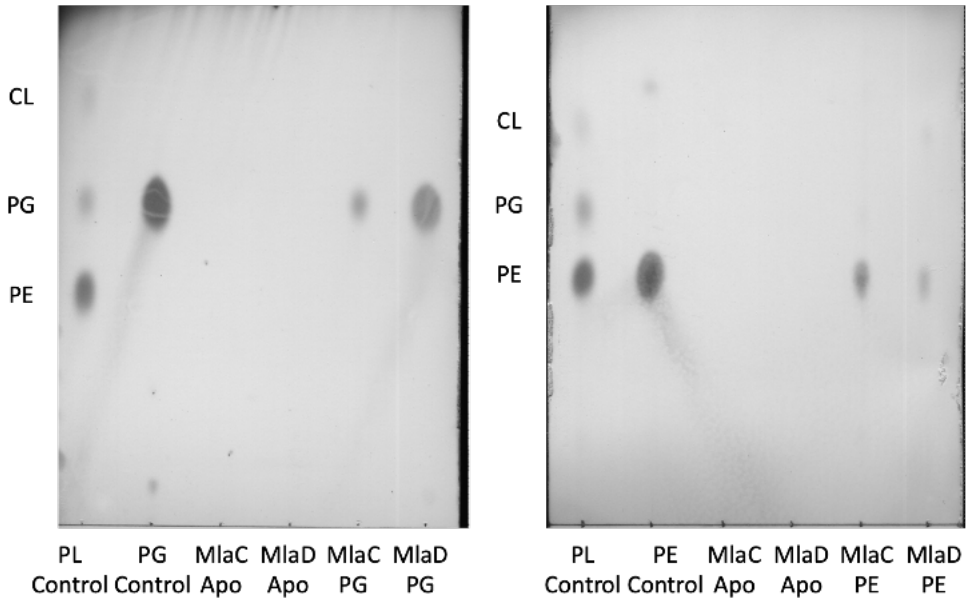

Figure 3.15 - TLCs of Mla proteins loaded with single species lipids alongside relevant controls. MlaD-PG and MlaD-PE were prepared by reflux of POPG/POPE liposomes over column bound MlaD-apo. MlaC-PG and MlaC-PE were prepared by reflux of MlaC-apo over an excess of column bound MlaD bound with POPG/POPE. Chloroform-methanol extracts of the lipid loaded proteins clearly show the presence of bound lipids. While TLC is not a quantitative technique it is worth noting that loading MlaD with POPE did not appear to be as successful as the loading of MlaD with POPG.

3.3.8.Native Mass Spectrometry of Lipid Loaded Proteins

We conducted native mass spectrometry with the intent of providing supporting evidence that removal of lipids from MlaDΔTM with β-OG does not result in β-OG binding to MlaDΔTM. [Figure 3.16A] shows m/z peaks associated with the 102951 Da MlaDΔTM hexamer as well as peaks associated with a 178 Da modification. [Figure 3.16B] shows m/z peaks of the monomer and monomers with a similar 178 Da modification. The Mw of this modification does not correspond with the Mw of β-OG (292.37 Da) or with the Mw of any cellular lipid. It was determined that the modification is associated with the non-enzymatic modification of the His-tag with the cellular metabolite 6-phosphoglucono-1,5-lactone followed by enzymatic activity of host cell phosphatases to produce a gluconoylated His-tag. Geoghegan et *al.* demonstrated that certain flanking residues may result in susceptibility of His-tags to gluconoylation resulting in a +178 or +258 Da modification[60]. Despite this modification, no evidence of residual β-OG is present bound to the hexamer. Loading of MlaDΔTM with POPG results primarily in m/z peaks associated with 3 and 4 bound POPG molecules. [Figure 3.17] shows m/z shifts of +2308 and +3056 in the POPG loaded samples corresponding approximately to the Mw of 3 and 4 molecules of POPG. This data suggests that the $POPG_4$ state is preferred.

We also verified that lipid loading of MlaC was successful. [Figure 3.18] shows m/z shifts associated with a POPE loaded MlaC sample. The m/z shift from the -apo state corresponded to a 718 Da change in mass, which is consistent with the Mw of POPE. We were able to confirm that loading of MlaC resulted in binding of a single POPE molecule, suggesting that the presence of POPE in TLC results was not the caused by erroneous POPE binding or transient aggregation to the protein. Native mass

spectrometry as well as data processing was carried out by Dr. Jeddidiah Bellamy-Carter.

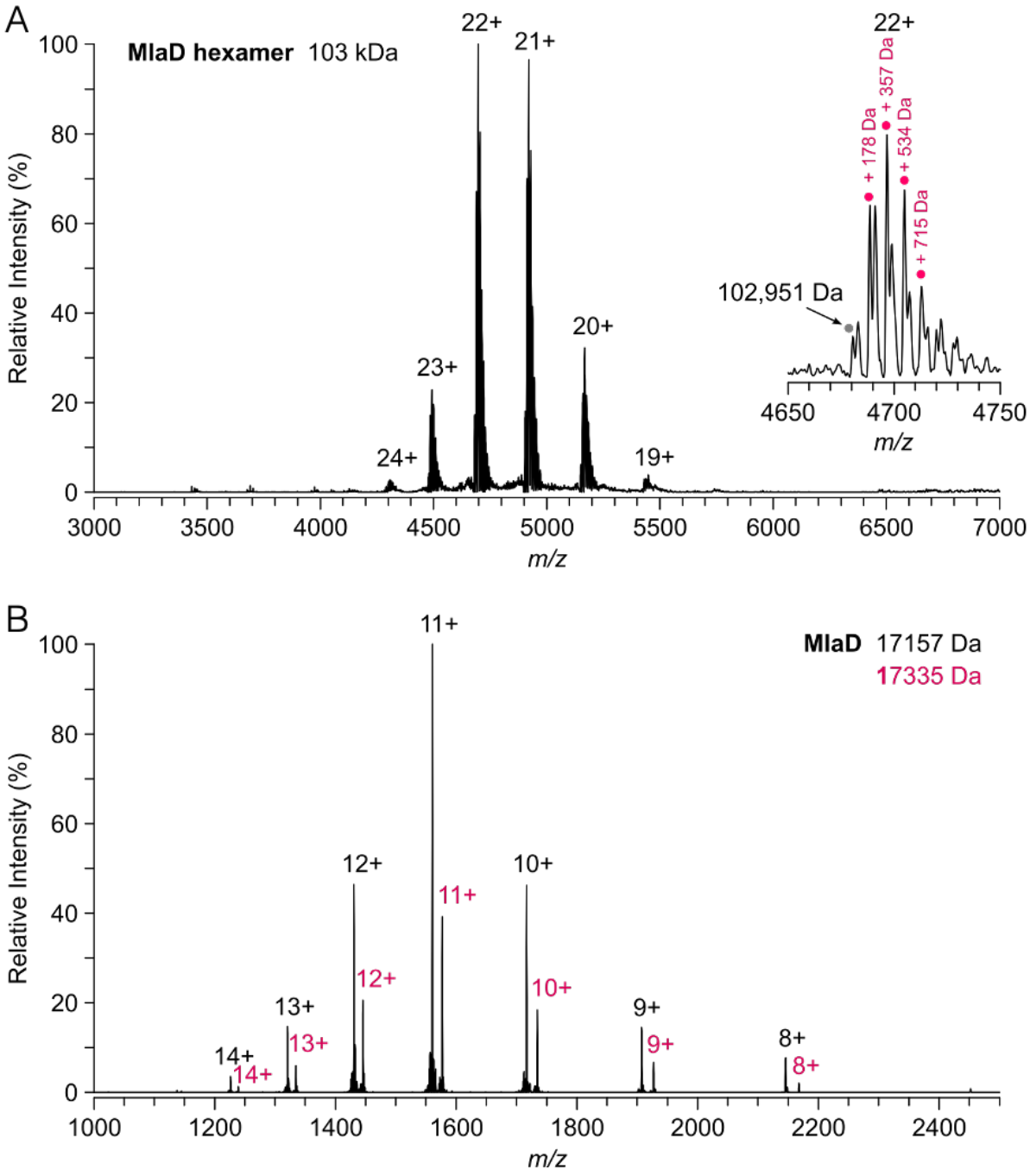

Figure 3.16 - Native mass spectrometry of MlaDΔTM-apo. (A) Shows the MlaDΔTM hexamer with labelled charge states. A zoomed in view of the 22+ charge state has been included to show peaks associated with the hexamer as well as peaks associated with successive addition of a +178 Da modification. This modification has been determined to be the result of gluconoylation of the His-tag. (B) The dissociated monomer units with labelled charge states, showing a mixed sample of gluconoylated and un-gluconoylated sample.

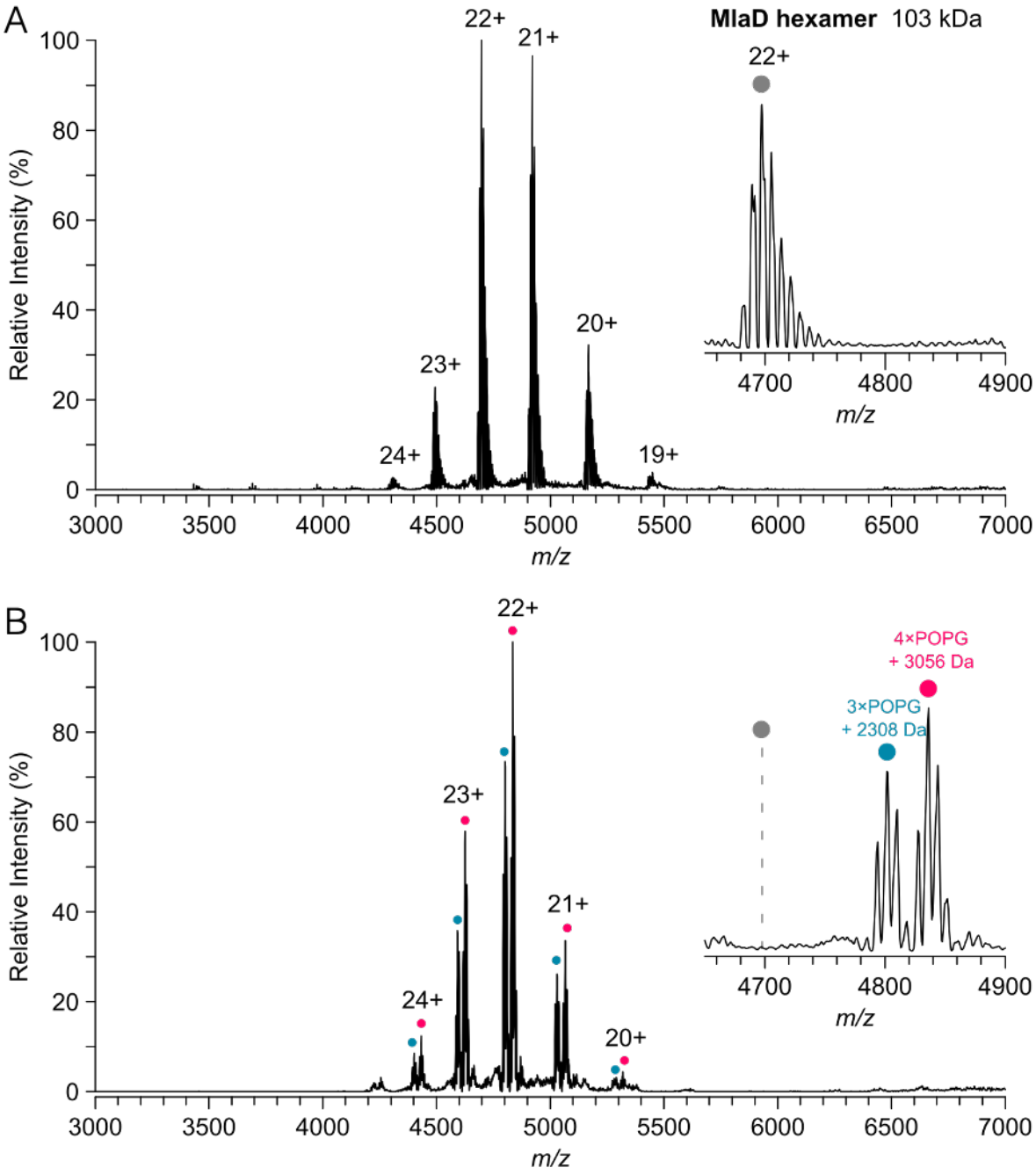

Figure 3.17 - Comparison of native mass spectrometry of (A) MlaΔTM-apo and (B) MlaDΔTM-POPG. There is a difference in m/z peaks between the two samples corresponding to a change of +2308 and +3056 Da, suggesting the loading of 3 and 4 POPG molecules respectively.

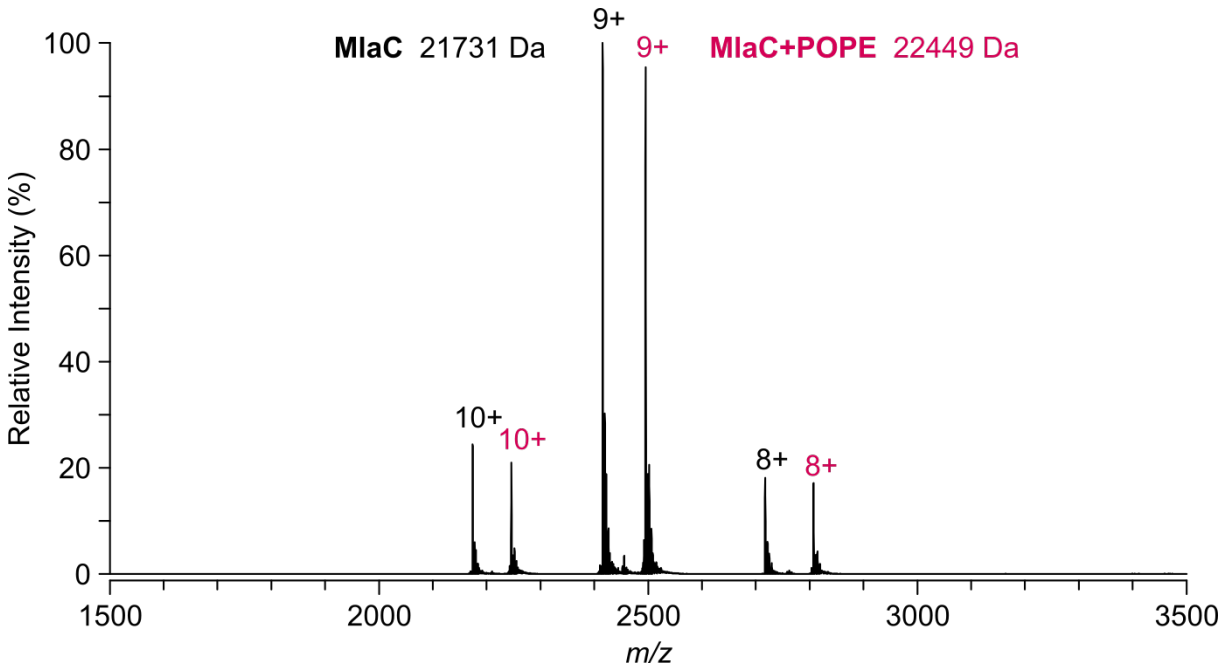

Figure 3.18 - Native mass spectrometry of MlaC-POPE, showing a m/z shift of +718 Da from the -apo state. This is consistent with the binding of 1 POPE molecule.

3.3.9.Determining the Structure of MlaC Loaded With a Single Species Lipid

Having convincingly determined that MlaC can be loaded with POPG and POPE we moved to screening crystallization conditions in an attempt to identify the specifics of the interaction between MlaC and the lipids. We hoped that we would be able to determine the specificities of the head group interaction. We were, however, unsuccessful in generating diffraction quality crystals of POPE loaded MlaC. We were, however, successful in crystallising POPG loaded MlaC. POPG loaded MlaC was crystallized from a sample at 45 mg/ml in a 4 μL sitting drop with a 1:1 mixture of protein sample to precipitant solution [Figure 3.19A]. The precipitant solution contained 2.2 M ammonium sulphate and 20% (w/v) glycerol. Crystals formed in these conditions were determined to be inappropriate for diffraction and instead used for streak seeding. Streak seeded POPG loaded MlaC was crystallized from a sample at 45 mg/ml in a 4 μL hanging drop with a 1:1 mixture of protein sample to precipitant

solution [Figure 3.19B]. The precipitant solution contained 2.0 M ammonium sulphate and 20% (w/v) glycerol. The entirety of the streak seeded drop crystallised into a single crystal. The crystal was fragmented and a segment was fished.

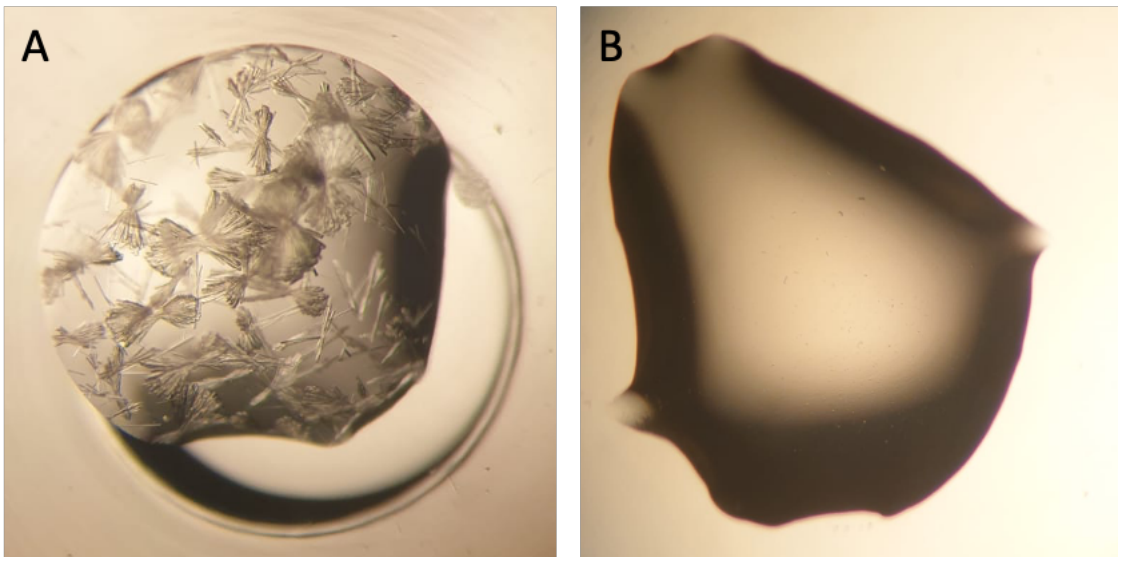

Figure 3.19 - (A) POPG loaded MlaC crystals formed in 2.2 M ammonium sulphate and 20% (w/v) glycerol. The crystals formed into bundled needles and were determined to be too fine to diffract well. Streak seeding using crystals from (A) into a 1:1 mixture of protein and precipitant solution crystallized in 2.0 M ammonium sulphate and 20% (w/v) glycerol. The entire drop (B) crystallised and was fragmented during fishing.

Data collection revealed that the crystal belonged to the H-centred trigonal space group, H3 and Matthews coefficient probabilities suggested a single copy of the protein in the asymmetric unit. The crystal structure of MlaC with POPG bound was determined at 1.93 Å resolution using chain A of the natively purified *E. coli* MlaC structure solved by Ekiert et *al.* (5UWA) with the lipid removed as a molecular replacement model. Experimental phasing produced a solution with a Log Likelihood Gain (LLG) of 758.3 and TF Z-score (TFZ) of 22.8. POPG was built into the density visible in the binding pocket during refinement. Refinement of the model to a final R_{work}/R_{free} of 0.2152/0.2632 resulted in clearly defined density for the POPG lipid in the binding site allowing for unambiguous positioning of the lipid head group [Figure 3.20, Figure 3.21, Table 3.1].

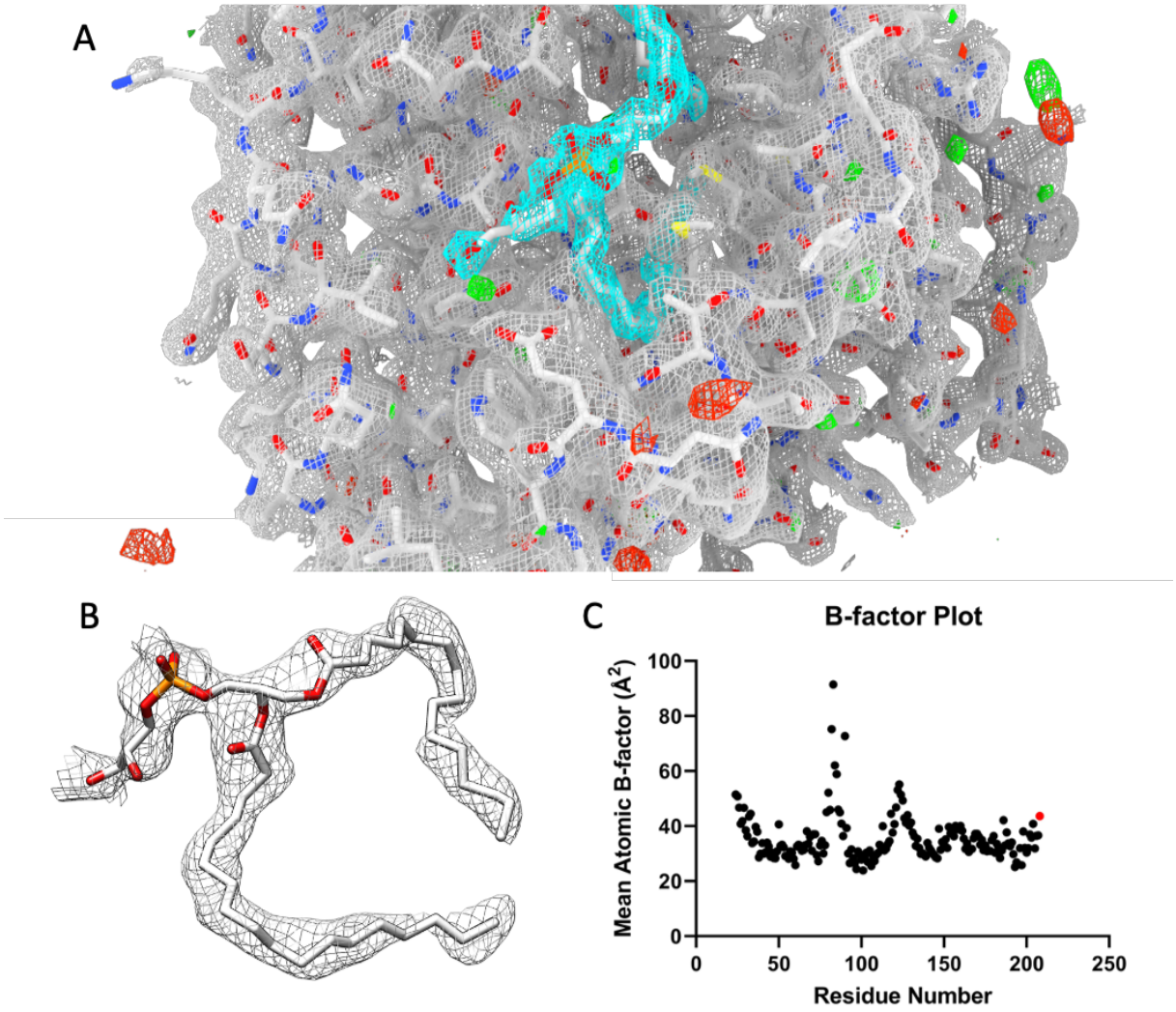

Figure 3.20 – (A) Electron density map determined from reflection data of crystallised POPG loaded MlaC. The 2Fo-Fc map (grey and blue) has contour level set at 1σ and Fo-Fc map (green and red) has contour level set at 3σ. Density associated with the lipid has been coloured blue for clarity. There is no negative density in the Fo-Fc map is associated with the POPG lipid. (B) 2Fo-Fc map showing only lipid. (C) B-factor plot by residue, POPG is plotted in red. The majority of B-iso values fall within the range to suggest that atomic positions are well defined.

The conformational state of the POPG bound protein matches well to the existing native structure, with less than 0.08 Å RMSD between the Cα trace for all residues aside from 6 on the β2 – β3 loop and 6 on the β3 – β4 loop, which exhibit minor deviations [Figure 3.22]. As such the POPG bound structure exhibits the same conformational differences to the apo state that the native protein exhibits.

The major difference between the POPG bound structure and the existing native model comes in the position of the lipid. Unlike previously modelled lipids, which were positioned based on the average density of a multitude of lipids, the density we observe suggests that POPG binds with one acyl chain in each of the two lipid binding pockets of MlaC [Figure 3.23]. Rather than both acyl chains sitting to one side of the Tyr105 that segments the binding site, our model suggests that the acyl tails sit to either side of the Tyr105 residue, resulting in that residue adopting a position closer to the -apo model [Figure 3.24]. It is worth acknowledging that the lipid modelled in the Ekiert structure is a di-palmitoyl lipid where the lipid we have docked has a C_{16} palmitoyl acyl tail and a C_{18} oleoyl acyl tail. It may be the case that MlaC accommodates longer lipid tails with the second frontal lipid binding pocket. Moreover, we have positional information for the lipid head, which appears to sit in a small cleft at the front of the protein rather than extending into the bulk solvent as was previously suggested in the structure presented by Kuzin et *al.*[90].

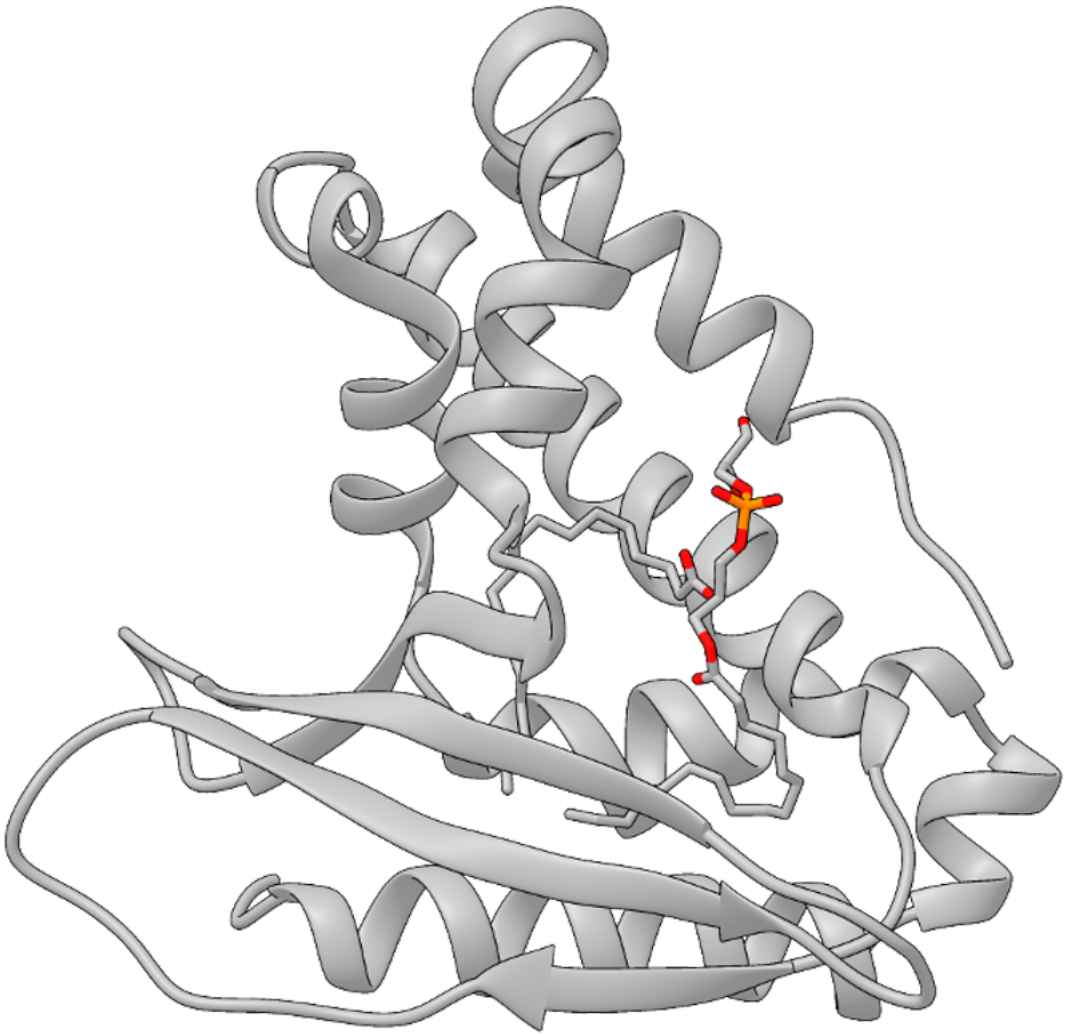

Figure 3.21 - Ribbon diagram of the final refined model of POPG bound MlaC showing lipid position.

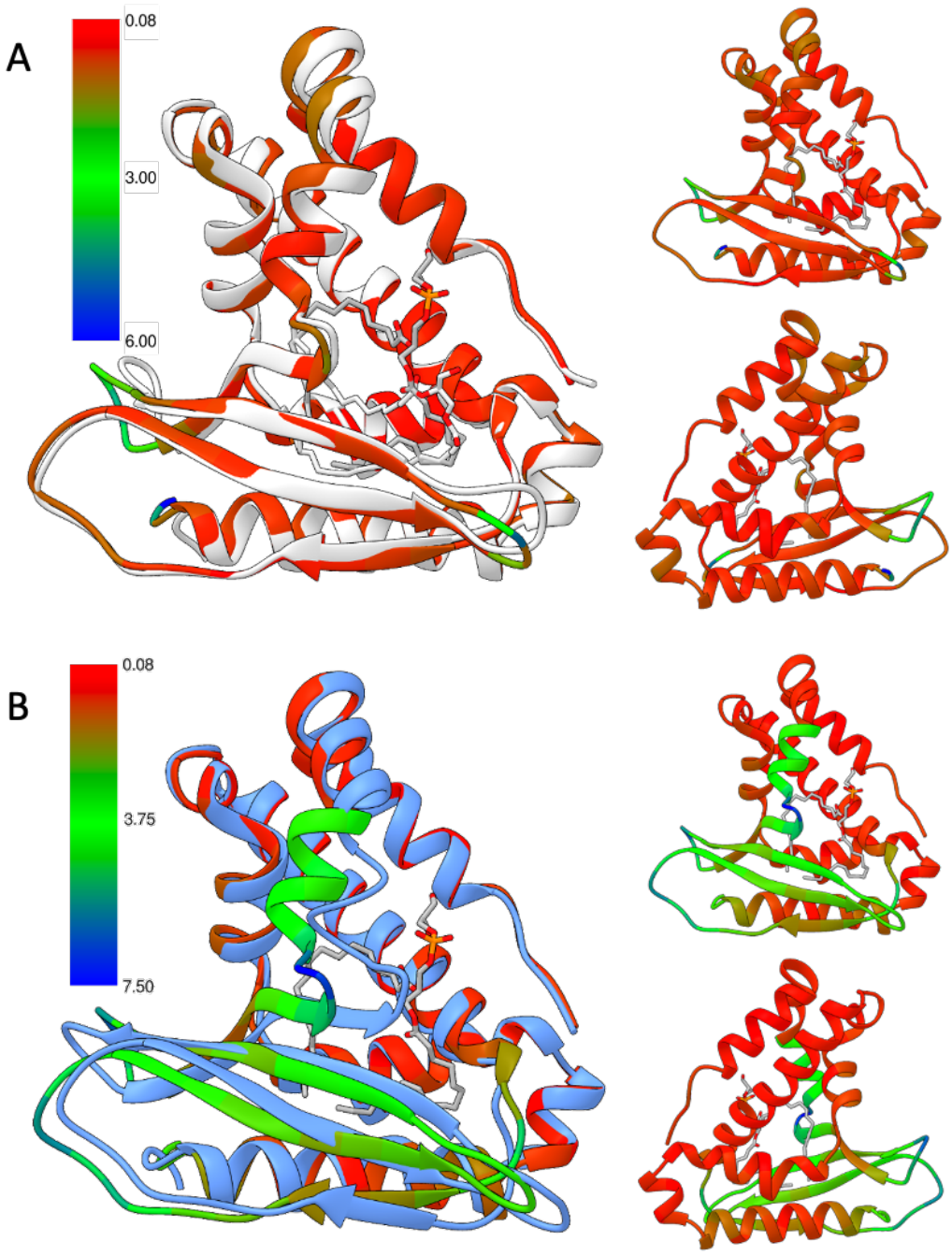

Figure 3.22 - The structure of POPG loaded MlaC is overlayed against (A) the 5UWA native structure of MlaC and (B) the 6GKI apo structure of MlaC. The structure of POPG loaded MlaC is coloured according to C-alpha RMSD from the structure it is overlayed against. The colour key is displayed on the far left. The POPG loaded MlaC structure has very little deviation from the native structure, with only minor deviations in the β2 – β3 and β3 – β4 loops (A). The POPG loaded MlaC deviated from the -apo structure in the same way that the natively purified MlaC does (B).

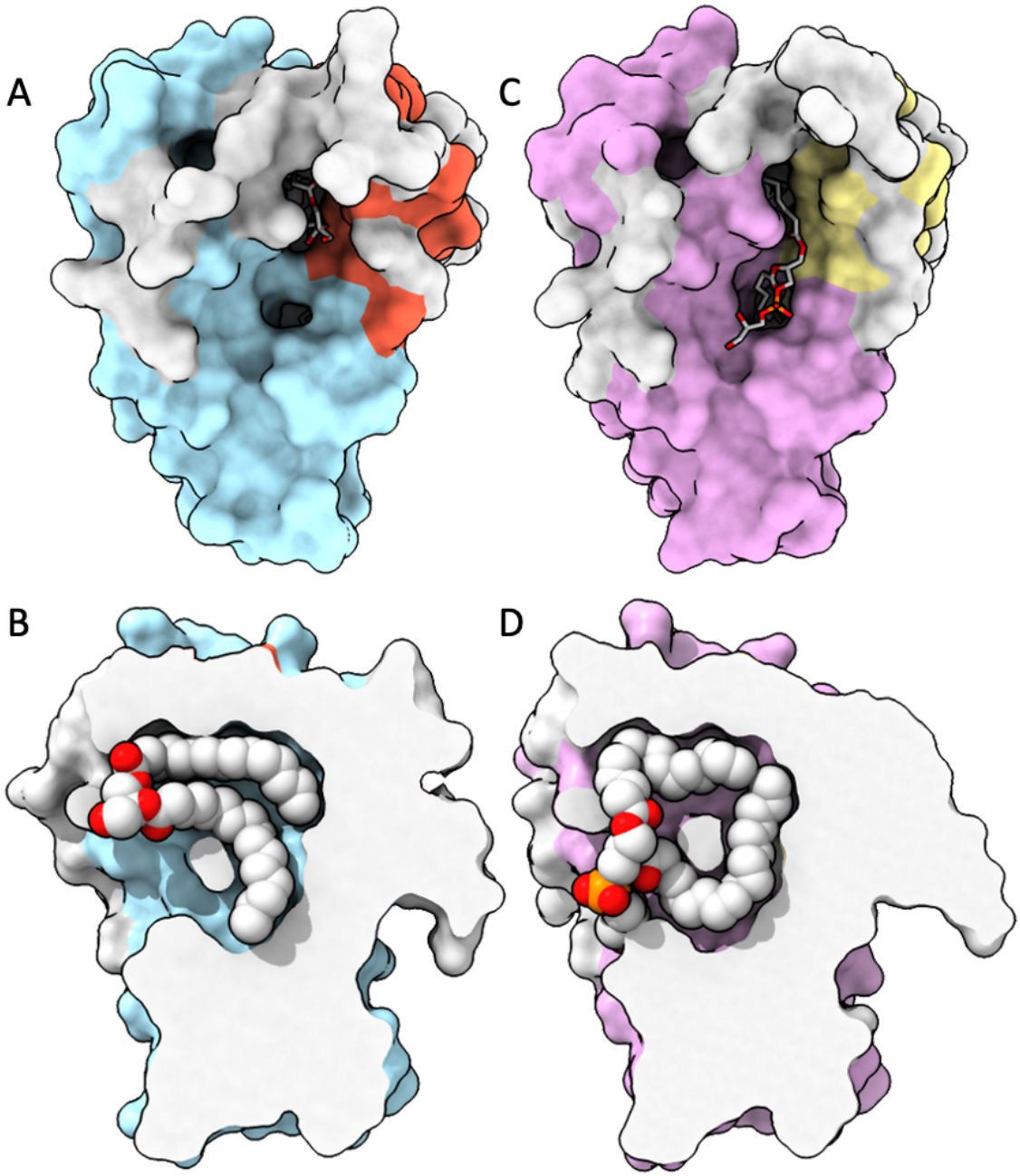

Figure 3.23 - Molecular surface model of the 5UWA structure of natively purified *E. coli* MlaC (A), alongside POPG loaded MlaC (C) . The two openings of the frontal binding pocket can be seen from the front as well as the side views presented below (B,D). The native structure suggests both lipid tails bind into a single binding pocket while the POPG bound structure we present suggests that the acyl tails bind into both frontal binding pockets.

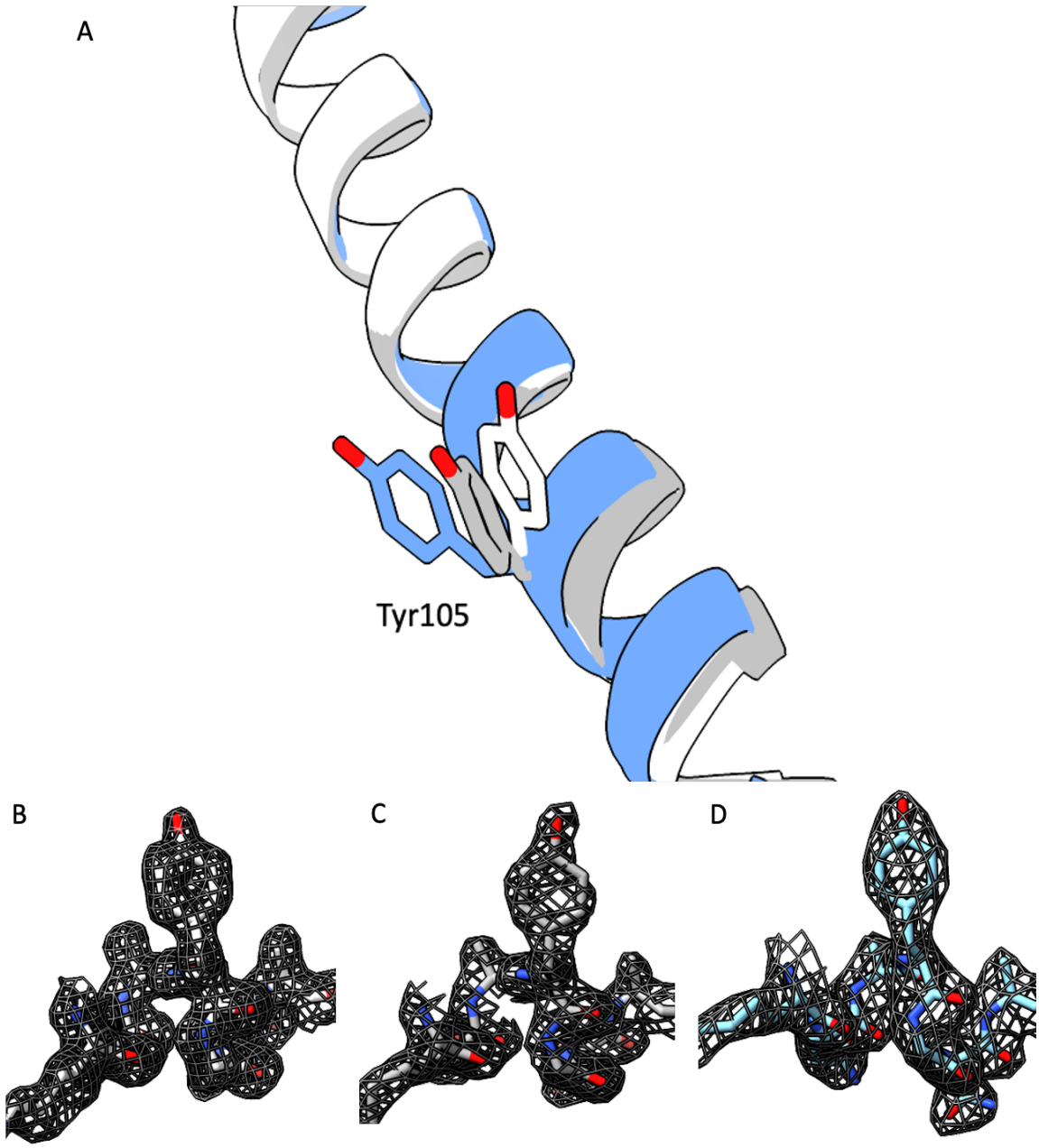

Figure 3.24 – (A) Ribbon diagram of the α6 helix of 5UWA (white), 6GKI (blue) and our POPG loaded structure (grey) showing differential deflection of the Tyr105 residue. In our structure the Tyr105 residue is deflected from the -apo position but not as far as the position of the native 5UWA structure. Map density in region around Tyr105 for 5UWA (B), our POPG loaded structure (C) and 6GKI (D).

The head group of the POPG lipid appears to interact with several residues on the α7 helix, which is disordered in the apo state from Lys177 onwards. It is believed that reorganisation of the α7 helix accommodates the conformational changes in the β1-5

sheet. The POPG lipid appears to be in hydrogen bonding range of this lysine residue, which in turn forms a hydrogen bond with the nearby Glu180. Tyr100 points towards the POPG head group but is not within hydrogen bonding distance. Comparatively the apo structure has Glu180 hydrogen bonded to Gln196 and Tyr100 hydrogen bonded to the Cα-O of Met173 [Figure 3.25]. If these interactions in the apo structure are involved in stabilising the conformation then the coordination of the lipid head group by these residues may be involved in the transition between the -apo and -holo states. This would suggest that interfering with the existing hydrogen bonds in the -apo state may be the mechanism by which other Mla proteins prime MlaC for loading. It stands to reason then that the interactions these residues make may be instrumental in the function of the system.

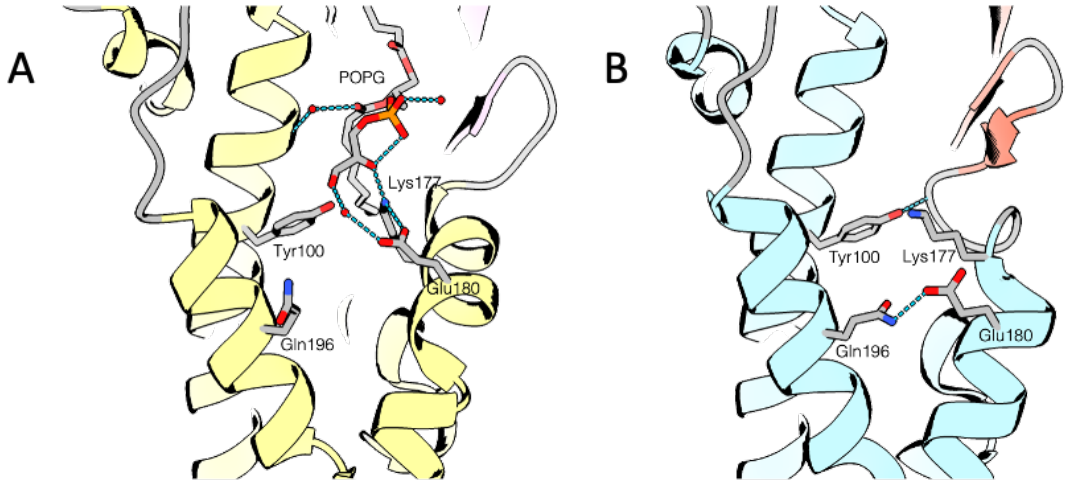

Figure 3.25 - Ribbon diagram showing interactions with the lipid head group in our POPG bound structure of MlaC (A) compared to the same region of the 6GKI -apo structure (B). Hydrogen bonding interactions are shown in blue. The POPG lipid head group forms several hydrogen bonding interactions with local residues that in the apo structure form interactions with other residues in the protein potentially resulting in reorganisation of the α7 helix.

	MlaC POPG
Wavelength	
Resolution range	34 - 1.93 (1.999 - 1.93)
Space group	R 3 :H
Unit cell	114.426 114.426 46.7459 90 90 120
Total reflections	
Unique reflections	17135 (1687)
Multiplicity	
Completeness (%)	99.95 (99.94)
Mean I/sigma(I)	
Wilson B-factor	33.26
R-merge	
R-meas	
R-pim	
CC1/2	
Reflections used in refinement	17129 (1686)
Reflections used for R-free	890 (100)
R-work	0.2153 (0.3093)
R-free	0.2632 (0.3680)
Number of non-hydrogen atoms	1681
macromolecules	1484
ligands	66
solvent	131
Protein residues	184
RMS(bonds)	0.010
RMS(angles)	1.27
Ramachandran favored (%)	98.90
Ramachandran allowed (%)	1.10
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	8.23
Average B-factor	37.37
macromolecules	35.83
ligands	64.57
solvent	41.12

Table 3.1 - Data collection and refinement statistics for MlaC POPG. Values in parentheses refer to the highest-resolution shell.

3.3.10.Determining the Structure of MlaDΔTM-apo

We anticipated that the apo structure of MlaDΔTM may provide useful information regarding the conformational changes that MlaD undergoes when accepting or transferring lipids. We expected that conformational changes that would otherwise be induced by MlaE may be reflected in differences in the -apo and -holo states of MlaD. As such, we endeavoured to acquire structural information of MlaDΔTM-apo. MlaDΔTM-apo was crystallized from a sample at 15 mg/ml in a 4 μL sitting drop with a 1:1 mixture of protein sample to precipitant solution. The precipitant solution contained 0.1 M MES pH 5.5 and 12% (w/v) polyvinylpyrrolidone [Figure 3.26A]. Crystals formed in these conditions were cracked and prone to dissolving when fished. Collection on these crystals resulted in low resolution datasets with high anisotropy. Optimisation and additive screening produced two conditions, which formed more stable crystals and generated less anisotropic data. These conditions were 0.1 M MES pH 5.5 and 12% (w/v) polyvinylpyrrolidone with either 10 mM betaine monohydrate or 5% DMSO [Figure 3.26B,C]. The data was still moderately anisotropic but ellipsoidal truncation of the reflections to H > 5.0 Å, K > 2.6 Å, L > 2.6 Å resulted in a solution that could be refined to an R_{free} below 0.4.

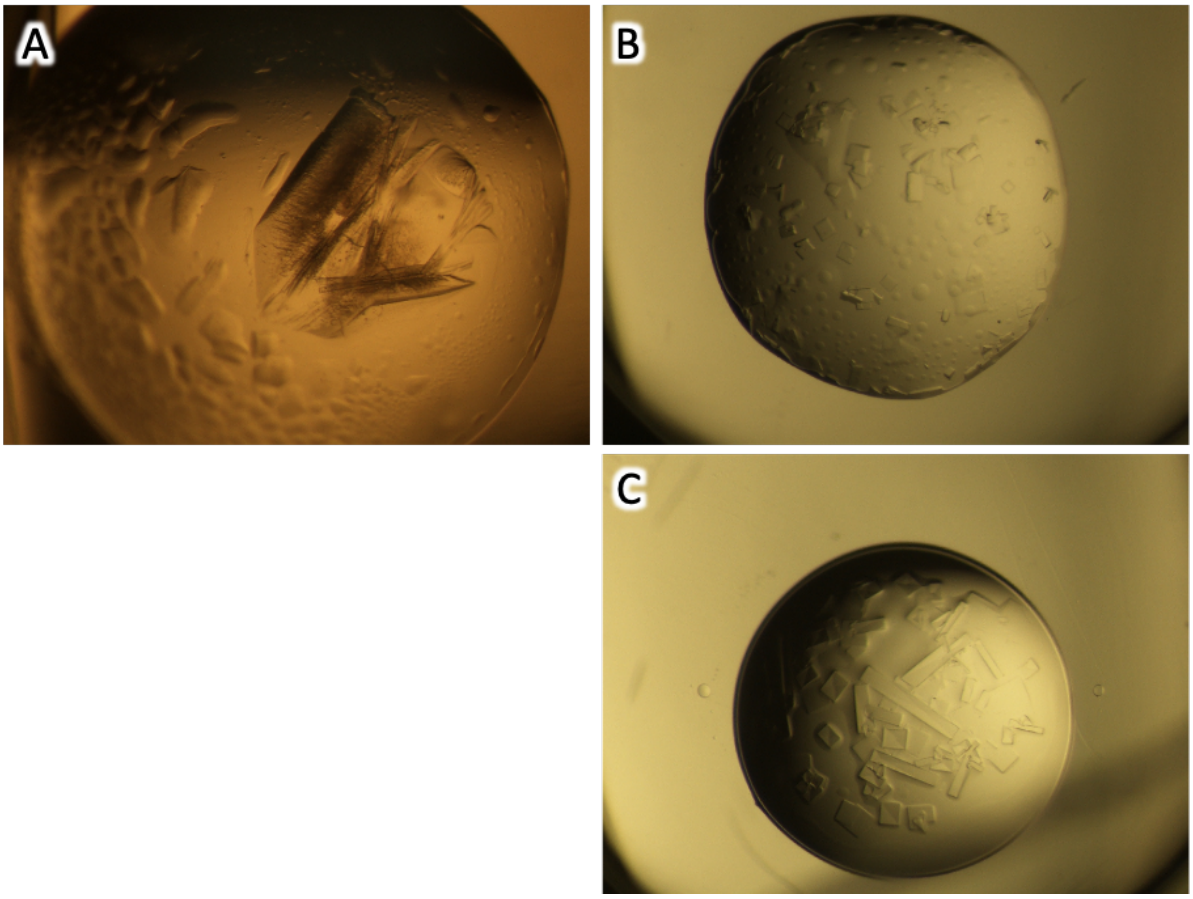

Figure 3.26 – (A) MlaD-apo crystals formed in 0.1 M HEPES pH 5.5 and 12% (w/v) polyvinylpyrrolidone. The crystals were large but diffracted poorly. More stable crystals formed with the addition of 10 mM betaine (B) and 5% DMSO (C).

Data collection revealed that the crystal belonged to the $P2_12_12$ space group and Matthews coefficient probabilities suggested 3 copies of the monomer in the ASU. A monomer of the 5UW2 crystal structure of MlaD was used as a molecular replacement model for phasing. Experimental phasing produced a solution with LLG of 1706 and TFZ of 40.5. The model was refined to a final R_{work}/R_{free} of 0.2450/0.3225 [Figure 3.27, Table 3.2]. Density for the Cα in the core fold was unambiguous, however, due to the quality of the data, it was not possible to determine the position of many side chains accurately. The final solved model showed very little significant difference from the native structure, with an RMSD of 0.571 Å across all atoms in each monomer [Figure 3.28]. No density was observed for residues in the 117 – 123 loop region or past Tyr152 in the central helix bundle. This is the same as with the native structure and

these regions may be too disordered to provide visible density. Despite the anisotropy and low completeness of the collected dataset, the refined model is sufficient to convincingly suggest no significant differences in the core fold between the native and delipidated structure.

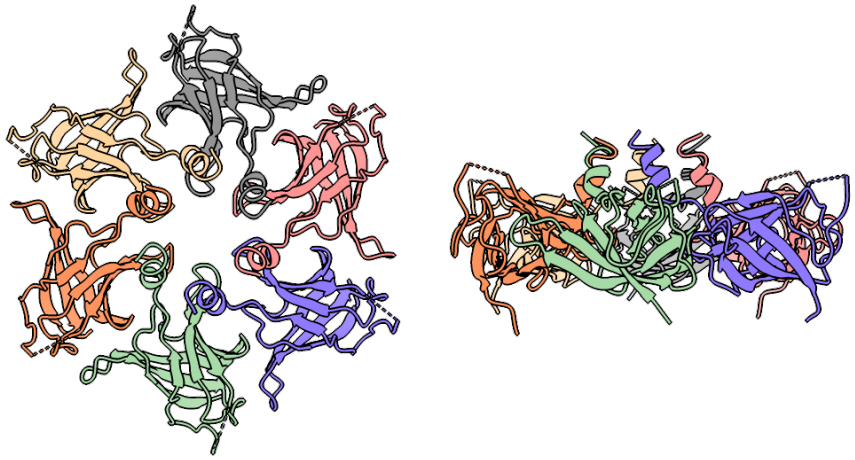

Figure 3.27 - Ribbon diagram of the final refined structure of MlaD-apo showing the ASU and the (X,-Y,-Z) symmetry copy to complete the biological assembly.

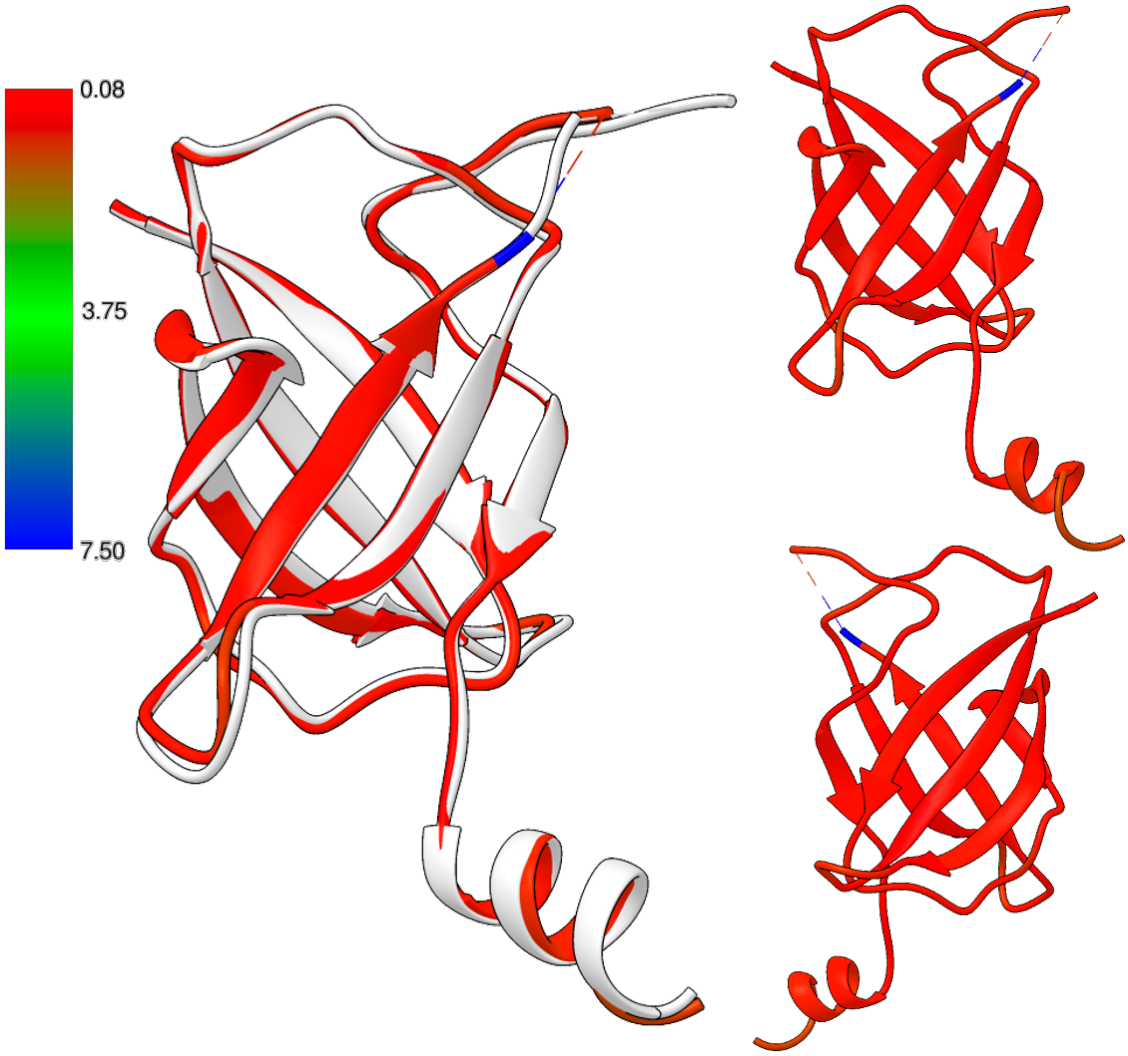

Figure 3.28 – Ribbon diagram of a monomer from the 5UW2 structure of natively purified *E. coli* MlaD (white) with the structure of apo-MlaD overlayed for comparison. The apo-MlaD structure is coloured by C-alpha RMSD from the native structure. The colour key is displayed at the far left. Front and back views of the structure of apo-MlaD are included to the right for clarity. There is no deviation between the structure of the two monomers.

	MlaDΔTM-apo
Wavelength	
Resolution range	57.1 - 2.669 (2.764 - 2.669)
Space group	P 2 21 21
Unit cell	40.93 106.78 114.2 90 90 90
Total reflections	
Unique reflections	8292 (210)
Multiplicity	
Completeness (%)	55.58 (14.31)
Mean I/sigma(I)	
Wilson B-factor	45.75
R-merge	
R-meas	
R-pim	
CC1/2	
Reflections used in refinement	8288 (210)
Reflections used for R-free	833 (23)
R-work	0.2445 (0.5200)
R-free	0.3218 (0.7814)
Number of non-hydrogen atoms	2489
macromolecules	2489
Protein residues	326
RMS(bonds)	0.009
RMS(angles)	1.27
Ramachandran favored (%)	91.08
Ramachandran allowed (%)	8.60
Ramachandran outliers (%)	0.32
Rotamer outliers (%)	0.00
Clashscore	16.13
Average B-factor	41.33
macromolecules	41.33

Table 3.2 - Data collection and refinement statistics for MlaDΔTM-apo. Values in parenthesis refer to the highest-resolution shell.

3.4.Discussion

Previous investigations suggest that MlaC primarily interacts with its glycerophospholipid substrates via hydrophobic interactions with GPL acyl chains. There is little evidence to suggest any specific interactions with the lipid head group in the existing models of MlaC. Models of the Ttg2d protein from *P.putida* and the phospholipid binding protein from *R.solanacearum* do have a modelled head group and the models suggest hydrogen bonding interaction between residues on the β2 – β3 loop and the phosphate, however, in both models the ethanolamine component of the head group is coordinated only by waters. Here we have presented the first structure suggesting that the head group, in this case a glycerol head group, may be bound into a pocket. It is worth considering that the positioning of the head group may be due to crystal packing in the lattice, which packs closely at the front face of the protein [Figure 3.29]. Although, due to the moderately positive electrostatic surface charge of that cleft and its close proximity to the acyl binding pocket it is more than likely that the lowest energy state conformation involves the glycerol head group binding in this region. Multiple sequence alignment as well as phmmer searches suggest that residues in this cleft are well conserved in the Uniprot reference proteomes database[136, 187]. The Glu180 and Lys177 residues, which we have identified to have potential interactions with the lipid head group, exhibited near 80% similarity over 14760 queried homologues, however, exact conservation of these residues is quite low, which may be in line with the assumption that MlaC has weak substrate specificity and follows from recent evidence that suggests interactions between MlaFEDB and lipids immobilised in its transmembrane domain are weak and non-specific [Figure 3.30].

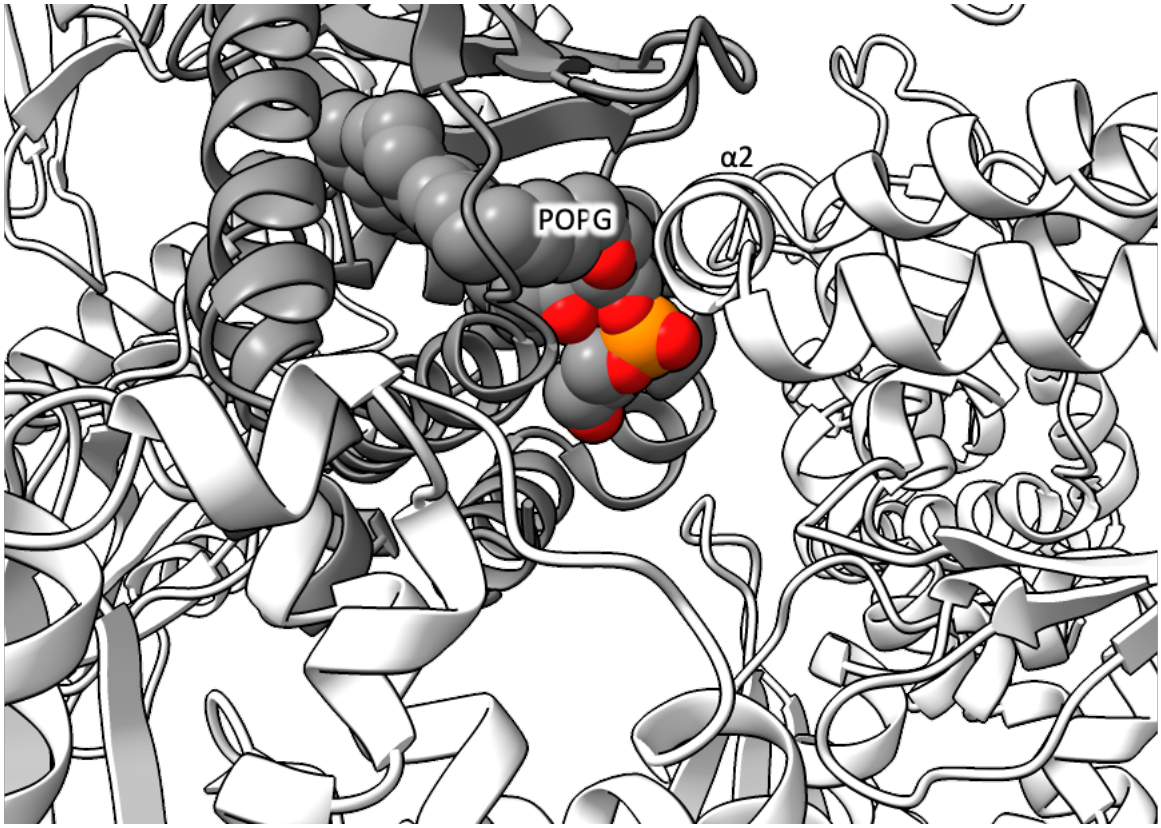

Figure 3.29 - Ribbon diagrams of the structure of our asymmetric unit (grey) and nearby symmetry copies (white). The α2 helix of a nearby symmetry copy is positioned close enough to the binding pocket to potentially prevent alternative head group positions.

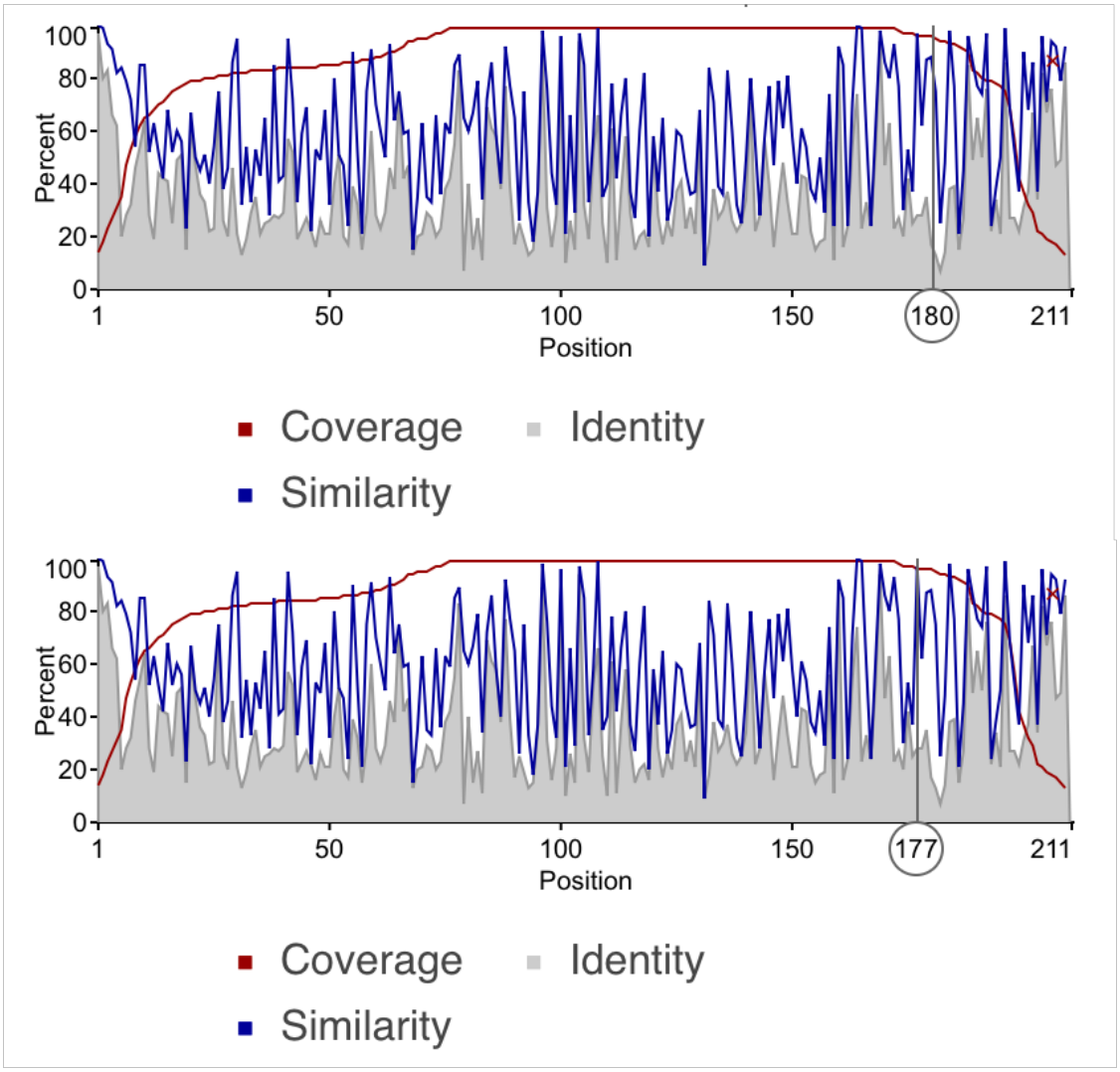

Figure 3.30 - Graphs showing conservation of residues across 14760 queried homologues of MlaC. Residues at position 180 and 177 are shown to have high similarity although low rates of exact conservation.

It can be suggested that any observed lipid specificity of MlaC native is an idiosyncrasy of the native preparation rather than a representation of the lipid binding capability of MlaC. This can be inferred from data presented in Hughes et *al.* wherein the lipid profile of MlaC loaded from MlaDΔTM native is exactly reflective of the lipids bound to MlaDΔTM native, rather than reflective of the lipid profile of MlaC-native[71]. The charge of this proposed binding cleft which is preferential for an anionic molecule

would conflict with this assumption that MlaC has no inherent preference for phospholipid head groups. It has been shown by Spencer and Brown (2015) that PldA has a specific activity for PE phospholipids nearly 10-fold greater than its specific activity for any other phospholipid[211]. It is possible that the lipid profile of natively purified MlaC is due to the high rate of PE turnover by PldA rather than any preference of the Mla pathway. It is not known if the phosphatidylethanolamine head group binds into the same cleft but it is reasonable to assume that the terminal NH_3 of the zwitterionic ethanolamine head group would not be positioned in such close proximity to Lys177 due to the opposing positive charges. We made attempts to crystallise POPE bound MlaC, however, we were not successful in producing crystals appropriate for diffraction. It is also worth considering if the length of the acyl tails have any effect on the positioning of the head group or if head group positioning is independent of acyl tail length. Crystallisation of MlaC with various lipid lengths may be a worthwhile follow-up to determine if phospholipid acyl tail length has an effect on head group binding.

We propose that the change in conformation between the -apo and -holo states may in some way be promoted or stabilised by an interaction between the head group and this binding pocket. Current proposed mechanisms for lipid exchange to/from the FEDB complex based on cryo-EM structures of lipids in the FEDB TMD cleft assume that the lipid may be exchanged in an extended state, where a single acyl tail is transferred followed by the head group and then the remaining acyl tail. This is based on the orientation of lipids observed by Ekiert et *al.* and Chi et *al.* in an extended conformation with one acyl tail projecting downwards into MlaE while the other projects upwards into MlaD[30, 71]. If we assume a anterograde direction of lipid transfer such an interaction would likely involve first transfer of a single lipid tail to MlaC, followed by a flipping of the second tail out from the MlaD binding pocket and into MlaC. It was suggested that a lateral channel between MlaD and MlaE may be associated in coordinating the head group during this process. Likewise it is possible

that once the lipid has been exchanged partially to MlaC, similar coordination of the head group is required to transfer the remaining lipid tail. In a retrograde system some interaction between MlaC and MlaD would likely trigger the selective release of one lipid tail without the release of the other. It stands to reason that interactions between MlaC and the lipid head group may be what allows this to happen. This also supports the idea of separate binding sites for each of the lipid tails which would allow for sequential binding or release.

With regards to MlaD, there does not appear to be any significant conformational changes observed in our soluble construct. Neither the core fold nor the relative position of monomers in the ring deviated substantially from the native structure. Moreover, the disordered regions between the β6 – β7 strands remain unresolvable. From what has recently been uncovered, it is assumed that conformational changes in MlaD as part of the FEDB complex are quite heavily reliant on the TM5 and TM3 helixes of MlaE, as well as the IF1 helix which runs parallel to both of these. As such no large conformational changes are observed between the soluble constructs, in response to lipid binding. Moreover, as the native structure of MlaD has no significant conformational differences with structures of MlaD as part of the MlaFEDB complex it can also be said that there is no significant conformational differences between the apo-MlaD structure and the structure of MlaD in complex with MlaFEB. This is of course not the case for the EQ_{tall} conformation presented by Chi et *al.*, which is cited as adopting a different conformation owing to stalled ATPase activity[30]. This does, however, suggest that at least *in vitro*, MlaD is capable of acting as a platform for lipid exchange even without the capability to make substantial conformational changes.

Mass spectrometry of our lipid loaded construct was in accordance with results previously presented by Thong et *al.*, with peaks corresponding to MlaD-$POPG_4$ and MlaD-$POPG_3$[184]. While it is possible that positional averaging could mask potential phospholipid states, from the structural data of lipid bound MlaFEDB presented by both Ekiert and Ximin it would appear that higher order stoichiometries above the

MlaD-$POPG_2$ state are not physiologically representative for centrally bound phospholipids. The peripherally bound lipids in the FEDB complex presented by Ximin et *al.* require the TMDs of both MlaD and MlaE which are not present in the soluble constructs[30]. However, Mann et *al.* present a structure with 7 bound lipids in and around the central helix bundle of MlaD[108]. If this is the case it is unexpected that we do not see higher order stoichiometries in the mass spectrometry data. More likely we believe that the binding of lipids to MlaD can occur with higher than twofold symmetry, however, binding is limited by the twofold symmetry of the FEDB complex. When soluble, the binding site becomes unrestricted by FEB but is still limited by steric hinderance and the maximum volume of the pore.

Overall the data presented here gives some further insight into aspects of the MlaC and MlaD lipid interaction but falls somewhat short of our initial expectations. We observed no significant changes in the conformation of MlaD attributable to lipid binding, as we had hoped to observe. While we were able to identify specific residue involvement in the coordination of a lipid head group in MlaC, the structural data of MlaDΔTM provided no further insight into the mechanisms of lipid exchange. As such the data we generated in this chapter is somewhat lacking to conclusively make any statements about the mechanisms of lipid exchange. However, the formation of samples loaded with specific lipids formed the basis for further investigations into the interaction that will be addressed in the following chapters.

4.Results: Creating a Stalled Intermediate of the Mla Pathway

4.1.Introduction

In this chapter we attempt to investigate the surface elements of MlaC and MlaD that permit complexation and transient lipid exchange interactions. We propose that the most indisputable evidence for determining the residues involved in this interaction may be provided by structural data of a stalled intermediate. However, generating a stalled intermediate and acquiring structural information on an unstable, potentially heterogeneous complex are both non-trivial endeavours. Even more so, due to the fact that all current evidence suggests MlaC readily dissociates from MlaD. Despite this, recent developments to methods of sample preparation and structural determination make this a feasible endeavour.

Conventional methods for stabilising transient interactions such as this include the use of chemical crosslinkers or cysteine mutations aimed at generating disulphide bonds and contemporary methods such as GraFix and AgarFix have been shown to improve sample preparation outcomes[2, 79]. We expect it may be possible to stabilise a transient intermediate of MlaC bound to MlaD by one of these methods. This still leaves the issue of acquiring high resolution structural information on said stalled intermediate, however, recent improvements in cryo-electron microscopy make this objective feasible.

Cryo-EM has been shown to have tolerance for moderate sample heterogeneity and recent advancements also permit structural determination of molecules as small as 50 kDa[55]. Moreover, as cryo-EM has historically been supplemented with crystallographic data[213] it is ideal in this situation wherein high-resolution crystallographic data is available. Chemical crosslinking is also often employed to

prevent subunit dissociation during grid prep of multicomponent complexes and glutaraldehyde crosslinking has been shown to be compatible with cryo-EM[112, 148, 183]. As such structural determination by cryo-EM should be robust to a chemically stabilised transient complex making this an ideal structural technique for our objectives.

4.2.Introduction to Methods

4.2.1.Surface Plasmon Resonance

Surface plasmon resonance (SPR) is a powerful surface-based technique for investigating protein-protein binding events and reaction kinetics. SPR relies on a phenomenon where electrons of a metallic surface can be excited by an incident beam of photons. Excited electrons propagate parallel to the surface of the metal and this induces a standing plasmon wave along the metal surface. Photons at the specific incident angle that cause this phenomenon are absorbed. The incident angle required to induce plasmon resonance is dependent on the refractive index of the surface. Adsorption to the metal surface modifies the refractive index and thus small changes on the surface can be measured by changes in the resonance angle[133].

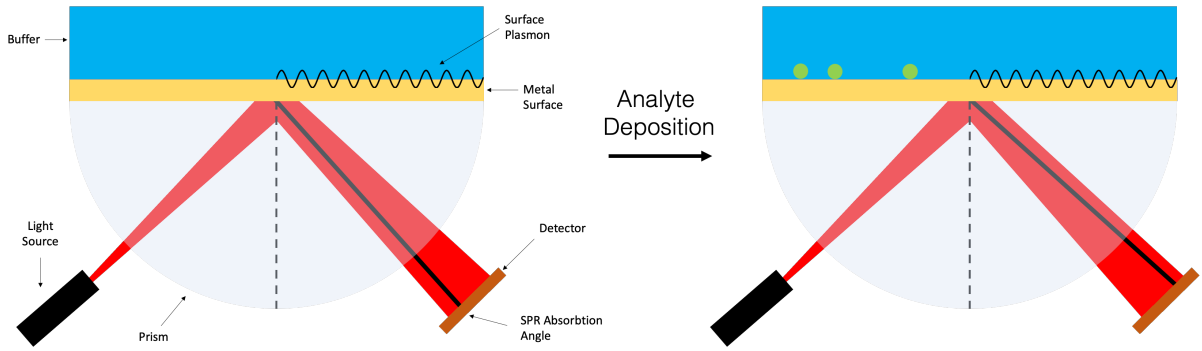

Figure 4.1 - Schematic diagram of the Kretschmann configuration, showing the experimental setup and how analyte deposition modifies the angle of plasmon resonance absorption. The transfer of energy from the incident beam to the surface is dependent on the incident angle at a fixed wavelength. This is because the propagation of the surface plasmon has a fixed vector and thus momentum of the incoming photons must match the polariton along that vector for an exchange of momentum to occur.

The general setup in an SPR experiment is known as the Kretschmann configuration, wherein a metal surface is deposited onto a glass block [Figure 4.1]. The incident beam travels through the block and induces a plasmon wave on the opposite side of the metal surface. Samples can be deposited on this surface and will modify the resonance angle. This surface can be functionalised with dextran or NTA to allow protein binding[88].

4.2.2.Cryo-Electron Microscopy

Cryo-electron microscopy (Cryo-EM) has recently become a popular method for the structural analysis of biological macromolecules. Since the introduction of this method of specimen preparation by Dubochet et *al.* in 1988, significant improvements have been made to sample preparation and data interpretation allowing for the determination of structures to sub-2 Å resolution[48]. These improvements have resulted in cryo-EM becoming an emergent method for the determination of macromolecular structures as small as 50 kDa[142].

The methodology begins with vitrification of an aqueous sample on a glow discharged mesh grid with a perforated support film, often made from gold or carbon. Glow discharging the grid creates a hydrophilic surface allowing for even distribution of aqueous samples while removing organic contaminants. The sample is pipetted onto the grid before being blotted off to leave a minimal aqueous film in the holes of the perforated support film. This is then plunge frozen in a thermally conductive liquid ethane or propane-ethane cryogen, precooled by liquid nitrogen. The result is a vitrified specimen suspended in a 10 – 50 nm thick film of amorphous ice. From this point transmission electron micrographs can be collected of the grid and the images can be processed through a pipeline of image processing techniques known as single particle reconstruction (SPA)[183].

A transmission electron microscope (TEM) uses an illuminating beam of electrons to image samples rather than an electromagnetic wave [Figure 4.2]. As the upper limit of imaging resolution is a function of the wavelength of the illuminating beam, sample imaging with electrons is theoretically capable of generating sub-1 Å resolution images. The De-Broglie wavelength of an electron is a function of its momentum and thus, electrons, which can be generated with momentums in excess of 100 KeV can have wavelengths below 0.04 Å. In a TEM electrons are generated by an electron gun. Lensing of the electron beam must be done with electrostatic lenses. The column of an electron microscope contains several condensing lenses to focus the beam at the sample plane where the specimen is loaded. After the sample plane there is an objective lens and several intermediate lenses before the projection lens which projects the image onto the detector.

The transformative breakthroughs in the field of cryo-EM have essentially come in three forms. Improvements to electron detectors and microscopes, breakthroughs in image processing algorithms and improvements to the automation of sample

preparation and collection. With regards to image collection, over the past 4 decades, the field has come from collecting electron micrographs on photographic film to the use of CCD cameras to the advent of direct electron detectors capable of collecting at high frame rates. Newer detectors made improvements to the detective quantum efficiency, allowing for collection of higher resolution information at higher spatial frequencies. The development of high frame rate direct detectors has also allowed for the collection of image stacks, which provides a means of attenuating the effects of beam induced motion and radiation damage on resolution through a process known as motion correction. Improvements in image alignment have advanced from maximum likelihood estimations to methods more appropriate to data with low signal-to-noise. The push for a Bayesian approach to minimalize overfitting noise came as late as the 2010s, far later than the introduction of such methods in other structural methods such as crystallography[153].

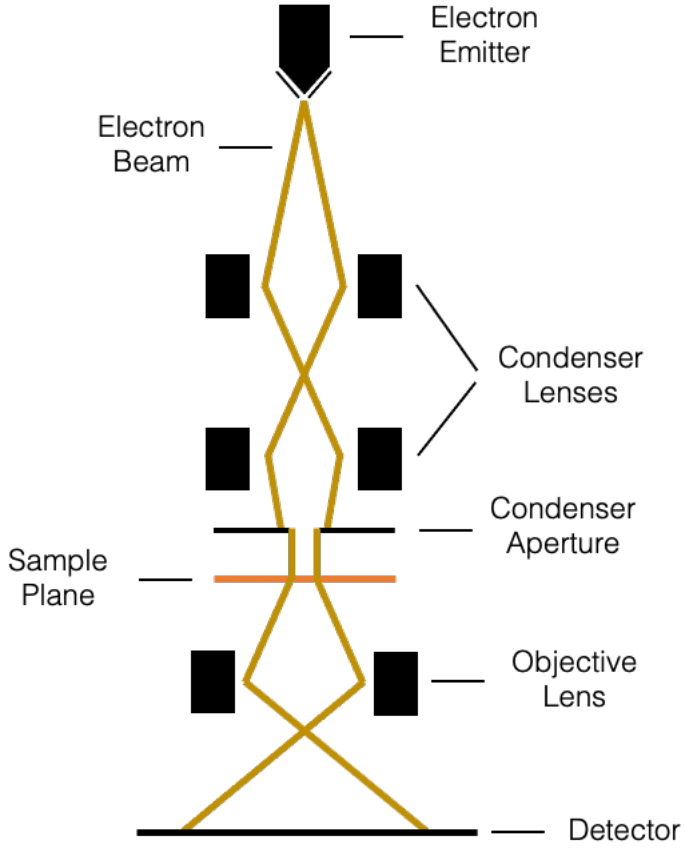

Figure 4.2 - Simplified schematic diagram of a TEM showing lensing of the electron beam. Additional objective apertures and projecting lenses may be present after the objective lens.

4.2.3.Single Particle Reconstruction

The pipeline of image processing begins with motion correction, wherein motion estimation algorithms estimate per-pixel movement and correct for beam induced motion. Then CTF correction is conducted, followed by particles selection, 2D classification, initial model generation, 3D classification, 3D refinement and finally post-processing to produce a finalised map.

The CTF or contrast transfer function refers to the power spectrum of the micrographs, visualised as Thon rings. They come about as a result of the scattered component of

the incident electrons interfering with the directly transmitted component. The result is a sinusoidal curve of intensity against spatial frequency, wherein as you go to higher spatial frequencies the phase difference of the scattered component goes from interfering constructively to having no effect on the intensity to interfering destructively. This difference in phase at different spatial frequencies comes about as a result of different path lengths the scattered components travel before they are refocused by the objective lens. CTF correction essentially inverts the intensity of the power spectrum for spatial frequencies where the interference is not constructive. The need for CTF correction comes about due to the need for image contrast. Biological material produces very little contrast in vitreous ice, and so images are collected at various defocus ranges to provide contrast, however, the periodic CTF phase interference is a function of defocus. Due to phase interference, the collected image is not truly representative of the sample and to recover the undistorted image CTF correction is required[155].

Once corrected images have been produced, determining the coordinate location of particles of interest in each micrograph and extracting them is the next phase of the pipeline. This stage can most simply be done by visually inspecting the micrographs and manually determining the location of particles. However, as large particle sets are generally required to produce high resolution maps, manually picking particles is often impractically time intensive. As such, algorithms for auto-picking particles have become a core component of the cryo-EM pipeline. A simple algorithm commonly employed for auto-picking takes advantage of the Laplacian of Gaussian (LoG) edge detection algorithm. LoG simply defines edges using the second derivative of a pixels intensity with regards to its coordinate position convoluted with a Gaussian smoothing filter. More sophisticated particle picking algorithms make use of referenced based machine learning methods and can be much more selective.

2D classification has a critical role in curating a dataset. It serves to increase the signal-to-noise ratio of particles by grouping and averaging them and to remove "junk" particles. Particles suspended in vitreous ice are immobilised in multiple different orientations, which will produce different 2D projections. However, even projections of particles in the same orientation can vary by in-plane rotations and translations, so they must be aligned prior to grouping. Intuitively, aligning images by in plane rotation is a process that may seem trivial, however, programming a computer to align images with in plane rotation is less trivial, especially when each image is incomplete due to noise. For particles with low signal-to-noise this becomes an error prone, non-trivial process. As such, statistics based alignment methods are usually employed such as maximum cross-correlation or maximum likelihood methods. Scheres (2012) describes the method for optimising particles alignment implemented in RELION 3.0[154].

An initial model is generated from the 2D tomogram projections. As the projection vector relating each 2D projection to the original object is unknown, determination of the particle orientation that resulted in each 2D projection is the computationally intensive part of single particle reconstruction. Conventionally, this is done by iterative projection matching, where an initial estimation of the 3D map is generated and theoretical projections are determined from it. Observed projections are then matched to theoretical projections and assigned the projection vector of their best matching reference. Projection vectors are assigned to all 2D particle images and then the 3D map is reconstructed from the assigned 2D particles. This is done iteratively to improve the map. In accordance with Fourier slice theory, projection matching can and is often done in Fourier space. At this stage of 3D reconstruction, particles that do not match can be differentially classified. This system of particle reconstruction underpins the main calculative component of cryo-EM[154].

4.3.Results

4.3.1.Determination of the Binding affinity of MlaC to MlaD

As a preliminary stage of assessing the conditions under which we could form a complex of our two proteins we decided to determine the binding affinity of MlaC for our soluble construct of MlaD. As we had a His-tagged MlaD construct and an un-tagged MlaC construct we decided that surface plasmon resonance would be a good technique to determine the binding affinity, owing to issues with sample aggregation during ITC. We bound MlaDΔTM to a nickel functionalised NTA SPR chip and recorded the response of a dilution series of MlaC. To determine if the presence of lipids on either protein had an effect on the binding affinity we repeated this for every combination of native and apo MlaC/MlaDΔTM [Figure 4.3].

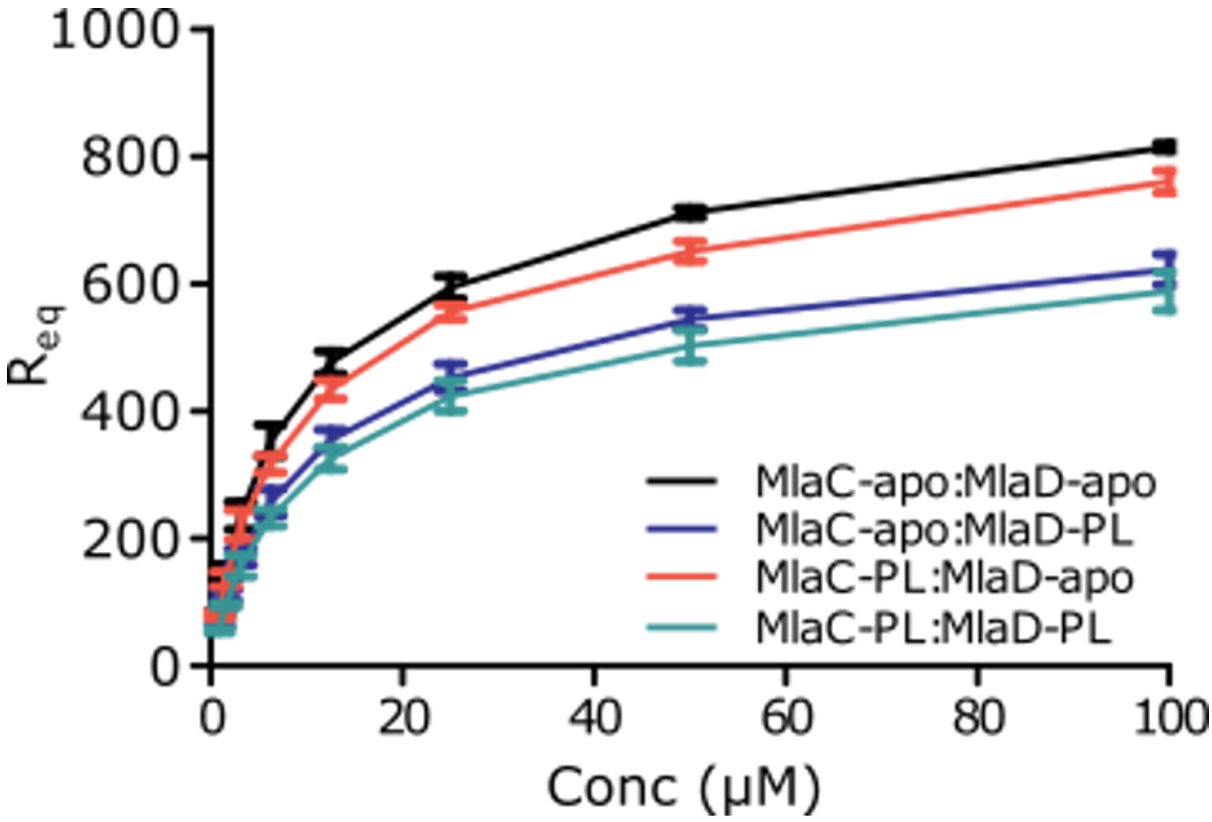

Figure 4.3 - Response change of interactions between MlaC-apo/-holo binding to surface immobilised MlaD-apo/holo plotted to determine kinetic data of the interaction. Data points are calculated from the mean of 3 repeats and error bars represent the standard deviation.

The K_D of all lipid bound states was in the range of 10 – 15 μM [Table 4.1]. These values provided us with a good estimate of the required concentrations to promote complexation in the following attempts to stabilise the interaction. This data was submitted for publication as supplementary material for Hughes et *al.* (2019)[71].

Interaction	Dissociation Constant (μM)
MlaC-apo:MlaDΔTM-apo	11.71 ± 0.95
MlaC-native:MlaDΔTM-native	13.70 ± 1.39
MlaC-apo:MlaDΔTM-native	12.99 ± 1.12
MlaC-native:MlaDΔTM-apo	12.63 ± 1.04

Table 4.1 - Dissociation constants for the interaction between MlaC and MlaDΔTM with various lipid bound states.

4.3.2.Generating a Disulphide Bonded Complex

We had initially made attempts to generate a stable intermediate complex by introducing cysteine mutants at residues identified by Ercan et *al.*[53]. We expected that the residues identified by Ercan et *al.* to result in the greatest amount of *p*Bpa crosslinking would be most likely to co-localise to the same region during interaction. As such we initially chose Val171 of MlaC and Met141 of MlaD to mutate to cysteine.

Samples of MlaC Val171Cys and MlaD Met141Cys were purified under reducing conditions. Purification of MlaC Val171Cys resulted in a sample that acted very much like WT MlaC during purification [Figure 4.4]. Some aggregation was observed in the highest concentration samples during initial Ni-NTA purification, however, upon final SEC purification, no aggregates were observed. Purification of MlaD Met141Cys appeared significantly different from purification of WT MlaD, with multiple peaks observed in the SEC trace not observed during WT purification [Figure 4.5]. SDS-PAGE of boiled samples from SEC Peaks 1 and 2 suggest that the major protein present in both peaks is the MlaD mutant, however, bands associated with the assembled hexamer are present even in boiled and reduced conditions suggesting the presence of highly stable aggregates. We pooled samples from Peak 2, which corresponded with the elution volume of the WT MlaD hexamer.

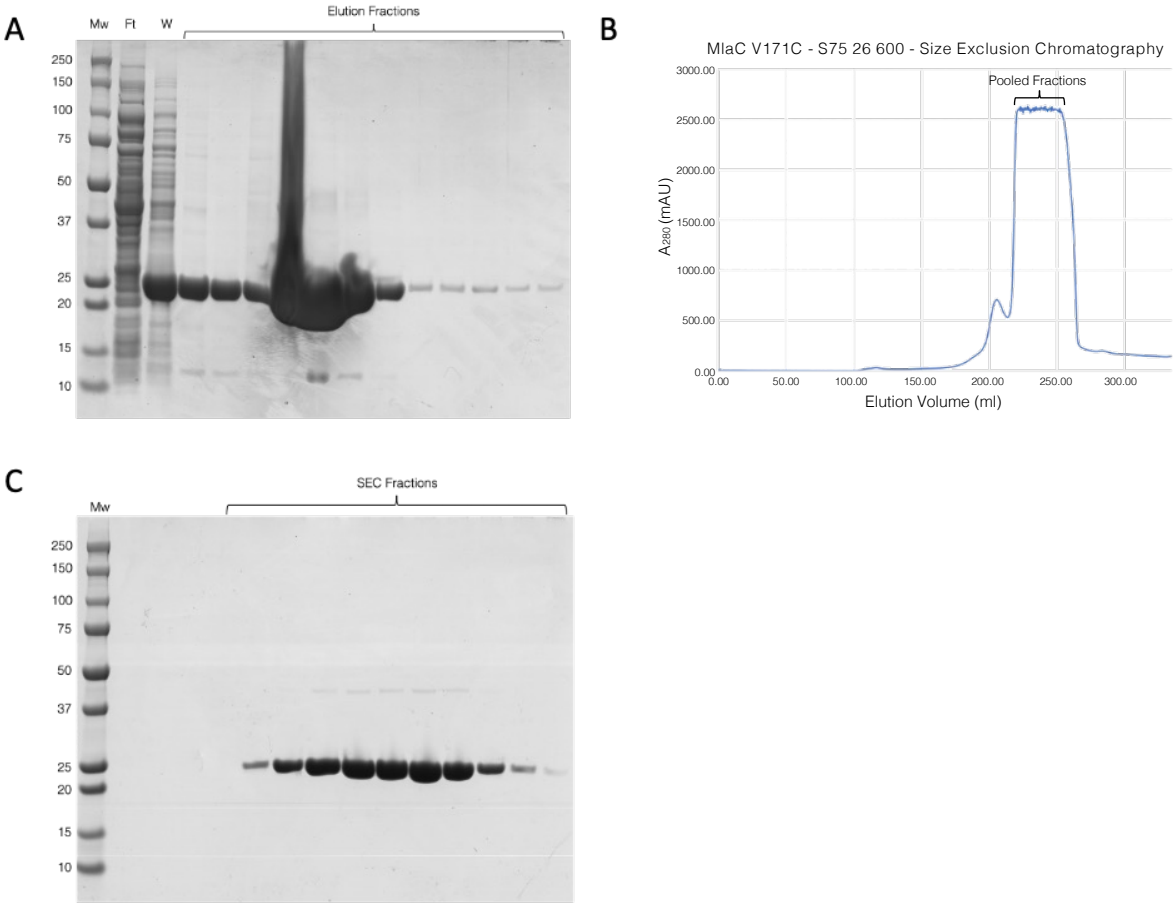

Figure 4.4 - (A) SDS-PAGE of MlaC Val171Cys Ni-affinity purification fractions. The major band present in the elution fractions corresponds with prior purifications of WT MlaC. (B) SEC trace of pooled MlaC Val171Cys elution fractions showing a single elution peak, with peak elution volume corresponding to similar purifications of WT MlaC on this column. (C) SDS-PAGE of SEC elution fractions show a band between 20 and 25kDa with no major contaminants. A band is faintly visible between 37 and 50kDa likely corresponding to a small population of dimerising MlaC Val171Cys due to unwanted disulphide bonding.

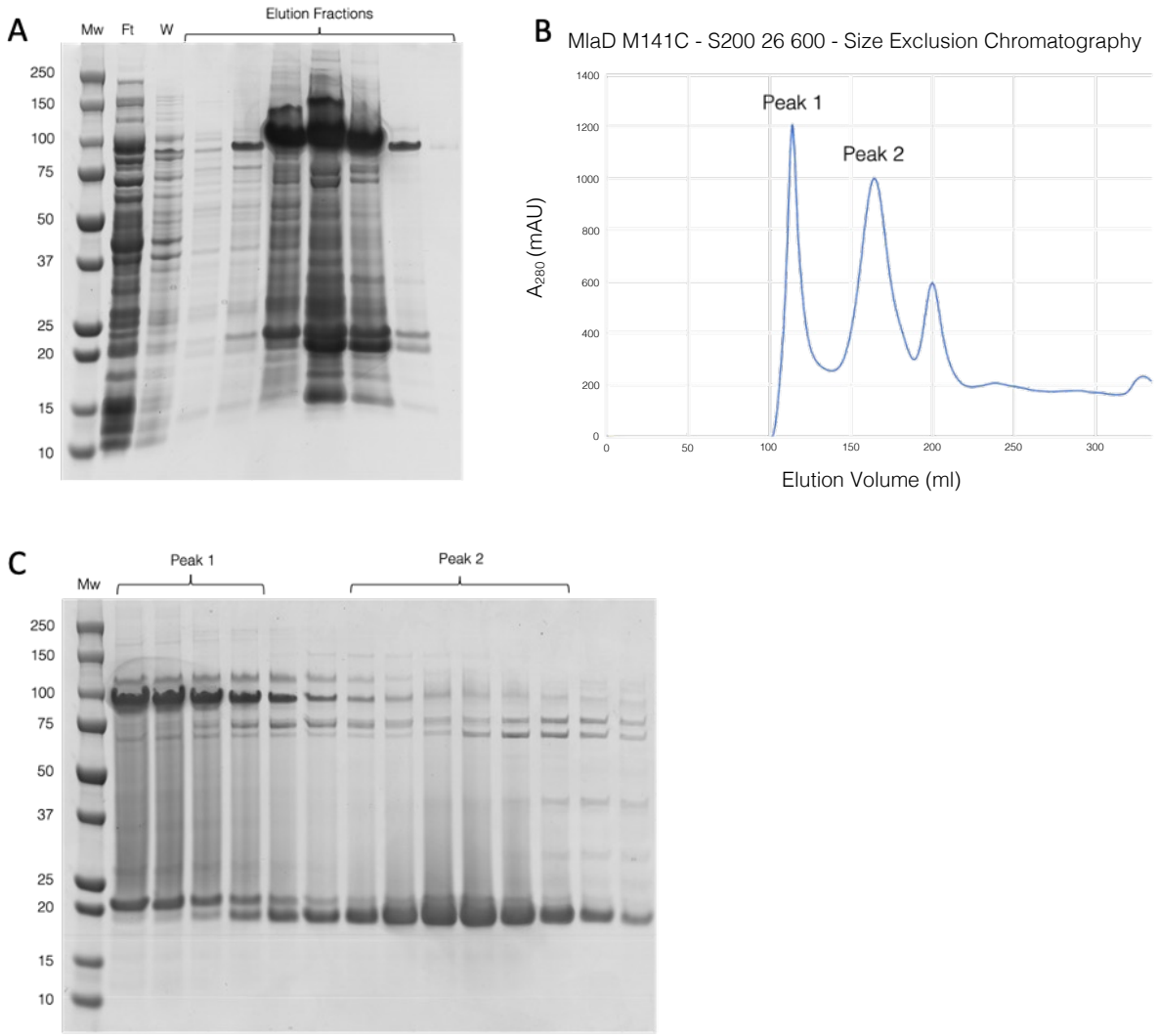

Figure 4.5 – (A) SDS-PAGE of MlaD Met141Cys Ni-affinity purification fractions. The major band present in the elution fractions corresponds with prior purifications of WT MlaD, however, the gel shows substantial contaminant bands as well as a strong band corresponding to the hexameric assembly even when the sample has been boiled in a sample buffer containing 2-mercaptoethanol. (B) SEC trace of pooled MlaD Met141Cys elution fractions showing several elution peaks. Samples from the first two major peaks, labelled as Peak 1 and Peak 2 were run on SDS-PAGE. (C) SDS-PAGE of samples from Peak 1 and Peak 2. The presence of bands associated with the hexameric assembly in samples from Peak 1 suggest that this peak contains aggregates. The bands present in the samples from Peak 2 are more similar to what is seen from purifications of WT MlaD.

To encourage the formation of disulphide bonds between MlaC Val171Cys and MlaD Met141Cys we incubated MlaC Val171Cys at 50 μM with 50 μM MlaD Met141Cys and dialysed the sample into a non-reducing buffer. We used SDS-PAGE run under non-reducing conditions to determine if we had successfully induced disulphide bonding

between MlaC Val171Cys and MlaD Met141Cys [Figure 4.6]. While we did see the presence of bands with Mw between 37 and 50 kDa in the mixed sample corresponding bands were also visible in control samples of MlaC Val171Cys and MlaD Met141Cys in non-reducing conditions. We determined that there was not sufficiently clear evidence to suggest the formation of a disulphide bonded complex between MlaC Val171Cys and MlaD Met141Cys against the background of complexes formed from these proteins with themselves. We expected that separation of any physiologically representative sample from aggregates and higher order complexes may have been difficult and as such decided to use another method of complex formation discussed later rather than pursue alternative mutations and methods to improve homogenous disulphide complex formation.

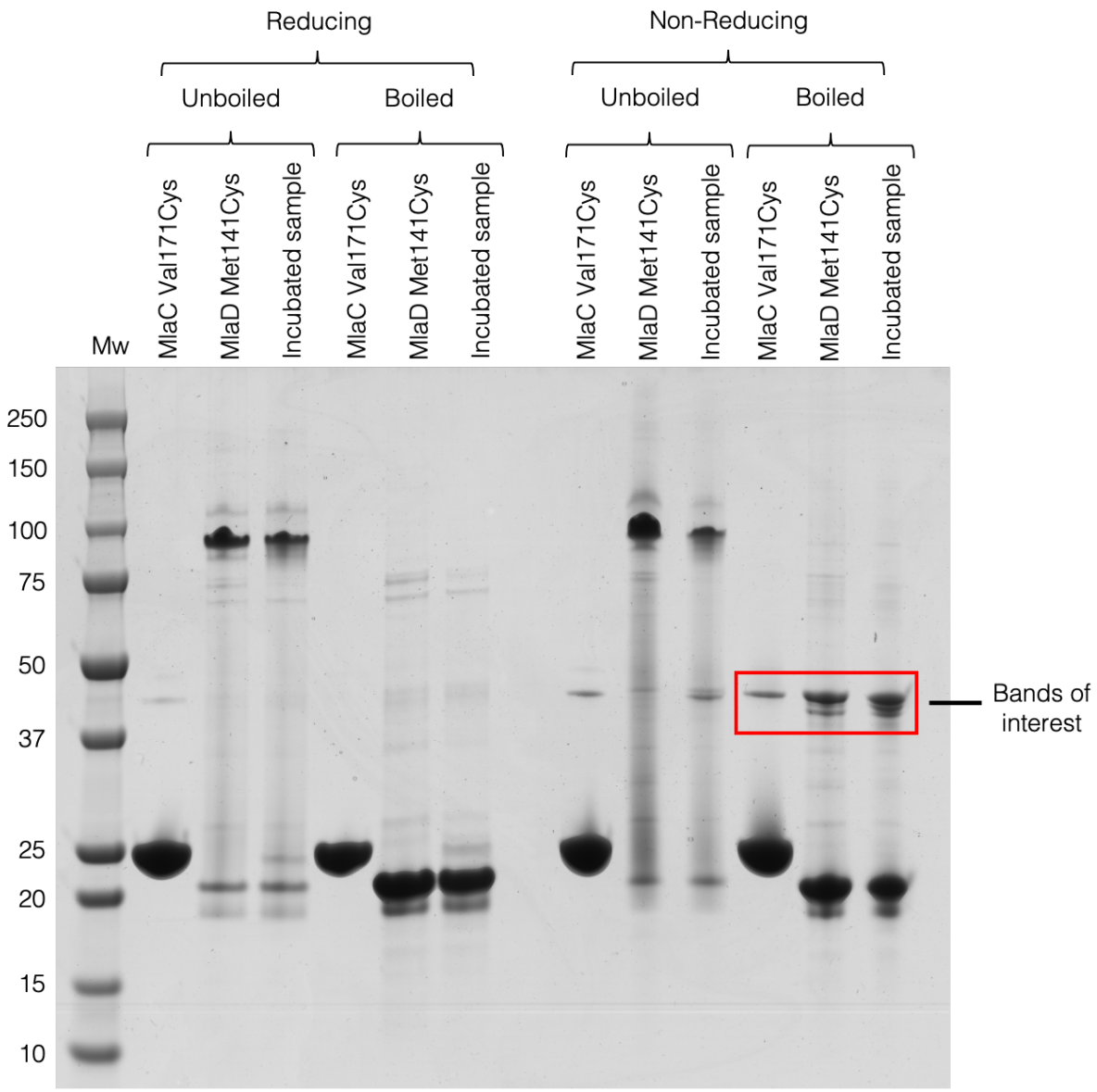

Figure 4.6 - SDS-PAGE of MlaC Val171Cys and MlaD Met141Cys incubated in non-reducing conditions alongside control samples of the individual proteins in the same conditions. The gel lanes run in non-reducing conditions show 3 bands in the incubated sample corresponding to 2 bands seen in the MlaD Met141Cys sample and 1 band seen in the MlaC Val171Cys sample. There is no clear evidence of successful crosslinking between MlaC Val171Cys and MlaD Met141Cys.

4.3.3.Glutaraldehyde Crosslinking of MlaC to MlaD

Initial attempts at producing a glutaraldehyde stabilised complex involved incubating samples of apo-MlaC and apo-MlaDΔTM together at a concentration minimally above the proposed 10-15 μM K_D of the interaction prior to the addition of glutaraldehyde. This resulted in samples that were too large to run into a 4-20% gradient gel [Figure 4.7]. It was determined that complexes substantially above 250 kDa were more likely to be attributed to random aggregation rather than the formation of any sort of physiologically representative reaction intermediate. We determined that simple glutaraldehyde crosslinking of the sample was an inappropriate method to produce a stalled intermediate suitable for structural determination by cryo-EM.

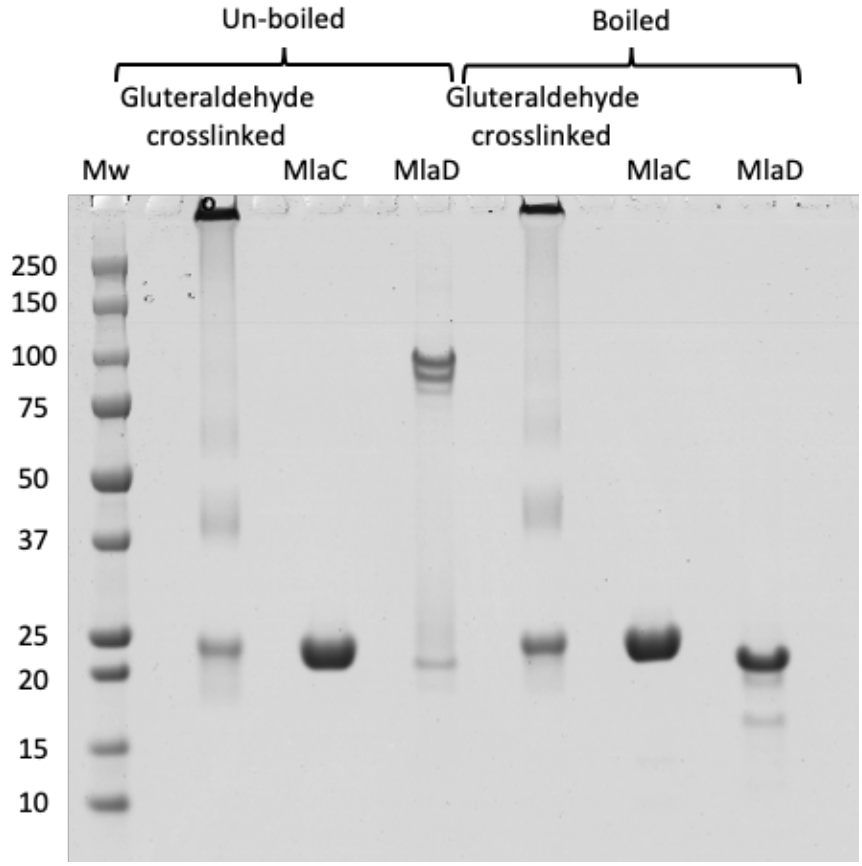

Figure 4.7 - SDS-PAGE of a glutaraldehyde crosslinked sample of MlaC and MlaD compared with control samples of MlaC and MlaD. While there are faint bands visible in the crosslinked sample, the majority of the sample appears to form aggregates too large to run into the gel.

4.3.4.Cardiolipin Stabilisation of MlaC/MlaD

As we have established in Chapter 3, MlaC and MlaDΔTM dissociate readily during SEC. This has been shown prior to this in Hughes et *al.* and is consistent with what we observe with our purified samples[71]. This remains true even with native lipids bound to MlaDΔTM. However, during trials to transfer various lipids to MlaC via MlaD (Chapter 3) we identified that MlaC is not so readily dissociated from MlaD in the presence of BDO-CL as observed via SEC. [Figure 4.8] shows elution fractions of SEC separation of BDO-CL bound MlaDΔTM from MlaC-apo. [Figure 4.8.A] shows two separate SEC runs of MlaC incubated with MlaDΔTM with and without preincubation with BDO-CL. In the presence of BDO-CL, the larger Mw peak, labelled as Peak 2 had a higher absorbance with the same amount of MlaD loaded initially. Figure 4.8.B shows the SDS-PAGE of the relevant peaks, with a 23 kDa band associated with MlaC clearly present in the lane associated with Peak 2 for the +BDO-CL sample run but not the -BDO-CL run. The difference is more clear in the boiled gel samples, where MlaDΔTM does not cause smearing due to partial unfolding. This inability to completely separate samples of BDO-CL loaded MlaDΔTM from MlaC-apo was the first indication of a complex stabilised by cardiolipin. We observe a third peak, most prominent in the +BDO-CL sample labelled Peak 3. SDS-PAGE suggests the peak contains both MlaC and MlaD in the +BDO-CL sample and is likely some larger aggregate of the two proteins.

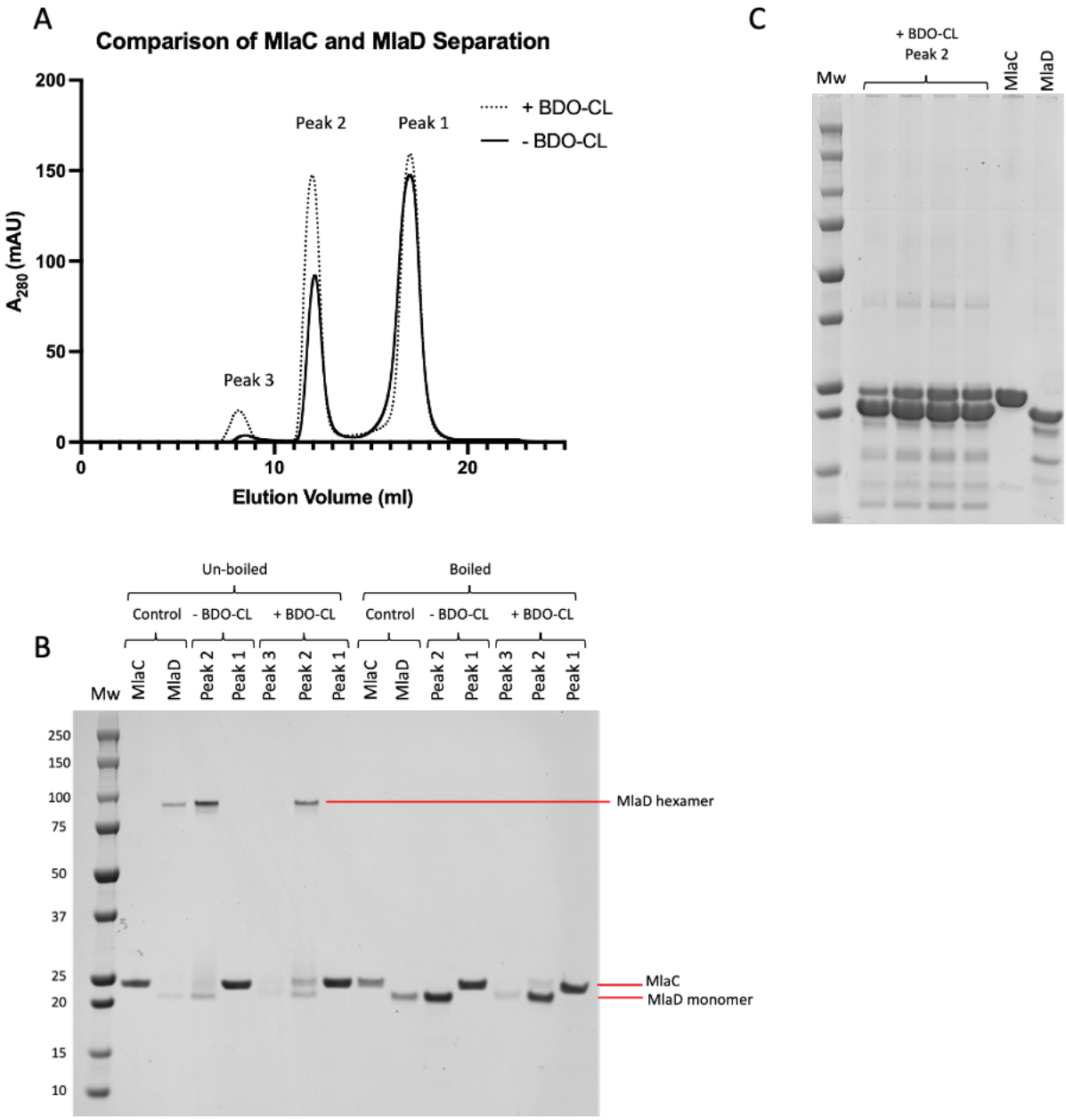

Figure 4.8 - (A) SEC trace comparing separation of MlaC from MlaD in the presence (+BDO-CL) or absence (-BDO-CL) of BDO-CL, alongside (B) SDS-PAGE of samples taken from each peak compared to MlaC and MlaD controls. For comparison 2μg of protein was loaded into each well. A band corresponding to the Mw of MlaC is only present in Peak 2 of the +BDO-CL sample. (C) SDS-PAGE of samples from Peak 2 of the +BDO-CL separation with 10 μL of undiluted sample loaded.

To further validate the presence of a stabilised complex we used sedimentation velocity analytical ultracentrifugation to more accurately assess the molecular weights of our samples [Figure 4.9]. Analytical ultracentrifugation was carried out by Pooja Shridhar at the University of Birmingham Analytical Ultracentrifugation Facility. Samples of MlaDΔTM sedimented with a calculated Mw of approximately 103.3 kDa corresponding to the expected Mw of a hexameric assembly. The MlaDΔTM sedimentation also had a very minor peak of the approximate Mw of an MlaDΔTM monomer.

Samples of BDO-CL incubated MlaDΔTM sedimented with approximate Mw of 109.7 kDa. The difference in Mw of 6.4 kDa accounts for 4 molecules of BDO-CL with Mw 1501 Da. This is in accordance with prior evidence that the preferred binding stoichiometry of lipids to MlaDΔTM is 4 per hexamer and suggests that despite being a polyanionic phospholipid, BDO-CL binds with the same ratio as POPG. A minor Mw peak at approximately 225.9 kDa in the BDO-CL loaded sample suggested possible dimerization of the lipid loaded sample and may explain the larger Mw peak observed in SEC.

The sample of the putative MlaCD complex sedimented with an approximate Mw of 127.5 kDa, corresponding to a Mw difference of 24.2 kDa with the apo MlaDΔTM and a difference of 17.8 kDa with the BDO-CL loaded sample. MlaC is shown to run with an approximate Mw of 19.9 kDa, which is reasonably consistent with its known Mw of 27.7 kDa. As MlaC appears to run with a Mw of approximately 20 kDa, the Mw difference between the putative complex and BDO-CL bound MlaD is sufficient to suggest the binding of a single molecule of MlaC.

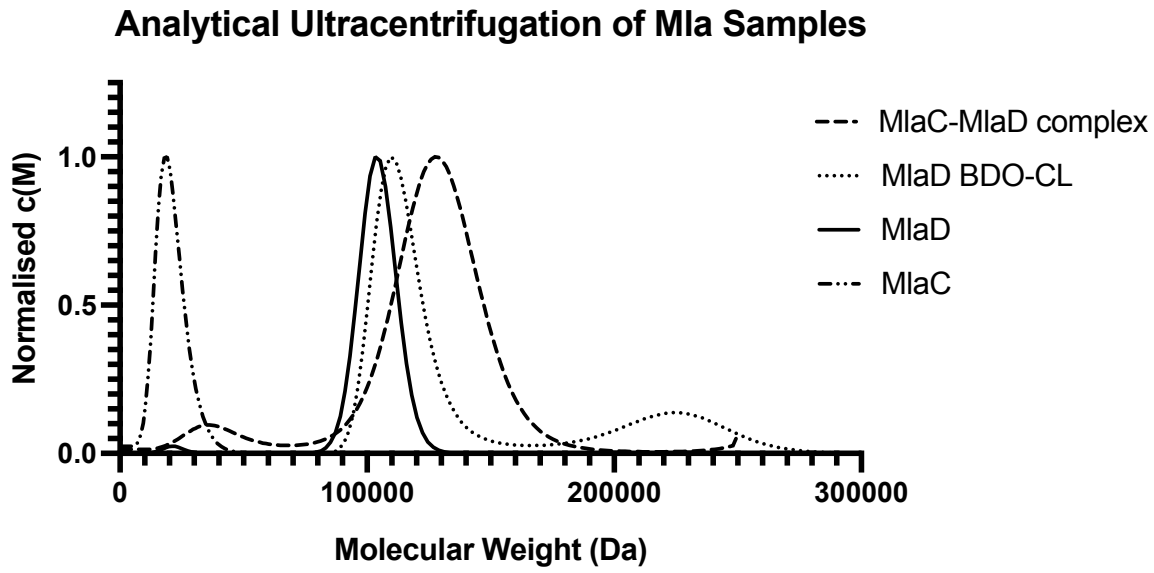

Figure 4.9 - AUC traces of samples of MlaC and MlaD compared to lipid loaded MlaD and the MlaC-MlaD complex stabilised by BDO-CL. Peak normalised c(M) suggests the complex of MlaC-MlaD consists of a single MlaD hexamer and 1 molecule of MlaC.

4.3.5.Cardiolipin as a Substrate of Mla Proteins

We had previously observed in Chapter 3 that we could load MlaD with POPE and POPG as well as exchange those lipids from MlaC to MlaD. As shown in [Figure 4.10] it was also possible to load MlaD with BDO-CL, and our attempts to transfer BDO-CL from MlaD to MlaC by reflux on a His-trap column showed moderate success. However, having observed that BDO-CL causes reduced ability to dissociate during SEC, we expect that the rate of exchange for cardiolipin lipids may be substantially lower than the rate of exchange for other di-acyl lipids. While we did not successfully confirm the reduced rate of BDO-CL transfer we expect that it could be done directly by SPR. Here we simply present that transfer of BDO-CL from MlaD to MlaC appears to be possible given enough time under reflux.

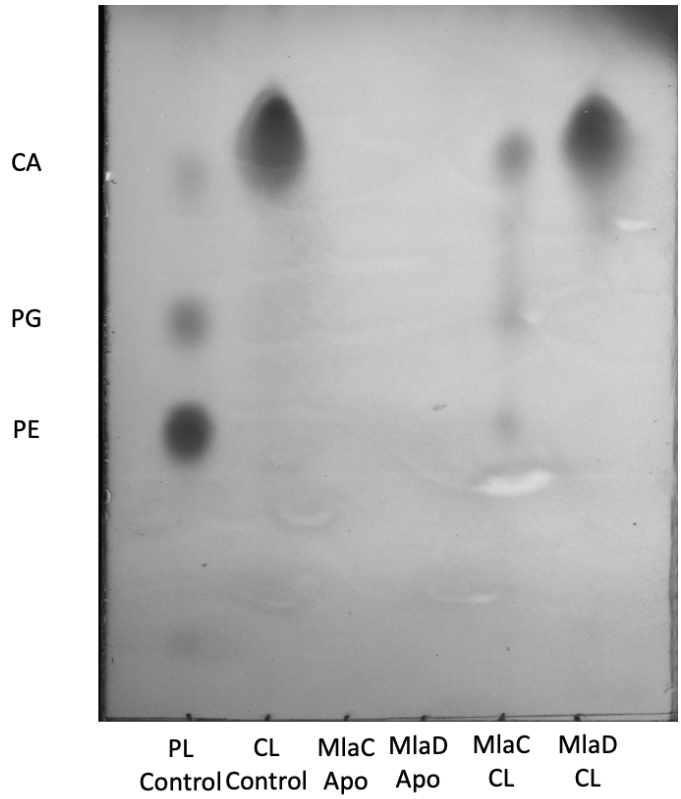

Figure 4.10 - TLC of lipid extracted from Mla proteins before and after attempts to load them with single species BDO-CL alongside relevant controls.

4.3.6.Preliminary EM Screening

We conducted preliminary screening of our stalled complex to determine the suitability of our construct for structural determination through cryo-EM. A small dataset was collected on R1.2/1.3 300 mesh gold grids treated with graphene oxide. Our sample was deposited on the grid at a concentration of 0.02 mg/ml. A small dataset of approximately 400 micrographs was collected and 2D classes were calculated from approximately 60000 picked particles [Figure 4.11].

A distinct hexagonal particle was easily recognisable in the 2D classes, however, the 2D classes strongly suggested a preferential orientation for our samples on the grid. Moreover, while MlaDΔTM was clearly visible we could not distinctly determine the

presence of MlaC. Overall, from the resolution of 2D classes, it was determined that data collection could produce a reasonably high resolution model if the issues of preferential orientation could be resolved.

The lack of visible MlaC in the top views led us to re-explore the possibility of chemical crosslinking to avoid potential issues of dissociation at the low concentrations the sample was deposited on the grid at.

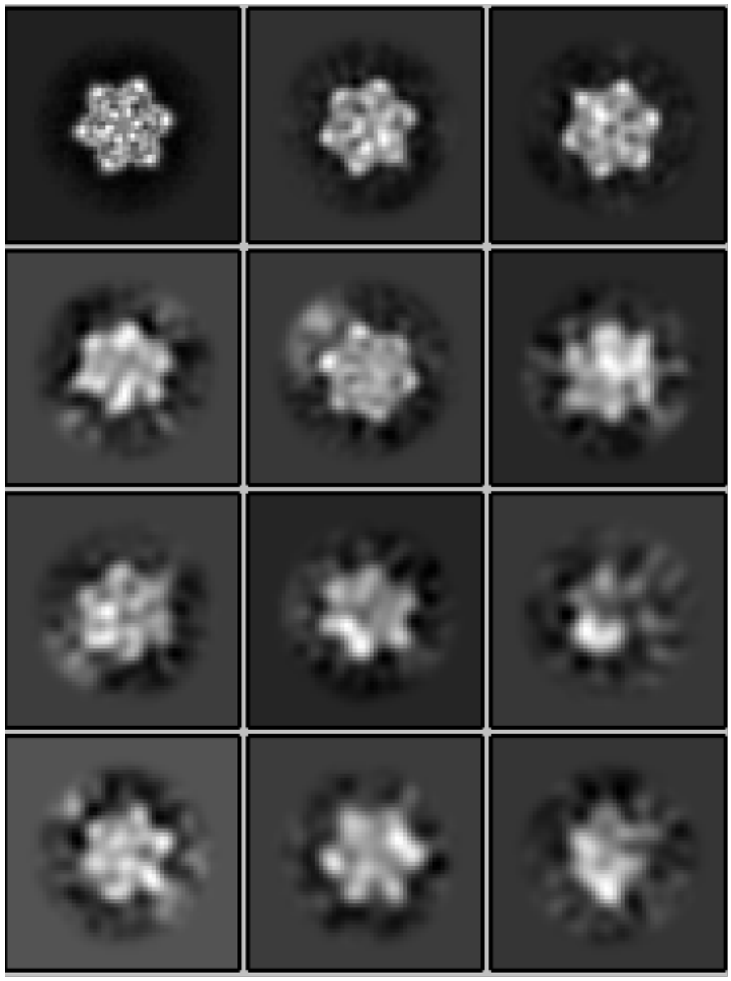

Figure 4.11 - Top 12 2D classes of a preliminary cryo-EM dataset from a sample of BDO-CL stabilised MlaC-MlaD complex ordered by estimated resolution. It is reasonably clear that there is an orientational preference for the top view of the MlaD hexamer. It is difficult to determine from these top views alone if there is a molecule of MlaC in complex with MlaD.

4.3.7.Glutaraldehyde Crosslinking of a Stabilised Complex

As it appeared BDO-CL by some mechanism stabilised the transient interaction between MlaC and MlaD, we attempted glutaraldehyde crosslinking with the stabilised complex in hopes that there would be less aggregation if a complex had already stably formed. Figure 4.12, shows the stabilised MlaC/MlaD complex after incubation with glutaraldehyde for different length of time varying from 30 seconds to 30 minutes. The crosslinked sample has no free bands associated with MlaC when boiled and appears to be a complex of greater than 250 kDa suggesting either 6 MlaC molecules are capable of binding an MlaD hexamer or that a complex of an MlaD hexamer and 1 MlaC has formed then multimerised. As we largely prioritised generating a homogenous sample and we observed no clear evidence of additional aggregation in samples crosslinked for a longer time, we decided that the sample crosslinked for 30 mins would be most appropriate to continue further EM screening with. To remove any potential aggregates that may have been generated we conducted a further stage of size exclusion [Figure 4.13]. The resulting SEC peak had a minor shoulder suggesting some sample in-homogeneity, however, we expected that the sample may be appropriate for further screening attempts.

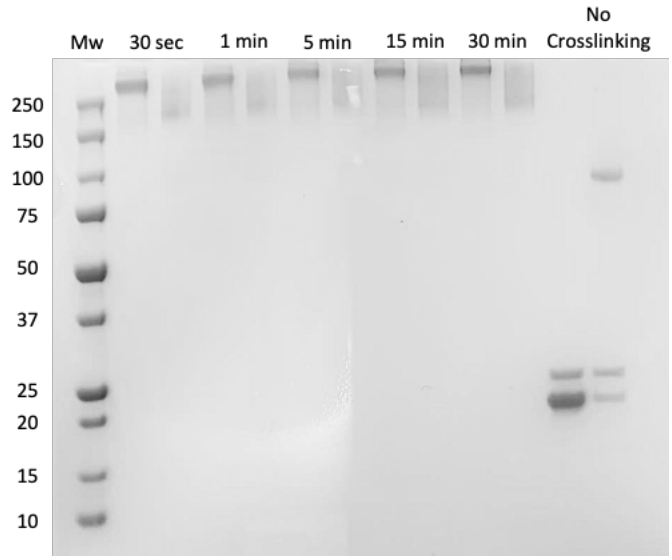

Figure 4.12 - SDS-PAGE of MlaC-MlaD complexes crosslinked by glutaraldehyde for various different lengths of time. A boiled (left) and un-boiled (right) sample was run for each time point. While minimal differences are observed in the boiled sample the smeared band present in the un-boiled sample becomes less diffuse in the samples that have been crosslinked for longer.

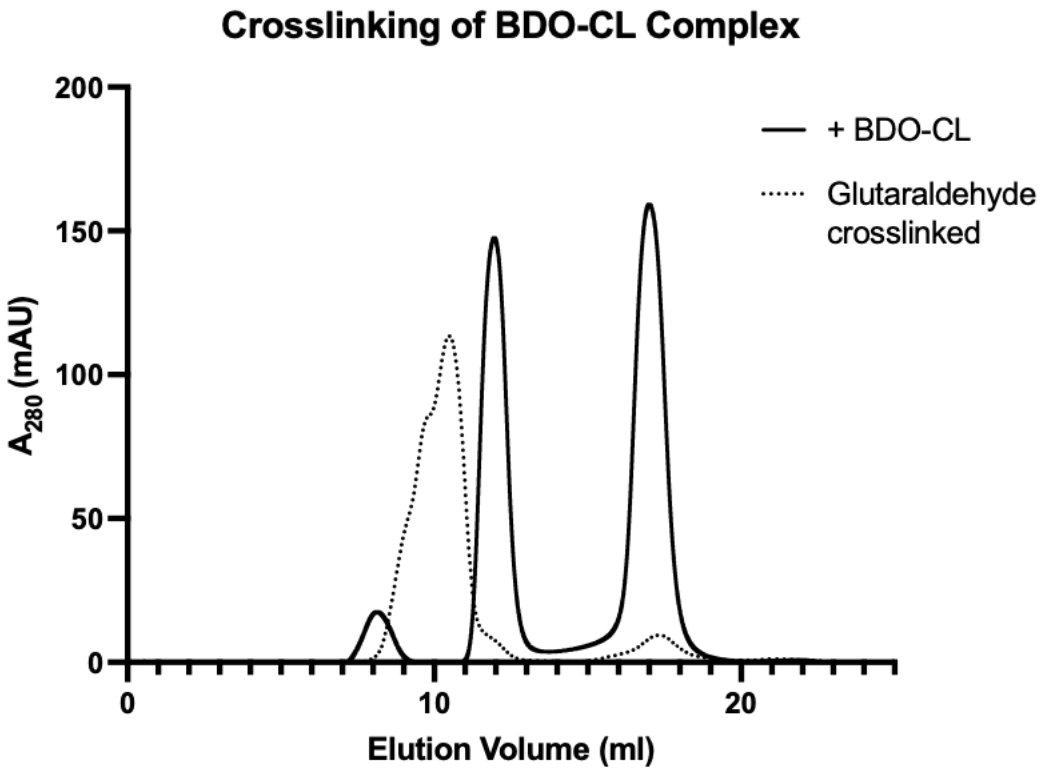

Figure 4.13 - Comparison of SEC traces of pre- and post- crosslinked MlaC-MlaD complex stabilised by glutaraldehyde run on an S200 10 300 size exclusion column. Peaks suggest crosslinking results in the formation of a complex larger what was previously observed by AUC.

4.3.8.Cryo-Electron Microscopy

Having generated what we believed to be a complex that would not dissociate during grid preparation and convincingly represented a physiological interaction between MlaC and MlaD we began to optimise grid conditions for a full dataset collection. Optimisation of grid conditions led us to collect on gold Quantifoil R2/2 holey carbon grids at a sample concentration of 0.8 mg/ml. We collected a dataset of slightly over 9000 micrographs on a Krios G3 microscope with a Gatan K3 detector at the Midlands Regional Cryo-EM facility (University of Leicester). Through a pipeline of manual picking into reference based auto-picking generated an initial set of ~1600000 particles.

Several rounds of screening particles through 2D alignment and classification resulted in a reduction of the original particle set to 525892 particles [Figure 4.14]. Side views were visible in the class averages, however, it was clear that crosslinking had resulted in the formation of what appeared to be MlaDΔTM crosslinked heterogeneously back to back. Side views also suggested the presence of MlaC, however there was evidence of complexes with both 1 and 2 MlaC bound per hexamer. We selected specifically for particles with 1 MlaC per hexamer due to a larger abundance of particles in these classes and higher resolution 2D class averages.

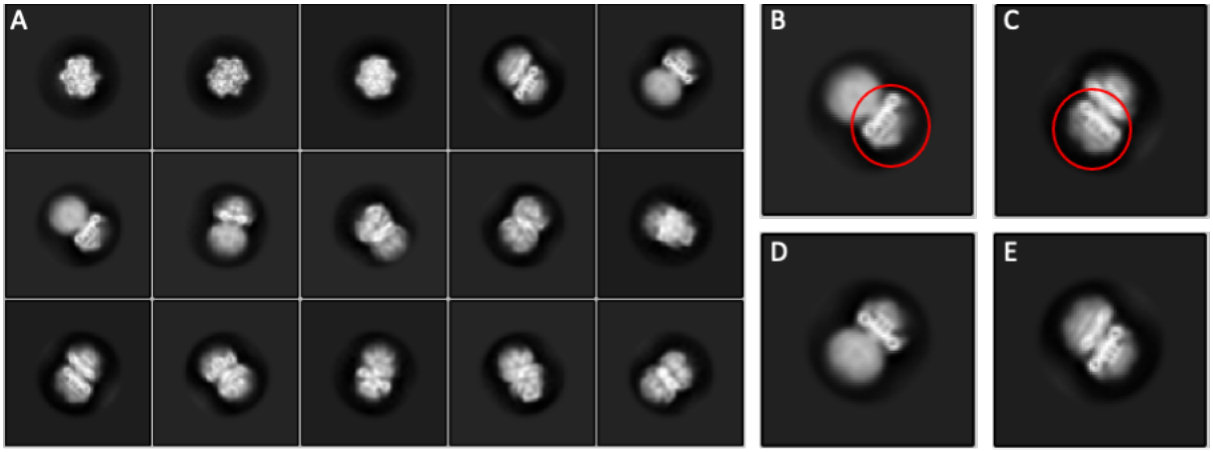

Figure 4.14 - (A) Top 15 2D classes generated from samples of BDO-CL and glutaraldehyde stabilised MlaC-MlaD complex ordered by estimated resolution.(B-E) Highlighted 2D classes showing side views of what appears to be MlaD with MlaC bound. The region circled red in (B) appears to show MlaD bound to a single MlaC molecule and the region circled red in (C) appears to show MlaD bound to two MlaC molecules.

The initial model generated from this particle set was at a resolution of 29 Å and as expected depicted an elongated complex of MlaDΔTM bound back to back [Figure 4.15A]. This assembly is likely due to the large non-polar patches on the bottom surface of MlaDΔTM, which would ordinarily interface with MlaE. As this assembly is

not physiologically representative we masked out the bottom half in an attempt to improve the resolution [Figure 4.15B].

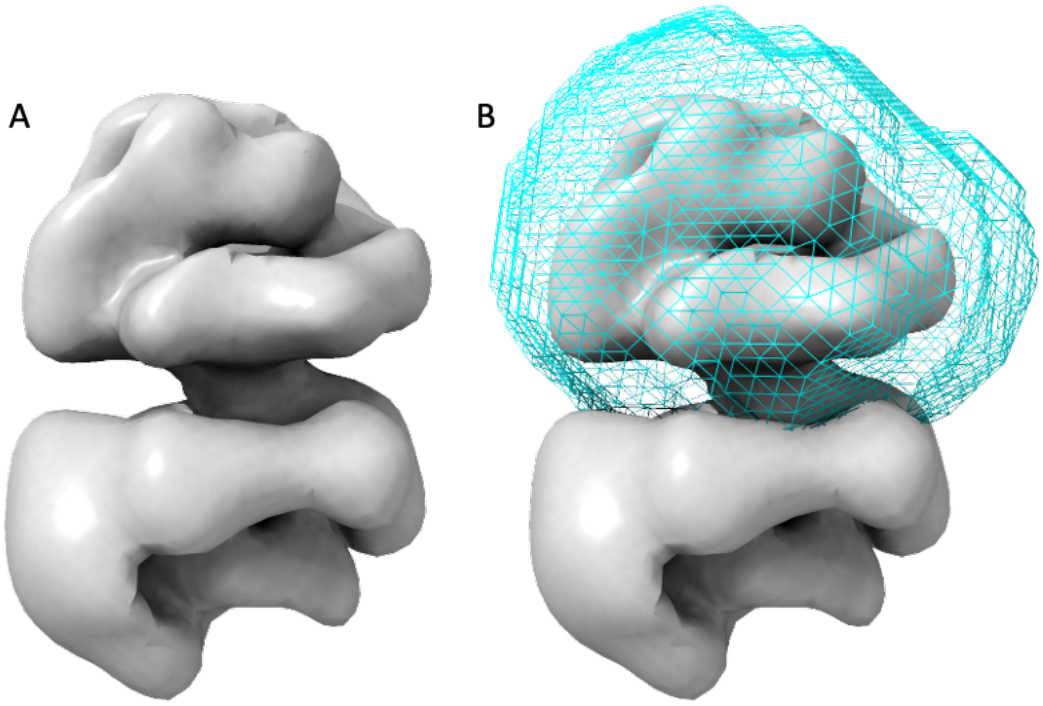

Figure 4.15 - (A) Initial model generated from curated particle set. (B) A mask (cyan) was generated to exclude one half of the assembly, which we believe to have formed as a result of glutaraldehyde crosslinking. Further 3D classification and refinement was conducted within the mask.

3D classification and refinement of the masked region resulted in iterative improvement in resolution from 12.35 Å to 5.43 Å [Figure 4.16]. Over several stages of refinement we improved the mask to better exclude the bottom complex.

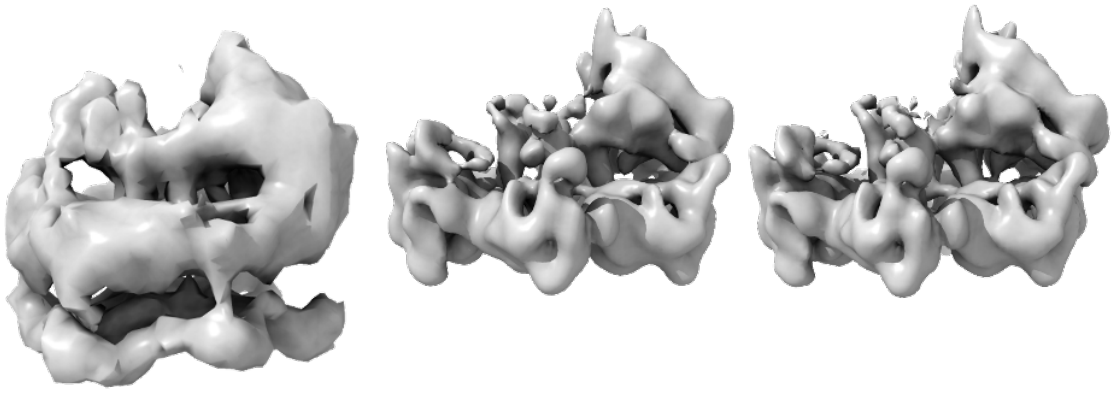

Figure 4.16 - Improvement of model resolution over several stages of iterative refinement (left to right).

Postprocessing and subsequent Bayesian polishing improved the resolution of the final model to 4.46 Å [Figure 4.17]. At this resolution it became possible to clearly trace the Cα through the map.

.

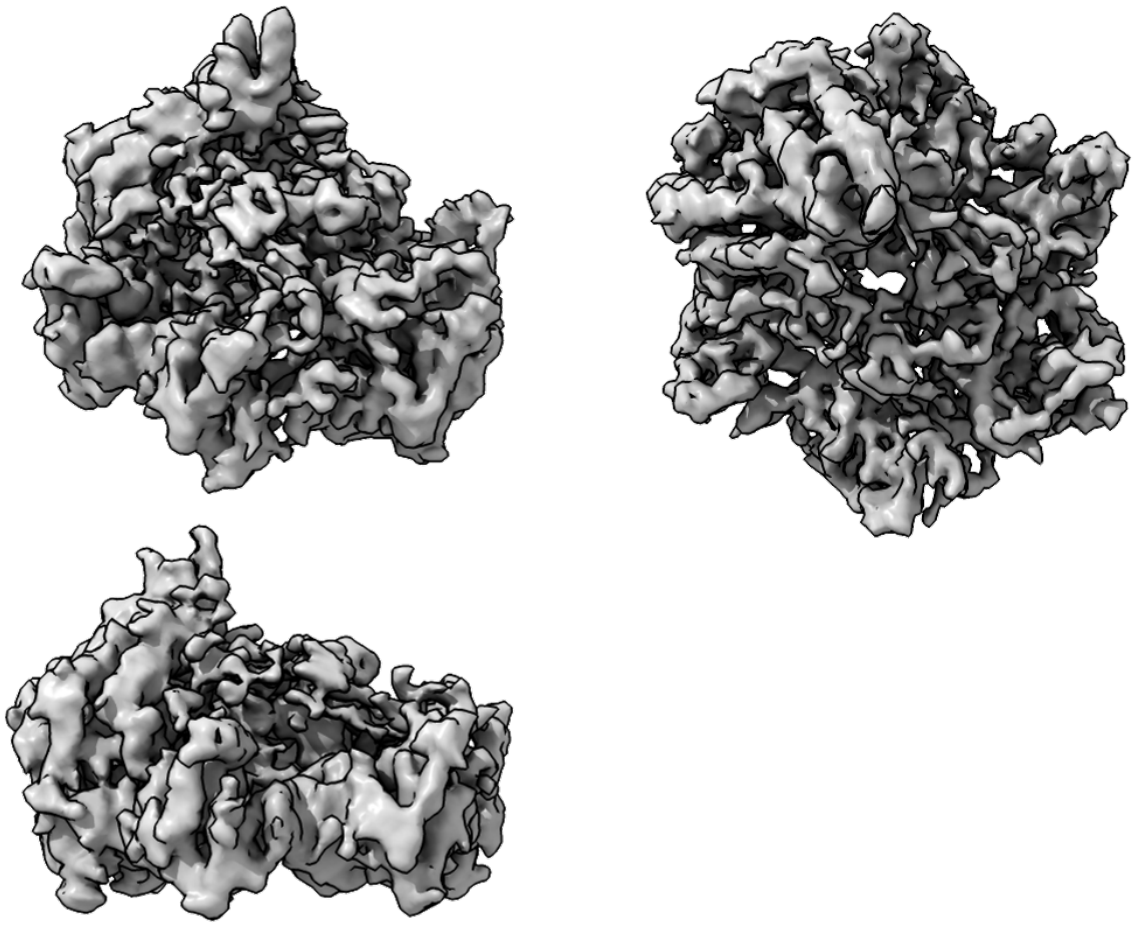

Figure 4.17 - Final density map generated from cryo-EM dataset of stabilised MlaC-MlaD complex showing various orientations.

4.3.9.Interpretation of Cryo-EM Map

While the final map was not of ideal resolution to build a model *de novo* it was of a sufficient resolution to superimpose existing models and perform real-space refinements. Here we used the 5UWA[50] model of MlaC and 7CGE[30] model of MlaD as initial models and refined atomic positions in Coot[52] to fit our EM map [Figure 4.18]. There is some unmodelled density in the map associated with the central helix bundle, however, it was not defined enough to build into. There was a defined patch of unmodelled density present at higher contour threshold at the C-terminus of MlaC allowing for the modelling of 4 C-terminal residues not present in previous crystal structures. The defined density for this usually disordered region may suggest that it becomes ordered as part of the interaction or that this conformation is stabilised by the glutaraldehyde crosslinking.

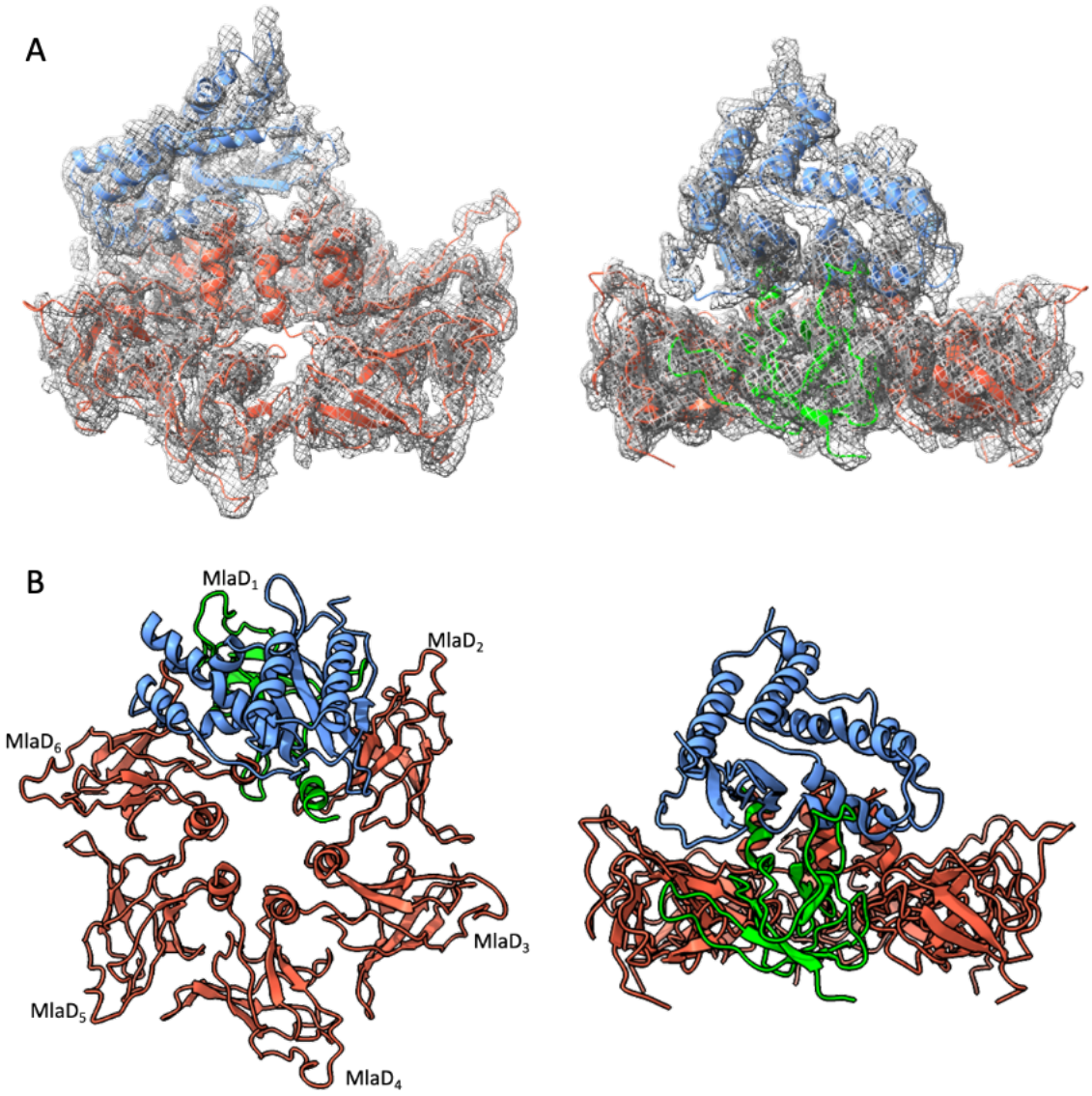

Figure 4.18 - (A) Model of MlaC-MlaD complex overlayed with cryo-EM map. The model is presented as a ribbon diagram and was generated by fitting existing models of MlaC and MlaD to the cryo-EM map. The model has been differentially coloured, with MlaC presented in blue, the MlaD monomer directly beneath MlaC coloured in green and the remainder of MlaD coloured in red. (B) Model of MlaC-MlaD complex without overlayed cryo-EM map. MlaD monomers have been numbered in the top view. Subsequent references to MlaD monomers will be consistent with this numbering.

The generated model suggests that MlaC sits offset, above a single monomer of MlaD with its lipid binding pocket directed towards the interface between two of the helices in the central helix bundle. The MlaD monomer directly beneath MlaC will henceforth be referred to as $MlaD_1$. The major interaction surface of MlaC is composed of

residues from β4 – 5, α4 and α7 as well as their connecting loops. The major interaction surface of MlaDΔTM is composed of residues from β5, β6, β8 and α1 as well as the α1 helix of the counter-clockwise adjacent MlaD monomer, which will henceforth be referred to as $MlaD_6$.

There appears to be a significant interaction between $MlaD_1$ and MlaC at the β6 – β7 loop of $MlaD_1$, which sits in a small pocket made by the α4 helix of MlaC [Figure 4.19]. The surface charge of the $MlaD_1$ loop is highly negative while the MlaC pocket is positively charged suggesting the interaction in this region may be stabilised by electrostatic forces. Glu119, Asp120 and Glu122 from $MlaD_1$ potentially interact with the backbone amines of MlaC or with nearby lysine residues like Lys83. It is also possible that the Lys83 residue of MlaC interacts with the Zn^{2+} ion suggested to be coordinated by the His92 residue of $MlaD_1$ in the 5UW2 structure of MlaD presented by Ekiert et *al.*[50]. While the map does not give sufficient detail to accurately determine the location of these side chain atoms, however, the proximity of the Cα chain would suggest that these residues interact.

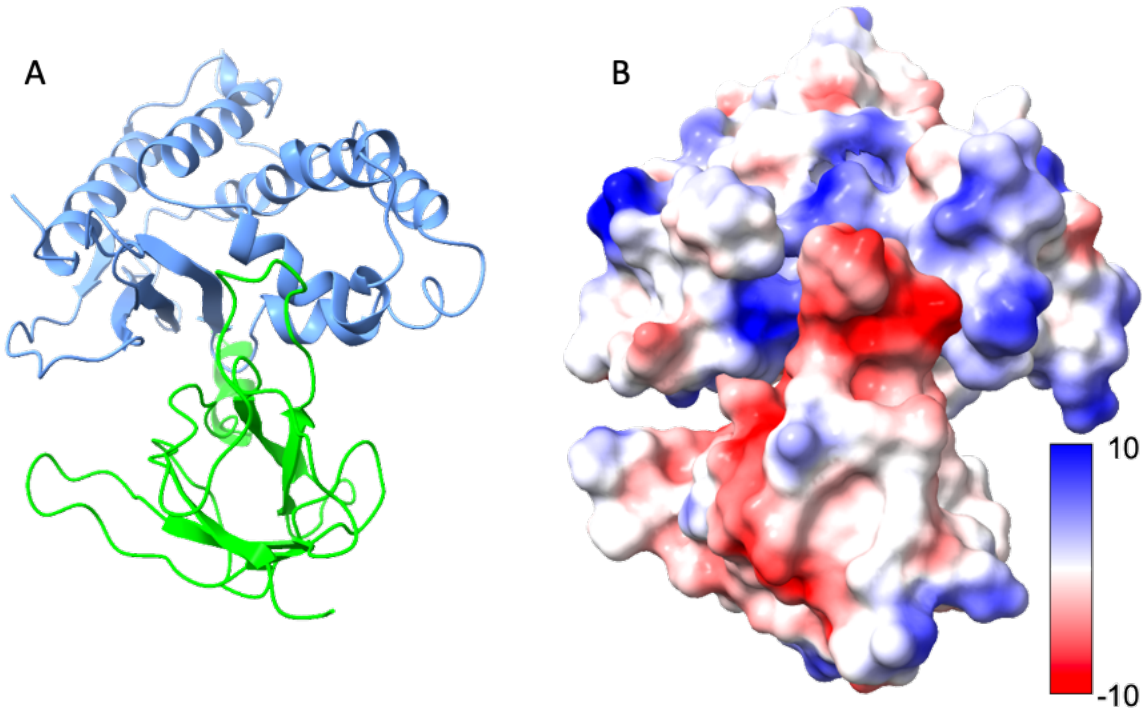

Figure 4.19 - (A) Ribbon diagram of the interaction between the β6 – β7 loop of MlaD and a pocket of MlaC alongside (B) surface diagram coloured by electrostatic potential measured in kcal/(mol·e). The β6 – β7 loop of MlaD is negatively charged and appears to sit in a positively charged pocket of MlaC.

When comparing the conformation of lipid bound MlaC to MlaC in this stalled complex, it is clear that the conformational differences are relatively minor. Almost all residues comprising core secondary structure elements deviate by less than 1 Å. The significant differences occur in the positions of loops regions between the β1 – 4 strands as well as in the α2, α3 and α7 helices [Figure 4.20]. The loops between β1 – 2 and β3 – 4 do not appear to interact with any MlaD residue and differences may be due to inherent deviation in the positioning of these loops due to a lack of defined structure. The β2 – 3 and β4 – 5 loops on the other hand appear to interact with the central helix bundle. As repositioning of the β4 – 5 loop is significant in the transition from the -apo conformation to the -holo conformation[71] this interaction may be involved in priming MlaC-apo before lipid transfer. The α7 helix is potentially disordered due to interactions between Arg186 and residues of $MlaD_6$ or due to its

interaction with the central helix bundle. Rearrangements in the α2 helix are possibly due to its interaction with the C-terminal residues and rearrangements in the α3 helix are possibly due to other conformational changes in MlaC, as this region has no interaction with MlaD.

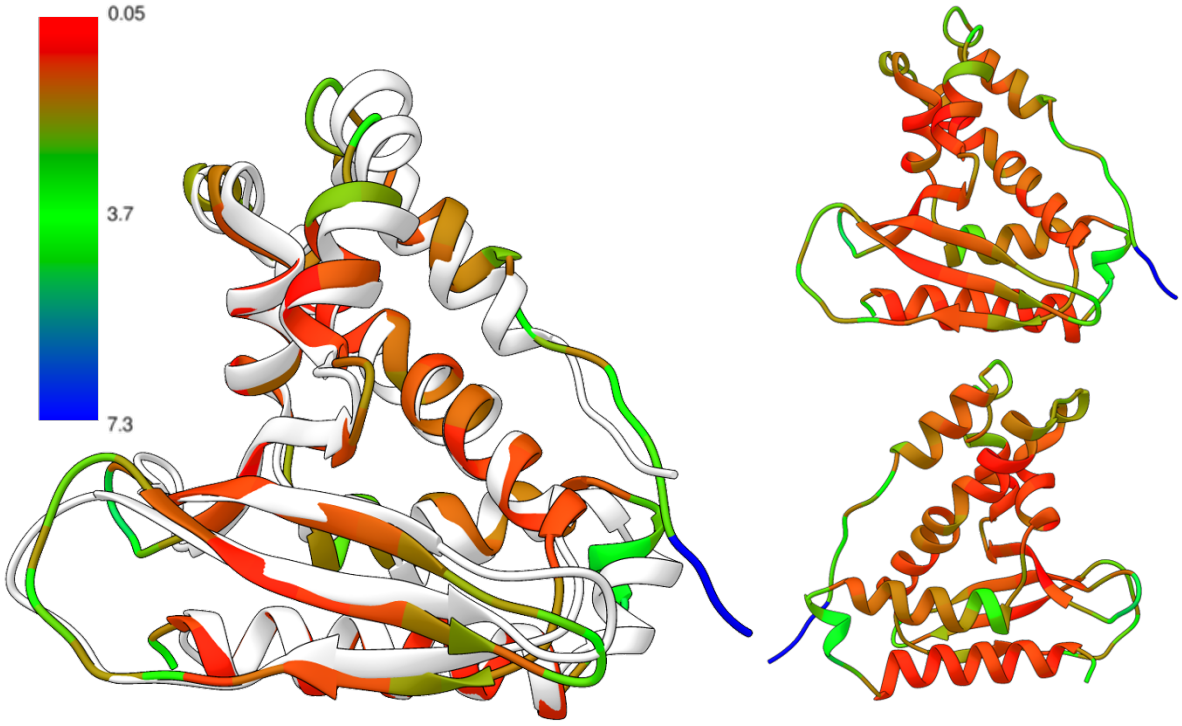

Figure 4.20 - Structure of holo-MlaC (white) overlayed with MlaC fit to the cryo-EM map of the MlaC-MlaD complex. The cryo-EM model of MlaC is coloured by C-alpha RMSD from the -holo structure. The colour key is shown on the far left. Front and rear views of the EM model are included to the right for clarity. The most significant differences between the two structures appears to be in the position of residues in the α2 and α3 helices.

There are quite clear conformational differences between the complexed MlaDΔTM and the conformation of MlaD as part of the nucleotide free MlaFEDB complex. The channel through the central helix bundle is moderately expanded in the complexed conformation, as the central Phe150 shifts back and the α1 helix of the counter-clockwise adjacent monomer is moved outwards from its original position [Figure 4.21]. Conformational changes to the core fold of each monomer are minimal, however, the relative positions of each monomer change moderately.

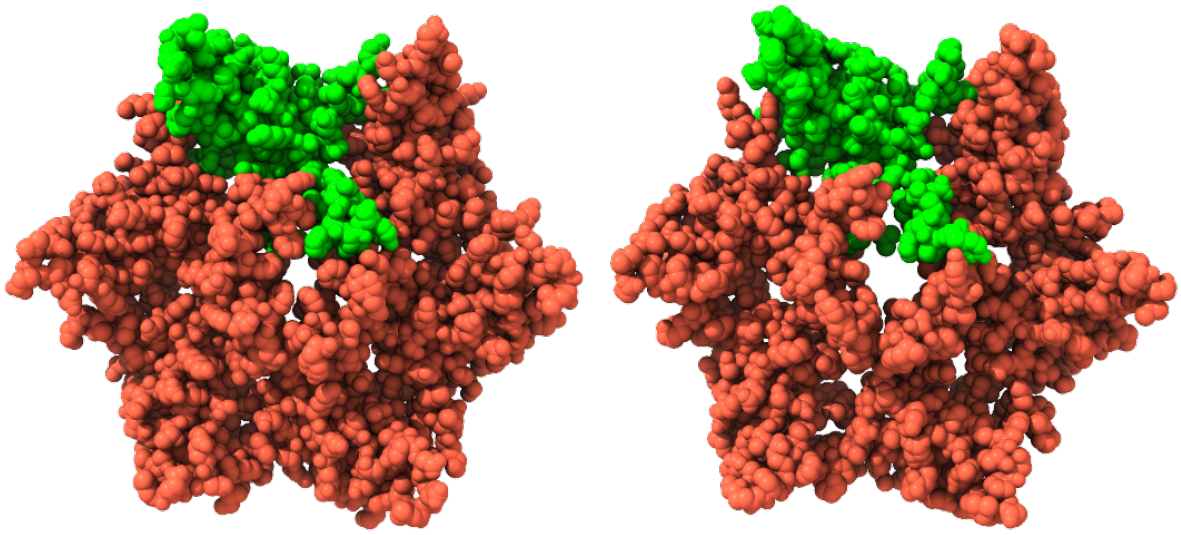

Figure 4.21 - Space filling model of the 7CGE structure of the MlaD hexamer (left) alongside our model (right), showing expansion of the central pore. $MlaD_1$ is coloured green to indicate the monomer unit which interacts most extensively with MlaC.

Considering that MlaDΔTM transfers lipids spontaneously to MlaC, it would suggest that these conformational changes are sufficient to induce an exchange. In that case it stands to reason that the interactions of these helices with MlaC are highly important to the transfer of lipids. The interaction between MlaC and the α1 helix of $MlaD_6$ appears to involve residues of the α7 MlaC helix. The Lys177 and Glu180 residues of MlaC are outward facing in this region of the α7 MlaC helix. The Glu180 residue is possibly capable of interacting with Gln149 or Asp145 of $MlaD_6$ and Lys177 of MlaC possibly interacts with Tyr152 of $MlaD_6$ or residues beyond what has been built in the model [Figure 4.22]. The α1 helix of the main interacting $MlaD_1$ monomer as previously stated interacts with β2 – 3 and β4 – 5 loops of MlaC. Glu169 from the β4 – 5 loop of MlaC likely forms hydrogen bonds with Gln149 of the α1 helix of $MlaD_1$. Arg143 of MlaC potentially interacts with the Tyr152 or $MlaD_1$ or unmodelled regions of the α1 helix beyond Tyr152 [Figure 4.23].

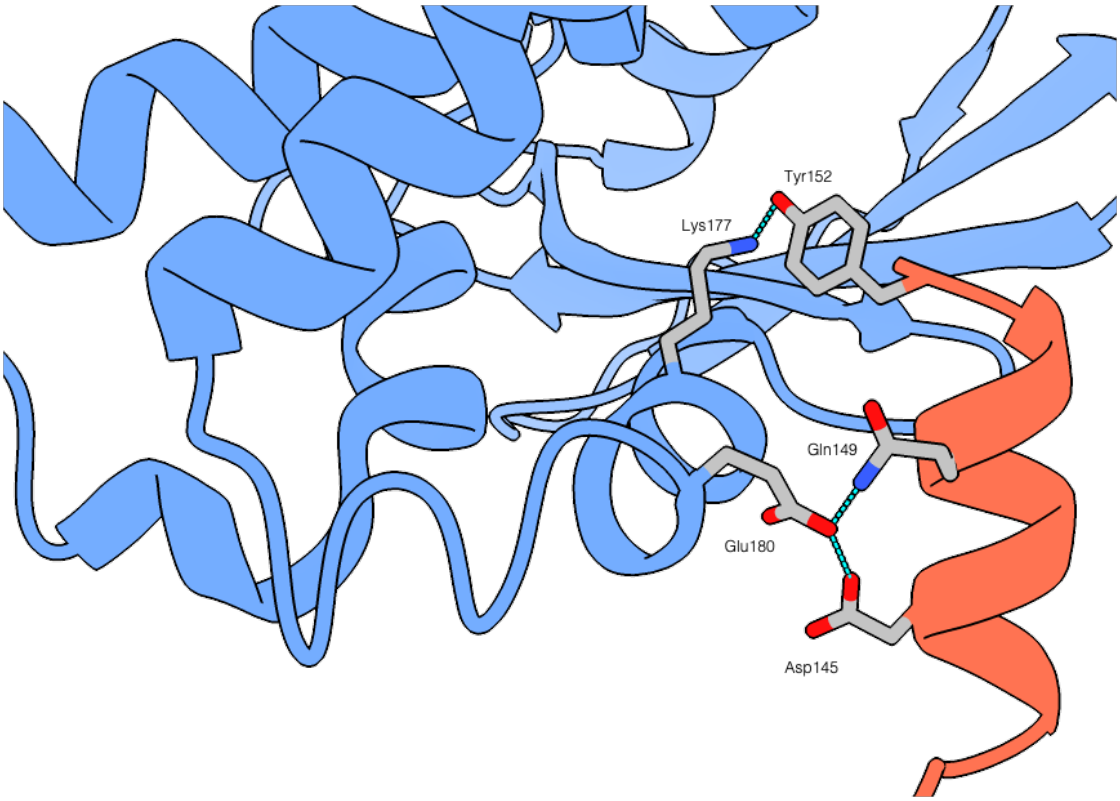

Figure 4.22 - Speculative interactions between charged residues of MlaC (blue) and $MlaD_6$ (red). Side chain atoms are positioned based on possible rotamers which best fit the local map as the map does not have sufficient resolution to accurately position side chain atoms. Potential hydrogen bonding interactions that may contribute to stabilising the interaction or inducing conformational changes involved in lipid exchange have been suggested in blue.

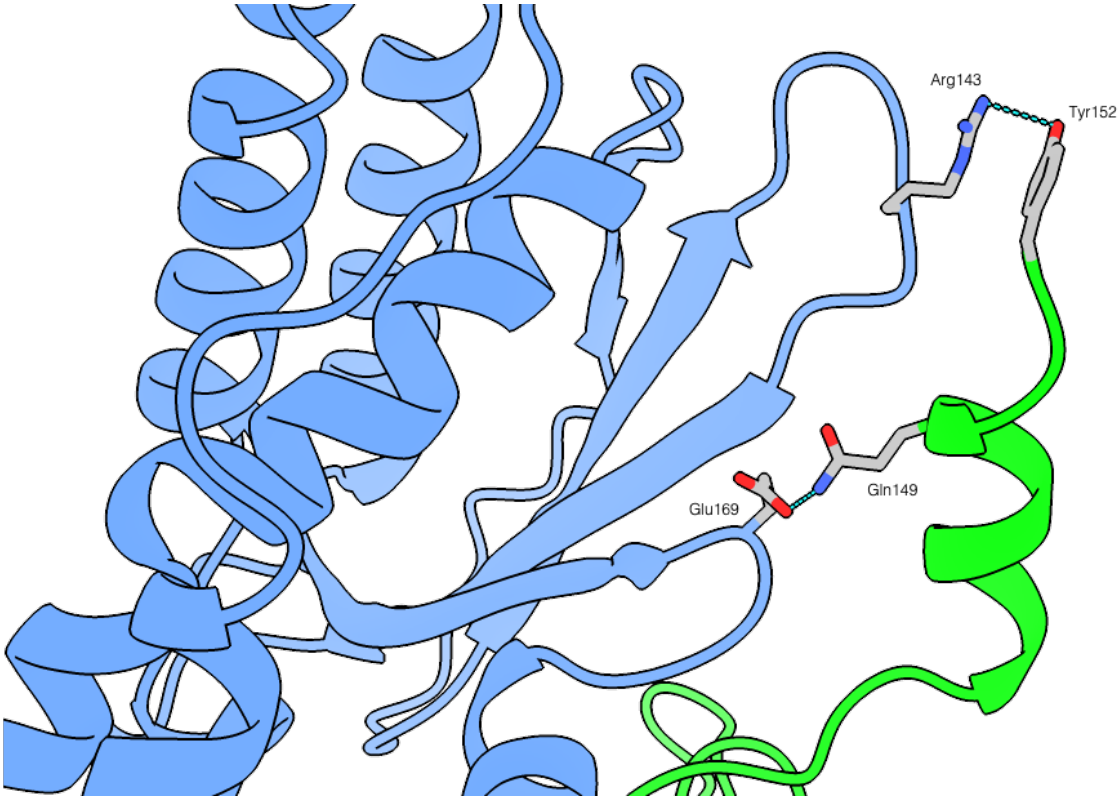

Figure 4.23 - Speculative interactions between charged residues of MlaC (blue) and $MlaD_1$ (green). Side chain atoms are positioned based on possible rotamers which best fit the local map as the map does not have sufficient resolution to accurately position side chain atoms. Potential hydrogen bonding interactions that may contribute to stabilising the interaction or inducing conformational changes involved in lipid exchange have been suggested in blue.

4.4.Discussion

In this chapter we present a methodology for stabilising the transient interaction between MlaC and MlaD for structural studies using cardiolipin. We have observed, that in the exchange direction of MlaD to MlaC, cardiolipin may not be effectively transferred and actually stalls and stabilises the complex. As it stands literature suggesting cardiolipin is a substrate of the Mla pathway varies between investigated orthologues. Tandem mass spectrometry carried out by Chi et *al.* in 2020 only identified phosphatidylglycerol and phosphatidylethanolamine bound to the FEDB complex, however, prior mass spectrometry investigation by Yero et *al.* into substrates

of the MlaC orthologue Ttg2D from *P.aeruinosa* identified *m/z* that corresponded to cardiolipin[30, 205]. Interestingly, Yero et *al*. noted that many of the *m/z* peaks they believed to be cardiolipin had lost a single acyl chain. They did not note any lyso-PE or lyso-PG substrates. Ercan et *al*. also reported on the substrates they found bound to *E. coli* MlaC using ^{31}P-NMR and their results presented a lack of bound cardiolipin in accordance with Chi et *al*. and suggested there is near equal representation of bound PE and PG[30, 53]. The difference in observed substrate specificity may have to do with structural differences observed between *E. coli* MlaC and the *P.aeruinosa* orthologue. Ttg2D has a larger binding pocket than *E. coli* MlaC and is modelled with two natively bound PG phospholipids in available structures. It is reasonably clear from the results of these three groups that there is a difference in the binding profiles of these two orthologues, likely due to differences in the volume of the binding pocket. Our results seem to agree with results presented by Chi and Ercan et *al.*, suggesting that *E. coli* does not traffic cardiolipin through the Mla pathway, as we see tentative evidence of cardiolipin exchange by TLC and evidence of a reduced capacity to dissociate with cardiolipin present. However, what we present here is largely circumstantial evidence as *in vitro* lipid exchange may not be entirely representative of the system *in vivo* and we do not present comparative quantitative evidence of reduced lipid exchange. Should it be necessary to provide more evidence suggesting that cardiolipin is not a native or preferred substrate of the *E. coli* Mla pathway, it could be done by tracking lipid exchange directly with a method such as SPR, although the existing body of evidence is quite conclusive as far as substrates of the *E. coli* Mla system is concerned.

We have also observed the formation of a non-physiological interaction between two MlaD hexamers as a result of the procedure of crosslinking the stabilised complex to prevent dissociation at low concentrations during EM grid preparation. This assembly of MlaD in the crosslinked sample was not observed by AUC of the stabilised complex

prior to crosslinking. However, AUC of BDO-CL bound MlaD does show evidence of a population of molecules that sediment with a Mw of slightly more than two times the Mw of the MlaD hexamer. We speculated that, considering the binding location of lipids in the cryo-EM structure of MlaFEDB presented by Thong et *al.*[184], dimerization of MlaD hexamers in the presence of cardiolipin may result from exposed acyl chains at the bottom of the complex encouraging dimerization through the sharing of acyl chains between two hexamers. If this is the reason for the formation of this back-to-back MlaD assembly then it is worth considering why we no longer observe this assembly of MlaD after exposure to MlaC. If the binding location of lipids to a soluble construct of MlaD is consistent with their location in published structures of MlaFEDB and we observe the reformation of dimerised MlaD hexamers only in the absence of MlaC after size exclusion it may suggest that the exposed acyl tails at what would be the MlaD/MlaE interface are not present when MlaC is bound. Speculatively, this could suggest that binding of MlaC causes lipids to move further up the MlaD binding channel, however, as we did not observe any map density we could convincingly attribute to a lipid in the cryo-EM model this remains to be speculation.

While the map generated from cryo-EM of the stalled MlaC–MlaD complex was of sufficient resolution to unambiguously position both MlaD and MlaC, the modelled positions of side chains are predominantly speculative. Taking that into consideration, we initially focussed on interacting regions of opposite surface charge to identify patches that may be core to the initial binding event, in accordance with known tenants of transiently interacting proteins. The binding groove formed by the α4, α5 and lower α6 helices of MlaC appear to have surface residues with low conformational variability according to the crystal data collected and presented in the previous chapter. Even residues that often adopt multiple orientations such as Arg have defined density for single orientations at 2σ in this region. While this may not necessarily be reflective of the conformational variability of the protein in solution it is

still worth acknowledging that residues in this region may have properties associated with anchor residues as described by Rajamani et *al.*[140]. The β6 – β7 loop of MlaD, however, is disordered in the crystal structure of the soluble construct. Regardless, we speculate that residues in this region are core to the binding event and alterations to the surface charge of this loop may significantly reduce the affinity of this binding interaction.

We also expected, based on the position of the opening of the MlaC lipid binding pocket that exchange of lipids between MlaC and MlaD likely occurred through the central helix of MlaD or between the helices. The structure of MlaC we report as part of the stabilized complex shares the greatest conformational identity with structures in the native conformation. As it was generated with an apo preparation of MlaC, we looked at residues interacting with loop regions that differentiate the -apo and -holo conformations in hopes of determining the residues involved in transitioning to the open conformation. As a 4.46 Å resolution does not permit the accurate determination of atomic positions we focussed more on regions of the model where at an arbitrary contour level we began to see points of contact between map density associated with MlaC and density associated with MlaD. We took specific note of the residues identified by Ercan et *al.* which we discussed in the introduction of this chapter[53]. Considering the proposed importance of Val171 on the β4 – β5 loop of MlaC, and the significance of this loop in the transition between the -apo and -holo states, we expected residues interacting in the region of Val171 to be in some way significant to the mechanism of exchange. We noticed that Leu151 and Tyr152 of $MlaD_6$ in conjunction with local residues from MlaC, notably Met173, form a hydrophobic bridge between the binding pocket of MlaC and the central channel of MlaD [Figure 4.24]. As mutation of Val171 to *p*Bpa had no significant effects on cell viability on SDS/EDTA plates, it stands to reason that any hydrophobic residues in this position is sufficient for activity. We believe that while the Val171 residue contributes to the

mechanism of exchange, it is likely not the only residue in this region involved in the interaction. Ercan et *al*. had identified several other residues at this interaction interface in their investigation. As they had previously identified, it would appear from our model that Glu169 of MlaC as well as Gln149 and Tyr152 from MlaD play some role in the interaction in this region. We believe that Glu169 interacts with Gln149 of $MlaD_1$ and this is the core interaction that pulls the β4 – β5 loop into the open position. However, the SDS/EDTA sensitivity phenotype of a *p*Bpa substitution at Glu169 is not reflective of the same substitution at Gln149, as you would expect it to be if these residues are an interacting pair. That said, the proximity of these residues in our model suggests they interact. It is possible that the due to its physiochemical properties, *p*Bpa substitution was capable of approximating the hydrogen bonding function of Glu in Ercan et *al*.'s investigation and thus substitution of Glu with *p*Bpa was not as impactful on SDS/EDTA resistance as the mutation to this putative interacting pair in MlaD. The Tyr152 residue of MlaD potentially forms interactions in both the $MlaD_1$ and $MlaD_6$ monomers. Lys177 of MlaC is positioned to potentially hydrogen bond with Tyr152 of $MlaD_6$, while Arg143 is positioned to potentially hydrogen bond with Tyr152 of $MlaD_1$. We speculate that these interactions may be involved in any positional rearrangements of the central helix bundle of MlaD.

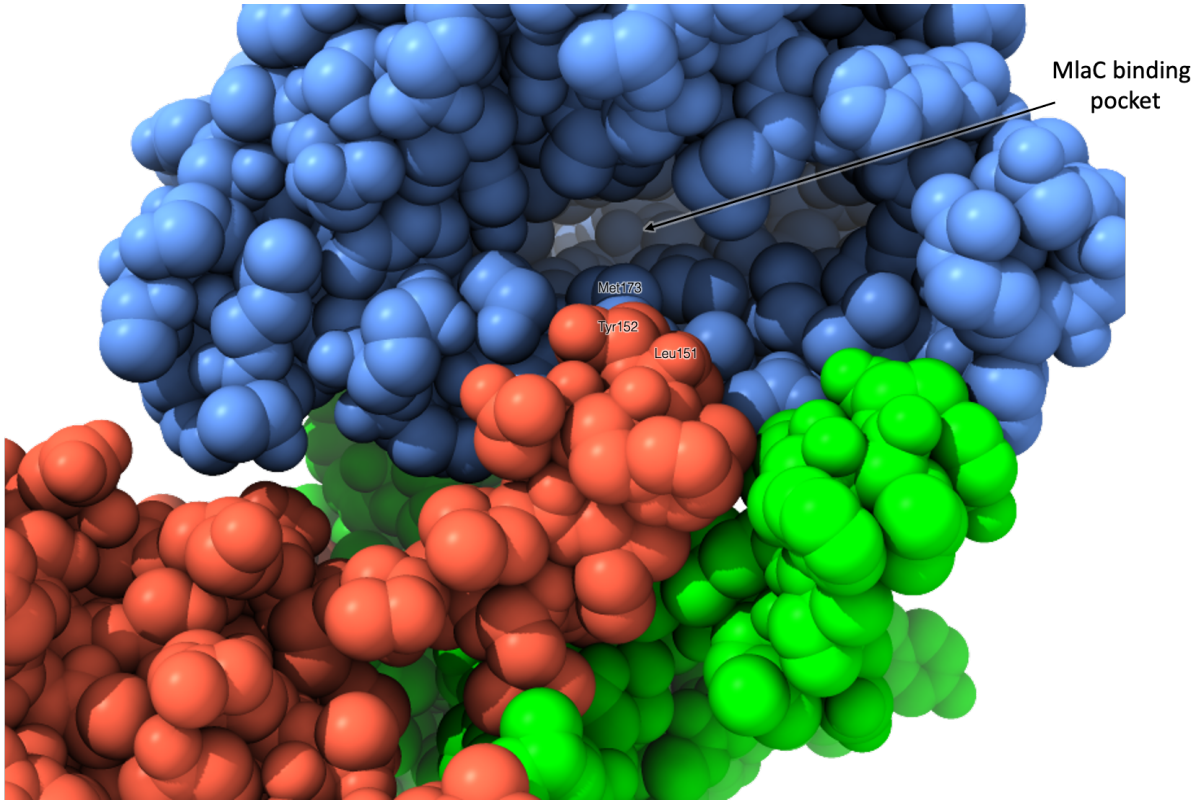

Figure 4.24 - The lipid binding pocket of MlaC (blue) during the MlaC-MlaD interaction. The residues (Met173 of MlaC along with Leu151 and Tyr152 of MlaD) involved in forming a hydrophobic bridge between MlaC and $MlaD_6$ (red) have been labelled. $MlaD_1$ (green) is also shown. The other monomers of MlaD have been hidden.

The other amino acids of note presented by Ercan et *al.* in this region, those being the Ser172 of MlaC and Met141 of MlaD, appear to be involved in interactions between the periplasmic facing top of MlaD and the interacting surface of MlaC formed from β3, β4, β5, α4, α5 and α7[53]. Ser172 is positioned to potentially interact with Lys138 of MlaD and it would appear that this is the only residue in the region that would be capable of photo-crosslinking if Ser172 were to be substituted for *p*Bpa in our model. Regarding the Met141 residue of MlaD, it does not appear that there are any nearby aromatic residues and thus no possible S-aromatic interactions. Potentially, burial of this methionine in a hydrophobic region is another component of stabilising the interaction. Methionine is not often surface exposed due to its highly hydrophobic properties conferred by the terminal methyl group and thus its burial in the binding

event and subsequent conformational change of MlaC to the open state may be enthalpically favoured.

Ercan et *al.* also identified several residues not associated with binding at the central helix bundle[53]. Phe118 from MlaD and Pro124 from MlaC also crosslinked when substituted with *p*Bpa. Evidence of interactions between the proteins involving these two residues provide insight into the interaction that is not immediately evident from our model. Phe118 is present on the β6 – β7 loop of MlaD, which we have previously suggested may be involved in the initial binding event. The Pro124 residue of MlaC, however, is far from any clearly interacting region with $MlaD_1$. The only nearby residues belong to the clockwise adjacent monomer of MlaD, henceforth referred to as $MlaD_2$. Pro124 being involved in the interaction would suggest that residues from the β6 – β7 loop of MlaD are not only involved in binding into the MlaC pocket created by the α4 helix but also coordinating the position of the β1 – β2 MlaC loop. Functionally this would sandwich the entire β-sheet region of MlaC between the β6 – β7 loops of $MlaD_2$ and $MlaD_1$. This provides evidence that 3 MlaD monomers are involved in the binding interaction with MlaC and explains why we see only 1 or 2 MlaC molecules binding per MlaD hexamer. This is also more in line with the symmetry constraints imposed on MlaD by the FEB components of the FEDB complex, which impose 2-fold symmetry restraints on the otherwise 6-fold symmetrical hexamer.

Considering that the interactions proposed in this investigation are primarily speculative, determining the accuracy of these speculations should form the basis of follow-up investigation. Acquiring higher resolution structural data could contribute to a much more accurate understanding of the interaction. The main roadblock to high resolution in our dataset is sample inhomogeneity. The dimerization of our sample as well as variability in MlaC binding stoichiometry is a large source of inhomogeneity. The issue here is that the back to back dimerization of our complex was a major

contributing factor to resolving our issues of preferential orientation. There is a possibility that, as the FEDB complex has been resolved to in the best case sub 3 Å, preparing a stalled complex with the entire FEDB complex could solve both issues. Although, generating a stalled complex of MlaFEDB in the same way we generated a stalled complex of MlaD alone may present its own issues.

5.Results: Investigation of Mla Mutants

5.1.Introduction

The fundamental aim of this chapter was to validate speculations we had made regarding the residues significant to the MlaC-D interaction in the previous chapter. To do this we determined that it was necessary to develop a more sensitive assay for lipid exchange between MlaC and MlaD. As such, we also aimed to develop a rapid fluorescence assay for lipid exchange that was both capable of identifying potential changes in the rate of lipid exchange as a result of modifying key residues and functional as a screen for inhibitors in the future.

5.2.Introduction to Methods

5.2.1.Microscale Thermophoresis

Microscale thermophoresis (MST) is a technique for investigating the binding interaction of biomolecules in a liquid suspension. The technique leverages the difference in thermophoretic properties of a molecule in its bound and unbound state. Liquid state thermophoresis was a phenomenon originally observed in 1856, and describes the diffusion of particles across a temperature gradient[103]. In the case of positive thermophoresis it is observed that molecules exhibit a net force away from the highest concentration of temperature owing to the increased average momentum of particles at higher temperature compared to those at lower temperature. The rate of thermophoresis of a particle is dependent on, given constant buffer conditions, its size, charge and solvation entropy and this allows the thermophoretic effect to be a determinant of ligand binding. MST uses fluorescent tags to track thermophoresis as a function of fluorescence intensity.

The experimental setup involves a solution of ligand and fluorescently tagged target in a glass capillary[75]. A small section of the capillary, approximately 50 μm in diameter, is irradiated with an infrared laser creating a temperature gradient of approximately 2 – 6°C. Throughout this process the fluorescent tag is exposed to light at its excitation wavelength. The intensity of the fluorescent emissions from the region exposed to the infrared laser correlates with the number of tagged molecules in that region and is thus used as a measure of thermophoresis. There are several stages in the thermophoresis trace as a result of using fluorescence in this manner. Most notably is the fluorescent response observed immediately after the sample is exposed to the infrared laser, which causes a phenomenon referred to as the temperature jump or T-jump. The T-jump is an immediate change in the fluorescence intensity of the fluorophore associated predominantly with a change in its photo-physical properties in response to an increase in temperature. Typically heating by the infrared laser is carried out for 30 seconds before it is deactivated, the fluorescent changes associated with T-jump occur within the 1 – 2 seconds of heating. The magnitude of the change in fluorescence intensity is dependent on the local environment of the fluorophore and thus changes in T-jump can be an indicator of a binding event. The remainder of the fluorescent changes are due to the thermophoretic effect. When the infrared laser is turned off, a reverse T-jump occurs followed by “back-diffusion” of particles into the heated region as it cools [Figure 5.1]. MST can be used as a technique to probe K_D given a set of solutions wherein only the concentration of ligand varies. Differences in the population of bound vs free tagged target molecule will be reflected by differences in the rate and magnitude of fluorescence change at both the T-jump stage and the thermophoresis phase.

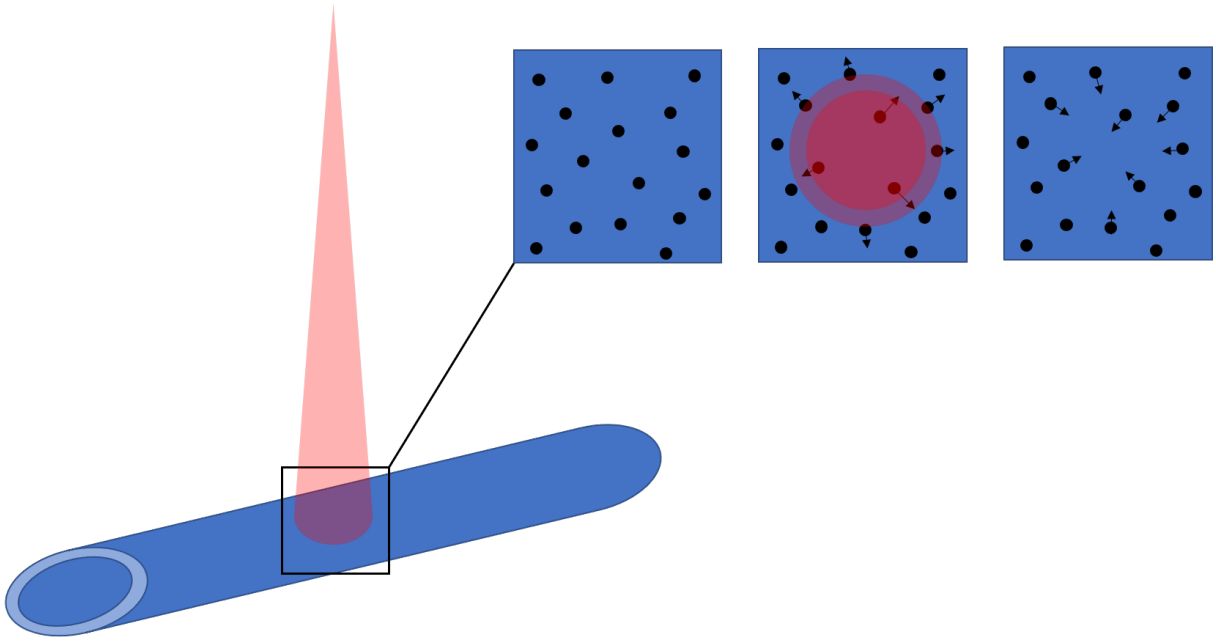

Figure 5.1 - MST relies primarily on the thermophoretic effect, wherein a temperature gradient causes a net motion of macromolecules in solution from a region of higher temperature to an area of lower temperature due to their higher average momentum. Particles that are larger or complexed tend to move less. Thus it is possible to report on complexation in a solution as a function of fluorescence depletion upon heating if the particles are fluorescent or fluorescently tagged.

5.2.2.Förster resonance energy transfer

Förster resonance energy transfer (FRET) is a technique based on the 1948 observations of Theodore Förster, wherein energy can be non-radiatively transferred between two chromophores[57]. FRET does not rely on fluorescence for energy transfer, but instead energy is transferred via dipolar coupling. If both the donor and acceptor chromophores are fluorescent, emissions of the acceptor chromophore can be used as a reporter of intermolecular distance.

FRET can occur when the energy gap of the excited donor chromophore is equivalent to the excitation energy gap of the acceptor [Figure 5.2]. Practically, this is the case for fluorophore pairs where the excitation wavelength range of the acceptor overlaps

the emission range of the donor. FRET efficiency is heavily reliant on the separation distance, r, between the donor and acceptor and falls off proportional to r^{-6}, with a maximum Förster distance of about 10 nm for most donor-acceptor pairs[98].

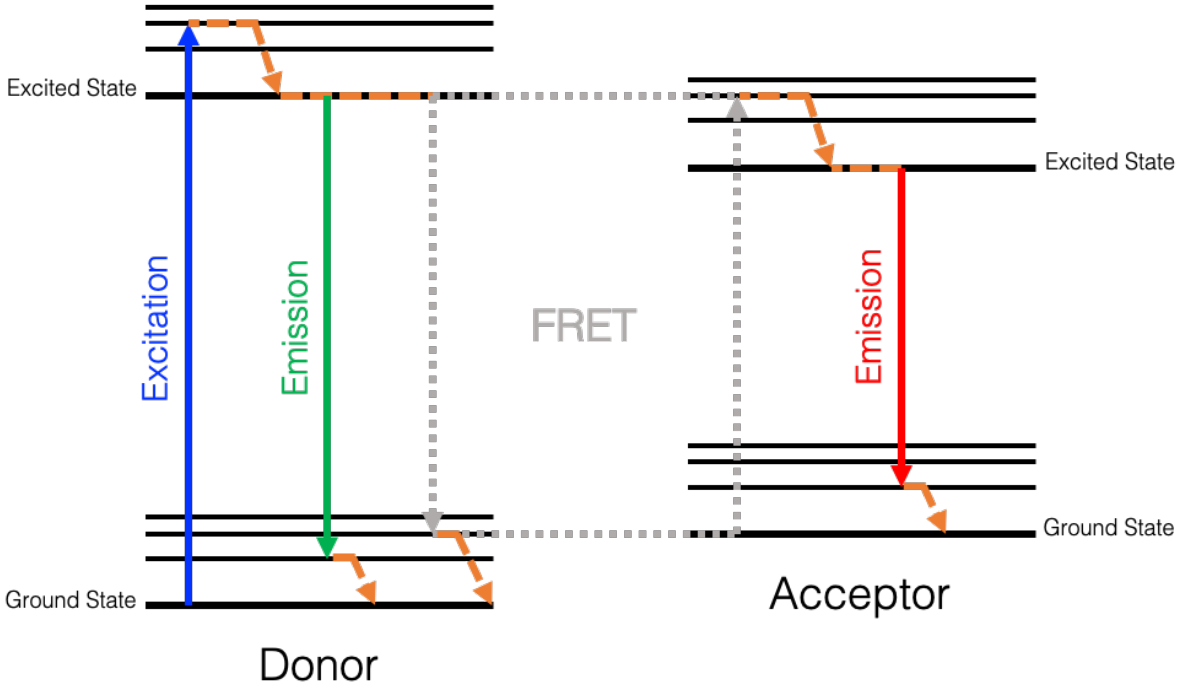

Figure 5.2 - Jablonski diagram demonstrating the requirements for FRET. Absorption of a photon of light at a specific wavelength results in electron excitation from the ground state to the excited state followed by subsequent vibrational relaxation (orange) within the same excited state. At this point to return to the ground state there must be some release of energy. This can be in the form of an emitted photon but if there is a nearby acceptor fluorophore de-excitation to the ground state can instead be coupled to excitation of the acceptor fluorophore through a dipole interaction. The acceptor fluorophore then returns to the ground state by the emission of a photon.

5.3.Results

5.3.1.A Fluorescence Assay for Lipid Exchange

Our method for determining the ability of MlaD to transfer lipids to MlaC in prior chapters has been by extracting lipids at the endpoint of lipid exchange and imaging them by TLC. We found this method to be sufficient for determining whether lipid transport was occurring, however, it would not be appropriate for determining the rate of transport. To study transport within the Mla pathway, we investigated the potential of utilising a FRET based assay to directly monitor PL transport rates. We focused on the MlaC/MlaD interaction. We wanted to use a FRET system that does not require any tag or fluorophore to be attached to the protein. As such the donor fluorophore was predetermined to be the Trp residues intrinsically incorporated into our proteins. Several papers have previously employed biological molecules functionalised with 5-(dimethylamino)naphthalene-1-sulfonyl (dansyl) fluorophores as an acceptor molecule for native Trp fluorescence. The dansyl fluorophore is reactive with primary amines and thus capable of tagging a wide variety of molecules. A 1998 paper by Cordat et al. used a dansyl functionalised melibiose disaccharide to determine the proximity of Trp residues in *E. coli* MelB from the melibiose binding site[32].

The dansyl fluorophore has an excitation peak at 335 nm and an emission peak at 518 nm, making it an ideal acceptor for Trp, which has an excitation peak at 280 nm and emission peak generally at 350 nm. Dansyl labelled fluorescently labelled PL headgroups are also readily available (Avanti). Futhermore as the soluble construct of MlaD we generated lacks any Trp residues and MlaC has 3 Trp residues, our hypothesis was that any PL exchange of the dansyl tagged lipid from MlaD to MlaC will come with an increase in FRET emissions.

We expect that dansyl-Trp FRET may also function as an assay for lipid exchange even in the case of the complete FEDB complex as the system of dansyl tagged lipids moves from being distributed randomly in a liposome at various distances from a Trp residue to being bound to MlaC at a fixed sub-4 nm distance from multiple Trp residues [Figure 5.3]. We believe that this can function as a rapid and effective assay to track and timescale lipid exchange between MlaFEDB and MlaC. If this assay shows comparable results to those presented by Tang et *al*. regarding the activity and transport directionality it may be useful as an assay for rapid in tray inhibitor screening without the need to set up a two-liposome system[182].

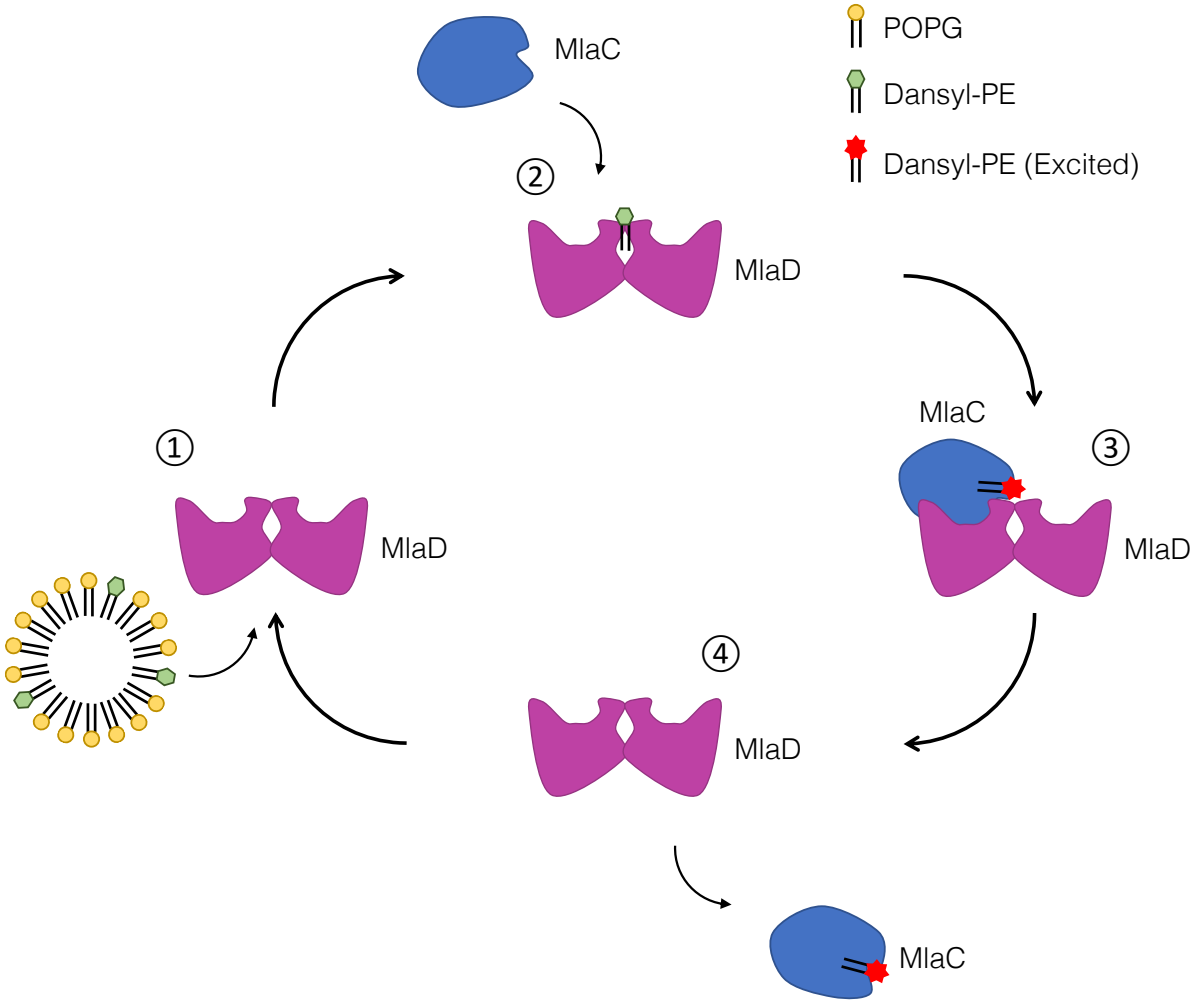

Figure 5.3 - Dansyl-PE assay for binding and lipid exchange activity between MlaC and MlaD. (1) Dansyl-PE lipids begin in liposomes, and are sequestered by MlaD. (2) When bound to MlaD, the lack of Trp residues result in very low levels of FRET emissions. (3) Binding of MlaC results in lipid exchange. When bound to MlaC, dansyl-Trp FRET results in fluorescence emissions. (4) MlaC bound to dansyl-PE dissociates and continues to emit at 518 nm. MlaD is then free to bind more lipids.

We began by determining the excitation and emission maxima of the fluorescent molecules in the conditions of the assay. We observed no significant fluorescence peak deviations in the spectra of the dansyl fluorophore, with excitation/emission peaks at 337/518 nm, which is in line with the expected fluorescent peaks from published literature [Figure 5.4]. The fluorescent spectra of MlaD has excitation/emission peaks at 280/300 nm and the spectra of MlaC has peaks at 280/350 nm. MlaD has an emission peak at 300 nm consistent with Tyr and Phe fluorescence, while MlaC has an emission peak in the region of 350 nm consistent with Trp fluorescence with low solvent exposure. The 300 nm peak associated with Tyr and Phe fluorescence is most likely masked by the wide 350 nm Trp fluorescence peak in the MlaC spectra.

While there does appear to be some spectral overlap between the emission wavelengths of MlaD and the excitation wavelengths of dansyl-PE, the peak excitation of dansyl-PE is substantially different from the peak emission of MlaD. FRET may occur between MlaD and dansyl-PE but the FRET efficiency of this FRET would be substantially lower than the FRET efficiency of the exchange between MlaC and dansyl-PE. This means that any exchange of dansyl-PE from MlaD to MlaC would still result in an increase in FRET emission, as quantum yield is proportional to the overlap integral of the normalised excitation/emission spectra of the donor/acceptor fluorophore.

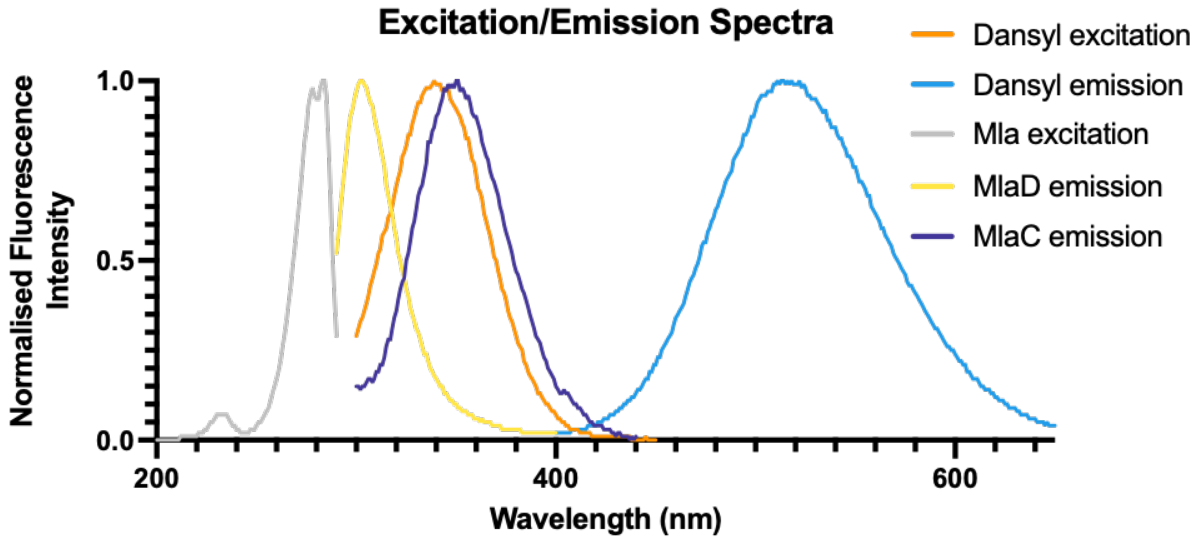

Figure 5.4 - Superimposed excitation/emission spectra of MlaC, MlaD and dansyl-PE. The excitation spectra of MlaC and MlaD overlay closely and thus a representative excitation peak generated from MlaD has been used. Spectra have been normalised to fraction of peak fluorescence for purposes of comparison between spectra.

Next we investigated the baseline FRET emissions of a system with only MlaD and fluorescent lipids and compared it against a system with MlaC present. We have substantial evidence showing that lipids are readily transferred to MlaC (Chapter 3) and expect a time dependent increase in FRET emissions as more lipid is transferred from MlaD to MlaC and consequently to within the Förster distance of the Trp residues in MlaC. To determine if there was a time dependent increase in FRET emission we collected FRET emissions in the 400 – 550 nm range as a time course following addition of MlaC [Figure 5.5]. We saw a clear increase in fluorescence at the 518 nm peak over the 120 minute time-course from approximately 57,000 AU to approximately 83,000 AU. It is reasonable to assume that the baseline FRET observed in the presence of only MlaD is attributed to the spectral overlap of the MlaD emission peak and the dansyl fluorophore excitation peak allowing for some amount of baseline excitation. The main purpose of collecting these spectra was to ensure that any increase in the peak at 518 nm was attributed entirely to the dansyl emission peak

and not due to any spectral overlap from the Trp emission peak at 350 nm. It is reasonably clear that emissions from the 350 nm peak do not have significant overlap with the dansyl emission peak.

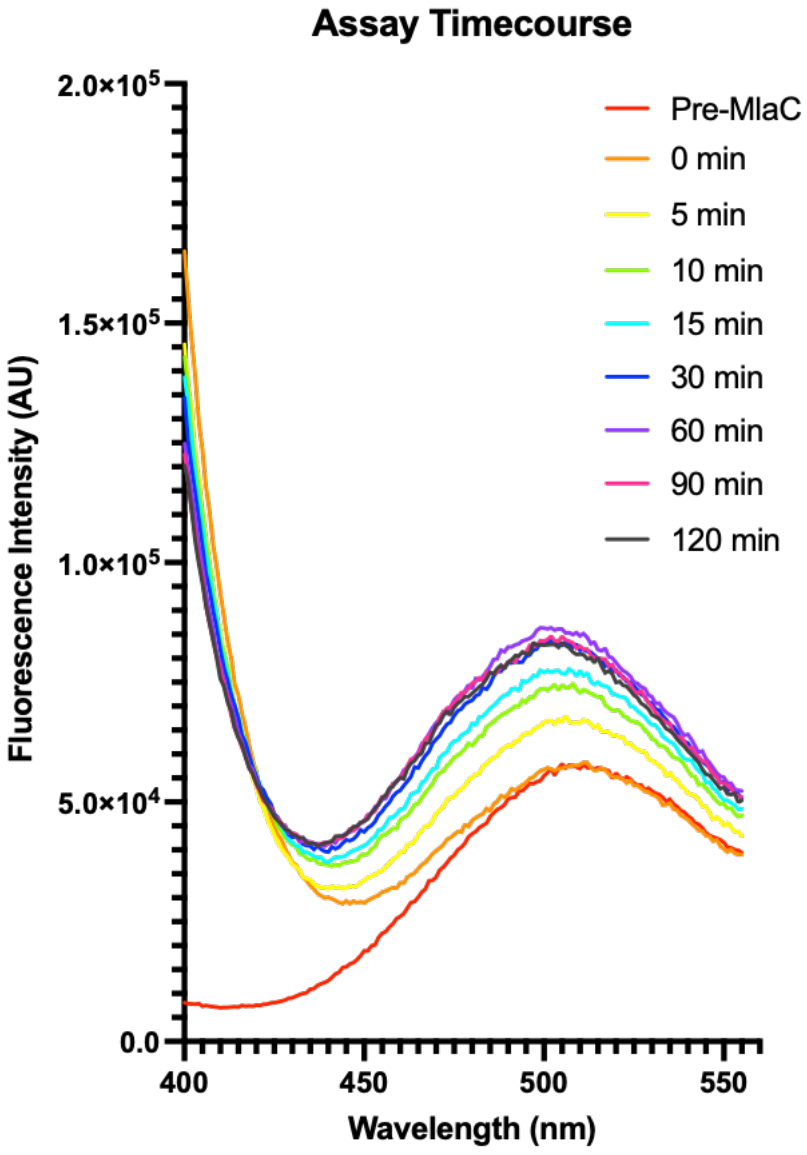

Figure 5.5 - Fluorescence spectra showing the FRET emission peak of dansyl-PE before the addition of MlaC as well as at various time-points afterwards. We see a clear increase over baseline fluorescence at time-points after the addition of MlaC with no significant interference or overlap between the MlaC emission peak and the dansyl-PE emission peak.

We next carried out time-course lipid exchange assays using FRET emission at 518 nm as a reporter of lipid exchange between various preparations of MlaC and MlaD. We expected that a natively purified sample of MlaC would be unable to accept lipids from MlaD due to the lipids already bound to the protein. We thus compared lipid exchange between MlaD and MlaC-apo with lipid exchange between MlaD and MlaC-native, expecting MlaC-native to act as a negative control showing no lipid transfer. We also included relevant MlaD and MlaC negative controls [Figure 5.6].

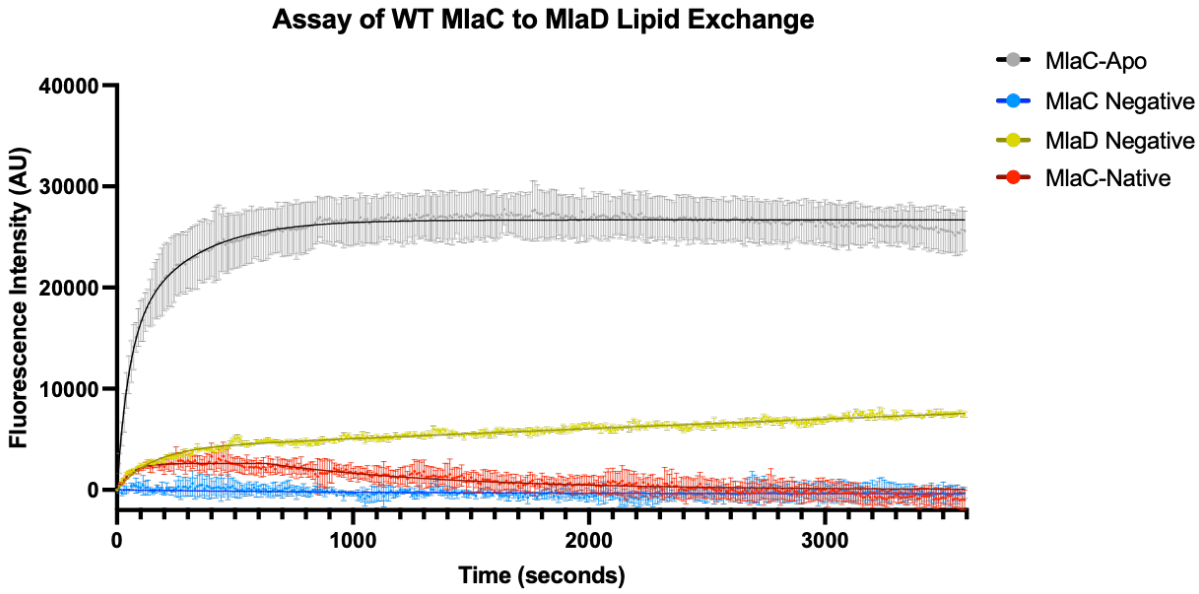

Figure 5.6 - Time course assay of lipid exchange from MlaD to MlaC as well as negative controls. An assay with MlaC-native was also included to show that fluorescent change does not occur when MlaC is not capable of accepting fluorescent lipids from MlaD. Each assay was repeated 3 times and the mean fluorescence change from baseline was plotted at each timepoint, with error bars showing standard deviation. The rate of fluorescence change for each assay has been modelled as a function of time using a two-phase association model or a liner model where appropriate.

We predicted the initial rate of fluorescence increase in the MlaC-apo exchange assay to be significantly higher than the negative controls. As expected, we saw an initially high rate of linear fluorescence increase over the first ~100 seconds at which point the rate of fluorescence change began to plateau, reaching a stable plateau at approximately the 900 second mark. We expect that this plateau represented either

the saturation point of MlaC-apo in the exchange or a stable equilibrium of backwards and forwards exchange. Both the MlaC negative and MlaC-native control assays showed no net increase in fluorescence over the full time-course of the assay. However, the MlaC-native sample exhibits a minor initial shift in fluorescence and a biphasic curve over the assay time-course. The fluorescence change is minor compared to the Mla-apo assay and likely due to a small population of MlaC-apo in the MlaC-native prep.

Interestingly we saw a small linear increase in the MlaD negative sample, that did not plateau over a period of 3600 seconds. This seems to suggest that MlaC-apo is capable of picking up some lipids in the absence of MlaD. Rate wise, this is somewhat contrary to the observations of Knowles et *al.*, who showed by NMR that MlaC-apo exhibits less than 10% lipid uptake over a period of 24 hours[71], however, differences in sample concentration likely has some effect on the rate of uptake in the absence of MlaD.

We see an initial fluorescence change of approximately 229 ±16 AUs^{-1} for the MlaC-apo assay. Comparatively over the same range the MlaC negative assay has a rate of 2 ± 9 AUs^{-1}, the MlaC-native assay has a rate of 21±12 AUs^{-1} and the MlaD negative assay has a rate of 21 ± 5 AUs^{-1} [Table 5.1]. The MlaC-apo assay had a final fluorescence of approximately 26000 AU while the MlaC-native and negative assays had a final fluorescence equivalent to or slightly below the initial fluorescence value. The MlaD negative assay had a final fluorescence of approximately 7600 AU. The results, therefore, suggest that the assay could be used to directly monitor lipid exchange between MlaD and MlaC.

Lipid Exchange Assay	Initial Rate of Fluorescence Change (AUs^{-1})
MlaC-apo (positive assay)	229 ±16
MlaC-native (control assay)	21±12
MlaC negative (control assay)	2 ± 9
MlaD negative (control assay)	21 ± 5

Table 5.1 - Initial rates of fluorescence change for lipid exchange assays and relevant controls as determined from experimental results presented in [Figure 5.6].

5.3.2.Preparation and Purification of MlaC and MlaD Mutants

Having determined that the assay works to directly monitor lipid transport between MlaC and MlaD we moved on to assess the viability of a set of mutants we expected to modulate activity. We generated several MlaC and MlaD mutants we anticipated would reduce the effectiveness of spontaneous lipid transport between the two proteins, based on structural information we had previously elucidated (Chapter 4). We determined that mutations to residues 177 and 180 of MlaC may have an effect on its ability to effectively transport lipids due to the involvement of these residues in the transition from the -apo to -native states as well as their interactions with MlaD as discussed in Chapter 4. We mutated Lys177 to a Met residue, hoping to modify the electrostatic properties by replacing Lys with a similar but non-polar amino acid. By several metrics, such as the Sneath's dissimilarity index[171], Met is a close non-polar amino acid to Lys. In the case of Glu180, we also expected that a charged interaction with MlaD may be involved in lipid exchange. We elected to substitute this residue to Ala instead of a more similar residue like Gln to eliminate the potential for hydrogen bonding at this position. For MlaD we determined that mutations to residues 119, 120 and 122 may prevent or inhibit lipid exchange [Figure 4.19]. As we expected the

exchange interaction to be surface charge dependent, we made mutations with the goal of inverting the surface charge of MlaD. The following individual mutations were made; Glu119Lys, Asp120Lys and Glu122Lys as well as the Glu119Lys/Asp120Lys/Glu122Lys triple mutant. As we had also observed conformational change at the central helix bundle [Figure 4.21] we expected that preventing parting of the central helix would also have some effect on lipid exchange. As such we designed a mutant with cysteine residues substituted at positions 149 and 151 with the intention of crosslinking adjacent MlaD monomers of the central helix bundle and interrupting the transit of lipids through the gap. Finally we designed two mutants to currently unmodelled regions of MlaD that we suspect are likely to interact in some way with regions of MlaC based on the relative positions of modelled regions in the complexed structure presented in Chapter 4. We predicted that MlaD residues 155, 157 and 158 to have potential interactions with loop regions of MlaC [Figure 4.23] and thus, as with the mutations to Glu119, Asp120 and Glu122, we designed mutations aimed at inverting surface charge. We generated a single mutant, Lys155Glu, and a double mutant with Asp157Lys/Asp158Lys that we predicted may have an effect on lipid exchange. We speculated that an inversion of charge in this region may interfere with potential interactions between these residues and Arg143 of MlaC. This is based on the speculative positioning of Arg143 in the stabilised structure presented in [Figure 4.23].

Purification of the apo-mutants was carried out in the same way that the WT proteins were purified. Nickel affinity gels for the purification of MlaD mutants are presented below in [Figure 5.7]. It is worth noting that the un-boiled elution fraction of all mutants bar the double Cys mutant lack a characteristic band between 100 and 150kDa. As this band in a standard MlaD purification corresponds to a population of MlaD hexamers that are not fully unfolded by SDS and reducing conditions, the lack of this band may suggest that the mutations have somewhat destabilised the hexameric

arrangement in the presence of unfolding agents. Otherwise, the bands present between 15 and 20kDa appear as we have previously observed in purifications of the WT. Expression of the Lys155Glu mutant is notably lower than the other mutants purified, which is quite possibly due to oversights in codon selection regarding frequency bias, co-occurrence bias or pair bias when deciding to use the GAG codon for the Glu mutation. However, the lower rate of expression did not pose an issue to purification and did not appear to impact proper folding of the hexamer. Despite a lack of evidence of the hexameric structure in the SDS-PAGE, all mutants eluted from an S200 16 60 column at the same volume we have observed with the WT to be consistent with the hexameric configuration [Figure 5.8].

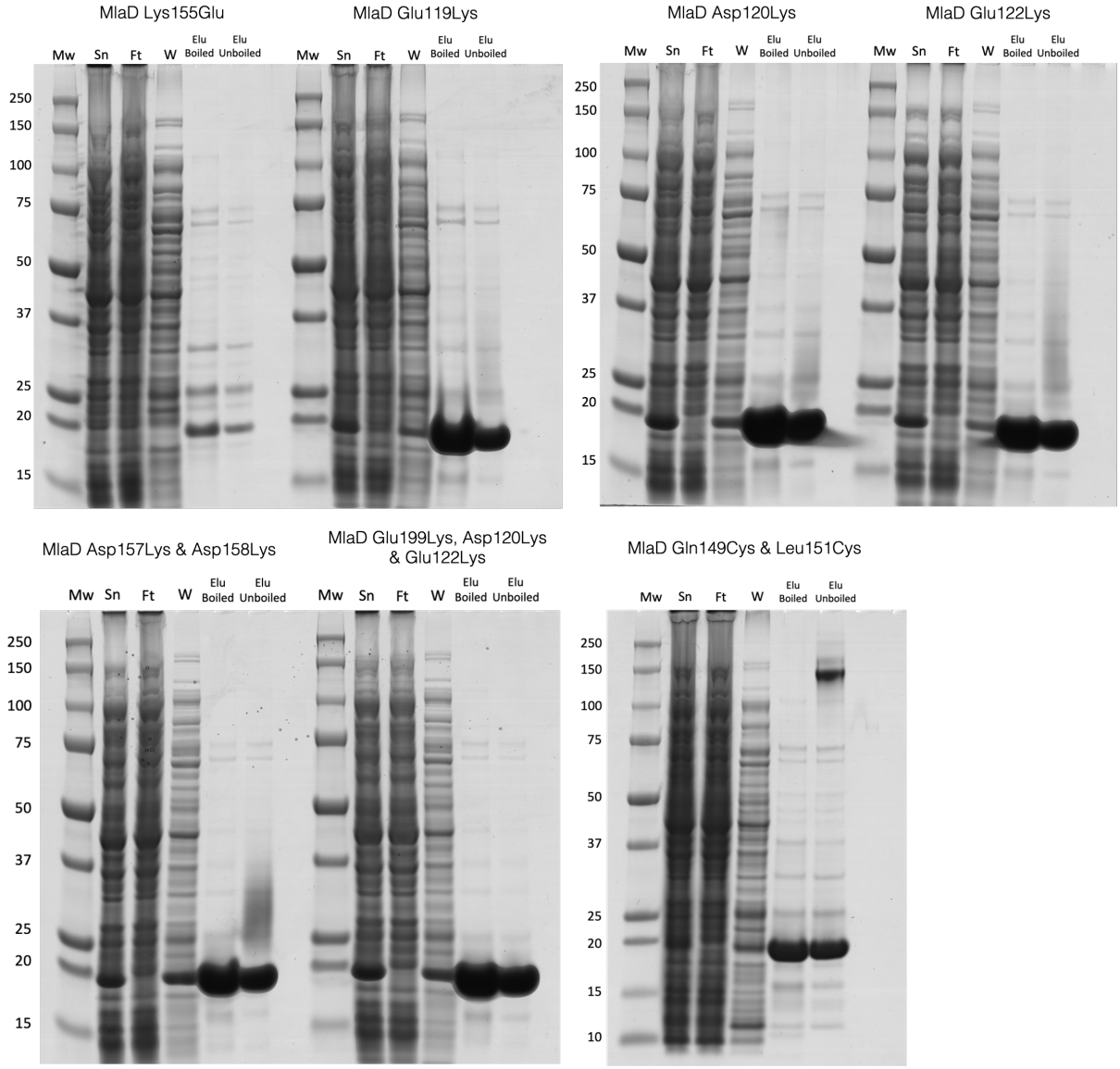

Figure 5.7 - SDS-PAGE of Ni-affinity purification fractions from MlaD mutant purifications. Samples from the pre-column supernatant (Sn), the column flowthrough (Ft) and a 50 mM imidazole wash (W) were run alongside a boiled (Elu Boiled) and un-boiled (Elu Unboiled) sample of the 500 mM imidazole elution. The elution fractions for all mutants contain a distinct band at 20 kDa, which is consistent with the MlaD band observed in previous purifications of WT MlaD.

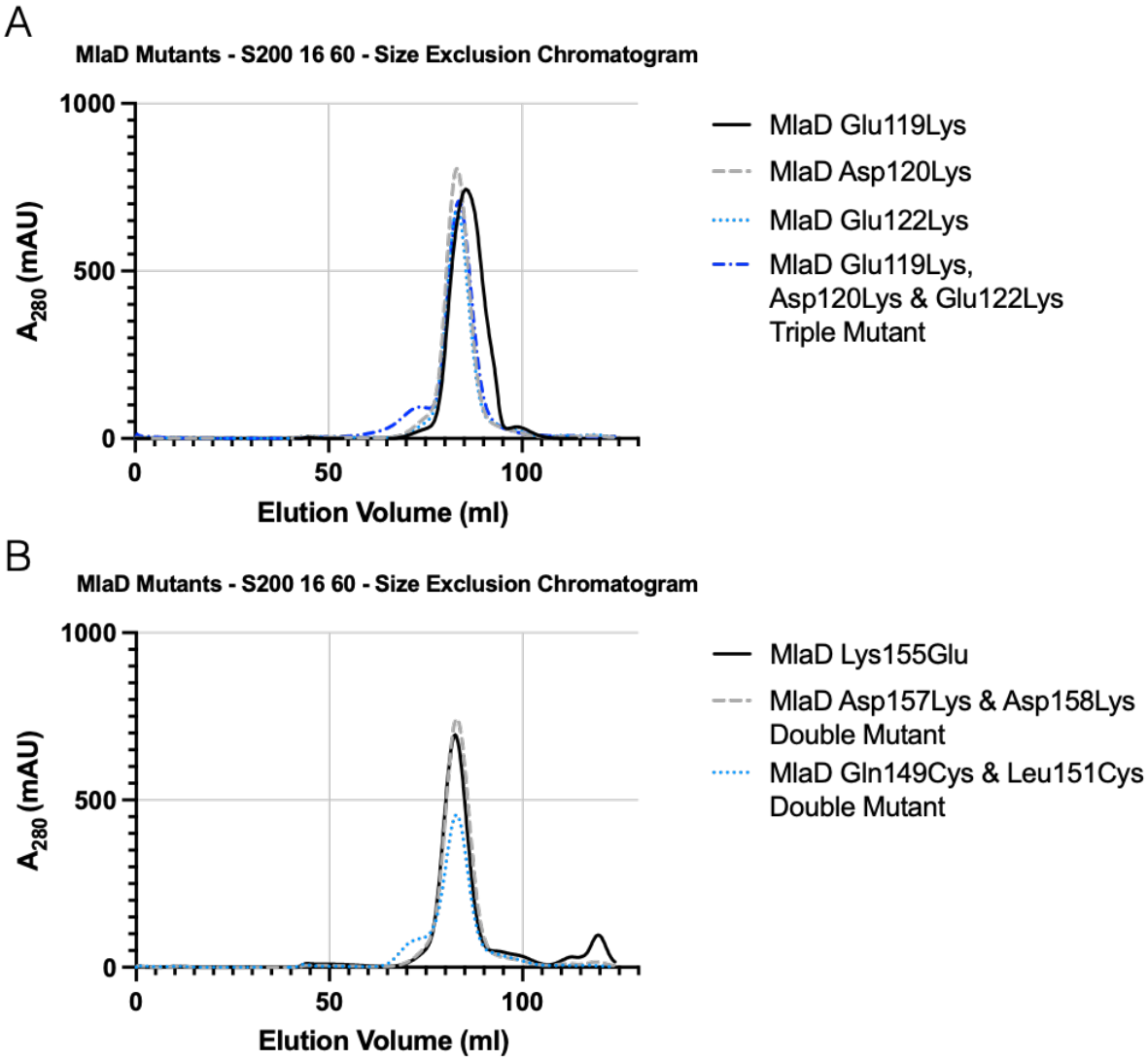

Figure 5.8 - SEC traces of MlaD mutant elution profiles, showing a single major peak for all mutants. All mutants eluted from an S200 16 60 column with a peak elution volume of approximately 85 ml, as we had previously observed with WT purifications. Traces have been split into two graphs to improve clarity of overlayed traces. (A) Shows SEC traces for the Glu119Lys, Asp120Lys and Glu122Lys single and triple mutants and (B) shows SEC traces for the Lys155Glu single mutant, the Asp157Lys & Asp158Lys double mutant and the Gln149Cys & Leu151Cys double mutant.

Nickel affinity gels for the purification of MlaC mutants are also presented below[Figure 5.9]. The major band in the elution fractions are within the range expected from MlaC WT and the mutants elute from an S75 16 60 column at the same volume as was observed with the WT [Figure 5.10]. All mutants purified with lipids bound, as shown in [Figure 5.11], suggesting retention of lipid binding capacity in all

mutants. All mutants were purified with lipids present and were made apo by the same method used for the WT protein.

As the purpose of the MlaD Gln149Cys/Leu151Cys double mutant was to cause disulphide bridging and prevent re-organisation of the central helix bundle during lipid exchange it was also necessary to confirm the formation of disulphide bonds. We determined that the simplest way to show correct formation of disulphide bonds between adjacent MlaD monomers in the hexamer assembly was to show that a boiled sample of MlaD runs as a hexamer in non-reducing conditions. [Figure 5.12] shows several samples of MlaD Gln149Cys/Leu151Cys run on SDS-PAGE in two different sample buffers. The sample run in a sample buffer lacking a reducing agent runs primarily as a band between 100 and 150 kDa even when boiled which is good indication of disulphide bridging.

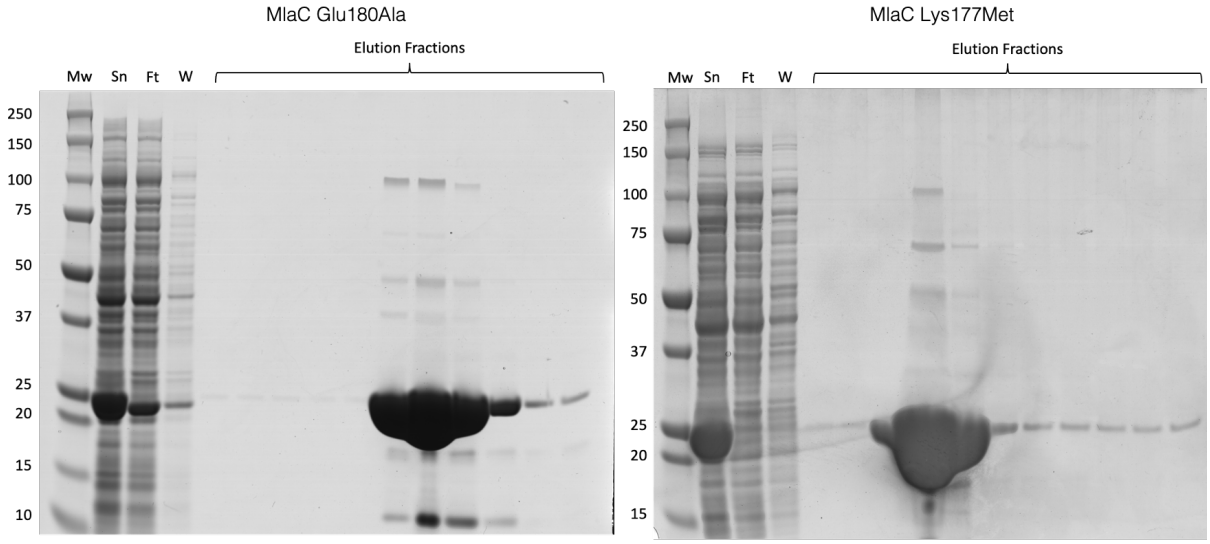

Figure 5.9 - SDS-PAGE of Ni-affinity purification fractions from MlaC mutant purifications. Samples from the pre-column supernatant (Sn), the column flowthrough (Ft) and a 20 mM imidazole wash (W) were run alongside fractions eluted with 500 mM imidazole. A major band with molecular weight between 20 and 25kDa is present in the elution fractions, which is consistent with the MlaC band observed in previous purifications of WT MlaC.

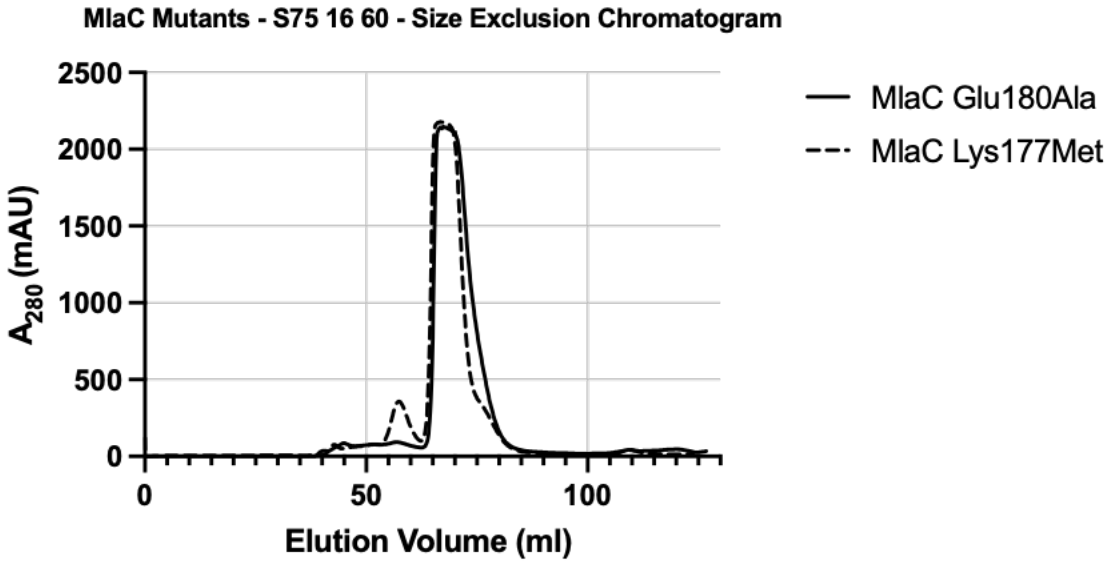

Figure 5.10 - SEC traces of MlaC mutant elution profiles. Both mutants elute in a single major peak. Minor contaminant peaks are visible in the MlaC Lys177Met mutant, however, they are distinct from the major sample peak and were avoided.

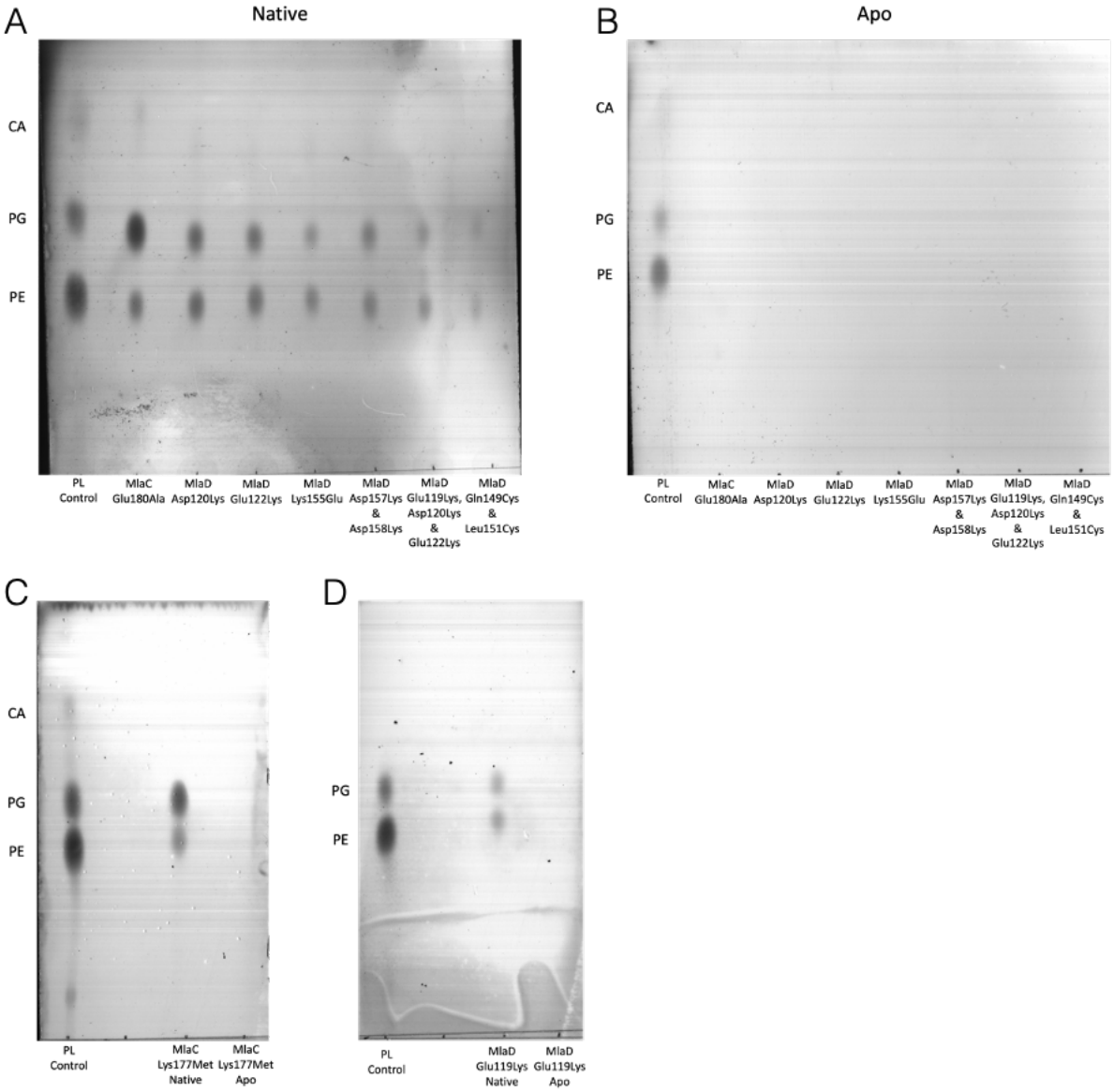

Figure 5.11 - TLC of lipids extracted from MlaC and MlaD mutants before (A) and after (B) a detergent wash. Due to size limitations of the TLC tank the MlaC Lys177Met (C) and MlaD Glu119Lys (D) mutants were run separately. TLC does not detect any lipids present after the detergent wash in any of the purified mutants. It is also worth noting that both MlaC mutants retain the preference for PG lipids observed in the WT.

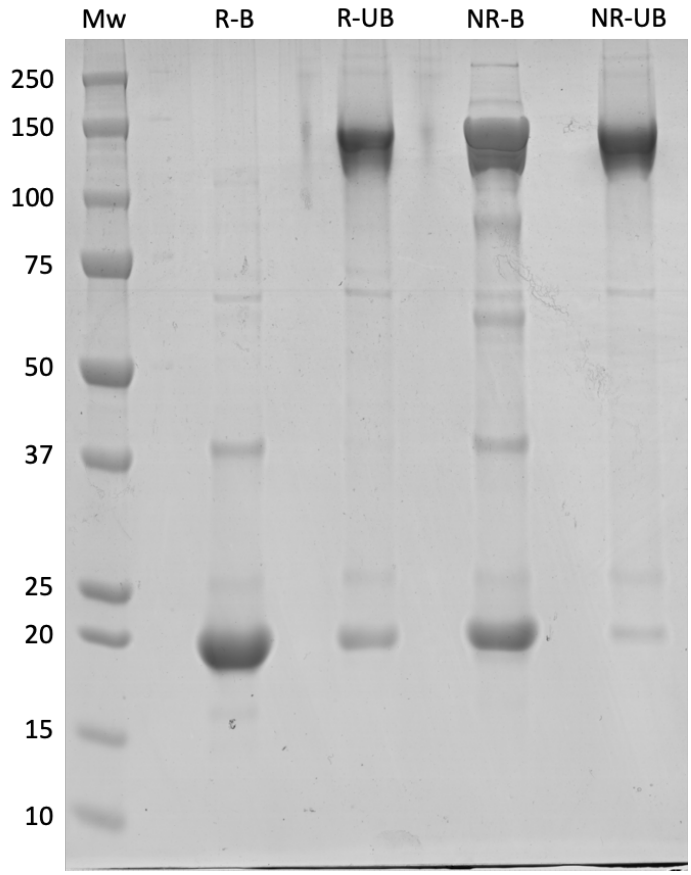

Figure 5.12 - SDS-PAGE of MlaD Gln149Cys/Leu151Cys run under 4 different conditions. Samples labelled -B were boiled at 95 °C for 10 mins prior to loading while the samples labelled -UB were not boiled prior to loading. Samples labelled R were run with a sample buffer containing the reducing agent β-mercaptoethanol, while samples labelled NR were run with a sample buffer containing no reducing agent. The boiled and reduced sample runs primarily as a monomer. The sample boiled without reducing agent runs primarily as a hexamer like the un-boiled samples suggesting disulphide bonding is stabilising the hexamer.

5.3.3.Assessing the Impact of MlaC and MlaD Mutants on Lipid Transport

Having established an assay to determine the rate of lipid exchange, we assessed the effect of mutations on lipid exchange *in vitro* by this method. [Figure 5.13] shows the rate of lipid exchange for the WT assay compared with MlaD mutants over a 3600 second period and [Figure 5.14] shows the initial rate over the first 100 seconds. We observed a degree of rate changes, from those that completely inhibited transport to

those that surprisingly seemed to enhance the rate of transport [Table 5.2]. Of the mutants, the ones that saw the most significant decrease in initial rate from the WT were the Asp120Lys, Glu119Lys/Asp120Lys/Glu122Lys, Lys155Glu and Gln149Cys/Leu151Cys mutants which have initial rates of 63 ± 9 AUs^{-1}, 2 ± 12 AUs^{-1}, 70 ± 30 AUs^{-1} and 72 ± 26 AUs^{-1} respectively compared to the WT rate of 229 ± 16 AUs^{-1}. The Asp157Lys/Asp158Lys mutant on the other hand saw an increase in rate over WT, with an initial rate of 398 ± 10 AUs^{-1}. The Glu119Lys and Glu122Lys mutants did not have initial rates significantly different from WT, with rates of 333 ± 108 AUs^{-1} and 208 ± 45 AUs^{-1} respectively.

Mutant Lipid Exchange Assay	Initial Rate of Fluorescence Change (AUs^{-1})
Control Assay	229 ±16
MlaD Glu119Lys	333 ± 108
MlaD Asp120Lys	63 ± 9
MlaD Glu122Lys	208 ± 45
MlaD Glu119Lys/Asp120Lys/Glu122Lys	2 ± 12
MlaD Lys155Glu	70 ± 30
MlaD Asp157Lys/Asp158Lys	398 ± 10
MlaD Gln149Cys/Leu151Cys	72 ± 26
MlaC Glu180Ala	21 ± 19
MlaC Lys177Met	214 ± 45

Table 5.2 - Initial rates of fluorescence change for lipid exchange assays run with MlaC and MlaD mutants as determined from experimental results presented in [Figure 5.14 & Figure 5.16]

Beyond the initial rate of exchange we saw differences in the steady state endpoint of the exchange assay when compared to the WT. While the WT holds a stable endpoint fluorescence, we observed a decrease in fluorescence over the full course of the assay in most of the mutants. The most significant of which is observed in the

Glu119Lys mutation, which over the course of the assay peaks and then drops back to baseline fluorescence. Of the remaining mutants investigated, only the Asp120Lys and Lys155Glu mutants verifiably hold a stable endpoint fluorescence like the WT. All other mutants show a peak followed by fluorescence decrease but this decrease is much less pronounced than the decrease observed in the Glu119Lys mutation. In the cases of the Glu122Lys and Asp157Lys/Asp158Lys mutants the decrease does not result in a reduction of fluorescence significantly below the WT end-point fluorescence. It should be noted that the magnitude of fluorescence change is different among mutants with the Asp157Lys/Asp158Lys mutant peaking at 44600 AU compared to the WT which peaks at 28200 AU. The significance and likely cause of this higher peak fluorescence will be addressed in the discussion. Overall, from the results presented here, it would seem that all mutations likely result in some form of dysregulation in lipid exchange aside from the Glu122Lys mutant, which does not significantly deviate from the WT fluorescence change at any point over the assay. The Asp157Lys/Asp158Lys mutant may have a positive effect on the ability for MlaC and MlaD to exchange lipids and all other mutants likely result in a reduced capacity to exchange lipids between the two proteins.

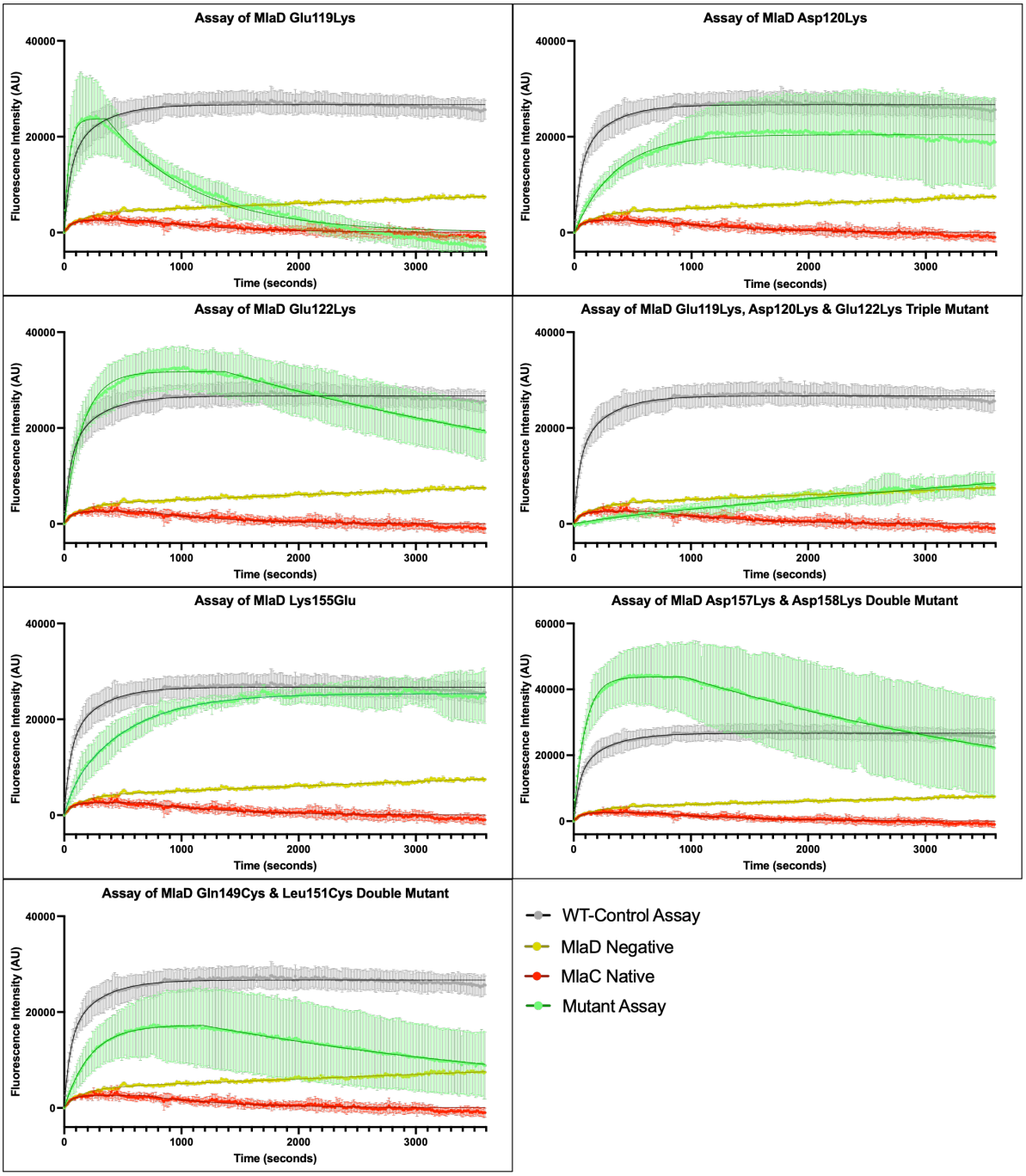

Figure 5.13 - Lipid exchange assays for MlaD mutants compared to WT. Each mutant assay is plotted on a separate graph alongside the WT assay and relevant control assays for comparison. Each assay was run in triplicate, and the mean fluorescence and standard deviation were plotted at each time-point. A model of fluorescence change as a function of time was also plotted alongside the observed data for each mutant as a trendline. The Asp120Lys and Lys155Glu mutant assays were modelled using a two-phase association model. The Glu119Lys/Asp120Lys/Glu122Lys mutant assay was fit to a linear model. The remaining mutant assays were modelled using an association/dissociation model.

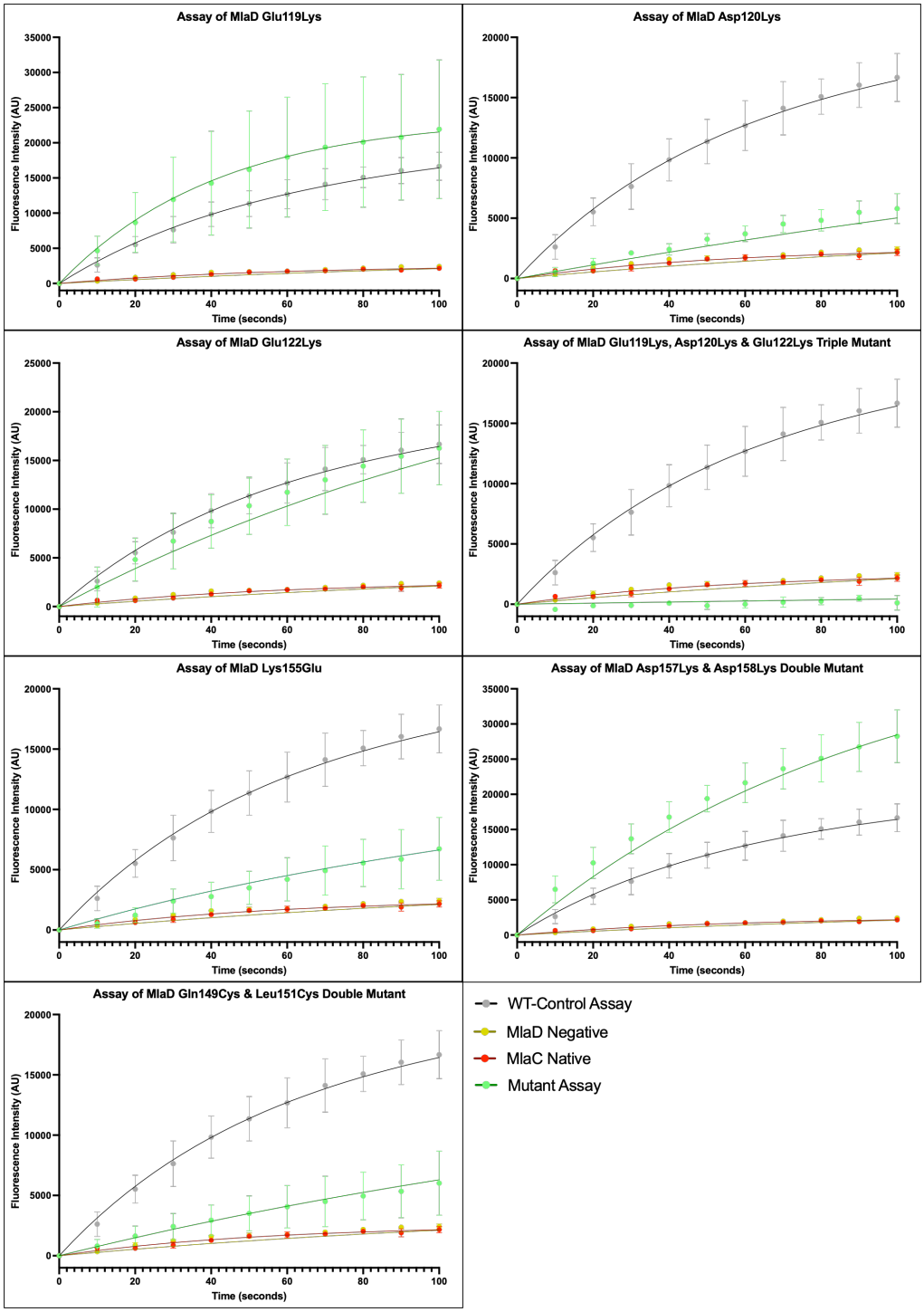

Figure 5.14 - Fluorescence assays for lipid exchange as presented in [Figure 5.13], showing only the time points between 0 and 100 seconds to better show the differences in initial rate of fluorescence change.

[Figure 5.15] shows the lipid exchange rates for the WT assay compared with MlaC mutants over a 3600 second period and [Figure 5.16] shows the initial rate over the first 100 seconds. The Glu180Ala mutant has a marked effect on initial transfer rate, reducing the rate of lipid transfer to a tenth of WT transfer rate, to 21 ± 19 AUs^{-1}. The Lys177Glu mutant on the other hand has an initial rate of 214 ± 45 AUs^{-1} which is not significantly different from the WT. However, much like the Glu119Lys MlaD mutant, the fluorescence does not reach a stable endpoint and falls back to baseline. From these results, it would appear that both these mutants have significant negative effects on lipid exchange.

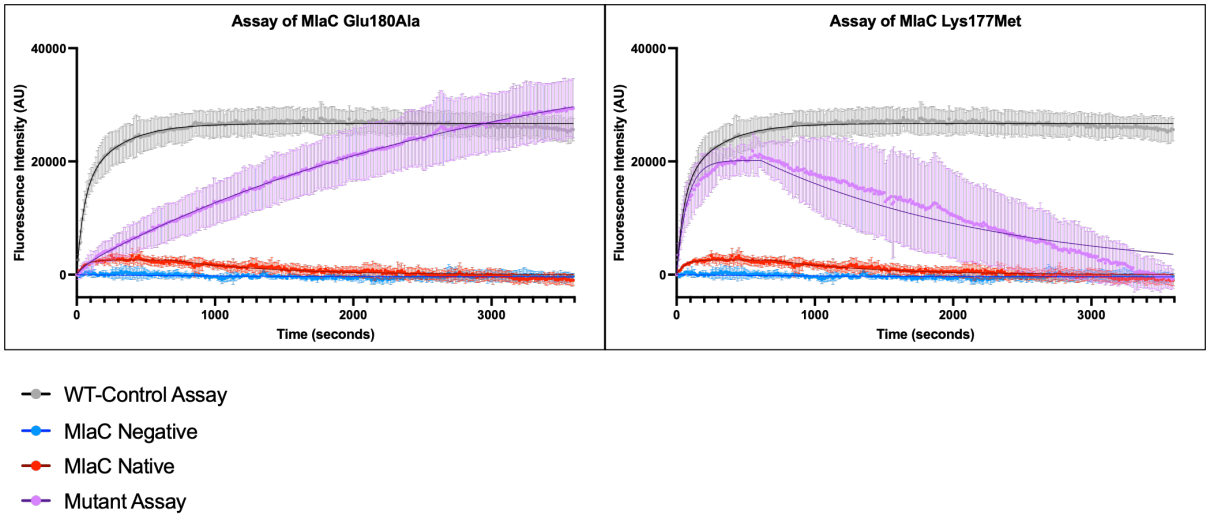

Figure 5.15 - Lipid exchange assays for MlaC mutants compared to WT. Each mutant assay is plotted on a separate graph alongside the WT assay and relevant control assays for comparison. Each assay was run in triplicate, and the mean fluorescence and standard deviation was plotted at each time-point. A model of fluorescence change as a function of time was also plotted alongside the observed data for each mutant as a trendline. The Glu180Ala mutant assay was fit to a two phase association model. The Lys177Met mutant assay was fit to an association/dissociation model.

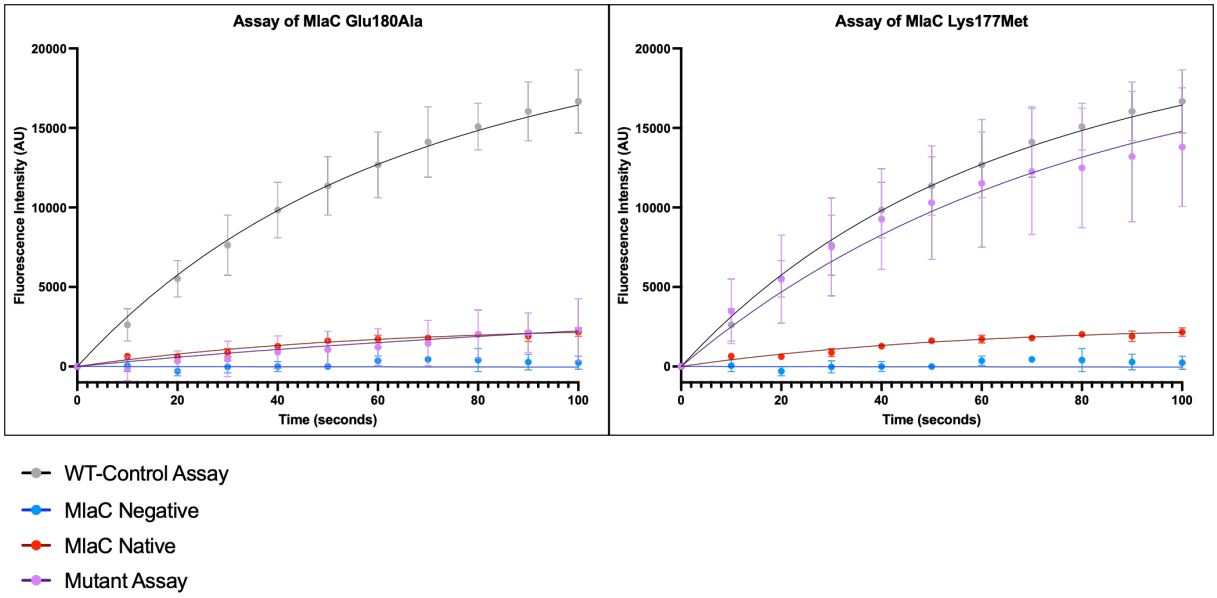

Figure 5.16 - Fluorescence assays for lipid exchange as presented in [Figure 5.15], showing only the time points between 0 and 100 seconds to better show the differences in initial rate of fluorescence change.

5.3.4.Kinetics of the MlaC to MlaD interaction

The assay also presented a new opportunity to further validate a result we had acquired previously. We were sceptical about the kinetic data we had previously acquired [Figure 4.3] that suggested the affinity of MlaC for MlaD was unchanged irrespective of the presence of lipids in either protein. We determined the rate of lipid exchange at various concentrations of MlaC and calculated K_D by plotting concentration against activity determined by fluorescence change [Figure 5.17]. By this method we determined an approximate K_D of MlaC for a lipid bound MlaD to be ~21 μM which is of a similar order to what we had determined previously by SPR (~12 ± 1 μM). As we have shown through our fluorescence assay, the initial lipid exchange event is fast and we believe that this is the best method by which we could approximate a true K_D for PG/PE bound, as other methods would rapidly result in depletion of lipids bound to MlaD. However, we noticed a drastic decrease in lipid

exchange at higher concentrations of MlaC. Likely, this corresponds to either some auto-inhibitory effect, allosteric binding or molecular crowding effects. We see a similar binding profile for apo-MlaC to apo-MlaD determined by MST [Figure 5.18]. MST results suggest that with no lipids present the K_D of the interaction is 43.8 μM which is similar to the results determined by SPR and through our fluorescence assay.

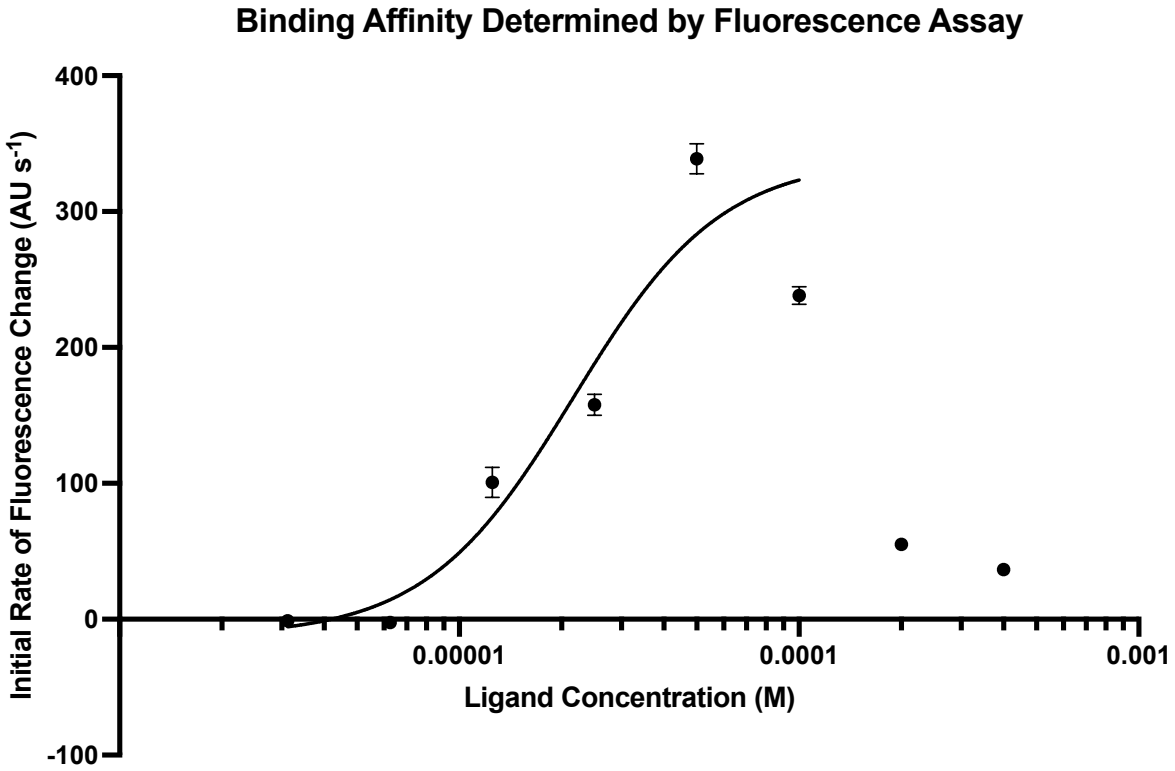

Figure 5.17 – Binding affinity curve for the interaction between MlaC and MlaD. MlaC was treated as the ligand and the concentration of MlaC was varied over the experiment. Response at each concentration was determined from the initial rate of fluorescence change using our assay for lipid exchange. Data plotted here represents the average of 3 repeat experiments and error bars represent the standard deviation. A Hill model with Hill coefficient of ~2 was fit to the data. The two highest concentration points excluded from the model fitting as outliers to allow for the fitting of this model. We expect that this data is more representative of the binding affinity of a PE/PG loaded MlaD for MlaC than previous data determined by SPR.

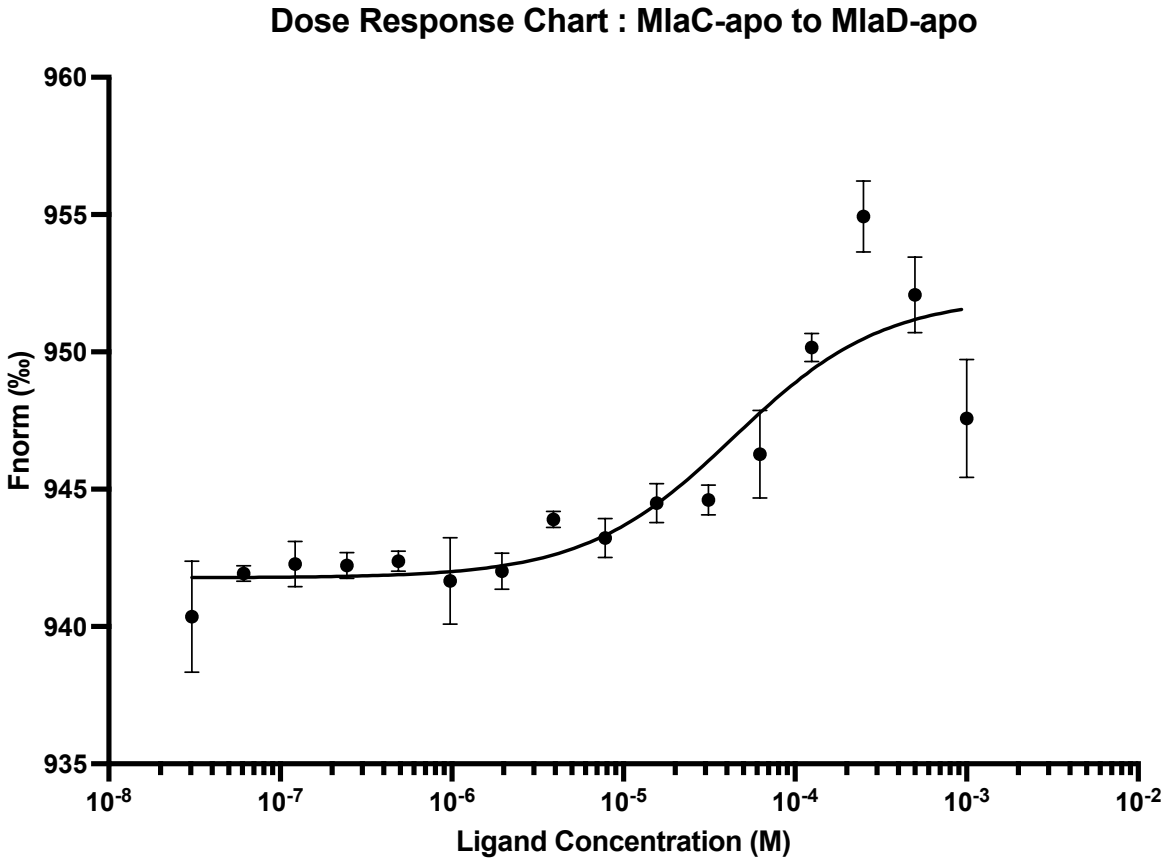

Figure 5.18 - Dose response curve for the interaction between MlaC-apo and MlaD-apo determined by MST. MlaC was treated as the ligand and the concentration of MlaC was varied over the experiment. Data plotted here represents the average of three repeat experiments and error bars represent the standard deviation. A standard K_D model was fit to the data. At higher concentrations we see a decrease in response rather than a plateau, which is consistent with the dose response profile we had determined through our fluorescent assay.

We next investigated the effect of preloading MlaD with cardiolipin on the affinity of MlaC for MlaD [Figure 5.19]. As we had previously utilised cardiolipin to stabilise the transient interaction between these two proteins, we expected that cardiolipin would not be rapidly exchanged to MlaC during the dose response experiment. The dose response profile was significantly different and suggests that this is indeed the case. However, contradictory to our expectations we were unable to reach saturation even with a 1mM peak ligand concentration and were thus unable to determine a K_D, although, it would suggest that the presence of cardiolipin actually substantially decreases the affinity of MlaC for MlaD or potentially results in non-specific

aggregation at higher concentrations of MlaC which would also result in a non-converging dose-response curve.

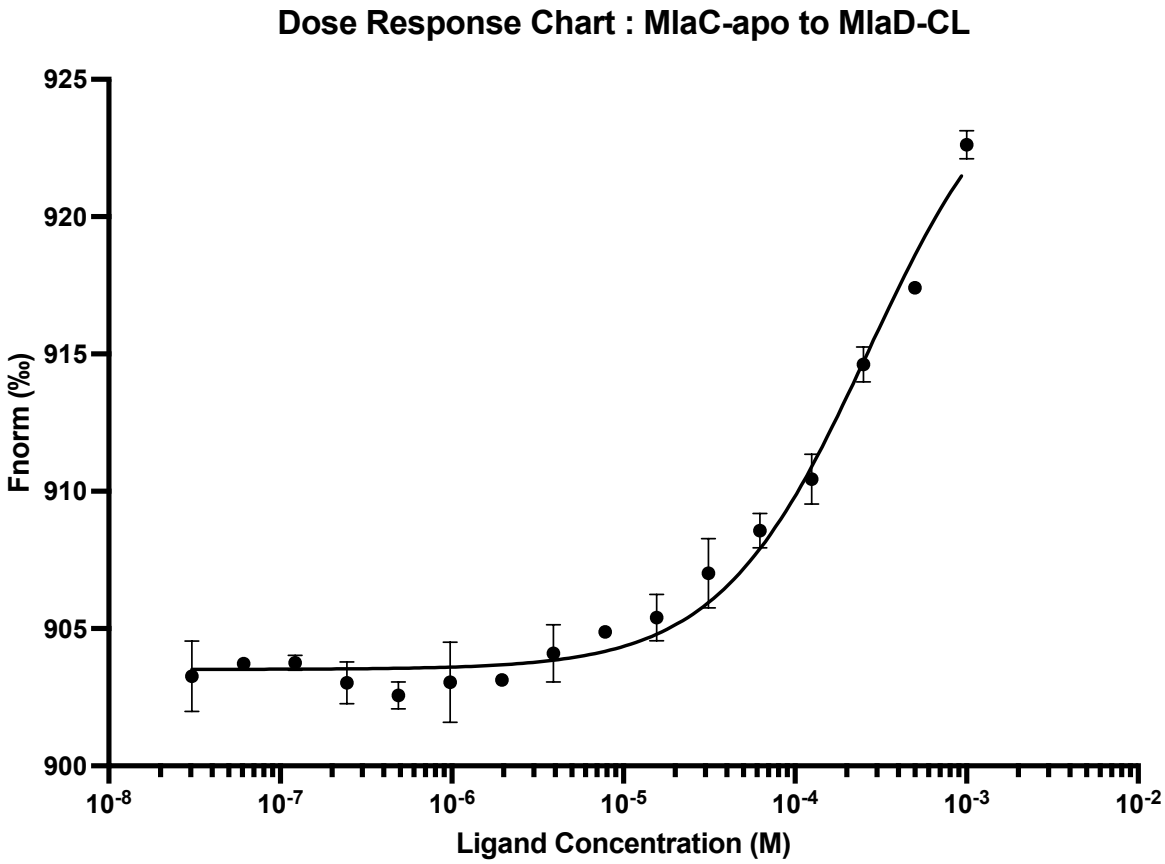

Figure 5.19 - Dose response curve for the interaction between MlaC-apo and MlaD-BDO-CL determined by MST. MlaC was treated as the ligand and the concentration of MlaC was varied over the experiment. Data plotted here represents the average of three repeat experiments and error bars represent the standard deviation. A standard K_D model was fit to the data. There is no evidence of an upper plateau and as such accurate determination of K_D is not possible from this data.

5.3.5.A Fluorescence Assay For MlaFEDB Lipid Exchange

Having determined that the assay is effective for investigating the transport of lipids from MlaD to MlaC we attempted a similar assay using MlaFEDB in a fluorescent liposome.

We anticipated that this assay, which makes use of intrinsic Trp fluorescence could be repurposed to track lipid exchange between MlaFEDB and MlaC despite the presence of Trp residues in the MlaFEDB complex. We expected we could take advantage of the strong reliance of FRET efficiency on FRET pair separation. We anticipated that the average distance between a fluorescently tagged lipid in a liposome and a viable FRET partner will be substantially greater than the distance between the fluorescent tag and a FRET partner when bound in the active site of MlaC. We thus expected a time dependent increase in FRET signalling as lipids are removed from the liposome and transferred to MlaC.

Samples of MlaFEDB in fluorescent liposomes were prepared from detergent extracted samples of MlaFEDB exchanged into liposomes via rapid dilution [Figure 5.20].

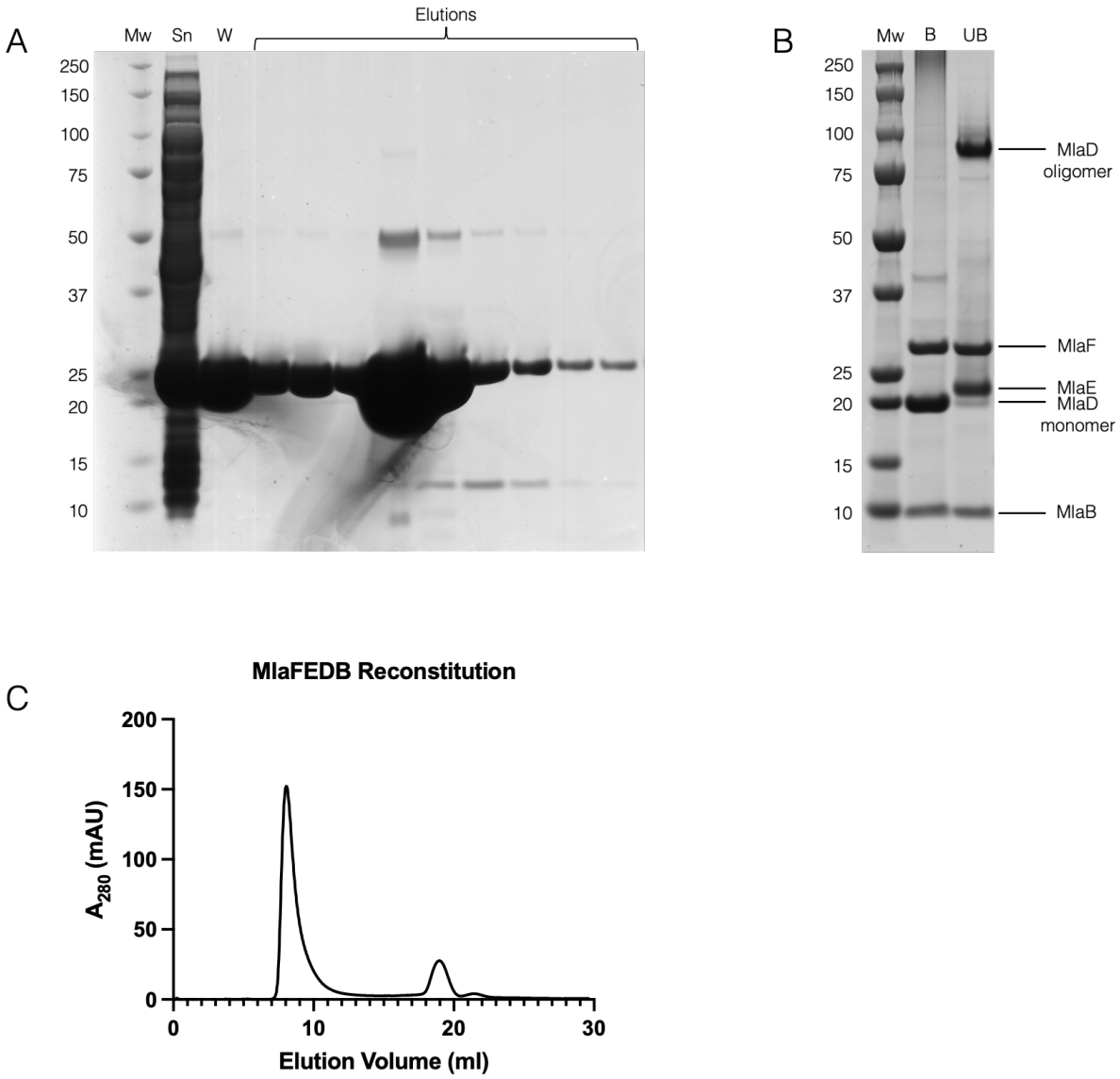

Figure 5.20 - (A) SDS-PAGE of MlaFEDB Ni-NTA purification fractions. Samples from the pre-column supernatant (Sn) and 20 mM imidazole wash (W) and were run alongside the elution fractions. All fractions were boiled prior to being run. (B) SDS-PAGE of purified MlaFEDB sample after SEC. The samples were either boiled prior to running or left un-boiled. Components of the complex are labelled. As noted by Thong et al., MlaE is not visible in the boiled sample. (C) SEC trace of on column reconstitution of MlaFEDB from detergent micelles into liposomes. Samples were taken from the large peak at ~8 mL elution volume.

[Figure 5.21] shows the fluorescence changes over 300 seconds for the assay with MlaFEDB as well as a negative control without the addition of MlaC-apo. Similar to the assay conducted with our soluble construct of MlaD, we see time dependent fluorescence increase upon addition of MlaC.

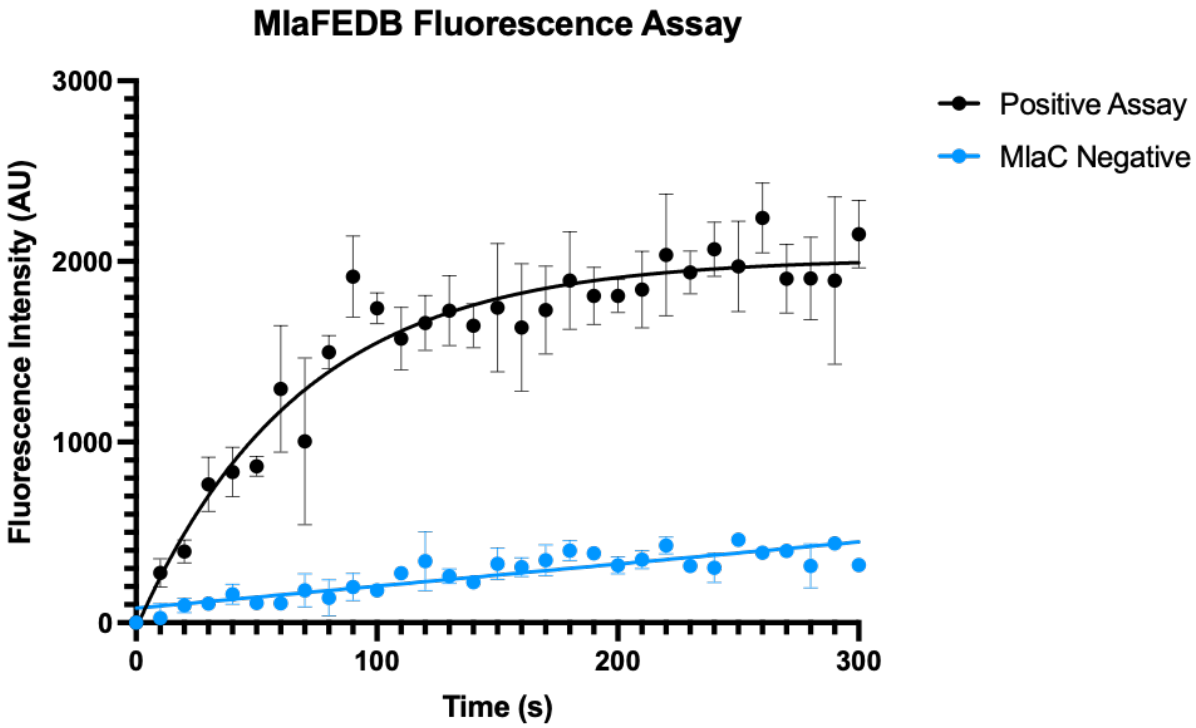

Figure 5.21 - Preliminary data from our fluorescence assay applied to the full MlaFEDB complex exchanged into a fluorescent liposome. Data plotted here represents the average of three repeat experiments and error bars represent the standard deviation. Data suggests that a Trp-dansyl FRET pair can be effective at tracking lipid transport between MlaFEDB and MlaC.

5.4.Discussion

In this chapter we have presented a sensitive assay for tracking real-time lipid exchange between two components of the Mla system with preliminary evidence also suggesting the same assay is functional for tracking lipid exchange from liposomal MlaFEDB to MlaC. This assay has conditional advantages over an existing assay

presented by Tang et *al*. for studying the transfer of lipids in the Mla system[182]. The Tang assay sets up a two liposome system with MlaFEDB and MlaA in separate liposomes. One liposome contains both nitrobenzoxadiazole tagged and rhodamine tagged PE lipids which act as a FRET pair. They used reduction of FRET signal intensity as a measure of lipids transported out of the fluorescent liposome. Comparatively the assay described here uses less material and only tracks lipid transport between two components of the system. It may be more useful when transport to MlaC is the subject of investigation and can potentially be used for rapid screening of inhibitors of lipid exchange between MlaD and MlaC. In fact, it is likely more appropriate for a high-throughput screening than existing assays in literature due to the use of primarily soluble components that are robust to cryo-freezing and storage, unlike he full FEDB complex, which is not robust to freezing and tends to lose activity. However, as this assay monitors a transport direction that is not representative of the physiological direction of lipid transport it may be inappropriate for screening inhibitors. It would, however, be possible to modify the assay to investigate transport from MlaC to MlaD by simply pre-loading MlaC with a fluorescent lipid and observing the reduction in FRET signal upon the introduction of MlaD. For this reason we believe that the assay presented here forms a good basis for a method of high-throughput inhibitor screening for MlaC/MlaD lipid exchange and this presents a good direction for immediate follow-up study.

There is, also, an unresolved discrepancy between what we observed in the negative controls of the FRET assay and existing literature that should likely be investigated. We observed an increase in fluorescence intensity in the MlaD negative assay compared to baseline fluorescence. This was not observed in the MlaC-native control, which showed no increase over baseline fluorescence. This increase is minor when compared to the positive result and as such, has no significant impact the results of the FRET assay, however, it does suggest a slow interaction between dansyl-PE and

MlaC in the absence of MlaD. It would be worth determining the post assay lipid binding states of the MlaC-apo protein in the MlaD negative control, as this result directly contradicts evidence suggesting that MlaC is incapable of spontaneous lipid uptake. There is a possibility that under the conditions of this assay spontaneous lipid uptake by MlaC can occur or that some non-specific interaction between the fluorescent lipid and MlaC results in an increase in FRET signal.

5.4.1.Effect of Mutants on Lipid Transport

We observe that the rate of fluorescence increase for the WT proteins can be modelled as an association of two molecules proportional to e^{-kt}, however, with many of the mutants that is not the case. The reaction endpoint is not stable for many of the mutants, most notably the MlaD Glu119Lys mutant and the MlaC Lys177Met mutant, which over the course of the reaction return to baseline fluorescence. Why this occurs is unclear, although, it suggests that either the direction of lipid transport between the two proteins reverses spontaneously or the lipids are not effectively transferred into the MlaC binding site and dissociate from the protein.

It is unlikely that a lipid bound to MlaC would dissociate into the bulk solvent. Considering the energetics, it has been established that the energy input required to remove a lipid from a membrane into the bulk solvent is approximately 100 $kJmol^{-1}$, and lipids are able to spontaneously transfer through MlaD to MlaC. It would be reasonable to assume that a lipid bound to MlaC is in a more energetically stable state than one in a membrane. As such it would be unlikely for a lipid to spontaneously be released from MlaC into aqueous solvent.

It is more likely that the directionality of lipid transport is situationally reversed. The liposomes used in this fluorescence assay are made using POPG lipids spiked with

dansyl labelled dioleoyl-PE. We decided to use POPG as the base lipid because PG is one of the major substrates of the Mla system and is more amenable to forming liposomes than PE lipids. We used labelled PE mainly due to availability of dansyl labelled PE, with the expectation that the base lipid would make little difference when the head group was heavily modified by the dansyl fluorescent tag. It may be the case that, in several of the mutants, once MlaC has reached saturation the dansyl tagged lipids are selectively exchanged for POPG due to a greater affinity for the binding pocket. If this were to be the case, there would be no preference for dansyl-PE over POPG until MlaC is saturated, but once saturated, exchange of dansyl-PE for POPG is favoured.

It would be possible to determine if the mutations are causing preferential loading of MlaC by conducting a simple follow-up experiment wherein a POPE loaded MlaC WT/mutant is differentially incubated with WT/mutants of MlaD. If these mutations are causing post-saturation exchange of PE lipids for POPG lipids, TLCs of the mutant proteins will show an exchange of PE for PG. This does assume that the differentiating factor is not the dansyl tag. Assuming the assumptions we have made here on the reason for this strange binding profile are correct, we are still unable to explain why, in the case of the Glu119Lys mutant, there is an effect on the binding preference of MlaC. The Glu119Lys mutant is far from the binding pocket of both proteins, however, it is clear from the result that this region of MlaD has some profound effect on lipid exchange beyond simply the binding interaction of MlaD to MlaC.

Of the two mutants that cause the most pronounced deviation from the WT activity profile, the Lys177Met MlaC mutation is situated directly within the lipid binding pocket. The other, Glu119Lys MlaD mutation, is situated far from the MlaC binding pocket. We initially expected that the Glu119Lys MlaD mutation would affect the ability for MlaC to properly bind MlaD and anticipated that this may be reflected in a

reduction in the initial rate of fluorescence change. We did not expect that a mutation in this region would have an impact on preferential lipid binding or lipid retention. This unexpected result may suggest that the pocket of MlaC that the MlaD β6 – β7 loop region binds into has some allosteric effect on lipid binding selectivity or is in some way important for the release of lipids from MlaC to MlaD.

To a lesser degree we also see some evidence of reversed lipid transport after the reaction has plateaued in several other mutants. The MlaD Glu122Lys, MlaD Asp157Lys/Asp158Lys and MlaD Gln149Cys/Leu151Cys mutants all do not come to a stable plateau. How this varied set of mutations all result in what we believe to be a preference for POPG post saturation of MlaC is unclear, however, it would suggest that to some degree these residues play an active part in the transport of lipids between the two proteins.

Considering the initial rates of fluorescence change, we have observed that MlaD Glu122Lys, Glu119Lys and MlaC Lys177Glu do not result in a significant change in rate. The remainder of the mutants appear to either reduce initial rate, or in the case of the MlaD Asp157Lys/Asp158Lys mutant increase initial rate. We expected the compound Glu119Lys/Asp120Lys/Glu122Lys mutant to significantly impede the interaction between MlaD and MlaC and it would appear to have done so, reducing initial lipid transport rate by more than 100-fold to below even the level of the MlaD negative mutant. The Glu119Lys/Asp120Lys/Glu122Lys mutant has an initial rate comparable to the MlaC negative mutant despite the presence of MlaC in the assay, suggesting that it may even impede interactions between MlaC and the free lipids, presumably by sequestering them itself. The assay suggests that this mutation is deleterious to the lipid exchange function of MlaD. As was initially planned, mutating the local residues on MlaC to recover the lipid exchange is necessary to show that

surface charge is the determining factor for interaction and this should be an immediate follow-up experiment to validate this assumption.

The other mutation that severely inhibited initial lipid exchange rate was the Glu180Ala mutation, a residue we expected to have some significant interaction with the Asp145 or Gln149 residues of MlaD. The assay suggests that this single interaction plays a large role in lipid transport. It is potentially this charge interaction that fixes the MlaC binding pocket in the correct position during lipid exchange, however, this is a highly speculative interpretation. Molecular dynamics simulations may elucidate the exact role in exchange.

The MlaD Asp157Lys/Asp158Lys mutant appears to increase the initial rate of fluorescent change. Moreover, the Asp157Lys/Asp158Lys mutant has a peak fluorescence substantially higher than the peak fluorescence of all other assays, which appear to peak at a similar fluorescence value. As this assay approximates total lipid exchange by measuring only the transfer of the tagged PE lipids there is the possibility that any dysregulation in the transfer of POPG compared to PE may result in a higher maximum fluorescence. It is possible that modifying these unstructured residues that potentially localise to a region that interacts with the lipid head group to invert their charge may affect their ability to release and exchange POPG lipids. If that is the case, it is possible that initial rate and maximum fluorescence is misreported due to a reduced rate of POPG transfer but an unaffected or moderately increased rate of PE transfer.

The MlaD Lys155Glu, MlaD Asp120Lys and MlaD Gln149Cys/Leu151Cys mutations all reduce initial activity to a similar degree. We had expected that if lipids were transferred through the central helix bundle then the MlaD Gln149Cys/Leu151Cys would completely attenuate lipid transfer, possibly to the degree that the

Glu119Lys/Asp120Lys/Glu122Lys mutant does. However, as the MlaD Gln149Cys/Leu151Cys mutant is still capable of lipid transport it is possible that our initial assumption regarding parting of the central helix bundle was wrong. This result would suggest that rather than transferring through the helices of the central helix bundle the lipid transports out the top of the channel. The remaining two mutants likely have some effect on the binding affinity as they are at the interface between the two proteins, and this is as we have anticipated. Overall, from the results of these fluorescent assays it would appear that all mutants except the Glu122Lys and Asp157Lys/Asp158Lys. However, it may be necessary to provide additional supporting evidence for the results presented here as we have not yet provided a clear link between reduced lipid exchange capacity in this assay and lipid exchange *in vivo*. Thus I believe it would be necessary to conduct *in vivo* SDS-EDTA sensitivity assays on cells transformed with these mutants to further validate that these results are representative *in vivo*. Should these results be representative *in vivo* it would be fair to say that we have successfully provided structural and functional evidence underpinning the key residues involved in the interaction between MlaC and MlaD as well as producing a high-throughput assay to screen for Mla inhibitors.

5.4.2.Assessment of Kinetic Data

Regarding the binding affinity determined through our fluorescent assay by a dose response experiment, we have fit the data to a standard Hill model assuming the maximum response to be the largest observed initial rate of fluorescence change [Figure 5.17], as we have structural evidence to suggest a binding ratio above 1:1. While this is clearly not a good fit to the data, which sees significantly reduced response at MlaC concentrations above 50-100 μM, we have done this for comparability with previous experiments to determine the binding affinities for these proteins. We expect that there is some degree of allosteric binding occurring at higher

MlaC concentrations resulting in this reduced capacity for lipid exchange. This appears to corroborate the dose response results determined through MST, which also sees reduced response at MlaC concentrations exceeding 100 μM. However, this assessment is somewhat conflicting with previous results, as prior attempts to determine binding affinities through SPR showed no reduced response at higher MlaC concentrations. Unlike the fluorescence assay, both MST and SPR should show some binding response even if the binding is allosteric and not conducive to lipid exchange, which suggests that reduced binding at higher MlaC concentrations is due primarily to an MlaC to MlaC interaction preventing subsequent interaction with MlaD. It is not clear why there is no reduction in response at higher concentrations of MlaC in the SPR experiment, although perhaps the conditions of SPR help to prevent MlaC to MlaC interactions. Overall, it is reasonably clear that more work needs to be done to provide kinetic evidence to determine if the binding stoichiometry determined structurally is indicative of cooperative binding.

6.Final Discussion and Conclusion

We began this investigation with the intent of improving our mechanistic understanding of the Mla system, with specific regard to the interaction between the *mce* domain protein MlaD and the periplasmic chaperone MlaC. Over the course of this report we have provided evidence of previously un-observed interactions between MlaC and its phospholipid substrate, visualised an intermediate complex of MlaC and MlaDΔTM, identified several residues of significant importance to activity and provided a high-throughput assay for phospholipid exchange. While we have not presented clear evidence of the mechanistic relationship between ATP hydrolysis at MlaF and the conformational changes that encourage phospholipid transport from MlaC to MlaD, the data we have provided here along with prior investigations into the MlaFEB complex is sufficient to formulate a likely mechanism for the complex.

6.1.Speculations on the Mechanisms of Mla Proteins

From the prior work of Coudray et *al.* and Chi et *al.* we know that MlaFEDB is outward open, and capable of binding phospholipids at the interface between MlaD and MlaE in the nucleotide-free conformation. Overtly, this suggests a Type II importer mechanism, however, from the results suggested here and in Hughes et *al.* we can see that the substrate binding protein, MlaC, readily dissociates under normal conditions. While this is uncharacteristic of Type II ABC proteins and none of the conformations discussed by Chi and Coudray present an open channel through the TMD, it is clear from the observed ATP independent transfer of phospholipids from MlaFEDB to MlaC that the nucleotide-free conformation has access to the membrane. As this is the case, it is quite possible that MlaFEDB is an irregular Type II ABC transporter where MlaD performs the role of substrate binding protein and is permanently incorporated into the complex. I would expect that if this is the case, it is

so that the conformational changes resulting from ATP binding can be used for removing phospholipids from MlaC rather than dissociating a substrate binding protein from the ABC complex. If the large conformational shifts in MlaD observed by Chi et *al*. as it transitions from the nucleotide-free conformation to the EQ_{close} conformation through the EQ_{tall} state occur while MlaC is bound to MlaD at the position we have shown, there is a possibility that varying distances between the α1 helix of $MlaD_6$ and the β6 – 7 loop of $MlaD_1$ are sufficient to induce a substantial enough conformational change in MlaC to partially collapse the binding pocket and effectively "squeeze" out an acyl tail of the bound substrate. This would both explain the large rotational movement of MlaD observed by Chi and rationalise the binding location of MlaC presented in [Figure 4.18].

From what is known and what we have presented here, I hypothesise that the cycle of ATPase activity in the MlaFEDB complex relates to substrate binding in the following manner [Figure 6.1]. When in the nucleotide-free state, phospholipids, henceforth PL_1 and PL_2 (potentially originating from the IM) sit in the outward open binding pocket of MlaE. Binding of MlaC induces minor conformational changes that enhance the affinity of MlaF for ATP. Binding of ATP results in conformational change first to the EQ_{tall} state, where a single acyl chain of the phospholipid bound to MlaC, henceforth PL_3, is ejected from the binding site in response to conformational change caused by MlaD (as described above). The acyl chain passes over the hydrophobic bridge between the two proteins, formed by Leu151 and Tyr152, into the central MlaD binding pocket [Figure 4.24], which is somewhat expanded in the EQ_{tall} state. Then as the complex shifts from the transient EQ_{tall} state to the EQ_{close} state PL_1 and PL_2 are expelled into the outer leaflet of the IM laterally while PL_3 is still bound partially between MlaC and MlaD. ATP hydrolysis occurs causing relaxation back to the nucleotide-free state and PL_3 is pulled into the MlaE binding pocket as it expands. Dissociation of PL_3 from MlaC is further encouraged by the loss of head-group coordination by Lys177

and Glu180, which now interact with MlaD residues [Figure 3.25 & Figure 4.22]. MlaC, now -apo dissociates and is replaced with a phospholipid bound MlaC. This is consistent with the observation by Hughes et *al.* that the presence of MlaC-apo enhances ATPase activity and the orientation of lipids in the MlaE binding site shown in [Figure 1.26] presented by Chi et *al.* as well as with the mechanisms by which LPS ejection is handled in other complexes, like the Lpt pathway.

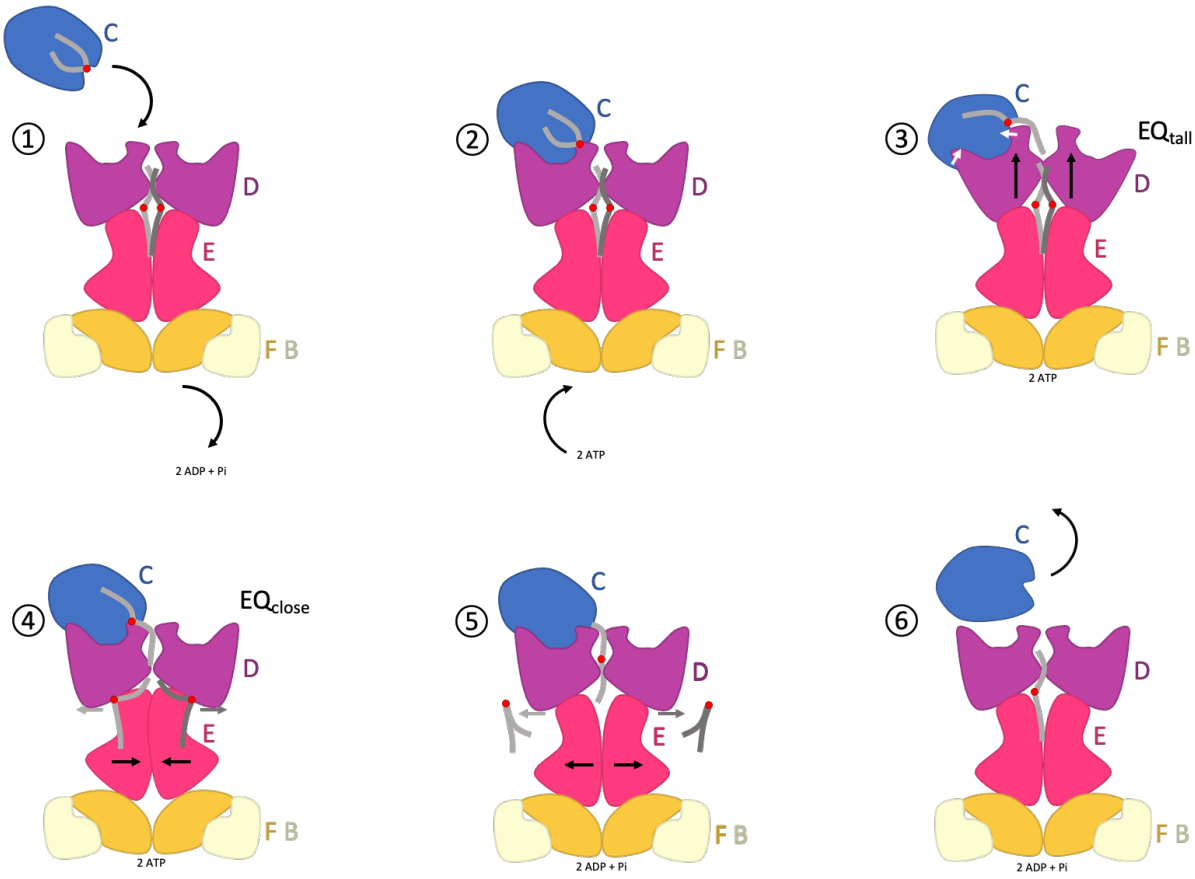

Figure 6.1 – Proposed mechanism for MlaFEDB activity. (1) MlaC binds to the periplasmic face of MlaD. (2) Binding of MlaC causes conformational changes that increase the affinity of MlaF for ATP. (3) Binding of ATP causes tight dimerization of the ATPase domain resulting in a shift from the unbound state through the transient EQ_{tall} state, which sees a moderate conformational change in MlaD that partially collapses the MlaC lipid binding pocket. This releases a lipid tail to MlaD. (4) The complex transitions out of the transient EQ_{tall} state to the EQ_{close} state where the MlaE lipid binding pocket collapses. This forces lipids in the MlaE binding pocket laterally into the membrane. (5) ATP hydrolysis occurs and the complex transitions back (possibly through the EQ_{tall} state) to the nucleotide free state and the MlaE binding pocket expands. (6) Expansion of the MlaE binding pocket allows the lipid, bound transiently between MlaC and MlaD to fall into the empty MlaE binding pocket. MlaC then dissociates.

I expect that the constant outward open state of MlaFEDB which, several groups have observed to permit ATP independent exchange of phospholipids to MlaC, would not be an issue for OM integrity assuming that all periplasmic MlaC is phospholipid bound and the equilibrium of lipid exchange between MlaC and MlaA heavily favours MlaC. The basis for this assumption is the near absence of the -apo state in native purifications of MlaC and the rationale that expending ATP to maintain a population of -apo MlaC in the periplasm, wherein the reverse exchange is spontaneous, would be a constant and unnecessary tax on cellular ATP. Given that the molar amount of periplasmic MlaC is substantially less than the cellular population of lipids it would simply be energetically favourable to allow the saturation of periplasmic MlaC. This mode thereby assumes that unlike most ABC proteins, the presence of a -holo substrate binding protein is not the main impetus for activity. Instead I hypothesise that increases in cellular ATP concentration above a threshold, the likes of which have been observed to occur transiently as part of bacterial stress response may be what encourages upregulated MlaFEDB activity[214, 215]. If this is the case then the purpose of MlaB may be as part of regulatory systems to inhibit Mla activity during instances of cellular ATP increase not associated with OM dysregulation. This model also rationalises observations we have made regarding the binding affinity of MlaC-apo/holo for MlaD in chapter 4. The minimal difference in affinity observed between the two states would not impede the lipid transport function if binding of MlaC-holo to FEDB is driven by a vast difference in concentration between MlaC-holo and MlaC-apo.

The mechanisms proposed here are of course highly speculative, but reflect my best interpretation of what is known about the pathway. Assumptions made in this section may be incorrect but represent an interesting direction for further investigation.

6.2.Topics for Further Investigation and Contributions to the Field of Study

With regards to necessary follow-up experiments, there are two which would significantly enhance the validity of the results presented in this report. The mutations to residues 119, 120 and 122 of MlaD to Lys were intended to show that the surface charge of that region was significant to the interaction between MlaC and MlaD. To ensure that the result we observed was due to an inversion of surface charge rather than steric hinderance due to a Glu to Lys mutation, it may be necessary to mutate residues of MlaC to recover phospholipid exchange activity close to WT. Potentially the Lys71, Lys83 and Arg90 residues of MlaC which are positioned closely to the mutated residues of MlaD are the residues that would recover close to WT phospholipid exchange if our predictions regarding the significance of charged interactions in this region are correct.

It would also significantly improve the validity of the claims made regarding the effect of MlaC and MlaD mutants on lipid exchange if the *in vivo* effects of those mutants are consistent with the results we observed. This can be determined by complementing knockouts of MlaC and MlaD with plasmid constructs of WT/mutants and assessing the capacity to grow on LB agar supplemented with SDS/EDTA, as was originally conducted by Malinverni and Silhavy to assess Mla knockouts. If our assay is an accurate representation of phospholipid exchange *in vivo*, we would expect only the Glu122Lys and the Asp157Lys/Asp158Lys to exhibit SDS/EDTA resistance comparable to WT.

With regards to investigation topics that are potentially worth pursuing, there are several avenues made more achievable due to the results presented here. We have demonstrated that BDO-CL can stabilise an interaction between MlaC and MlaD, there is a possibility that this is also applicable to MlaD as part of the MlaFEDB

complex. It is potentially worth investigating whether exchanging MlaFEDB into a primarily BDO-CL liposome allow for stabilisation of a complex of MlaC and MlaFEDB. Even if BDO-CL does not stabilise an interaction between MlaFEDB and MlaC, structural information presented in this report would make designing a disulphide bonded complex much more likely to be successful. There is definite potential for replicating structural studies carried out by Chi et *al.* with a disulphide stabilised MlaFEDB-MlaC complex to determine, what, if any conformational changes ATP hydrolysis at MlaF cause to MlaC when it is present.

Possibly the most immediately impactful avenue for follow-up research may be the validation of clorobiocin as an inhibitor of Mla activity. We have already demonstrated an assay, wherein, natively purified MlaC with an occupied binding pocket shows activity comparable to negative controls. If clorobiocin does compete for the phospholipid binding pocket of MlaC, as Huang *et al.* suggested[219], the assay we have presented in this report should be effective for validating that claim. Clorobiocin has known antibacterial activity as an aminocoumarin, which inhibits DNA gyrase. If clorobiocin can have off-target effects conducive to OM destabilisation, design of clorobiocin derivatives with improved antimicrobial effects, like was attempted by Galm *et al.* may be more directed[216].

The assay we present here also presents a valuable high-throughput assay for screening compound libraries for activity against the Mla complex. Screening methods for general antibacterial activity, such as disk diffusion or microdilution antibiotic susceptibility testing, may be ineffective at identifying compounds targeted to the Mla system because Mla inhibition is, as previously discussed, not immediately cytotoxic. We believe the assay presented here may allow for the identification of chemical agents with supplementary effects to existing antibacterial agents.

6.3.Conclusion

Overall, in this report we have presented some rather interesting finding regarding the mechanics of lipid exchange in the Mla pathway. We believe that, the structural and functional data we have present here while not sufficient to provide a complete understanding of the molecular mechanism and regulatory aspects of Mla, does provide a basis for further investigation. We believe that some of the tools developed and presented in this report have tangible value to the field of antimicrobial research, however moderate and we believe that with regards to the understating of the Mla pathway we have made a somewhat significant contribution.

7.References

1. Abellón-Ruiz J, Kaptan SS, Basle´ A, Claudi B, Bumann D, Kleinekathöfer U, et al. Structural basis for maintenance of bacterial outer membrane lipid asymmetry. Nat Microbiol. 2017;2(12):1616-23.

2. Adamus K, Le SN, Elmlund H, Boudes M, Elmlund D. AgarFix: Simple and accessible stabilization of challenging single-particle cryo-EM specimens through crosslinking in a matrix of agar. J Struct Biol. 2019;207(3):327-31.

3. Alpi E, Griss J, da Silva AW, Bely B, Antunes R, Zellner H, et al. Analysis of the tryptic search space in UniProt databases. Proteomics. 2015;15(1):48-57.

4. Anderson MS, Bulawa CE, Raetz CR. The biosynthesis of gram-negative endotoxin. Formation of lipid A precursors from UDP-GlcNAc in extracts of Escherichia coli. J Biol Chem. 1985;260(29):15536-41.

5. Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW. Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. Science. 1993;261(5127):1454-7.

6. Baarda BI, Zielke RA, Le Van A, Jerse AE, Sikora AE. Neisseria gonorrhoeae MlaA influences gonococcal virulence and membrane vesicle production. PLoS Pathog. 2019;15(3):e1007385.

7. Bakelar J, Buchanan SK, Noinaj N. The structure of the β-barrel assembly machinery complex. Science. 2016;351(6269):180-6.

8. Barenholz Y, Gibbes D, Litman BJ, Goll J, Thompson TE, Carlson RD. A simple method for the preparation of homogeneous phospholipid vesicles. Biochemistry. 1977;16(12):2806-10.

9. BARTHOLOMEW JW, MITTWER T. The Gram stain. Bacteriol Rev. 1952;16(1):1- 29.

10. Bayer ME. Zones of membrane adhesion in the cryofixed envelope of Escherichia coli. J Struct Biol. 1991;107(3):268-80.

11. Begic S, Worobec EA. Regulation of Serratia marcescens ompF and ompC porin genes in response to osmotic stress, salicylate, temperature and pH. Microbiology (Reading). 2006;152(Pt 2):485-91.

12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235-42.

13. Bernier SP, Son S, Surette MG. The Mla Pathway Plays an Essential Role in the Intrinsic Resistance of Burkholderia cepacia Complex Species to Antimicrobials and Host Innate Components. J Bacteriol. 2018;200(18).

14. Bishop RE, Gibbons HS, Guina T, Trent MS, Miller SI, Raetz CR. Transfer of palmitate from phospholipids to lipid A in outer membranes of gram-negative bacteria. EMBO J. 2000;19(19):5071-80.

15. Blair JM, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJ. Molecular mechanisms of antibiotic resistance. Nat Rev Microbiol. 2015;13(1):42-51.

16. Bohl HO, Shi K, Lee JK, Aihara H. Crystal structure of lipid A disaccharide synthase LpxB from Escherichia coli. Nat Commun. 2018;9(1):377.

17. Bos MP, Tommassen J. Biogenesis of the Gram-negative bacterial outer membrane. Curr Opin Microbiol. 2004;7(6):610-6.

18. Botos I, Majdalani N, Mayclin SJ, McCarthy JG, Lundquist K, Wojtowicz D, et al. Structural and Functional Characterization of the LPS Transporter LptDE from Gram-Negative Pathogens. Structure. 2016;24(6):965-76.

19. Bragg WL, West J. X. A technique for the X-ray examination of crystal structures with many parameters. Zeitschrift fu̇r Kristallographie-Crystalline Materials. 1929;69(1-6):118-48.

20. Bragg WH, Bragg WL. The reflection of X-rays by crystals. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character. 1913;88(605):428-38.

21. Braun V, Wu H. Lipoproteins, structure, function, biosynthesis and model for protein export. New comprehensive biochemistry: Elsevier; 1994. p. 319-41.

22. Bravais A. Les Systemes Formes par des Points Distribues Regulierement sur un Plan ou Dans Lespace. Journal de lEcole Polytechnique, XIX. 1850;1.

23. Brock DJ, Kass LR, Bloch K. Beta-hydroxydecanoyl thioester dehydrase. II. Mode of action. J Biol Chem. 1967;242(19):4432-40.

24. Buetow L, Smith TK, Dawson A, Fyffe S, Hunter WN. Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis. Proc Natl Acad Sci U S A. 2007;104(11):4321-6.

25. Buskiewicz I, Deuerling E, Gu SQ, Jȯckel J, Rodnina MV, Bukau B, et al. Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. Proc Natl Acad Sci U S A. 2004;101(21):7902-6.

26. Carty SM, Sreekumar KR, Raetz CR. Effect of cold shock on lipid A biosynthesis in Escherichia coli. Induction At 12 degrees C of an acyltransferase specific for palmitoleoyl-acyl carrier protein. J Biol Chem. 1999;274(14):9677-85.

27. Casali N, Konieczny M, Schmidt MA, Riley LW. Invasion activity of a Mycobacterium tuberculosis peptide presented by the Escherichia coli AIDA autotransporter. Infect Immun. 2002;70(12):6846-52.

28. Casali N, Riley LW. A phylogenomic analysis of the Actinomycetales mce operons. BMC Genomics. 2007;8:60.

29. Chen J. Molecular mechanism of the Escherichia coli maltose transporter. Curr Opin Struct Biol. 2013;23(4):492-8.

30. Chi X, Fan Q, Zhang Y, Liang K, Wan L, Zhou Q, et al. Structural mechanism of phospholipids translocation by MlaFEDB complex. Cell Res. 2020;30(12):1127-35.

31. Choi KH, Heath RJ, Rock CO. beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. J Bacteriol. 2000;182(2):365-70.

32. Cordat E, Mus-Veteau I, Leblanc G. Structural Studies of the Melibiose Permease of Escherichia coli by Fluorescence Resonance Energy Transfer: II. Identification of the

Tryptophan Residues Acting as Energy Donors. Journal of Biological Chemistry. 1998;273(50):33198-202.

33. Costerton JW, Ingram JM, Cheng KJ. Structure and function of the cell envelope of gram-negative bacteria. Bacteriol Rev. 1974;38(1):87-110.

34. Coudray N, Isom GL, MacRae MR, Saiduddin MN, Bhabha G, Ekiert DC. Structure of bacterial phospholipid transporter MlaFEDB with substrate bound. Elife. 2020;9.

35. Cronan JE, Rock CO. Biosynthesis of Membrane Lipids. EcoSal Plus. 2008;3(1).

36. Cronan JE, Waldrop GL. Multi-subunit acetyl-CoA carboxylases. Prog Lipid Res. 2002;41(5):407-35.

37. Davidson AL, Dassa E, Orelle C, Chen J. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev. 2008;72(2):317-64, table of contents.

38. Dekker N. Outer-membrane phospholipase A: known structure, unknown biological function. Mol Microbiol. 2000;35(4):711-7.

39. Delcour AH. Outer membrane permeability and antibiotic resistance. Biochim Biophys Acta. 2009;1794(5):808-16.

40. Deuerling E, Schulze-Specking A, Tomoyasu T, Mogk A, Bukau B. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature. 1999;400(6745):693-6.

41. Doerrler WT. Lipid trafficking to the outer membrane of Gram-negative bacteria. Mol Microbiol. 2006;60(3):542-52.

42. Dong H, Zhang Z, Tang X, Huang S, Li H, Peng B, et al. Structural insights into cardiolipin transfer from the Inner membrane to the outer membrane by PbgA in Gram-negative bacteria. Sci Rep. 2016;6:30815.

43. Donohue-Rolfe AM, Schaechter M. Translocation of phospholipids from the inner to the outer membrane of Escherichia coli. Proc Natl Acad Sci U S A. 1980;77(4):1867-71.

44. Dormán G, Prestwich GD. Benzophenone photophores in biochemistry. Biochemistry. 1994;33(19):5661-73.

45. Dowhan W. A retrospective: use of Escherichia coli as a vehicle to study phospholipid synthesis and function. Biochim Biophys Acta. 2013;1831(3):471-94.

46. Drulyte I, Johnson RM, Hesketh EL, Hurdiss DL, Scarff CA, Porav SA, et al. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. Acta Crystallogr D Struct Biol. 2018;74(Pt 6):560-71.

47. du Plessis DJ, Nouwen N, Driessen AJ. The Sec translocase. Biochim Biophys Acta. 2011;1808(3):851-65.

48. Dubochet J, Adrian M, Chang JJ, Homo JC, Lepault J, McDowall AW, et al. Cryo-electron microscopy of vitrified specimens. Q Rev Biophys. 1988;21(2):129-228.

49. Dunstan RA, Hay ID, Wilksch JJ, Schittenhelm RB, Purcell AW, Clark J, et al. Assembly of the secretion pores GspD, Wza and CsgG into bacterial outer membranes does not require the Omp85 proteins BamA or TamA. Mol Microbiol. 2015;97(4):616-29.

50. Ekiert DC, Bhabha G, Isom GL, Greenan G, Ovchinnikov S, Henderson IR, et al. Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. Cell. 2017;169(2):273-85.e17.

51. Emptage RP, Daughtry KD, Pemble CW, Raetz CR. Crystal structure of LpxK, the 4'-kinase of lipid A biosynthesis and atypical P-loop kinase functioning at the membrane interface. Proc Natl Acad Sci U S A. 2012;109(32):12956-61.

52. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.

53. Ercan B, Low WY, Liu X, Chng SS. Characterization of Interactions and Phospholipid Transfer between Substrate Binding Proteins of the OmpC-Mla System. Biochemistry. 2019;58(2):114-9.

54. Falagas ME, Bliziotis IA. Pandrug-resistant Gram-negative bacteria: the dawn of the post-antibiotic era? Int J Antimicrob Agents. 2007;29(6):630-6.

55. Fan X, Wang J, Zhang X, Yang Z, Zhang JC, Zhao L, et al. Single particle cryo-

EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. Nat Commun. 2019;10(1):2386.

56. Faruqi A, McMullan G. Direct imaging detectors for electron microscopy. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2018;878:180-90.

57. Forster T. Energiewanderung und fluoreszenz. Naturwissenschaften. 1946;33(6):166-75.

58. Galloway SM, Raetz CR. A mutant of Escherichia coli defective in the first step of endotoxin biosynthesis. J Biol Chem. 1990;265(11):6394-402.

59. Gama-Castro S, Jimeńez-Jacinto V, Peralta-Gil M, Santos-Zavaleta A, Peñaloza- Spinola MI, Contreras-Moreira B, et al. RegulonDB (version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. Nucleic Acids Res. 2008;36(Database issue):D120-4.

60. Geoghegan KF, Dixon HB, Rosner PJ, Hoth LR, Lanzetti AJ, Borzilleri KA, et al. Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. Anal Biochem. 1999;267(1):169-84.

61. Goñi FM. The basic structure and dynamics of cell membranes: an update of the Singer-Nicolson model. Biochim Biophys Acta. 2014;1838(6):1467-76.

62. Grimm J, Shi H, Wang W, Mitchell AM, Wingreen NS, Huang KC, et al. The inner membrane protein YhdP modulates the rate of anterograde phospholipid flow in. Proc Natl Acad Sci U S A. 2020;117(43):26907-14.

63. Hahn T, Shmueli U, Arthur JW. International tables for crystallography: Reidel Dordrecht; 1983.

64. Hara T, Matsuyama S, Tokuda H. Mechanism underlying the inner membrane retention of Escherichia coli lipoproteins caused by Lol avoidance signals. J Biol Chem. 2003;278(41):40408-14.

65. Heath RJ, Rock CO. Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in Escherichia coli. J Biol Chem. 1996;271(18):10996-1000.

66. Heinrichs DE, Yethon JA, Whitfield C. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of Escherichia coli and Salmonella enterica. Mol Microbiol. 1998;30(2):221-32.

67. Hirai K, Aoyama H, Irikura T, Iyobe S, Mitsuhashi S. Differences in susceptibility to quinolones of outer membrane mutants of Salmonella typhimurium and Escherichia coli. Antimicrob Agents Chemother. 1986;29(3):535-8.

68. Hong M, Gleason Y, Wyckoff EE, Payne SM. Identification of two Shigella flexneri chromosomal loci involved in intercellular spreading. Infect Immun. 1998;66(10):4700-10.

69. Hsu L, Jackowski S, Rock CO. Isolation and characterization of Escherichia coli K-12 mutants lacking both 2-acyl-glycerophosphoethanolamine acyltransferase and acyl-acyl carrier protein synthetase activity. J Biol Chem. 1991;266(21):13783-8.

70. Huang YM, Miao Y, Munguia J, Lin L, Nizet V, McCammon JA. Molecular dynamic study of MlaC protein in Gram-negative bacteria: conformational flexibility, solvent effect and protein-phospholipid binding. Protein Sci. 2016;25(8):1430-7.

71. Hughes GW, Hall SCL, Laxton CS, Sridhar P, Mahadi AH, Hatton C, et al. Evidence for phospholipid export from the bacterial inner membrane by the Mla ABC transport system. Nat Microbiol. 2019;4(10):1692-705.

72. Isom GL, Davies NJ, Chong ZS, Bryant JA, Jamshad M, Sharif M, et al. MCE domain proteins: conserved inner membrane lipid-binding proteins required for outer membrane homeostasis. Sci Rep. 2017;7(1):8608.

73. Jackman JE, Raetz CR, Fierke CA. UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase of Escherichia coli is a zinc metalloenzyme. Biochemistry. 1999;38(6):1902-11.

74. Jeeves M, Knowles TJ. A novel pathway for outer membrane protein biogenesis in Gram-negative bacteria. Mol Microbiol. 2015;97(4):607-11.

75. Jerabek-Willemsen M, Andre´ T, Wanner R, Roth HM, Duhr S, Baaske P, et al. MicroScale Thermophoresis: Interaction analysis and beyond. Journal of Molecular Structure. 2014;1077:101-13.

76. Jones NC, Osborn MJ. Translocation of phospholipids between the outer and inner membranes of Salmonella typhimurium. J Biol Chem. 1977;252(20):7405-12.

77. Kamischke C, Fan J, Bergeron J, Kulasekara HD, Dalebroux ZD, Burrell A, et al. The Acinetobacter baumannii Mla system and glycerophospholipid transport to the outer membrane. eLife. 2019;8:e40171.

78. Kaplan E, Greene NP, Crow A, Koronakis V. Insights into bacterial lipoprotein trafficking from a structure of LolA bound to the LolC periplasmic domain. Proc Natl Acad Sci U S A. 2018;115(31):E7389-E97.

79. Kastner B, Fischer N, Golas MM, Sander B, Dube P, Boehringer D, et al. GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods. 2008;5(1):53-5.

80. Keenleyside WJ, Whitfield C. A novel pathway for O-polysaccharide biosynthesis in Salmonella enterica serovar Borreze. J Biol Chem. 1996;271(45):28581-92.

81. Kellenberger E. The 'Bayer bridges' confronted with results from improved electron microscopy methods. Mol Microbiol. 1990;4(5):697-705.

82. Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat Rev Microbiol. 2009;7(3):206-14.

83. Koebnik R, Locher KP, Van Gelder P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol Microbiol. 2000;37(2):239-53.

84. Koh YS, Roe JH. Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus in Escherichia coli. J Bacteriol. 1995;177(10):2673-8.

85. Kolich LR, Chang YT, Coudray N, Giacometti SI, MacRae MR, Isom GL, et al. Structure of MlaFB uncovers novel mechanisms of ABC transporter regulation. Elife. 2020;9.

86. Kong H, Akakin HC, Sarma SE. A generalized Laplacian of Gaussian filter for blob detection and its applications. IEEE transactions on cybernetics. 2013;43(6):1719-33.

87. KOPMANN HJ, JANN K. Biosynthesis of the 09 Antigen of Escherichia coli: The Polysaccharide Component of E. coli 09: K29–. European Journal of Biochemistry. 1975;60(2):587-601.

88. Kretschmann E, Raether H. Radiative decay of non radiative surface plasmons excited by light. Zeitschrift für Naturforschung A. 1968;23(12):2135-6.

89. Kumar A, Chandolia A, Chaudhry U, Brahmachari V, Bose M. Comparison of mammalian cell entry operons of mycobacteria: in silico analysis and expression profiling. FEMS Immunol Med Microbiol. 2005;43(2):185-95.

90. Kuzin AP, Chen Y, Jayaraman S, Chen C, Fang Y, Cunningham K, et al. Three-dimensional structure of the phospholipid-binding protein from Ralstonia solanacearum Q8XV73_RALSQ in complex with a phospholipid at the resolution 1.53 A. Northeast Structural Genomics Consortium target RsR89 (2QGU). 2007 [cited 2022]. Available from: https://www.rcsb.org/structure/2QGUhttps://www.rcsb.org/structure/2QGU.

91. Kuzin A, Su M, Seetharaman J, Sahdev S, Xiao R, Ciccosanti C, et al. Crystal Structure of Toluene-tolerance protein from Pseudomonas putida (strain KT2440), Northeast Structural Genomics Consortium (NESG) Target PpR99 (4FCZ). 2013 [cited 2022]. Available from: https://www.rcsb.org/structure/4FCZ.

92. Langley KE, Hawrot E, Kennedy EP. Membrane assembly: movement of phosphatidylserine between the cytoplasmic and outer membranes of Escherichia coli. J Bacteriol. 1982;152(3):1033-41.

93. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. 2005;69(1):101-23.

94. Lee S, Sowa ME, Watanabe YH, Sigler PB, Chiu W, Yoshida M, et al. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell. 2003;115(2):229-40.

95. Lewinson O, Lee AT, Locher KP, Rees DC. A distinct mechanism for the ABC transporter BtuCD-BtuF revealed by the dynamics of complex formation. Nat Struct Mol Biol. 2010;17(3):332-8.

96. Li X, Keskin O, Ma B, Nussinov R, Liang J. Protein–protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. Journal of molecular biology. 2004;344(3):781-95.

97. Li Y, Orlando BJ, Liao M. Structural basis of lipopolysaccharide extraction by the LptB2FGC complex. Nature. 2019;567(7749):486-90.

98. Liao JY, Song Y, Liu Y. A new trend to determine biochemical parameters by quantitative FRET assays. Acta Pharmacol Sin. 2015;36(12):1408-15.

99. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol. 2019;75(Pt 10):861-77.

100. Linton KJ, Higgins CF. The Escherichia coli ATP-binding cassette (ABC) proteins. Mol Microbiol. 1998;28(1):5-13.

101. Liu D, Cole RA, Reeves PR. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J Bacteriol. 1996;178(7):2102-7.

102. Low WY, Thong S, Chng SS. ATP disrupts lipid-binding equilibrium to drive retrograde transport critical for bacterial outer membrane asymmetry. Proc Natl Acad Sci U S A. 2021;118(50).

103. Ludwig C. Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösung: Aus der KK Hof-und Staatsdruckerei, in Commission bei W. Braumüller ...; 1856.

104. Luo Q, Yang X, Yu S, Shi H, Wang K, Xiao L, et al. Structural basis for lipopolysaccharide extraction by ABC transporter LptB2FG. Nat Struct Mol Biol. 2017;24(5):469-74.

105. Madan Babu M, Sankaran K. DOLOP--database of bacterial lipoproteins. Bioinformatics. 2002;18(4):641-3.

106. Malinverni JC, Silhavy TJ. An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. Proc Natl Acad Sci U S A. 2009;106(19):8009-14.

107. Malinverni JC, Werner J, Kim S, Sklar JG, Kahne D, Misra R, et al. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. Mol Microbiol. 2006;61(1):151-64.

108. Mann D, Fan J, Farrell DP, Somboon K, Andrew Muenks S, Tzokov S, et al. Structural basis for lipid transport by the MLA complex. bioRxiv. 2020:2020.05. 30.125013.

109. May JM, Owens TW, Mandler MD, Simpson BW, Lazarus MB, Sherman DJ, et al. The Antibiotic Novobiocin Binds and Activates the ATPase That Powers Lipopolysaccharide Transport. J Am Chem Soc. 2017;139(48):17221-4.

110. McConkey EH. Molecular evolution, intracellular organization, and the quinary structure of proteins. Proc Natl Acad Sci U S A. 1982;79(10):3236-40.

111. McPherson A, Gavira JA. Introduction to protein crystallization. Acta Crystallographica Section F: Structural Biology Communications. 2014;70(1):2-20.

112. Migneault I, Dartiguenave C, Bertrand MJ, Waldron KC. Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. Biotechniques. 2004;37(5):790-802.

113. Miller WH. A treatise on crystallography: Deighton; 1839.

114. Mitra D, Saha B, Das D, Wiker HG, Das AK. Correlating sequential homology of Mce1A, Mce2A, Mce3A and Mce4A with their possible functions in mammalian cell entry of Mycobacterium tuberculosis performing homology modeling. Tuberculosis (Edinb). 2005;85(5-6):337-45.

115. Moser R, Aktas M, Fritz C, Narberhaus F. Discovery of a bifunctional cardiolipin/phosphatidylethanolamine synthase in bacteria. Mol Microbiol. 2014;92(5):959-72.

116. Moussatova A, Kandt C, O'Mara ML, Tieleman DP. ATP-binding cassette transporters in Escherichia coli. Biochim Biophys Acta. 2008;1778(9):1757- 71.

117. Mühlradt PF, Golecki JR. Asymmetrical distribution and artifactual reorientation of lipopolysaccharide in the outer membrane bilayer of Salmonella typhimurium. Eur J Biochem. 1975;51(2):343-52.

118. Nagamori S, Smirnova IN, Kaback HR. Role of YidC in folding of polytopic membrane proteins. J Cell Biol. 2004;165(1):53-62.

119. Nakayama T, Zhang-Akiyama QM. pqiABC and yebST, Putative mce Operons of Escherichia coli, Encode Transport Pathways and Contribute to Membrane Integrity. J Bacteriol. 2017;199(1).

120. Narita SI, Tokuda H. Bacterial lipoproteins; biogenesis, sorting and quality control. Biochim Biophys Acta Mol Cell Biol Lipids. 2017;1862(11):1414-23.

121. Nikaido H. Restoring permeability barrier function to outer membrane. Chem Biol. 2005;12(5):507-9.

122. Nikaido H, Nakae T. The outer membrane of Gram-negative bacteria. Adv Microb Physiol. 1979;20:163-250.

123. Nikaido H, Vaara M. Molecular basis of bacterial outer membrane permeability. Microbiol Rev. 1985;49(1):1-32.

124. Nooren IM, Thornton JM. Diversity of protein-protein interactions. EMBO J. 2003;22(14):3486-92.

125. Okada C, Wakabayashi H, Kobayashi M, Shinoda A, Tanaka I, Yao M. Crystal structures of the UDP-diacylglucosamine pyrophosphohydrase LpxH from Pseudomonas aeruginosa. Sci Rep. 2016;6:32822.

126. Okuda S, Freinkman E, Kahne D. Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. Science. 2012;338(6111):1214-7.

127. Osborn MJ, Munson R. Separation of the inner (cytoplasmic) and outer membranes of Gram-negative bacteria. Methods Enzymol. 1974;31:642-53.

128. Oursel D, Loutelier-Bourhis C, Orange N, Chevalier S, Norris V, Lange CM. Identification and relative quantification of fatty acids in Escherichia coli membranes by gas chromatography/mass spectrometry. Rapid Commun Mass Spectrom. 2007;21(20):3229-33.

129. Overath P, Pauli G, Schairer HU. Fatty acid degradation in Escherichia coli. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants. Eur J Biochem. 1969;7(4):559-74.

130. Parsons JB, Broussard TC, Bose JL, Rosch JW, Jackson P, Subramanian C, et al. Identification of a two-component fatty acid kinase responsible for host fatty acid incorporation by Staphylococcus aureus. Proc Natl Acad Sci U S A. 2014;111(29):10532-7.

131. Parsons JB, Rock CO. Bacterial lipids: metabolism and membrane homeostasis. Prog Lipid Res. 2013;52(3):249-76.

132. Patterson AL. A direct method for the determination of the components of interatomic distances in crystals. Zeitschrift fu̇r Kristallographie-Crystalline Materials. 1935;90(1-6):517-42.

133. Pattnaik P. Surface plasmon resonance: applications in understanding receptor-ligand interaction. Appl Biochem Biotechnol. 2005;126(2):79-92.

134. Pérez-Cruz C, Delgado L, López-Iglesias C, Mercade E. Outer-inner membrane vesicles naturally secreted by gram-negative pathogenic bacteria. PLoS One. 2015;10(1):e0116896.

135. Pettit RK, Judd RC. The interaction of naturally elaborated blebs from serum- susceptible and serum-resistant strains of Neisseria gonorrhoeae with normal human serum. Mol Microbiol. 1992;6(6):729-34.

136. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200-W4.

137. Powers MJ, Simpson BW, Trent MS. The Mla pathway in Acinetobacter baumannii has no demonstrable role in anterograde lipid transport. Elife. 2020;9.

138. Powers MJ, Trent MS. Expanding the paradigm for the outer membrane: Acinetobacter baumannii in the absence of endotoxin. Mol Microbiol. 2018;107(1):47-56.

139. Raetz CR. Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli. Microbiol Rev. 1978;42(3):614-59.

140. Rajamani D, Thiel S, Vajda S, Camacho CJ. Anchor residues in protein-protein interactions. Proc Natl Acad Sci U S A. 2004;101(31):11287-92.

141. Ray TK, Cronan JE. Activation of long chain fatty acids with acyl carrier protein: demonstration of a new enzyme, acyl-acyl carrier protein synthetase, in Escherichia coli. Proc Natl Acad Sci U S A. 1976;73(12):4374-8.

142. Renaud JP, Chari A, Ciferri C, Liu WT, Rémigy HW, Stark H, et al. Cryo-EM in drug discovery: achievements, limitations and prospects. Nat Rev Drug Discov. 2018;17(7):471-92.

143. Rest RF, Cooney MH, Spitznagel JK. Susceptibility of lipopolysaccharide mutants to the bactericidal action of human neutrophil lysosomal fractions. Infect Immun. 1977;16(1):145-51.

144. Robert V, Volokhina EB, Senf F, Bos MP, Van Gelder P, Tommassen J. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. PLoS Biol. 2006;4(11):e377.

145. Roier S, Zingl FG, Cakar F, Durakovic S, Kohl P, Eichmann TO, et al. A novel mechanism for the biogenesis of outer membrane vesicles in Gram-negative bacteria. Nat Commun. 2016;7:10515.

146. Ruch FE, Vagelos PR. The isolation and general properties of Escherichia coli malonyl coenzyme A-acyl carrier protein transacylase. J Biol Chem. 1973;248(23):8086-94.

147. Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ. Identification of two inner- membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. Proc Natl Acad Sci U S A. 2008;105(14):5537-42.

148. Sabatini DD, Bensch K, Barrnett RJ. Cytochemistry and electron microscopy: the preservation of cellular ultrastructure and enzymatic activity by aldehyde fixation. The Journal of cell biology. 1963;17(1):19-58.

149. Saha SK, Nishijima S, Matsuzaki H, Shibuya I, Matsumoto K. A regulatory mechanism for the balanced synthesis of membrane phospholipid species in Escherichia coli. Biosci Biotechnol Biochem. 1996;60(1):111-6.

150. Salafsky J, Groves JT, Boxer SG. Architecture and function of membrane proteins in planar supported bilayers: a study with photosynthetic reaction centers. Biochemistry. 1996;35(47):14773-81.

151. Samuelson JC, Chen M, Jiang F, Möller I, Wiedmann M, Kuhn A, et al. YidC mediates membrane protein insertion in bacteria. Nature. 2000;406(6796):637-41.

152. Sankaran K, Wu HC. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. J Biol Chem. 1994;269(31):19701-6.

153. Scheres SH. A Bayesian view on cryo-EM structure determination. J Mol Biol. 2012;415(2):406-18.

154. Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180(3):519-30.

155. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife. 2014;3:e03665.

156. Schmeing TM, Ramakrishnan V. What recent ribosome structures have revealed about the mechanism of translation. Nature. 2009;461(7268):1234- 42.

157. Schmidt G, Fromme I, Mayer H. Immunochemical studies on core lipopolysaccharides of Enterobacteriaceae of different genera. Eur J Biochem. 1970;14(2):357-66.

158. Schmidt H, Hansen G, Singh S, Hanuszkiewicz A, Lindner B, Fukase K, et al. Structural and mechanistic analysis of the membrane-embedded glycosyltransferase WaaA required for lipopolysaccharide synthesis. Proc Natl Acad Sci U S A. 2012;109(16):6253-8.

159. Schreiber G, Fersht AR. Energetics of protein-protein interactions: Analysis ofthe Barnase-Barstar interface by single mutations and double mutant cycles. Journal of molecular biology. 1995;248(2):478-86.

160. Selkrig J, Belousoff MJ, Headey SJ, Heinz E, Shiota T, Shen HH, et al. Conserved features in TamA enable interaction with TamB to drive the activity of the translocation and assembly module. Sci Rep. 2015;5:12905.

161. Selkrig J, Mosbahi K, Webb CT, Belousoff MJ, Perry AJ, Wells TJ, et al. Discovery of an archetypal protein transport system in bacterial outer membranes. Nat Struct Mol Biol. 2012;19(5):506-10, S1.

162. Shechtman D, Blech I, Gratias D, Cahn JW. Metallic phase with long-range orientational order and no translational symmetry. Physical review letters. 1984;53(20):1951.

163. Sherman DJ, Xie R, Taylor RJ, George AH, Okuda S, Foster PJ, et al. Lipopolysaccharide is transported to the cell surface by a membrane-to- membrane protein bridge. Science. 2018;359(6377):798-801.

164. Shrivastava R, Chng SS. Lipid trafficking across the Gram-negative cell envelope. J Biol Chem. 2019;294(39):14175-84.

165. Shruthi H, Anand P, Murugan V, Sankaran K. Twin arginine translocase pathway and fast-folding lipoprotein biosynthesis in E. coli: interesting implications and applications. Mol Biosyst. 2010;6(6):999-1007.

166. Sijbrandi R, Urbanus ML, ten Hagen-Jongman CM, Bernstein HD, Oudega B, Otto BR, et al. Signal recognition particle (SRP)-mediated targeting and Sec- dependent translocation of an extracellular Escherichia coli protein. J Biol Chem. 2003;278(7):4654-9.

167. Silhavy TJ, Kahne D, Walker S. The bacterial cell envelope. Cold Spring Harb Perspect Biol. 2010;2(5):a000414.

168. Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. Science. 1972;175(4023):720-31.

169. Sklar JG, Wu T, Gronenberg LS, Malinverni JC, Kahne D, Silhavy TJ. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. Proc Natl Acad Sci U S A. 2007;104(15):6400-5.

170. Sklar JG, Wu T, Kahne D, Silhavy TJ. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. Genes Dev. 2007;21(19):2473-84.

171. Sneath PH. Relations between chemical structure and biological activity in peptides. J Theor Biol. 1966;12(2):157-95.

172. Sperandeo P, Martorana AM, Polissi A. The lipopolysaccharide transport (Lpt) machinery: A nonconventional transporter for lipopolysaccharide assembly at the outer membrane of Gram-negative bacteria. J Biol Chem. 2017;292(44):17981-90.

173. Stoffler-Meilicke M, Stoffler G, Odom OW, Zinn A, Kramer G, Hardesty B. Localization of 3' ends of 5S and 23S rRNAs in reconstituted subunits of Escherichia coli ribosomes. Proc Natl Acad Sci U S A. 1981;78(9):5538-42.

174. Struyvé M, Moons M, Tommassen J. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J Mol Biol. 1991;218(1):141-8.

175. Su Z, Wu Y. Computational studies of protein-protein dissociation by statistical potential and coarse-grained simulations: a case study on interactions between colicin E9 endonuclease and immunity proteins. Phys Chem Chem Phys. 2019;21(5):2463-71.

176. SUBBAIAH TV, STOCKER BA. ROUGH MUTANTS OF SALMONELLA TYPHIMURIUM. I. GENETICS. Nature. 1964;201:1298-9.

177. Sutterlin HA, Shi H, May KL, Miguel A, Khare S, Huang KC, et al. Disruption of lipid homeostasis in the Gram-negative cell envelope activates a novel cell death pathway. Proc Natl Acad Sci U S A. 2016;113(11):E1565-74.

178. Suzuki T, Murai T, Fukuda I, Tobe T, Yoshikawa M, Sasakawa C. Identification and characterization of a chromosomal virulence gene, vacJ, required for intercellular spreading of Shigella flexneri. Mol Microbiol. 1994;11(1):31-41.

179. Suzuki Y, Osawa M, Hatta A, Sueʹtaka W. Mechanism of absorption enhancement in infrared ATR spectra observed in the Kretschmann configuration. Applied surface science. 1988;33:875-81.

180. Swapna LS, Bhaskara RM, Sharma J, Srinivasan N. Roles of residues in the interface of transient protein-protein complexes before complexation. Sci Rep. 2012;2:334.

181. Tam HK, Foong WE, Oswald C, Herrmann A, Zeng H, Pos KM. Allosteric drug transport mechanism of multidrug transporter AcrB. Nat Commun. 2021;12(1):3889.

182. Tang X, Chang S, Qiao W, Luo Q, Chen Y, Jia Z, et al. Structural insights into outer membrane asymmetry maintenance in Gram-negative bacteria by MlaFEDB. Nat Struct Mol Biol. 2021;28(1):81-91.

183. Taylor KA, Glaeser RM. Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. Journal of structural biology. 2008;163(3):214-23.

184. Thong S, Ercan B, Torta F, Fong ZY, Wong HY, Wenk MR, et al. Defining key roles for auxiliary proteins in an ABC transporter that maintains bacterial outer membrane lipid asymmetry. Elife. 2016;5.

185. Tomasek D, Rawson S, Lee J, Wzorek JS, Harrison SC, Li Z, et al. Structure of a nascent membrane protein as it folds on the BAM complex. Nature. 2020;583(7816):473-8.

186. Institute U-D. UCLA Diffraction Anisotropy Server. [cited 2022]. Available from: https://srv.mbi.ucla.edu/Anisoscale/.

187. Consortium U. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021;49(D1):D480-D9.

188. Valvano MA. Export of O-specific lipopolysaccharide. Front Biosci. 2003;8:s452-71.

189. Vollmer W, Bertsche U. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. Biochim Biophys Acta. 2008;1778(9):1714- 34.

190. Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J. Role of a highly conserved bacterial protein in outer membrane protein assembly. Science. 2003;299(5604):262-5.

191. WAKIL SJ, PORTER JW, GIBSON DM. Studies on the mechanism of fatty acid synthesis. I. Preparation and purification of an enzymes system for reconstruction of fatty acid synthesis. Biochim Biophys Acta. 1957;24(3):453- 61.

192. Wang L, Tonggu L. Membrane protein reconstitution for functional and structural studies. Sci China Life Sci. 2015;58(1):66-74.

193. Wang Z, Wang J, Ren G, Li Y, Wang X. Influence of Core Oligosaccharide of Lipopolysaccharide to Outer Membrane Behavior of Escherichia coli. Mar Drugs. 2015;13(6):3325-39.

194. Weber HJ. Stoichiometric measurements of 30S and 50S ribosomal proteins from Escherichia coli. Mol Gen Genet. 1972;119(3):233-48.

195. White SW, Zheng J, Zhang YM, Rock. The structural biology of type II fatty acid biosynthesis. Annu Rev Biochem. 2005;74:791-831.

196. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallographica Section D: Biological Crystallography. 2011;67(4):235-42.

197. Winter G. xia2: an expert system for macromolecular crystallography data reduction. Journal of applied crystallography. 2010;43(1):186-90.

198. Wood C, Burnley T, Patwardhan A, Scheres S, Topf M, Roseman A, et al. Collaborative computational project for electron cryo-microscopy. Acta Crystallogr D Biol Crystallogr. 2015;71(Pt 1):123-6.

199. Woodward R, Yi W, Li L, Zhao G, Eguchi H, Sridhar PR, et al. In vitro bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. Nat Chem Biol. 2010;6(6):418-23.

200. Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell. 2005;121(2):235-45.

201. Xie F, Li G, Zhang W, Zhang Y, Zhou L, Liu S, et al. Outer membrane lipoprotein VacJ is required for the membrane integrity, serum resistance and biofilm formation of Actinobacillus pleuropneumoniae. Vet Microbiol. 2016;183:1-8.

202. Yao J, Rock CO. Phosphatidic acid synthesis in bacteria. Biochim Biophys Acta. 2013;1831(3):495-502.

203. Yao J, Rock CO. Exogenous fatty acid metabolism in bacteria. Biochimie. 2017;141:30-9.

204. Yeow J, Tan KW, Holdbrook DA, Chong ZS, Marzinek JK, Bond PJ, et al. The architecture of the OmpC-MlaA complex sheds light on the maintenance of outer membrane lipid asymmetry in. J Biol Chem. 2018;293(29):11325-40.

205. Yero D, Díaz-Lobo M, Costenaro L, Conchillo-Solé O, Mayo A, Ferrer-Navarro M, et al. The Pseudomonas aeruginosa substrate-binding protein Ttg2D functions as a general glycerophospholipid transporter across the periplasm. Commun Biol. 2021;4(1):448.

206. Yokota N, Kuroda T, Matsuyama S, Tokuda H. Characterization of the LolA- LolB system as the general lipoprotein localization mechanism of Escherichia coli. J Biol Chem. 1999;274(43):30995-9.

207. Young T. I. The Bakerian Lecture. Experiments and calculations relative to physical optics. Philosophical transactions of the Royal Society of London. 1804(94):1-16.

208. Zanetti G, Riches JD, Fuller SD, Briggs JA. Contrast transfer function correction applied to cryo-electron tomography and sub-tomogram averaging. J Struct Biol. 2009;168(2):305-12.

209. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell. 1999;98(6):811-24.

210. Zhou Z, White KA, Polissi A, Georgopoulos C, Raetz CR. Function of Escherichia coli MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. J Biol Chem. 1998;273(20):12466-75.

211. Spencer C, Brown HA. Biochemical characterization of a Pseudomonas aeruginosa phospholipase D. Biochemistry. 2015;54(5):1208-18.

212. Chin JW, Martin AB, King DS, Wang L, Schultz PG. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. Proc Natl Acad Sci U S A. 2002;99(17):11020-4.

213. Wang GJ, Porta C, Chen ZG, Baker TS, Johnson JE. Identification of a Fab interaction footprint site on an icosahedral virus by cryoelectron microscopy and X-ray crystallography. Nature. 1992;355(6357):275-8.

214. Mathis RR, Brown OR. ATP concentration in Escherichia coli during oxygen toxicity. Biochim Biophys Acta. 1976;440(3):723-32.

215. Soini J, Falschlehner C, Mayer C, Böhm D, Weinel S, Panula J, et al. Transient increase of ATP as a response to temperature up-shift in Escherichia coli. Microb Cell Fact. 2005;4(1):9.

216. Galm U, Heller S, Shapiro S, Page M, Li SM, Heide L. Antimicrobial and DNA gyrase-inhibitory activities of novel clorobiocin derivatives produced by mutasynthesis. Antimicrob Agents Chemother. 2004;48(4):1307-12.

217. Doyle MT, Jimah JR, Dowdy T, Ohlemacher SI, Larion M, Hinshaw JE, et al. Cryo-EM structures reveal multiple stages of bacterial outer membrane protein folding. Cell. 2022;185(7):1143-56. e13.

218. Yang C, Peng S, Chen C, Zhao XS. Molecular mechanism of networking among DegP, Skp and SurA in periplasm for biogenesis of outer membrane proteins. Biochemical Journal. 2020;477(16):2949-65.

219. Huang, Y.M.M., Munguia, J., Miao, Y., Nizet, V. and McCammon, J.A., 2019. Docking simulation and antibiotic discovery targeting the MlaC protein in Gram-negative bacteria. Chemical biology & drug design, 93(4), pp.647-652.