Structure-function characterisation of *Chlamydia pneumoniae's* Major Outer Membrane Protein

PhD Structural Biology
School of Biological Sciences

Amy Elizabeth Danson

September 2018

Abstract

Chlamydia pneumoniae is a Gram negative bacterium responsible for a number of human respiratory diseases and linked to some chronic inflammatory diseases. The major outer membrane protein (MOMP) of *Chlamydia* is a conserved immunologically dominant protein located in the outer membrane, which together with its surface exposure and abundance, has led to MOMP being the main focus for vaccine and antimicrobial studies in recent decades. In the absence of structural data for MOMP, the fatty acid transporter of *E. coli*, FadL, was used for the development of a plausible homology model, which suggested a 14-stranded β-barrel. X-ray crystallography has revealed a low-resolution crystal structure at 4 Å, from non-optimised crystals, that showed a β-barrel structure with an occluded pore.

Initial research focussed on producing recombinant stable and active MOMP for higher resolution structural studies. In this work, MOMP's yield was greatly improved through the use of SB3-14 for solubilisation, identified through detergent screening using FSEC and Western blotting. Optimisation of the purification procedure led to crystallisation screening of DDM exchanged MOMP. The resulting crystals were analysed using a synchrotron radiation source. Diffraction patterns indicated that a small selection of crystals were protein, yielding new crystallisation conditions that can be optimised in future research.

Confirmed as a β-barrel protein, MOMP is often considered to have a porin-like function. Homology modelling with FadL, as well as the low resolution structure, indicate an occluded barrel and is suggestive of a more active transport role for MOMP, namely for fatty acids which are known to be scavenged from the host. The ability of MOMP to transport fatty acids was assessed using a novel functional growth assay in *E. coli*, developed in this work. Although MOMP was shown to be expressed in the outer membrane, using fluorescence microscopy, interestingly, transport of fatty acids was not observed.

However, MOMP has a more widely accepted role in the *Chlamydial* outer membrane complex through the formation of intermolecular disulphide bonds, although the exact interactions formed are currently unknown. It was hypothesised in this work, due to the large number of cysteines available for disulphide bonding, that interactions occur with cysteine rich pockets as opposed to individual residues. Molecular dynamics simulations in conjunction with the crystal structure provided an impetus for rational cysteine mutant design. The localisation of MOMP in the *E. coli* membrane was assessed with super resolution fluorescence microscopy, which in support of this novel hypothesis showed a decrease in membrane clustering with cysteine rich regions containing two mutations, indicative of a compensatory mechanism never before conceived.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged

Amy Elizabeth Danson

Acknowledgements

I sincerely thank my supervisors Prof. Kimberly Watson, Dr Sheila MacIntyre, Dr Isabel de Moraes, and Dr Martin Walsh for their guidance, support, and motivation throughout the duration of my PhD research. Thank you all for providing me with ample opportunities to develop my career and for continually inspiring me. Thank you to my supervisory committee Dr Craig Hughes and Dr David Leake, for their fresh perspectives and interesting discussions.

Special thanks go to the number of researchers I have had the pleasure of collaborating with. Thank you to Dr Alice Pollitt at the University of Reading, and Dr Lin Wang and Alex McStea at the Central Laser Facility, for your time and expertise. I gratefully appreciate the support from Dr Anna Duncan in collaboration with Dr Mark Sansom at the University of Oxford. Thank you to my colleagues at the Membrane Protein Lab and my Reading University labs for creating a happy and supportive environment.

Finally, I would like express my sincere gratitude towards my friends and family, for their continual encouragement and advice. I am infinitely grateful to my partner Thomas, who shared all the highs and lows, for his unwavering support and attention.

Conference attendance and presentations

<u>School of Biological Sciences PhD symposium</u>, University of Reading

Presentation: "Using super resolution fluorescence microscopy to characterise the clustering behaviour of *Chlamydia pneumoniae* MOMP" June 2018

Presentation: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein" June 2017. **Awarded prize for best second year biomedical sciences presentation**

Presentation: "The major outer membrane protein of *Chlamydia pneumoniae*" May, 2016

<u>UK-Korean Structural Biology Symposium</u>, Diamond Light Source, August 2018

<u>South West Structural Biology Consortium</u>, University of Exeter, June 2018

Presentation and Poster: "Using super resolution fluorescence microscopy to characterise the clustering behaviour of *Chlamydia pneumoniae* MOMP" **Awarded first place poster prize**

<u>British Crystallographic Association Spring Meeting</u>, University of Warwick, March 2018

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein"

<u>CCP4 Study Weekend</u>, University of Nottingham, January 2018

<u>24th Congress of the International Union of Crystallography</u>, India, August 2017

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein"

<u>CCP4 Study Weekend</u>, University of Nottingham, January 2017

<u>Membrane Protein Symposium</u>, Harwell Science and Innovation Campus, November 2016

<u>8th Meeting of the European Society for *Chlamydia* Research</u>, Oxford, September 2016

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein"

<u>South West Structural Biology Consortium</u>, University of Portsmouth, June 2016

<u>Microbiology Society Annual Conference</u>, Liverpool, March 2016

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein"

<u>CCP4 Data Collection and Structure Solution Workshop</u>, Diamond Light Source, December 2015

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein"

<u>Frontiers in Integral Membrane Protein Structural Biology</u>, University of Oxford, September 2015

Poster: "Structure-function characterisation of *Chlamydia pneumoniae* major outer membrane protein" **Awarded third place poster prize**

<u>Diamond Synchrotron Radiation User Meeting</u>, Diamond Light Source, September 2015

Table of Contents

List of Abbreviations

A	aa	Amino acids
	ASR	Anabaena sensory rhodophsin
	ATP	Adenosine triphosphate
	AU	Arbitrary units
B	BC	Before centrifugation
	BCA	Bicinchoninic acid
	bp	Base pairs
	BSA	Bovine serine albumin
C	C12E8	Octaethylene glycol monododecyl ether
	cAMP	Cyclic adenosine monophosphate
	CAP	Catabolite activator protein
	CD	Circular dichroism
	CDs	Constant domains
	CG	Course grained
	CLF	Central Laser Facility
	ClusDoC	Cluster detection with degree of colocalisation
	CMC	Critical micellar concentration
	COMC	*Chlamydial* outer membrane complex
	CV	Column volume
D	DA-DA	D-alanyl-D-alanine
	DBSCAN	Density based spatial clustering of applications with noise
	DDM	n-Dodecyl-β-D-Maltopyranoside
	DED	Direct electron detector
	DLS	Diamond Light Source
	DNA	Deoxyribonucleic acid
	dNTP	Deoxynucleotides
	DOPG	Dioleoylphosphatidylglycerol
	DPPE	Dipalmitoyl phosphatidylethanolamine
	DTT	Dithiothretol
E	E	Eluted
	EB	Elementary body
	ECL	Electrochemiluminescence
	EGFR	Epidermal growth factor receptor
	EM	Electron microscopy
	ESRF	European Synchrotron Radiation Facility
F	Fad	Fatty acid degradation
	FC-12	Fos-Choline 12

	FPALM	Fluorescence photoactivated localisation microscopy
	FSEC	Fluorescence-detection size exclusion chromatography
G	GFP	Green fluorescent protein
	GlcNAc	N-acetylglucosamine
	GOI	Gene of interest
	GUI	Graphical user interface
H	hpi	Hours post infection
	HPLC	High performance liquid chromatography
	HRP	Horse radish peroxidase
I	IB	Intermediate body
	IGF2	Insulin-like growth factor 2
	IMAC	Immobilised metal affinity chromatography
	IPTG	Isopropyl-thiogalactopyranose
K	Kb	Kilobase
	KDO	3-deoxy-D-manno-octulosonic acid
L	LB	Luria-Bertani broth
	LCFAs	Long chain fatty acids
	LC-MS/MS	Liquid chromatography-mass spectrometry/mass spectrometry
	LDAO	Lauryldimethylamine-N-Oxide
	LMNG	Lauryl maltose neopentyl glycol
	LPS	Lipopolysaccharide
	LOS	Lipooligosaccharide
M	M6P	Mannose-6-phosphate
	MAbs	Monoclonal antibodies
	MALDI-MS	Matrix assisted laser desorption/ionisation mass spectrometry
	MD	Molecular dynamics
	MM	Minimal media
	MOMP	Major outer membrane protein
	MR	Molecular replacement
	MS	Mass spectrometry
	MurNac	N-acetylmuramic acid
	MW	Molecular weight
	MX	Macromolecular crystallography
N	NAD	Nicotinamide adenine dinucleotide
	NADH	Reduced nicotinamide adenine dinucleotide
	nH_2O	Nanopure water
	NLP	Nanolipid particle
	nMOMP	Native MOMP
	NMR	Nuclear magnetic resonance

	NTA	Nitrilotriacetic acid
O	OD	Optical density
	OG	n-Octyl-β-D-glucopyranoside
	OGNG	Octyl glucose neopentyl glycol
	OM	Outer membrane
	OMP	Outer membrane protein
P	P	Pellet
	PALM	Photoactivated localization microscopy
	PBS	Phosphate buffered saline
	PCR	Polymerase chain reaction
	PDB	Protein data bank
	PG	Peptidoglycan
	pI	Isoelectric point
	Pmp	Polymorphic membrane protein
	POI	Protein of interest
	PVDF	Polyvinylidene difluoride
Q	Q-PCR	Quantitative polymerase chain reaction
R	RB	Reticulate body
	ROI	Region of interest
	rMOMP	Recombinant MOMP
	RMSD	Root mean square deviation
S	Sarkosyl	n-Lauroylsarcosine sodium salt
	SB3-12	n-Dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate
	SB3-14	n-Tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate
	SDM	Site directed mutagenesis
	SDS-PAGE	Sodium dodecyl sulphate polyacrylamine gel electrophoresis
	SE	Standard error
	SEC	Size exclusion chromatography
	SF	Soluble fraction
	SiR	Silicon rhodamine
	SRCD	Synchrotron radiation circular dichroism
	STORM	Stochastic optical reconstruction microscopy
T	T3SS	Type III secretion system
	TCA	Tricarboxylic acid
	TCEP	Tris(2-carboxyethyl)phopshine
	TIRF	Total internal reflection fluorescence
U	U	Unbound
	UDM	n-Undecyl-β-D-Maltopyranoside

	UoR	University of Reading
	UV	Ultraviolet
V	VD	Variable domain
	VDAC	Voltage-dependent anion channel
	vdW	van der Waals
X	XRD	X-ray diffraction
	θ_C	Critical angle

Appendices

List of Tables

Chapter 5

Chapter 1

Introduction

1. Introduction

1.1. *Chlamydia pneumoniae*

Chlamydia pneumoniae is a Gram negative obligate intracellular bacterium (Grayston, 2000) responsible for 10% of community-acquired pneumonia in humans (Grayston *et al.*, 1989) as well other respiratory diseases such as sinusitis and bronchitis. *C. pneumoniae* also has an often-debated role in inflammatory diseases such as atherosclerosis (Belland *et al.*, 2004; Saikku *et al.*, 1988; Sorrentino *et al.*, 2015), reactive arthritis (Carter *et al.*, 2012; Rahman *et al.*, 1992; Zeidler and Hudson, 2016) and asthma (Hahn and McDonald, 1998; Webley and Hahn, 2017). The majority of *C. pneumoniae* infections occur from early childhood through to teenage years, leading to a high prevalence within the population with antibodies found in 75% of a Tawainese population and 63% in Spanish citizens (Wang and Grayston, 1990). This highlights the importance of early intervention for adequate control and in reducing *Chlamydia* prevalence worldwide. Due to the poor efficacy observed with antibiotic treatments, this is most likely achievable through development of a robust vaccine. The overall aim of this project is toward the elucidation of the structure and functional properties of the major outer membrane protein (MOMP) of *C. pneumoniae*, which has been identified as a promising vaccine candidate despite little being known of its structure and function.

1.2. The biphasic lifecycle of *Chlamydia*

Unique to *Chlamydia* is its biphasic lifecycle, which is depicted in **Figure 1.1**. Starting with the smaller electron dense infectious form known as the elementary body (EB), *Chlamydia* circulates the host system until contact with a mucosal epithelial cell occurs. The type III secretion system (T3SS), which is considered to be essential for virulence in a range of Gram

negative bacteria, is often described as a 'molecular syringe' which acts to insert bacterial effector proteins directly into the host (Betts-Hampikian and Fields, 2010). The metabolically inactive EB, which contains a nucleus condensed by histone proteins HctA and HctB, becomes internalised following delivery of a T3SS protein into the host cell which initiates actin recruitment (AbdelRahman and Belland, 2005). The EB is then initially surrounded by a protective membrane known as the *Chlamydial* vacuole which then matures within the vicinity of the host nucleus, Golgi apparatus and endoplasmic reticulum in order to obtain components necessary for development, such as sphingomyelin, cholesterol and iron (Carabeo *et al.*, 2002; Schoborg, 2011) eventually forming an inclusion. Within this inclusion EBs transition to an intermediate body (IB) before finally differentiating into the metabolically active form, the reticulate body (RB), where the cell size begins to increase as metabolic proteins are upregulated and cells initiate division by binary fission (Wyrick, 2010). These three forms are shown in the electron micrograph of *C. trachomatis* in **Figure 1.2**. After around 60-72 hours for *C. pneumoniae* and 9-11 regenerations, the RB start to redifferentiate back into the EB, which accumulate and eventually exit the cell via exocytosis or whole cell lysis, facilitating infection of more host cells (Brunham and Rey-Ladino, 2005; Hatch *et al*., 1984). Under stress conditions, the RB can enter into a reversible persistent state known as aberrant bodies and, whilst unable to replicate, the cells remain viable until the stimulus is removed, therefore leading to chronic infection (Matsumoto and Manire, 1970). Interferon gamma (IFNγ) is the most notable stimulant, but also nutrient deprivation, the herpes simplex virus, and iron starvation have all been shown to induce this persistent form (Wyrick, 2010). The genetic regulation of these persistent forms is particularly difficult to study, as the cell phenotype can vary hugely depending on the

stimulus. Due to this biphasic lifecycle, it is necessary for a comprehensive vaccine to target all stages of the cell cycle in order to produce an effective immune response.

Mucosal epithelial cell

Inclusion
formation

48h

RB differentiation
back to EB

Differentiation
into RB

12-24h

60-72h

Exocytosis

EB entry

96h
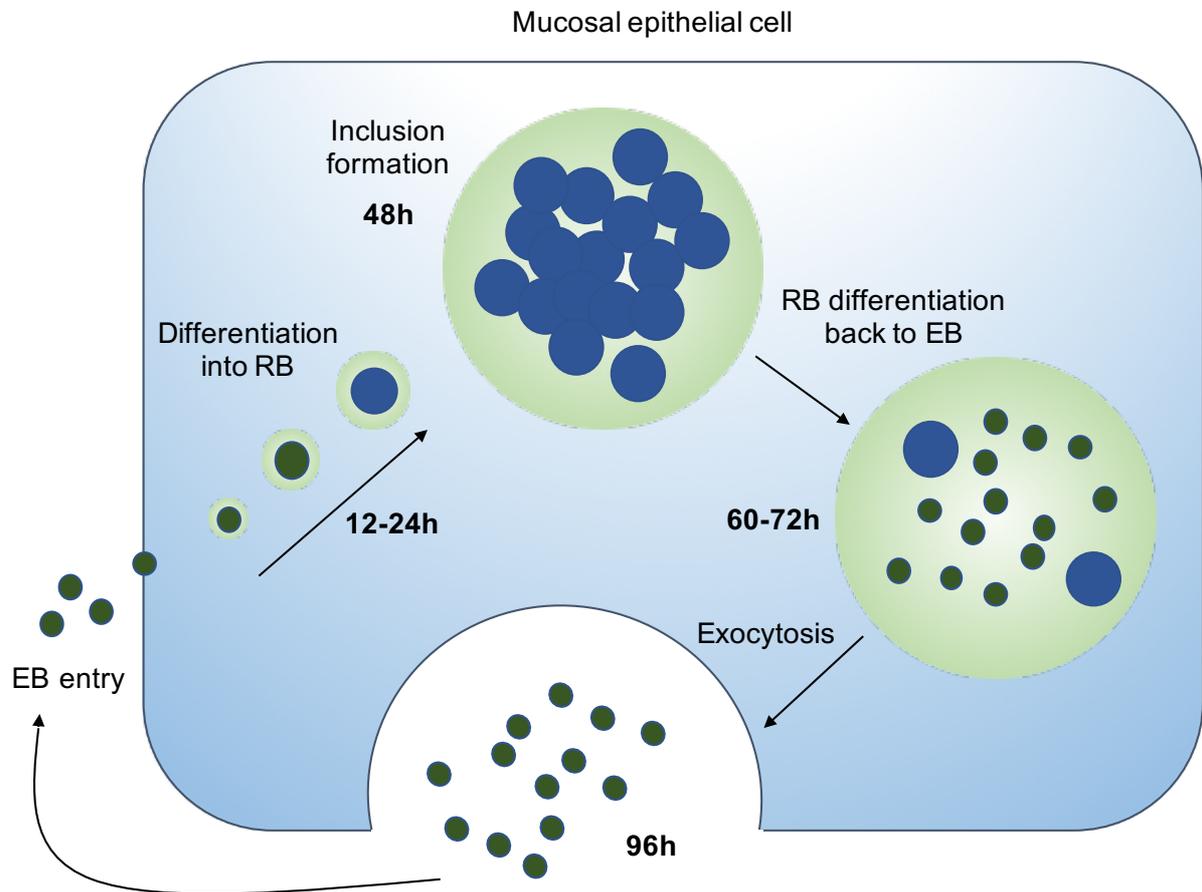
Figure 1.1. The biphasic developmental lifecycle of *Chlamydia pneumoniae.*
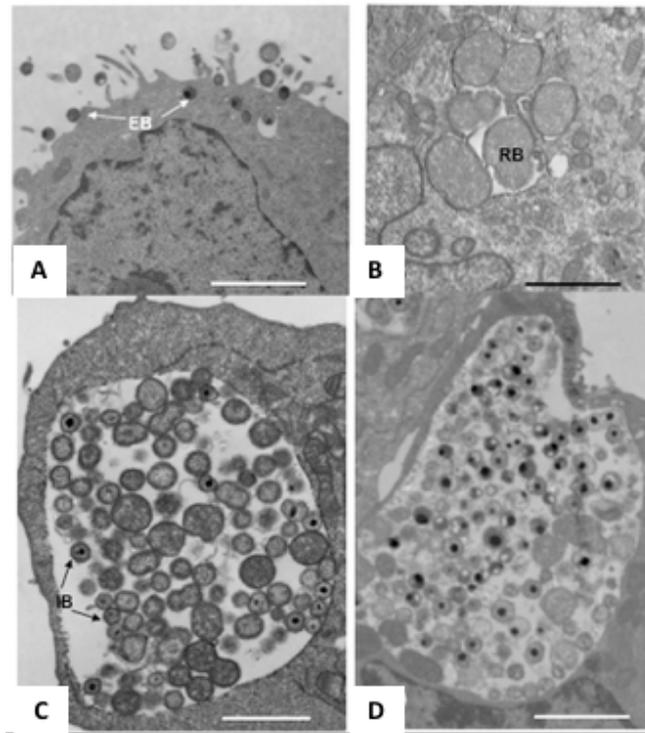
Figure 1.2. The stages of the *C. trachomatis* developmental cycle. **A**, One hour post infection (hpi) the EB attaches to infected host endometrial epithelial cells. **B**, 8-12 hpi an early inclusion forms containing RB. **C**, Mid-development cycle containing intermediate bodies (IB). **D**, Late developmental inclusion containing mainly EB. Scale is 2 μm. (Wyrick, 2010)

1.3. *Chlamydia* metabolism

Chlamydia gene expression is temporally regulated between the two stages of the developmental lifecycle due to the differing roles of the EB and RB cells, in survival and replication respectively, which has been previously investigated with microarray analysis and quantitative-PCR (Q-PCR). Typically, the RBs are recognised as the more metabolically active form, whilst the EB's function is in host dissemination. *Chlamydial* genes belong to one of four main classes; early, mid, late and tardy phase expression, with an additional persistent stage for aberrant bodies, for which the regulatory cycles for *C. pneumoniae* are comprehensively displayed in **Figure 1.3** (Roulis *et al.*, 2013). *C. pneumoniae* takes between 60-96 hours to fully complete a cell cycle, compared to just 36-48 hours for *C. trachomatis*. *Omp1*, the gene encoding MOMP also referred to in some instances as *ompA*, along with

those encoding other outer membrane proteins such as OmcA and PorB, is downregulated following internalisation of the EB allowing the RB to focus on expression of proteins required for cell replication and metabolism (Roulis *et al.*, 2013). During this early stage genes that are involved in DNA and protein biosynthesis, such as *groEL3*, *dnaX2*, and *xerC*, are expressed, followed by those required for biotin synthesis (*bioB*) in the mid-phase and glycolysis (*pyk*) in the late stage (Roulis *et al.*, 2013). In preparation for release during the tardy stage at 60-72 hpi, *hctB* for DNA condensation and *omp1* are then upregulated along with *omcA*, *omcB*, and *porB*, signifying that MOMP's main role is focussed around the function of the EB rather than the RB. *Omp1* is also clustered with thirteen other genes, namely other OMPs and a component of the T3SS, *yscC*, which is suggestive of linked transcription regulation (Nicholson *et al.*, 2003). As *Chlamydia* have a reasonably small genome size of just one million base pairs (bp) compared to 4.6 million in *E. coli* (Stephens *et al.*, 1998), some biosynthetic pathways are truncated, such as the tricarboxylic acid (TCA) cycle, whilst others are missing altogether, as in the case of the nicotinamide adenine dinucleotide (NAD) biosynthetic pathway (Käding *et al.*, 2014). As a result, *Chlamydia* is in possession of Npt1 and Npt2 which are able to scavenge NAD/NADH and ATP from the host cells, with the former also present in the EB stage of the lifecycle indicating that acquisition can occur during both stages of the biphasic lifecycle, although it is currently unknown how these substrates cross the inclusion membrane (Fisher *et al.*, 2013). *Chlamydia* has also been shown to drive host cell metabolism in order to obtain the desired substrates, for example through the upregulation of tryptophan biosynthesis enzymes in the early stage and glucose metabolism genes in the mid stage resulting in a 2-3 fold increase in substrates (Käding *et al.*, 2014).

	Early ↑	Early ↓	Mid ↑	Mid ↓	Persistent ↑	Persistent ↓	Late ↑	Late ↓	Tardy ↑	Tardy ↓
Membrane	pmp5.2 pmp14	ompA omcA omcB porB	pmp19 incC	pmp11 pmp14 pbp2	omcB (IFN) pmp19	omcB (Fe) ompA (Fe) porB pmp11	pbp2	pmp19 pmp5.2 incC	ompA omcA omcB porB	pmp19 pbp2
Energy/ Co-factors	euo htrB **yciA** pfkA2 **pgsA1**	pyk	htrA **_bioB_** **_phhA_** atpI dskA	pgi pfkA2 hctB **_guaB_**	htrA (Pen) tyrP (IFN) gspD (Fe) euo (Fe) atpB	htrB gspD (IFN) hctB euo (IFN) atpI	pyk	dskA htrB htrA yciA **pgsA1**	**pgi** hctB **_guaB_**	euo **_bioB_** **_phhA_** atpI
Transcription/ Translation	groEL3 **dnaX2** **xerC**	RpoA infC groEL1 groES	RpoA infC infB	dnaK groEL1 rpsD	RpoA (IFN) infC ptsH gatA groEL(IFN)	RpoA (Fe) dnaK groEL hflX infB	groES **groEL1** infB	**dnaX2** **xerC** groEL3	dnaK rpsD	RpoA infC infB
Cell division	murE	murA ftsK	ftsW rodA	ftsH	ftsK (Fe) ftsW murA	ftsK (IFN) ftsH	murA murB ftsK	ftsW murE rodA	rodA ftsH	ftsK murA murB
Secretion	tsp gspD	secA2 lcrH1 lcrH2	lcrH2	yscC tsp lcrE gspD	secA1 yscC sctN	secA2 lcrH1 lcrH2	secA2 gspD		yscC lcrH1 tsp **lcrE**	secA2 lcrH1 lcrH2 gspD
Transport	oppA4 oppB1	oppA1 dppF1	dppF1	oppB1	oppA4 oppB1 dppF1 ytgA	oppA1 oppB2 dppF2 ytgD	oppB2 dppF2	oppA1	oppA1	oppA4 oppB2 dppF1 dppF2

TRENDS in Microbiology

Figure 1.3. Temporal gene expression in _C. pneumoniae_. Genes in bold indicate those with different expression to _C. trachomatis_, and bold underlined genes are indicative of variations between _C. pneumoniae_ strains. (Roulis _et al._, 2013)

1.4.   The *Chlamydia* outer membrane

Similar to other Gram negative bacteria, *C. pneumoniae* consists of an inner membrane (IM),

separated from the outer membrane (OM) by the periplasmic space (Stephens and Lammel,

2001). Within the OM of *Chlamydial* species, the lipopolysaccharide (LPS) is truncated,

resulting in a terminal 3-deoxy-D-manno-octulosonic acid (KDO) trisaccharide, identified by

nuclear magnetic resonance (NMR) spectroscopy (Rund *et al*., 1999) as shown in **Figure 1.4A**

and instead referred to as a lipooligosaccharide (LOS). The typical structure of LPS in Gram

negative bacteria, as shown in **Figure 1.4B**, extends beyond the KDO groups to include a

core containing a variety of sugar moieties and the O-antigen region, which is often

associated with pathologies such as sepsis (Steimle *et al*., 2016) and acts to protect the

cellular integrity.
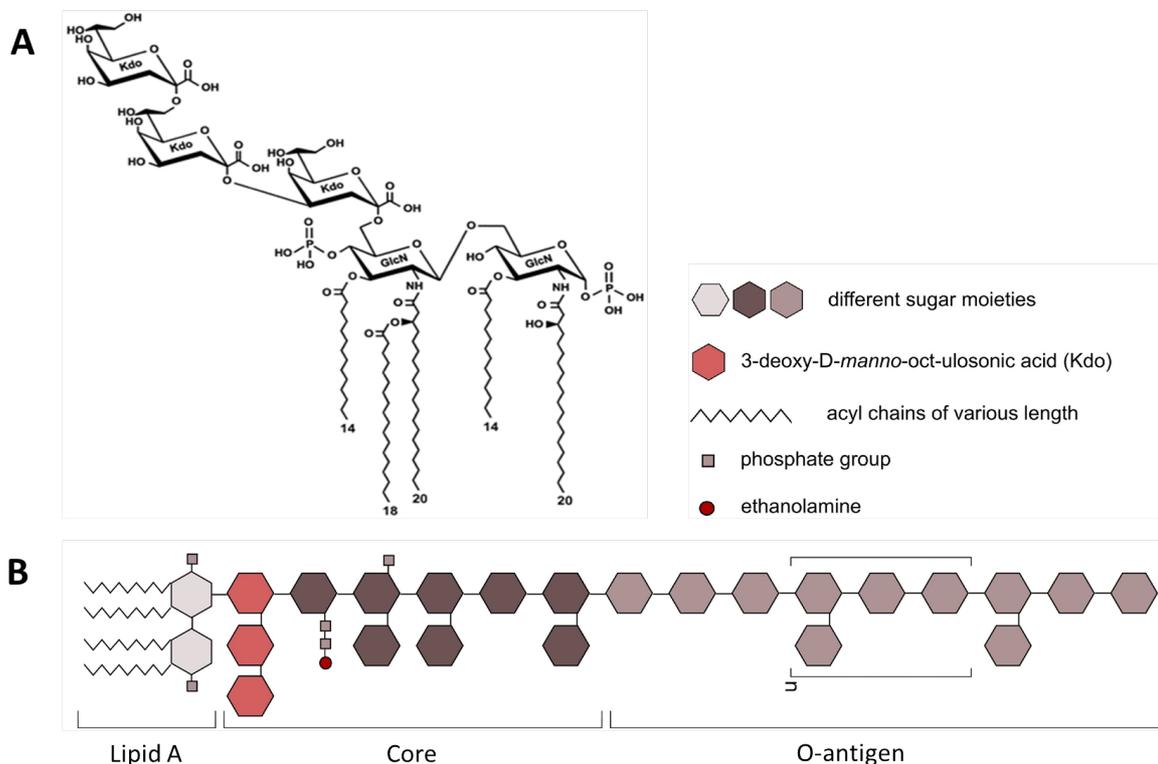
Figure 1.4. A, the chemical structure of LOS in *C. trachomatis* L2. NMR was used to determine the phosphorylated carbohydrate backbone and matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS) to determine the acylation pattern (Rund *et al*., 1999). **B, a schematic displaying the typical structure of LPS in Gram negative bacteria**. Note the orange hexagons at the beginning of the core section are where *Chlamydial* LOS is truncated. Adapted from Steimle *et al*., 2016.

Peptidoglycan (PG), typically found between the two membranes of Gram negative bacteria, is a sugar polymer consisting of alternate glycan residues; N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNac), whose purpose is to confer cell structural stability. These two residues are connected via β-(1,4)-glycosidic bonds which then becomes inserted into the growing polymer chain through a transglycosylation reaction, before transpeptidation between the pentapeptide side chains and the MurNaC crosslinks the strands to complete the net-like layer within the periplasmic space (Ghuysen and Goffin, 1999). During pentapeptide chain assembly both L- and D-amino acids, such as D-alanine and D-glutamic acid, are incorporated. As mammalian cells can only utilise L-amino acids, the enzymes that synthesise and integrate these D-amino acids into PG act as unique targets to antibiotics such as D-cycloserine and β-lactams (Liechti *et al*., 2014).

However, in the past three decades scientists have debated heavily over the '*Chlamydia* anomaly', whereby *Chlamydia* was both susceptible to penicillins and PG-targeting enzymes such as lysozyme and lysostaphin (How *et al*., 1984), as well as being in possession of the genetic elements required for PG synthesis (McCoy and Maurelli, 2006), but where contrarily only minimal levels of PG had actually been detected. This perplexing incongruity was recently clarified by Liechti *et al* using fluorescent microscopy. This experiment employed a click-chemistry reaction whereby a component essential for PG pentapeptide crosslinking, D-alanyl-D-alanine (DA-DA), was labelled with an azide group that was then selectively captured in conjunction with an Alexa fluorophore (Liechti *et al*., 2016). This technique revealed that the majority of the labelled PG of *C. trachomatis* L2 was localised at the cellular division plane (the septum) of the dividing RBs, as shown in **Figure 1.5**, indicative of a novel role for PG in growth and cytokinesis in *Chlamydia* (Liechti *et al*., 2016). Upon entry into the host cells an inclusion body forms creating an osmotically stable environment,

inside which the RBs undergo metabolic processes and division, and therefore a rigid envelope is not required during the RB stage of the lifecycle. The limitation of PG to the septal growth regions of RBs and the apparent absence of a protective PG layer encompassing the cell within EBs has now refocussed efforts into understanding how *Chlamydia* may obtain its structural stability, particularly in the circulating EBs of the biphasic lifecycle, since the PG layer is essential in other bacteria for the maintenance of cell shape as well as for protection against osmotic stress.
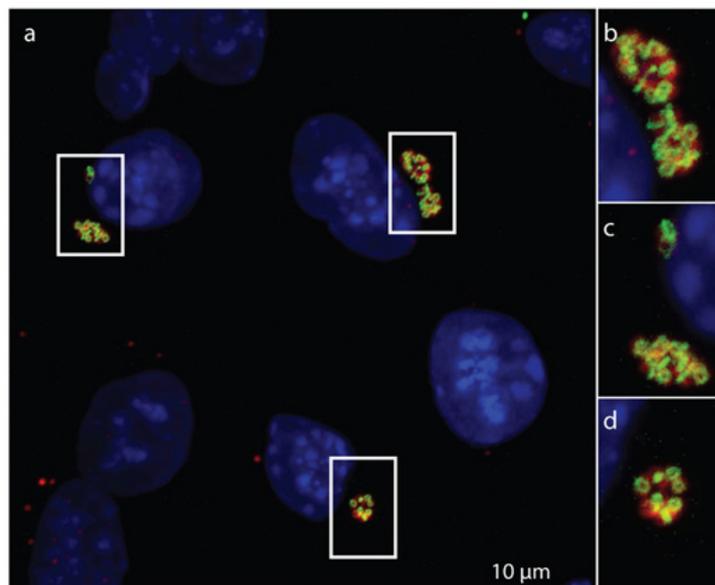
Figure 1.5. Localisation of PG in *C. trachomatis* infected mouse fibroblast cells. Labelling of PG with alkyne- DA-DA analogue probe EDA-DA. Click-chemistry was used to bind the probe to Alexa 488 (green), DAPI (blue) used to stain the nuclei, and antibody (red) to label MOMP for identification of EB and RB cells. **A,** Boxes show the location of the *Chlamydial* inclusion. **B-D,** Magnification of *Chlamydial* inclusion areas. PG can be seen as a ring bisecting the MOMP labelled *Chlamydial* cells. (Liechti *et al*., 2016)

For a number of years, it has been known that the *Chlamydia* species possess a *Chlamydial* outer membrane complex (COMC), a mesh of cysteine rich OMPs linked through intermolecular disulphide bonding, also referred to as the P-layer within the outer membrane of the EB cells (Bavoil *et al*., 1984; Hatch *et al*., 1984; Hatch, 1996; Newhall *et al*.,

1983), 61% of which is made up of MOMP (Caldwell *et al.*, 1981). It is the COMC that is likely to replace PG in providing *Chlamydia* with its required structural stability, during the EB stage of the lifecycle. Recently, a total of 17 cysteine-rich proteins were identified in the *C. trachomatis* L2 COMC, using mass spectrometry (**Table 1.1**), in which MOMP and OmcB were reportedly most abundant (Liu *et al.*, 2010). To further demonstrate the significance of disulphide bonding in formation of the mesh layer, another OMP lacking cysteine residues, PmpD, did not form part of the COMC despite being expressed in the EB outer membrane (Liu *et al.*, 2010).

L2a Gene	Protein	Mass (Da)	Number of amino acids	Number Cysteine residues	Cysteine content (%)	Average spectral counts
ctl0050	MOMP	43,036	394	9	2.28	200
ctl0702	OmcB	60,783	553	24	4.34	134
ctl0250	PmpG	108,032	1012	15	1.48	46
ctl0251	PmpH	108,119	1006	15	1.49	35
ctl0248	PmpE	105,500	965	15	1.55	18
ctl0043	PulD/YscC	100,574	921	7	0.76	15
ctl0626	OprB	49,929	442	9	2.04	12
ctl0887	CTL0887	48,560	174	4	2.30	11
ctl0082	PorB	37,839	340	8	2.35	9
ctl0703	OmcA	10,013	88	13	14.77	9
ctl0071	PmpB	184,085	1749	19	1.09	9
ctl0671	PmpC	187,729	1774	14	0.79	8
ctl0249	PmpF	113,209	1032	13	1.26	8
ctl0541	CTL0541	42,241	381	8	2.10	8
ctl0493	OMP85	88,943	792	4	0.51	4
ctl0645	CTL0645	47,018	408	13	3.19	2
ctl0863	Pal	21,532	188	4	2.13	1

Table 1.1. The 17 COMC proteins identified in *C. trachomatis* L2. Proteins were identified using liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) are ordered based on their abundance (average spectral counts). Expanded from Liu *et al.*, 2010.

As MOMP has been identified as one of the leading proteins in the COMC, it is almost certainly forming disulphide bonds, either with other MOMPs or other cysteine rich OMPs.

Currently, there are no available structures for MOMP in the Protein Data Bank (PDB) from any of the *Chlamydia* species, and therefore structural predictions are largely based on homology modelling. This in turn proves difficult when hypothesising which cysteine residues are likely to have key roles in disulphide bond formation. Sequence data can aid identification of cysteine proximity for determination of intramolecular disulphide bonds, but as there is a paucity of 3D structural data for both MOMP and the additional sixteen proteins identified in the COMC, bioinformatics analysis of any likely intermolecular contacts is extremely challenging.

1.5.   Genetic studies in *Chlamydia*

The study of structures such as the COMC, as well as the structure-function relationships of *Chlamydia's* prominent proteins has been hindered by the complex biphasic developmental cycle. In order to investigate the COMC, the EB phase of the life cycle requires isolation from the morphologically distinct RB stage, which in itself is a complex process due to their overlapping expression patterns. Not only this, but the presence of the stabilising COMC in EBs is likely to prevent the uptake of foreign DNA. Additionally, due to *Chlamydia*'s parasitic nature, host cells are required for bacterial growth and proliferation, again adding a layer of complexity to *Chlamydial* research not only due to infection and growth, but also in facilitating the transfer of DNA across a multitude of membranes, including that of the host, the inclusion, and the *Chlamydial* outer and inner membranes. Although inefficient, DNA transformation is now achievable via calcium chloride treatment and electroporation, even in EBs although at an extremely low frequency, and improvement upon this procedure will enable the application of other techniques such as transposon mutagenesis, which can provide insight into the function of the *Chlamydial* genes (Bastidas and Valdivia, 2016).

Random chemical mutagenesis is an additional method being explored (Heuer *et al.*, 2007) which coupled with whole genome sequencing can identify the genetic variants produced leading to the creation of a mutant library (Kokes *et al.*, 2015). However, despite these successes, genetic manipulation is still lagging. One obstacle is the restrictions placed on antibiotic selection in research for the generation of recombinant strains, due to their importance in treating clinical cases of *Chlamydial* infection, with an additional setback being the requirement to surpass more than one lipid bilayer as well as to limit toxicity to the host cells (Bastidas and Valdivia, 2016). Transformation with a recombinant plasmid is also problematic due to the presence of native plasmids in some species, with over five copies per cell, and as a result plasmid replication factors favour the smaller native plasmids thus decreasing the transformation efficiency.

1.6. The Major Outer Membrane Protein

MOMP of *C. pneumoniae*, also commonly and incorrectly referred to as OmpA in *Chlamydia*, is a 40 kDa β-barrel protein accounting for 60% (w/w) of the total protein mass found in the OM of EBs (Perez-Melgosa *et al.*, 1991). Although 2-3% of the Gram negative bacterial genome encodes transmembrane β-barrel proteins, they account for only 1% of the solved protein structures from Gram negative bacteria (Freeman *et al.*, 2011). Being the most abundant OMP in the *Chlamydial* membrane, it is suspected that MOMP plays an important role in the human immune response and, therefore, in recent years research has been heavily focussed on determining the exact structure of MOMP in order to design more effective treatments against *Chlamydia* (Kari *et al.*, 2009; Olsen *et al.*, 2015; O'Meara *et al.*, 2017; Xia *et al.*, 2013). MOMP appears to be unique to *Chlamydia* due to the paucity of homologous proteins and is considered conserved among the three human species, with

between 64-98% identity and only minor differences in sequence lengths, which range from 389 to 402 amino acids (aa) (as shown in **Appendix 1 Figure A1.1**). As a result, research conducted on MOMP from one *Chlamydial* species is likely to be significantly beneficial to the study of the other MOMPs. However, due to the difficulties in expression and purification of stable and folded recombinant MOMP, resulting from the high cysteine content and the requirement to extract MOMP from its native bilayer with detergents, there are currently no published crystal structures available for this protein across the whole species, leading researchers to rely on other means for structural determination, including *in silico* prediction methods.

1.6.1. Secondary structure

It has long been hypothesised that MOMP is a β-barrel outer membrane protein, due to its porin-like function identified in 1984 by Bavoil *et al*. This was confirmed much later with the help of investigative analytical technique circular dichroism (CD) (Cai *et al*., 2009; Sun *et al*., 2007; Wen *et al*., 2016; Wyllie *et al*., 1998), with experiments in the far-UV providing valuable information regarding the protein secondary structure, which in this case was predominantly β-sheet consistent with a β-barrel structure. The CD acquired data was compared, by a series of algorithms, with a database of model reference proteins to assign the most probable features; Sun *et al* predicted a β-sheet content of 38-44% for native MOMP (nMOMP). Wen *et al* analysed both native and recombinant MOMP (rMOMP) using CD, which revealed a peak at 215 nm on both spectra, indicative of β-sheet secondary structure (Wen *et al*., 2016). This data supports other predictive methods used to discern MOMP's secondary structure, such as hydrophobicity plotting and artificial neural networking (Rodríguez-Marañón *et al*., 2002). However, the number of β-strands is still

contested and will continue to be until a high-resolution crystal structure emerges. Nevertheless, it is widely believed to be in the region of 14-18 strands. This is similar to other well characterised OMPs, such as *E. coli* proteins; FadL, a 14-stranded barrel, which functions as a fatty acid transporter (van den Berg *et al*., 2004); OmpF, a 16-stranded trimeric porin (Cowan *et al*., 1992); and LamB, an 18-stranded trimeric maltose porin (Schirmer *et al*., 1995). Shared homology to these proteins may provide some insight into MOMP's possible function.

1.6.2. Variable domains

Whilst MOMP is considered conserved, different serovars and species have been classified based on regions within the MOMP sequence known as variable domains (VDs), for which there are four, flanked by the conserved constant domains (CDs). Addition of monoclonal antibodies (MAbs) to *C. trachomatis* serovar F, analysed with an immuno-dot blot assay, revealed that VDII and VDIV are surface accessible on MOMP (Wang *et al*., 2006) and, therefore, are therefore likely to promote mutation in order to evade the immune system. Statistical analysis of the genetic differences observed in the VDs of *C. trachomatis* serovars revealed that almost 94% of mutations were nonsynonymous, as opposed to approximately half of the mutations in the CDs which were silent, confirming that the variability is mainly targeted to the VDs (Nunes *et al*., 2009). *C. trachomatis* serovar C MOMP was later analysed by a number of bioinformatics methods, which predicted that the VDs exist external to the transmembrane β-strand region (Feher *et al*., 2013). Mutation of surface exposed regions of MOMP may be an evolutionary mechanism adopted by *Chlamydia* to assist in immune evasion, but until a high resolution crystal structure emerges the exact location of these domains within the barrel structure will remain speculative.

Additionally, MOMP is characteristically cysteine rich, containing nine cysteine residues, of which eight are conserved between *C. pneumoniae* and *C. trachomatis*, constituting 2.3% of the total protein. Whilst both species have the same number of cysteines, *C. pneumoniae* has one at position C342 whilst *C. trachomatis* has one at C101. The formation of intermolecular disulphide bonds between the cysteine residues has been suggested to be responsible for the high osmotic stability observed in EBs, as discussed in **Section 1.4**, which may compensate for the apparent lack of PG (Perez-Melgosa *et al.*, 1991). The amino acid sequence of *C. pneumoniae* AR39 MOMP is displayed in **Figure 1.6A**, where these key regions are highlighted and also displayed on the MOMP homology model (**Figure 1.6B**), which is discussed in greater depth in **Section 1.6.4**.

A

10	20	30	40	50	60	
MKKLLKSALL	SAAFAGSVGS	LQALPVG<u>NPS</u>	DPSLLIDGTI	WEGAAGDP**C**D	P**C**ATW**C**DAIS	

70	80	90	100	110	120	
LRAGFYGDYV	FDRILKV**DAP**	**KTFSMGAKPT**	**GSAAANYTTA**	**VDRP<u>NPA</u>YNK**	HLHDAEWFTN	VD I

130	140	150	160	170	180	
AGFIALNIWD	RFDVF**C**TLGA	SNGYIRGNST	AFNLV**GLFGV**	**KGTTVNANEL**	**PNVSLSNGVV**	VD II

190	200	210	220	230	240	
ELYTDTSFSW	SVGARGALWE	**C**G**C**ATLGAEF	QYAQSKPKVE	ELNVI**C**NVSQ	**FSVNKPKGYK**	VD III

250	260	270	280	290	300	
GVAFPLPTDA	**GVATATGTK**S	ATINYHEWQV	GASLSYRLNS	LVPYIGVQWS	RATFDADNIR	VS IV

310	320	330	340	350	360	
IAQPKLPTAV	**LNLTAW<u>NPS</u>L**	**LGNATALSTT**	**DSFSDFM**QIV	S**C**QINKFKSR	KA**C**GVTVGAT	VD IV

370	380	
LVDADKWSLT	AEARLINERA	AHVSGQFRF

B

Figure 1.6. The predicted variable domains of MOMP. A, Amino acid sequence of *C. pneumoniae* AR39 MOMP. NPA/NPS motifs, underlined; cysteine residues, red; putative VD I-IV, bold; five CDs located between VDs. **B, Mapped variable domains on the *C. pneumoniae* MOMP homology model**. Homology model is described in greater depth in **Section 1.6.4**. VDs are shown in green. (Atanu *et al.*, 2013).

17

1.6.3. Oligomeric state

MOMP's native oligomeric state is generally hypothesised as trimeric, as previously determined by migration patterns on SDS-PAGE gels (Feher *et al.*, 2013; Sun *et al.*, 2007). Although the band observed by Sun *et al* at 67 kDa does not correlate with a trimeric molecular weight for MOMP, whose monomeric MW is approximately 40 kDa, they have identified other OMPs with well characterised trimers exhibiting similar migration patterns. These include OmpF and PhoE from *E. coli*, both with MW's of 39 kDa, which migrate in a similar manner (Sun *et al.*, 2007). Supported by the research conducted by Newhall *et al* in 1983, which indicated that trimeric MOMP was stabilised with disulphide bonds through reducing and non-reducing SDS-PAGE gels, this complements the suggestion that MOMP forms clusters through intermolecular disulphide bonding, which is believed to provide structural stability to the *Chlamydial* cell during its infectious stage (discussed in greater depth in **Chapter 5**).

1.6.4. Homology modelling of MOMP

Homology modelling for MOMP has been somewhat problematic over the years, mainly due to the paucity of highly homologous proteins, exacerbated by the lack of structural data for the most closely related protein PorB of *Chlamydia*, which shares 22% identity. Therefore, it is necessary to extend the search to a wider field of β-barrel proteins with a lower degree of similarity and identity. One such homology model proposed by Atanu in 2013 was based on the *E. coli* fatty acid transporter, FadL, with 16.9% identity and 39.6% similarity (sequence alignment shown in **Appendix 1 Figure A1.2**), selected due to the availability of high quality structural data to 2.6 Å along with the greatest homology to MOMP. The main features of this model, as shown in **Figure 1.7A**, are a 14-stranded barrel with an externally exposed α-

helical loop, an N-terminal occluded region of 37 aa at the interior of the barrel at the periplasmic side, which has been named the 'plug domain', and the NPS and NPA motifs common to transport proteins, the location of which are also shown in **Figure 1.7A**. Additionally, all nine cysteine residues are predicted to be within β-strands, supporting the hypothesis that MOMP forms intermolecular disulphide bonds (Atanu *et al.*, 2013). According to this model, only VDII is externally located in the α-helical loop region, with the remaining three domains situated within the barrel (**Figure 1.6B**). Whilst it is generally considered that the VDs are responsible for antigenicity and are therefore surface exposed, research conducted by Baehr *et al* on *C. trachomatis* MOMP indicated that only VDII and VDIV were susceptible to protease treatment, which suggests otherwise (Baehr *et al.*, 1988), and thus the model was not fit around these constraints. This model is somewhat different to that suggested by Feher *et al*, published in the same year, which instead features 16 β-strands, with neither a protruding α-helical loop region nor a plug domain occluding the barrel (**Figure 1.7B**). As opposed to the method adopted by Atanu *et al*, which identified a single homologous *E. coli* protein through structure-based prediction software packages IntFold and Phyre to derive a model, Feher *et al* utilised the topology prediction software TMBpro, which in this instance resulted in some ambiguity in the C-terminal region. Nevertheless, TMBpro is a valuable tool as the pipeline pulls information solely from transmembrane protein data and is not reliant on sequence similarity. It is imperative that all prediction and modelling algorithms are updated regularly as previous constraints restricted β-barrels to an even number of strands. However, based on experimental evidence, the crystal structure of a voltage-dependent anion channel (VDAC) was published in 2008 showing a structure with 19 β-strands (PDB code: 2JK4), thus throwing these widely followed constraints into question. The differences between these two models, published

within a few months of each other, again highlights the difficulties faced in structure prediction of proteins with no reliable or biologically relevant homologues. In this work, the homology model produced by Atanu *et al* will be referenced and used, as it specifically relates to *C. pneumoniae* AR39 MOMP, which is the MOMP to be used throughout this thesis work.
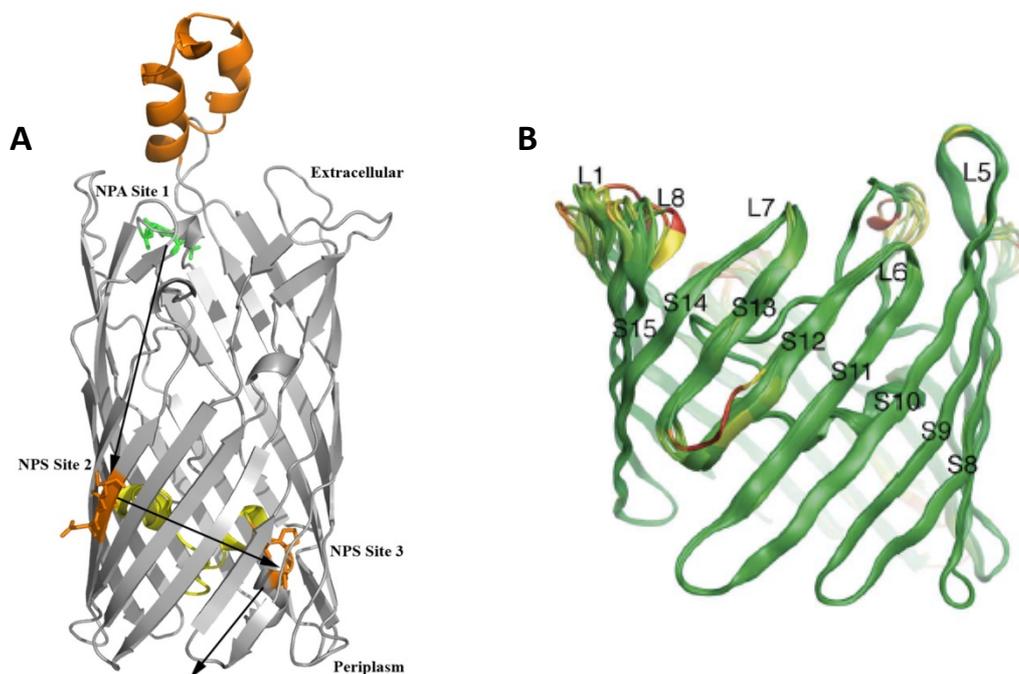
Figure 1.7. A, *C. pneumoniae* **AR39 MOMP homology model based on FadL of** *E. coli***.** A 14-stranded β-barrel with the alpha helical loop shown on the extracellular side in orange and the barrel occlusion in yellow at the periplasmic side. NPA and NPS motifs are also labelled, with arrows suggesting a possible solute transport route. (Atanu *et al*., 2013). **B,** *C. trachomatis* **MOMP homology model constructed using TMBpro.** A 16-stranded β-barrel showing the root mean square deviations (RMSD) for backbone heavy atoms to the average chains, with low as green and high as red. Strand (S) and loop (L) numbers are indicated. (Adapted from Feher *et al*., 2013).

1.6.5. Current structural information regarding MOMP

Currently a low-resolution crystal structure for rMOMP of *C. pneumoniae* has been solved using molecular replacement (MR) to 4 Å (**Figure 1.8**) (Atanu, 2014), following collection of small wedges of data which were then merged using crystallographic software BLEND (Foadi *et al*., 2013). The collection of small 10-15 ° wedges prevents large scale radiation damage,

20

permitting the collection of a greater number of data sets before the crystal lattice becomes irreparably damaged, which is especially useful for delicate samples or microcrystals that are smaller than 10 μm. Previous synchrotron radiation circular dichroism (SRCD) experiments established MOMP's β-sheet secondary structure (Atanu, 2014), which was confirmed through X-ray diffraction (XRD), with the barrel shape clearly visible within the density (**Figure 1.9A**). The structure was solved by MR using the FadL homology model, previously described in **Section 1.6.4**, with **Figure 1.9A** demonstrating a view into the barrel from what would be the periplasmic side, where it is clear that the barrel contains an occluded region near the N-terminus of MOMP. Not only is this structure supportive of the 14-stranded homology model but also of the hypothesis that MOMP does not act as a general porin but may have a more specific role in ligand transport. As hypothesised from the homology model, MOMP also has a large extracellular loop, possibly involved in ligand transport as observed in *E. coli's* FadL and TodX, the toluene transporter of *Pseudomonas putida* (Hearn *et al*., 2008), in addition to a number of smaller loop regions. From **Figure 1.9B**, a number of regions with undefined density have been identified, indicative of a number of flexible regions such as the external loops, previously modelled. Being low resolution data, it could be argued that MR with a homology model largely biases the resulting model found. However, MR on data merged with BLEND also was conducted with a number of known *E. coli* contaminants, such as OmpF, OmpA and OmpC, whereby the fit to the density and the resulting packing were deemed unsuitable (Atanu, 2014). Mass spectrometry was also used to analyse the crystal drop composition, which determined approximately 70% MOMP, with the remainder OmpF (Atanu, 2014). As OmpF does not possess an occluded barrel, and the MR electron density fit and cell packing was unsuitable,

it was highly likely the crystal structure observed arose from MOMP rather than any contaminant protein.

Whilst this data confirms the presence of the plug domain within the lumen of the barrel, in addition to the 14-stranded structure, as opposed to the 16-stranded model hypothesised by Feher *et al* in 2013, the structure was only resolved to 4 Å and, therefore, the side chain orientations were not completely characterised. Detailed information regarding the side chain orientations, as well as the external loop regions and internal plug will provide essential information regarding MOMP's possible function, as well as enhance the design of vaccines and antimicrobials targeted towards MOMP. Therefore, this research will focus on obtaining a higher resolution crystal structure to infer greater detail of MOMP's structure and function.
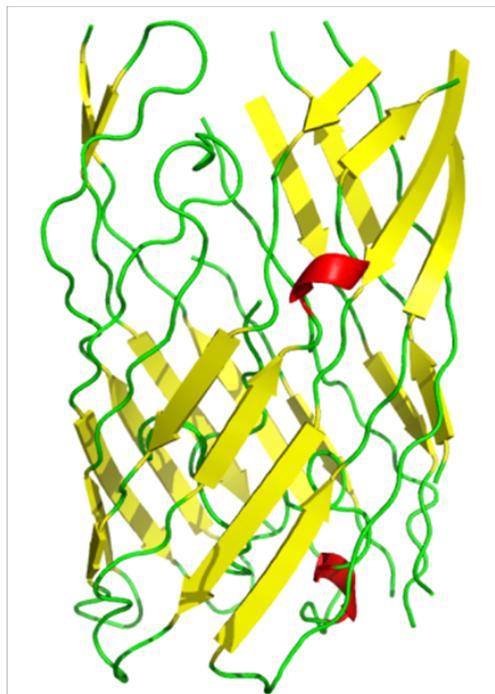
Figure 1.8. 4 Å crystal structure solved by MR using the FadL homology model. MOMP was isolated with sarkosyl before solubilisation in DDM and buffer exchange into OG for crystallisation. Data was collected in 10-15 ° wedges from crystals grown in 11.5 % PEG 4000, 0.1 sodium chloride, 0.1 lithium sulphate, 0.1 ADA, at pH 6.5 on the I04 beamline at DLS, and merged using BLEND. The structure suggests a 14-stranded β-barrel, as hypothesised.
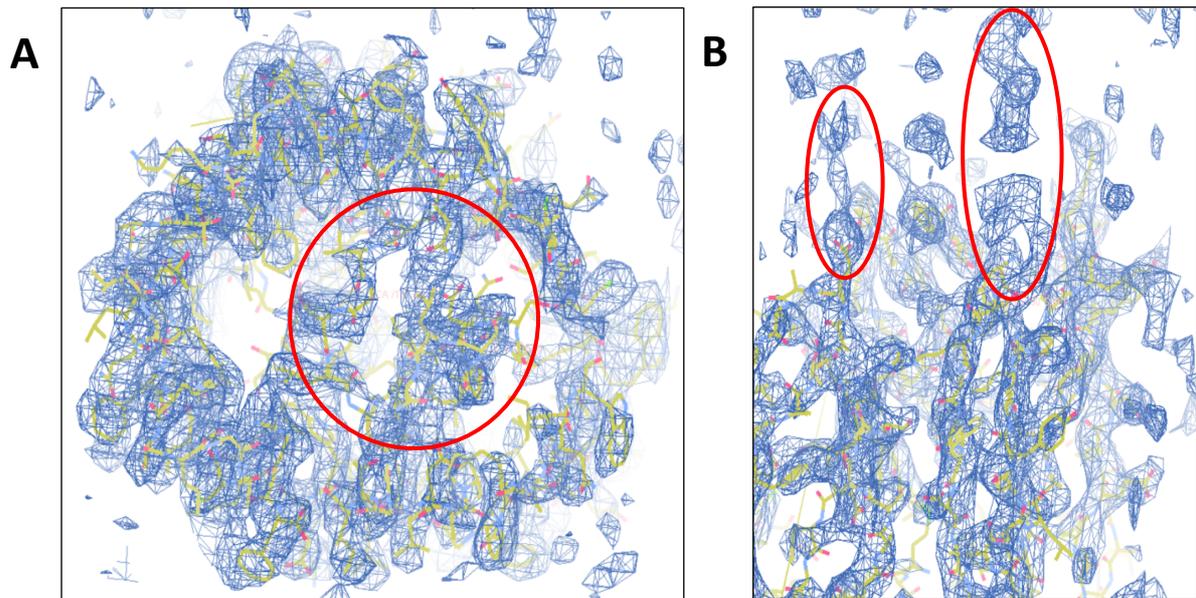
Figure 1.9. Electron density map (2Fo-Fc = 2σ) of the low resolution MOMP structure. **A, periplasmic view into the barrel.** The plug domain occluding the barrel lumen is indicated by the red circle. **B, side view of the undefined density at the extracellular side.** The two red ovals indicate the undefined density, likely due to flexible loop regions. (Atanu, 2014).

1.6.6. Functional studies of MOMP

The functional role of MOMP remains to be determined, but as previously mentioned, both the surface exposure and the abundance of MOMP within the *Chlamydial* membrane is suggestive of a possible role for MOMP in host-*Chlamydia* interactions, such as adhesion and consequently in stimulation of a host immune response (Caldwell and Schachter, 1982; Hafner *et al*., 2008; Redgrove and McLaughlin, 2014). Adhesion of EBs to a viable host cell is a vital step in the *Chlamydial* pathogenesis pathway and likely occurs via clathrin-mediated endocytosis (Hybiske and Stephens, 2007). Building a comprehensive understanding of the components involved in this process will pave the way to more effective disease prevention, treatment, and management. MOMP's potential role as an adhesin was identified by Su *et al* in 1990 through the study of two externally exposed variable domains, VDII and VDIV, which were suggested to contribute to cell electronegativity. The addition of MAbs, specific for these two domains, disrupted the electrostatic interaction between EBs and host cells,

whilst heat denaturation prevented the hydrophobic interaction between a conserved hydrophobic nonapeptide on VDIV and the host cell, indicating that MOMP might function as an adhesin (Su *et al.*, 1990). It was also possible that the MAbs blocked other interactions between MOMP and the host membrane, or that they sterically inhibited another important component in adhesion. A later study determined that *C. pneumoniae* utilises the mannose-6-phosphate/insulin-like growth factor 2 (M6P/IGF2) receptor on epithelial cells for attachment and, although not yet determined, a likely candidate for this interaction is MOMP due to its high-mannose oligosaccharide moiety, which is similar to the native ligand (Puolakkainen *et al.*, 2005). Whilst MOMP may be a prime candidate as an adhesin molecule, due to its abundance, since this research in 2005 the focus has shifted to other *Chlamydial* OMPs as the main adhesins, namely OmcB, the second most abundant OMP, which binds heparin sulphate-like glycosaminoglycans to form electrostatic interactions during the initial stage of EB invasion (Mölleken and Hegemann, 2008) followed by the irreversible interaction of polymorphic membrane protein 21 (Pmp21) with the host epidermal growth factor receptor (EGFR) on epithelial cells to mediate invasion (Mölleken *et al.*, 2013). Determining the mechanism by which *Chlamydia* attaches and infects host cells will prove invaluable in disease prevention strategies.

MOMP has long been considered a porin due to its β-barrel structure, which is an inherent feature for this function. Porins are a type of channel located in the outer membrane of Gram negative and some Gram positive bacteria that allow the passive diffusion of small solutes (<600 Da), including nutrients and waste products. Archetypal porins, such as OmpF of *E. coli*, form trimers and due to MOMP's suggested trimeric state it is feasible that MOMP may also possess porin activity. Although it had been shown by Bavoil *et al* in 1984 that

Chlamydia contained water filled pores, it wasn't until 1998 and the research of Wyllie *et al* that suggested this activity was due to MOMP, whereby addition of a neutralising MAb reduced the amplitude of the channel openings in planar lipid bilayers incorporated with nMOMP. A liposomal assay of rMOMP of *C. trachomatis* expressed in *E. coli* demonstrated the porin-like ability of MOMP to transport all 20 of the amino acids (Jones *et al*., 2000). Later, Sun *et al* examined the porin activity of nMOMP from *C. trachomatis* reconstituted into proteoliposomes, for which pore forming activity was indicated but at a rate fifty times lower than that of OmpF from *E. coli*, when testing the small sugar L-arabinose (Sun *et al*., 2007). PorB, which is also a contributor to the COMC, has been identified as a *Chlamydial* porin using a liposomal swelling assay which revealed the ability to transport small monosaccharides such as arabinose and glucose (Kubo and Stephens, 2002). Alternatively, it has been hypothesised that MOMP is an active transporter based on its homology with FadL, the fatty acid transporter of *E. coli*, as discussed previously in **Section 1.6.4** (Atanu *et al*., 2013). In this model, both the NPS and NPA motifs, which have been demonstrated to have roles in hydrophobic solute transport and water selective permeation in aquaporins (Guan *et al*., 2010), respectively, are suggested to have a role in solute coordination through the barrel, as shown with arrows in **Figure 1.7A** (Atanu *et al*., 2013). Whilst it was initially hypothesised that the NPA motif had a role in the FadL plug domain flexibility, knockout of this sequence did not affect the plug orientation in X-ray crystal structures (Hearn *et al*., 2007). This was supportive of the lateral diffusion model for FadL, whereby fatty acids enter the barrel and are then extruded through a hole in the barrel side into the membrane. However, as there is currently no indication of a lateral opening in MOMP, mutagenesis studies in conjunction with transport assays of these motifs will be useful for determining their importance in MOMP's functionality. Valuable knowledge would also be gained from

functional assays in order to determine which solutes could be transported by MOMP and, if essential to *Chlamydial* survival, could be blocked by targeting inhibitor molecules to MOMP's extracellular side. Whilst MOMP's main function may be as a major component of the COMC for maintenance of EB structural stability, which will be investigated and discussed further in this research, it is also possible that MOMP has a dual role, similar to PorB.

1.7. Aims and Objectives

Currently, there is very little literature surrounding structure-function characterisation of *Chlamydial* MOMPs due to difficulties in producing recombinant protein in a folded, active state. This research has been focussed on the production of rMOMP and the use of X-ray crystallography to solve MOMP's structure to a higher resolution, in an ongoing effort to discern the finer details of MOMP's β-barrel structure, including the loop and plug domains, as well as to identify the correct register of β-strands and hence accurately localise cysteine residues particularly with an external orientation that may contribute to intermolecular disulphide bonding. Knowledge of the orientation and position of the NPA and NPS motifs also could provide some insight into MOMP's possible function as a ligand transporter. Due to MOMP's homology with the fatty acid transporter of *E. coli*, FadL, a ligand transport assay using *E. coli* knockout mutants is developed, as part of this thesis work, to determine if MOMP can act as a fatty acid transporter. The role of the many cysteine residues is also investigated through mutagenesis studies, designed to both improve MOMP's solubility for crystallography efforts and also to observe the effect on MOMP clustering in the membrane, using high resolution fluorescence microscopy. These experiments will enhance

our understanding of MOMP's possible transport activity and mechanism, and its role in the structural integrity of *Chlamydia*.

Each of the subsequent research chapters will introduce specific literature, methods, and hypotheses in greater depth.

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 Materials

2.1.1 *E. coli* strains

Strain name	Genotype	Parent strain
BL21 (DE3)	F- *ompT* hsdS$_B$(r$_B$-m$_B$-) *dcm gal* λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*])	B834
BW25113	Δ(*araD-araB*)567 Δ*lacZ4787*(::*rrnB-3*) λ- *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	BD792
BW25113 Δ*fadL* Δ*fadR*	Δ(*araD-araB*)567 Δ*lacZ4787*(::*rrnB-3*) λ- *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514 Δ*fadL*::kan Δ*fadR*::chlor	BW25113
C41 (DE3)	F- *ompT* hsdS$_B$(r$_B$- m$_B$-) *dcm gal* λ(DE3 [*lacI lac-T7 gene 1 ind1 sam7 nin5*])	BL21 (DE3)
JW2341	F- Δ(*araD-araB*)567 Δ*lacZ4787*(::*rrnB-3*) λ- Δ*fadL*752::kan *rph*-1 Δ(*rhaD-rhaB*)568 *hsdR*514	BW25113
LS6164	*fadR- ΔfadL5*	K12
Stellar	F- *endA1 supE44 thi-1 recA1 relA1 gyrA96 phoA* Φ80d *lacZ ΔM15* Δ(*lacZYA-argF*) U169 Δ(*mrr-hsdRMS-mcrBC*) Δ*mcrA* λ–	HST08
Top10	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80lacZ *ΔM15* Δ*lacX74 recA1 araD139* Δ(*ara leu*) 7697 *galU galK rpsL* (*StrR*) *endA1 nupG*	MC1061

Table 2.1. Host *E. coli* strains used for recombinant protein production and expression.

2.1.2 Plasmids

Vector maps for each plasmid displayed in **Table 2.2** are detailed in **Appendix A2**.

Vector	Vector reference	Insert	Antibiotic resistance	Tag	Designated R number
pBADrha	Ford *et al.*, 2014	MOMP$_{H6}$	Cm^R	His-6	R1407
pBADa	This work	MOMP$_{H6}$	Cm^R and Ap^R	His-6	R1411
		FadL		N/A	R1414
pET101/D-TOPO	Invitrogen	MOMP$_{H6}$	Ap^R	His-6	R1185
		MOMP$_{H8}$		His-8	R1401
		MOMP$_{H10}$		His-10	R1402
		MOMP C136A		His-6	R1420
		MOMP C201A		His-6	R1409
		MOMP C203A		His-6	R1419
		MOMP C201A/C203A		His-6	R1410
		MOMP C136A/C201A		His-6	R1421
		MOMP C136A/C203A		His-6	R1423
		MOMP C226A		His-6	R1422
pOPIN-3C-Halo	Berrow *et al.*, 2009	MOMP	Ap^R	Halo and His-6	R1408
		MOMP C201A			R1416
		MOMP C203A			R1417
		MOMP C201/3A			R1418

Table 2.2. **Plasmids used for expression.** Cm^R, carbenicillin; Ap^R, ampicillin. An R number is assigned after the construct has been correctly sequenced. The original pET101/D-TOPO vector containing MOMP$_{H6}$ was kindly gifted by Dr E. Oveido-Orta, University of Surrey. The linearised pOPIN-3C-Halo vector was kindly gifted by the Oxford Protein Production Facility, Research Complex at Harwell.

2.1.3 Antibodies

Antibody	Type	Dilution	Application	Source
Anti-his HRP conjugated monoclonal mouse	Primary	1/10,000	Western blotting	Roche
6X his epitope tag monoclonal mouse IgG	Primary	1/1500	Western blotting	Pierce
Anti-mouse IgG (H+L) -HRP conjugated	Secondary	1/4000	Western blotting	Jackson ImmunoResearch
6X his-tag polyclonal rabbit	Primary	1/500	TIRF	AbCam
Anti-*E. coli* OmpA polyclonal rabbit	Primary	1/500	TIRF and STORM	Antibody Research Corporation
Alexa-488 anti-rabbit	Secondary	1/500	TIRF and STORM	LifeTechnologies
Alexa-647 anti-his	Primary	1/100	STORM	LifeTechnologies
Alexa-647 anti-rabbit	Secondary	1/200	STORM	Invitrogen

Table 2.3. **List of antibodies used and their application.**

2.1.4   Solutions and buffers

Solution name	Composition	Purpose
X4 SDS-PAGE sample buffer	250 mM Tris-HCl pH 6.8, 8% (w/v) SDS, 40% (v/v) glycerol, 0.024% (w/v) bromophenol blue, 4% (v/v) β-mercaptoethanol	SDS-PAGE
X1 SDS-PAGE running buffer	25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, pH 8.3	SDS-PAGE
Coomassie blue stain	0.25% (w/v) Coomassie Blue R250, 45% (v/v) methanol, 10% (v/v) glacial acetic acid	SDS-PAGE
Coomassie blue destain	30% (v/v) methanol, 10% (v/v) glacial acetic acid	SDS-PAGE
X1 NuPage Mes SDS running buffer	50 mM MES, 50 mM Tris Base, 0.1% (w/v) SDS, 1 mM EDTA, pH 7.3	SDS-PAGE
Tween-20 and tris buffered saline (TTBS) buffer	0.05 M Tris pH 7.5, 0.15 M NaCl, 0.1% (v/v) Tween-20	Western blotting
TAE agarose gel buffer	40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA	Agarose gel electrophoresis
Lysis buffer	20 mM Tris, 300 mM NaCl, 1 mM $MgCl_2$, 100 mg/ml lysozyme, 50 mg/ml DNase I, protease inhibitor, pH 8.0 at 4 °C	Large scale expression
Resuspension buffer	20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM EDTA, pH 8.0 at 4 °C	Large scale expression

FSEC solubilisation buffer	20 mM Tris, 300 mM NaCl, 10 mM DTT	Detergent screening using FSEC
FSEC flow buffer	20 mM Tris, 300 mM NaCl, 0.03% (w/v) DDM, pH 8.0	Detergent screening using FSEC
Solubilisation buffer	20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM DTT, X% (w/v) detergent	Purification
Buffer A	10 mM imidazole, 20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM DTT, 0.1% (w/v) DDM	His-trap affinity chromatography
Buffer B	500 mM imidazole, 20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM DTT, 0.1% (w/v) DDM	His-trap affinity chromatography
Resin equilibration buffer	20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP, 20 mM imidazole	Ni-gravity flow purification
Imidazole wash/elution buffer	20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP, X mM imidazole, X% (w/v) detergent	Ni-gravity flow purification and detergent exchange
Dialysis buffer	20 mM Tris, 300 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 1 mM TCEP, 3X CMC of relevant detergent	Dialysis
Detergent exchange buffer	20 mM Tris, 150 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 1 mM TCEP, 1.17% (w/v) OG	Detergent exchange with HiPrep column

Table 2.4. **List of solutions and buffers used, their composition and application.** Concentrations given as X are variable and specified in the relevant chapters.

2.1.5 Equipment

Equipment	Purpose	Location
Innova 44 shaker incubator	Bacterial culture	DLS
Amersham Biosciences Ultrospec 10 cell density meter	OD_{600} measurement	DLS
Helios α Thermospectronic spectrophotometer	OD_{600} measurement	UoR
Heraeus Megafuge 1.0R centrifuge	Small scale cell pelleting	UoR
Sorvall RC5B Centrifuge	Medium scale cell pelleting	UoR
Sorvall evolution RC centrifuge	Large scale cell pelleting	DLS
Stansted Fluid Power FPG12800 pressure cell homogeniser	Cell lysis	UoR
Constant Cell Disruptor Systems	Cell lysis	DLS
Beckman Coulter L-90k ultracentrifuge	Medium scale membrane preparation	UoR
Beckham Coulter-Optima L-100 XP ultracentrifuge	Large scale membrane preparation and purification	DLS
GE Peristaltic pump	Column loading	DLS
GE 5 ml HisTrap column	Purification	DLS
GE ÄKTA purifier	Purification	DLS
Spectromax microplate reader	BCA assay protein detection	DLS
Beckman Coulter-Optima Max-TL benchtop ultra-centrifuge	Detergent screening	DLS
Shimadzu HPLC	FSEC for detergent screening	DLS
GE Superose 6 10/300 L column	FSEC for detergent screening	DLS
GE Superdex 200 column	SEC for detergent exchange	DLS
GE HiPrep 26/10 desalting column	Detergent exchange	DLS
Beckman Coulter Allegra X-22R centrifuge	Protein concentration	DLS

TTP Labtech Mosquito robot	Crystallisation	DLS
Molecular Dimensions incubator	Crystallisation plate incubator	DLS
Leica DFC295 light microscope	Monitoring crystal growth	DLS
Invitrogen iBind	Western blotting	DLS
Invitrogen iBlot	Western blotting	DLS
Image Quant LAS 4000	Imaging SDS-PAGE gels	DLS
Syngene G.box v1.5.2.0	Imaging SDS-PAGE gels and spot plates	UoR
Eppendorf MasterCycler	PCR	UoR
NanoDrop spectrophotometer	Protein/DNA quanitification	UoR and DLS
SnapGene Viewer v4.1.1	Plasmid map creation	Freely available online
PyMOL v1.7.4	Visualisation of 3D protein structures	Freely available online
VMD v.1.9.2	Visualisation of MD simulations	University of Oxford
Gromacs v4.6	MD simulations	University of Oxford
Nikon Eclipse Ti inverted microscope	Epifluorescence and TIRF microscopy	UoR
NIS-Elements AR 4.5	Epifluorescence and TIRF microscopy image acquisition	UoR
ImageJ v1.51	Epifluorescence and TIRF image analysis	Freely available online
Zeiss Elyra microscope	STORM data collection	CLF, Harwell
Zeiss Black 2012	STORM data collection and analysis	CLF, Harwell
ClusDoc	Clustering analysis of STORM data	Freely available online

Table 2.5. **List of equipment, their purpose and location.**

All chemicals and reagents used were supplied by Sigma Aldrich unless otherwise stated. Method A refers to protocols followed at the University of Reading, whilst Method B refers to those followed at Diamond Light Source.

2.2 Methods

2.2.1 Molecular biology

2.2.1.1 Primer design

Forward and reverse primers used in construct design and mutagenesis are listed in **Table 2.6**. For the selection of primers, an optimum melting temperature (Tm) of between 58-65 °C, with a difference less than 4 °C between the forward and reverse primers, is the most important parameter to consider. The gene of interest (GOI) homologous sequence should also be between 18-25 bp in length, with a GC content ideally between 40-60%. Calculation of Tm and GC% were only in regard to the GOI homologous region. For In-Fusion, it is essential to have 15 bp of overlap with the desired construct, with the remaining oligonucleotide homologous to the GOI. The primers shown in **Table 2.6** were purchased from Eurofin MWG-Biotech and resuspended in nanopure water (nH$_2$O) to a final concentration of 200 pmol/µl. Working stocks were used at 10 pmol/µl. All primer properties were calculated using A Plasmid Editor (ApE) software V2.0.51.

Vector	Name	Direction	Primer	Size (bp)	Tm (°C)	GC (%)	Purpose
pET101	MOMP C136A	For	GGGATCGCTTTGATGTTTTC**GCT**ACTTTAGGAGCTTCTAATGG	43	79	44	Mutagenesis
		Rev	CCATTAGAAGCTCCTAAAGT**AGC**GAAAACATCAAAGCGATCCC				
pET101	MOMP C201A	For	CTCGTGGAGCCTTATGGGAA**GCC**GGTTGTGCAACTTTG	38	81	55	Mutagenesis
		Rev	CAAAGTTGCACAACC**GGC**TTCCCATAAGGCTCCACGAG				
pET101	MOMP C203A	For	GCCTTATGGGAATGCGGT**GCT**GCAACTTTGGGAG	34	79	56	Mutagenesis
		Rev	CTCCCAAAGTTGC**AGC**ACCGCATTCCCATAAGGC				
pET101	MOMP C201A/C203A	For	GCCTTATGGGAAGCCGGT**GCT**GCAACTTTGGGAG	34	80	59	Mutagenesis
		Rev	CTCCCAAAGTTGC**AGC**ACCGGCTTCCCATAAGGC				
pET101	MOMP C226A	For	GTTGAAGAACTTAATGTGATC**GCT**AACGTATCGCAATTCTCTGTAAAC	48	79	38	Mutagenesis
		Rev	GTTTACAGAGAATTGCGATACGTT**AGC**GATCACATTAAGTTCTTCAAC				
pET101	MOMP His 8	For	CTACGCGTACCGGT<u>CACCAC</u>CATCATCACCATCAC	35	86	57	Mutagenesis
		Rev	GTGATGGTGATGATG<u>GTGGTG</u>ACCGGTACGCGTAG				
pET101	MOMP His 10	For	CTACGCGTACCGGT<u>CATCACCACCAC</u>CATCATCACCATC	39	87	56	Mutagenesis
		Rev	GATGGTGATGATG<u>GTGGTGGTGATG</u>ACCGGTACGCGTAG				
pBADrha	MOMPH6	For	CAGCAGGATCACATA*TGAAAAAACTCTTAAAGTCGGCGTTAT*	27 + 15	58	33	Infusion
		Rev	GGGTACCATGGCATA*TCAATGGTGATGGTGATGATGAC*	15 + 23	57	43	
pBADrha	Ampicillin	For	CATTTTGCCTTCCTGTTTTTGCTC	24	57	42	Infusion
		Rev	TGTAGATAACTACGATACGGGAGGG	25	58	48	
pBADa	FadL	For	CAGCAGGATCACATA*TGAGCCAGAAAACCCTGTTTACA*	15 + 23	58	43	Infusion
		Rev	GGGTACCATGGCATA*TCAGAACGCGTAGTTAAAGTTAGTACC*	15 + 27	58	41	
pOPINE3CHalo	MOMP	For	AGGAGATATACCATG*AAAAAACTCTTAAAGTCGGCGTTATTATC*	15 + 29	56	31	Infusion
		Rev	CAGAACTTCCAGTTT*GAATCTGAACTGACCAGATACG*	15 + 22	54	45	
N/A	T7	For	TAATACGACTCACTATAGGG	20	48	40	Sequencing
		Rev	CTAGTTATTGCTCAGCGGT	19	53	47	
N/A	rhaB	For	CTGTCAGTAACGAGAAGGTCGC	22	58	55	Sequencing
N/A	pBAD (pTrcHis)	Rev	CTTCTGCGTTCTGATTTAATCTG	23	53	39	Sequencing
N/A	pOPIN3CHalo	Rev	TTACCGTGCAGGAACAGCACAG	22	61	55	Sequencing

Table 2.6. **Oligomers used for mutagenesis, In-Fusion, and sequencing.** Bold, mutated bases; italics, non-homogeneous regions for In-Fusion; underlined, inserted regions. All Tms and GC content calculated using ApE software, except for site directed mutagenesis which used the following formula to account for mismatched bases: Tm = 81.5 + 0.41(%GC) − (675/N) - % mismatch, as specified by the QuikChange XL II handbook.

2.2.1.2 QuikChange II XL site directed mutagenesis using polymerase chain reaction (PCR)

PCR was performed using Agilent Technologies QuikChange II XL site directed mutagenesis (SDM) kit. The reaction mixture was made up to 25 µl using both 50 ng and 100 ng DNA templates, X10 reaction buffer, forward and reverse primers, and QuikSolution. The reaction mixture was mixed, pulse centrifuged then supplemented with Pfu ultra DNA polymerase before being immediately placed in the PCR cycler (Eppendorf MasterCycler). The extension time was adjusted to 2 minutes per Kb. Cycling parameters are shown in **Table 2.7**. The resulting PCR products were incubated at 37 °C with *DpnI* endonuclease (0.5 µl per 15 µl PCR reaction) for 2h and analysed via agarose gel electrophoresis.

PCR products of the correct size were then transformed into Top10 or Stellar *E. coli* cells and plated onto LB plates containing 0.6% (w/v) glucose and the appropriate antibiotic. Plasmids from three to five transformants were prepared and sent for sequencing with Eurofins MWG to confirm the correct mutation. Overnight cultures were used to make glycerol stocks in 15% (v/v) glycerol, which were frozen at -80 °C.

Step	Temperature (°C)	Time	Cycles
Initial denature	95	2 min	1
Denature	95	1 min	
Anneal	60	1 min	18
Extension	68	2min/Kb	
Final extension	68	7 min	1

Table 2.7. **Modified PCR cycling parameters for QuikChange II XL mutagenesis.**

2.2.1.3 Amplification of insert

CloneAmp™ HiFi PCR premix was used to amplify the GOI for insertion into the vector of interest. Two concentrations of template DNA were used to obtain the optimum PCR product; 20 ng and 50 ng. The master mix preparation is shown in **Table 2.8**. An initial five cycle temperature gradient for annealing of the GOI was run before 25 cycles at a fixed temperature; full details of the cycling parameters are shown in **Table 2.9**. Samples were treated with *DpnI* for 2h at 37 °C before analysis on an agarose gel.

Reagent	Volume (µl)	
	20ng template	50ng template
CloneAmp HiFi PCR premix	12.5	12.5
Primer 1 (10pmol/µl)	0.75	0.75
Primer 2 (10pmol/µl)	0.75	0.75
Template (20ng/µl)	1	2.5
nH$_2$O	10	8.5

Table 2.8. **PCR master mix recipe for amplification of inserts.** Two different template DNA concentrations were used to obtain the optimum PCR product. Each 25 µl reaction mixture was done in triplicate to test three different annealing temperatures.

Step	Temperature (°C)	Time (seconds)	Cycle
Denaturing	98	10	
Annealing	-2, 0, +2 difference to primer Tm	10	5
Extension	72	30/Kb	
Denaturing	98	10	
Annealing	62	10	25
Extension	72	30/Kb	
Extension	72	600	1

Table 2.9. **PCR cycling conditions used for amplification of the GOI**. Initially, to improve annealing to the GOI homologous region, three annealing temperatures were chosen; at the primer Tm, and 2 °C above and below, which was cycled five times. The following 25 cycles were done at a higher annealing temperature of 62 °C. A final ten minute extension at 72 °C was performed to ensure complete filling of any protruding strands.

2.2.1.4 Gel and PCR purification

PCR products and linearised vectors were initially verified using 0.7% (w/v) agarose gels (**Section 2.2.1.7**). If present as a single band, samples were PCR purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). If multiple bands were present, the DNA was loaded onto an agarose gel and run at 60 v for 2h. The gel was imaged using a UV protective screen before excision of the band of interest under UV light. This was then purified using the same kit, following the gel purification procedure. The DNA concentration was measured using a NanoDrop spectrophotometer at 260 nm.

2.2.1.5 Inverse PCR

Inverse PCR was used to linearise DNA without the use of restriction digestion enzymes. Primers were designed to flank the desired insertion point, following the criteria set out in **Section 2.2.1.1**, and the PCR reaction was prepared following the CloneAmp HiFi Premix procedure. The PCR product was *DpnI* treated, checked on an agarose gel for the correct size, and purified.

2.2.1.6 In-Fusion Cloning Reaction

The In-Fusion HD cloning kit (Clonetech) was used for the In-Fusion cloning reaction. Linearised vector was created via inverse PCR. CloneAmp™ HiFi PCR premix was used to amplify the GOI. Both linearised vector and PCR produced inserts were purified prior to use. The In-Fusion cloning reactions were set up at room temperature using molar ratios calculated using Clonetech's online tool. These reaction mixtures were then incubated at 50 °C for 15 minutes before placing on ice until transformation. 2.5 µl of reaction mix was transformed with Stellar cells (Clonetech), which were plated onto LB plates containing 0.6%

(w/v) glucose and the appropriate antibiotic (pOPINE3CHalo, 50 µg/ml ampicillin; pBADrha, 34 µg/ml chloramphenicol; pBADa, 100 µg/ml ampicillin), and incubated overnight at 37 °C.

2.2.1.7 Agarose gel electrophoresis

DNA was analysed with 0.7% (w/v) agarose gels for both plasmid templates and PCR products. Agarose was resuspended in TAE buffer, dissolved with gentle heating, and supplemented with 0.2X GelRed. Hyperladder 1 Kb (Bioline) was used as a molecular weight marker. Template plasmid (2 µl) or PCR product (4 µl) was mixed with 1 µl BlueJuice loading dye (Invitrogen) and subsequently loaded onto the horizontal gel. The DNA was separated for 80 minutes at 80 v (Biorad powerpack 300) and visualised using Syngene G.box V1.5.2.0.

2.2.1.8 Nucleic acid quantification using a Nanodrop spectrophotometer

DNA concentration was estimated using a Nanodrop Spectrometer (ThermoScientific) by absorbance of UV-visible light at 260 nm.

2.2.1.9 Transformation

Competent cells, stored at -80 °C in microcentrifuge tubes, were thawed on ice. 1.5 µl of DNA (50-100 ng) was added to 25 µl of competent cells in fresh microcentrifuge tubes and mixed by gentle tapping. The resulting mixtures were incubated on ice for 30 minutes and then transferred to a 42 °C water bath for 45 seconds, then were immediately returned to ice for 5 minutes. 225 µl of LB media was added to the transformed cells, which were then incubated at 37 °C for 1h. 10-100 µl of each transformed cell culture was spread on LB agar plates supplemented with 0.6% (w/v) glucose and the appropriate antibiotic (100 µg/ml ampicillin, 34 µg/ml chloramphenicol, 50 µg/ml kanamycin) and incubated overnight at 37 °C.

2.2.1.10   Plasmid mini prep

10 ml of LB supplemented with the appropriate antibiotic and 0.6% (w/v) glucose was inoculated with a single colony and grown overnight at 30 °C, in a shaking incubator at 225 rpm. The plasmid miniprep was then prepared using the QiaPrep Spin Miniprep kit (Qiagen) and checked on an agarose gel. The DNA concentration was measured using a NanoDrop spectrophotometer at 260 nm.

2.2.1.11   Glycerol stocking

Glycerol stocks of positively transformed cells were generated from overnight cell cultures with a final concentration of 15% (v/v) glycerol. Stocks were mixed in sterile cryogenic tubes until an even suspension was achieved before being frozen at -80 °C.

2.2.1.12   Sequencing

Constructs were sequenced by Eurofins Genomics, using the sequencing primers specified in **Table 2.6**.

2.2.2   Protein expression and purification

2.2.2.1 Competent cell prep

Competent cell stocks were streaked onto plain LB plates and incubated overnight at 37 °C. A single colony was inoculated into 10 ml LB and incubated overnight at 30 °C in a shaking incubator at 225 rpm. The overnight culture was diluted 100-fold into fresh LB and incubated at 37 °C in a shaking incubator at 225 rpm, until an OD_{600} of 0.4-0.5 had been reached, as measured using a Helios α Thermospectronic spectrophotometer. Cells were immediately placed on ice and pelleted at 3,000 rpm, using a Heraeus Megafuge 1.0R centrifuge with swing out rotor, for 15 minutes at 4°C. The supernatant was removed and

the pellet resuspended in 5 ml of ice cold 100 mM $CaCl_2$. The cells were incubated on ice for 15 minutes before being centrifuged again under the same conditions. The pellet was then resuspended in 1 ml of 100 mM $CaCl_2$ and incubated on ice for 1.5-2h. Ice cold glycerol was added to a final concentration of 15% (v/v) before 100 µl aliquots were dispensed into sterile microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80 °C.

2.2.2.2 Growth curve

A single colony of transformed *E. coli* cells was inoculated into 10-20 mL of LB supplemented with the appropriate antibiotic and 0.6% (w/v) glucose and incubated overnight at 30 °C, shaking at 225 rpm. The overnight culture was diluted 100-fold into 20 ml of fresh LB supplemented with the appropriate antibiotic, in triplicate, and incubated in a shaking incubator at 37 °C until an OD_{600} of ~0.5. Cultures were induced with 1 mM (final concentration) of isopropylthiogalactoside (IPTG) and grown at 25 °C, shaking at 225 rpm for an additional 6h to 10h, with OD_{600} readings taken every hour following induction.

2.2.2.3 Medium scale protein expression (at Reading)

A single colony of transformed *E. coli* cells was inoculated into 10 mL of LB supplemented with 100 µg/ml ampicillin and 0.6% (w/v) glucose and incubated overnight at 30 °C, shaking at 225 rpm. The overnight culture was diluted 100-fold into 250-500 mL of LB in 2 L FisherScientific flasks supplemented with 100 µg/ml ampicillin and incubated at 37 °C, shaking at 225 rpm until an OD_{600} of ~0.5 was reached. Cultures were induced with 1 mM of IPTG and grown at 25°C, shaking at 225 rpm for between 3h and 16h. The OD_{600} was measured and cells pelleted for 20 minutes, 10,000 rpm (15,191 x g) at 4 °C, using a Sorvall RC5B centrifuge and F16/250 rotor. The pellet was resuspended in lysis buffer before passage through a pressure cell homogeniser (Stansted Fluid Power FPG12800) twice at

20,000 psi. The lysed cells were centrifuged for 20 minutes at 12,000 rpm (17,210 x g) 4°C, using a Sorvall RC5B centrifuge and SS34 rotor to sediment inclusion bodies. The supernatant was then pelleted at 15,000 rpm (26,890 x g) 4 °C for 45 minutes, using a Sorvall SS-34 rotor in a Beckman Coulter L-90k ultracentrifuge. The outer membranes were then resuspended in resuspension buffer, according to their OD_{600} measurement at the time of harvesting; 0.5 ml of buffer per OD unit per 100 ml of original culture. Samples were then analysed by SDS-PAGE and Western blotting, and stored at -20 °C.

2.2.2.4 Large scale expression (at Diamond Light Source)

A single colony of transformed *E. coli* cells was inoculated into a 500 ml Erlenmeyer flask containing LB media supplemented with 100 μg/ml ampicillin and 0.6% (w/v) glucose and incubated overnight at 30 °C in a shaking incubator at 225 rpm. The overnight culture was diluted 100-fold into 2 L baffled Erlenmeyer flasks, each containing 500 ml of LB media supplemented with 100 μg/ml ampicillin. Cultures were incubated in an Innova 44 incubator shaker at 37 °C until an OD_{600} of ~0.5 was reached. Cultures were induced with a final concentration of 1 mM IPTG and grown at 25 °C for an additional 16h to 20h.

Cells were harvested by centrifugation for 10 minutes at 5,500 rpm (6,600 x g) 4 °C using a Sorvall evolution RC centrifuge (ThermoScientific) with an F8-6x1000y Fiberlite Piramoon tech inc. rotor. The pellet was then resuspended in lysis buffer and the cells passaged twice through a cell disruptor (Constant Cell Disruptor Systems) at 25,000 psi. The lysed cells were centrifuged for 20 minutes at 12,000 rpm (16,740 x g), 4 °C using a Beckham Coulter-Optima L-100 XP ultracentrifuge with a Type 45 Ti fixed-angle rotor, to sediment inclusion bodies. The supernatant was transferred to new centrifuge tubes and ultracentrifuged for 2h at 41,000 rpm (195,426 x g) 4°C, using the same centrifuge and rotor. The pellet was

resuspended and homogenised with resuspension buffer in a 1:5 pellet to buffer ratio using a glass homogeniser. Membranes were then flash frozen using liquid nitrogen and stored at -80 °C.

2.2.2.5 Solubilisation

Membranes were removed from storage at -80 °C and thawed at room temperature. In order to remove contaminant soluble proteins, membranes were stirred slowly for 1h at 4 °C in 2% sarkosyl solubilisation buffer. The solubilised solution was ultracentrifuged for 45 minutes at 41,000 rpm (195,426 x g), 4 °C using a Beckham Coulter-Optima L-100 XP ultracentrifuge with a Type 45 Ti fixed-angle rotor. The supernatant, containing the sarkosyl soluble proteins, was discarded and the insoluble pellet containing the membrane portion was resuspended in solubilisation buffer using a glass homogeniser. The solution was then supplemented with 1% (w/v) SB3-14 and again stirred slowly for 2h at 4 °C. After ultracentrifugation under the same conditions, the supernatant containing the detergent soluble fraction was collected and used for nickel gravity flow purification.

2.2.2.6 Nickel gravity flow purification

Nickel Superflow resin (Qiagen) was pre-equilibrated using 10 CV of equilibration buffer at 4 °C. SB3-14 solubilised protein was supplemented with approximately 1 ml resin per 1 L of original culture and 20 mM imidazole, followed by gentle stirring at 4 °C for 1 h. This was applied to a glass Econo-Column (Bio-Rad) and washed with 10 CV of 30 mM imidazole wash buffer and 10 CV of 50 mM imidazole wash buffer. MOMP was then eluted with 2-5 CV of 200 mM imidazole elution buffer.

2.2.2.7 Desalting and buffer exchange into OG using a HiPrep column

Following elution from the nickel resin, up to 15 ml of protein was filtered using a 0.45 µm syringe filter before being manually loaded onto a HiPrep 26/10 desalting column (GE Healthcare Life Sciences) preequilibrated with 2 CV of detergent exchange buffer. A flow rate of 2-4 ml/min, depending on the pressure (max 0.5 MPa), was set for 1 CV (53 ml) with the fraction of interest collected in a 50 ml falcon tube. Consecutive samples were run in a similar matter after observation of the conductivity peak at approximately 30-50 ml.

2.2.3   Protein Analysis and Quantification

2.2.3.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Method A: 1 mm 12.5% SDS-PAGE gels were prepared following the Laemmli recipe (Laemmli, 1970). Samples were supplemented with X4 SDS-PAGE sample buffer, then heated to 100 °C for 10 minutes and briefly centrifuged before being loaded onto a gel. Gels were run with X1 SDS-PAGE running buffer at 200 v for 35 minutes (Bio-Rad Powerpac 300). Proteins were compared with PageRuler prestained protein ladder (ThermoScientific). Analysis of the gels was done using Coomassie blue and/or Western blotting. Gel images were captured using a Syngene G.box V1.5.2.0.

Method B: Samples were supplemented with X4 NuPage LDS sample buffer (Novex) and X10 NuPage sample reducing agent (Novex), then heated to 100 °C for 10 minutes and briefly centrifuged before being loaded onto a 1 mm 4-12% Bis-Tris precast gel (Novex). Gels were run with X1 NuPage Mes SDS running buffer at 200 V for 35 minutes (Invitrogen Life Technologies PowerEase 500). Proteins were compared using a SeeBlue Plus2 prestained

standard (Novex). Analysis of the gels was done using Coomassie blue and/or Western blotting. Gels were imaged using an Image Quant LAS 4000 (GE Healthcare Life Sciences).

2.2.3.2 Coomassie blue staining

SDS-PAGE gels were stained for 1h or overnight in Coomassie blue stain and then de-stained twice for 45 minutes each time. Gels were left in distilled H_2O until imaged. Alternatively, InstantBlue (C.B.S. Scientific) was used to fix and stain for 15 minutes.

2.2.3.3 Western blotting analysis

SDS-PAGE gels were prepared as described (**Section 2.2.3.1**). For each blot, the antibodies used are described in the corresponding figure legends. A full list of antibodies is provided in **Table 2.3**.

Method A: Two sponges, two pieces of filter paper and a nitrocellulose membrane were soaked in transfer buffer for 10 minutes to remove air bubbles. The stack for transfer was organised as follows; sponge, filter paper, SDS-PAGE gel, nitrocellulose membrane, filter paper, sponge. The tank was filled with 1 L of transfer buffer and run at 46 V (150 mA) for 1h. The nitrocellulose membrane was blocked overnight at 4 °C with 20 ml TTBS containing 5% (w/v) non-fat milk powder (Marvel). The blocked membrane was incubated for 30 minutes at room temperature and then washed with TTBS. The primary antibody was applied to the membrane for 1h with gentle rocking, before washing again with TTBS. If primary antibody was unconjugated, a secondary conjugated antibody step was included with incubation for 1h, followed by washing. The membrane was then exposed to EMSA stable peroxide solution and luminol enhancer solution at a 1:1 ratio for 5 minutes. Membranes were imaged using a Syngene G.box V1.5.2.0.

Method B: Proteins were transferred from SDS-PAGE gels onto an iBlot polyvinylidene difluoride (PVDF) membrane, using the Invitrogen iBlot 2 dry blotting system, and run at 25 V for 7 minutes. Immunodetection of the PVDF membrane was then done, using the automated iBind Western system, using appropriate primary and secondary antibodies. The membrane was incubated for a duration of 3h to overnight before being washed in distilled water. Proteins were detected with ECL reagents (Pierce) in a 1:1 ratio for 2 minutes and imaged using an Image Quant LAS 4000 (GE Healthcare Life Sciences).

2.2.3.4 The bicinchoninic acid (BCA) Assay

Total protein concentration was quantified using a Pierce BCA protein assay kit (ThermoScientific) following the microplate procedure. The absorbance was recorded at 562 nm using a Spectromax microplate reader. A standard curve of the bovine serum albumin (BSA) standards was plotted, and the protein concentration of the membrane sample derived from the linear equation.

2.2.3.5 Quantification of protein concentration using a Nanodrop

Purified protein concentration was estimated using a Nanodrop Spectrometer (ThermoScientific) by absorbance at 280 nm. The molar extinction coefficient ($M^{-1}cm^{-1}$), molecular weight (kDa) and Abs 0.1% (=1 g/l) of his-tagged MOMP (UniProtKB P27455.1) was computed to be 15.9 $M^{-1}cm^{-1}$ using the ExPasy ProtParam tool found at: http://web.expasy.org/protparam/.

2.2.4 Spot plate method for fatty acid functional assay

A single colony from each of the transformed strains were grown overnight in MM with 0.4% (w/v) glucose at 37 ˚C with shaking at 225 rpm. The overnight cultures were subcultured 1/10 into fresh MM containing 0.4% (v/v) glycerol and grown until an OD_{600} of 0.5. Cells were induced with 0.02% (w/v) rhamnose for 3h before cells equating to an OD_{600} value of 0.5 were harvested with centrifugation at 8,000 rpm for 3 minutes. The cells were then washed twice with MM, to remove the glycerol carbon source, before diluting in a 10-fold series down to -7 on ice. On a variety of MM plates (MM; MM + 0.4% w/v glucose; MM + 5mM sodium palmitate + 0.25% w/v Brij-58; MM + 0.25% w/v Brij-58, MM + 5mM sodium oleate + 0.25% w/v Brij-58) 5 μl of each dilution for each strain was spotted (one plate containing 10^1 to 10^{-3}, the other 10^{-4} to 10^{-7}) and incubated at 37 ˚C for 96h. A schematic of the layout can be seen in **Figure 2.1**. Images of the plates were taken after 96h using the manual capture function on the Syngene G:box.
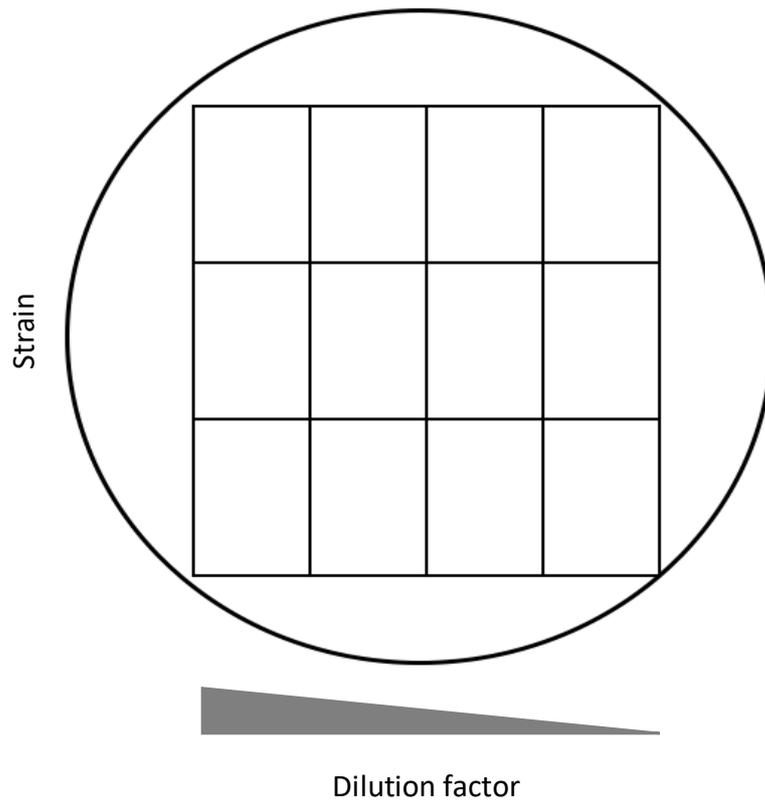
Figure 2.1. Schematic of the grid system used for the spot plate method. The y-axis lists strains, the x-axis the decreasing cell concentration.

2.2.5   Role of cysteine mutants in clustering of MOMP

2.2.5.1 Molecular dynamics simulations

In MD simulations, two 4 Å MOMP crystal structures were inserted into a box of size x=15 nm, y=15 nm, z=9 nm at a distance of 42 Å from each other. An externally exposed residue Trp55 was labelled red, in the absence of the alpha helical loop structure, for visualisation purposes. One model remained stationary, whilst the other was rotated 90 ° about the Z axis for four different orientations to allow different faces of the protein to interact. A lipid bilayer consisting of DOPG and DPPE, in equal proportions, was then inserted into the box using python script Insane (Wassenaar *et al*., 2015). The lipid bilayer was equilibrated for 100 ns and sodium ions and water added. Simulations were prepared and run using

Gromacs v4.6 (Berendsen *et* al., 1995) for a duration of 10 µs, equating to 5×10^8 steps. Parameters were obtained from the Martini force field (Wassenaar *et al.*, 2015). Results were visualised using VMD v.1.9.2 (Humphrey *et al.*, 1996).

2.2.5.2 Cysteine mutagenesis in pETMOMP$_{H6}$ and pOPINHaloMOMP

Cysteine residues were mutated to alanine in pETMOMP$_{H6}$, using the QuikChange XL II SDM kit with the primers detailed in **Table 2.6** and full protocol given in **Section 2.2.1.2**. As the pOPINHaloMOMP construct was not amenable to SDM, In-Fusion mutagenesis was utilised to insert the cysteine mutated sequences from the new pETMOMP$_{H6}$ constructs (the full protocol is provided in **Section 2.2.1.6**). All pETMOMP$_{H6}$ mutants were sequenced with the commercially available T7 forward and reverse sequencing primers, whilst the pOPINHaloMOMP mutants were sequenced using the T7 forward primer and, in the reverse direction, with a specially designed pOPIN3CHalo reverse primer (**Table 2.6**).

2.2.5.3 Bacterial cultures

C41(DE3) *E. coli* cells, transformed with mutated pOPINHaloMOMP or pETMOMP$_{H6}$ constructs, were inoculated into 10 ml of LB media containing 0.6% (w/v) glucose and 100 µg/ml ampicillin and grown overnight at 30 °C with shaking at 225 rpm. Overnight cultures were subcultured into 20 ml of fresh LB with antibiotics at a 1/100 dilution and grown at 37 °C. At an OD$_{600}$ of ~0.5, cells were induced with 1 mM IPTG and the temperature reduced to 25 °C. 1h and 3h post-induction cells were harvested and resuspended to an OD$_{600}$ of 0.5 in PBS.

2.2.5.4 Ibidi slide preparation and immunostaining

An Ibidi 8-well glass bottom μ-slide (170 μm +/- 5 μm) was incubated with 0.01% poly-L-lysine for 10 minutes, then washed three times with PBS. Bacterial cells, resuspended in PBS, were immobilised on the slides for 1h, followed by additional washing with PBS. Cells were fixed with 2% (v/v) formaldehyde-PBS for a further 10 minutes before being washed three times with PBS. The slide was stored in PBS at 4 ˚C until the day of imaging.

For immunostaining, storage PBS was aspirated off and the slide incubated in 50 mM ammonium chloride for 10 minutes in order to quench any residual fixative. After washing with PBS, cells were permeabilised with 0.1% (v/v) Triton X-100 for 5 minutes, followed again by washing. Cells were then incubated for 45 minutes with 100 μg/ml lysozyme and washed with PBS. The slide was blocked for 1h in 0.5-3% (w/v) BSA-PBS blocking buffer before incubation with primary antibody or conjugated antibody in 0.5-3% (w/v) BSA-PBS for 1h. Wells were washed and incubated for 1h with secondary antibody. A set of control cells, not expressing MOMP, were not immunostained to assess for bacterial auto-fluorescence. Wells were washed and stored in fresh PBS for imaging.

2.2.5.5 TIRF and epifluorescence imaging conditions

Samples plated on Ibidi slides were imaged by TIRF and epifluorescence, using a Nikon Eclipse Ti inverted microscope with a 100X oil-immersion objective and 1.50X NA TIRF objective. Samples were excited with both 488 (green) and 647 (red) lasers. Images were captured with an EMCCD camera (Andor) after a 20-200 ms exposure time and visualised using NIS-Elements AR 4.5 imaging software. Individual imaging conditions are stated in corresponding figure legends. Images were post processed with software packages ImageJ

(version 1.51) (Schneider *et al.*, 2012) and its associated distribution Fiji (version 1.0) (Schindelin *et al.*, 2012).

2.2.5.6 STORM imaging conditions

STORM experiments were carried out using the Octopus facility, at the Central Laser Facility (CLF), Harwell Campus, UK. Data were collected using a Zeiss Elyra microscope fitted with an Objective Plan-Apochromat 100x/1.46 oil DIC M27 objective lens using Immersol 518 F immersion oil (Zeiss). Fluorophores on both Ibidi and Superfrost microscope slides were excited using the 642 nm laser, raised to 4-8% to achieve a blinking state, before being lowered to 2% for data collection and detected with a LP 561/642 filter. 20,000 images were recorded with an exposure time of 20 ms, a camera gain of 100, and drift correction every 500 frames, using an EMCCD camera (Andor).

2.2.5.7 Cluster analysis from STORM images

Collected data were processed in Zen Black 2012 software (Zeiss), using the PALM module where the peak intensity to noise ratio was set to reject dimly emitting fluorophores and overlap of molecules was accounted for, through a multi-object fitting algorithm. Model based drift correction was applied, using automatic segmentation and trails were grouped for Alexa-647 as follows; max on time 5 frames, max off gap 10 frames, and capture radius of 2 pixels, based on previously calibrated data (Willson *et al.*, 2016; Pageon *et al.*, 2016). Localisation precision was selected based on the precision histogram for each data set, typically in the range of 1-25 nm. The ASCII text files detailing spatial localisation information for each molecule were saved for each data set and used in clustering analysis with ClusDoC software (Pageon *et al.*, 2016). Using the ClusDoC GUI, regions of interested (ROI) were manually defined by the user. The density-based spatial clustering of applications

with noise (DBSCAN) algorithm was implemented on the ROIs using the following parameters; epsilon 20 nm, min points 3, plot cut off 10, threads 2, L (r) – r 50 nm, smooth radius 14 nm.

Chapter 3

Expression, purification and crystallisation of recombinant MOMP in *E. coli*

3   Expression, purification and crystallisation of recombinant MOMP in *E. coli*

3.1   Introduction

MOMP has often been the target of *Chlamydial* vaccines and antimicrobials, not only due to its surface exposed regions and high abundance within the OM of circulating EBs, but also because of its observed immunodominance (Wolf *et al*., 2001; Molina *et al*., 2010). Of the *Chlamydia* species, namely human pathogens *C. trachomatis*, *C. psittaci*, and *C. pneumoniae*, MOMP is highly conserved with over 63% sequence identity (see **Appendix A1 Figure A1.1**). However, due to the difficulties in the expression and purification of active and correctly folded recombinant protein, there are currently no 3D structures for any MOMP proteins in the PDB. Recombinant expression has largely been favoured over native protein production in purification attempts due to the difficulty of both cultivating *Chlamydia* within host HeLa cells and also in extracting the EBs, where the majority of MOMP expression occurs. Much of the difficulty surrounding recombinant expression and purification has arisen due to both the high cysteine content (on average 9 cysteine residues over a total of approximately 395 aa) and the requirement to extract MOMP from its native membrane environment with detergent. As documented, MOMP forms intermolecular disulphide bonds within the COMC in order to provide structural stability (Liu *et al.,* 2010); these disulphide bonds are likely problematic in obtaining stable monomeric MOMP for further structural analysis with techniques such as macromolecular crystallography (MX), cryo-electron microscopy (cryo-EM), and NMR spectroscopy, discussed below.

3.1.1   Methods of structure determination

Cryo-EM, first developed in the 1980s, extends the capabilities of standard EM through the rapid vitrification of samples in their native conditions, be it single proteins, large

complexes, or whole cells, preventing the crystallisation of water and permitting more detailed data collection through improved signal-to-noise (Murata and Wolf, 2018). In standard EM, a sample is plated onto a small grid before illumination with an electron beam, whereby electrons are either able to pass through to the detector or become scattered, creating a shadow image. Whilst cryo-EM has certainly advanced in recent decades, due to the introduction of a more powerful detector system known as a direct electron detector (DED), the improvement in sample preparation for example graphene coated carbon grids to control protein adsorption, and the developments in image processing, it still has its limitations (Murata and Wolf, 2018). The most significant limitation with regard to structure solution of MOMP is the limit on protein size, which currently stands at around 100 kDa, with the resolution limits occasionally surpassing 3 Å. However, recently haemoglobin, a protein of 64 kDa, was solved to 3.2 Å using cryo-EM, although a high resolution structure previously resolved with single crystal XRD to 2.5 Å was used in model fitting (Khoshouei *et al*., 2017). Therefore, solely due to MOMP's small size (approximately 40 kDa), cryo-EM is a not a suitable technique at this stage, although will potentially prove useful in future studies to assess the clustering behaviour arising from intermolecular disulphide bonding (discussed in **Chapter 5**).

Conversely, solution NMR is typically limited to small proteins below 35 kDa, due to the difficulties in characterising the high number of atoms present in larger structures, which again limits the applicability for MOMP. Significantly, 1 mM of protein is typically required for solution NMR (Berg *et al*., 2002), which for MOMP equates to an extremely high concentration of 40 mg/ml. Given the difficulties associated with expression and purification of membrane proteins, solution NMR is very likely an unfeasible structural solution method

for MOMP solely due to the quantity required. However, recently solid-state NMR has been used to successfully study membrane proteins within a lipid bilayer or cell membrane without any restrictions on size, with the 81 kDa Anabaena sensory rhodophsin (ASR) protein recently solved using this method (Wang *et al*., 2013). NMR works on the basis that certain atomic nuclei possess a spin, which enables them to behave as a magnet. Upon application of an externally applied magnetic field the lower energy spin state is excited to a higher energy state, referred to as resonance, the spectrum of which can be recorded (Berg *et al*., 2002). As the naturally occurring ^{12}C isotope is not NMR active due to the equal number of protons and neutrons and ^{14}N produces a broad signal, this technique also requires isotopic labelling with ^{13}C and ^{15}N to obtain higher resolution data, typically achieved by supplementing *E. coli* growth media with isotopically labelled nutrients such as ^{15}NH$_4$Cl and ^{13}C-glucose (Cai *et al*., 1998). Despite successful reports of membrane protein structure solution with solid-state NMR, a number of limitations and challenges remain. Notably, reconstitution into either a lipid bilayer or detergent micelles results in a heterogeneous environment, which in turn contributes extra signals and reduces sensitivity, as well as taking up valuable space resulting in a lower intensity protein signal (Ladizhansky, 2017). Additionally, as the name suggests, samples are required to be in a solid state for example as crystals, such as those used in the recent structure determination of the membrane protein YadA, whereby researchers utilised poorly diffracting microcrystals originally intended for X-ray diffraction (Shahid *et al*., 2012). Alternatively, proteins reconstituted into lipid bilayers can be spread across glass plates, as was the case in the structure solution of the membrane protein phospholamban involved in muscle excitation (Verardi *et al*., 2011).

In MX, concentrated protein samples are encouraged to crystallise using a variety of multicomponent screens, the 3D structures of which are then resolved using XRD. Data are typically collected using a high powered synchrotron source, such as Diamond Light Source (DLS) in the UK or the European Synchrotron Radiation Facility (ESRF) in France, whereby electrons emitted from a cathode gun are propagated to extremely high speeds using a series of particle accelerators before the GeV beam is then directed around a storage 'ring' by a series of magnets (Materlik *et al.*, 2015). High quality data collection is also heavily reliant on the quality of the crystals, as their geometry and well-ordered nature creates a regular arrangement of atoms that are able to diffract incoming X-rays, producing specific diffraction patterns. Measurement of the intensity and angle of these diffracted spots using a range of tools relays information on the atom positioning, with more highly ordered crystals producing higher resolution data. A larger unit cell results in a greater quantity of reflections (Wlodawer *et al*., 2013). Resolutions of 2 Å are frequently documented in protein crystallography, whereby the backbone and most of the sidechains are clearly defined within the resulting electron density, although there are also over 600 structures within the PDB at resolutions better than 1 Å. However, MX is not without its limitations. Not unsurprisingly, in order for membrane proteins to crystallise they must be stable in solution and, therefore, must be transferred from their native bilayer environment to a detergent micelle. For crystals to form, the solution must be supersaturated for the protein to overcome its limit of solubility and form a critical nucleus. Typically, non-covalent crystal contacts, such as van der Waals (vdW) forces, hydrophobic interactions, ionic and hydrogen bonds, then begin to form, however, the hydrophobic regions on membrane proteins are mostly shielded by the detergent micelle, resulting in weaker protein-protein interactions and as such more fragile crystals (Moraes *et al.*, 2014). Most importantly, the protein-

detergent complex must be monodisperse, as incorrectly formed crystal contacts from contaminants will prevent further extension of the lattice and thus terminate growth. Typically, vapour diffusion is the most common method of crystallisation, whereby purified protein is mixed in a droplet of precipitant containing buffer and allowed to equilibrate in close proximity to a larger buffer reservoir, either in a well (sitting drop) or suspended from a coverslide (hanging drop) (Dessau and Modis, 2011), as shown in **Figure 3.1**, although lipidic cubic phase, bicelles, and microbatch are also procedures useful in membrane protein crystallisation (Moraes *et al.*, 2014).
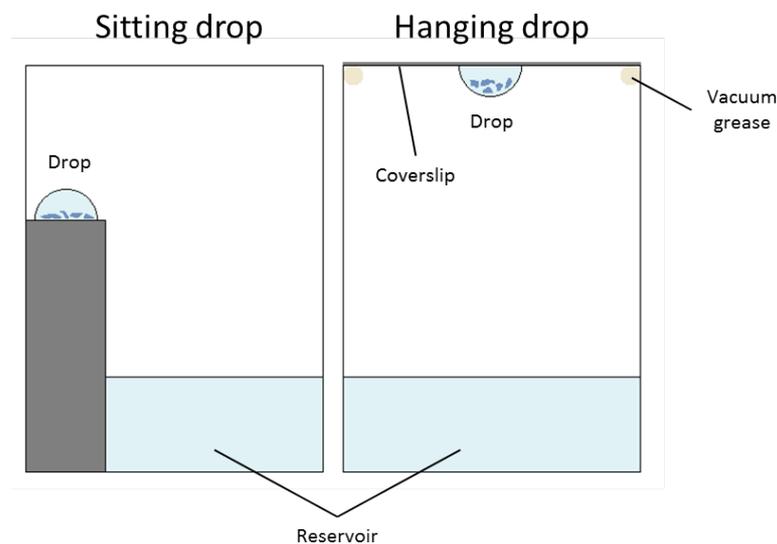
Figure 3.1. Diagram showing sitting drop and hanging drop vapour diffusion set up. In sitting drop vapour diffusion, the drop containing protein and precipitant is elevated above the reservoir and secured with sealing tape. In hanging drop vapour diffusion, the drop is suspended upside down on a coverslip and sealed with high vacuum grease.

As a result of the limitations associated with both NMR and cryo-EM, particularly for the study of 40 kDa MOMP, MX is the biophysical technique of choice in this research that is most likely to produce high-quality structural details for both side chain and β-sheet

orientation for this protein. Despite the huge number of variables to consider, such as detergent, crystallisation screening (pH, buffer, additives, precipitant, temperature), phasing and data collection strategy, MX is applicable to any protein able to form crystals. Additionally, with advancements of nanolitre crystallisation robotics, lower sample volume is required, which is a particularly important consideration when studying membrane proteins that are difficult to express and purify.

3.1.2 Optimisation of the purification and crystallisation of rMOMP

Previous expression screening of rMOMP in a variety of *E. coli* strains, media, and temperatures was conducted in order to determine the optimum conditions for MOMP expression within the membrane (Atanu, 2014). A crucial factor was MOMP's export to the OM, as incorrectly folded protein can become trapped as insoluble inclusion bodies within the cell cytosol, which requires refolding, a complex and rarely successful process. It was determined that the pET101-D/TOPO construct containing MOMP expressed in C41 (DE3) cells, under the control of the bacteriophage T7 RNA polymerase system, was optimal for rMOMP OM expression. This was likely due to the fact that C41 (DE3) is a strain derived from BL21 (DE3) containing an uncharacterised mutation that inhibits cell death during toxic protein expression (Miroux and Walker, 1996), which in combination with the pET101 vector designed for high yields resulted in successful expression.

As membrane proteins by nature are highly hydrophobic, their purification requires amphiphilic detergents in order to extract them from their native lipid environment and maintain them in a soluble and stable state (Moraes *et al*., 2014). This is achieved through aggregation of the detergent monomers into a micelle around each protein, be it monomer, dimer, trimer etc. The concentration at which detergents form this micelle is known as the

critical micellar concentration (CMC), and therefore detergents are typically used at a concentration much higher than this in order to ensure micelle formation during solubilisation. Detergents typically fall into three main categories; ionic, non-ionic, and zwitterionic, based on their hydrophilic head groups. Detergents such as SDS and sarkosyl are classed as ionic and are typically harsh due to their disruption of protein-protein interactions, which often results in protein denaturation. More gentle detergents, which disrupt protein-lipid interactions instead, include glucosides and maltosides which are classed as non-ionic due to their uncharged head groups. Finally, detergents with both a positive and negative charge but which are overall neutral are referred to as zwitterionic and include LDAO, FC-12, and CHAPS. As these detergents behave in a similar manner to ionic detergents, they are also capable of disrupting protein-protein interactions, although tend to be less harsh. **Table 3.1** details the detergents used in this research, along with their CMC and hydrophilic group class.

Name	Abbreviation	Type	CMC
CHAPS	N/A	Zwitterionic	8 mM (0.49%)
CYMAL-5	N/A	Non-ionic	2.4-5 mM (0.12%)
Elugent*	N/A	Non-ionic	N/A
Fos-choline-12	FC-12	Zwitterionic	1.5 mM (0.047%)
Lauryl Maltose Neopentyl Glycol	LMNG	Non-ionic	0.01 mM (0.001%)
Lauryldimethylamine-N-Oxide	LDAO	Zwitterionic	1-2 mM (0.023%)
n-Dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate	SB3-12	Zwitterionic	2.8 mM (0.094%)
n-Dodecyl-β-D-Maltopyranoside	DDM	Non-ionic	0.17 mM (0.0087%)
n-Lauroylsarcosine sodium salt	Sarkosyl	Ionic	14.4 mM (0.42%)
n-Octyl-β-D-glucopyranoside	OG	Non-ionic	18-20 mM (0.53%)
n-Tetradecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate	SB3-14	Zwitterionic	0.16 mM (0.007%)
n-Undecyl-β-D-Maltopyranoside	UDM	Non-ionic	0.59 mM (0.029%)
Octaethylene Glycol Monododecyl Ether	C12E8	Non-ionic	0.09 mM (0.0048%)
Octyl Glucose Neopentyl Glycol	OGNG	Non-ionic	1 mM (0.058%)
Triton X-100	N/A	Non-ionic	0.23 mM (0.01-0.016%)

Table 3.1. **Detergents used in MOMP solubilisation, for both Western blotting and fluorescence-detection size exclusion chromatography (FSEC) screening, and their properties.** *Mixture of alkyl glucosides (e.g. Triton X-100, OG). Data sourced from Anatrace in 2016.

Previously, MOMP was isolated with sarkosyl (Atanu, 2014), a detergent which is not capable of solubilising outer membranes (Filip *et al*., 1973) and thus removes a multitude of contaminant soluble proteins that may either co-associate with the protein of interest (POI) or become more difficult to remove in subsequent steps, leaving the membranes intact. Therefore, an additional detergent is required to extract MOMP from the membrane for solubilisation. As many detergents capable of extracting proteins from their native environment have not been successful in crystallisation attempts, it is also anticipated that a third detergent will be required for MOMP crystallisation. One of the most vital steps in membrane protein purification is solubilisation, as this greatly impacts the protein stability as well the final yield, which is especially important for membrane proteins due to their associated toxicity and the limited volume of the cell membrane. Previously, both recombinant and native MOMP have been solubilised in a range of detergents including DDM, OG, LDAO, SB3-14, and CHAPS, typically at a concentration of 1% (Atanu, 2014; Findlay *et al*., 2005; Pal *et al*., 2005), although success has also been reported with a cell-free system using nanolipid particles (NLPs) (He *et al*., 2017). Owing to the huge range of detergents available, with many different properties, it is essential to screen a representative sample in a high throughput manner. One method of doing so involves fluorescent labelling of the POI, however, MOMP previously tagged with green fluorescent protein (GFP) was not correctly expressed in *E. coli*, and instead formed inclusion bodies (Atanu, 2014). It was concluded that the large size of the GFP tag (27 kDa) likely affected protein folding and export to the membrane, especially as MOMP itself had already demonstrated toxic behaviour. However, recently a 3 kDa tag referred to as P3NTA has been developed (Backmark *et al*., 2013). This P3NTA tag is a peptide based fluorescent probe consisting of a fluorescein label attached to a Strep peptide containing three terminal

cysteine residues linked to three nickel charged nitrilotriacetic acid (NTA) groups, as shown

in **Figure 3.2** (Backmark *et al.*, 2013).
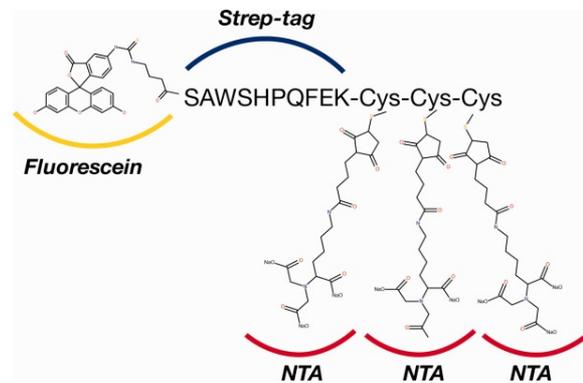
Figure 3.2. Diagram showing the chemical structure of the fluorescent P3NTA tag. (Backmark *et al.*, 2013)

P3NTA binds to his-tagged proteins through the same principle applied during immobilised

metal affinity chromatography (IMAC), which allows any his-tagged protein to be monitored

by fluorescence-detection size exclusion chromatography (FSEC) without the need to co-

express a fluorescent label, thus preserving the POI's structure and activity. Size exclusion

chromatography (SEC) is a technique routinely used in protein purification to determine the

heterogeneity of a sample. After equilibration, the protein sample is applied to the column,

a porous matrix typically made up of cross-linked agarose and/or dextran, which can then

be eluted based on size. From a SEC trace, a single symmetric Gaussian peak represents

stable, folded, monodisperse protein, in contrast to numerous asymmetric peaks which are

typically due to unfolded, unstable, and polydisperse protein (Kawate and Gouaux, 2006).

The location of the peak can also assist in determining the protein's MW, which is based on

diffusion of the protein through the pores of the stationary phase. Larger proteins are

prohibited from most of the pores and are thus eluted first, as opposed to the smaller

proteins which diffuse into and are retained within the pores, eluting later. One of the main benefits of FSEC is the ability to identify small amounts of protein in crude preparations as opposed to purified protein, such as solubilised membranes, to a high sensitivity and specificity (Backmark *et al.*, 2013). This permits screening during the earlier stages of protein production, such as during expression or solubilisation with detergents, to assist in small scale optimisation. As a result, this technique along with the small P3NTA probe will be used together with standard Western blotting in order to determine which detergent is the most effective at solubilising MOMP.

Following successful solubilisation, it is then necessary to isolate the POI from any other contaminant proteins that may also co-crystallise or indeed hinder crystallisation of the POI. This can be done with a number of tags through affinity chromatography, and in this case MOMP is hexahistidine tagged for purification using IMAC, which is a very well documented and relatively simple technique (Block *et al.*, 2009). In IMAC, a transition metal ion is chelated to an insoluble matrix, such as agarose, via an NTA group which binds to the metal with four valencies. The imidazole rings within the his-tag contain electron donor groups which are able to form two more coordination bonds with the immobilised metal, creating a stable and reversible interaction, as shown in **Figure 3.3** (Bornhorst and Falke, 2000). As a result, only histidine containing proteins are immobilised on the resin, with the string of histidines within the tag forming a more specific and stronger interaction due to the two coordination bonds. Imidazole can be used to competitively bind the metal ion, as the imidazole ring will also form the same interaction as histidine, which at low concentrations can prevent non-specific contaminant binding during protein loading and at higher concentrations can be used to elute the POI. A number of different divalent cations can be

used in IMAC, such as Ni^{2+}, Cu^{2+} and Zn^{2+}, each with different binding properties, with nickel

columns most routinely used to obtain a high yield, and cobalt to obtain a higher purity. As a

sarkosyl step is included to remove contaminant proteins, his-tagged MOMP will be purified

by Ni-NTA affinity chromatography in order to optimise the yield.
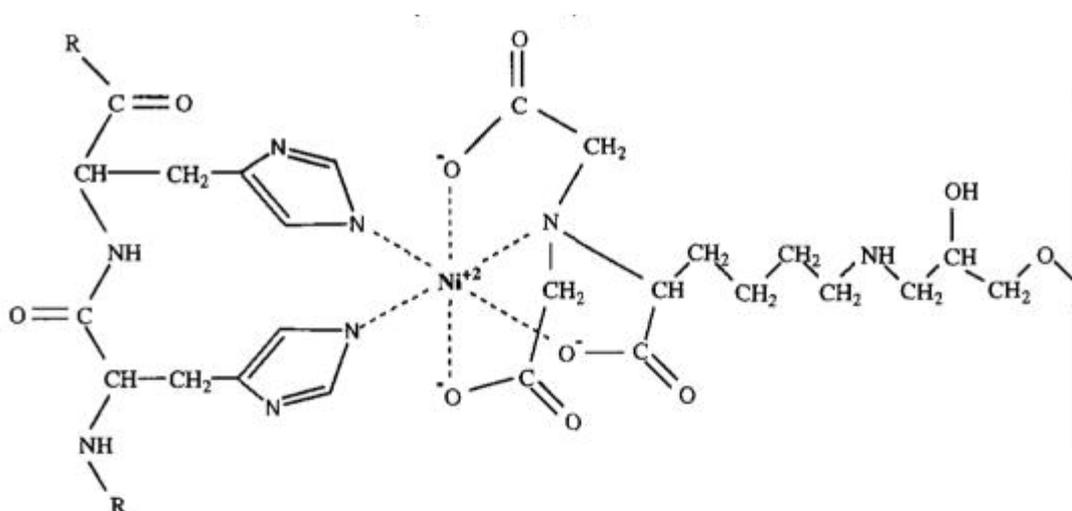
Figure 3.3. Interaction between the his-tag and Ni^{2+} cation in IMAC. From left to right; two histidines, Ni^{2+}, NTA group. Dotted lines represent the coordination bonds. (Bornhorst and Falke, 2000)

Following purification, as previously stated, it is imperative that the protein be

monodisperse for successful crystallisation. In order to do so, SEC is utilised, which follows

the same principles as described for FSEC, whereby a singular narrow peak represents stable

and monodispersed protein. The fractions corresponding to the peak of interest, also

identified through the protein's MW, can be collected, pooled, and purified. However, an

additional benefit of SEC is that the buffer in which the sample exists can be exchanged

during elution. This feature is ideal for crystallographic studies of membrane proteins, as the

detergent used for solubilisation can be exchanged for a number of detergents that are

more amenable to crystallisation. This method can also remove excess detergent micelles,

as well as alter other buffer components that may affect crystallisation such as the salt content, pH etc. Previously, 10 mg/ml of purified MOMP was sufficient for crystallisation (Atanu, 2014) and, therefore, this concentration will be used as a target for future purifications.

As previously mentioned, successful crystallisation requires monodisperse and solubilised protein to be supersaturated within a solution in order for crystal contacts to form. An additional unknown variable within membrane protein crystallisation is the use of a detergent to maintain the protein in a stable and soluble state, where it is likely that crystal contacts are inhibited by the detergent, particularly those with long acyl chains that engulf more of the protein (Sonoda *et al*., 2010). Analysis of membrane protein structures solved with XRD, as deposited in the PDB, revealed the crystallisation conditions most common in producing strongly diffracting crystals. These data were collated and used in the design of two crystallisation screens specifically created for membrane proteins, known as MemGold I and II (Parker and Newstead, 2016), which were successful in producing the rMOMP crystals obtained previously (Atanu, 2014). The aim of screening is to cover the largest crystallisation space (the conditions used to produce crystals) possible, whilst conserving the amount of protein required. One method of comparing the chemical similarity of different screens is through the freely available C6 Web Tool, an online tool which compares two screens providing a metric between zero and one, with 0 indicating two identical screens and 1 relating to those that are completely different (Newman *et al*., 2010). Comparison of the two MemGold screens reveals a score of 0.54, with a score above 0.4 suitable for screening, indicating that these two screens are an appropriate starting point for MOMP crystallisation optimisation. One of the most successful detergents in membrane protein crystallisation is

DDM, accounting for almost 50% of structures in the PDB although typically exhibiting a lower resolution (Parker and Newstead, 2016). However, success has also been demonstrated with some shorter chain detergents, such as OG and LDAO, crystals of which tend to diffract to higher resolutions, with OG accounting for the highest mean resolution of structures currently available (Parker and Newstead, 2016). Therefore, both due to the previous success with MOMP solubilisation and crystallisation in DDM, OG, and LDAO, these detergents will be the focus of MOMP crystallisation in the first instance. The low resolution crystal structure for rMOMP (Atanu, 2014) is described in **Chapter 1 Section 1.6.5**.

3.2   Aims and Objectives

Membrane proteins play vital roles in many cellular processes including but not limited to drug influx/efflux, signalling, cell recognition and adhesion. Despite this, only 4173 of the ~120,000 protein structures solved with XRD in the PDB are membrane proteins, with only a quarter of these at a resolution better than 2 Å. This further highlights the difficulties associated with membrane protein expression, purification, and crystallisation, particularly as it has been estimated that membrane proteins account for approximately 20-30% of all proteins (Krogh *et al.*, 2001). Due to the many roles of membrane proteins, detailed information regarding their structures and functions will prove invaluable in developing our understanding of not only normal cellular physiology but also disease mechanisms in both humans and bacteria.

Previously, unoptimized MOMP crystals grown for over three months only diffracted to 4 Å. Whilst the overall size and shape of the barrel, and the presence of the internal plug region were determined, a higher resolution data set is required to confidently identify the side chain orientations, the exact composition of the β-strands, and the precise locations of the

extracellular loops. As a result, the aim of this research is to optimise the purification protocol for the production of *C. pneumoniae* rMOMP in *E. coli* and for subsequent crystallisation optimisation. A number of different detergents will be tested during solubilisation in order to maximise the total protein yield and decrease the volume of growth culture required. Purified protein will be used in sitting drop vapour diffusion crystallisation trials, in the first instance with commercial screens MemGold I and II due to their previous success in MOMP crystallisation. Any crystallisation hits from the screening will be optimised in an effort to obtain higher diffraction quality crystals.

Continuing advancements in cryo-EM may soon permit data collection of smaller proteins, such as MOMP, which will prove particularly useful in membrane protein analysis due to the ability to image within whole cells, eliminating the complex purification and crystallisation procedure required for MX studies. However, at present XRD remains an excellent method for structural solution, producing high resolution data which can detail atomic positioning and reveal insights into not only MOMP's structure, but also how this may relate to its as yet unknown function.

3.3 Methods

For final optimised methods, refer to **Chapter 2**.

3.3.1 Difference between cultures grown at the University of Reading and Diamond Light Source

A single colony of C41 (DE3) *E. coli* cells, transformed with pETMOMP$_{H6}$, was inoculated into either 10 ml (50 ml falcon tube) or 150 ml (500 ml flask) LB, with 100 µg/ml ampicillin and 0.6% (w/v) glucose. Both of these cultures were growth overnight at 37 °C in a shaking incubator at 225 rpm, with an additional 150 ml culture grown at a lower temperature of 30 °C. Following overnight growth, OD_{600} measurements were taken before each overnight culture was then used to subculture 10 ml fresh LB at a 1/100 dilution. This was done in triplicate for each overnight condition, along with supplementation with 0%, 0.5%, or 1% (w/v) glucose. Growth at 37 °C was monitored with OD_{600} measurements every hour until passing a value of 0.5. This comparative study was conducted at Diamond Light Source.

3.3.2 Solubilisation with DDM

Prior to optimisation, membranes, as prepared in **Chapter 2 Section 2.2.2.4**, were solubilised following the protocol described in **Chapter 2 Section 2.2.2.5**, initially with 2% (w/v) DDM in place of SB3-14.

3.3.3 His-trap affinity chromatography

Protein solubilised in DDM was purified by HisTrap affinity chromatography. Using a peristaltic pump (GE Healthcare Life Sciences, Pump P1, 3.1 mm tubing) the HisTrap column (HP 5 ml Global, GE Healthcare Life Sciences) was washed with nH_2O for approximately 5 minutes at a flow rate of 2 ml/min to remove the ethanol used for storage. The column was then equilibrated with Buffer A (**Chapter 2 Table 2.4**) for approximately 5 CV at 2 ml/min. 10

mM imidazole was added to the solubilised protein sample before loading onto the pre-equilibrated HisTrap column at a flow rate of 2 ml/min. The ÄKTA purifier (GE Healthcare Life Sciences) was prepared by washing with nH$_2$O before equilibration with Buffers A and B (**Chapter 2 Table 2.4**) in their respective lines. The HisTrap column was then applied to the ÄKTA system and run at a segmented gradient of 6.1% Buffer B for 10 CV followed by 100% Buffer B for 20 CV, producing a gradient of 30-500 mM imidazole. Fraction volumes of 4 ml were collected at a flow rate of 1 ml/min with a max pressure of 0.5 MPa. Based on the UV trace, fractions from peaks of interest were collected and analysed by SDS-PAGE.

3.3.4 Site directed mutagenesis for his-tag extension

Extension of the hexahistidine tag was completed using the commercial QuikChange XL II SDM kit, as described in **Chapter 2 Section 2.2.1.2.** MOMP His-8 and MOMP His-10 primers are detailed in **Chapter 2 Table 2.6**. PCR products were *DpnI* treated and checked on an agarose gel prior to transformation. Plasmids were prepared following the method described in **Chapter 2 Section 2.2.1.10** before being sequenced with T7 reverse sequencing primers.

3.3.5 Small scale detergent screening

Using microcentrifuge tubes (eppendorfs), membranes were solubilised under a number of conditions (4 °C for 2h, 4 °C overnight, or room temperature for 1h) in the presence of a selected detergent at 1% (w/v) (DDM, OG, LDAO, FC-12, CHAPS, or v/v Triton X-100). Samples were taken before high speed centrifugation (BC) at 50,000 rpm (153,700 x g) for 45 minutes at 4 °C, using a Beckman Coulter-Optima Max-TL benchtop ultra-centrifuge with a TLA-55 fixed-angle rotor. Samples of the soluble fraction (SF) were taken following

centrifugation. BC and SF samples were then analysed by Western blotting to compare solubility.

3.3.6 Detergent screening using fluorescence-detected size exclusion chromatography (FSEC)

Membrane samples were thawed and diluted to a final concentration of 3 mg/ml in solubilisation buffer (**Chapter 2 Table 2.4**) to a total volume of 1 ml, with each aliquot containing one of the following detergents at 1%; DDM, OG, LDAO, SB3-12, SB3-14, LMNG, OGNG, UDM, Cymal 5, C12E8, or 5% Elugent (full details of detergent properties are shown in **Section 3.1.2 Table 3.1**). The membranes were solubilised at 4 °C for 1h with gentle agitation, before centrifugation at 50,000 rpm (153,700 x g) for 45 minutes at 4 °C. In glass vials, 102.6 μl of solubilised protein was mixed with 7.4 μl of P3NTA probe, creating a final concentration of 1.5 μM probe, before being loaded in the Shimadzu high-performance liquid chromatography (HPLC) machine equipped with fluorescence detection. Samples were resolved on a Superose 6 10/300 L column (GE Healthcare Life Sciences) with a flow rate of 0.3 ml using degassed FSEC flow buffer (**Chapter 2 Table 2.4**), washing in between samples with degassed nH$_2$O.

3.3.7 Small scale Nickel resin binding trial

Nickel Superflow resin (Qiagen) was added to SB3-14 solubilised membranes along with varying concentrations of imidazole (10-90 mM in 20 mM increments) and stirred at 4 °C for 1h. The samples were then pelleted at 8,000 rpm for 3 minutes and the supernatant stored as 'unbound' sample (U), before being washed in elution buffer (**Chapter 2 Table 2.4**) containing 300 mM imidazole, and named 'eluted' sample (E). The unbound and eluted samples were then analysed with SDS-PAGE and Western blotting.

3.3.8 Detergent exchange

During nickel gravity flow purification, increasing concentrations of OG were introduced into the wash and elution buffers in order to detergent exchange from SB3-14, which is not amenable to crystallography, into OG. Along with 0.01% (w/v) SB3-14, the 30 mM and 50 mM imidazole wash buffers were supplemented with 0.6% and 0.9% (w/v) OG, respectively. In the 200 mM elution buffer, SB3-14 was excluded and instead replaced with 1.17% (w/v) OG, equating to two times CMC.

3.3.9 Dialysis and concentration

Following SDS-PAGE analysis of the fractions from HisTrap affinity chromatography, fractions were selected and pooled together if they displayed a band at the appropriate MW, or alternatively the eluted protein from the nickel gravity flow procedure was used. Dialysis tubing (Spectra/Por Dialysis Membrane, MWCO 6-8,000) was soaked for a few minutes in nH$_2$O. One end was clamped tightly shut, the protein solution added from the opposite end, and then clamped forming a sealed tube with minimal air. This was placed inside a beaker containing 1 L of dialysis buffer (**Chapter 2 Table 2.4**) and left gently stirring at 100 rpm overnight at 4 °C.

The dialysed protein was concentrated using an appropriate Amicon Ultracentrifugal concentrator (100 kDa for DDM purified MOMP, and 50 kDa for SB3-14 solubilised MOMP and OG detergent exchanged MOMP), using a Beckman Coulter Allegra X-22R centrifuge at 3,500 rpm (2,800 x g) 4 °C for 5 minutes. The protein was resuspended after each centrifugation until an appropriate concentration and volume was obtained. Protein concentration was estimated by absorbance at 280 nm using a Nanodrop Spectrometer.

3.3.10 Size exclusion chromatography (SEC)

Purified and concentrated protein, buffer exchanged from SB3-14 into DDM, was centrifuged for 10 minutes, 13,000 rpm (~18,000 x g) at 4 °C to pellet any precipitants. The protein was then buffer exchanged into 1.17% (w/v) OG using an ÄKTA system fitted with a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) after loading into a 500 μl injection loop, using detergent exchange buffer (**Chapter 2 Table 2.4**). A flow rate of 0.4 ml/min was set for 1 CV (25 ml) with 0.2 ml fractions collected in a 96 well microtube plate. Fractions located within the peak trace were pooled and concentrated with a 50 kDa Amicon protein concentrator.

3.3.11 Crystallisation Screening

Purified protein buffer exchanged into DDM, concentrated to ~8 mg/ml, was used for crystallisation trials in hydrophobic CrystalQuick 96-well sitting drop plates (Greiner). Prior to set up, protein samples were centrifuged at 13,000 rpm (~18,000 x g) for 10 minutes in order to pellet any aggregates. Sitting drop vapour diffusion screens using 100 nl protein and 100 nl reagent from the crystal screens MemGold I and MemGold II (Molecular Dimensions) were set up using a Mosquito robot (TTP Labtech). The plates were incubated at 20 °C in a Molecular Dimensions incubator and crystal formation was monitored periodically using a Leica DFC295 microscope.

3.4   Results

3.4.1 Optimisation of *E. coli* growth conditions between DLS and UoR

Previously, a low resolution crystal structure for MOMP was resolved from recombinant protein expressed in *E. coli* (Atanu, 2014). In order to improve the resolution and gain more information regarding MOMP's structure, this method of expression was replicated at the

two collaborating institutions. However, between laboratories at Diamond Light Source (DLS) and the University of Reading (UoR) there was an observable difference in the time taken for C41 (DE3) *E. coli* cells transformed with pETMOMP$_{H6}$ to reach an OD$_{600}$ of ~0.5, at which induction is deemed optimal. Despite following the same protocol, cells at DLS took approximately seven hours to reach an appropriate OD$_{600}$, as opposed to those at UoR which took just three hours, the latter being more typical of this cell strain. At DLS, the OD$_{600}$ values of different overnight culture conditions were examined before subculture into 10 ml LB with varying glucose concentrations (0%, 0.5%, 1%, w/v). As shown in **Table 3.2**, each of the overnight cultures had very similar OD$_{600}$ values.

Condition	Temp (°C)	Volume (ml)	OD$_{600}$
A	37	150	1.60
B	37	10	1.55
C	30	150	1.55

Table 3.2. **Comparison of overnight culture conditions.** C41 (DE3) *E. coli* cells transformed with pETMOMP$_{H6}$ were grown overnight in LB media following inoculation with a single colony. OD$_{600}$ measurements were taken after 16 h of overnight growth.

After subculture into fresh LB, an OD$_{600}$ measurement was taken every hour until the culture passed the induction stage of ~0.5, shown in **Figure 3.4**. From **Figure 3.4** it was clear that a reduced overnight temperature of 30 °C (condition C) resulted in a shorter period of growth, of around three and a half hours, before reaching an OD$_{600}$ of ~0.5. Cultures subcultured with condition A, 150 ml at 37 °C, demonstrated a considerably longer growth period of approximately seven and a half hours, with condition B, a smaller volume of 10 ml at 37 °C, also taking over six hours to reach an OD$_{600}$ of ~0.5. Differing concentrations of

glucose did not appear to have a significant effect upon growth under any of the conditions tested. Due to improved growth, an overnight temperature of 30 °C was used in all further protein expression studies at DLS. Additionally, due to the slight improvement in growth between 10 ml and 150 ml cultures, subsequent large scale cultures were grown in 2 L Erlenmeyer flasks with just 500 ml of culture, as opposed to 1 L in order to improve aeration.
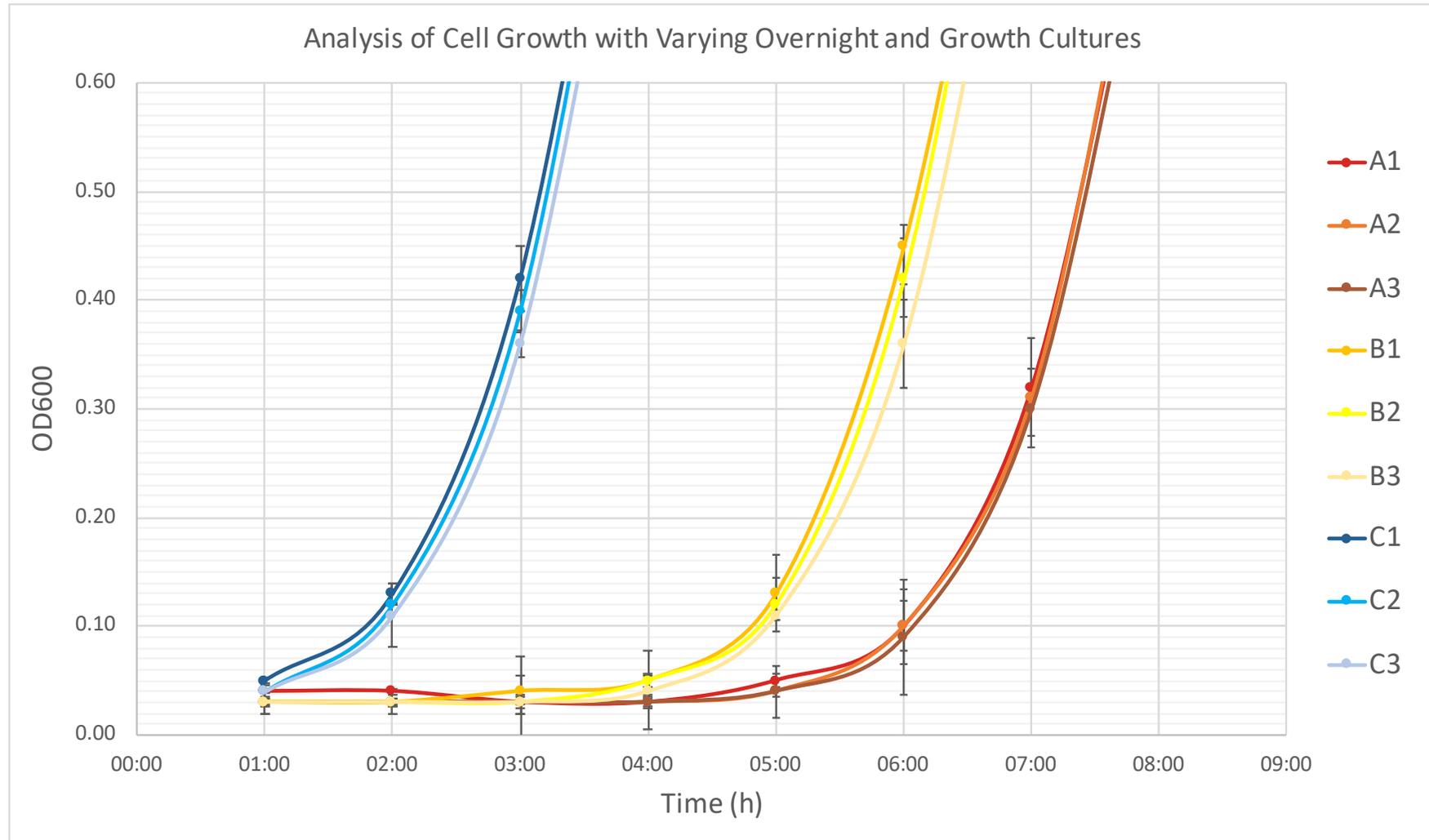
Figure 3.4. **Graph showing the time taken to reach an OD$_{600}$ of ~0.5 following subculture with different overnight cultures.** C41 (DE3) *E. coli* cells transformed with pETMOMP$_{H6}$ were grown overnight under a variety of temperatures and volumes (A-C) and glucose concentrations (1-3), before being subcultured into 10 ml fresh LB and grown at 37 °C, 225 rpm. Measurements were taken hourly until an OD$_{600}$ of ~0.5-0.6 was passed. A (red), 37 °C 150 ml overnight culture; B (yellow), 37 °C 10 ml overnight culture; C (blue), 30 °C 150 ml overnight culture. 1, 0% glucose; 2, 0.5% glucose; 3, 1% glucose. Error bars represent standard deviation.

3.4.2 Poor solubility of MOMP with detergent DDM

Following optimisation of the large-scale growth conditions at DLS, crude membranes were prepared from 6 L of culture, as described in **Chapter 2 Section 2.2.2.4**, then solubilised with 2% (w/v) DDM following the his-trap affinity chromatography procedure (**Section 3.3.3**) in order to replicate the procedure previously successful in the production of rMOMP crystals (Atanu, 2014). The chromatograph obtained from the IMAC purification revealed an extremely small peak in fractions B1-B3, as shown in **Figure 3.5A**, eluted with an imidazole gradient between 30-500 mM. At each stage of the process, samples were taken for analysis with Western blotting, shown in **Figure 3.5B**. From these results, it was evident that the majority of MOMP was DDM insoluble, as indicated by the arrow in well 8, which shows a considerably thicker band in comparison to the soluble DDM fraction in well 7. Additionally, the sarkosyl isolation step, included in order to solubilise contaminant inner membrane and cytosolic proteins, whilst isolating the outer membranes in the insoluble fraction (Filip *et al*., 1973), showed a strong band for MOMP in well 6 corresponding to the insoluble pellet. Repetition of this procedure, scaled up to 24 L of culture, did not increase the protein yield with a similar peak observed within the UV trace beginning at fraction B1, as shown in **Figure 3.5C.**
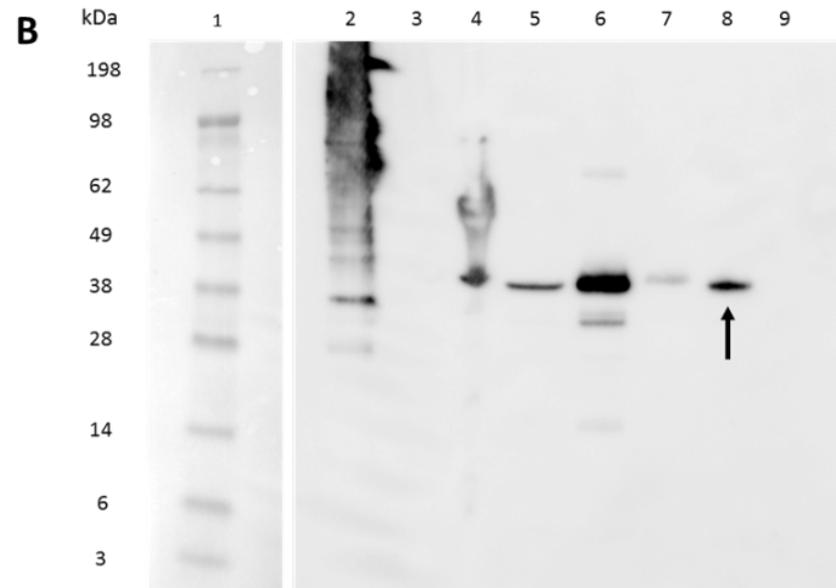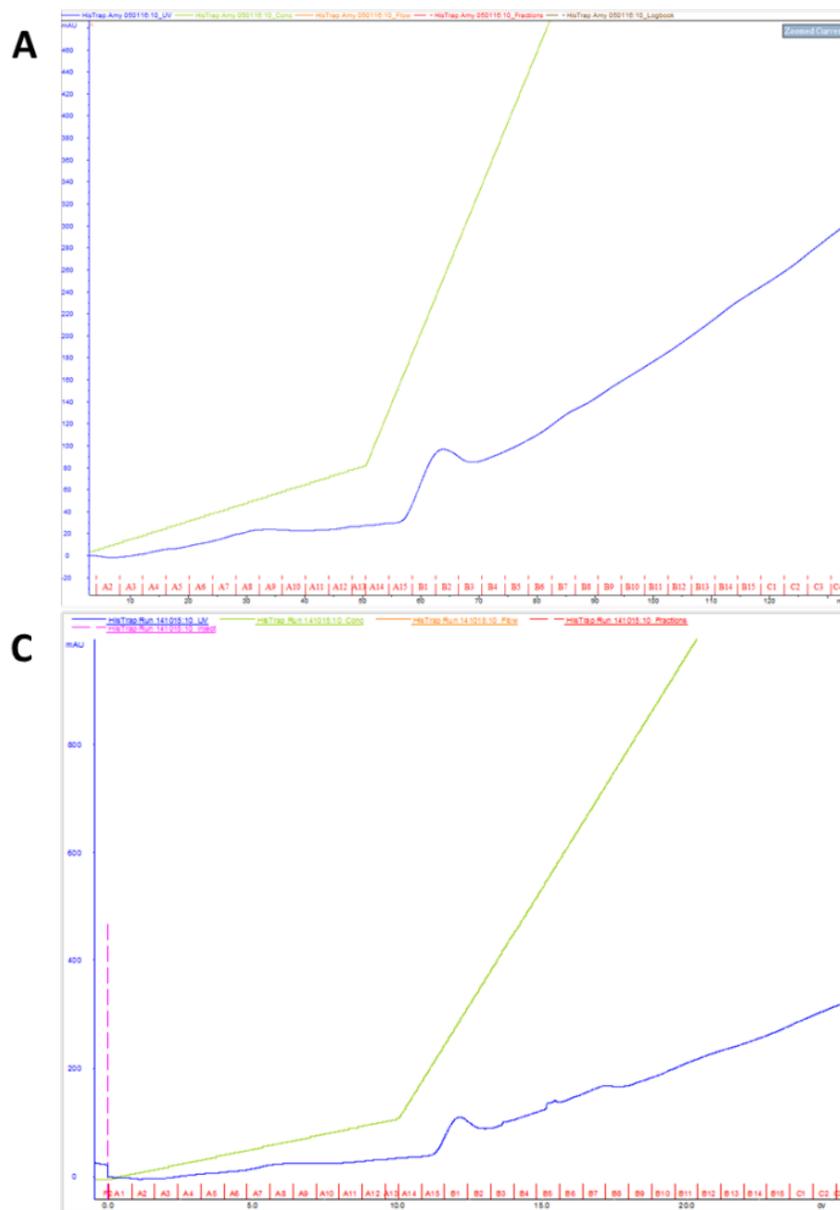
Figure 3.5. **Purification of pETMOMP$_{H6}$ membranes expressed in C41 (DE3) cells with DDM solubilisation. A, IMAC UV trace from 6 L of original culture.** Only a small peak was eluted in fractions B1-B3. Dark blue, UV absorbance; green, imidazole concentration gradient. **B, Western blot of samples taken during the IMAC purification of a 6 L membrane prep.** 1, MW protein standard; 2, his-tagged protein standard; 3, blank; 4, crude membranes; 5, sarkosyl soluble fraction; 6, sarkosyl insoluble fraction; 7, DDM soluble fraction; 8, DDM insoluble fraction; 9, HisTrap column flow through. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody. **C, IMAC UV trace from 24 L of original culture.** No improvement was observed between peak volumes for 6 L and 24 L membrane preps.

3.4.3   His-tag extension of the pETMOMP$_{H6}$ construct

The hexahistidine tag at the C-terminus of MOMP in the pETMOMP$_{H6}$ construct was extended to His-8 and His-10 by SDM in an effort to improve binding to the Ni-NTA column during protein purification. The mutations were confirmed through sequencing with the T7 reverse sequencing primer (**Chapter 2 Table 2.6**) before constructs were transformed into both C41 (DE3) and BL21 (DE3) *E. coli* cells to test protein expression. **Figure 3.6** shows a Western blot comparing whole cell expression of the three constructs, whereby it was evident that both MOMP$_{H8}$ and MOMP$_{H10}$ greatly reduced expression in both strains in comparison to the positive control, MOMP$_{H6}$. MOMP$_{H8}$ expressed in BL21 (DE3) cells had comparable levels to that of MOMP$_{H6}$ in C41 (DE3), while MOMP$_{H10}$ was not expressed in either strain. The Coomassie blue stained loading control is shown in **Figure A3.1** in **Appendix 3**. As the pETMOMP$_{H10}$ construct demonstrated no protein expression, it was again sequenced in the opposite direction using the T7 forward sequencing primer (**Chapter 2 Table 2.6**) in order to identify any untoward mutations that may have been introduced into the MOMP sequence, however, none were detected.
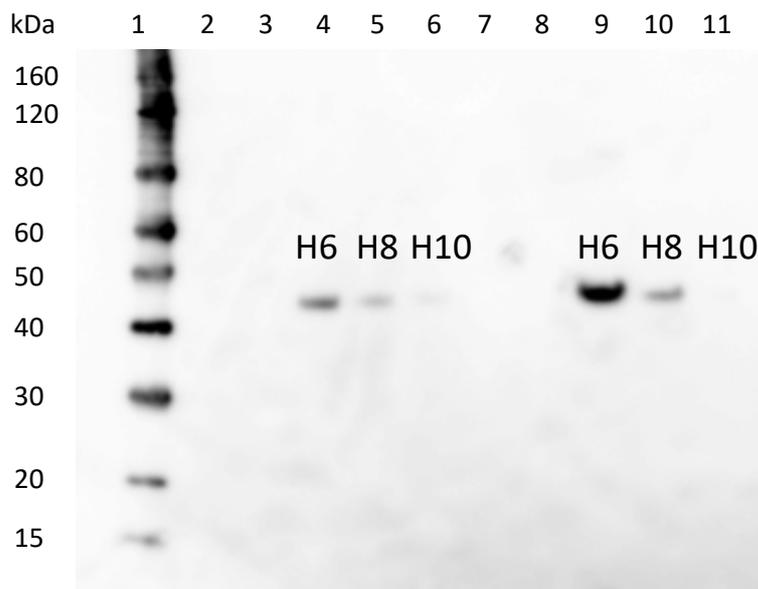
Figure 3.6. **Western blot showing expression of His-6, His-8, and His-10 pETMOMP constructs in whole cells.** Constructs were transformed into both C41 (DE3) and BL21 (DE3) *E. coli* cells. 1, His-tagged protein standard; 2, Blank; 3, C41 (DE3) control; 4, C41 (DE3) His-6; 5, C41 (DE3) His-8; 6, C41 (DE3) His-10; 7, Blank; 8, BL21 (DE3) control; 9, BL21 (DE3) His-6; 10, BL21 (DE3) His-8; 11, BL21 (DE3) His-10. Protein was detected using anti-his HRP conjugated monoclonal mouse antibody.

3.4.4 Detergent screening for optimal MOMP solubilisation using Western blotting

Due to MOMP's low expression, Western blotting was required to visualise the desired MW band and, despite the advice from Newby *et al* in 2009, the membranes were not diluted prior to solubilisation, again due to MOMP's low expression levels. Additionally, the sarkosyl isolation step was not included in initial detergent screens, due to the variability observed between the starting concentrations of MOMP in each sample in preliminary tests (data not shown). Three conditions were tested; room temperature for one hour (**Figure 3.7A**), 4 °C for two hours (**Figure 3.7B**), and 4 °C overnight (**Figure 3.7C**), with the following detergents at 1%; DDM, OG, LDAO, FC-12, CHAPS, and Triton X-100. The only wells indicating solubilised

MOMP are shown in **Figure 3.7A**, in wells 9 and 13, corresponding to FC-12 and Triton X-100, respectively. In all of the BC samples solubilised at room temperature (**Figure 3.7A**), there were multiple bands for MOMP in addition to the expected monomeric band at ~40 kDa, suggestive of both a dimeric form at ~80 kDa and degradation at the lower MW bands of ~25 kDa and 15 kDa. These additional bands were also faintly visible in **Figure 3.7B** and **Figure 3.7C**.
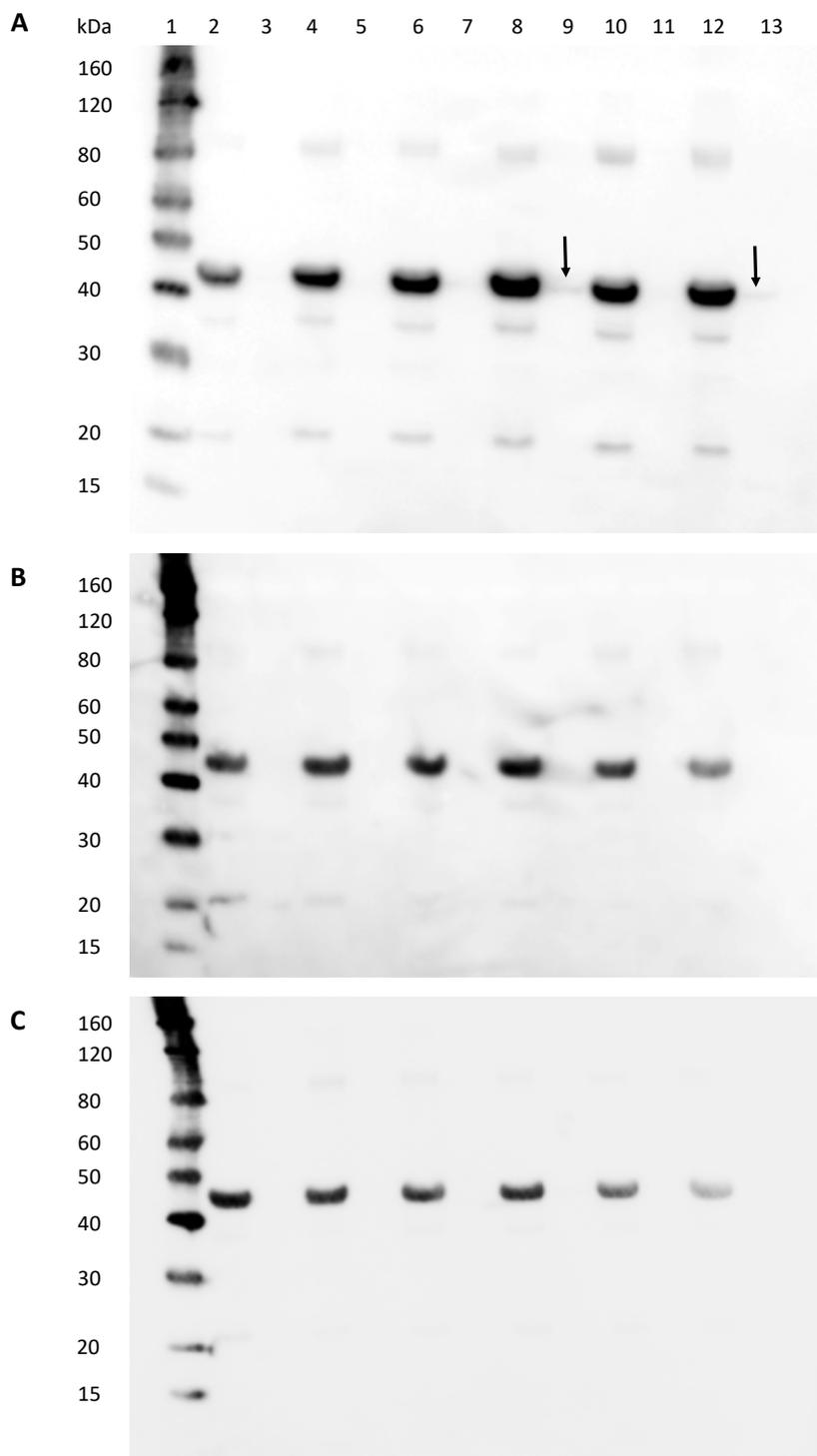
Figure 3.7. **Western blots of the detergent solubility screens under the following conditions; A, Room temperature for 1h, B, 4 °C for 2h, and C, 4 °C overnight.** 1, His-tagged protein standard; 2, DDM BC; 3, DDM SF; 4, OG BC; 5, OG SF; 6, LDAO BC; 7, LDAO SF; 8, FC-12 BC; 9, FC-12 SF; 10, CHAPS BC; 11, CHAPS SF; 12, Triton X-100 BC; 13, Triton X-100 SF. BC refers to 'before centrifugation' and includes soluble and non-soluble proteins, and SF refers to solely the 'soluble fraction'. Bold arrows indicate presence of MOMP. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

Due to the poor solubility observed, **Figure 3.7A-C**, the sarkosyl isolation step was reintroduced in an attempt to remove as many contaminating proteins as possible, to allow more mild detergents, such as DDM, access to MOMP. The optimal sarkosyl concentration was investigated by testing a gradient of 2-10% (v/v), in 2% increments. Unlike before, the SF was not analysed, but instead the insoluble pellet containing MOMP, denoted P, was resuspended and analysed. As shown in **Figure 3.8**, all increments of sarkosyl successfully isolated MOMP, with equal protein levels observed between BC and P samples. Therefore, to avoid unnecessarily harsh treatment of MOMP, the lowest concentration of sarkosyl, 2% (v/v), was used with stirring at 4 °C for 1h.
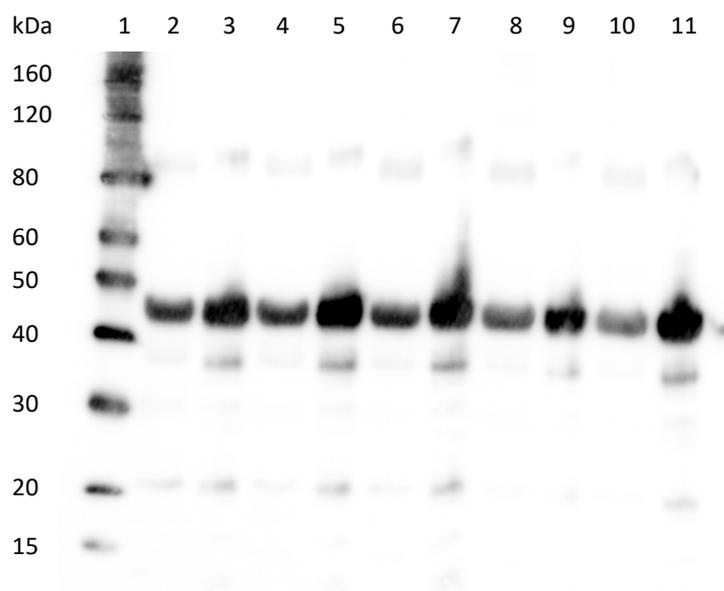
Figure 3.8. **Western blot of a 2-10% sarkosyl isolation gradient.** 1, His tagged ladder; 2, 2% BC; 3, 2% P; 4, 4% BC; 5, 4% P; 6, 6% BC; 7, 6% P; 8, 8% BC; 9, 8% P; 10, 10% BC; 11, 10% P. BC refers to before centrifugation and includes soluble and non-soluble proteins and P refers to insoluble protein in resuspended pellets. All % (v/v) values are in relation to sarkosyl. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

A number of additional conditions were tested along with the sarkosyl isolation step in order to encourage solubilisation. The pH of the Tris buffer solution was increased from 7.5 to 8 (room temperature readings, which equate to 8.1 and 8.58 at 4 °C) in order to raise the

pH above MOMP's isolelectric point (pI) of 8.29, calculated using the online tool ProtParam to include the v5 epitope and hexa-histidine tag which are not cleaved off during the purification procedure (Gasteiger *et al.*, 2005). The previous six detergents, shown in **Figure 3.7**, were again used for solubilsation for 2h at 4 °C although no protein bands were observed in the soluble fractions with Western blotting, apart from that with FC-12 as before (data not shown). An additional detergent of zwitterionic nature, SB3-12, was used alongside DDM and FC-12 at a higher temperature of 30 °C overnight, and the sarkosyl isolation step was increased to 2h at room temperature, but again only FC-12 solubilised any MOMP (data not shown).

3.4.5 Detergent screening using the novel P3NTA tag with fluorescence-detection size-exclusion chromatography (FSEC)

For a higher throughput method of detergent screening, FSEC was carried out on crude membranes produced from pETMOMP$_{H6}$ transformed C41 (DE3) cells grown in 12 L of LB. A range of detergents, some previously shown to be promising by Atanu in 2014 such as DDM, OG and LDAO, were tested in order to determine the most promising candidates for MOMP solubilisation.

Figure 3.9 shows the results of this screening where the unbound probe peak at 19.5 ml is clearly visible for all samples (**Figure 3.9A**), along with the void volume at 9 ml, shown more clearly in **Figure 3.9C** for the LMNG sample, although not observed with DDM (**Figure 3.9A**). Within every trace, it was notable that the unbound P3NTA probe was not completely washed out from the previous sample, due to the truncated peak at 19.5 ml, which results in a gradual decrease back to baseline between 0 ml and 5 ml at the start of each trace, most clearly visible in **Figure 3.9B**.

A

Detergent screening of MOMP using FSEC

B

Detergent screening of MOMP using FSEC

C

Detergent screening of MOMP using FSEC
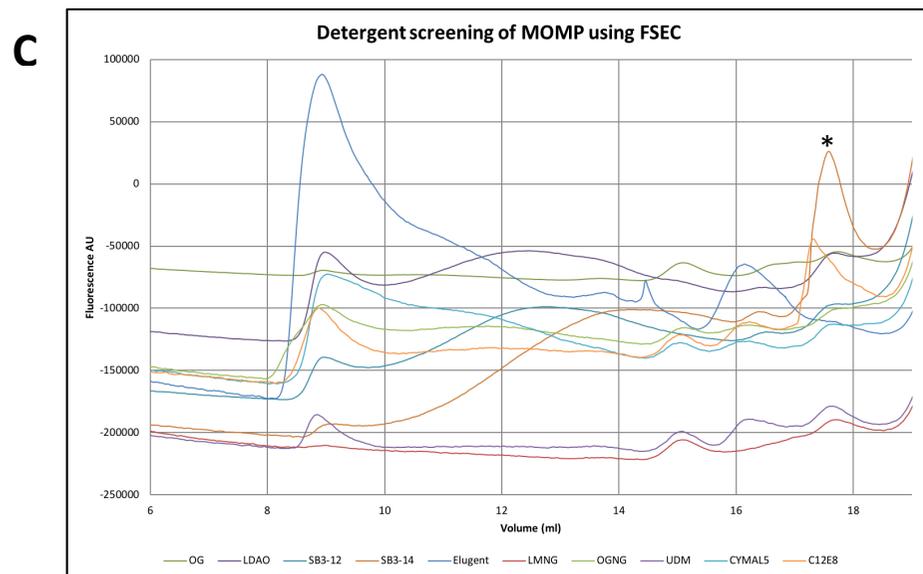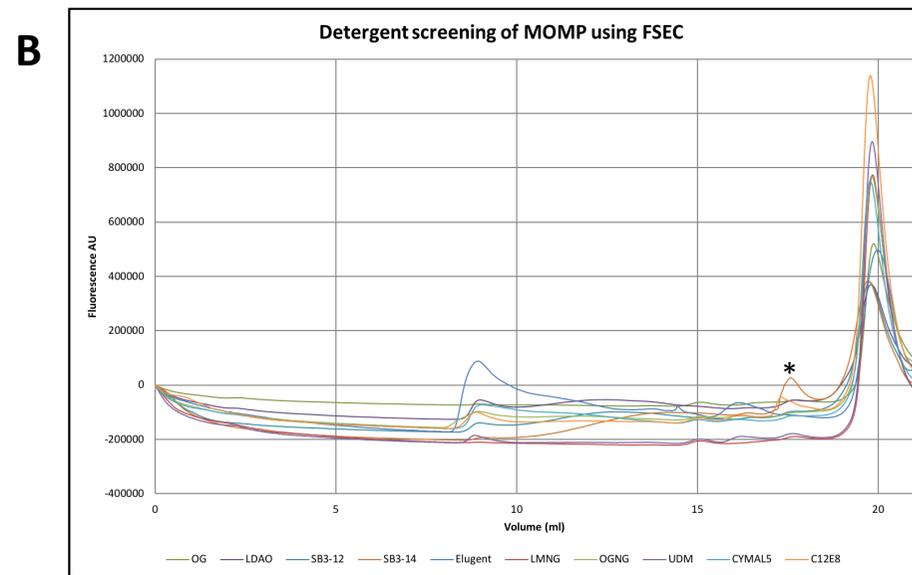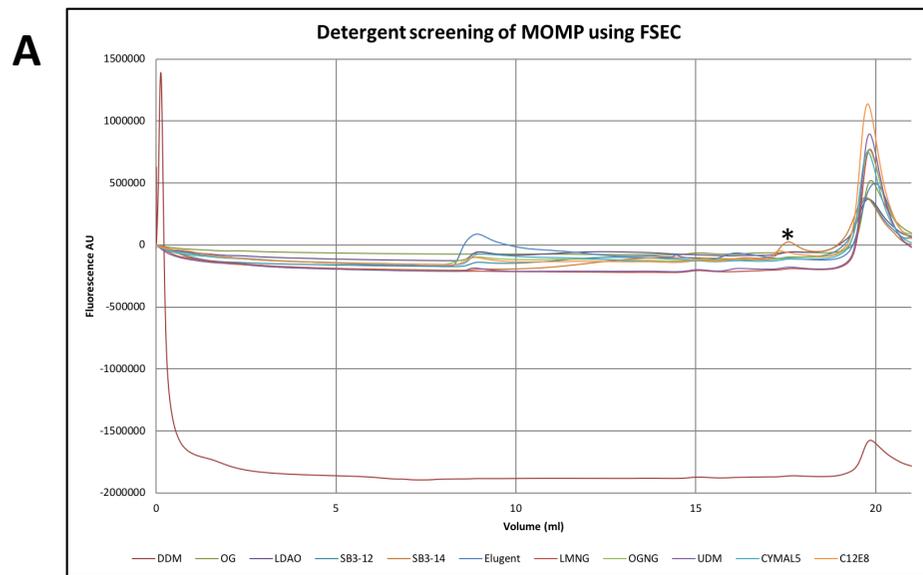
Figure 3.9. **FSEC traces of detergent solubilised MOMP labelled with P3NTA probe**. **A, All detergents screened, B, omission of DDM trace, and C, omission of P3NTA peak at 19.5 ml.** Membranes were diluted to 3 mg/ml and solubilised with 1% detergent (Elugent at 5%) for 1h at 4 °C. Soluble fractions were supplemented with 1.5 μmol P3NTA and analysed by FSEC. The graph shows the fluorescence absorption for each sample. The peak visible at 17.5 ml for SB3-14 (*) corresponds to calibrated values for a protein of 36 kDa. Free P3NTA probe elutes at 19.5 ml. The void volume of the column was 9 ml.

Notably, the DDM sample which is only shown in **Figure 3.9A** was at a considerably lower absorbance unit than the other samples, had no void volume, and only a small peak for unbound probe. Additionally, there is a very large inexplicable initial peak.

A small peak was visible for SB3-14, a detergent used in both *Campylobacter jejuni* major outer membrane protein solubilisation (Ferrara *et al.,* 2016) and in *C. trachomatis* MOMP solubilisation (Findlay *et al*., 2005; Pal *et al*., 2005), at 17.5 ml (**Figure 3.9C**, indicated by *) which corresponds to a MW of approximately 36 kDa from previously calibrated values for the P3NTA probe, and therefore was likely due to MOMP. A shorter peak at this location also was present in the C12E8 trace, partially overlapping the peak of SB3-14, which also was likely due to monomeric MOMP although at a much reduced quantity.

The void volume present at 9 ml for the majority of samples was due to aggregated protein, with the 5% Elugent sample producing the highest and broadest peak, most likely resulting from the increased detergent concentration. There were a number of additional small peaks for other detergents in the trace, which were likely due to multimeric protein. After the peak located at 17.5 ml, the second most consistent protein peak was seen at approximately 15 ml for the following samples; LMNG, UDM, Cymal 5, C12E8, OGNG, and OG, and appears to correlate to a protein of 95 kDa.

Despite the promising monomeric peak for the SB3-14 sample, a broader peak appeared just after the void volume between 10 ml and 16 ml, which was likely due to aggregated multimeric protein. Optimisation of these conditions, such as detergent concentration, temperature, duration of solubilisation etc. may reduce aggregation.

Due to the novelty of the P3NTA tag, it was necessary to analyse these results through Western blotting to act as a comparison with previous detergent screening experiments,

along with an SDS-PAGE gel imaged with a fluorescent detector for visualisation of any

P3NTA tagged protein, as shown in **Figure 3.10**.
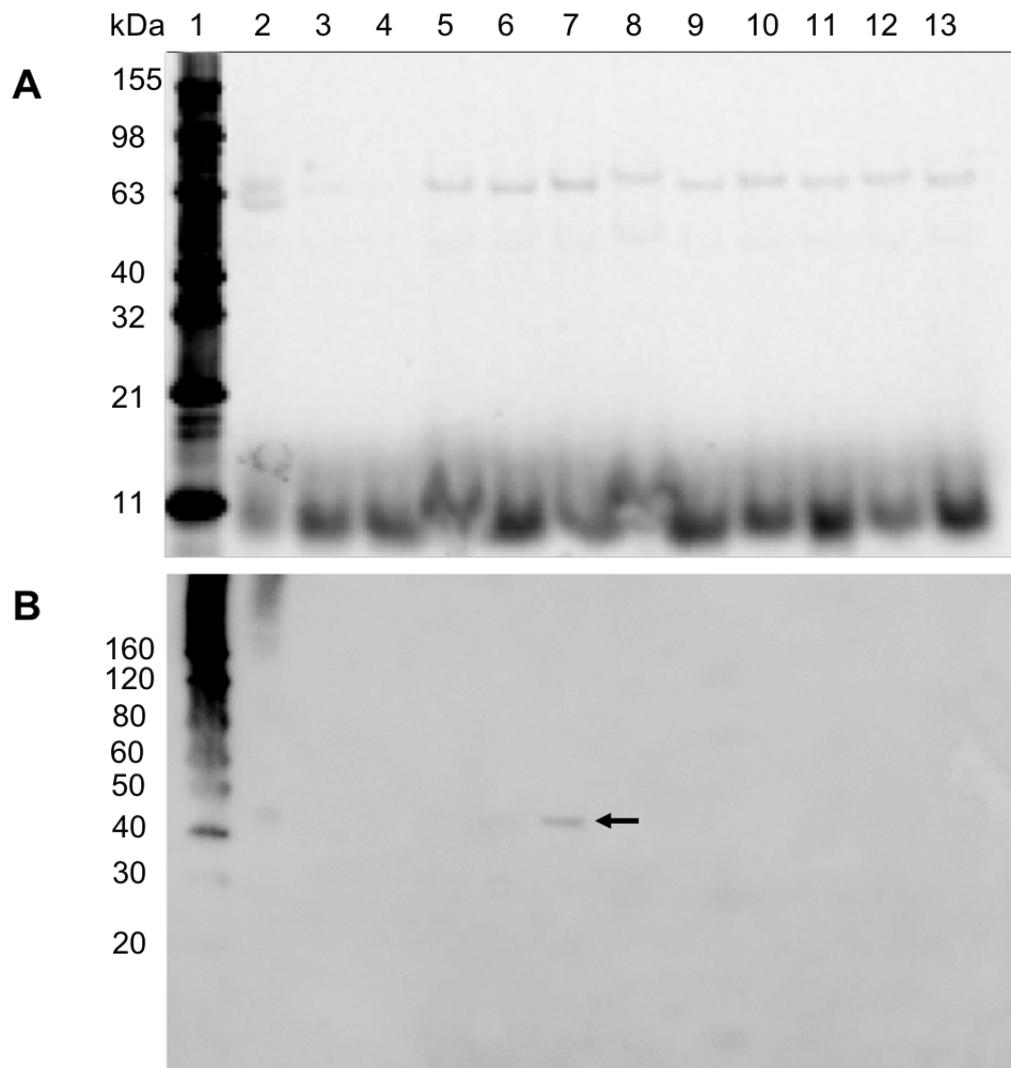
Figure 3.10. Detergent solubilised membrane samples previously analysed by FSEC. A, SDS-PAGE with 1/200 P3NTA fluorescent probe and B, Western blot with anti-his antibody. Gels were run at 100 v for 120 minutes at 4 °C to prevent degradation of the P3NTA probe. 1, Protein standard; 2, 1 mg/ml membranes; 3, DDM; 4, OG; 5, LDAO; 6, SB3-12; 7, SB3-14; 8, 5% Elugent; 9, LMNG; 10, OGNG; 11, UDM; 12, Cymal 5; 13, C12E8. All detergents were at 1% unless otherwise stated. The arrow indicates the location of MOMP. For Western blotting, protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

In **Figure 3.10A** there was a strong band at approximately 63 kDa, with another weaker

band situated below at 45 kDa. The latter is likely due to monomeric MOMP, in addition to

the 3 kDa tag, and the former possibly due to a histidine rich contaminant protein. The

highly fluorescent band present in the lower molecular weight region of the gel was a dye front caused by excess P3NTA. **Figure 3.10B** shows a Western blot of the same samples labelled with anti-his antibody, where only a single band corresponding to MOMP was visible, as indicated by the arrow in well 7. MOMP in this sample was solubilised with SB3-14, the same detergent from the FSEC trace in **Figure 3.9** that showed a reasonable peak at approximately 36 kDa. Interestingly, the 63 kDa bands that were clearly present in **Figure 3.10A** were not visible in the Western blot. Over exposure of the blot resulted in a very faint band in well 7 for SB3-12 (not shown here). Additionally, in the crude membranes sample in well 3 there was a large smear near the top of the well, which was likely due to aggregated protein and a small faint band at approximately 40 kDa corresponding to MOMP.

3.4.6   MOMP solubilisation with SB3-14 assessed with Western blotting

After establishing SB3-14 as the most suitable detergent for native rMOMP solubilisation with FSEC, an additional solubilisation experiment was conducted in order to test two detergent concentrations of 1% and 2%, and two novel cysteine mutants, C201A and C201/203A. Western blotting was used as before in order to create a comparable measure. From the Western blot shown in **Figure 3.11** it was evident that SB3-14 had dramatically improved MOMP solubility for both concentrations, compared to the detergents tested in **Section 3.4.4**. Notably, the same detergent failed to solubilise the two cysteine mutants, with 2% SB3-14 resulting in a higher molecular weight band for MOMP at approximately 50 kDa in both mutants prior to centrifugation. Due to the adverse effect observed in the mutants, and the negligible difference observed in solubilisation between 1% and 2% SB3-14 for wild type MOMP, a concentration of 1% SB3-14 was used in future solubilisations of native MOMP.
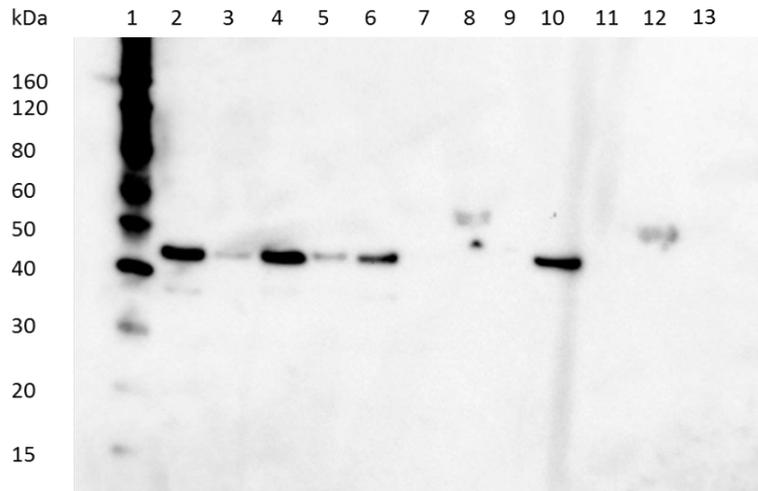
Figure 3.11. Western blot of SB3-14 solubility screen with native and mutant MOMP. 1, protein standard; 2, native 1% BC; 3, native 1% SF; 4, native 2% BC; 5, native 2% SF; 6, C201A 1% BC; 7, C201A 1% SF; 8, C201A 2% BC; 9, C201A 2% SF; 10, C201/3A 1% BC; 11, C201/3A 1% SF; 12, C201/3A 2% BC; 13, C201/3A 2% SF. BC refers to before centrifugation and includes soluble and non-soluble proteins, and SF refers to solely the soluble fraction. All percentages refer to SB3-14 (w/v) concentration. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

3.4.7   Optimisation of MOMP purification

A 3 L membrane prep from C41 (DE3) cells expressing native MOMP was again purified following the his-trap affinity chromatography procedure outlined in **Section 3.3.3**, instead with solubilisation by zwitterionic detergent SB3-14. However, at 4 °C buffers A and B (**Chapter 2 Table 2.4**) containing 0.1% (w/v) SB3-14, in place of DDM, became extremely viscous in the thin tubing of the ÄKTA system which resulted in a gradual increase in pressure and concurrently required a decreased flow rate (from 2ml/min to 0.5ml/min). Despite warming the buffers to reduce viscosity, due to the narrow diameter of the tubing and therefore the rapid cooling of the fluid in the lines at 4 °C, the pressure issue could not be overcome, with the flow rate reaching a low of 0.5 ml/min, which quadrupled the predicted run time. Instead, the SB3-14 concentration was further decreased to 0.01% (w/v) and MOMP was washed and eluted from the his-trap column using a peristaltic pump with

30 mM and 40 mM imidazole wash buffers and 250 mM imidazole elution buffer. However, the resulting protein was not pure, as analysed by SDS-PAGE (**Figure 3.12**).
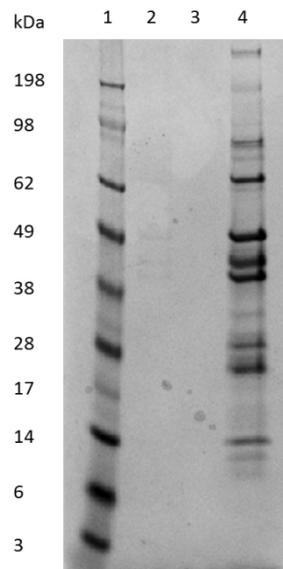
Figure 3.12. Coomassie blue stained SDS-PAGE gel demonstrating the low purity achieved during initial purification of rMOMP solubilised with SB3-14. 1, Protein standard; 2, dialysed protein; 3, concentrator flow through; 4, concentrated sample. Due to the precipitation of SB3-14 at low temperatues, rMOMP bound to the his-trap column was eluted manually with a peristaltic pump.

As a result, further purifications were carried out using nickel Superflow resin to avoid high pressure damage to the his-trap column. As the elution profile of the POI could not be visualised by this method, a small-scale trial was first conducted to determine the appropriate concentrations of imidazole to be used for the equilibration, washes, and final elution in order to achieve a higher purity.

A low concentration of imidazole is typically added to the nickel resin binding buffer in order to prevent non-specific binding of contaminant proteins while still permitting near optimal binding of the POI. In order to monitor the binding efficiency of his-tagged MOMP, a small-scale nickel resin binding trial was conducted whereby imidazole concentrations were increased incrementally, between 10 mM and 90 mM. The results are shown in **Figure 3.13**. From the Coomassie blue stained SDS-PAGE gel in **Figure 3.13A** it was evident that no

supplementation with imidazole (well 4) resulted in the binding of many contaminant proteins. Cross validation with the Western blot, shown in **Figure 3.13B**, revealed that only imidazole concentrations of 10 mM and 30 mM still permitted MOMP binding, due to the bands located at approximately 40 kDa in wells 6 and 8. However, at the lower concentration of 10 mM, contaminant proteins were still visible in the SDS-PAGE gel (**Figure 3.13A**). As the band corresponding to MOMP in the 30 mM sample appeared to be weaker, an intermediate imidazole concentration of 20 mM was used in future equilibrations with nickel resin.
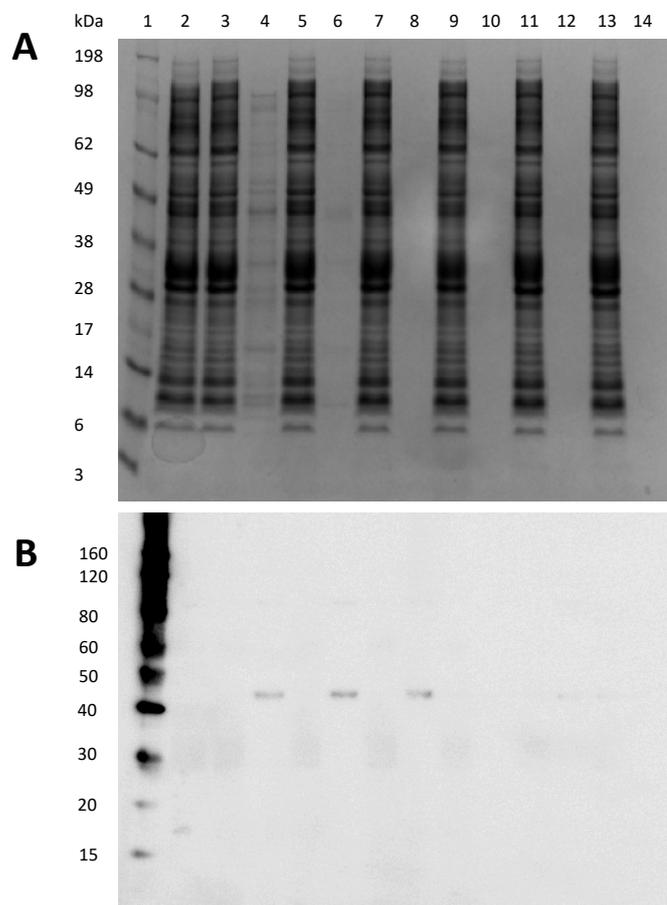
Figure 3.13. Small scale screen of appropriate imidazole concentration for equilibration with nickel resin. A, Coomassie blue stained SDS-PAGE gel and B, Western blot. 1, Protein standard; 2, Before centrifugation; 3, U0; 4, E0; 5, U10; 6, E10; 7, U30; 8, E30; 9, U50; 10, E50; 11, U70; 12, E70; 13, U90; 14, E90. Where U refers to unbound protein and E to eluted protein. All numbers refer to imidazole concentration in mM. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

Another small-scale purification of 3 L was carried out using nickel resin, whereby two wash steps of 40 mM and 60 mM imidazole were included, prior to elution in three stages with 100 mM, 200 mM and 300 mM imidazole. Fractions were taken at each stage, with three taken for each elution concentration. The results of this test are shown in the Western blot in **Figure 3.14**.
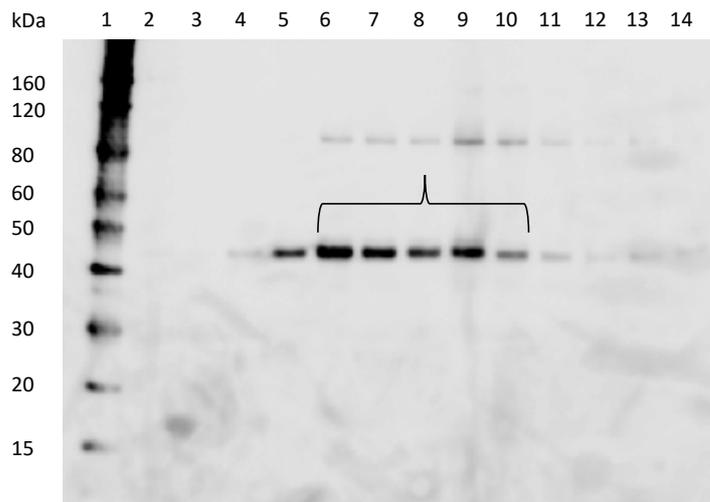
Figure 3.14. Small scale purification to determine the optimal imidazole concentration in 40-60 mM wash and 100-300 mM elution buffers. 1, Protein standard; 2, Blank; 3, resin flow through; 4, 40 mM wash; 5, 60 mM wash; 6-8, 100 mM elutions 1-3; 9-11, 200 mM elutions 1-3; 12-14, 300 mM elutions 1-3. All values in mM refer to the imidazole concentration. Bracket indicates pooled MOMP. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

MOMP appeared to be binding to the resin, with no unbound protein present in the flow through fraction in well 3. From **Figure 3.14**, it is also evident that the majority of MOMP was eluted in the 100 mM fractions and the first two 200 mM fractions and, therefore, the 300 mM elution step was omitted in future purifications. Additionally, a substantial amount of protein was lost during the 60 mM wash step (well 5) and a small quantity lost in the 40

mM wash. Therefore, wash steps of 30 mM and 50 mM were tested in the subsequent small-scale purification, the results of which are shown in **Figure 3.15**.
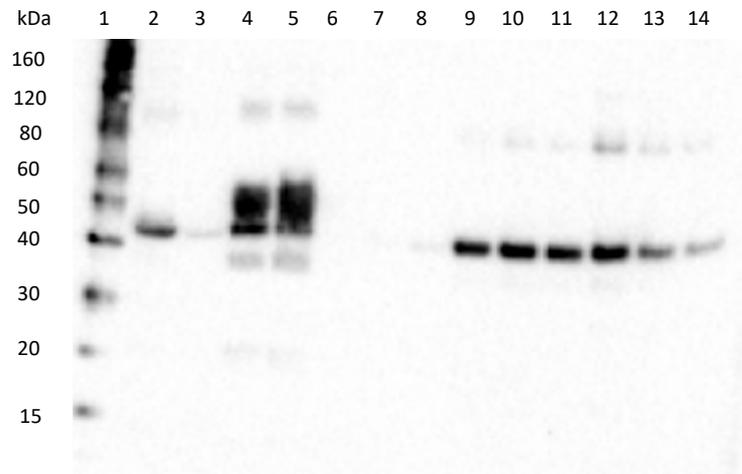
Figure 3.15. Small scale purification to determine the optimal imidazole concentration in 30-50 mM wash buffers. 1, Protein standard; 2, sarkosyl BC; 3, sarksoyl SF; 4, SB3-14 BC; 5, SB3-14 SF; 6, resin flow through; 7, 30 mM wash; 8, 50 mM wash; 9-11, 100 mM elutions 1-3; 11-14, 200 mM elutions. All values in mM refer to the imidazole concentration; BC, before centrifugation; SF, soluble fraction. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody.

From this Western blot, there appeared to be effective solubilisation of MOMP with SB3-14, due to the two equal bands in wells 4 and 5, corresponding to samples before centrifugation and the soluble fraction, respectively. Additionally, protein was not lost during the two new wash steps of 30 mM and 50 mM imidazole. As the majority of the protein was eluted during the two elution steps of 100 mM and 200 mM imidazole, of 5 CV each, a single step of 200 mM imidazole with a decreased elution volume of 2 CV was employed in future purifications. Decreasing the elution volume also served a dual purpose in decreasing the volume of dialysis buffer required and, therefore, conserving expensive detergent, as well as the time taken to concentrate MOMP. **Figure 3.16** shows a Coomassie blue stained SDS-

PAGE gel of purified MOMP obtained from solubilisation and purification with SB3-14, with a strong band indicated at approximately 40 kDa.
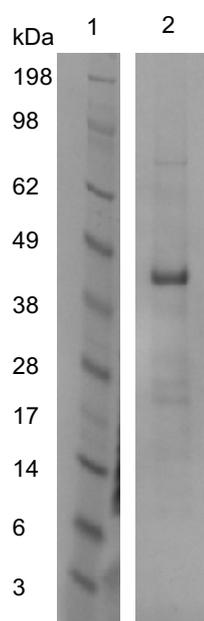
Figure 3.16. Coomassie blue stained SDS-PAGE gel of MOMP purified with SB3-14. 1, protein standard; 2, purified MOMP. Protein was concentrated in a 20ml 50 kDa MW cut off concentrator to 0.2mg/ml from a 6L prep.

3.4.8 Buffer exchange of SB3-14 solubilised MOMP into a crystallisation suitable detergent

Following optimisation of the nickel resin purification procedure, it was important that MOMP was buffer exchanged into a detergent suitable for crystallisation and concentrated to approximately 10 mg/ml for crystal screening. Following elution from the nickel gravity flow column, MOMP was dialysed (dialysis buffer, **Chapter 2 Table 2.4**) to remove excess imidazole with 0.01% (w/v) SB3-14. The protein was then concentrated, using a 50 kDa centrifugal concentrator, which accounted for the micellar size of SB3-14 at ~30 kDa, with the intention of subsequent buffer exchange into OG using SEC. However, excess SB3-14 was concentrated together with MOMP and resulted in ruinous denaturation, as observed through a lack of increase in concentration beyond 1 mg/ml of protein. To overcome this,

MOMP was buffer exchanged into OG during the wash and elution steps on the nickel gravity flow column, whereby further supplementation of SB3-14 was omitted after the 30 mM wash step and the OG concentration (w/v) increased to 0.6%, 0.9%, and 1.17% during the two washes and the final elution step, respectively. The micellar size of OG is ~25 kDa and, therefore, a 30 kDa centrifugal concentrator was selected in order to both avoid concentration of OG and loss of MOMP with a 50 kDa MW cut off. However, precipitation began to occur after a concentration of 1 mg/ml and, therefore, centrifugations were limited to 5 minutes at 3,500 rpm (2,800 x g), as opposed to 10 minutes, to allow protein to be resuspended more frequently to prevent dramatic increases in concentration. Additionally, to prevent loss of protein and to limit the concentration of detergent, the purification was repeated but instead utilising a two-step concentration, initially using a 30 kDa cut off to prevent protein loss before switching to a 50 kDa cut off in order to remove the majority of the excess detergent micelles. However, after concentrating to 200 μl, only 220 μg of protein was obtained from a 12 L membrane prep.

In order to limit the detergent micelle accumulated during the protein concentration step, a detergent of a higher molecular weight was used for buffer exchange, in this case DDM which is approximately 70 kDa. DDM was supplemented to the 30 mM and 50 mM wash buffers at 0.012% and 0.018%, respectively, and the elution buffer at 0.018% which equated to three times CMC, with 0.01% SB3-14 maintained in the two wash buffers only. Concentration to 200 μl, as before, resulted in 170 μg of protein from a 6 L prep. Hypothetically, scaling this up to compare with the 12 L membrane prep volume used previously indicated a ~35% increase in protein yield. Due to this, 18 L of membranes were purified following this method, which resulted in sufficient protein for two crystallisation

screens (MemGold I and II, Molecular Dimensions) in 96 well CrystalQuick plates (Greiner) at

a final concentration of 7.6 mg/ml, both of which were incubated at 20 °C.

As DDM is not an ideal detergent for crystallisation, an additional large sale membrane prep

from 24 L of culture was purified in order to buffer exchange into OG using SEC, following

buffer exchange and concentration in DDM during IMAC. The protein was concentrated to a

volume of 500 µl (365 µg) before centrifugation for 10 minutes to pellet any precipitate that

may have blocked the column. Although a monodisperse peak was observed and collected,

it was very small (**Figure 3.17**) and, therefore, an extremely low concentration of protein

was obtained with a 50 kDa MW cut-off concentrator (0.27 mg/ml in 60 µl; 17 µg). In order

to obtain the 10 mg/ml required for crystallisation trials, this purification procedure would

require scaling up by a factor of 30, which was simply not feasible.
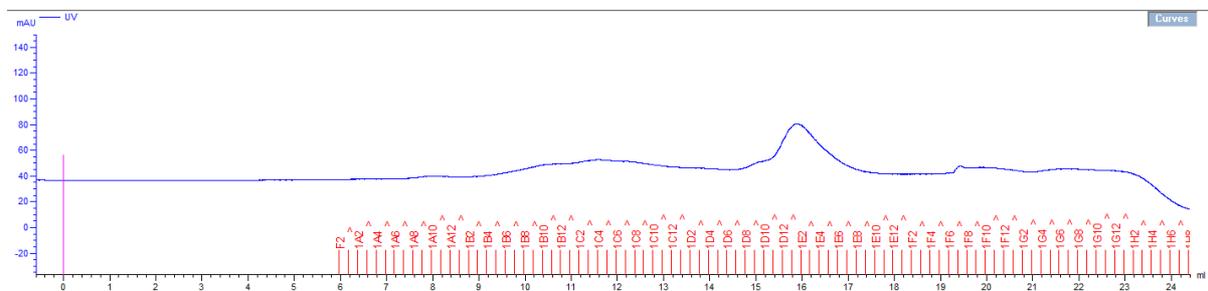
Figure 3.17. SEC UV trace of rMOMP buffer exchanged into 1.17% (w/v) OG having been solubilised in SB3-14 and buffer exchanged into DDM during IMAC purification. Notably, no void volume was observed, with a single monodisperse peak at ~16 ml corresponding to rMOMP. An ÄKTA system was fitted with a Superdex 200 column (GE healthcare, 10/300 GL) and sample loaded using a 500 µl injection loop. A flow rate of 0.4 ml/min was set for 1 CV (25 ml) with 0.2 ml fractions collected in a 96 well microtube plate.

Consequently, an alternative method was adopted. Instead of overnight dialysis to remove excess imidazole, a HiPrep desalting column was selected to both buffer exchange MOMP into OG from SB3-14 after nickel gravity flow purification, as well as to remove excess imidazole and decrease the NaCl concentration by half to limit salt crystal formation during crystallisation screening. The column permitted 15 ml of eluted protein per CV (53 ml) and a fast flow rate allowed buffer exchange to occur in under 30 minutes. From the UV trace shown in **Figure 3.18,** it is evident that the protein is eluted in the first 15 ml, as indicated by the broad peak (E1). The second lower absorbance hump between 15-25 ml (E2) also was collected and concentrated, but due to the speed at which the volume decreased, in addition to a very low protein concentration recorded by Nanodrop, it was concluded that this region did not contain MOMP. **Figure 3.19** shows both the Coomassie blue stained SDS-PAGE gel (**Figure 3.19A**) along with the Western blot (**Figure 3.19B**), which also indicates that the majority of MOMP is eluted in E1 (well 11) as opposed to E2 (well 14).

To maintain a low pressure over the HiPrep column, it was disconnected from the ÄKTA system, during manual loading of the eluted protein via a syringe and, therefore, the 15 ml void volume was not recorded and is not visible in **Figure 3.18**. From the conductivity trace (shown in cyan, **Figure 3.18**) the excess salt and detergent were eluted between 30-50 ml. From a 6 L membrane prep, a final concentration of 0.1 mg/ml (100 µg total) was achieved using a 50 kDa MW cut off concentrator.
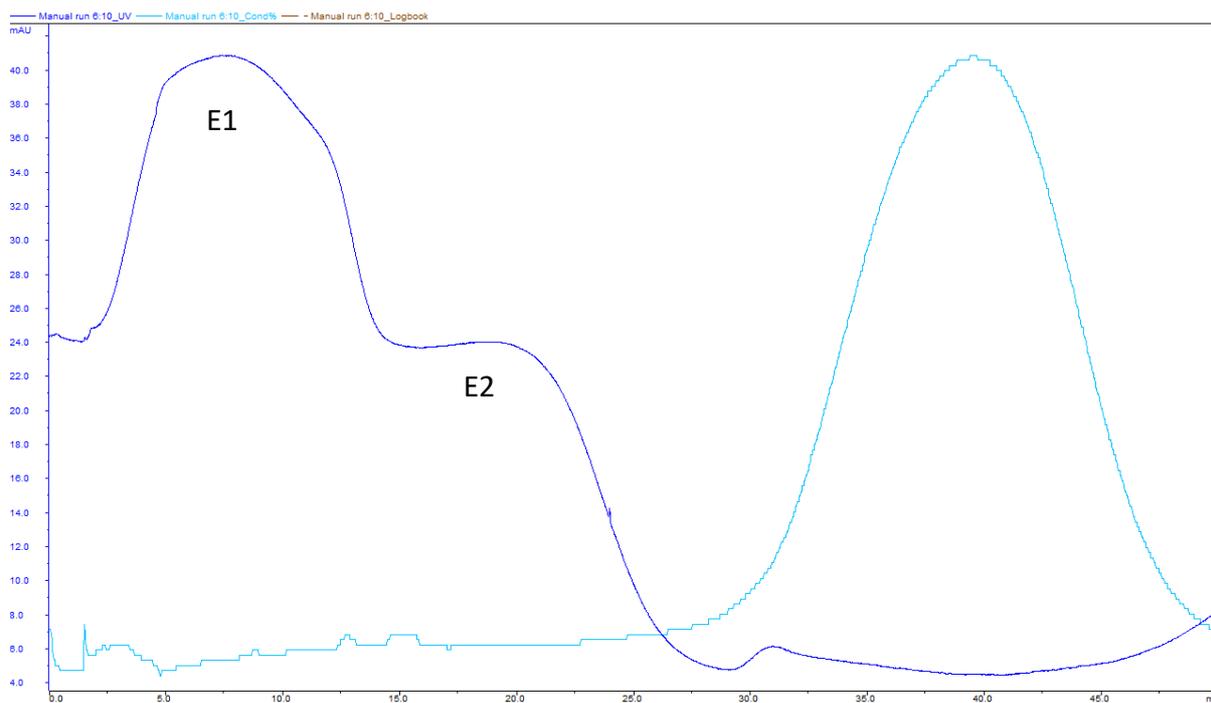
Figure 3.18. Buffer exchange of rMOMP into 1.17% (w/v) OG using a HiPrep desalting column. 15 ml of SB3-14 solubilised rMOMP buffer exchanged into DDM during IMAC purification was manually applied to a HiPrep 26/10 desalting column (GE Healthcare Life Sciences). A flow rate of 2-8 ml/min, was set for 1 CV (53 ml) with the fraction of interest collected in 15 ml falcon tubes. UV trace, dark blue; conductivity trace, cyan. rMOMP membranes were purified from a 6 L membrane prep. Elution 1 (E1) and elution 2 (E2) are indicated.
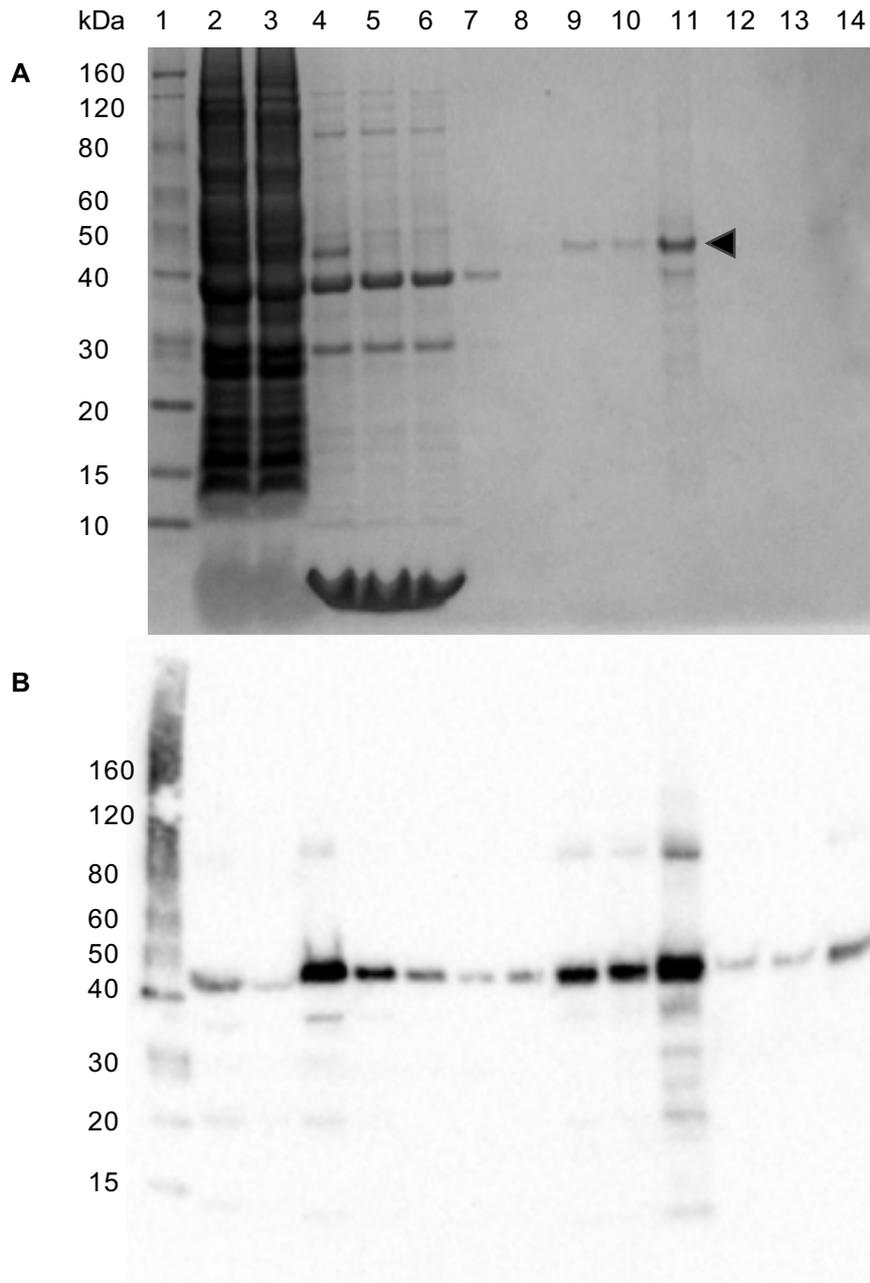
Figure 3.19. Coomassie blue and Western blot of OG buffer exchanged rMOMP. 1, Protein standard; 2, sarkosyl BC; 3, sarksoyl SF; 4, SB3-14 BC; 5, SB3-14 SF; 6, resin flow through; 7, 30 mM wash; 8, 50 mM wash; 9, 200mM elution; 10, HiPrep E1; 11, HiPrep E1 concentrated; 12, HiPrep E1 concentrator flow through; 13, HiPrep E2; 14, HiPrep E2 concentrated. All values in mM refer to the imidazole concentration; BC, before centrifugation; SF, soluble fraction. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L) -HRP conjugated secondary antibody. Arrow head indicates rMOMP.

Due to this success, the same procedure was scaled up to 24 L, and the HiPrep elutions pooled and concentrated. Once the protein was concentrated to 1.37 mg/ml (250 μg), it was left at 4 °C overnight and concentrated again the following day. Despite this, the concentration did not increase as expected and washing of the membrane revealed precipitated protein, likely due to remaining stationary overnight. As a result, the procedure was again repeated, instead storing the eluted fractions at 4 °C overnight and concentrating in three separate concentrators the following morning, both to speed up the process, thus limiting the time the POI was at 4 °C, and to prevent excessive concentration which may have led to aggregation. Despite these additional precautions, only a concentration of 0.3 mg/ml (180 μg) was obtained.

3.4.9 Analysis of DDM purified MOMP crystal screens using the microfocus beamline, I24, at Diamond Light Source

Following incubation of DDM purified MOMP in the two MemGold crystal screens at 20 °C for 6 weeks, a number of potential hits were detected, an example of which is shown in **Figure 3.20**. These samples were taken to the I24 beamline at DLS which offers *in situ* plate screening using a microfocus beam suitable for crystals down to a size of 1.5 μm. Grid scanning of regions of interest produced diffraction patterns in the low-resolution region, a small subset of which appeared ordered and possibly due to protein (**Figure 3.21A**), whilst the majority exhibited disordered spots with smearing and curves (**Figure 3.21B**), likely arising from detergent crystals.
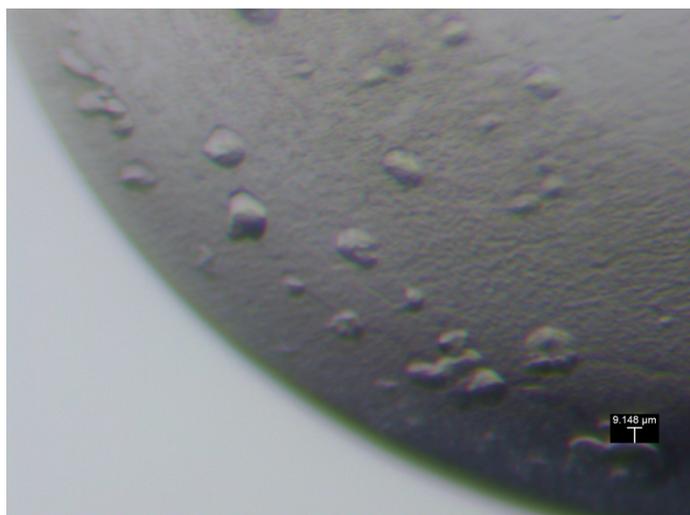
Figure 3.20. Crystals observed in MemGold I after 6 weeks. The reservoir composition of well H1, pictured, was as follows; 0.1 M sodium chloride, 0.1 M sodium phosphate, pH 7.0, 33% (v/v) PEG 300.
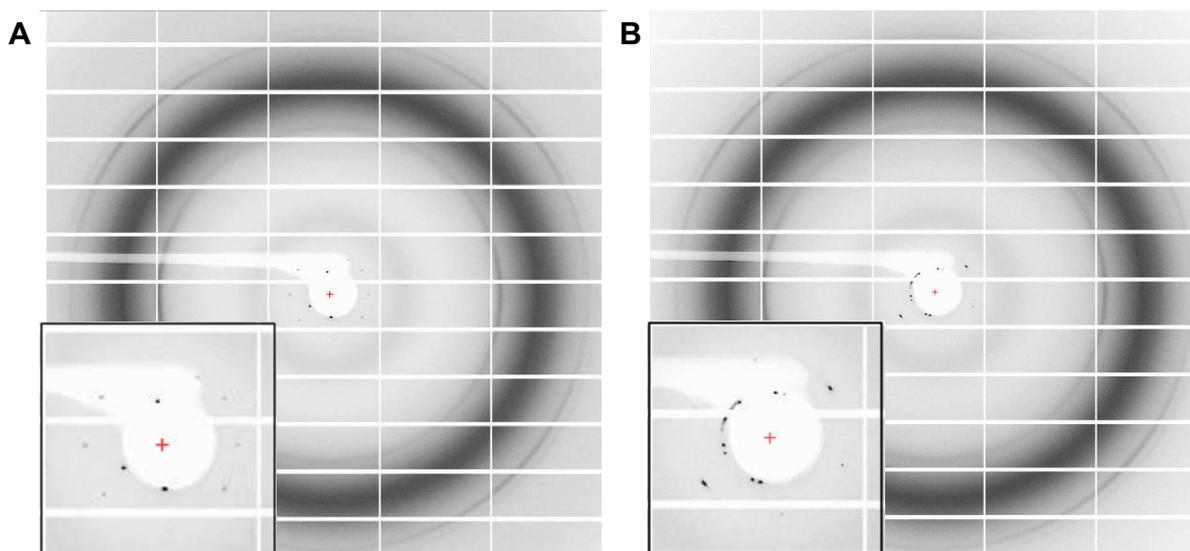
Figure 3.21. Diffraction spots observed from *in situ* plate screening on beamline I24, DLS. A, well ordered diffraction spots, possibly due to protein crystals and **B, disordered and curved diffraction pattern likely due to detergent crystals**. rMOMP was solubilised with SB3-14 and detergent exchanged into DDM during IMAC purification before concentration to 7.6 mg/ml. A number of crystals observed within both MemGold I and II crystallisation screens were grid scanned. **A** is from well D11 in MemGold II (0.02 M lithium chloride, 0.02 M glycine, pH 10, 33% (w/v) PEG 1000) and **B** is from well H1 in MemGold I (0.1 M sodium chloride, 0.1 M sodium phosphate, pH 7.0, 33% (v/v) PEG 300).

3.5   Discussion

3.5.1   Optimisation of *E. coli* growth conditions between DLS and UoR

Previously, 10 mg/ml rMOMP was produced in *E. coli* from approximately 48 L of culture, which was solubilised with DDM and subjected to crystallisation trials with a number of different detergents (Atanu, 2014). As a result of the previous success of this large scale expression and purification procedure, it was replicated in order to duplicate the conditions from which crystal growth was observed and a low resolution structure obtained. This would then permit optimisation of crystallisation conditions in order to produce more robust crystals able to diffract to a higher resolution.

Initially, small-scale expression tests were conducted at UoR in order to confirm MOMP expression with *E. coli* strains C41 (DE3) and BL21 (DE3). Under these conditions, C41 (DE3) cells transformed with MOMP grew to an OD_{600} of 0.5 within three hours of subculturing with a 1 in 100 dilution of overnight culture. However, when these growth conditions were replicated at the collaborating institute, DLS, growth following subculturing was significantly slower, taking an extra four hours to reach an OD_{600} of 0.5. It was initially hypothesised that this effect may have, in part, been due to the culture volumes used, with those at DLS typically being of a larger scale with less aeration and, therefore, overnight cultures were grown in volumes of 10 ml and 150 ml. However, it also was important to consider the types of incubators used between the two institutes. At DLS, small Innova 44 shaker incubators were used, as opposed to the 37 °C incubated room used at UoR, in which the shaker was placed at ground level. Therefore, it was anticipated that the incubator room at UoR was less accurately regulated and in fact independent measurements revealed an actual temperature of 29.5 °C at ground level. Therefore, to investigate this, one of the overnight

cultures with a volume of 150 ml was incubated at 30 °C. Additionally, basal expression of the protein, which is known to be toxic, could have reduced the growth rate. To investigate this, the growth media was supplemented with either 0.5% or 1% (w/v) glucose, in order to encourage catabolite repression through maintenance of low levels of cyclic adenosine monophosphate (cAMP). In glucose rich environments, a by-product of glucose metabolism is able to repress cyclic adenosine, the enzyme responsible for converting ATP to cAMP, resulting in low cAMP expression. However, in the absence of glucose, cAMP is freely produced and able to interact with catabolite activator protein (CAP), which forms a complex capable of binding upstream of the *lac* operon, thus stimulating transcription via the RNA polymerase and resulting in basal protein expression (Novy and Morris, 2001).

While the OD_{600} measurements of the overnight growth cultures were all very similar, the growth curve indicated a dramatic difference in growth, with the culture grown at the lower temperature of 30 °C reaching an inducible level of growth after three hours. Both cultures inoculated with overnight cultures grown at 37 °C grew at a much slower rate, with the 150 ml condition growing the slowest, likely due to reduced aeration. As a result, it was concluded that a reduced overnight culture temperature of 30 ˚C prevented progression into the stationary and/or death phase, instead resulting in rapid entry into the exponential phase upon sub-culturing, which permitted the cells to reach the OD_{600} of 0.5 required for induction in a shorter amount of time. It is noteworthy that the OD_{600} of the three overnight cultures did not appear to be significantly different. As OD_{600} is a measure of cell biomass and not cell number, it is possible that the overnight cultures grown at the higher temperature had either larger cells or a greater ratio of dead:live cells, as opposed to those at the lower temperature. The proximity of the curves in each condition group (A-C, **Figure**

3.4) indicated that glucose supplementation had minimal effect upon the time taken to reach an OD_{600} of 0.5 and, as a result, suggested that basal protein expression was likely non-existent or low enough to not induce a hindrance to cell growth.

Consequently, a reduced temperature of 30 °C was used for all subsequent overnight cultures at DLS, with UoR cultures grown under the same conditions as before. Additionally, at the UoR, cultures of 250-500 ml were grown in 2 L flasks, compared to the previous large scale expression protocol used by Atanu (2014) at DLS, which contained 1 L of culture in the same type of flask. Instead, in order to increase aeration and thus healthy growth, only 500 ml of culture was grown in 1 L flasks for large scale expression at DLS.

3.5.2   Poor solubility of MOMP with detergent DDM

rMOMP was previously solubilised with 2% (w/v) DDM, following the removal of contaminant proteins by sarkosyl treatment, which led to the purification of stable and monodisperse protein suitable for crystallisation trials (Atanu, 2014). As a result of this previous success, this protocol was replicated on a small scale with just 6 L of LB growth media, in comparison to the ~48 L culture volume used in large scale experiments, in order to assess the robustness of the protocol. From purification with a HisTrap column, only a minimal amount of protein was eluted, as indicated by the small peak in the UV trace. During each step of the procedure, small fractions were collected for analysis with Western blotting. From this, it was evident that a small proportion of protein was present in the sarkosyl soluble fraction, which was likely a result of inclusion body formation. However, the majority of rMOMP was insoluble in sarkosyl, as demonstrated by the strong band in the Western blot corresponding to 40 kDa, indicating that MOMP was expressed within the membrane, as desired. Despite the previous success reported with the use of DDM in

MOMP solubilisation, the majority of MOMP unexpectedly appeared to be within the DDM insoluble fraction. It was, therefore, anticipated that the low yield obtained from IMAC purification was due to poor solubility and as such lower than expected application of protein onto the column. It was hypothesised that the low yield was due to the small culture volume of just 6 L, as previous success required 48 L of culture. Therefore, the procedure was again repeated but with 24 L of LB media, in an attempt to increase the total quantity of rMOMP. However, interestingly the total protein yield remained the same as reported in the small-scale 6 L study. One hypothesis investigated was that, due to the size of DDM, the hexahistidine tag was not available to the Ni-NTA column, resulting in low retention and thus low yield. Testing of this hypothesis is addressed below (**Section 3.5.3**). Primarily the low solubility in DDM, as indicated by the Western blot, was deemed to be a major concern due to the high loss of protein. As a result, a number of alternative detergents were screened, in an effort to optimise MOMP solubility and possibly reduce the volume of culture required to obtain suitable quantities of protein appropriate for crystallisation screening.

3.5.3 His-tag extension of the pETMOMP$_{H6}$ construct

Following the low yield obtained from both 6 L and 24 L of culture medium, the hexahistidine tag of the pETMOMP$_{H6}$ construct was extended to His-8 and His-10 in an attempt to improve association with the Ni-NTA resin during IMAC. Following successful creation of the his-tag extensions using SDM, expression was tested in both C41 (DE3) and BL21 (DE3) *E. coli* cells. The Western blot of whole cells expressing each construct indicated that BL21 (DE3) cells with pETMOMP$_{H6}$ construct produced the most rMOMP. It was expected that a stronger band corresponding to MOMP would be observed in BL21 (DE3)

cells, as previous experiments indicated that these cells were capable of much higher expression in comparison to mutant C41 (DE3), although much of this protein was found in inclusion bodies. The MOMP constructs containing the His-6 tag acted to serve as controls for expression, with both His-8 and His-10 tagged constructs resulting in lower levels of MOMP, although pETMOMP$_{H8}$ expressed in BL21 (DE3) cells produced protein at comparable levels to that of pETMOMP$_{H6}$ in C41 (DE3) cells. Notably, the His-10 tagged construct did not express MOMP in either strain. To determine if any additional mutations had been introduced that may have disrupted expression, MOMP was sequenced from the N-terminal direction with the T7 forward sequencing primer, although no additional mutations were detected. It was concluded that the additional two histidine residues likely terminated MOMP expression, as even incorrectly folded protein would result in inclusion bodies. This further highlighted the requirement to analyse whole cells and not just the membrane region, which in this instance indicated that His-10 tagged MOMP was not expressed. Analysis of solely the membrane fraction, whilst useful to determine correct export of MOMP, would not have demonstrated the detrimental effect of His-10 tag extension on expression. Despite the reduced expression of His-8 tagged MOMP in C41 (DE3) cells, if interaction with the Ni-NTA column is improved then this reduced expression may be a suitable compromise for a higher yield of purified protein. Alternatively, His-8 tagged MOMP expressed in BL21 (DE3) also may be a reasonable construct to pursue, due to comparable expression of MOMP$_{H6}$ in C41 (DE3), although further analysis would be required in order to determine if MOMP was exported to the OM.

3.5.4 Detergent screening for optimal MOMP solubilisation using Western blotting

Despite extension of the his-6 tag, it was reasoned that improved solubility would likely confer the greatest improvement in yield due to the large difference observed with DDM soluble and insoluble samples. As a result, membranes were prepared for small scale detergent screening, detected through Western blotting due to both the low expression of MOMP and the presence of *E. coli* contaminant proteins at the same MW. Initially, the sarkosyl isolation step was not included in the initial detergent screens, due to slight variations observed between samples, likely due to disparities in resuspension of the insoluble pellets, which may have distorted the effect of the solubilisation detergent. A range of conditions also was tested, to include two temperatures for different durations. It was hypothesised that solubilisation at room temperature may encourage a more fluid cellular membrane, permitting easier extraction of MOMP, although the duration was limited to just one hour due to the risk of denaturation. Solubilisation at 4 °C was assessed, as a typical temperature for the majority of membrane protein procedures, again to ensure stability and prevent denaturation. In order to examine if the detergent simply required a longer duration, in order to successfully solubilise MOMP, one of the 4 °C experiments was incubated overnight for at least 16 hours. Introducing an overnight step also would prove useful in the overall expression and purification procedure, whereby freshly prepared membranes could be isolated with sarksoyl and then solubilised with detergent overnight, removing the requirement to flash freeze membranes. This may have improved retention of stable and active protein through avoidance of thawing, although contrarily solubilisation for such a long duration may counteract this benefit.

As previously discussed, DDM, OG and LDAO were previously lauded for their success in membrane protein crystallisation. Typically, due to their expense, OG and LDAO are not routinely used in solubilisation but were tested in this instance in the interest of obtaining highly coveted purified MOMP. If successful, the additional detergent exchange step into a crystallisation suitable detergent would also not be required, further ensuring MOMP's stability. CHAPs is a zwitterionic detergent with an unusual structure containing a steroid ring, similar to that found in bile salts, which is capable of disaggregating protein complexes while maintaining the secondary and tertiary structure (Rodi *et al.*, 2014). Thus, CHAPs was also tested in preliminary detergent screens to determine its efficiency at extracting disulphide bonded MOMP clusters. FC-12, also a zwitterionic detergent containing twelve hydrophobic methylene groups, often is viewed negatively in membrane protein research with suggestions that, due to its harsh nature, proteins soluble in FC-12 and SDS but not DDM are likely misfolded (Thomas and Tate, 2014). However, there has recently been a report detailing the success achieved with FC-12 in both solubilisation and crystallisation of *E. coli* OmpF, although at low resolution of between 3.5-5 Å (Kefala *et al.*, 2010). Despite this, due to the limited success of FC-12 solubilisation in membrane proteins, the advice of Thomas and Tate was followed herein, whereby FC-12 solubilised protein was considered most likely misfolded and thus served as a comparison to correctly folded and exported protein soluble in alternative detergents. Similarly, Triton X-100 is a detergent well known for its ability to permeabilise the membrane at low concentrations of just 0.1% (v/v) and, therefore, a higher concentration of 1% was expected to be able to extract MOMP, although again, due to its harsh nature, was expected to result in denaturation.

From initial detergent screening with the described conditions and detergents, only FC-12 and Triton X-100 showed a band corresponding to MOMP in the soluble fraction using Western blotting analysis and only with screening conducted at room temperature. Furthermore, within this blot were additional bands at 80 kDa, likely due to dimeric MOMP, and also at lower MWs of 35 kDa and 20 kDa, which were likely due to degradation products. Due to the presence of both aggregates and degradation products, it was likely that the solubilised protein observed was misfolded protein, which possibly begun to denature under the warmer room temperature conditions. The weak band (at ~ 40 kDa) also indicated that only a small amount of MOMP exported to the OM was misfolded.

It was hypothesised that, due to the large number of contaminant proteins, the milder detergents such as OG and LDAO were solubilising other, more easily accessible proteins and as such resulted in no solubilisation of MOMP under any of the conditions tried. As a result, the sarkosyl isolation step was reintroduced, instead with one batch prepared and then aliquoted in order to limit variability with pellet resuspension in individual samples. Firstly though, the sarkosyl concentration was tested in order to optimise MOMP isolation, with concentrations between 2-10% (v/v). This again was analysed using Western blotting, where it was determined that all concentrations were successful at isolating MOMP and therefore the lowest concentration of 2% (v/v) was used to prevent unnecessarily harsh treatment, as well as to limit residual sarkosyl, as visualised through smears in the samples with higher concentrations.

A number of additional conditions were tested, following sarkosyl isolation, including; changes in pH to raise it above MOMP's pI of 8.29, increase in temperature to 30 °C to increase membrane fluidity, increase in sarksoyl isolation to two hours at room temperature

to remove more contaminant proteins, and testing of an additional zwitterionic detergent, namely, SB3-12. However, again only FC-12 was capable of solubilising MOMP. Although reports suggested that proteins only solubilised by FC-12 were most likely misfolded (Thomas and Tate, 2014), the soluble band visible in Western blotting was extremely weak and did not account for the observed high ratio of MOMP in comparison to the band visible before centrifugation. This suggested that much of MOMP remained inaccessible to the detergents within the membrane and thus a higher throughput technique, allowing an analysis of a greater number of detergents, was required.

3.5.5 Detergent screening using the novel P3NTA tag with FSEC

For a higher throughput method of detergent screening, FSEC was adopted. As previously discussed, MOMP previously tagged with GFP was incorrectly folded and expressed as inclusion bodies. To overcome the requirement for a fluorescence tag, the novel P3NTA probe was used, which binds to the his-tag under the same principle as IMAC. Being just 3 kDa and supplemented after solubilisation, it was unlikely that any adverse effects would be observed on MOMP's structure or stability. As a high throughput technique, a number of additional detergents could be tested, which along with DDM, OG, LDAO, and SB3-12 as tested before with Western blotting, included SB3-14, Elugent, LMNG, OGNG, UDM, CYMAL5, and C12E8.

Despite DDM being the detergent of choice for MOMP solubilisation which successfully led to crystals (Atanu, 2014), the FSEC trace for this detergent revealed an unusual initial peak, possibly due to a high number of aggregates, with no void volume and a small peak at 20ml for unbound probe. On the contrary, all other detergents examined exhibited a peak around 9 ml, correlating to aggregated protein which was too large to infiltrate the pores of the

column and thus was eluted first, commonly referred to as the void volume. Notably, the void volume for low cost detergent Elugent was particularly large, possibly due to a combination of the higher concentration of 5% and the non-ionic detergent mix of octyl-glucosides and Triton X-100, and thus lower purity. Elugent was used at a higher concentration due to successful reports of its use between 1-5% for membrane protein solubilisation (Noinaj *et al*., 2013; Norell *et al*., 2014; Zabara *et al*., 2018).

Additionally, each trace displayed a large peak at 20 ml, correlating to the 3 kDa unbound P3NTA probe. Due to the novel nature of this probe, especially in use with membrane proteins, the optimal concentration was unknown, although it became clear from the size of the peak that the concentration could be comfortably reduced from 1.5 µm in future.

Most importantly from the trace, it was evident that whilst the peaks between the void volume and the unbound probe were all small, there was a distinctive peak at 17.5 ml correlating to SB3-14, a detergent previously used in the successful solubilisation of *C. jejuni* MOMP, as well as both native *C. trachomatis* MOMP (Tifrea *et al.,* 2011) and rMOMP expressed in *E. coli* (Findlay *et al*., 2005). MOMP of *C. jejuni* shares both a name and a β-barrel structure with *C. pneumoniae* MOMP, although exists as a trimeric protein lacking any cysteine residues with no regions homologous with MOMP. Despite this, the use of SB3-14 to solubilise this *Campylobacter* β-barrel in *E. coli* has also proven useful in the solubilisation of *Chlamydial* MOMP. SEC columns are typically calibrated with standard proteins and Stokes radii to help identify the location of an uncharacterised protein (Nozaki *et al*., 1976). However, these calibrations are based on soluble proteins which are typically globular in shape with similar partial specific volumes, and are therefore at best an estimate for membrane proteins due to the accompanying detergent micelle (Slotboom *et al*., 2008).

The peak at 17.5 ml was deemed to result from a 36 kDa protein from previously calibrated data using the P3NTA tag, however membrane proteins analysed with this specific system were reported to elute with a 10 kDa difference. Therefore, MOMP being a 40 kDa protein, raised to 43 kDa with the addition of the fluorescent tag is likely the cause of the peak at 36 kDa, although confirmation was required through gel electrophoresis.

There were also a number of additional peaks of very small sizes, possibly due to MOMP degradation products or contaminant histidine containing proteins, which would be in higher quantities due to the omission of the sarkosyl isolation step. The most common peak other than that observed at 17.5 ml was that at 15 ml, correlating to a protein of 95 kDa and therefore possibly trimeric MOMP, although at extremely low levels.

Whilst SB3-14 clearly exhibited the strongest peak correlating to MOMP, just prior to this was a broad peak spanning 5 ml, likely due to a large aggregate of disulphide bonded MOMP. Whilst care would be required to eliminate this aggregate in such close proximity to the peak of interest in future, SB3-14 was undoubtedly the most promising detergent for MOMP solubilisation. As the P3NTA tag was a novel probe, a fluorescent SDS-PAGE gel was compared to a Western blot of the same samples in order to draw comparisons with the blotting method adopted during the early stages of solubility screening. From the P3NTA labelled gel, the most visible band was one at 63 kDa, which previously has been reported as the size of trimeric MOMP (Kari *et al*., 2009), although the reasoning behind this is unclear, especially due to MOMP's monomeric MW of 40 kDa, although OMPs PhoE and OmpF of *E. coli*, both with monomeric MWs of 39 kDa, have also been shown to migrate to a similar MW as trimers (Sun *et al*., 2007). It was certainly possible that this 63 kDa band was a result of non-specific binding of P3NTA to a native *E. coli* protein, such as 66.8 kDa protein GlmS

(Bolanos-Garcia and Davies, 2006) which contains almost 4% histidine residues, which would likely become exposed upon denaturation with SDS-containing loading buffer. Despite the presence of this 63 kDa band within every sample in fluorescent SDS-PAGE analysis, such a peak common to all detergents was not visible within the FSEC traces. This suggested that the histidine residues binding to the P3NTA tag only became exposed upon denaturation for gel electrophoresis and thus were not detectable during the FSEC experiment. Interestingly, these 63 kDa bands were not visible with Western blotting, which suggested that the P3NTA probe, which only needs to bind to just two histidine residues, was far less specific than the anti-his$_6$ antibody.

Below the 63 kDa band were additional fainter bands within each sample correlating to MOMP at the correct MW size, which was only visible in one well with Western blotting, which correlated to SB3-14. This strongly indicates that the peak of interest identified at 17.5 ml determined to be due to monomeric MOMP was in fact correct. Most importantly, no other bands corresponding to monomeric MOMP were visible within the Western blot, and therefore the solubilising effect of SB3-14 was clearly distinguishable from the remaining detergents, as opposed to in the fluorescent SDS-PAGE gel.

As a result, whilst FSEC with the novel P3NTA tag proved extremely useful in screening a large number of detergents, validation of results through Western blotting was preferred over fluorescence gel electrophoresis.

3.5.6 MOMP solubilisation with SB3-14 assessed with Western blotting

After SB3-14 was established as the most promising detergent for rMOMP solubilisation, Western blotting was again used to determine the optimum concentration. In FSEC experiments, solubilisation with 1% (w/v) SB3-14 resulted in the most significant peak

relating to MOMP, which was later confirmed with Western blotting. Again, a 1% concentration was tested, in addition to 2%, along with two novel cysteine mutants C201A and C201/203A as well as native MOMP. The broad aggregation 'peak' demonstrated in the SB3-14 FSEC trace was hypothesised to result from intermolecular disulphide bonding within MOMP, which is discussed and investigated in greater depth in **Chapter 5**. Briefly, *Chlamydia* uses a network of cysteine rich proteins, connected through intermolecular disulphide bonds, to provide circulating EBs with structural stability, of which MOMP is a dominant protein. It was hypothesised that mutation of particular cysteine residues believed to play major roles in such interactions would prevent MOMP clusters from forming and thus improve MOMP solubility. However, supplementation with 1% SB3-14 had no effect on solubility in both mutants, as opposed to a concentration of 2% which resulted in an inexplicable shift in the size of the band corresponding to MOMP up to approximately 50 kDa, in comparison to the expected 40 kDa observed in the remaining samples. As a result of this peculiar shift, the lack of soluble protein observed for any of the mutants, and the good solubility of native MOMP with 1% SB3-14, only the latter was taken forwards into large scale purification optimisation. Solubilisations were carried out at 4 °C for two hours, due to the increase in denaturation and aggregation observed at room temperature.

3.5.7   Optimisation of MOMP purification

After successful optimisation of MOMP solubilisation, a small scale purification following the his-trap affinity chromatography procedure was again followed. A lower concentration of 0.1% (w/v) SB3-14 was used in the two buffers in order to reduce excess detergent not required for micelle formation, with the CMC of SB3-14 being 0.007%. However, an artefact of SB3-14 not previously realised was its eventual precipitation upon storage at 4 °C. Due to

the fragility of MOMP, previously observed during room temperature experiments, all buffers required for purification were chilled prior to use. However, after successful loading of the his-trap column using a peristaltic pump, the buffers containing SB3-14 began to precipitate within the thin tubing of the ÄKTA system, which was over three times narrower in diameter compared to that used in sample loading. Despite warming the buffer to decrease its viscosity, once drawn inside the tubing within the 4 °C cabinet, the viscosity returned, in turn increasing the pressure within the ÄKTA system. The buffers were exchanged for those containing just 0.01% SB3-14 in an attempt to alleviate this effect, but to no avail. Whilst the flow rate was reduced, the continued increase in viscosity resulted in an eventual run time of over 12 hours, by which time MOMP may have become degraded or pressure damage occurred to the ÄKTA system or column. Instead, the his-trap column was manually washed using the peristaltic pump and eluted in a single step with 250 mM imidazole. However, as this method was not optimised in terms of the imidazole wash steps and elution, a number of contaminant proteins and potential MOMP degradation products were observed upon analysis of the eluted fraction with an SDS-PAGE gel.

Due to the negative effect of SB3-14 on pressure within the ÄKTA system, free nickel resin was used for future purifications, as opposed to a his-trap column. Whilst it is possible to purify MOMP, as before, using a his-trap column and peristaltic pump, as the pressure is not measured there is a high risk of damage to the column. This alternative method required optimisation at each step, starting with the ratio of resin to protein. Whilst a large quantity of resin may in the first instance appear beneficial for optimum isolation of the POI, this actually can increase the non-specific binding of contaminant proteins and decrease the overall purity. For MOMP, 1 ml of resin per 1 L of original culture was used, with the resin

capable of binding up to 50 mg/ml. Another vital step in prevention of contaminant binding was in the addition of a low concentration of imidazole during the resin binding step. As discussed previously, the imidazole rings found in both histidine and imidazole competitively bind to the Ni^{2+} ion, and therefore weakly interacting contaminant proteins can be replaced by low levels of imidazole. Concentrations of imidazole between 10 mM and 90 mM were tested and analysed with both Coomassie stained SDS-PAGE gels and Western blotting, whereby from the former it was evident that no imidazole supplementation resulted in retention of a multitude of contaminant proteins. Use of 10 mM imidazole decreased the contaminant binding, visualised with Coomassie blue staining, although a number of faint bands remained visible. With Coomassie staining, the 30 mM imidazole supplementation appeared to result in no protein within the eluted fraction, however cross validation with the Western blot indicated a band corresponding to MOMP. Due to the presence of a small number of contaminant proteins still present with 10 mM imidazole on the SDS-PAGE gel, but the weaker band corresponding to MOMP with 30 mM imidazole on the Western blot, a concentration of 20 mM imidazole was used during nickel resin binding. Solubilised protein was incubated with nickel resin with gentle agitation for a duration of one hour at 4 °C.

A small scale purification following this protocol analysed with Western blotting revealed the absence of a band corresponding to MOMP within the nickel resin flow through fraction. This suggested that all available MOMP was successfully isolated on the resin and that both the duration of binding was sufficient and that the supplementation of 20 mM imidazole did not compete with MOMP-nickel interactions. Two wash steps were also introduced in order to remove any contaminant proteins prior to elution. Concentrations of 40 mM and 60 mM

imidazole were used, whereby it was evident from the Western blot that MOMP was also removed but most noticeably with the higher concentration. The 40 mM wash also appeared to remove the majority if not all of the contaminant proteins, as visualised with Coomassie blue staining, which suggested that only a low concentration of imidazole was required to observe the desired effect. Due to the detrimental effect of contaminant proteins in crystallisation, a trade-off between loss of MOMP and removal of contaminants was essential. In an additional purification, wash steps of 30 mM and 50 mM imidazole were also tested, whereby only a very faint band corresponding to MOMP was visible with Western blotting of the higher concentration fraction and thus as a result, these two wash steps were included in future large scale purifications. In initial purification attempts, three elution steps of 100 mM, 200 mM and 300 mM imidazole were assessed with three fractions taken of each, whereby it was apparent that the majority of MOMP was eluted within the 100 mM fractions, and the first two fractions of the 200 mM step. At this stage, an additional band was visible at a higher MW of approximately 80 kDa, corresponding to dimeric MOMP. It was likely that dimeric MOMP was present during the earlier stages of the purification, however due to its low proportion to monomeric MOMP, only became visible at higher concentrations during the elution steps. As much of MOMP was eluted at the lower concentration of 100 mM, confirmed through an additional small scale purification, a single step elution of 200 mM was used, whereby the higher concentration would permit elution in a smaller volume and thus decrease both the volume of dialysis buffer required and the volume of protein to be concentrated. It was important to remove imidazole through either dialysis or SEC due to its detrimental effect on crystallisation, although in some cases imidazole can act to stabilise particular proteins. Additionally, due to the

aromatic ring, the presence of imidazole is capable of distorting the true protein concentration measured using spectroscopy at 280 nm.

From the Western blot of all steps, including the sarkosyl isolation and SB3-14 solubilisation, it was evident that only a small quantity of MOMP was lost to the sarkosyl soluble fraction, which was likely to be inclusion bodies and thus undesirable for crystallisation. SB3-14 exhibited good solubility, with much of MOMP retained although as a large smear between 40-60 kDa, a likely artefact of the detergent. Samples taken for Western blotting were also stored at -20 °C and not flash frozen, and therefore were subjected to freeze-thaw degradation and aggregation. The additional lower MW bands were most probably degradation products, again emphasising the requirement to both limit the duration of the purification procedure and to conduct all steps at a low temperature. No MOMP was detectable in the resin flow through, again signifying efficient binding, or in the first 30 mM imidazole wash step, with trace amounts lost during the higher concentration wash. As discussed, a single step elution containing 200 mM imidazole was used for future purifications.

Due to MOMP's low expression and toxicity, it was essential to optimise each stage of the purification procedure in order to limit protein loss and increase overall yield. In this case, due to the low expression, Western blotting proved invaluable in assessing each stage of the procedure, although Coomassie blue staining was also essential in determining the presence and quantity of contaminant proteins.

3.5.8 Buffer exchange of SB3-14 solubilised MOMP into a crystallisation suitable detergent

SB3-14, whilst the most useful detergent for extraction and solubilisation of MOMP, has not demonstrated previous success in crystallisation with many researchers instead opting to buffer exchange into more amenable detergents such as DDM and OG (Beis *et al*., 2006). This is typically undergone using SEC, whereby not only can detergent exchange occur, but the salt concentration can be reduced and the sample assessed for its monodispersity. Following elution from the nickel resin, fractions were dialysed in order to both remove imidazole, which may have detrimental effects upon MOMP's stability, and to decrease the SB3-14 concentration down to 1.5 times CMC (0.01%) in order to limit the presence of excess detergent micelles whilst also maintaining MOMP's solubility. In selecting a suitable concentrator, the micellar size of the detergent was an important consideration as it envelopes the POI thus contributing to the overall MW. SB3-14 has a micellar size of 30 kDa, which combined with MOMP was a total size of 70 kDa and as a result a centrifugal concentrator of 50 kDa was selected, whereby it was intended that concentrated MOMP would then be buffer exchanged into a crystallisation suitable detergent such as OG using SEC. However, unexpectedly, excess detergent micelles were concentrated alongside purified protein, the increase in concentration of which resulted in denaturation of MOMP. Despite the presence of the 50 kDa filter, it was likely that gradual build up detergent blocked the filter and permitted concentration of excess detergent alongside MOMP. In order to overcome this, the SB3-14 concentration required reducing but was already at a very low CMC of just 1.5 times, whereas it is more typical to use three times CMC to ensure protein stability. An alternative method was adopted, whereby MOMP was detergent exchanged into OG during the IMAC purification procedure. The OG concentration was

gradually increased to a two times CMC of 1.17% (w/v), in 0.6% and 0.9% steps, whilst concurrently reducing the SB3-14 concentration. However, with a micellar size of 25kDa, a smaller MW cut off concentrator of 30 kDa was required for concentration of MOMP buffer exchanged into OG, as much of MOMP became trapped within the filter of a 50 kDa MW cut off concentrator used in a small scale test run. However, again precipitation occurred after the protein concentration reached 1 mg/ml, which initially was believed to be due to rapid concentration. The centrifugation duration was reduced by half, to just five minutes, in order to permit regular resuspension and prevent localised increases in protein concentration. However, it was then hypothesised that due to the size of the concentrator, it was likely that both OG and excess SB3-14 were also becoming concentrated. Despite buffer exchange into OG, it was highly likely that some SB3-14 still remained, and potentially also traces of sarkosyl used during membrane isolation, all contributing to an uncharacterised detergent 'soup'. To limit the concentration of detergent and the loss of protein within the higher MW membrane, a two-step concentration procedure was devised whereby a 30 kDa cut off was used before switching to a filter of 50 kDa. This was in attempt to prevent protein loss in the first instance, with exchange into the higher MW cut off to remove concentrated detergent. However, whilst the protein stability was improved, the total yield was poorer than anticipated so again an alternative method was adopted. MOMP was again buffer exchanged during gravity flow IMAC but instead with DDM, a detergent with a higher MW of 70 kDa, resulting in a total MW of 110 kDa. This then permitted the use of a 100 kDa centrifugal filter, which would remove excess DDM as well as low MW SB3-14 and any residual sarkosyl. Due to the lower cost of DDM compared with other high purity detergents, a concentration of three times CMC was used in the final elution with concentration resulting in a theoretical increase in yield by 35%, compared to

that obtained from detergent exchange with OG with two-step concentration. From a scaled up 18 L membrane preparation, 7.6 mg/ml of protein was obtained following this procedure, equating to ~460 µg which was enough for two crystallisation trials in MemGold I and II. Whilst successes have been observed in crystallisation with DDM, previous attempts by Atanu in 2014 produced small and poorly diffracting crystals. Additionally, due to the difficulty in producing stable, purified MOMP, the sample was not analysed with SEC for monodispersity at the risk of losing sample. Instead, the procedure was again repeated with a 24 L membrane preparation, whereby protein buffer exchanged into DDM was concentrated to a volume of 500 µl for injection onto a SEC column. Given the volume, the yield of protein (365 µg) was as expected, however disappointingly only a small monodisperse peak was observed, which upon collection and concentration resulted in minimal protein. It was uncertain as to where such drastic protein loss occurred, as no other peaks corresponding to aggregated protein were observed. In order to obtain enough protein for crystallisation trials following this method, a culture volume of over 700 ml was required, which was simply unfeasible.

As a result, an alternative method was sought, whereby a HiPrep desalting column was used to both remove excess imidazole in place of dialysis, whilst also buffer exchanging MOMP into OG from SB3-14, in addition to reducing the salt content. Not only would this method limit the time MOMP spent at 4 °C, but would also remove excess SB3-14 and OG, preventing denaturation during the concentration procedure. Initially, the system required optimisation, whereby a multitude of fractions were collected in order to determine the location of MOMP within the trace. What appeared to be a baseline, soon dropped away to form a large broad flat topped 'peak', which was later discovered to contain MOMP. As 15

ml of eluted sample from the IMAC procedure could be applied, in the absence of a large enough loop, sample was injected directly onto the column using a syringe. Due to the pressure within the ÄKTA system, the tubing connected to the UV reader was disconnected whilst sample was applied. With a void volume of 15 ml, the sample was eluted almost immediately after reconnection of the tubes and commencement of flow. After determination of MOMP's location with Coomassie blue staining and Western blotting, and successful concentration to 0.1 mg/ml in 1 ml from a 6 L prep, the procedure was repeated to a larger scale which involved multiple sample applications of 10-15 ml from the 50 ml (2 CV) elution step. By increasing the culture volume four fold from 6 L to 24 L, it was anticipated that a concentration of 0.4 mg/ml in 1 ml would be achieved, which could then be concentrated down to 10 mg/ml for a number of crystallisation trials, which require 12 µl of sample per screen. Due to the fast flow rate of the HiPrep column, the eluted fractions were pooled and concentrated immediately in order to limit the duration at which the protein was left at 4 °C. Promisingly, concentration to a 1 ml volume revealed a concentration of 0.3 mg/ml, which whilst lower than calculated theoretically, was a marked improvement from previous attempts. However, the large volume eluted, approximately 50 ml, hugely increased the time taken to concentrate and at 1.37 mg/ml (~250 µg) MOMP was stored at 4 °C for eight hours before re-commencement of concentration. Due to the high concentration, precipitation occurred overnight as indicated by the lack of increase in concentration and the presence of precipitant upon washing of the membrane. Not only this, but it was also observed that the concentration of protein was not increasing as expected, with a higher total quantity of 300 µg measured at 1 ml compared to 250 µg measured following concentration down to ~200 µl. This suggested that loss of protein was not solely due to storage at 4 °C. The procedure was again repeated, instead with

concentration the following day after storage of the eluted product at 4 °C. The large volume of sample was split into three concentrators in order to limit both the time required to concentrate (from five hours in the previous preparation) and to prevent the potential aggregation from highly concentrated samples. However, despite these precautions, the final yield was too low for crystallisation screening and it was hypothesised that with the introduction of two extra concentrators, more protein was lost within the membrane.

Additionally, the HiPrep column is marketed as being high throughput, with flow rates of 15 ml/min achievable. In initial experimentation of this column with a 6 L prep, the flow rate was limited to 2 ml/min by the back pressure on that particular ÄKTA. However, two large scale experiments were able to be run at a flow rate of 8 ml/min, due to the use of a different ÄKTA system where back pressure was no longer an issue. Therefore, it was also hypothesised that the fast flow rate limited the success of detergent exchange into OG, resulting in unstable protein that did not concentrate as well as expected, and as a result the flow rate should be restricted to 2 ml/min in future.

Despite the complications associated with concentration, this method offers a novel and higher throughput method for MOMP purification. Further optimisation is required in the vital concentration step, with the addition of extra TCEP following overnight incubation at 4 °C likely to prevent aggregation arising through intermolecular disulphide bonding. This would permit lengthy concentration in a single concentrator as opposed to three, whereby it was likely that MOMP became trapped within the membrane.

3.5.9 Analysis of DDM purified MOMP crystal screens on microfocus beamline, I24

MOMP buffer exchanged into DDM during IMAC purification resulted in a concentration of 7.6 mg/ml. Due to the difficulties associated with obtaining stable and purified MOMP, this was put into two crystallisation screens, MemGold I and II, using the sitting drop vapour diffusion method and incubated at 20 °C. After incubation for six weeks, a number of crystalline structures were detected in both screens. In initial screening, it was imperative to consider the future analysis of any hits, and therefore plates compatible with the I24 microfocus beamline at DLS were used. Following identification of a number of hits, crystals were analysed with *in situ* plate screening, whereby areas of interest were grid scanned. The purpose of this was to determine the presence of proteinaceous material, as it is certainly not uncommon for salt crystallisation to occur. Additionally, as this sample was not analysed with SEC, excess buffer was not removed except through concentration with a higher MW cut off and therefore the formation of detergent crystals was also possible. All spots present within the diffraction patterns were in the low resolution region, although no high resolution spots corresponding to ice rings or salt were detectable. Whilst a small subset of data was indicative of a regular crystalline lattice, the majority exhibited 'smeared' or double spots, likely arising due to detergent crystals.

3.6 Conclusion

Currently, there are no known structures for any *Chlamydial* OM proteins, with very little definitive information regarding their function, aside from the established proteins of the COMC (Liu *et al*., 2010), porin PorB (Kubo and Stephens, 2002), and adhesin, OmcB (Mölleken *et al*., 2008). Many of the *Chlamydial* OMPs have hypothetical roles, based on their homology to other Gram negative membrane proteins, with Omp85 suspected to

mediate porin translocation (Jacob-Dubuisson *et al*., 2009) and the many Pmps alleged to have various roles as autotransporters and invasins (Mölleken *et al*., 2013). While the majority of the *Chlamydial* OMPs contribute to the COMC, it is interesting to note that a number of these proteins are also suggested to have additional roles in *Chlamydial* biology. This, along with knowledge on *Chlamydia*'s small genome size, indicates that a number of OMPs have multiple functions, again supportive of both MOMP's role within the COMC and as a potential transporter.

MOMP was previously solubilised in DDM prior to buffer exchange into OG and LDAO, both of which produced diffraction quality crystals, with the former resulting in a low resolution (4 Å) structure solved by molecular replacement (Atanu, 2014). As a starting point for this work, the same protocol was followed in order to optimise the crystallisation process, with the main objective of obtaining a higher resolution crystal structure, which would provide novel insight into not only MOMP's structure but also its potential function, whether that be solely in contributing to formation of the COMC or, additionally, in ligand transport.

However, replication of the protocol revealed poor solubility with detergent DDM, with the majority of rMOMP lost to the insoluble fraction. Purification attempts at both small and large scale, surprisingly, did not result in any difference in yield, which was strongly suggestive of suboptimal solubilisation. As MOMP is a major contributor to the COMC, forming intermolecular disulphide bonds both with other MOMPs and other cysteine rich OMPs (Liu *et al*., 2010; Newhall *et al*., 1983), it was speculated that the improved expression of MOMP, achieved simply through increased aeration, resulted in increased clustering within the *E. coli* membrane. This, in turn, reduced the effect of the mild detergent DDM to efficiently extract MOMP from the membrane, despite the inclusion of reducing agents such

as DTT and TCEP. Cysteine mutants of MOMP designed to decrease this clustering behaviour, discussed in greater depth in **Chapter 5**, exhibited unusual migration with Western blotting and thus their solubilisation and purification was not pursued at this time.

Instead, the solubilisation of native MOMP was optimised. Previous studies have utilised a range of different detergents including OG, LDAO, SB3-14, and CHAPs (Findlay *et al*., 2005; Pal *et al*., 2005; Tifrea *et al.,* 2011) for MOMP solubilisation, and therefore a range of detergents were examined. In the first instance, due to low expression, MOMP solubility was analysed through a series of small scale screens using Western blotting. However, introduction of the novel P3NTA fluorescent probe, which was designed to bind to any his-tagged proteins, permitted a higher throughput method known as FSEC, which also was able to readily identify monomeric MOMP. Whilst the P3NTA tag proved useful in FSEC experiments, identifying SB3-14 as the most promising detergent, its reliability in SDS-PAGE gels was somewhat debateable due to the non-specific binding of the 63 kDa protein that was not observed in Western blotting using the anti-his antibody. Fluorescent gels also did not identify a discernible difference between the detergents; all appeared to solubilise MOMP protein, although both FSEC and Western blotting clearly demonstrated SB3-14 to be the more promising lead for the solubilisation of monomeric MOMP, in support of previous studies which used SB3-14 during MOMP purification (Findlay *et al*., 2005; Tifrea *et al*., 2011). Therefore, it was still important to use Western blotting not only to correctly identify MOMP but also to distinguish the most effective detergent in solubilisation experiments. Nevertheless, FSEC with the P3NTA tag was an extremely useful combination for screening a large number of detergents in a high throughput manner.

The concentration of SB3-14 and the conditions under which solubilisation of MOMP was to occur were optimised, however, it was not possible to follow MOMP's elution profile using an ÄKTA system, due to precipitation of SB3-14 at low temperatures. Instead, MOMP was purified with free Ni-NTA resin in a gravity flow column, in which each stage of the binding, washing, and elution steps were optimised for maximum yield. A number of complications arose upon concentration of MOMP, with increases in free detergent micelles denaturing MOMP and causing it to precipitate out of solution. Instead, MOMP was buffer exchanged into DDM during the IMAC procedure. In the presence of this detergent, MOMP was successfully concentrated to 7.6 mg/ml and used in sitting drop vapour diffusion crystallisation screens. Following a period of incubation, a number of crystalline structures were observed and scanned on the microfocus beamline on I24, DLS using in-plate screening. Many of the resulting diffraction patterns of low resolution were attributed to detergent crystals, although a small selection exhibited regular patterns of reflections suggestive of protein diffraction.

Due to the difficulties in producing pure and stable MOMP, MOMP buffer exchanged and concentrated in DDM was not assessed for monodispersity with SEC for fear of losing the protein through aggregation. In order to buffer exchange MOMP into a more crystallisation suitable detergent, such as OG, the protein eluted from the nickel resin (still solubilised with SB3-14) was applied to a desalting column. Initially, in small scale trials, this new method appeared to hold significant promise, although it did not yield as much protein as expected once scaled up, again with difficulties observed during the concentration process.

Presented here is a novel method for MOMP purification, in which each stage has been rigorously optimised in order to increase the total protein purity and yield. Due to MOMP's toxic expression and high cysteine content, each unit expressed within the *E. coli* membrane

is precious. Comparable levels of protein are now obtainable with half the culture volume previously required (Atanu, 2014), following the protocols described in **Chapter 2 Sections 2.2.2.4** to **2.2.2.7**, and final optimisation of the concentration procedure will certainly and reliably lead to stable and pure MOMP for crystallisation optimisation.

Chapter 4

Investigating the functional role of MOMP as a fatty acid transporter

4 Investigating the functional role of MOMP as a fatty acid transporter

4.1 Introduction

4.1.1 MOMP as a potential fatty acid transporter

Due to the paucity of high quality structural data available for *Chlamydial* MOMPs, a functional role for this protein has been difficult to infer. Due to its high β-strand content of between 38-44% (Sun *et al*., 2007) and modelled amphipathic barrel shape, (as discussed in **Chapter 1, Section 1.6.4**), it has been suggested that MOMP exhibits classic porin activity, typical of this OMP bacterial class. This has been demonstrated through liposomal swelling assays, whereby rMOMP from *C. trachomatis* was able to transport all twenty of the amino acids (Jones *et al*., 2000), and nMOMP was able to transport the sugar L-arabinose, although at a considerably reduced rate in comparison to the well characterised porin OmpF of *E. coli* (Sun *et al*., 2007). However, porins (also known as non-specific channels) are not capable of transporting large solutes over 600 Da, and therefore transporter proteins are required in order to supply the bacterial cell with additional nutrients and carbon sources. Additionally, organisation of the residues within the *E. coli* OmpF pore, detailed in the crystallographic structure (PDBid: 2ZFG), creates a highly ordered water filled area, which makes passage of hydrophobic molecules, such as fatty acids, energetically unfavourable (Schulz, 2002; van den Berg, 2010). As fatty acids are principle sources of metabolic energy, as well as being vital components of the cellular membrane, their transport into the cell is essential and occurs via OM transporter proteins, such as FadL in *E. coli*, TodX in *Pseudomonas putida,* and TbuX in *Ralstonia pickettii* (Black *et al*., 1987; Hearn *et al*., 2008).

Since *Chlamydia* have a small genome (1 million bp), it is imperative that they scavenge the host cell for essential nutrients that cannot be synthesised independently. *Chlamydia* possesses a *de novo* biosynthetic pathway for the synthesis of its membrane phospholipids,

which involves the uptake of the host building blocks serine, glucose, and isoleucine (Yao, Dodson *et al.*, 2015). However, lipidomic analysis of *C. trachomatis* cells indicated that 8% of the OM phosphatidylethanolamine consisted of fatty acids 18:1/15:0. Furthermore, *C. trachomatis* is unable to synthesise unsaturated fatty acids (Yao, Cherian *et al.*, 2015), unlike *E. coli* which possesses a gene encoding an acyl-carrier protein dehydratase/isomerase (*fabA*) used for the insertion of a double bond (Feng and Cronan, 2009). Therefore, it is hypothesised that oleic acid (18:1) was derived from the host HeLa cells (Yao, Cherian *et al.*, 2015). Furthermore, mass tracing of isotopically labelled fatty acids identified incorporation of lauric, palmitic, and myristic acid into the *C. trachomatis* phospholipids (Yao, Dodson *et al.*, 2015). As fatty acid biosynthesis is an energy intensive process, it is not unexpected that *Chlamydia* may instead scavenge these fatty acids from the host, which are then modified for use in the *Chlamydial* membrane. The mechanism by which host fatty acids are transferred to the *Chlamydial* inclusion remains unknown, and therefore in this research MOMP was investigated as a potential fatty acid transporter due to its homology with FadL, which formed the basis for structural homology modelling.

4.1.2 MOMP's homology to *E. coli* fatty acid transporter, FadL

Homology modelling of MOMP by Atanu *et al* in 2013 revealed a 14-stranded β-barrel structure, described in greater depth in **Chapter 1 Section 1.6.4**, which was based on the fatty acid transporter of *E. coli*, FadL, shown in **Figure 4.1A**. FadL permits cell growth with long chain fatty acids (LCFAs), fatty acids with over 14 carbons, which can act as substrates for the fatty acid degradation (Fad) enzymes, unlike fatty acids with a carbon chain length of less than eleven (Nunn and Simons, 1978). Specifically, FadL has the highest binding affinity to oleic acid (18:1) and palmitic acid (16:0), it binds myristic acid (14:0) with a lower affinity,

and with no affinity for decanoic acid (10:0) (Black, 1990). In *E. coli*, LCFAs are transported across the OM in an energy independent manner by FadL. A specific transporter is required due to the restrictions placed on spontaneous diffusion across the lipid bilayer by the LPS permeability barrier, which can extend to up to 30 Å in thickness (van den Berg, 2010). FadL has an internal N-terminal plug domain within its 14-stranded β-barrel, as shown in **Figure 4.1B**, which would require a spontaneous conformational change in order to allow classical transport of solutes. Instead, structural and functional growth studies have provided substantial evidence that FadL exhibits a lateral opening within the barrel due to a kink in the S3 strand (**Figure 4.1C**) (Hearn *et al.*, 2009), which interacts with the plug domain and creates a hole of approximately 8 x 10 Å, which is large enough for the passage of LCFAs. The detergent LDAO, used in the solubilisation process, was found to be positioned at the lateral opening indicative of a high affinity site for LCFAs, with many of the proximal residues hydrophobic in nature (Hearn *et al.*, 2009). The hydrophobic core of the barrel, in addition to the lateral opening, would then permit transport of LCFAs into the OM region, where these then could diffuse into the periplasmic space (Hamilton, 2007; van den Berg, 2010). The LCFAs are then transferred to the IM associated protein FadD (acyl-CoA synthase). Once transported into the cystosol by FadD the LCFAs then enter into the β-oxidation cycle, where a number of enzymes are involved in FA catabolism, such as FadE, FadB and FadA (Fujita *et al.,* 2007). Expression of these β-oxidation proteins, along with FadL and FadD, are repressed by FadR, which binds to promoter sites in the operon; this interaction is reversed by accumulation of long chain acyl-CoA esters, resulting from fatty acid uptake (Janßen and Steinbüchel, 2014). As FadL and FadD are required for fatty acid uptake, these proteins are only partially repressed by FadR (Feng and Cronan, 2012). In mutant strains, lacking a functional FadR, *E. coli* can grow on medium chain length fatty

acids, as well as LCFAs, since all Fad genes are no longer repressed (Janßen and Steinbüchel, 2014). Experiments with a *fadL* knockout had no effect on medium chain fatty acid uptake (6-13 carbons), which are typically more polar than LCFAs, indicating an alternative transport mechanism for these fatty acids (van den Berg, 2010).
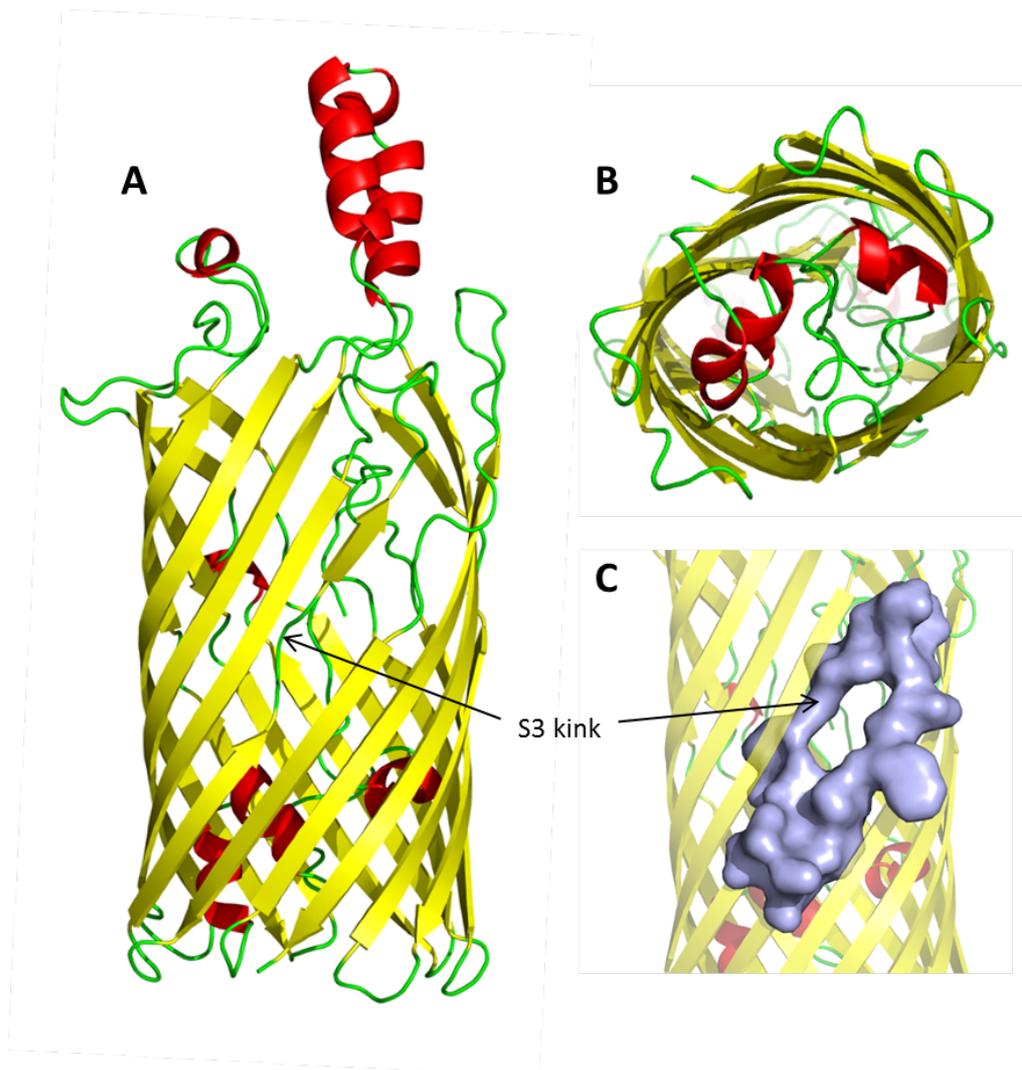
Figure 4.1. FadL crystal structure showing the occluded barrel and lateral opening. A, 14-stranded β-barrel structure; B, view up into the barrel of the N-terminal plug domain; C, the lateral opening at the S3 kink. 2.6 Å crystal structure of FadL of *E. coli* (PDBid: 1T16), coloured according to secondary structure (yellow, beta sheet; red, alpha helix; green, random loop) with the lateral opening highlighted in violet. Imaged using PyMOL (van den Berg *et al.*, 2004)

4.1.3 Increased thermal stability of MOMP in the presence of fatty acids

Circular dichroism (CD) spectroscopy records the unequal absorption of right and left-handed circularly polarised light from a chromophore, a light absorbing group such as the Cα carbons of the polypeptide backbone. SRCD spectroscopy extends the applications of CD spectroscopy as the elevated flux of a synchrotron permits the collection of data at considerably lower wavelengths and in much more dilute samples in smaller volumes (Wallace and Janes, 2010); a distinct advantage for membrane proteins, which are notoriously difficult to obtain in high quantities and at high concentrations. Due to the sensitivity of protein secondary structures, SRCD also permits data collection in the presence of small additives, such as ligands, or under environmental changes such as pH or temperature. The secondary structure of solubilised and purified rMOMP was analysed with SRCD in the far-UV, 190-260 nm, along with a number of potential ligands including the fatty acids palmitate and oleate. The midpoint of denaturation, also known as the Tm, is often taken as a measure of stability, which for native MOMP was 57.7 °C following temperature ramping between 10 °C and 80 °C (**Figure 4.2A**). Fatty acids palmitate and oleate increased the Tm of native MOMP to 59.3 °C and 72 °C (**Figure 4.2B**), respectively (Atanu, 2014). It is hypothesised that these two fatty acids contribute to MOMP's thermal stability, through binding of a hydrophobic hatch region at the externally exposed surface of MOMP, the site at which ligand transport would likely be initiated.
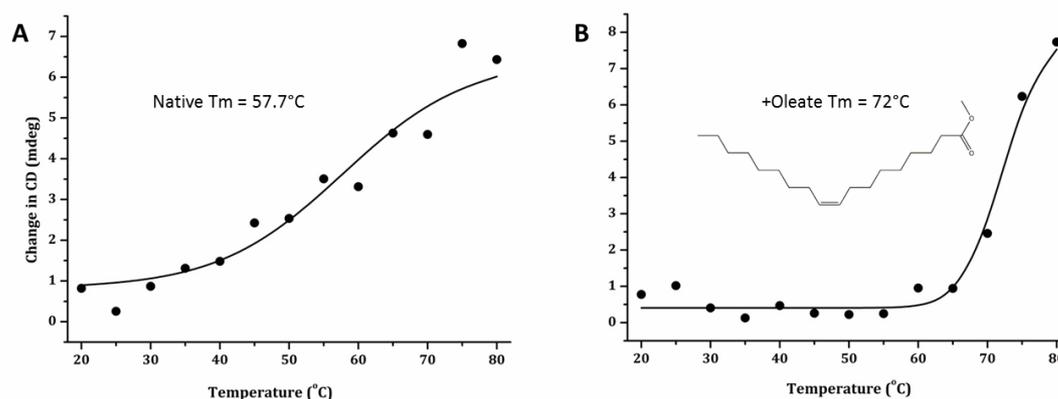
Figure 4.2. SRCD spectroscopic traces of A, native MOMP and B, MOMP supplemented with oleate. rMOMP solubilised in FC-12 and buffer exchanged into 0.026% (w/v) DDM was used at a concentration of 12 μm, in 20 mM Tris, 150 mM NaCl, 5 mM EDTA at pH 8.0. CD spectroscopic changes observed in the far-UV (at 218 nm) showing the increased Tm of 71.98 °C in the presence of five molar excess of oleate, the chemical structure of which is displayed (inset). (Atanu, 2014)

4.1.4 *E. coli* as a model system

E. coli is an excellent model organism for a fatty acid functional assay, as the transport proteins, mechanism, and degradation enzymes required for the use of LCFAs as a sole carbon source are all extremely well characterised, as described above. This enables knockout mutants of *fadL* to be used, in complementation with the POI, to determine if the POI can transport fatty acids and reinstate *E. coli* growth. Additionally, unlike liposomal swelling assays, pure protein is not required; the production of which is a costly and time consuming process, especially for membrane proteins. Instead, the protein can be expressed recombinantly from the pBADrha vector through induction with rhamnose (plasmid maps are shown in **Appendix A2**). Rhamnose, a naturally occurring sugar, is taken up by the RhaT transport system and is under the control of two activator proteins, RhaR and RhaS. In the presence of rhamnose, RhaR binds to the *rha*P_BAD promoter in order to activate transcription of both RhaR and RhaS, to which rhamnose and RhaS then together

initiate transcription of the downstream genes, resulting in a regulatory cascade responsive to differing rhamnose concentrations (Wegerer *et al.*, 2008). This makes rhamnose an especially useful inducer as protein expression is tuneable, a feature particularly important for toxic proteins, whereby a lower level of correctly folded and active protein is more desirable than higher yields of potentially misfolded protein trapped within inclusion bodies. The promoter is also controlled by catabolite repression (as described in **Chapter 3 Section 3.5.1**) and is therefore inactive in the presence of glucose, again in order to limit toxic basal expression. Additionally, unlike the pET101/D-TOPO vector, the bacterial strain is not required to contain the DE3 lysogen, improving the flexibility in mutant strain selection and removing the time-consumptive process of introducing the prophage.

An important consideration for all growth assays is the media to which the bacteria are exposed. In this instance, the functional assay exploits *E. coli*'s ability to grow on a sole carbon source of LCFAs and, therefore, it was imperative that all other potential carbon sources were omitted from the growth media. For this reason, minimal media (MM) is utilised because, as the name suggests, it only contains the minimal ingredients required to sustain culture; inorganic salts and water, supplemented with a carbon source of choice. Additionally, the agar used in the MM growth plates must be free of any carbon sources. Initially, sodium palmitate was used in the set-up of the functional assay, as it had been previously confirmed that FadL was the designated transporter of this fatty acid (Hearn *et al.*, 2009), whilst it also has been demonstrated that *Chlamydia* scavenge this fatty acid with high affinity from the host cells (Yao, Dodson *et al.*, 2015). Experiments assessing the binding affinity of FadL also indicate that LCFAs, oleate and myristate, which are scavenged by *Chlamydia*, are FadL substrates and therefore can be tested. Previous experimental data concluded that host-derived lauric acid also was scavenged by *C. trachomatis* cells,

however, as a medium chain fatty acid of twelve carbons it was not possible to include laurate as a fatty acid carbon source in this functional assay, as it is transported through the *E. coli* OM via an alternative mechanism to the FadL system, most likely through a porin due to its increased water solubility (Hearn *et al.*, 2009).

4.2   Aims and Objectives

MOMP's similarity to the *E. coli* fatty acid transporter FadL, together with the increased thermal stability of native MOMP observed in the presence of fatty acids, and the presence of host fatty acids in the *Chlamydial* outer membrane provides rationale for a functional growth assay in *E. coli*. From the low resolution 4 Å structure shown in **Chapter 1 Section 1.6.5**, MOMP does not seem to possess a lateral hole in the barrel wall, like FadL, and therefore remodelling of the N-terminal plug domain may be required for fatty acid transport. If MOMP is shown to act as a fatty acid transporter, then efforts can be focussed on designing an inhibitor of MOMP as an alternative to a vaccine, for which the former would result in decreased fatty acid uptake from the host, with the potential to both decrease *Chlamydia*'s infectivity and its ability to survive and proliferate within the host cells, since a stable membrane is essential for bacterial survival. Thus, based on previous research of *E. coli* FadL and on *Chlamydia*'s scavenging behaviour, uptake of LCFAs such as palmitic acid and oleic acid by rMOMP is worthy of investigation.

4.3   Materials and Methods

4.3.1 Materials

4.3.1.1 Minimal media (MM)

Liquid MM, pH 7.4 contained; 42 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 19 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$. The media was supplemented with either 0.4% carbon source such as glucose (w/v) or glycerol (v/v), or 5 mM fatty acid and 0.25% (w/v) Brij-58, plus appropriate antibiotics.

Fatty acids were prepared from their sodium salts, e.g. sodium oleate for oleic acid, in order to improve solubility. Autoclaved nH_2O was used to create a 20 mM fatty acid solution containing 1% (w/v) Brij-58, heated to 70 °C in a water bath with frequent stirring to promote solubility. MM plates containing fatty acids were then prepared immediately with liquid MM supplemented with 1.5% (w/v) Oxoid no. 1 Agar.

4.3.1.2 Oligomers

Oligomers for all DNA work are given in **Chapter 2 Table 2.6**.

4.3.1.3 Bacterial strains

A list of bacterial strains used can be seen in **Chapter 2 Table 2.1**.

4.3.2 Methods

The following methods described herein relate specifically to the functional assay set up and optimisation. The standard and finalised protocol is detailed in **Chapter 2 Section 2.2.4**.

4.3.2.1 pBADMOMP$_{H6}$ expression vector design

For construction of a rhamnose inducible expression system, hexahistidine tagged MOMP from the pETMOMP$_{H6}$ vector (R1185) was amplified using the CloneAmp HiFi PCR premix, as described in **Chapter 2 Section 2.2.1.3**. The pBADrha vector was linearised using restriction enzyme digestion at a single *NdeI* site, four bases downstream of the Shine-Dalgarno sequence. 0.5 μl of FastDigest *NdeI* (ThermoFisher Scientific) was incubated with 2 μl of pBADrha DNA (~100 ng) along with 1X FastDigest buffer (ThermoFisher Scientific) in a total volume of 10 μl for 1h at 37 °C. The enzyme was inactivated by heat treatment at 65 °C for 5 minutes and the linearised vector size checked with agarose gel electrophoresis. Both the linearised vector and the amplified insert were *DpnI* treated and purified before the In-Fusion reaction. The resulting vector was sequenced in both directions at the In-Fusion sites using sequencing primers rhaB and pBAD (**Chapter 2 Table 2.6**) and named pBADMOMP$_{H6}$ (R1407).

4.3.2.2 Addition of ampicillin resistance to the pBADrha vector

Due to the use of additional *E. coli* knockout mutants that contained the same antibiotic resistance as the pBADrha expression vector (chloramphenicol), an ampicillin resistance gene was added to the vector. The ampicillin resistance gene was amplified from the pETMOMP$_{H6}$ (R1185) vector, which was then inserted by In-Fusion into empty vector pBADrha as well as the pBADMOMP$_{H6}$ vector, both of which had been linearised by inverse PCR (**Chapter 2 Section 2.2.1.5**). The two new vectors were checked for resistance by growth on LB ampicillin plates and named pBADa and pBADaMOMP$_{H6}$ (R1411), respectively (**Appendix A2.4**).

4.3.2.3 Complementation vector design

For complementation of the functional assay, *fadL*, the knockout gene, was inserted into the pBADa vector. Genomic DNA was extracted from wild type BW25113 *E. coli* cells using a QuickDNA kit (ZymoResearch), and used as a template for *fadL* amplification with primers pBADa FadL for and rev displayed in **Chapter 2 Table 2.6**. The pBADa vector was linearised as before, with restriction enzyme *NdeI*, and the purified PCR product with *fadL* gene inserted by an In-Fusion reaction. The plasmid, designated pBADaFadL (R1414, **Appendix A2.3**), was isolated from an *E. coli* Top10 transformant and sequenced in both directions using sequencing primers rhaB and pBAD (**Chapter 2 Table 2.6**).

4.3.2.4 Initial rhamnose concentration testing

A small-scale growth experiment of BW25113 *E. coli* cells transformed with either pBADMOMP$_{H6}$ or empty vector, pBADrha, was conducted to identify the optimum rhamnose concentration for limiting cell toxicity. Transformed cells from LB plates were inoculated in triplicate into 10 ml LB with 34 μg/ml chloramphenicol where appropriate and 0.6% (w/v) glucose and incubated overnight at 37 °C with shaking at 225 rpm. Overnight cultures were diluted 100-fold into fresh LB and incubated at 37 °C with shaking until an OD$_{600}$ of ~0.5. At this point, cultures were induced with either 0.2%, 0.02%, or 0.002% (w/v) rhamnose and grown for a further 7h at a lower temperature of 25 °C. At 5h and 7h post induction, OD$_{600}$ measurements were taken. This was later repeated in triplicate for knockout strain LS6164 transformed with pBADaMOMP$_{H6}$ testing the following rhamnose concentrations; 0%, 0.005%, 0.01%, and 0.02% (w/v), with OD$_{600}$ measurements taken every hour following induction for a total of 6h.

4.3.2.5 Medium-scale expression test to determine optimal rhamnose concentration

A medium-scale expression test of BW25113 *E. coli* cells transformed with both pBADMOMP$_{H6}$ and empty vector pBADrha was conducted in order to identify the optimum rhamnose concentration for both protein expression and toxicity limitation. Membranes were prepared following the medium-scale expression testing protocol described in **Chapter 2 Section 2.2.2.3**, with the following amendments; use of 34 μg/ml chloramphenicol in place of ampicillin and, based on the results described in the experiment conducted in **Section 4.4.1**, cultures were induced with either 0.1% or 0.02% (w/v) rhamnose as opposed to IPTG and grown for a further 3h at 25 °C. Harvested membranes were analysed by SDS-PAGE and Western blotting.

4.3.2.6 Brij-58 toxicity assay

A Brij-58 toxicity assay was conducted to determine the level of Brij-58 at which cell growth begins to be adversely affected. BW25113 cells grown from MM plates were inoculated into liquid MM supplemented with 0.4% (w/v) glucose and incubated for 16h at 37 °C with shaking at 225 rpm. This culture was then diluted 100-fold into fresh MM containing 0.4% (w/v) glucose with varying concentrations of Brij-58 (0 to 1% w/v) and incubated with shaking for 20h at 37 °C. An OD$_{600}$ measurement was taken every hour for 8h with a final measurement taken at 20h.

4.3.2.7 Super-resolution microscopy

E. coli knockout strain LS6164 cells transformed with pBADaMOMP$_{H6}$ were grown overnight in MM with 0.4% (w/v) glucose at 37 °C in a shaking incubator. Overnight cultures were diluted 1/10 into fresh MM containing 0.4% (v/v) glycerol and grown at 37 °C in a shaking

incubator until an OD_{600} of ~0.5. Cells were induced with 0.02% (w/v) rhamnose for 3h, before harvesting and resuspension to an OD_{600} value of 0.5 in PBS. Cells were then prepared following the Ibidi slide method developed in this research (**Chapter 2 Section 2.2.5.4**) and labelled with 1/100 Anti-his Alexa 647 fluorophore. LS6164 control cells were also prepared in MM, with OmpA labelled with 1/500 anti-*E. coli* OmpA polyclonal rabbit primary antibody and 1/500 Alexa-488 anti-rabbit secondary antibody. All cells were imaging following the standard STORM imaging conditions described in **Chapter 2 Section 2.2.5.6**.

4.4 Results

4.4.1 Induction with 0.02% (w/v) rhamnose is optimal for MOMP expression from pBADrha vector

For tuneable expression, the pBADrha vector was utilised in the functional growth assay in order to finely control protein expression with titratable inducer rhamnose. Additionally, this vector does not require the DE3 lysogen, unlike the pET101/D-TOPO vector, and therefore permitted the use of a wider variety of strains. A small-scale expression test was conducted in BW25113 *E. coli* cells transformed with empty pBADrha vector or pBADMOMP$_{H6}$ (details on plasmid construction in **Section 4.3.2.1** and plasmid map in **Appendix A2.4**) in order to determine the rhamnose concentration at which cell toxicity begins to occur due to expression of MOMP. Toxicity was measured by OD_{600}, a monitor of cell density. Typically, between 0.002% and 0.2% (w/v) final concentration of rhamnose is used with this expression system (Cardona and Valvano, 2005; Ford *et al*., 2014; Wegerer *et al.,* 2008) and, therefore, a broad range of concentrations of 0.002%, 0.02% and 0.2% were initially tested in this study. Cell cultures grown in LB were supplemented with rhamnose, after reaching an OD_{600} of ~0.5 and grown for an additional five and seven hours.

Construct	Rhamnose (w/v %)	OD_{600}	
		5h	7h
pBADrha	0.002	1.8	2.5
	0.02	1.5	2.1
	0.2	1.7	2.0
pBADMOMP$_{H6}$	0.002	1.3	1.6
	0.02	0.9	0.9
	0.2	0.7	0.8

Table 4.1. The effect of rhamnose concentration on cell density. BW25113 *E. coli* cultures transformed with either empty vector pBADrha or pBADMOMP$_{H6}$ were prepared in duplicate, grown in LB supplemented with 34 µg/ml chloramphenicol, and induced with rhamnose at an OD_{600} of ~0.5. Measurements were taken at 5h and 7h post induction and averaged. Rhamnose concentration is shown as w/v.

As indicated in **Table 4.1**, in contrast to cultures with the pBADrha vector control, MOMP expressing cells appeared to have stopped growing between the two timepoints, and had either grown more slowly or begun to die at rhamnose concentrations of 0.02% and 0.2%. Additionally, in conditions containing 0.02% and 0.2% rhamnose, cell density was considerably lower than that of the lowest rhamnose concentration (0.002%), at below OD_{600} 0.9 compared to over OD_{600} 1.3, respectively. Furthermore, in all instances, cell density in MOMP expressing cells was lower than the corresponding control cells transformed with only the empty vector, with the control cells also impacted by increasing rhamnose concentrations.

In order to compare protein expression, a larger scale expression test was conducted, whereby membranes were isolated from induced cells and analysed by SDS-PAGE. The low impact of cultures induced with 0.002% rhamnose likely corresponded to low MOMP expression, and therefore the two higher rhamnose concentrations (0.02% and 0.1%) were selected for protein expression studies. The duration of protein expression also was limited to just three hours, due to the cessation of growth observed between five and seven hours in the initial trial. At the time of harvesting, both cultures transformed with pBADMOMP$_{H6}$

had reached an OD_{600} of 1.2, which suggests that both rhamnose concentrations had a similar effect upon cell growth, although the cells transformed with the empty vector reached a higher OD_{600} of 2.00.

Membranes were isolated as described in **Chapter 2 Section 2.2.2.3** and were analysed by SDS-PAGE, with both Coomassie blue staining and Western blotting, in order to compare MOMP expression. MOMP (40 kDa) was rarely discernible using Coomassie blue, due to the presence of an *E. coli* protein (likely OmpF) at the same MW and, therefore Western blotting was typically required for visualisation of MOMP, however, Coomassie blue staining was useful for comparing total protein levels. This ensured that the normal physiological protein expression of the cell was as expected and that samples were appropriately standardised for comparison and, therefore, any differences in levels of MOMP could be attributed to the recombinant expression conditions. In **Figure 4.3A**, a faint band for MOMP, identified through its absence in the pBADrha control samples, is indicated by the arrow just above the 40 kDa MW marker. The Western blot in **Figure 4.3B** confirmed MOMP expression in these samples and it was evident that actually the lower concentration of rhamnose (0.02%, well 5) produced more protein than the higher concentration (0.1%, well 6), which also corresponds to a faint band with Coomassie blue staining. Consequently, 0.02% (w/v) rhamnose was used for induction of MOMP expression.
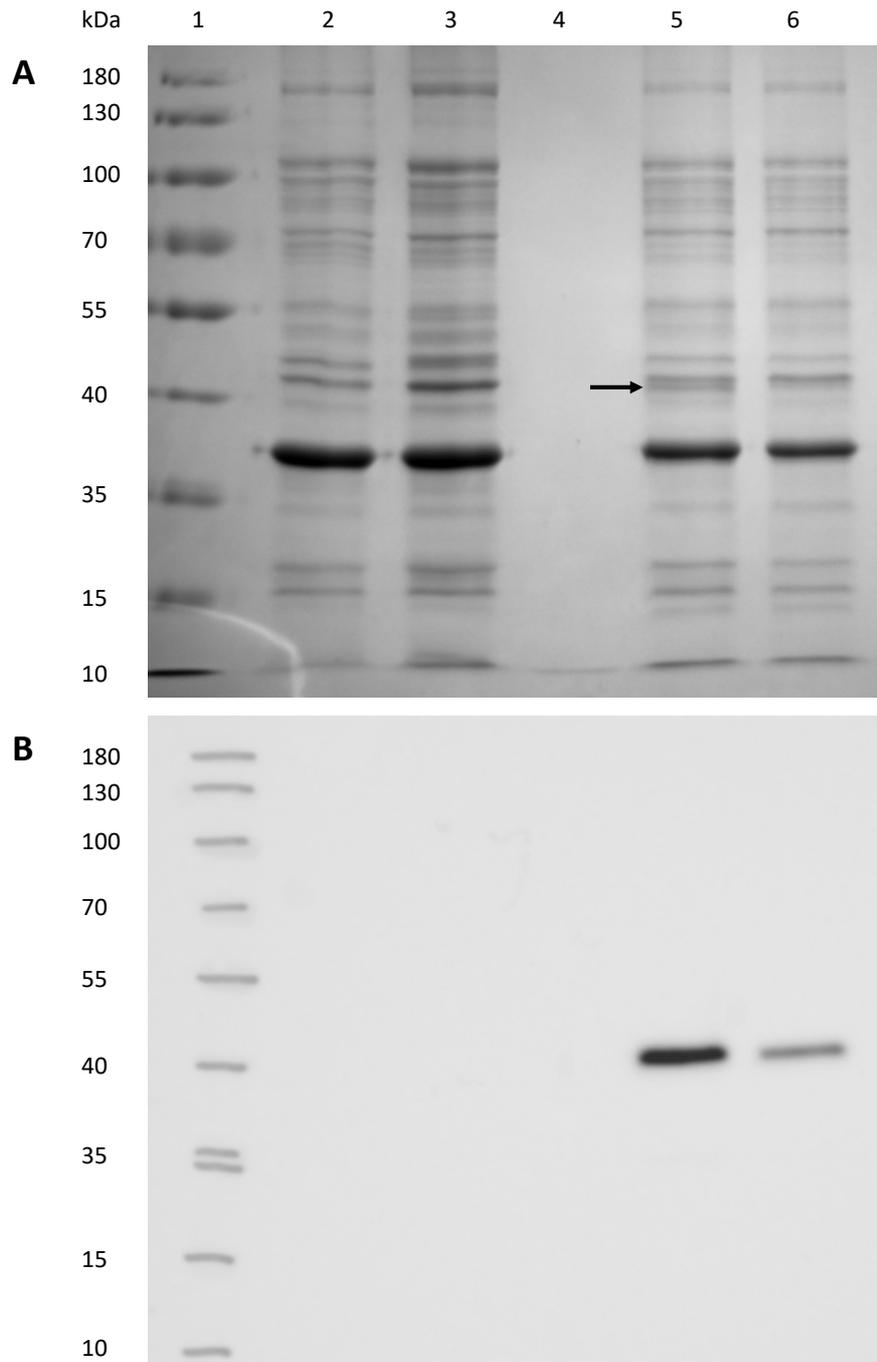
Figure 4.3. **SDS-PAGE analysis of pBADMOMP$_{H6}$ expression induced with 0.1% or 0.02% rhamnose.** Membranes were isolated from BW25113 *E. coli* cells expressing either empty vector pBADrha or pBADMOMP$_{H6}$ following the protocol described in **Chapter 2 Section 2.2.2.3** under the same conditions as per **Table 4.1**, for 3 hours. **A) Coomassie blue stained gel.** The suspected location of MOMP is indicated by the arrow. **B) Western blot.** For both samples, percentages refer to rhamnose concentration (w/v). 1, MW protein standard; 2, 0.02% pBADrha; 3, 0.1% pBADrha; 4, Blank; 5, 0.02% pBADMOMP$_{H6}$; 6, 0.1% pBADMOMP$_{H6}$. MOMP was detected with 1/10,000 anti-his HRP conjugated antibody.

4.4.2 Brij-58 toxicity is reduced with concentrations below 0.5% (w/v)

In order to establish the functional assay for fatty acid transport in *E. coli* (Hearn *et al*., 2009), BW25113 wild type and the two isogenic Δ*fadL* mutants, named JW2341 7-3 and 8-3, were grown with a sole carbon source of sodium palmitate. As fatty acids are insoluble in water, due to their hydrophobic carbon chains, supplementation with a detergent such as Brij-58 was required. Following the protocol described by Hearn *et al* in 2009, the strains were spotted onto MM plates, containing 5 mM sodium palmitate and 0.5% (w/v) Brij-58. Whereas the control cells grown on MM supplemented with 0.4% (w/v) glucose showed confluent colony formation, the sodium palmitate plates had spots of translucent agar, as shown in **Figure 4.4**.
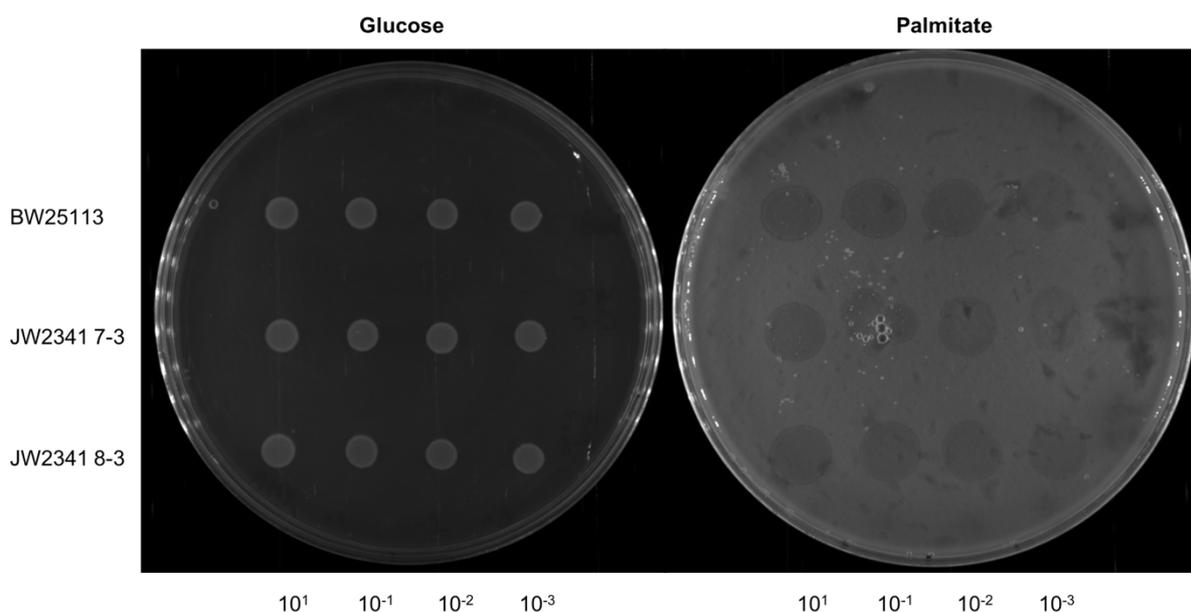
Figure 4.4. Initial spot plate assay to assess growth of wild type and *fadL* knockout *E. coli* strains on MM. A dilution series of wild type BW25113 and JW2341 Δ*fadL* cultures (7-3 and 8-3) were grown overnight in MM with 0.4% (w/v) glucose before being washed and plated onto MM plates containing either 0.4% (w/v) glucose or 5 mM sodium palmitate + 0.5% (w/v) Brij-58. Plates were incubated at 37 °C for 96h. Tenfold dilution factor is given along the horizontal axis.

As the knockout JW2341 cells, not expected to grow on sodium palmitate, exhibited the same spots of translucent agar as the wild type BW25113, it was first hypothesised that the Brij-58 (used to solubilise the fatty acid) was provoking a toxic effect on the cell, resulting in cell lysis and release of sodium palmitate metabolising enzymes. A quick investigation of this revealed that colonies plated onto glucose MM plates, with and without 0.5% (w/v) Brij-58 grew significantly better without Brij-58 supplementation, suggesting that this detergent was indeed detrimental to *E. coli* growth. To quantify the extent of this effect, a toxicity assay was conducted to assess the effect of various concentrations of Brij-58 on BW25113 growth, measured by optical density, in liquid MM over an eight hour period with a final measurement taken following overnight growth. The graph in **Figure 4.5** shows the growth curve over this period of time. However, further investigation to a lower dilution revealed that the cleared region within the opaque sodium palmitate plates was in fact a result of *E. coli* growth and healthy utilisation of the fatty acid, as visualised with individual colony formation at lower dilutions. Growth here was unexpected, due to the knockout of transporter FadL in the JW2341 strain, hence further testing of Brij-58 toxicity was conducted, as described below.
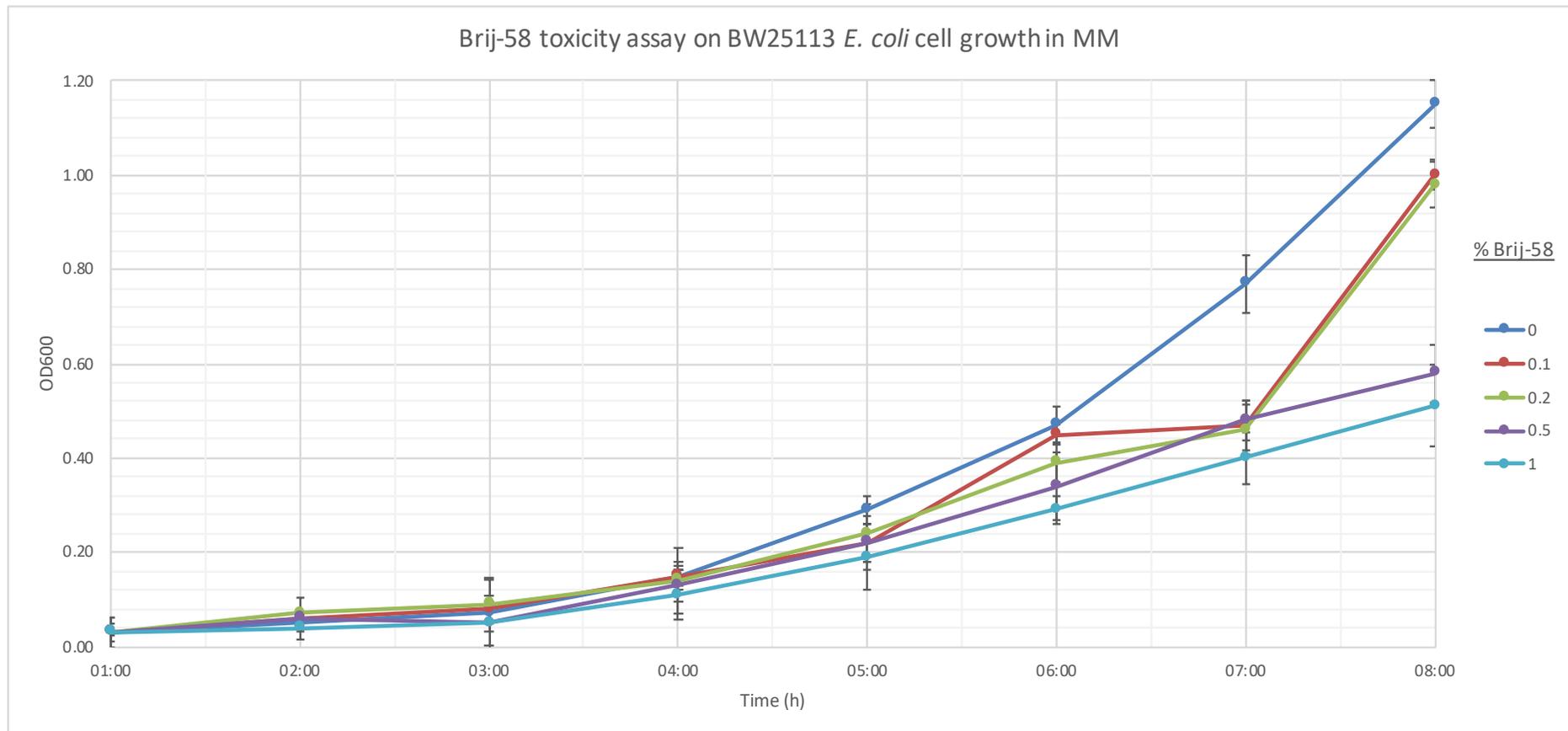
Figure 4.5. Growth curve of BW25113 cells to assess the toxicity of varying concentrations of Brij-58. Liquid MM cultures were supplemented with 0 to 1% (w/v) Brij-58 to assess toxicity. Measurements were taken hourly over an 8 hour period. Error bars represent standard deviation.

It was evident that by eight hours, the cultures containing the highest concentrations of Brij-58, 0.5% and 1%, were of a much lower density than those of the lower concentrations, between 0% and 0.2%. After twenty hours of growth, the OD_{600} measurement of the 0% control was 3.4, the 0.1% and 0.2% samples were 3.1 and 3.00 respectively, with the 0.5% and 1% samples, 2.7 and 2.2, respectively. It was clear that inclusion of 0.5% to 1% Brij-58 in the culture media reduced growth of BW25113. As Brij-58 was vital for the solubilisation of the fatty acids required in this functional study, reduction of the Brij-58 concentration may have hindered this solubility. However, experimentation revealed good solubility of 20 mM sodium palmitate in a 1% concentration of Brij-58, which allowed continued use of MM plates with a final concentration of 5 mM sodium palmitate and just 0.25% (w/v) Brij-58. Decreasing the Brij-58 concentration below 0.2% would slightly improve bacterial growth and survival, but be detrimental to sodium palmitate solubility and was therefore not used.

4.4.3 Confirmation of suitable knockout *E. coli* strains

Prior to optimising the spot plate growth assay, it was first necessary to monitor the growth of wild type BW25113 and the two knockout JW2341 Δ*fadL* strains on different carbon sources, before complementation with MOMP and FadL, to confirm that the physiological behaviour of the cells was as expected. As described in **Section 4.4.2** and shown in **Figure 4.4** the initial spot plate showed plaque-like areas of sodium palmitate metabolism, assumed to result from the toxic effect of Brij-58. An additional unconfirmed knockout mutant was sourced, also lacking the *fadR* fatty acid metabolism regulator protein gene in addition to *fadL*, which was called BW25113 Δ*fadL* Δ*fadR* (van den Berg, unpublished) for use in this assay. As this newly acquired knockout mutant was uncharacterised, it was important to confirm the expected mutations with preliminary spot plate testing.

After decreasing the Brij-58 concentration to a more tolerable level of 0.25%, cells were spotted to a lower dilution of 10^{-7} (equating to ~<4 cells per 5μl spot) in order to identify single colonies (**Figure 4.6A**), as the initial glucose control plates only demonstrated confluent growth in the spotted areas (**Figure 4.4**). As previously described in **Section 4.4.2**, it appeared that the plaque-like areas observed on the sodium palmitate plates in **Figure 4.4** were in fact due to confluent growth, due to the appearance of colonies at dilutions below 10^{-5} for both wild type BW25113 and knockout mutant JW2341 in **Figure 4.6B.** As the BW25113 knockout mutant showed comparable growth to the wild type strain on glucose plates, and no growth on sodium palmitate, as expected, this strain was taken forward for complementation with MOMP and FadL. Additional spot plates on MM, MM + 0.25% Brij-58, and MM + 0.02% rhamnose showed no growth for either strain, indicating that neither Brij-58, required for fatty acid solubilisation, or rhamnose, used for induction, were utilised as carbon sources (data not shown) and hence would not interfere with the assay.
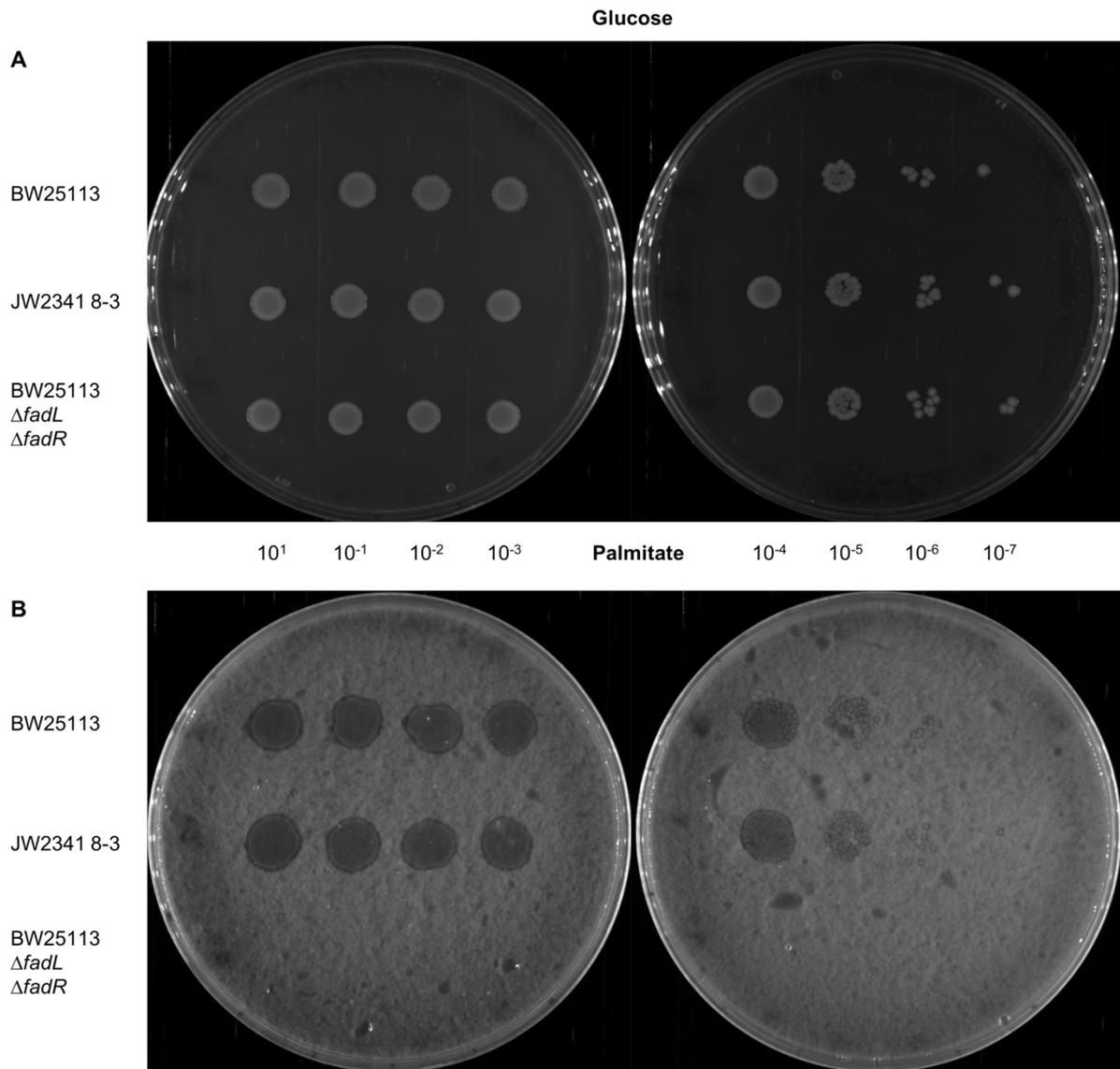
Figure 4.6. The BW25113 knockout strain was not able to grow on palmitate plates, as expected. A dilution series, down to 10^{-7}, was prepared for cells grown overnight in MM and spotted onto MM plates, containing either **A**, 0.4% (w/v) glucose or **B**, 5 mM sodium palmitate + 0.25% (w/v) Brij-58. Plates were incubated at 37 °C for 96h. Dilution is given in the centre. Individual colonies can be observed at the lower dilutions.

4.4.4 FadL complementation as a positive control

As the BW25113 knockout strain contained the same antibiotic resistance marker (chloramphenicol) as the pBADMOMP$_{H6}$ plasmid, an ampicillin resistance marker was introduced into the pBADrha vector to permit selectivity of the knockout strain transformed with pBADMOMP$_{H6}$. This new plasmid was called pBADaMOMP$_{H6}$. Additionally, the *fadL* gene from BW25113 wild type *E. coli* was amplified from genomic DNA and also inserted into the new pBADa vector and named pBADaFadL. Complementation of the knockout BW25113 cells with the pBADaFadL plasmid was intended to serve as a positive control for this transport assay, by reintroducing the gene encoding the knocked-out transporter protein, FadL. As performed previously, spot plates were set up from transformed cells cultured in MM. Both uninduced overnight cultures and cultures induced for three hours were spot plated onto sodium palmitate plates with either +/- rhamnose (0.02% w/v), in addition to the control glycerol plates. Glycerol was used as a carbon source in control plates for strains transformed with FadL or MOMP in order to prevent catabolite repression. However, knockout BW25113 cells transformed with pBADaFadL failed to grow under any conditions, except with glycerol as a carbon source (data not shown). As the knockout genes of this strain were unconfirmed, it is possible that other detrimental effects were unintentionally introduced. Consequently, a new knockout strain was acquired, named LS6164, again with Δ*fadL* Δ*fadR* knockouts (Ginsburgh *et al*., 1984). As this was a newly acquired strain originating from K12, optimal rhamnose concentration testing was repeated to confirm the optimum rhamnose concentration for minimisation of toxicity and optimisation of expression. The growth curve in **Figure 4.7** shows the results whereby only cultures without any rhamnose (dark blue, LS6164 control; orange, LS6164 pBADaMOMP)

grew normally. Any addition of rhamnose resulted in a decrease in growth shortly after

induction.
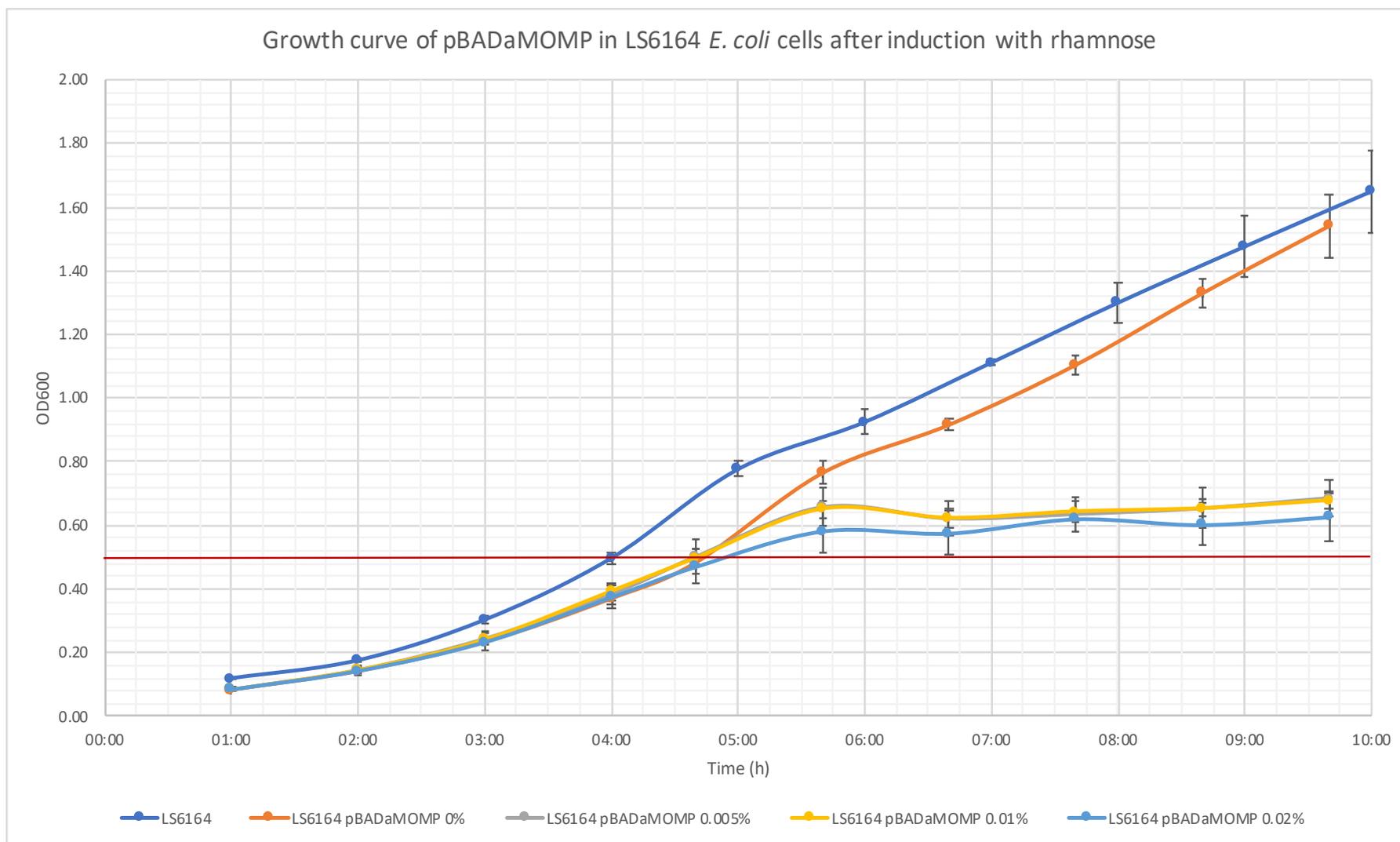
Figure 4.7. Growth curve of LS6164 *E. coli* cells induced with increasing concentrations of rhamnose. Liquid LB cultures were induced, in triplicate, with 0% to 0.02% (w/v) rhamnose upon reaching an OD_{600} of approximately 0.5 (red line). Cultures were then grown for an additional 6h, with OD_{600} measurements taken every 1h. Error bars represent the standard deviation.

Samples of whole cells harvested during growth at one, three, and five hours post induction, were pelleted, solubilised in SDS-PAGE sample buffer, and analysed by Western blotting for protein expression (**Figure 4.8**, loading control **Appendix 4 Figure A4.1**). The strongest reaction for MOMP$_{H6}$ was identified in well 13, corresponding to a rhamnose concentration of 0.02% at three hours expression. There was no detectable MOMP$_{H6}$ with 0% or 0.005% (w/v) rhamnose. As decreasing the rhamnose concentration did not improve cell growth, as indicated in **Figure 4.7**, all future experiments using the LS6164 knockout strain were induced with a concentration of 0.02% rhamnose, where applicable.
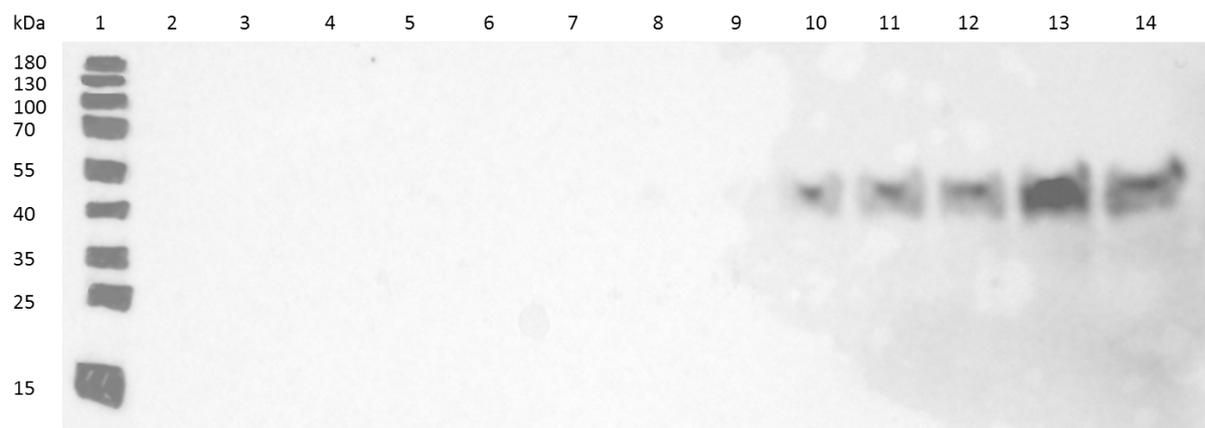
Figure 4.8. **Western blot of whole LS6164 knockout cells expressing pBADaMOMP$_{H6}$.** Cultures were grown as described in **Figure 4.7** and induced with a range of rhamnose concentrations (given as percentage w/v). Whole cells were harvested after 1h, 3h, and 5h. 1, MW marker; 2, LS6164 control cells; 3, 0% 1h; 4, 0% 3h; 5, 0% 5h; 6, 0.005% 1h; 7, 0.005% 3h; 8, 0.005% 5h; 9, 0.01% 1h; 10, 0.01% 3h; 11, 0.01% 5h; 12, 0.02% 1h; 13, 0.02% 3h; 14, 0.02% 5h. Samples 3-14 were all transformed with the pBADaMOMP$_{H6}$ construct. MOMP was detected with 1/10,000 anti-his HRP conjugated antibody.

Following additional rhamnose testing for the new knockout strain, a number of conditions for the MM plates were assessed including non-induced, induced for three hours, and induced with spotting onto rhamnose supplemented plates for wildtype BW25133, knockout LS6164, and LS6164 complemented with pBADaFadL, the results of which are

shown in **Figure 4.9**. As the pBAD promoter is under catabolite repression, as previously mentioned, glucose was substituted with glycerol in order to permit induction of FadL expression with rhamnose in the LS6164 strain transformed with pBADaFadL.
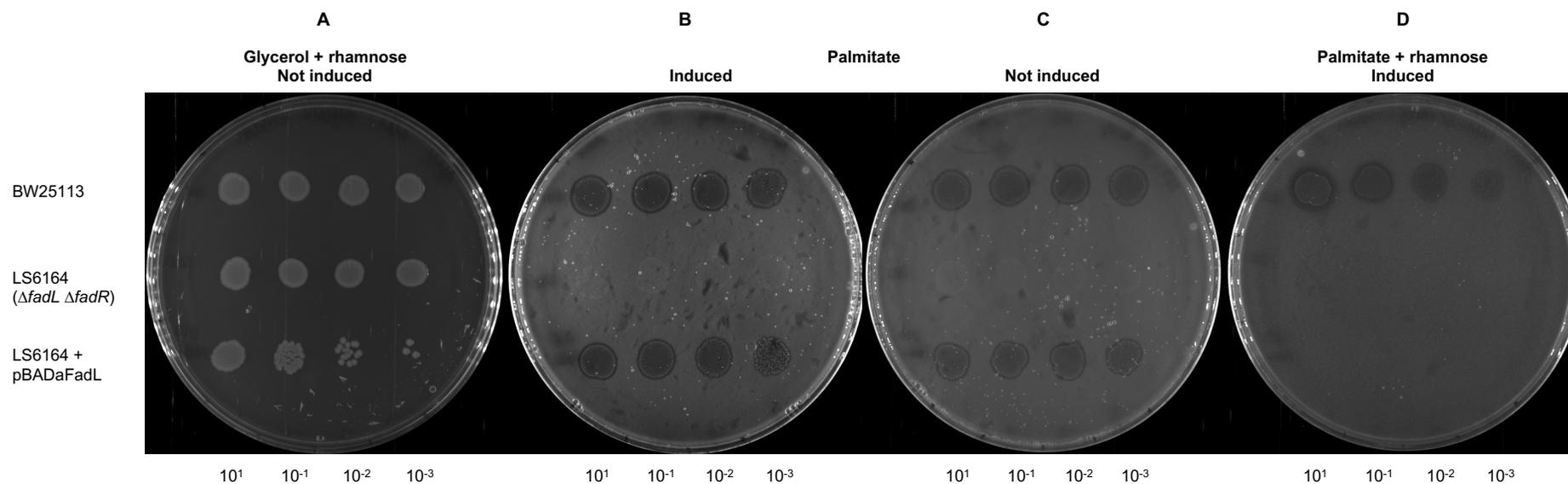
Figure 4.9. Spot plate assay demonstrating successful complementation with FadL in comparison to wildtype and LS6164 knockout cells. A, 0.4% (v/v) glycerol plates containing rhamnose, **B** and **C**, 5 mM palmitate plates, and **D**, 5 mM palmitate plates containing rhamnose. Dilutions from 10^1 to 10^{-3} were spotted on a range of MM plates. Induced cells were sub-cultured from overnight cultures and grown to an OD_{600} of 0.5 before induction with 0.02% (w/v) rhamnose for 3h at 25 °C. Uninduced cells were sub-cultured and grown at 37 °C for the same duration. Sodium palmitate was solubilised with 0.25% (w/v) Brij-58, as before.

From these results it is evident that the new LS6164 mutant cannot transport palmitate but is able to grow with glycerol (**Figure 4.9A**), as expected. Complementation of this knockout with pBADaFadL showed that the knockout strain was then able to resume transport of palmitate in both induced and non-induced cells, as indicated by translucent spots in **Figure 4.9B** and **4.9C**. However, it is apparent that rhamnose supplemented into the MM plates prevented growth of pBADaFadL transformed knockout cells, as indicated by the lack of colonies in **Figure 4.9D**. There appeared to be fewer colonies of these cells in the rhamnose containing glycerol plate (**Figure 4.9A**), at dilutions below 10^{-2}, in comparison to the LS6164 knockout cells. It was evident that plates supplemented with rhamnose were detrimental to the growth of cells transformed with pBADaFadL as well as wild type BW25113 cells and, therefore, in all future experiments rhamnose was omitted from MM plates.

4.4.5 Complementation with MOMP did not demonstrate cell growth on palmitate or oleate

After successful confirmation of the assay with the FadL positive control, pBADaMOMP$_{H6}$ was transformed into the knockout strain to determine if MOMP was capable of transporting sodium palmitate or sodium oleate. Cells expressing FadL grew under all conditions, as shown in **Figure 4.10A**, confirming the robustness of the assay. However, for cells expressing MOMP, colonies failed to form at any dilution (**Figure 4.10B**). As for the positive control, this experiment was repeated at two induction times of 1.5h and 3h, and with rhamnose induction at a lower concentration of 0.01% in an attempt to alleviate any potential over expression or toxicity. However, LS6164 cells expressing MOMP were still unable to grow with a fatty acid carbon source (data not shown).
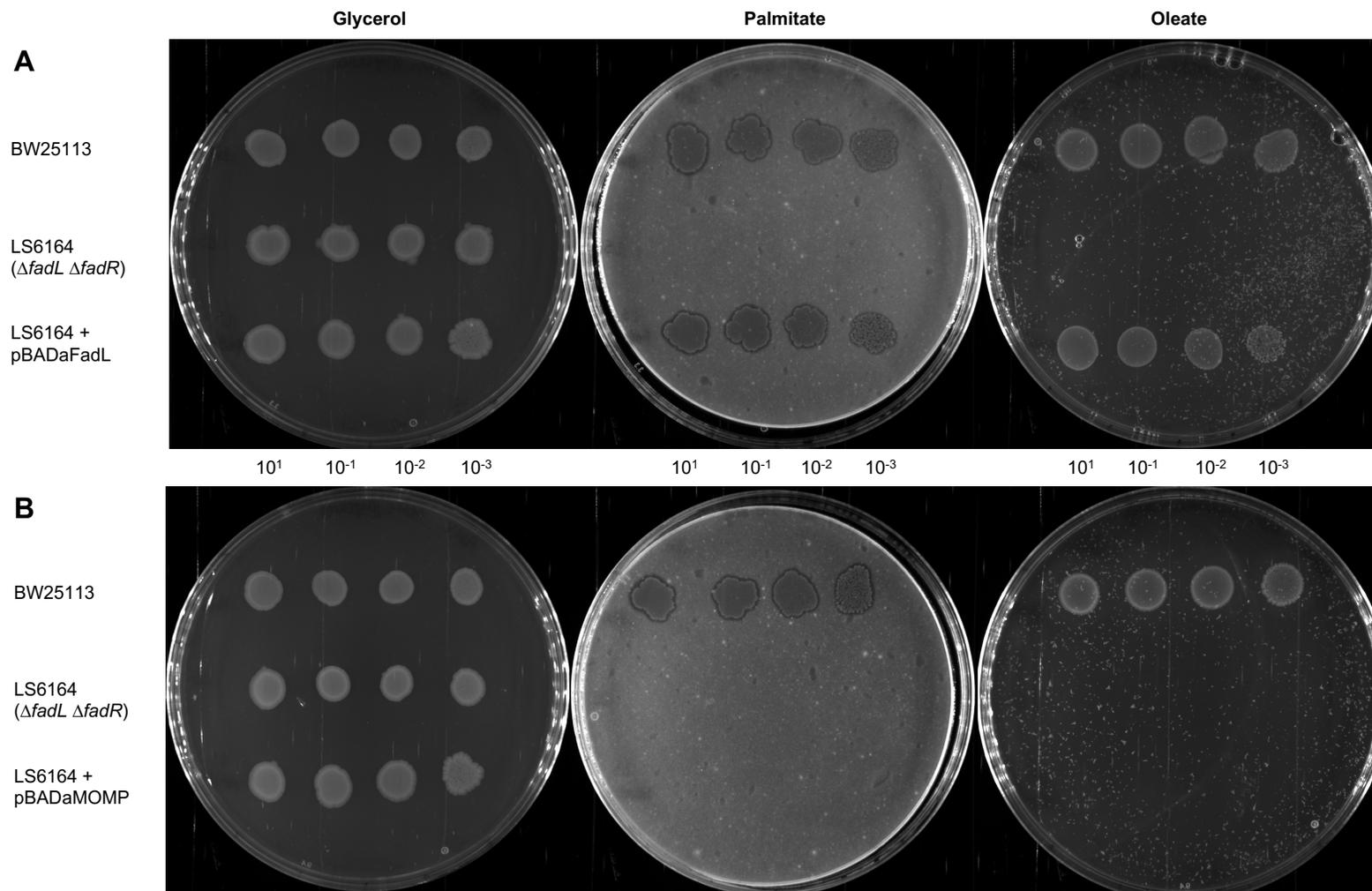
Figure 4.10. Complementation with MOMP does not permit cell growth on palmitate or oleate. Dilutions from 10^1 to 10^{-3} were spotted on a range of MM plates; 0.4% (v/v) glycerol, 5 mM palmitate, or 5 mM oleate, with none containing rhamnose. Induced cells were sub-cultured from ON cultures and grown to an OD_{600} of 0.5 before induction with a lower concentration of 0.01% (w/v) rhamnose for 3h at 25 °C. Fatty acids were solubilised with 0.25% (w/v) Brij-58, as before.

4.4.6 MOMP was located at the outer membrane with fluorescence microscopy

One hypothesis as to why complementation with MOMP did not permit growth of the knockout strain on a fatty acid carbon source was that MOMP was not assembled in the outer membrane. To determine the cellular location of MOMP, pBADaMOMP$_{H6}$ transformed into LS6164 was grown in MM and induced for three hours with 0.02% (w/v) rhamnose to replicate the conditions used in the spot plate growth assay. Anti-his Alexa-647 was used to label his-tagged MOMP before imaging with epifluorescence and STORM, two techniques described in **Chapter 5**. OmpA, a major protein of the *E. coli* outer membrane, was also fluorescently labelled with secondary antibody Alexa-488, as prepared in **Chapter 2 Section 2.2.5.5** for TIRF microscopy. The results are shown in **Figure 4.11**. OmpA (**Figure 4.11A)** is visualised as a distinctive and characteristic ring around the perimeter of the cell with epifluorescence imaging, and shown in greater detail with STORM, which can achieve resolutions of up to 20 nm (Rust *et al.*, 2006). **Figure 4.11B** shows the MOMP labelled cells. Again, the majority of the fluorescent signal was located at the perimeter of the cells, with both epifluorescence and STORM imaging. The fluorescent signal observed in the MOMP epifluorescence sample does not exhibit the same membrane continuity as OmpA and has some denser patches; **Chapter 5** is dedicated to further investigation of this observation. The brightfield images serve to demonstrate the location and morphology of the cells, but are shown with a wider field of view and, as such, have differing scale bars.
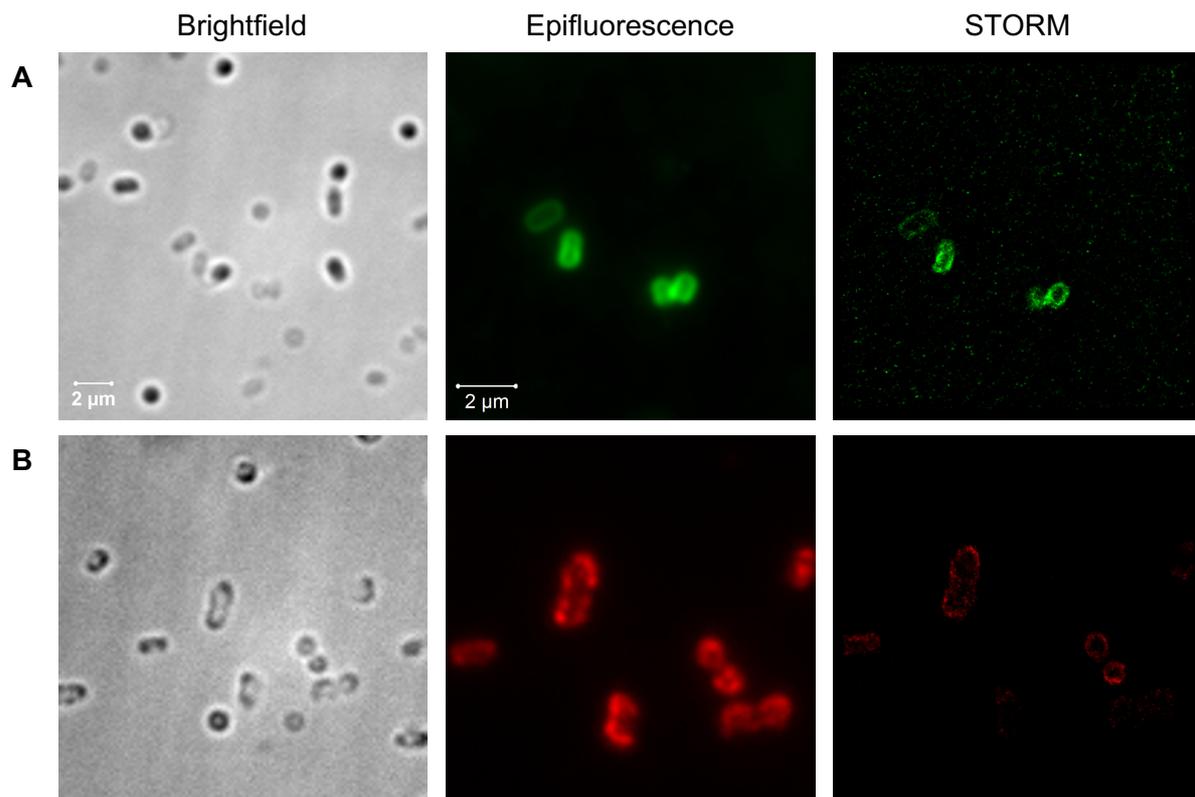
Figure 4.11. Brightfield, epifluorescence and STORM imaging of A, OmpA and B, MOMP labelled LS6164 cells. Cells transformed with pBADaMOMP were grown under the same conditions used in the spot plate assay and induced for 3h with 0.02% (w/v) rhamnose. OmpA was labelled with 1/500 anti-*E. coli* OmpA polyclonal rabbit primary antibody and 1/500 Alexa-488 anti-rabbit secondary antibody. MOMP was labelled with 1/100 anti-his Alexa-647 conjugated antibody. Data was collected using Octopus at the Central Laser Facility, Harwell.

4.5 Discussion

4.5.1 Determining the optimal rhamnose concentration

The pBADrha expression system is controlled by the titratable inducer rhamnose, which is typically used at concentrations between 0.002% and 0.2% (Cardona and Valvano, 2005; Ford *et al*., 2014; Wegerer *et al.,* 2008). Induction with rhamnose is highly dependent on the number of rhamnose permeases present in each cell (Rosano and Ceccarelli, 2014) and, therefore, a cell with a high number will over-produce the recombinant protein. As expression levels are dependent on the strain used, it is imperative that the optimum rhamnose concentration for expression of the POI be determined for each new expression system. BW25113 *E. coli* cells were induced during the log phase of growth, at an OD_{600} of approximately 0.5, using three concentrations of rhamnose. MOMP expressing cells had lower OD_{600} values, under all rhamnose concentrations, in comparison to similarly treated control cells and also exhibited a marked difference in OD_{600} at the lowest rhamnose concentration of 0.002% compared to both 0.02% and 0.2%. This decrease in growth, observed between MOMP expressing cells with higher rhamnose concentrations, was indicative of toxic recombinant MOMP expression. Additionally, whilst the control cells showed increased cell density, in comparison to those expressing MOMP, there were still observable differences between the OD_{600} readings for control cells induced with different rhamnose concentrations, most dramatically after seven hours, with the lowest concentration of rhamnose producing the greatest cell density. This suggested that it was not only the recombinant protein expression that was toxic to the cells, but also the concentration of rhamnose used and, therefore, a balance between the two is required for optimal cell growth and protein expression.

Based on the observed effects in OD, during medium-scale expression testing, two additional concentrations of rhamnose were assessed, 0.02% and 0.1%, selected to encourage rMOMP expression whilst limiting rhamnose associated toxicity. The duration of protein expression also was limited to just three hours, as opposed to five or seven hours, due to the limited increase in OD_{600} observed between these two time points. In this experiment, all four cultures reached an OD_{600} of ~0.5 at the same time, although the OD_{600} values of the control cells at the time of harvesting were higher than those expressing MOMP, which again confirmed the toxicity of MOMP in this cell strain. Membranes from this medium-scale expression test were prepared and analysed with both Coomassie blue staining and Western blotting, where it was clear that the lower rhamnose concentration of 0.02% produced a higher quantity of rMOMP. Solely based on the OD_{600} readings taken at the time of harvesting, which did not differ significantly, it appeared that the level of cell toxicity caused by MOMP expression was equal in both cultures. However, as OD is really a measure of turbidity, deceased cells no longer growing or expressing protein may still result in a similar OD value to those that expressed protein more slowly, as is likely with the lower rhamnose concentration in this case. Under the higher rhamnose concentration, it was likely that rapid over production of toxic MOMP resulted in premature cell death and, hence, lower overall expression levels. Some cells may only have a small number of the required permease and, therefore, produce considerably less protein whilst also maintaining healthy cell growth and subsequently more sustained expression. Therefore, it is imperative that for each recombinant protein expression system, the rhamnose concentration is tested in order to identify optimal conditions. As a result of these optimisation experiments, for this expression system, BW25113 cells containing the pBADMOMP$_{H6}$ construct were induced with a rhamnose concentration of 0.02% (w/v) for optimum protein yield.

4.5.2 Brij-58 toxicity assay

Prior to complementation with FadL and MOMP, it was imperative that both the wild type and knockout *E. coli* strains were exhibiting the expected growth; the ability to grow on glucose/glycerol and LCFA carbon sources, or solely glucose/glycerol, respectively. Sodium palmitate, solubilised with detergent Brij-58, was used to characterise growth behaviour due to previously confirmed reports of *E. coli* growth with this LCFA (Hearn *et al.*, 2009). The sodium salt of each LCFA was used (as opposed to the free fatty acid form) in order to improve solubility; polarity is reduced by replacement of the hydrogen in the carboxyl tail with a sodium ion, although detergent Brij-58 also was still required to solubilise the LCFAs to the desired concentration.

A dilution series for each strain was created and spotted onto MM plates containing either carbon source. As expected, both the wildtype BW25113 strain and the isogenic *fadL* knockouts JW2341 7-3 and 8-3 grew with glucose. However, each dilution of the strains spotted onto the palmitate plate showed a region of translucent agar. Upon solidification, palmitate supplemented plates were opaque, as opposed to translucent MM plates containing glucose, which suggested that the appearance of translucent spots in the palmitate test was due to palmitate utilisation. However, as the JW2341 knockout strain lacked *fadL*, the gene required to transcribe the fatty acid transporter FadL which has been shown to uptake palmitate, it could be hypothesised that the Brij-58 used to solubilise palmitate produced a toxic effect on the cell. This toxicity may have perturbed the cellular membrane, resulting in cell lysis and release of fatty acid metabolising enzymes, such as FadE, FadB, and FadA. Alternatively, being a detergent, it is possible that excess Brij-58 (not required for palmitate solubilisation) was permeabilising the membrane, permitting the LCFA direct access to the periplasm and FadD within the inner membrane. An investigation

into the potential toxicity of Brij-58 was conducted on MM glucose plates with and without 0.5% (w/v) Brij-58, in which colonies grew significantly better (more numerous and larger) without the presence of Brij-58. To further examine this effect, a toxicity assay was employed to determine the effect of varying Brij-58 concentrations on cell growth. From growth of wild type cells over an eight hour period, it was evident that concentrations over 0.5% (w/v) produced a detrimental effect on growth. It was only after five hours of growth that the effects of Brij-58 became apparent, likely due to the slow growth rate already associated with MM. Cultures containing Brij-58 at 0.1% and 0.2% experienced a lag in growth around seven hours, but recovered within the following hour to an OD_{600} value comparable to the control cells not supplemented with Brij-58. This suggested that, at low concentrations, cells were able to overcome the potentially toxic effect of Brij-58. As Brij-58 was essential for fatty acid solubilisation, a trade-off between optimising solubility and limiting toxicity was required. It was established that a final concentration of 0.25% Brij-58 still permitted suitable solubilisation of 5 mM palmitate required for the growth assay, whilst limiting the associated cell toxicity.

4.5.3 Confirmation of *E. coli* knockout strains

Following the appearance of plaque-like regions in the palmitate plates, assumed to arise from Brij-58 toxicity, an additional mutant strain was acquired. This new mutant strain also contained a knockout for the *fadR* gene, called BW25113 Δ*fadL* Δ*fadR*. The *fadR* gene encodes the fatty acid metabolism regulator protein, FadR, which partially represses FadL expression through binding of the operon, as previously discussed (**Section 4.1.2**). It was hypothesised that this *fadR* knockout also may be required in order to encourage growth on LCFAs as a carbon source by release of the repression on FadL expression. Following Brij-58

toxicity assays, this new mutant, in addition to the isogenic JW2341 mutants, were examined on glucose and palmitate plates to determine their growth behaviour. Cultures also were plated to a lower dilution of 10^{-7} in order to identify individual colonies. As expected, all strains grew with glucose as a carbon source, with no adverse effects observed with knockout mutants in comparison to the wild type, with individual colonies visible at a 10^{-6} dilution. However, the spots of translucent agar were still present in both the BW25113 and JW2341 strains, but not the new double knockout strain BW25113 Δ*fadL* Δ*fadR*. At the lower dilutions, namely 10^{-5} and 10^{-6}, it appeared that the translucent regions present in BW25113 and JW2341 spots were in fact due to colony formation and palmitate utilisation, as opposed to cell lysis or membrane disruption, as previously hypothesised. It was not anticipated that the JW2341 knockout strain would be able to grow on palmitate, due to the lack of fatty acid transporter protein FadL, however, due to the lack of growth observed with the double mutant BW25113 Δ*fadL* Δ*fadR* under the same conditions, it was hypothesised that knockout of both the *fadL* and *fadR* genes was essential for the prevention of fatty acid uptake. However, as FadR is a repressor, and no other LCFA transporter has been identified in *E. coli*, it was more likely that the JW2341 knockout was not correctly characterised, and perhaps only possessed a partial *fadL* mutant as opposed to a knockout. Both the wildtype and mutant BW25113 cells also were spotted onto plain MM, Brij-58, and rhamnose plates to determine if either of the strains were able to utilise the additional components of the assay as carbon sources. Whilst the BW25113 knockout strain mutations had not been formally characterised, its growth behaviour in the screens indicated that it was unable to utilise LCFAs or any other component in the plates as a carbon source, an essential trait for this functional assay and, therefore, this strain was taken forward for complementation with FadL.

To confirm the sustained knockout of the two genes, the mutant BW25113 strain contained two antibiotic resistance markers; chloramphenicol and kanamycin. The pBADrha vector expressing MOMP, which was to be complemented with this knockout strain in later experiments, also contained a chloramphenicol resistance marker and, therefore, it would not have been possible to select for successfully transformed cells. Instead, an ampicillin gene was inserted into the pBADrha vector. Within this vector, a fragmented ampicillin resistance gene was discovered and, therefore, the terminal ends of this region were used to insert the whole gene. In doing so, no other components within the vector were disturbed and newly transformed In-Fusion products were able to grow in the presence of ampicillin, confirming the correct insertion.

For the complementation control, the *fadL* gene was obtained from wildtype BW25113 genomic DNA and inserted into the newly created ampicillin resistant pBADa vector. It was anticipated that complementation of the knockout strain with the pBADaFadL construct would reinstate LCFA transport and permit growth on palmitate plates. Cells were grown as before and induced for three hours before spotting onto palmitate plates with and without rhamnose supplementation, to encourage FadL expression in subsequent colonies. However, transformed cells only grew in the presence of glycerol, which was used instead of glucose to avoid catabolite repression. As FadL is a native *E. coli* protein, this was unexpected as the cells possessed all the necessary mechanisms required for folding and export of FadL to the OM. Thus, it was hypothesised that induction with rhamnose resulted in over expression of FadL, which then became toxic and prevented growth, and therefore uninduced cells were also plated, as it was expected that basal expression would likely produce enough protein to observe growth. However, again growth was not observed in palmitate plates. As the knockout genes of this strain were unconfirmed, it was possible that

an additional detrimental effect was introduced, and therefore a new knockout mutant was acquired.

The new double knockout mutant named LS6164 had been characterised previously (Ginsburgh *et al*., 1984) and used in several studies investigating LCFA transport (Kumar and Black, 1993; van den Berg, 2004; Hearn *et al*., 2009), but again was plated on glycerol and palmitate plates to confirm the expected growth patterns. Originating from K12, it was necessary to repeat rhamnose concentration testing again, due to potentially different permease levels within the cytoplasm. This was first investigated using a growth curve, as before, (described in **Section 4.4.4**) whereby lower concentrations were tested due to the toxicity observed previously with concentrations over 0.1%. Transformed cells, not induced, grew at a slightly lower rate than control LS6164 cells, suggesting that some basal expression was occurring. Rhamnose concentrations of 0.05%, 0.01%, and 0.02% all grew in a similar manner, with no increase in OD_{600} observed after six hours, suggesting that MOMP expression was toxic to cell growth. As a quick measure of protein expression, whole cells were pelleted one, three, and five hours post induction for each rhamnose concentration, which were then analysed via Western blotting. As membranes were not extracted and only a small volume of cells were pelleted, only faint bands corresponding to MOMP were observed (see **Figure 4.8**). No detectable protein was expressed at below 0.005% rhamnose, with samples induced with 0.02% rhamnose harvested after three hours producing the strongest band, as before. Therefore, for this strain the same rhamnose concentration was used as before.

Upon acquisition of the more robust double knockout mutant LS6164, pBADaFadL complementation again was evaluated with both non-induced cells and cells induced for three hours, with the latter also spotted onto rhamnose supplemented palmitate plates to

encourage FadL expression in second generation cells. Complementation with pBADaFadL reinstated the cells' ability to transport LCFAs in all instances, except where rhamnose was included within the plate. In this case, even wild type cells struggled to grow, indicative of rhamnose associated toxicity, the effect of which was likely enhanced in cells that also were induced, indicating why no growth was observed in this plate for complemented cells. The success of pBADaFadL complementation with LS6164, along with the strain's inability to grow with palmitate as a carbon source, provided a robust assay for assessing the ability of MOMP to transport LCFAs.

4.5.4 MOMP complementation to assess LCFA transport

Following successful development of the functional assay with the control protein, FadL, pBADaMOMP$_{H6}$ was transformed into knockout LS6164 strain and examined for growth with fatty acids palmitate and oleate. In the first instance, transformed cells were not induced, instead permitting basal expression to occur to allow low levels of MOMP expression, unlikely to result in overloading of the OM. However, colonies did not form under these conditions, nor in cultures induced for varying durations, nor at a lower rhamnose concentration, again to limit expression levels and prevent toxicity.

Previous SRCD experiments on purified MOMP indicated an increase in thermal stability in the presence of both LCFAs oleate and palmitate (Atanu, 2014). However, this functional growth assay suggested that MOMP was unable to transport these ligands to support *E. coli* growth. It is therefore possible that instead of binding to the hatch domain (the external loop predicted to guide substrate entry) in preparation for transport, these two ligands were solely attracted to and interacting with the hydrophobic regions of MOMP's structure, resulting in the increase in stability, as seen in an increase in Tm.

One of the most important considerations for success of the growth assay was in the correct folding and export of rMOMP to the *E. coli* OM in order for fatty acids to be potentially transported and delivered to FadD at the inner membrane. Previously, C41 (DE3) *E. coli* cells transformed with pETMOMP$_{H6}$ successfully expressed rMOMP in the membrane, as shown through both the purification procedure (**Chapter 3**), whereby a sarkosyl isolation step indicated that the majority of MOMP was found in the insoluble outer membrane fraction, and STORM imaging (**Chapter 5**), which revealed fluorescence localisation to the external perimeter of the cell. Therefore, this expression system also was analysed with super-resolution STORM to visualise the cellular location of rMOMP. From comparison with the control *E. coli* protein OmpA, it was apparent that rMOMP was localised at the membrane region of the cell, as opposed to within the cytoplasm as inclusion bodies, which would present as fluorescence in the central region of the cell. Previous research using electron microscopy determined the width of the periplasmic space in Gram negative bacteria to be between 13-25 nm (Seltmann and Holst, 2002) and, as the maximum resolving power of STORM is 20 nm, it was not possible to resolve the difference between fluorescently labelled inner and outer membranes (data not published). However, MOMP does not possess the appropriate hydrophobic α-helices required for IM insertion (Schulz, 2002), with less than 5% of native MOMP's secondary structure content attributed to helices, as analysed with CD (Sun *et al.*, 2007). As a result, if MOMP was incorrectly folded and unable to complete export to the OM, it would be present within the cytoplasm as inclusion bodies and not identifiable at the membranous region of the cell. A low concentration of detergent Triton X-100 was used during sample preparation, in order to permeabilise the outer membrane and permit access of the fluorophore to the internally exposed his-tag. Due to such permeabilisation, any MOMP expressed within the cytoplasm would also be amenable

to labelling, although this was not observed. The absence of cytoplasmic inclusion bodies, in addition to the membrane localisation observed using STORM, strongly suggests that MOMP was located at the OM region of the LS6164 cells. Additionally, it must be noted that whilst expressed at the membrane, MOMP does not display homogeneous dispersion as seen in OmpA. It has been hypothesised that MOMP forms clusters within the *E. coli* membrane due to the large number of cysteine residues available for intermolecular disulphide bonding; this is investigated in greater detail in **Chapter 5**.

Biotinylation, commonly used in protein purification, has been used to label surface exposed loops of OM proteins (Voss and Cover, 2015), along with trypsin digestion of insoluble membrane pellets for analysis with mass spectrometry (MS) (Mirza *et al.*, 2007) and MAb binding assays (Baehr *et al.*, 1988). The low specificity of trypsin and the high cost associated with the production of MAbs has highlighted biotinylation as an alternative method to STORM imaging for confirmation of MOMP's insertion into the OM. Previously, biotin was deemed membrane permeable and required supplementation with oligosaccharides, such as stachyose, in order to limit the reaction to the bacterial membrane (Myers-Morales *et al.*, 2007), which is explicitly required in this research to confirm MOMP's external exposure. However, more recently, negatively charged and water soluble long-chain reagents such as Sulfo-NHS-LC-LC-Biotin (ThermoFisher) have been developed for cell surface labelling. Due to the large number of cysteines in MOMP, it is possible that these residues could be used as biotinylation targets, namely towards C52, C56, and C226 located near the top of the barrel (based on the low resolution structure shown in **Chapter 1 Figure 1.8**), which are the most likely residues to be externally exposed. Alternatively, biotinylation could be targeted towards free lysine residues, such as K161 located within the external loop region of MOMP. However, whilst residue specific, these interactions are not protein

specific and, therefore, any array of lysine or cysteine containing *E. coli* proteins would also become labelled with biotin. Extraction of membranes for Western blotting, as previously conducted in this research, would identify biotinylated proteins through labelling with streptavidin HRP conjugated antibody. However, it would be essential to include a biotinylated *E. coli* control sample, not containing MOMP, in order to identify any ~40 kDa proteins, such as OmpA (37 kDa), OmpN (39 Kda), OmpC (38 kDa) or HofQ (42 kDa) (Molloy *et al*., 2000), any of which also may become biotinylated and thus mask observation of MOMP. Should such a protein be detected within the control at this location, the band within the test sample would require excision and analysis with MS in order to identify MOMP. Whilst MOMP is cysteine rich, due to *Chlamydia*'s unique biology, it has been hypothesised that many of these residues form intermolecular disulphide bonds and, therefore, would not be freely available for biotinylation, particularly as the addition of reducing agents was not demonstrated to improve MOMP solubilisation (**Chapter 3**). As it is unknown which cysteine residues form these interactions and therefore which would be available for biotinylation, in addition to the high likelihood of *E. coli* protein contaminant detection through lysine targeted biotinylation, this method was not adopted, with STORM imaging favoured. A higher resolution structure characterising the residues situated within the external loop regions, not previously modelled into the 4 Å structural data due to poor density in this region (**Chapter 1 Section 1.6.5**), would facilitate the design of an antipeptide antibody to this surface exposed region, leading to a simpler detection method.

4.6   Conclusion

Since the super-resolution STORM images indicated that MOMP was likely correctly inserted into the OM, based on the functional assay it appeared that the recombinant MOMP expressed from the pBAD vector was unable to transport LCFAs palmitate and oleate, evidenced through lack of growth with these substrates as a sole carbon source. Alternatively, it is possible that MOMP either requires a partner protein in order to initiate the conformational change required within the plug domain (for binding and transport) in the absence of a lateral hole, or that MOMP is able to transport the LCFAs but is then unable to communicate with FadD, the *E. coli* protein required for inner membrane transport, leaving the carbon source trapped within the periplasmic space and inaccessible to the cytoplasmic β-oxidation proteins. It may be possible to test this hypothesis through fluorescent labelling of substrates, however, due to both the difficulties associated with distinguishing the inner and outer membranes with fluorescent microscopy, as well as the requirement for complex three-colour imaging, this method is likely to remain unfeasible until higher resolutions are achievable.

It has been demonstrated that *Chlamydia* is able to scavenge host-derived fatty acids such as oleic acid and palmitic acid (Yao, Cherian *et al*., 2015), most likely in order to limit the energy expenditure in intensive pathways such as phospholipid synthesis. As *Chlamydial* LOS is a truncated form of LPS (Rund *et al*., 1999), it may be possible for the LCFAs to 'flip-flop' across the outer membrane due to a reduced lipid barrier, however, it is also possible that a transport protein is involved in translocation of host derived fatty acids from the inclusion lumen to the OM and perhaps has a role in MOMP transport activation. MOMP's role in LCFA uptake could be investigated in *Chlamydia*, in a similar fashion to previous studies used to investigate uptake of radiolabelled host-derived LCFAs in *E. coli* (Lepore *et*

al., 2011), whereby mutations introduced into the lumen of MOMP's β-barrel to encourage a water-filled pore, as observed in *E. coli*'s OmpF, would restrict the potential transport of hydrophobic substrates such as LCFAs.

Given the increase in expression of MOMP during the EB stage of the *Chlamydial* lifecycle (Roulis *et al.*, 2013), in addition to the high number of cysteine residues, it is possible that MOMP's sole role is in providing structural stability within the COMC, although some typical porin activity has previously been reported (Jones *et al.*, 2000; Sun *et al.*, 2007). However, whilst down regulated in RBs, MOMP is still expressed during this metabolically active stage which is suggestive of a transport role. Additionally, it has been hypothesised that a number of the COMC proteins exhibit additional functions, such as porin PorB and the Pmps as adhesins, which also suggests that MOMP may have a dual function. Based on our current low resolution structure of MOMP, suggesting the occluded region at the periplasmic side of the lumen (**Chapter 1 Section 1.6.5**), it is highly likely that any ligand transport will require conformation changes in the plug to permit passage. Developing a greater understanding of both *Chlamydial* biology and MOMP's structure will provide further insight into the role of the abundant proteins of the COMC in cell proliferation and survival.

Chapter 5

Analysis of MOMP's clustering behaviour using fluorescence microscopy

5   Analysis of MOMP's clustering behaviour using fluorescence microscopy

5.1   Introduction

As discussed in **Chapter 1 Section 1.4**, all species of *Chlamydia* possess a COMC in order to compensate for the lack of PG between the inner and outer membranes in EBs. As MOMP is both cysteine rich and the most abundant of the 17 proteins recently identified in the COMC (**Chapter 1 Table 1.1**), it is likely to play a significant role in intermolecular disulphide bond formation. The lack of high resolution structural data for any of these OMPs makes mapping the potential interactions extremely difficult, and as a result there is currently a paucity of knowledge in this area. Whilst bioinformatics analysis of each of the 17 proteins could be pursued, homology modelling for each protein would not only be extremely time consuming, but also any slight variation from the true structure could adversely alter the position of vital cysteine residues, resulting in miscalculations with regard to the significance of any particular residues.

Aside from the two published homology models for MOMP (Atanu *et al*., 2013; Feher *et al*., 2013), only one other COMC protein homology model exists, which is for PorB (Pourhajibagher and Bahador, 2016), further highlighting the difficulty in producing reliable models for these proteins. Additionally, as mentioned previously, the two homology models for MOMP were produced in the same year and therefore both groups of researchers were likely to have access to the same or similar software. Despite this, the two models still demonstrate significant variability, with one suggesting a 14-stranded β-barrel (Atanu *et al*., 2013) and the other a barrel of 16-strands (Feher *et al*., 2013), which in turn would predict entirely different locations for the nine cysteine residues. However, these two models are based on *C. pneumoniae* and *C. trachomatis* MOMP, respectively and this, therefore, may

account for the differences observed. However, until high resolution crystal structures emerge for each COMC protein, homology modelling using a variety of techniques is likely the most promising *in silico* method for the identification of cysteine rich regions within the barrel, which can then be rationally targeted in mutagenesis experiments.

Some experimental research has been conducted for *C. trachomatis* MOMP, in an effort to identify the cysteine residues likely to be involved in disulphide bonding, however, these reports appear to be conflicting. One group has predicted that cysteine pairs C48-C55 (corresponding to C49-C56 in *C. pneumoniae* MOMP, as analysed with BLAST) and C201-C203 form intramolecular disulphide bonds, leaving the remaining cysteines free and available for intermolecular disulphide bonding (Yen *et al*., 2005). The following year, however, a different group also investigating *C. trachomatis* MOMP inferred that C26-C337 and C116-C208 (corresponding to C49-C353 and C136-C226 in *C. pneumoniae* MOMP, respectively) were forming intramolecular disulphide bonds (Wang *et al*., 2006). It seems highly unlikely that residues C201-C203 would form an intramolecular disulphide bond due to their close proximity and likely positioning on the same β-strand, which would not serve to enhance the stability of the barrel and, therefore, form a redundant interaction. Additionally, intramolecular disulphide bonding between C136 and C226 also seems improbable due to their location on opposite sides of the β-barrel, according to our low-resolution *C. pneumoniae* rMOMP crystal structure, although, a potential (albeit significant) frame shift of the β-strands could make this sterically feasible. The method employed in both studies involved trypsin digestion of MOMP and analysis of the fragments using mass spectrometry. Through the creation of fragments, it is feasible that cysteine residues previously not sterically available for intramolecular disulphide bonding (when localised in the β-barrel) then became exposed to one another. As a result of these apparent

discrepancies and as the method used for analysis of potential disulphide formation is questionable, and the research was not conducted on *C. pneumoniae* MOMP, these results did not heavily influence the choice of cysteine mutants selected in this study.

A low-resolution crystal structure of 4 Å, found using molecular replacement with the 14-stranded MOMP homology model (Atanu *et al*., 2013), which although incomplete (with the absence of the external loop region), contains all nine cysteine residues allowing more astute predictions to be made on the potential residues involved in intermolecular disulphide bonding between other MOMPs. Due to the high number of cysteine rich proteins contributing to the COMC, which between them contain 194 cysteine residues, it would be an extremely energy intensive process to locate the specific and desired cysteine residue for each protein amongst the milieu of proteins inside the periplasm that are awaiting oxidation and export to the OM. Additionally, the mechanism by which disulphide bonds are formed in *Chlamydia* has yet to be established and, so far, only one enzyme, the oxidoreductase DsbH, has been characterised (Mac *et al*., 2008). As the exact interactions between the proteins comprising the COMC (as shown in **Chapter 1 Table 1.1**) have not yet been established, predictions made, based on relative protein abundance, can be useful in deciphering the likely interacting proteins. Due to the lack of information regarding the COMC, a novel hypothesis is herein proposed whereby COMC proteins possess cysteine rich regions or 'pockets' that act as general target regions for intermolecular disulphide bond formation. As MOMP is the most abundant of the COMC proteins, it seems highly likely that much of MOMP's intermolecular disulphide bonding will occur with other MOMPs, with some cross linking to additional COMC proteins likely in order to form a complete network within the protein mesh. Therefore, in this chapter the hypothesis that rMOMP expressed in *E. coli* will display different OM localisations, depending on the cysteine residues present,

will be investigated through fluorescence microscopy. In native MOMP, it is expected that clusters will form due to intermolecular disulphide bonding, as opposed to cysteine mutated MOMP which will disrupt disulphide bonding and result in a more homogenous dispersion. Molecular dynamics simulations in conjunction with the low resolution crystal structure will be used to guide cysteine mutant design.

5.1.1 Molecular dynamics simulations

Molecular dynamics (MD) is a computer simulation technique developed in the 1970s utilised to assess the behaviour of biological molecules under certain conditions (Hospital *et al*., 2015). Since then, due to the advancement of both computers and key algorithms, calculations have dramatically improved from simulations of just ~300 atoms to well over 75,000 now, permitting the modelling of large macromolecular complexes and membrane proteins within a lipid environment. As only a low resolution 4 Å crystal structure is available for *C. pneumoniae* MOMP, it is possible that a higher resolution structure may reveal variations within the side chain orientations, which are most accurately characterised at resolutions above 2.5 Å. Due to this, atomistic modelling is inappropriate at this stage as this method employs a high level of accuracy, parameterising each atom within the molecule to detail the precise dynamic behaviour of the protein and, additionally, is extremely computationally expensive. Alternatively, coarse grained (CG) modelling offers a more simplistic representation, with 4-5 heavy atoms (excluding hydrogen atoms) grouped together as one particle, simplifying the system and greatly reducing the computational time and expense required to run a biologically significant (~10 μs) simulation (Bond *et al*., 2007). In this instance, MARTINI force-field equations (Marrink *et al*., 2007), derived from thermodynamic data, will be used in order to calculate the particles' potential energy and

group atoms together. As MARTINI has a special focus on lipid parameterisation, the behaviour of polar and non-polar environments has been well characterised and, therefore, this CG model system is particularly appropriate for simulating membrane proteins within a lipid bilayer.

An asymmetric lipid bilayer consisting of dipalmitoyl phosphatidylethanolamine (DPPE) and dioleoylphosphatidylglycerol (DOPG), the chemical and CG structures of which are shown in **Figure 5.1**, will be used in order to mimic the *Chlamydial* membrane as closely as possible.

As discussed previously in **Chapter 4 Section 4.1.1**, *C. trachomatis* was demonstrated to incorporate oleic acid (18:1) into the *Chlamydial* membrane, as well as lauric (C12:0), myristic (C14:0) and palmitic acids (C16:0) derived from the host (Yao, Dodson *et al.*, 2015). As the MARTINI lipid database does not explicitly include *Chlamydia*'s unique, truncated LPS (LOS), the two lipids that were selected, based on their chain lengths, are C16:0 for DPPE and C18:1 for DOPG, corresponding to both palmitic and oleic acid and to the reported lipid chain lengths on LPS, which correspond to chains of C14-16, C18, and C20 (Kosma, 1999) in an effort to provide a simplified *Chlamydial* bilayer.
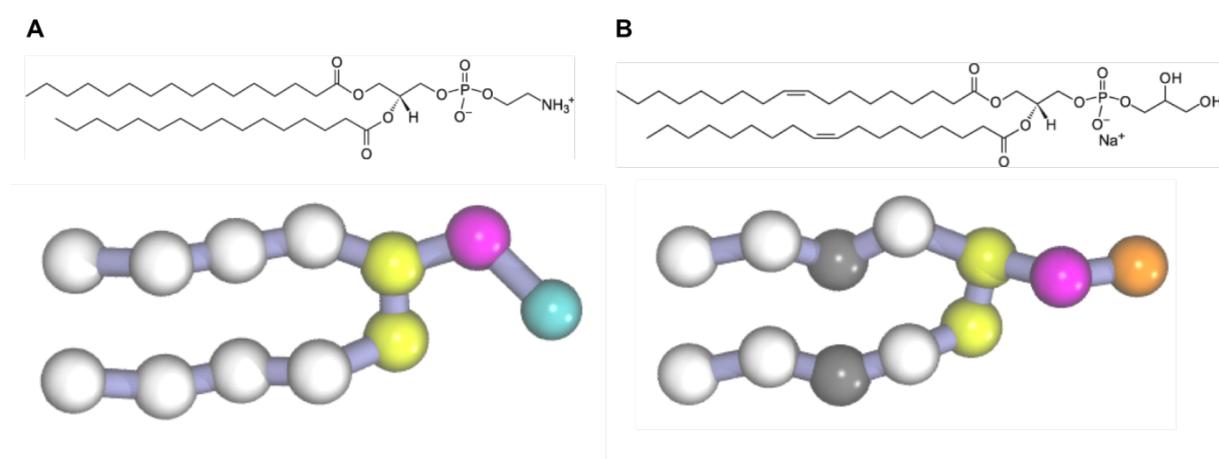
Figure 5.1. **Chemical and MARTINI CG structures for A, DPPE and B, DOPG.** Four heavy atoms, including their respective hydrogens, are incorporated into a CG particle. White particles, carbon backbone; black, carbon backbone with double bond; yellow, carboxylic acid; pink, phosphate; cyan, amine; and orange, hydroxyl.

To create the DPPE/DOPG lipid bilayer, the tool Insane (Wassenaar *et al*., 2015), an acronym for 'insert membrane', will be utilised. Insane implements CG lipid templates, such as those created by MARTINI, to create simple custom membranes in which membrane proteins can be embedded. Insane also allows the protein's membrane orientation to be shifted based on their defined Z location, permitting simulation of partially embedded proteins. The default for this system assures that the centre of the protein at Z=0 reflects alignment with the membrane, which will be implemented for MOMP modelling. Additionally, Insane is used to generate the solvent content of the system, in this instance solely consisting of water and sodium ions. The final MD simulation will be run in Gromacs (Abraham *et al*., 2015), a highly optimised software package used to simulate the Newton equations of motion for a defined system.

 The aim of these MD simulations is, firstly, to determine if two MOMP crystal structures inserted within an asymmetric lipid bilayer at a suitable distance and with multiple orientations will associate together after a duration of 10 µs and, secondly, if any associations are identified, to characterise the interacting residues and to determine if any are due to intermolecular disulphide bonding. It is intended that this *in silico* analysis will provide insight into which cysteine residues are responsible for or are most significant in MOMP clustering, which will help guide the planned mutagenesis studies for experimental research into MOMP's behaviour within the *E. coli* OM.

5.1.2   Fluorescence Microscopy

5.1.2.1 Epifluorescence and total internal reflection fluorescence (TIRF) microscopy

Since its first use, microscopy has been an important tool for the visualisation of physiological and morphological changes within cells. With technological advancements

come increasingly higher spatial and temporal resolutions, with fluorescence microscopy allowing us to selectively follow specific targets, such as proteins (Huang *et al.*, 2009). Fluorescence is a property typically associated with aromatic hydrocarbons, also known as fluorophores. Fluorescence occurs in a three-step process. Initially, a photon of energy is supplied to the fluorophore by a light source, such as a laser, the absorption of which results in an excited state. This excited state is short lived, during which the conformation of the fluorophore becomes altered and dissipates some energy, before the fluorophore emits a photon, returning itself to the ground state. Finally, as some of the energy has been dissipated, the emitted light has a longer wavelength, compared to the first stage of the absorption; this difference is referred to as the Stokes shift (Sanderson *et al.*, 2014).

Epifluorescence microscopy is typically perceived as the conventional method for visualising fluorophores and is often used to gather preliminary data on fluorophore labelling methods, prior to more technically advanced procedures such as confocal microscopy. In epifluorescence, the incident light illuminates the entire sample, as shown in **Figure 5.2A**, however, in doing so also captures the out of focus plane leading to a low signal-to-noise ratio. Given the high background signal, the limit of resolution is only approximately 200 nm and, therefore, two fluorophores within the same cell are extremely unlikely to be spatially resolved due to the small size of bacteria.
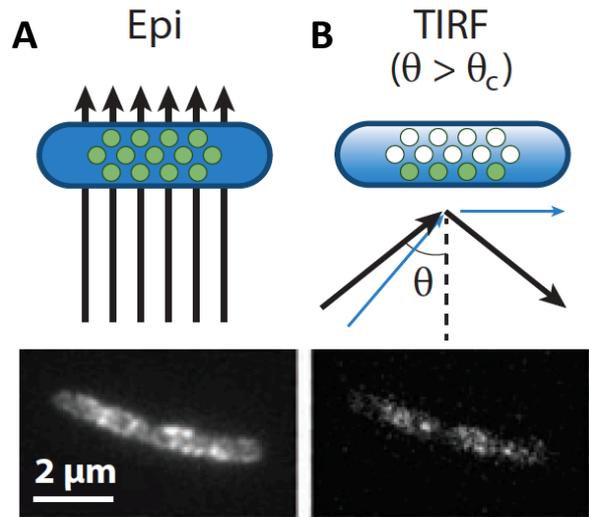
Figure 5.2. **A, Epifluorescence and B, TIRF microscopy.** Blue region, area of illumination within bacterial cell; green circles, excited fluorophores; white circles, non-excited fluorophores; black arrows, incident laser angle (θ); blue arrows, critical angle (θ_C). Bacteria shown are *Bacillus subtilis* labelled with green fluorescent protein (GFP). **A**. In epifluorescence mode, θ is approximately 0° and the cell is excited equally throughout. **B**. In TIRF mode, where $\theta > \theta_C$, the light is completely reflected creating an evanescent wave that illuminates molecules close to the cell surface. Adapted from Yao and Carballido-López, 2014.

As MOMP is situated within the bacterial OM, total internal reflection fluorescence (TIRF) microscopy will be used in complement with epifluorescence, due to the high background signal associated with the latter. TIRF is a near-field method of illumination, used to visualise fluorescently labelled molecules on the outer surface of cells. When a sample is exposed to a laser at a critical angle (θ_C), the differing refractive properties of the glass coverslide and the sample in buffer results in complete reflection of this light, which in turn propagates an exponentially decaying evanescent wave perpendicular to the coverslide that only illuminates the fluorophores situated within 60-100 nm of the glass/sample interface (Yao and Carballido-López, 2014). Unlike in epifluorescence, only a portion of the available fluorophores are illuminated within one focus plane, as shown in **Figure 5.2B**. This reduction in background fluorescence means that a resolution of at least 100 nm can be achieved with

this method. Whilst TIRF is mainly performed on mammalian cells, it is also possible to visualise bacterial cells using this method, such as *Bacillus subtilis* (Domínguez-Escobar *et al.*, 2011) and *E. coli* (Jovanovic *et al.*, 2014), despite being over ten times smaller at just 2 μm in length. Possibly due to their small size, methods for bacterial cell preparation for TIRF are underrepresented in the literature and, therefore, in this research a new method will be developed. TIRF microscopy will be used in this research to obtain preliminary experimental data on MOMP's membrane distribution whilst also optimising the imaging conditions including cell number, cell treatments, and fluorescent probe labelling. This knowledge will help ensure optimal imaging conditions for a higher resolution technique (see below), with the aim to obtain quantitative data for the analysis of MOMP clustering within the *E. coli* OM.

5.1.2.2 Stochastic optical reconstruction microscopy (STORM)

In order to acquire not only higher resolution images but also to obtain quantitative data for clustering analysis, stochastic optical reconstruction microscopy (STORM) will be adopted. In conventional epifluorescence microscopy, it is difficult to localise a large number of fluorophores within close proximity, which results in low resolution images. STORM is a fluorescent microscopy method based upon single molecule localisation, whereby the use of photoswitchable organic fluorophores for labelling is essential (Godin *et al.*, 2014). These fluorophores can be controlled with exposure to low intensity light and a reducing agent, creating an on/off cycling state, also known as 'blinking', whereby, at any given time, only a sparse subset of fluorophores are in the on state and, therefore, are not overlapping. Collection of a large number of frames, over 15,000, captures the localisation of each fluorophore within the sample for which the centroid positions can be calculated and the

positions reconstructed into a high-resolution image, as shown in **Figure 5.3**. As a result, the

resolution of STORM is not limited by diffraction, but rather the measured localisation

precision (Bates *et al.*, 2013) and is reported to be as good as 20 nm (Rust *et al.*, 2006). It is

not uncommon in STORM imaging for the same fluorophore to be present in a number of

frames due to the on/off cycling times. For example, previously calibrated data for Alexa-

647 shows that this fluorophore remains in the on state for five frames and the off state for

ten frames (Willson *et al.*, 2016). Therefore, in post imaging analysis such data, known as

trails, can be grouped according to the fluorophore from calibrated on/off times and known

capture radii. Concurrent with STORM's development (Rust *et al.*, 2006), additional research

groups established photoactivated localisation microscopy (PALM) (Betzig *et al.*, 2006) and

fluorescence photoactivated localisation microscopy (FPALM) (Hess *et al.*, 2006). All three

methods adhere to the same principle, although the term STORM will be used here.
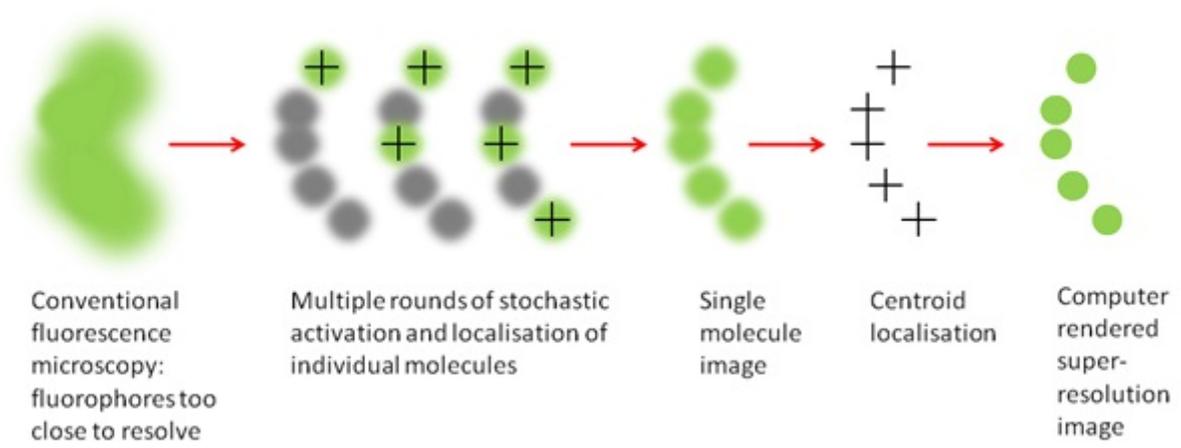
Figure 5.3. **The principle of STORM.** The use of photoactivatable fluorophores in STORM means that, at any one point, only a small subset of fluorophores are in the on state at one time. Collection of a large number of images of these blinks results in a single molecule image. The resolution can be further improved by identifying the centroid location of each fluorophore, resulting in a computer rendered super resolution image. (Thorley *et al.*, 2014)

It is anticipated that with super-resolution STORM, valuable quantitative data can be

obtained for the clustering behaviour of both native and cysteine mutated MOMP within

the *E. coli* OM. By determining the cysteine mutants that most significantly disrupt or alter MOMP's membrane localisation, insight can be gained into how MOMP, a major component of the *Chlamydial* COMC, may form these associations, enhancing our knowledge on how *Chlamydia* maintains its structural stability.

5.1.2.3 Fluorophore labelling methods

Within MOMP, both the N and C termini are located at the periplasmic side of the outer membrane, with the N terminus likely concealed inside the plug region of the β-barrel. Therefore, MOMP is labelled with a C-terminal his-tag in the pET vector, which is exposed to the periplasm as opposed to the extracellular side and as a result, for fluorescence imaging in *E. coli,* the fluorophore requires access to the internal side of the outer membrane. The novel P3NTA tag, used previously for detergent screening with FSEC in **Chapter 3 Section 3.1.2**, which binds to the his-tag will be tested as a one-step labelling method in initial epifluorescence and TIRF experiments. Additionally, Alexa-647 is currently one of the most widely used fluorophores for STORM due to its high photostability, strong fluorescent signal and good blinking activity (Dempsey *et al*., 2011). However, labelling of pETMOMP$_{H6}$ with both the P3NTA tag and anti-his Alexa-647 will require permeabilisation of the outer membrane with a low concentration of detergent, such as Triton X-100, which may disturb the structural integrity of the membrane.

In order to overcome this, a construct consisting of the pOPINe-3C-Halo vector containing the *C. pneumoniae* AR39 MOMP gene (**Appendix A2.2**) was created using the In-Fusion method (**Chapter 2 Section 2.2.1.6**). The Halo-tag is a 33 kDa protein derived from a bacterial haloalkane dehalogenase from the *Rhodococcus* species that covalently and irreversibly couples to Halo-tag ligands, such as HaloTag silicon rhodamine (SiR) or HaloTag

Alexa-488 (Promega) and permits fluorescent imaging (Los *et al.*, 2008). The highly stable covalent bond is achieved through mutation of the His272 residue, located within the active site, to a phenylalanine, which prevents the hydrolysis and release of the intermediate ester (England *et al.*, 2015). **Figure 5.4** displays a simple representation of the HaloTag protein bound to the HaloTag ligand. In lieu of a fluorescent ligand, the Halo-tag ligand can also be linked to resins for isolation of the protein during purification. Similar to the position of the tag within the pETMOMP$_{H6}$ vector, the HaloTag gene was inserted at the C-terminus of MOMP and, therefore, Halo-tagged near-infrared SiR is of particular interest as a fluorescent probe, in this case. This is due to SiR's membrane permeability meaning that detergent treatment for permeabilisation of the cell is not required, therefore, preserving membrane integrity.
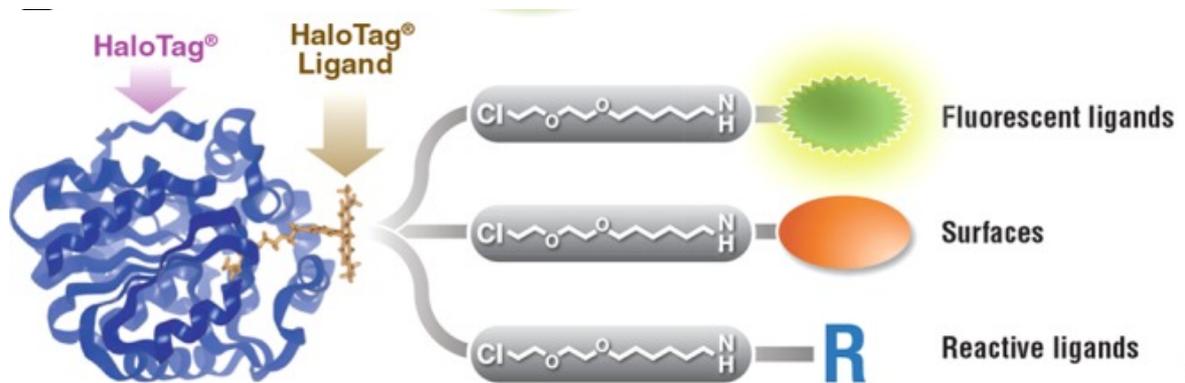
Figure 5.4. **The HaloTag principle.** The HaloTag is derived from a haloalkane dehalogenase, mutated at residue 272, within the binding region, in order to permit irreversible covalent bonding of the HaloTag ligand. The HaloTag ligand can be coupled with a number of functional groups for fluorescent imaging or protein immobilisation and purification. (England *et al.*, 2015)

5.2 Aims and objectives

E. coli, being a Gram negative bacterium that is easy to transform and grow, is an ideal model organism for this research. Not only is *C. pneumoniae* classed as a category two pathogen, but due to its unique biphasic lifecycle the EB cells containing the COMC require isolation from the RBs, which in addition to the difficulties acknowledged in mutagenesis studies in *Chlamydia* makes research into natively expressed MOMP extremely challenging. With the support of MD simulations to aid identification of the key cysteine residues involved in intermolecular disulphide bonding and MOMP clustering, cysteine mutants will be designed initially in the pETMOMP$_{H6}$ construct for analysis with epifluorescence and TIRF. These two lower resolution fluorescence microscopy techniques will provide preliminary information regarding the sample preparation and imaging conditions and will aim to indicate any initial qualitative differences observed between native MOMP and the derived cysteine mutants. As the HaloTag-SiR fluorophore is not commercially available and requires a time consuming and costly synthesis, the pOPINHaloMOMP mutant constructs will be created following initial epifluorescence and TIRF experiments, once the viability of this method to explore the clustering hypothesis has been confirmed; that native MOMP forms clusters within the *E. coli* OM and that the cysteine residues likely to be involved in disulphide bond formation increase homogenous dispersion. Ultimately, STORM will be used to quantitatively compare native and cysteine mutated MOMP in order to determine if disruption of intermolecular disulphide bonding occurred.

Very little is currently understood about the COMC and the interactions that occur between the cysteine rich proteins, or even the mechanism *Chlamydia* employs to form disulphide bonds. Therefore, this research aims to provide novel insight into the clustering behaviour of MOMP in bacterial cells.

5.3   Methods

The following methods described herein relate specifically to the preparation of super-resolution microscopy slides, which were not used in data collection. The standard and finalised protocols for MD simulations, TIRF, epifluorescence, and STORM microscopy are detailed in **Chapter 2 Section 2.2.5**.

5.3.1   Fluorescence microscopy

5.3.1.1 Superfrost microscope slide preparation and immunostaining

Cells harvested at 1h and 3h, stored in 1.5 ml microcentrifuge tubes, were washed once with PBS with centrifugation at 8,000 rpm (3,500 x g) for 3 minutes before being fixed in 1% (v/v) formaldehyde-PBS for 30 minutes. Fixed cells were then washed three times with PBS and stored overnight at 4 ˚C. On the day of imaging, only the pET constructs were incubated with 0.1% (v/v) Triton X-100 for 5 minutes, before all cells were incubated with 100 µg/ml lysozyme in PBS for 45 minutes, with wash steps between each treatment. Cells were then incubated with either 50 ng/ml HaloTag-SiR fluorescent dye (kindly gifted by the Eggeling group at the Weatherall Institute of Molecular Medicine, Oxford) for pOPINHaloMOMP constructs or a 1/100 dilution of Alexa-647 anti-his antibody for pETMOMP$_{H6}$ constructs for between 30 minutes to 1h in the dark, before washing three times with PBS. A control sample of C41 (DE3) cells were not stained, in order to check for autofluorescence. A precleaned 76x26 mm superfrost microscope slide (ThermoScientific) was incubated with 0.01% (v/v) poly-L-lysine-PBS for 10 minutes before being washed with PBS and left to dry. Once dry, cells were attached to the slide for 1h before being washed once with PBS. Cells were then post fixed with 2% (v/v) formaldehyde-PBS for 10 minutes, before three washes with PBS. 10 ul of 0.1 M DTT was pipetted onto the slide and a 0.17 mm (± 0.005 mm)

coverslip (Zeiss) overlaid, which was then sealed with clear nail varnish to prevent evaporation.

5.4   Results

5.4.1   Molecular dynamics simulations

5.4.1.1 Initial orientation of MOMP models

A number of CG simulations were run in parallel to determine the clustering behaviour of two MOMP models within a lipid bilayer. Two 4 Å crystal structures were placed a protein distance apart, ~42 Å, to limit bias with one model rotated 90° about the Z axis to create four initial conformations. **Figure 5.5** displays these orientations, with tryptophan 55 (Trp55) labelled in red for clarity of orientation, as the low-resolution crystal structure does not detail the external loop region.
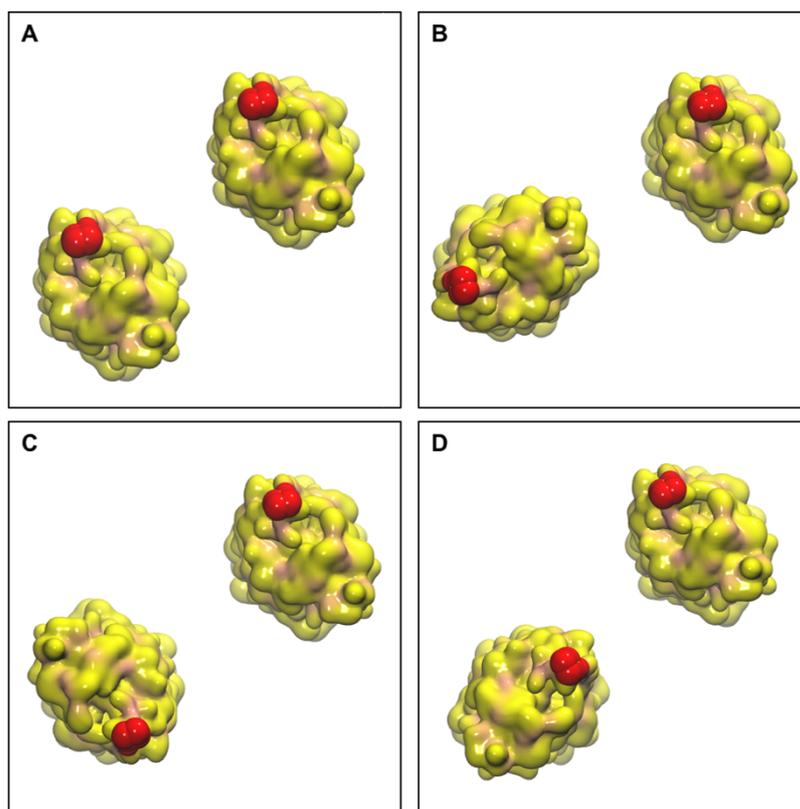
Figure 5.5. **A-D, the four orientations of the two 4 Å MOMP crystal structure models.** The MOMP model on the right-hand side remained stationary whilst the other MOMP model was rotated 90° about the Z axis to allow multiple faces of the β-barrel to interact. Trp55 is shown in red to indicate the top, external side of the barrel. Images were produced in VMD software v1.9.2.

5.4.1.2 MD simulation of MOMP in a DPPE/DOPG bilayer

Following a simulation time of 10 µs, the two MOMP molecules in **Figure 5.5B**, herein referred to as orientation B, migrated to a distance that enabled intermolecular contacts to occur between the two molecules, while inserted in the lipid bilayer, as indicated in **Figure 5.6A**. The Trp55 residue (red) was selected as a marker due to its proximity to the predicted location of the extracellular loop, and as shown in greater clarity in **Figure 5.6B** it is evident that this region of the MOMP molecule is externally exposed.
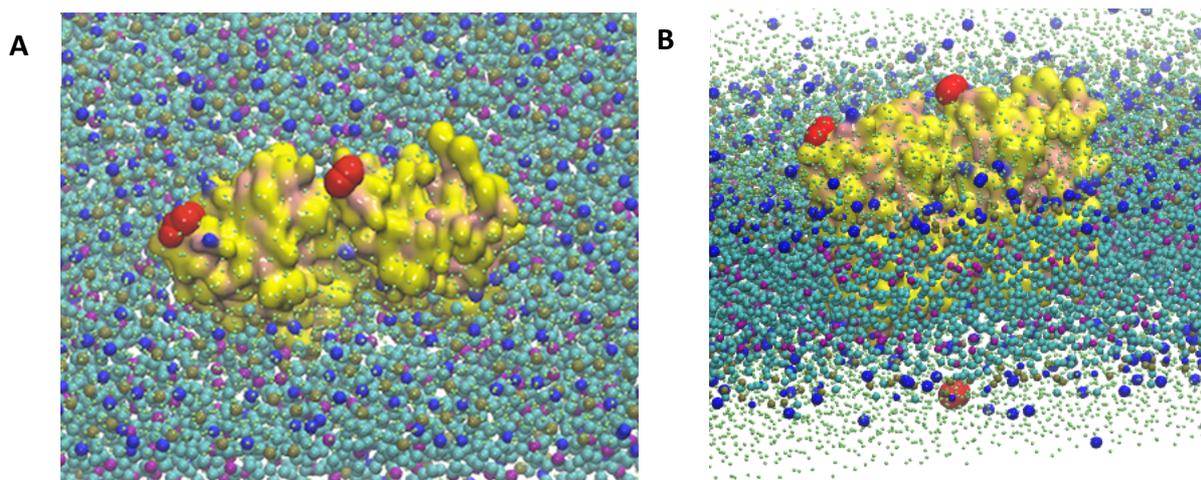
Figure 5.6. **MD simulation of two MOMP models in a DPPE/DOPG lipid bilayer. A, View of the top, extracellular face of the bilayer.** The final conformation of orientation B following a 10 μs simulation. The distance of ~42 Å between the two MOMP molecules was closed and intermolecular interactions formed. Residue Trp55 is also positioned on the external face of the bilayer in both MOMP models. **B, Cross sectional view of the bilayer**. Insertion of the two MOMPs within the lipid bilayer, is made clearer with a cross sectional view. Yellow, MOMP; red, Trp55; green, water; blue, cyan and purple, DPPE/DOPG. Images were produced in VMD software v1.9.2.

5.4.1.3 Analysis of intermolecular interactions between MOMPs from orientation B simulation

An additional script was implemented, in order to characterise the intermolecular interactions formed in the MD simulation of orientation B, whereby two cut off values for interaction were set, at both 45 Å and 80 Å, in order to ensure that all possible interactions were considered. From this, 37 interactions were identified for both cut off values. **Figure 5.7** depicts a graphical representation of these interactions, whereby the most frequently interacting residues are indicated by dark red bars, and the most likely interacting pairs by dark green blocks. From this graph, the most dominant residues involved in intermolecular contacts were as follows; Met85, Asn127, Arg195, Trp199, Thr262, Lys305, Pro318, Ser319, and Asp365, as indicated by a score above 3.0% (red; interactions accounted for by residue).
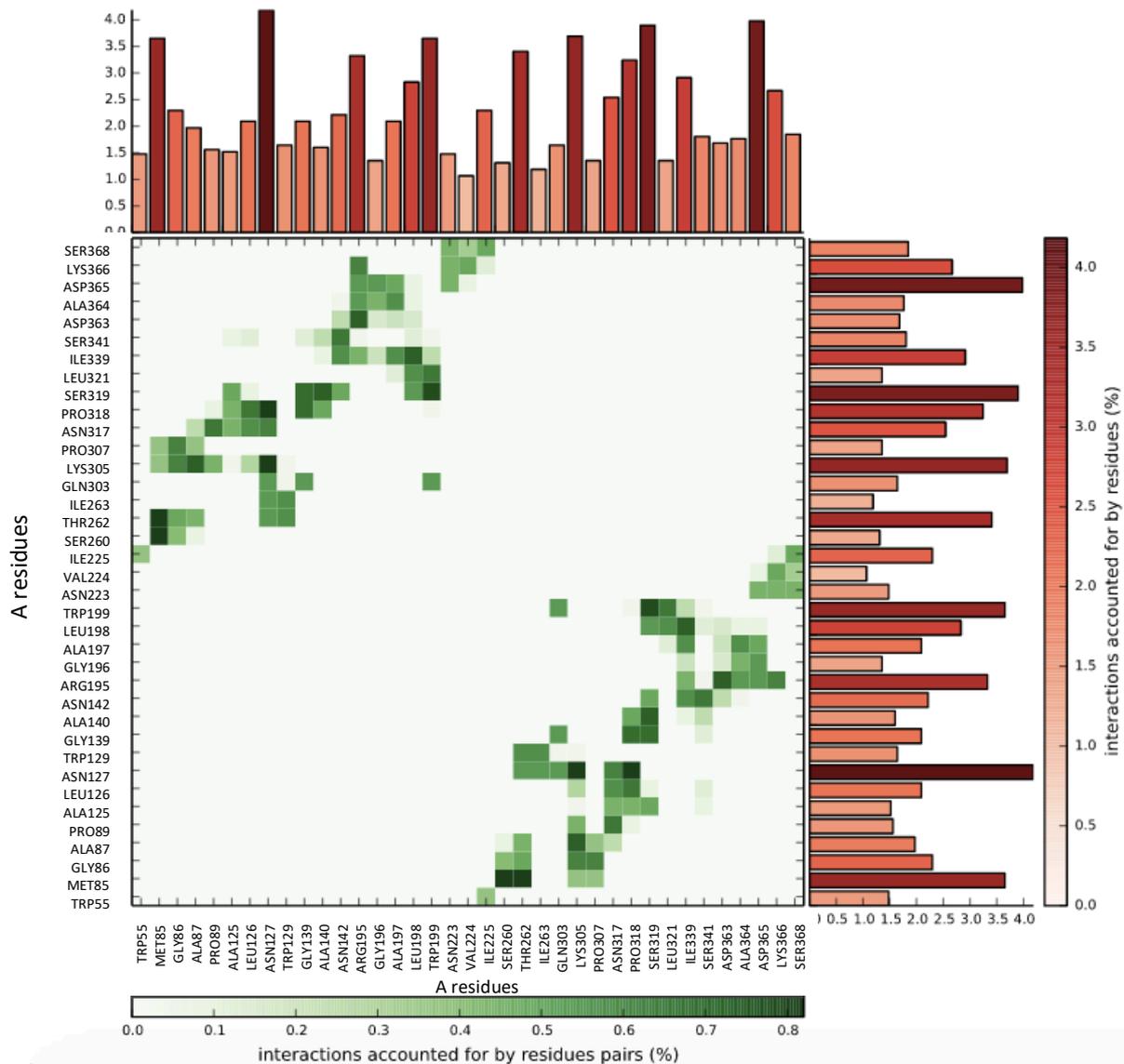
Figure 5.7. **Graph identifying the interacting residues from the MD simulation of orientation B.** The interactions between two MOMP crystal structures inserted into an asymmetrical bilayer were analysed. The red bars indicate the percentage of interactions accounted for by each residue, while the green squares indicate the percentage of interactions accounted for by each pair.

An additional output from the interaction assessment was a PDB model with residues coloured by B factor (a measure of relative vibrational motion), whereby blue reflects well-ordered residues with low displacement, red indicates highly displaced residues, and white is intermediate. As all residues are coloured on one model, **Figure 5.8A** depicts the model duplicated and aligned based on both the interacting residues suggested in **Figure 5.7** and the location of the blue coloured residues, which relates to the residues most likely to form an interaction. This figure indicates the likely alignment of the two MOMP molecules within the lipid bilayer, as simulated by MD. **Figure 5.8B** shows the outward orientation of the two interacting faces in order to obtain a clearer image of the interacting residues. From **Figure 5.8B**, the most dominant intermolecular contacts aligned to the complete MOMP protein sequence, in blue, are as follows; Arg195-Asp365, Trp199-Ser319, Asn127-Lys305, and Met85-Thr262.
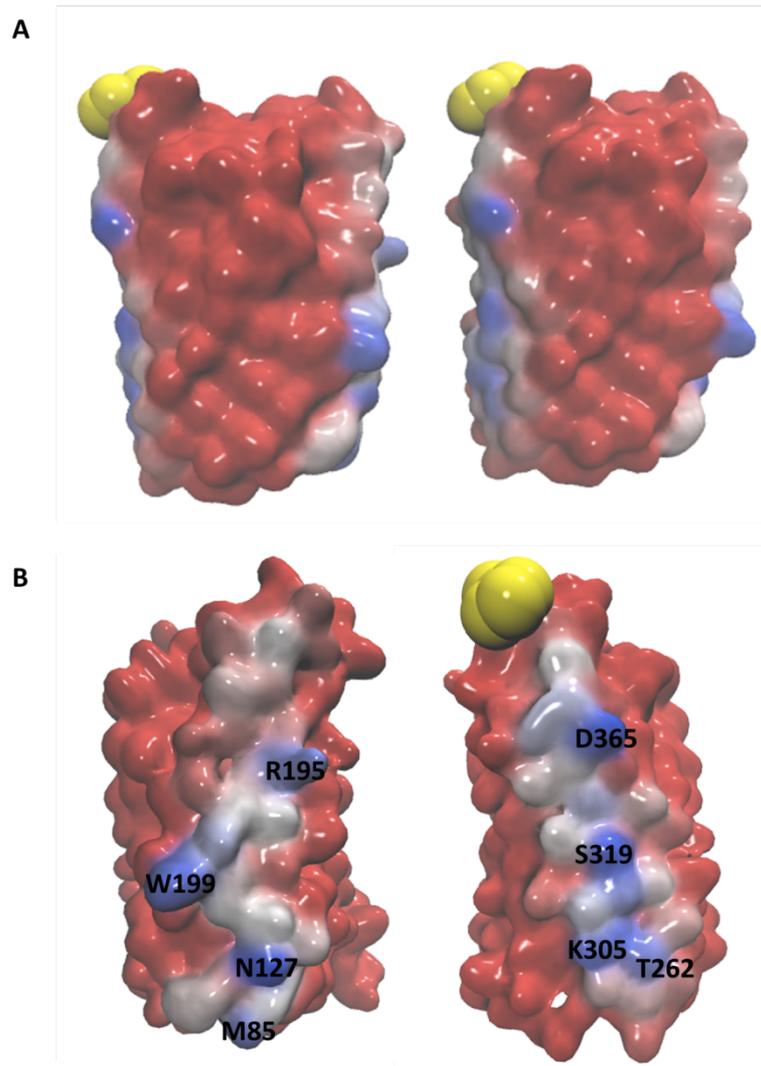
Figure 5.8. **MOMP crystal structure coloured according to B factor, arising from the MD simulation. A, Duplicated model to demonstrate the potential interaction between two MOMPs** and **B, The models orientated to show the labelled residues at the interaction interface.** Blue, well ordered; red, highly disordered; white, intermediate; yellow, Trp55. Images were produced in VMD software v1.9.2.

As none of the most important residues involved in the interaction occurring between the two MOMP models were cysteines, as expected, **Figure 5.9** shows the MOMP crystal structure with each well-ordered interaction residue highlighted, along with each of the nine cysteine residues. Interestingly, from this figure, it is evident that residues C136, C201, C203, and C342 on the opposing face are in close proximity to the interacting faces of the barrel as predicted by MD simulations.
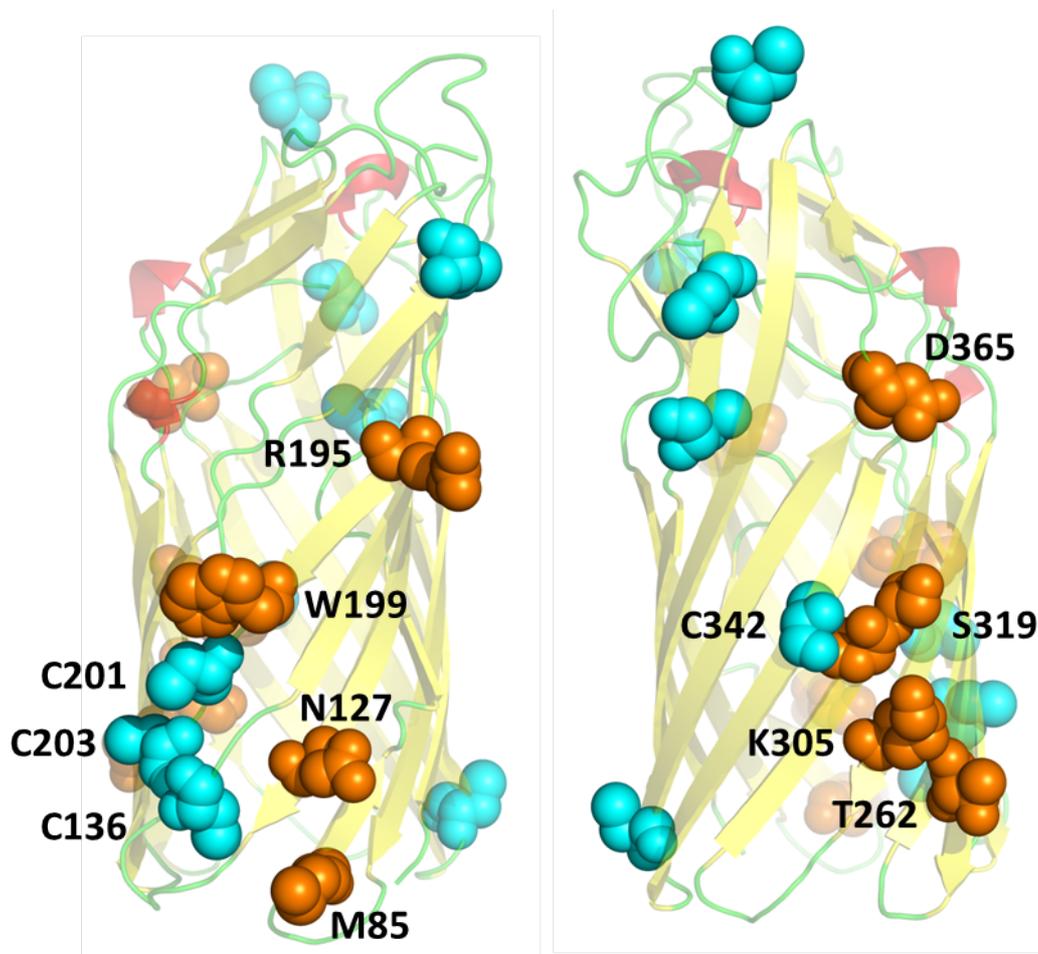
Figure 5.9. **Two MOMP crystal structures showing the interacting residues from MD simulations and their proximity to the cysteine residues.** Models are displayed in the same orientation as in **Figure 5.8B**. Well-ordered interacting residues identified through MD simulations are shown as orange spheres and labelled while the nine cysteines residues are shown in cyan, with key residues labelled.

5.4.1.4 Cysteine residue mutagenesis of MOMP

Figure 5.10 shows the current low-resolution crystal structure for *C. pneumoniae* MOMP, with four cysteine residues of interest highlighted in magenta; C136, C201, C203 and C226. In line with the novel hypothesis that cysteine rich pockets exist within the COMC proteins, the first three residues were of particular interest due to their close proximity and external orientation within the β-barrel, making these cysteine residues amenable for intermolecular disulphide bonding. Additionally, data collected from MD simulations indicated that these three residues, in addition to the single C226 residue, are in close proximity to the primary

residues involved in the interaction between two MOMP models within the asymmetric lipid bilayer and, therefore, are potentially involved in intermolecular disulphide bond clustering. Residue C226 was also mutated for completeness, in order to characterise all the cysteines on this interacting face of the barrel.

Initially, residues C201 and C203 were both mutated individually and together, as a double mutant, to alanines using SDM with the pETMOMP$_{H6}$ construct in order to observe their potential effect on intermolecular disulphide bonding, as characterised by MOMP's distribution in the *E. coli* outer membrane. As these residues may have been essential for correct folding and insertion of MOMP, a growth curve was conducted. As expected, and shown in **Figure**

Figure 5.10. **MOMP crystal structure with cysteine residues highlighted.** The four cysteines of interest C136, C201, C203, and C226 are labelled and shown as magenta spheres, with the remaining cysteines shown as cyan spheres.

5.11, the control cultures containing C41 (DE3) cells grew to a higher OD more quickly, reaching an OD$_{600}$ of 0.5 around 45 minutes earlier than transformed cells, as well as reaching an OD$_{600}$ of 2.12 after six hours of growth. By comparison, cells expressing MOMP and its mutants all took longer to reach an inducible level (compared to the control) with no substantial difference between the native and mutant proteins. All cells transformed with a variant of MOMP grew at a similar rate, with growth plateauing at an OD$_{600}$ of approximately 1.3.

Figure 5.11. **Growth curve of pETMOMP_H6 cysteine mutants compared to C41 (DE3) control cells.** All samples were transformed into C41 (DE3) *E. coli* cells. The standard growth curve procedure was followed, with measurements taken hourly post induction for a total of 6h. Dark blue, control C41 (DE3) cells; orange, native MOMP; yellow, MOMP C201A; grey, MOMP C203A; and light blue, MOMP C201/3A. Error bars represent standard deviation.

A medium scale growth test of 250 ml was conducted for each mutant in order to compare the expression levels of MOMP. As previously discussed, due to low expression in general, Western blotting was used to compare each mutant, as shown in **Figure 5.12**. From this figure, it is evident that MOMP expression is comparable (and roughly equal) between each construct, regardless of mutation. Coomassie stained loading controls are shown in **Figure A5.1A.**
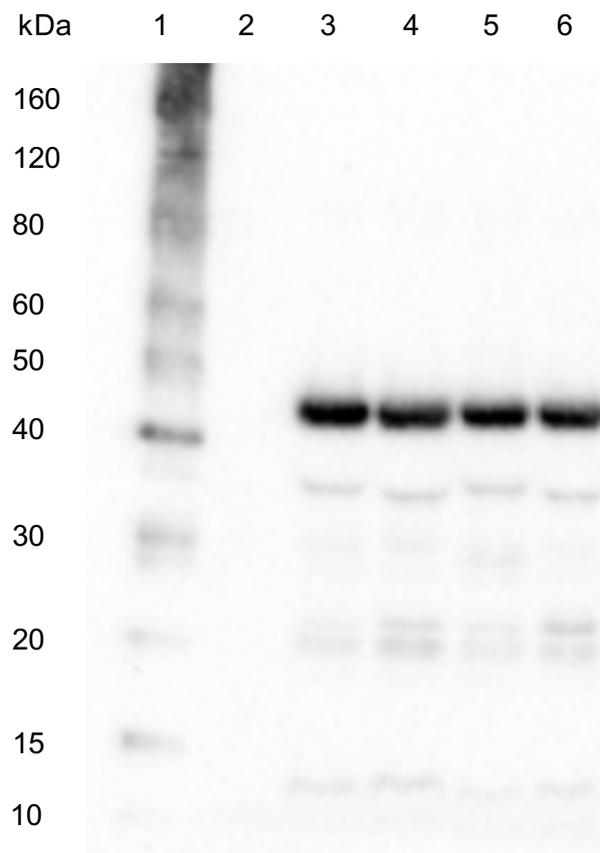
Figure 5.12. **Western blot comparing expression of MOMP cysteine mutants.** pETMOMP$_{H6}$ constructs were transformed into C41 (DE3) *E. coli* cells and membranes were prepared as per the standard medium-scale expression protocol (see **Chapter 2 Section 2.2.2.3**). 1, His-tagged protein standard; 2, C41 (DE3) control; 3, native MOMP; 4, MOMP C201A; 5, MOMP C203A; and 6, MOMP C201/3A. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L)-HRP conjugated secondary antibody.

5.4.2 Fluorescence Microscopy

5.4.2.1 Determining the optimal bacterial count

Initial experiments were aimed at determining the optimal cell number for adherence to the poly-L-lysine coated glass microscope slides, taking into account the potential loss of cells during the multitude of wash steps. Attachment of varying dilutions of cells indicated that an OD_{600} value of 0.5, measured at the time of harvesting and equating to approximately 4×10^8 cells, produced an even dispersion of bacteria with limited overlap.

5.4.2.2 Accounting for bacterial autofluorescence

To account for bacterial autofluorescence under excitation with the 488 laser, control cells of C41 (DE3) not expressing any $pETMOMP_{H6}$ were imaged in the first instance and the 488 nm laser settings adjusted so that autofluorescence was no longer visible. In order to limit false positive results, test samples were then imaged using these baseline settings so that any fluorescence observed was due to the fluorescent labelling and not bacterial autofluorescence.

5.4.2.3 Addition of a lysozyme treatment step to improve fluorophore access

A lysozyme treatment step was added to the labelling protocol, in order to improve access of the fluorophores to the C-terminal his-tag of MOMP, which is positioned in the periplasm between the inner and outer membranes. This step was introduced, following permeabilisation of the outer membrane with Triton X-100, in order to hydrolyse the linker bonds within the PG layer and improve accessibility around the his-tag. From **Figure 5.13,** it is evident that this additional treatment dramatically increased the efficiency of antibody labelling, with all cells from the brightfield image in the +lysozyme sample displaying

fluorescence in the respective epifluorescence and TIRF images, as opposed to the weak and

patchy fluorescence seen in -lysozyme non-treated cells.

| Brightfield | Epifluorescence | TIRF |
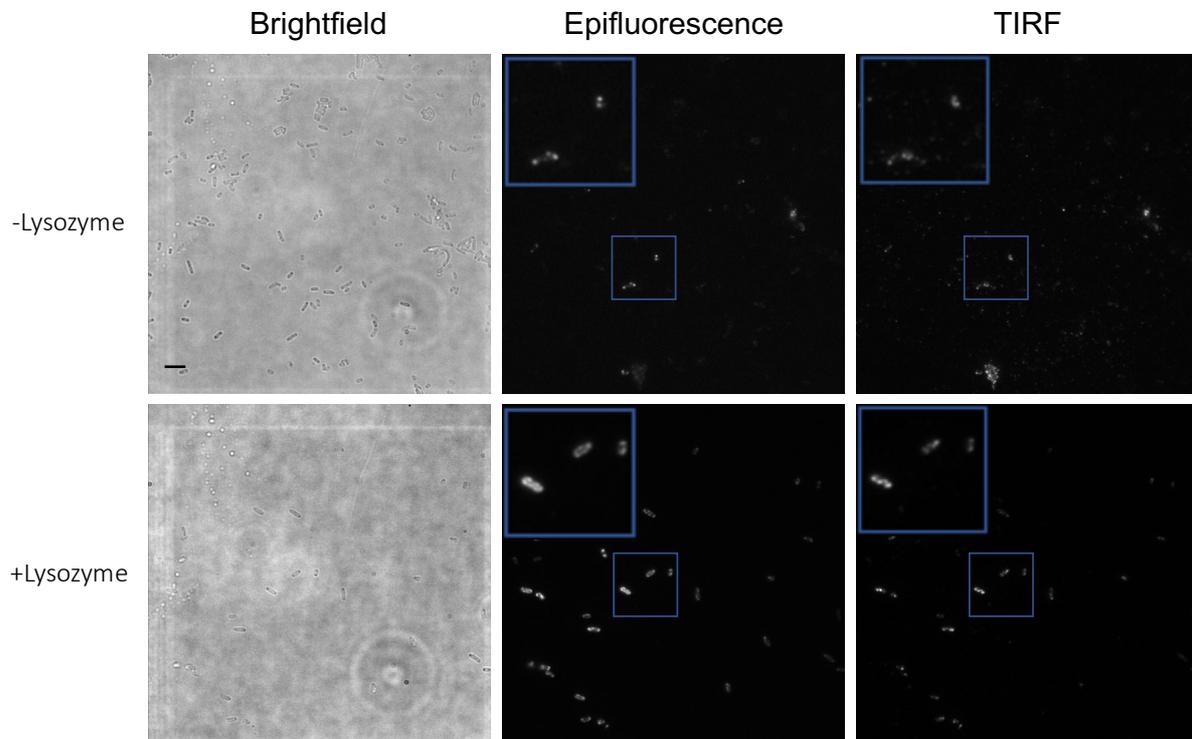
Figure 5.13. **Brightfield, epifluorescence, and TIRF microscopy images showing treatment with -/+ lysozyme.** C41 (DE3) cells expressing MOMP were stained using anti-his antibody conjugated to Alexa 488 secondary antibody and imaged at the University of Reading with the following conditions; 0.7% 488 laser, 200 ms exposure, 200 gain, and a 512 x 512 pixel ROI. Scale bar is 5 μm.

5.4.2.4 Alternative labelling approaches

A *C. pneumoniae* anti-MOMP antibody was obtained from the *Chlamydia* Biobank (CT603) in

an effort to directly label MOMP, without requiring membrane permeabilisation. A series of

dilutions between 1/250 to 1/1000 were used in cohort with anti-mouse Alexa-647

secondary antibody, however, fluorescent labelling was not observed. A Western blot was

also conducted on rMOMP membranes using this antibody, whereby again no signal was

detected (data not shown). Additionally, the novel P3NTA tag, as used in FSEC for detergent

screening, was used to tag pETMOMP$_{H6}$ in a single step process. Unfortunately, the strength of the attached fluorophore, fluorescein, was not sufficient enough to resist photobleaching and, therefore, it was not possible to capture both epifluorescence and TIRF images before this occurred.

5.4.2.5 Clustering differences between native and mutant MOMP

Following the development of a successful method for sample preparation and labelling, the cysteine mutants of MOMP were imaged, alongside native MOMP and the control protein OmpA, in order to determine if these residues had critical roles in the clustering of MOMP in the outer membrane. OmpA is a well characterised and abundant *E. coli* outer membrane protein and, therefore, was labelled with an anti-OmpA antibody in order to demonstrate the membrane localisation of an evenly distributed protein. Importantly, OmpA does not form intermolecular disulphide bonds, as expected with MOMP. From **Figure 5.14A,** it is clear that OmpA forms a distinct and solid ring with epifluorescence imaging, which presents as a solid block of fluorescence with TIRF, signifying a regular distribution throughout the membrane. **Figure 5.14B,** showing native MOMP, exhibits cells with a clustering pattern of fluorescence observed in both microscopy techniques, hypothesised to be due to intermolecular disulphide bond formation, which prevents even distribution of MOMP in the outer membrane. **Figures 5.14C, 5.14D** and **5.14E** represent the single C201A, C203A, and double C201/3A mutants respectively. **Figure 5.14C**, relating to mutant C201A, appears to exhibit a more homogenous distribution, as opposed to the C203A mutant in **Figure 5.14D** which still shows a high degree of clustering. Interestingly, the TIRF microscopy image of the double mutant in **Figure 5.14E** indicates highly disperse fluorescence within the membrane, through the observable solid and less speckled fluorescence.
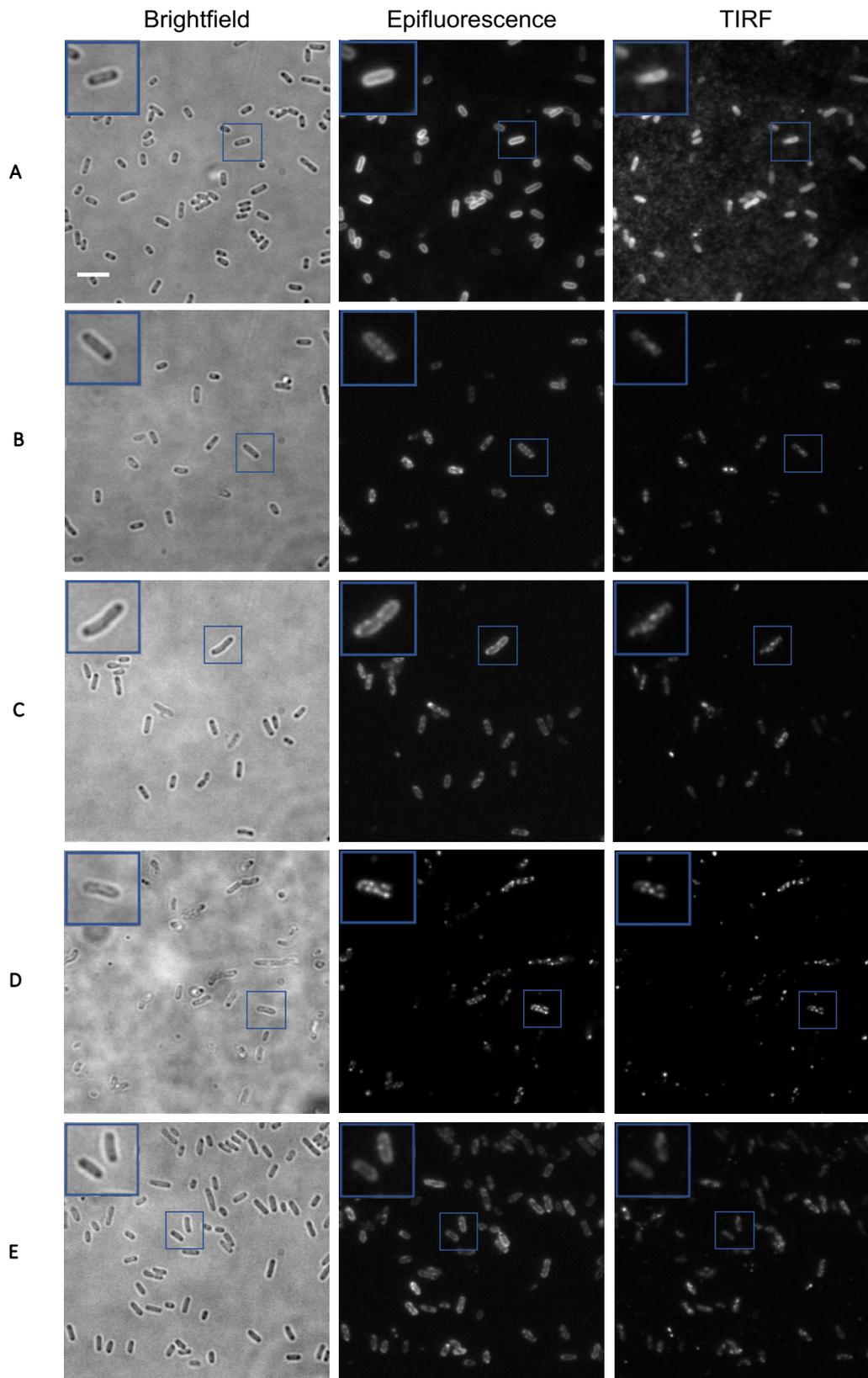
Figure 5.14. **Brightfield, epifluorescence and TIRF microscopy images of MOMP cysteine mutants in C41 (DE3)** *E. coli* **cells. A**, *E. coli* anti-OmpA; **B**, native MOMP; **C**, MOMP C201A; **D**, MOMP C203A; **E**, MOMP C201/3A. All cells were stained using anti-his antibody conjugated to Alexa-488 secondary antibody and imaged at the University of Reading with the following conditions; 0.7% 488 laser, 200 ms exposure, 200 gain, and a 256 x 256 pixel ROI. Scale bar is 5 μm.

5.4.2.6 Additional cysteine mutant design

Since the preliminary TIRF experiments indicated a difference between the native MOMP and the cysteine mutants C201A and C201/3A, an additional four mutants were designed, one of which was residue C226 near the top of the barrel, in agreement with the MD simulation predicted interaction face. The others included a single mutant C136A, due to its close proximity to the initial two mutants, and two associated double mutants C136/201A and C136/203A. Medium scale expression tests were conducted, whereby resuspension buffer volumes were calculated in proportion to the OD_{600} measurement taken at the time of harvesting. Prepared membranes were analysed by Western blotting, as shown in **Figure 5.15** (loading control **Figure A5.1B**), which indicates that all mutants produce comparable quantities of MOMP. The smear present in lane 4, corresponding to mutant C201A is an artefact of the gel, as this effect was not observed in the previous Western blot in **Figure 5.12** which analysed the same membrane preparation.
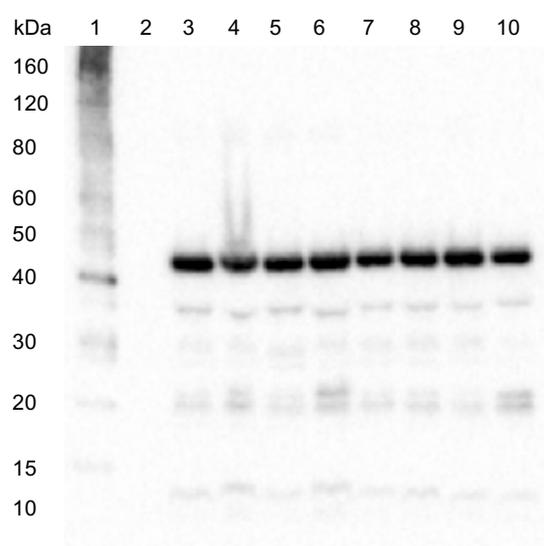
Figure 5.15. **Western blot comparing expression of additional MOMP cysteine mutants.** pETMOMP$_{H6}$ constructs were transformed into C41 (DE3) *E. coli* cells and membranes were prepared as stated in the standard medium-scale expression protocol (**Chaper 2 Section 2.2.2.3**). 1, His-tagged protein standard; 2, C41 (DE3) control; 3, native MOMP; 4, MOMP C201A; 5, MOMP C203A; 6, MOMP C201/3A; 7 MOMP C136A; 8, MOMP C136/201A; 9, MOMP C136/203A; and 10, MOMP C226A. Protein was detected using 6X his epitope tag monoclonal mouse IgG primary antibody and anti-mouse IgG (H+L)-HRP conjugated secondary antibody.

5.4.2.7 Initial STORM strategy

STORM was used to obtain both higher resolution images and, more importantly, quantitative data desired for an analysis of MOMP clustering. Initially, the superfrost microscope slide method developed in this research was adopted, being the established protocol for bacterial cell adhesion for STORM at the Central Laser Facility (CLF). However, this method required pelleting cells after each treatment and wash cycle. The cells plated during this method demonstrated poor labelling with the same antibodies used in TIRF microscopy. Alternatively, the ibidi slide method, as developed for and used in TIRF microscopy experiments, was adopted. Sample processing for this technique was considerably quicker and resulted in improved labelling. Additionally, as a coverslide was not required, the samples were viable for a longer period, as the DTT imaging buffer did not evaporate as quickly within the wells after exposure to the intense laser beam.

5.4.2.8 HaloTag-SiR fluorophore with the pOPINHaloMOMP construct

HaloTag-SiR was tested as a potential labelling method for MOMP expressed using the pOPINHalo vector due to its membrane permeability and negating the requirement of Triton X-100 for cell permeabilisation. The pOPINHaloMOMP construct was not amenable to SDM using the QuikChange II XL commercial kit (**Chapter 2 Section 2.2.1.2**), the previous method used to create cysteine mutants, as the PCR products presented as multiple bands as determined using agarose gel analysis. Instead, In-Fusion was used to extract the mutated sequences from pETMOMP$_{H6}$ for insertion into the linearised pOPINHalo vector. Insertion of the correct sequence was confirmed through sequencing in both directions with the commercially available T7 forward primer and the custom pOPIN3CHalo reverse primer (designed in this research). After single step labelling with HaloTag-SiR and incubation in

fresh imaging buffer, STORM data of native MOMP and the initial cysteine mutants at residues 201 and 203 were captured. Interestingly, unlike in preliminary epifluorescence and TIRF microscopy, the fluorescent signals were not localised to the cell membrane and instead appeared to form an angular criss-cross pattern in all samples, as seen in **Figure 5.16A**. From the brightfield image in **Figure 5.16B**, of the same region of interest, it is clear that there is only a single cell present, and not a number of cells, as might be expected.
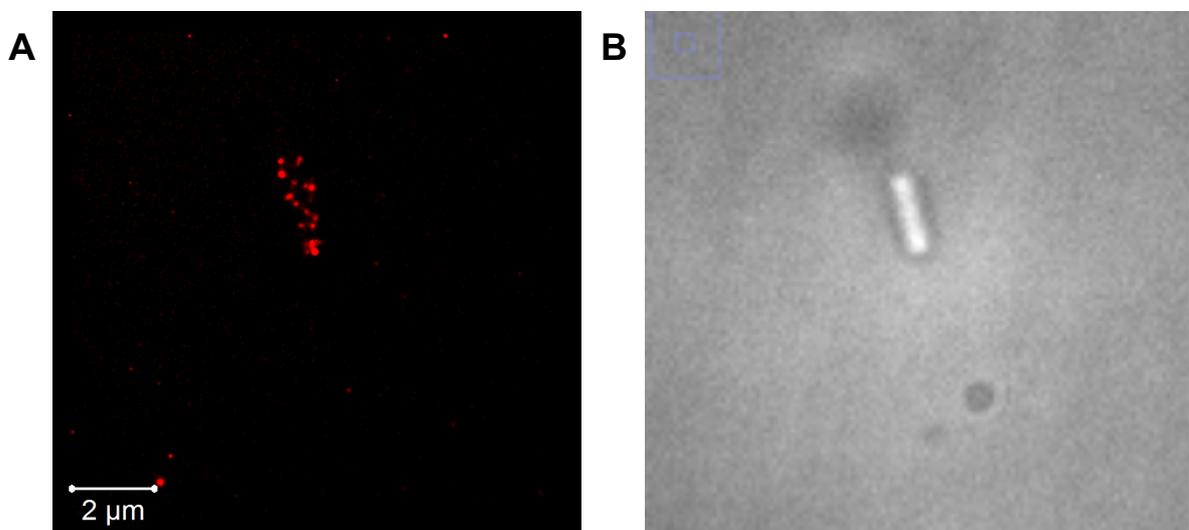
Figure 5.16. **C41 (DE3)** *E. coli* **cell expressing pOPINHaloMOMP C201/3A double mutant labelled with HaloTag-SiR. A, STORM image** and **B, Brightfield image.** The criss-cross pattern of fluorescence was visible in both native and mutant MOMP samples. Cells were prepared without permeabilisation and labelled with HaloTag-SiR. Images were captured using Octopus at the CLF, Harwell using a 647 nm laser, 20 ms exposure, 100 gain, and a 128 x 128 pixel ROI.

To further investigate this anomaly, 3D STORM imaging was utilised for this particular sample in order to obtain details on the axial fluorophore locations, in additional to the lateral locations, as in conventional STORM. The aim of this was to determine if the criss-cross pattern of fluorescence observed in **Figure 5.16A** was in fact a helical arrangement around the outside of the cells, possibly as a result of poor expression or inefficient export of MOMP to the outer membrane. The results of the 3D STORM imaging are shown in

Figure 5.17, whereby **5.17A** indicates the cellular morphology with brightfield imaging.

Figure 5.17B-D illustrates various perspectives of the cell using 3D STORM, whereby the Z-axis is highlighted in blue, the X-axis in red, and the Y-axis in green. **Figure 5.17C** depicts the same perspective as in **Figure 5.17A**, with only the X and Y axis. In **Figure 5.17B** and **5.17D**, the Z-axis view enables 3D visualisation of the fluorophore labelling, which shows the fluorescence as defined loops around the cell, as opposed to being more dispersed across the membrane.
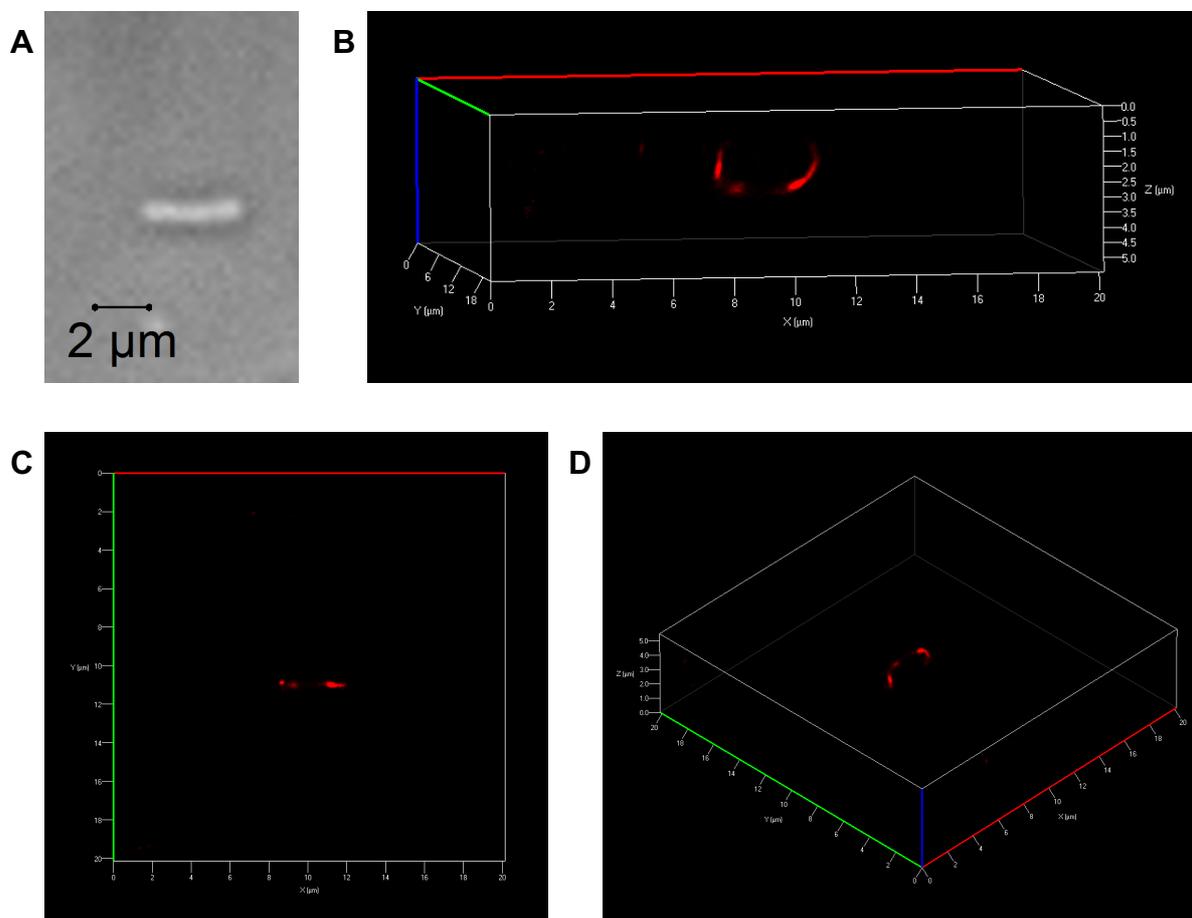
Figure 5.17. **3D STORM to investigate the behaviour of HaloTag-SiR fluorophore. A, brightfield image and B, C, and D, 3D STORM images.** C41 (DE3) *E. coli* cells expressing pOPINHaloMOMP C201/3A double mutant were prepared without permeabilisation and labelled with HaloTag-SiR. **A** and **C** both show the same perspective, with **B** a view along the Z plane, and **D** demonstrating the cell within all planes. X-axis, red; Y-axis, green; Z-axis, blue. Images were captured using Octopus at the CLF, Harwell and visualised with Zen Black 2012 software.

5.4.2.9 Anti-his Alexa-647 one-step labelling method

Due to the unusual results observed with the HaloTag-SiR fluorophore, a different labelling method was adopted. As in Western blotting detection, the C-terminal his-tag was targeted with a conjugated anti-his Alexa-647 dye. However, as this antibody is not membrane permeable, treatment with Triton X-100, as employed in the preliminary TIRF experiments, was reintroduced. Since Alexa-647 exhibits high photostability and a strong fluorescence signal, only a laser power of 2-3% was required for successful cell imaging, permitting the collection of 20,000 frames before photobleaching began to occur. For each sample, at least ten STORM data sets were collected. **Figure 5.18** shows a selection of the super-resolution STORM images for control OmpA, native MOMP and all seven of the cysteine MOMP mutants. Although the images in **Figure 5.18** showed significantly more detail, compared to the TIRF images in **Figure 5.14**, it remained that qualitative distinctions between the mutants was still a challenging and subjective process.
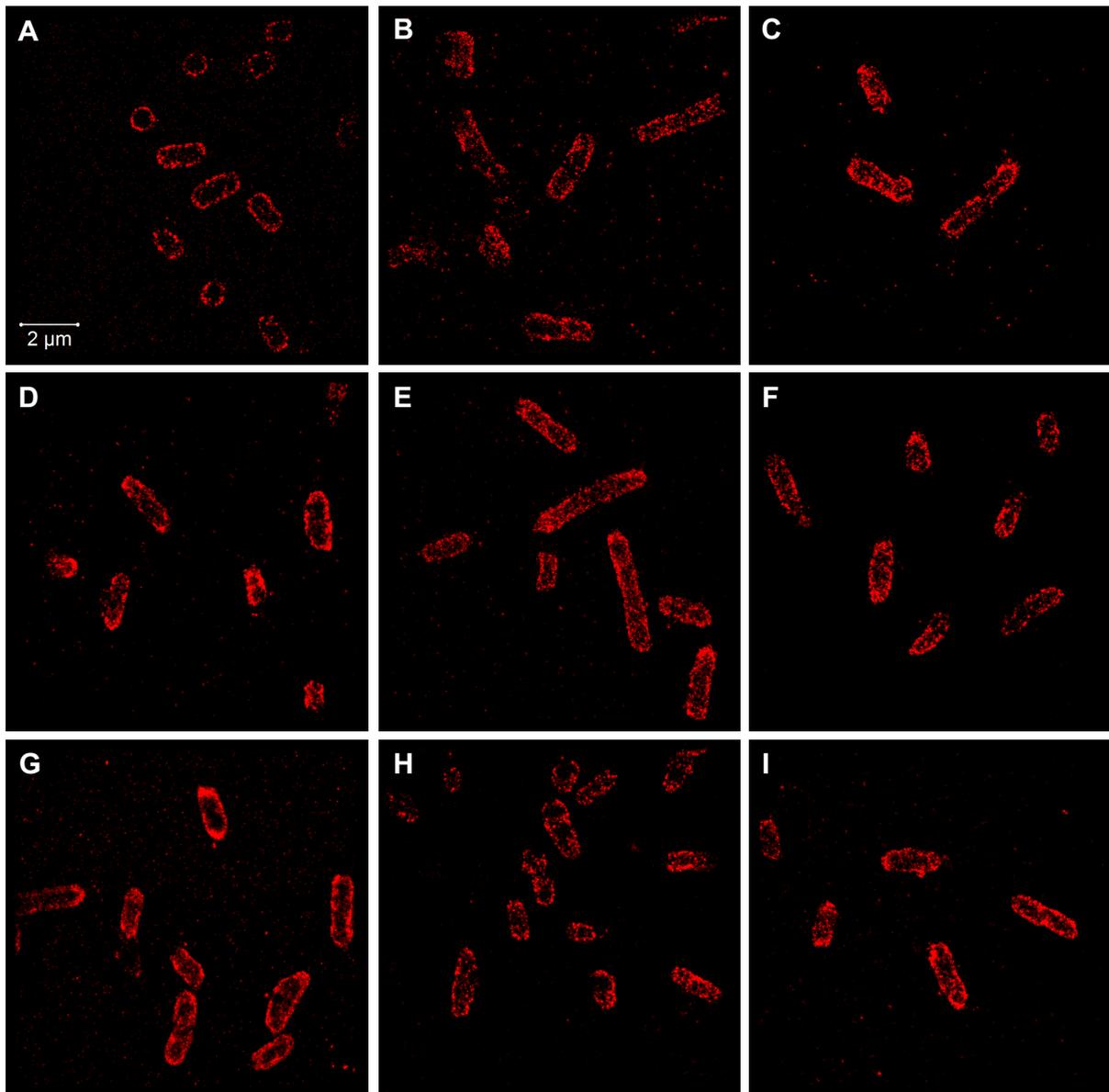
Figure 5.18. **Super-resolution STORM images of C41 (DE3)** *E. coli* **cells. A**, *E. coli* anti-OmpA; **B,** native MOMP; **C**, MOMP C201A; **D**, MOMP C203A; **E,** MOMP C201/3A; **F**, MOMP C136A; **G**, MOMP C136/201A; **H**, MOMP C136/203A; and **I**, MOMP C226A. All cells were induced for a total of 3h, with MOMP expressing cells labelled with 1/100 anti-his Alexa-647 antibody and OmpA first with anti-OmpA primary antibody coupled with anti-647 secondary antibody. 20,000 frames were collected at the Octopus facility (CLF, Harwell) under the following conditions; 2-3% 647 laser, 20 ms exposure, 100 gain, and a 128 x 128 pixel ROI.

Each STORM image (**Figure 5.19A**) is in fact comprised of a dataset of fluorophore localisations, which are first drift corrected to account for miniscule camera movements and grouped to normalise trails before quantitative analysis of clustering can be conducted, using ClusDoC software, which implements the density based spatial clustering of applications with noise (DBSCAN) algorithm (Ester *et al.*, 1996). The localisation text file from each STORM dataset was uploaded into the ClusDoC GUI (**Figure 5.19B**), whereby regions of interest (ROI), in this case whole bacterial cells, were manually selected with the ROI drawing tool. In order to utilise the DBSCAN algorithm for clustering a number of parameters were required, the most important being the radius of the fluorophore and the minimum number of points required for clustering (**Figure 5.19C**). Following analysis, the software provides an image of the ROI showing the normalised cluster density (**Figure 5.19D**), in addition to a data table containing quantitative clustering information.
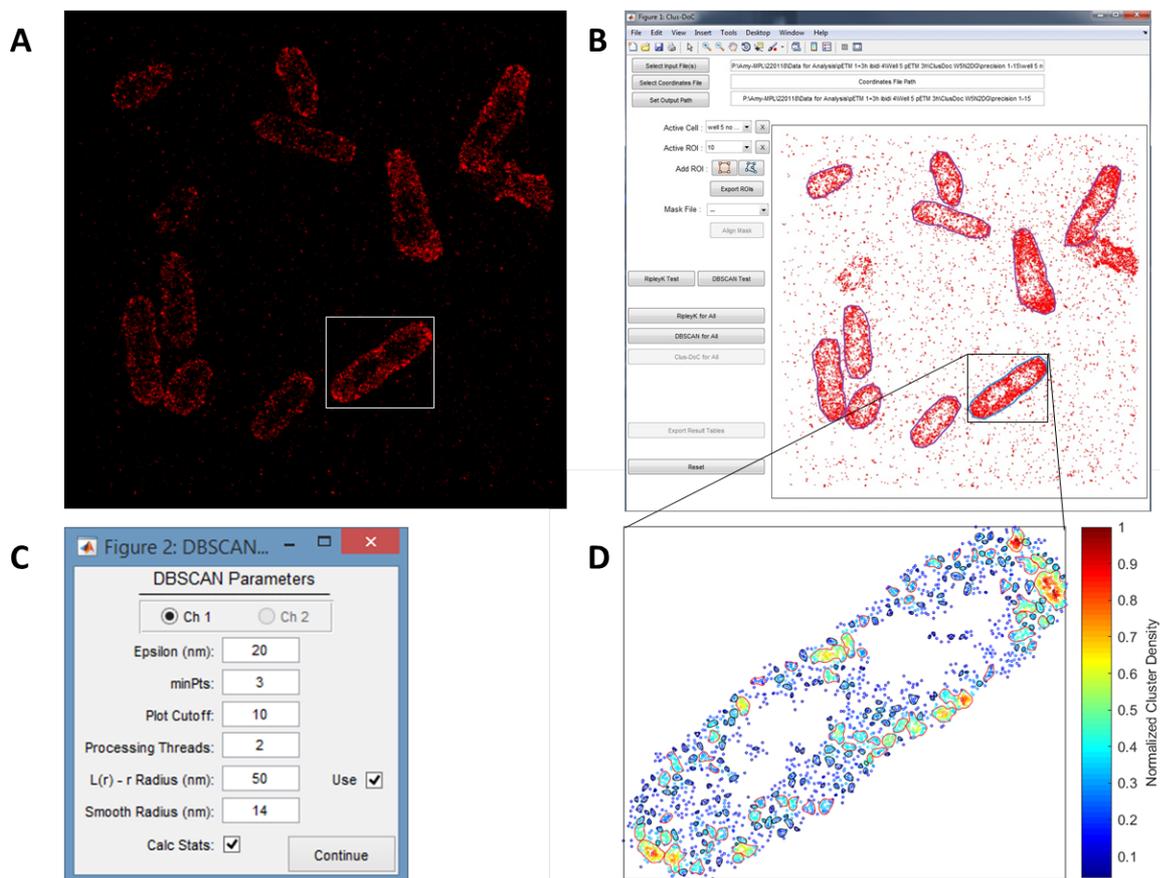
Figure 5.19. **Workflow for ClusDoC software for the analysis of STORM data. A,** Post-processing STORM image; **B,** ClusDoC GUI displaying imported STORM localisation data; **C,** Parameters required for DBSCAN algorithm; and **D,** Example of clustering assignment image for one ROI, colour coded according to the normalised cluster density.

For each control and mutant, the resulting analyses were combined and the mean and standard error (SE) calculated, the results of which are shown in **Figure 5.20** and **Table 5.1**. From **Figure 5.20,** it is evident that native MOMP was the most clustered, with an average relative density of clusters of approximately 3.7 arbitrary units (AU), compared to the OmpA control, which had an average density of 2 AU. The mutants with the most significant decrease in clustering were the two double mutants C201/203A and C136/201A and the single mutant C226A with average densities between 1.9 AU and 2.1 AU (**Table 5.1**). In order to further analyse the statistical significance of this data, a one-way ANOVA was conducted (**Table 5.2**) against the null hypothesis that there is no statistically significant difference

between groups. The results of this one-way ANOVA (F(8,630) = 35.8, P>0.0001) revealed that the null hypothesis can be rejected, with a P-value less than an alpha value of 0.05 as well as the F critical value of 1.96 being less than the value of F. In order to determine where these differences occurred, for groups with unequal sample sizes, a Tukey post-hoc test was conducted (**Table 5.3**). From this test, it was apparent that clustering differences between OmpA and native MOMP were statistically significant, as well as those between native MOMP and all seven of the cysteine mutants. This complements the graph in **Figure 5.20**, whereby SE error bars between these groups do not overlap.

A statistically significant difference was observed between OmpA and the following mutants; C201A, C203A, C136A, C136/203A (**Table 5.3**). This suggests that whilst these mutants exhibited reduced clustering from native MOMP, they were not as homogenous as control group OmpA. On the contrary, no statistically significant difference was observed between OmpA and mutants C201/203A, C136/201A, and C226A, (**Table 5.3**, highlighted in grey) suggesting that clustering in these mutants had been reduced to exhibit a homogenous dispersion similar to that observed with the OmpA control.

It is interesting that the single mutants C201A, C203A, and C136A did not exhibit a dramatic effect on clustering reduction, despite the effect observed in their double mutant counterparts, although their reduction in clustering compared to the native MOMP still revealed significant differences (**Table 5.3**). Additionally, the double mutant C136A/203A resulted in a decrease in average density, compared to the native protein, but not as substantial as that observed for the other two double mutants.
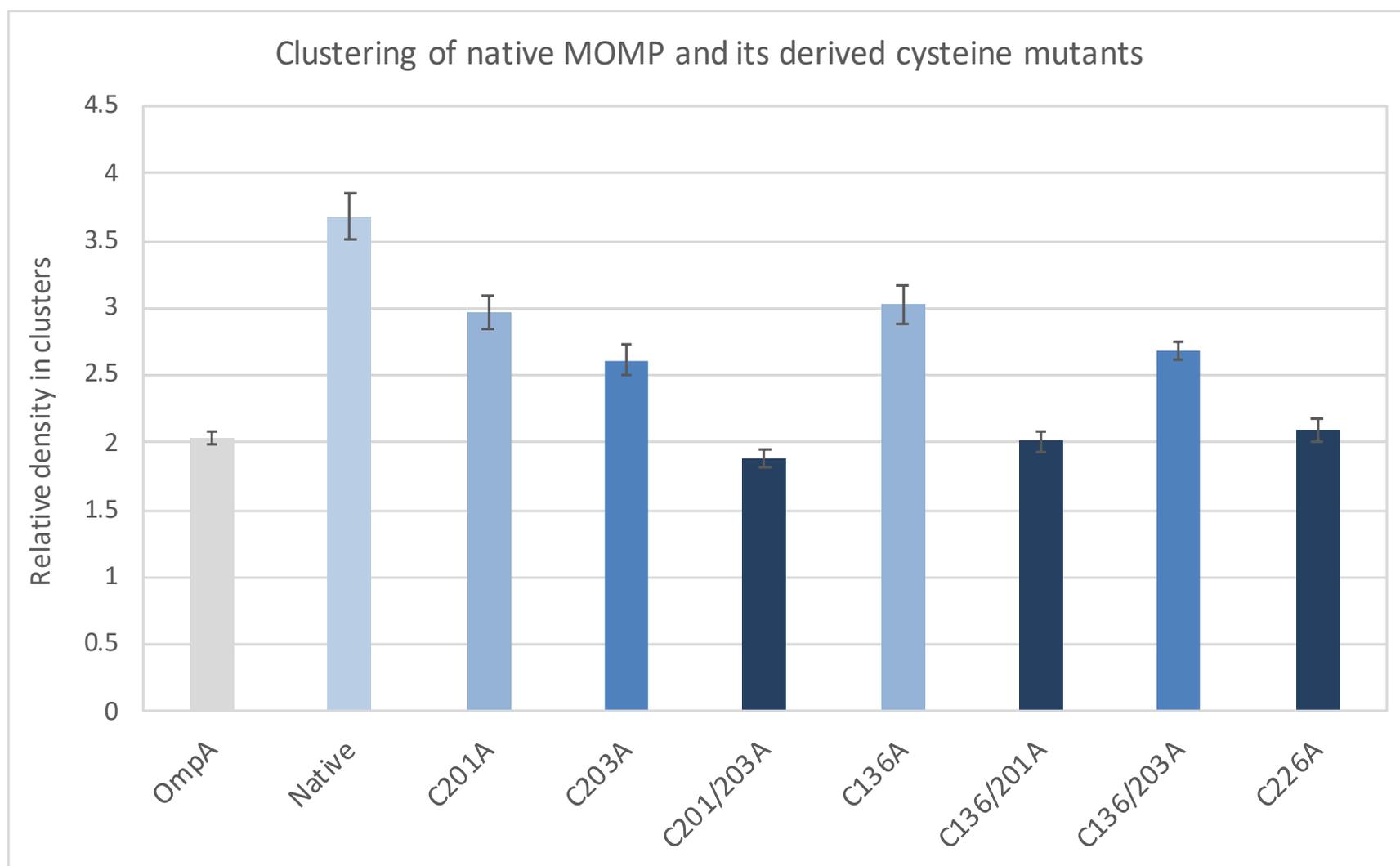
Figure 5.20. **Graph displaying the relative cluster densities of OmpA, native MOMP, and MOMP cysteine mutants.** STORM localisation data was analysed with ClusDoC software before the mean and standard error were calculated for each sample (**Table 5.1**). All data were collected and processed under the same conditions. The relative density in clusters is provided as a standardised arbitrary unit and error bars indicate the standard error.

Construct	Relative density in clusters	SE	n
OmpA	2.04	0.05	94
Native	3.68	0.17	59
C201A	2.97	0.13	51
C203A	2.61	0.11	62
C201/203A	1.88	0.07	71
C136A	3.02	0.14	69
C136/201A	2.01	0.08	78
C136/203A	2.69	0.07	90
C226A	2.10	0.08	65

Table 5.1. **The relative cluster densities and associated standard errors of OmpA, native MOMP, and MOMP cysteine mutants.** STORM localisation data was analysed with ClusDoC software before the mean and standard error were calculated for each sample. All data were collected and processed under the same conditions. SE, standard error; N, number of cells analysed.

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	191.86	8	23.98	35.80	<0.0001	1.95
Within Groups	422.09	630	0.67			
Total	613.95	638				

Table 5.2. **One-way ANOVA statistical analysis of STORM localisation data.** The null hypothesis, that there is no statistically significant difference between groups, can be rejected as both the P-value is <0.05 and the F critical value of 1.95 is less than F at 35.80. SS refers to sum-of-squares; df, degrees of freedom; and MS, mean squares.

Important pairs		Mean difference	N (group 1)	N (group 2)	SE	q-crit	Significant?
OmpA	Native	1.64	94	59	0.10	17.05	YES
Native	C201A	0.71	59	51	0.11	6.44	YES
Native	C203A	1.07	59	62	0.11	10.15	YES
Native	C201/203A	1.80	59	71	0.10	17.63	YES
Native	C136A	0.66	59	69	0.10	6.45	YES
Native	C136/201A	1.67	59	78	0.10	16.71	YES
Native	C136/203A	0.99	59	90	0.10	10.21	YES
Native	C226A	1.58	59	65	0.10	15.21	YES
OmpA	C201A	0.93	94	51	0.10	9.20	YES
OmpA	C203A	0.57	94	62	0.09	6.03	YES
OmpA	C201/203A	0.16	94	71	0.09	1.75	NO
OmpA	C136A	0.98	94	69	0.09	10.65	YES
OmpA	C136/201A	0.03	94	78	0.09	0.33	NO
OmpA	C136/203A	0.65	94	90	0.09	7.60	YES
OmpA	C226A	0.06	94	65	0.09	0.60	NO

Table 5.3. **Tukey post-hoc test to indicate where differences between groups arise.** Following rejection of the null hypothesis with a one-way ANOVA, a Tukey post-hoc test for groups of different sizes was conducted. Clustering between OmpA and native MOMP was deemed to be significantly different, as well as clustering between native MOMP and every cysteine mutant, as assessed with q-crit values above 4.39, the q value derived from a studentised q table with alpha as 0.05. Mean difference refers to the difference in clustering values in AU; N to the number of cells analysed; SE, standard error.

5.5   Discussion

5.5.1   Simulating MOMP interactions within a lipid bilayer using MD

5.5.1.1 MD simulation initial set up

Two MOMP crystal structures were inserted into a DPPE/DOPG asymmetric lipid bilayer at the equivalent of one protein distance away or more specifically 42 Å, the measurement of the widest point of the barrel, based on the low resolution MOMP crystal structure. The purpose of this was to allow the proteins to associate in their own manner, without bias. Placed too close together and bias may occur, whereby the proteins are already predisposed to forming a particular interaction at a preselected interface. A distance of 42 Å was deemed significant, as an overly expansive distance would likely have required a longer simulation time in order for interactions to form, increasing the computational expense.

Four orientations of the barrel (about the Z axis) were simulated to allow multiple faces of the protein barrel to interact, again endeavouring to reduce bias and sample the barrel more or less equally. It was hypothesised that, if suboptimal faces of the barrel were in alignment, then reorientation of the two models into the correct conformation would be a considerably more energy intensive process and may not prove successful. From the results, only the models in orientation B were successful in forming interactions while remaining embedded within the lipid bilayer, as indicated by the external exposure of the (red labelled) reference Trp55 residue (**Figure 5.6**).

5.5.1.2 MOMP insertion and association within the lipid bilayer

Interestingly, the initial alignment of the two MOMP models in orientation B was not the same orientation of the interacting pair after a 10 μs simulation, which showed a rotation of 90° about the Z axis of the stationary MOMP model (right-hand side), resulting in alignment

of the Trp55 residues. This resulting orientation was not modelled in the four initial orientations and, therefore, indicates that the distance between the two models and the duration of the simulation permitted ample time and movement for the models to form this novel interaction. Surprisingly, orientation D did not result in the same association or even correct insertion into the lipid bilayer, despite also only requiring a 90° shift in one of the models to align the Trp55 residues in the same manner as the successful interaction simulation.

Although only four orientations were modelled, due to the nature of proteins with their numerous sidechains, there are in fact countless minor orientations in alignment possible between these two protein models, however, it would be a huge undertaking to simulate every 1-5° degree shift in orientation/rotation around the barrel. For example, if we were to assume simple rectangular cuboid structures in place of the two β-barrel MOMP models, then the four faces of each structure would be able to align in sixteen different ways. However, due to the dynamic nature of MD simulations, as demonstrated here, proteins aligned in a particular manner are clearly able to move and align to form the most favourable interaction and, therefore, a small subset of initial orientations is suitable in this instance.

Additionally, orientation B was the only orientation that resulted in insertion of both MOMP models within the lipid bilayer, with the reference Trp55 residue indicating the location of the loop region, externally exposed to the extracellular side, as expected. In the remaining orientations, some models were able to form interactions but none were correctly inserted into the membrane. In one instance, one of the models was located lengthways along the external surface of the lipid bilayer and in another instance the MOMP models associated but were both lying on top of the membrane (orientation D). Based on experimental

knowledge of the COMC (Liu *et al*., 2010) and the typical behaviour of β-barrel proteins (Fairman *et al*., 2011), which both suggest OM insertion, these orientations were not considered to be biologically relevant and were consequently disregarded.

5.5.1.3 Interaction assessment of associating MOMP models

The interactions formed between the two MOMP models (from orientation B) were classified using an additional script, which identified 37 interacting residues with both 45 Å and 80 Å cut off distances, indicating that all residues were accounted for in both instances. It was clear from the resulting graph (**Figure 5.7**) that none of the key interacting residues were cysteines, as might be expected. In fact, not even the weakly associated residues, coloured white according to B-factor (**Figure 5.8**), were cysteines.

Further investigation of the initial MD simulation script revealed that intermolecular disulphide bonding as a potential interaction was not explicitly accounted for; only standardised van der Waals, hydrogen bonding, and salt bridges were fully parameterised. So even if two cysteines were aligned and disulphide bonding could feasibly occur between the two MOMP models, the algorithm wouldn't recognise -S-S- as an interaction, since the algorithm cannot compute the associated force fields and transition barriers arising from a disulphide linkage. As some proteins possess intramolecular disulphide bonding for structural stability, it is possible to account for this type of interaction in the initial CG protein set up, however, as the predicted bonds between MOMPs are intermolecular (between other proteins), to create such interactions would require two MOMP models to be reclassified as a single protein, with the disulphide bonds predefined as intramolecular interactions. Significantly, this approach would require biased alignment of the two models based on explicitly defining the cysteine residue location and their predicted

orientation/interaction and, therefore, the simulation would only serve to confirm if the models were correctly inserted within the membrane or not, as disulphide interactions would already be predefined. There are nine cysteine residues in *C. pneumoniae* MOMP, which creates a huge subset of potential combinations. Due to the extensive time required to map and define all these interactions, in addition to the huge bias and lack of real information achieved from running such a simulation, this option was not pursued. However, based on the crystal structure sensible pairings can be made relating to their location within the barrel, their outward orientation, and whether these are likely to be situated within the membrane region. Thus, potential interacting residues were mapped in conjunction with the cysteine residues to determine any correlation. Interestingly, the three cysteine residues at positions 136, 201, and 203 together with 226 near the top of the barrel were all situated in close proximity to one of the interacting faces (**Figure 5.9**), with residues C342 and C56, and also likely C52 and C49, all on the opposing side. Based on these observations, it is possible that the residues in the cysteine rich region of 136, 201 and 203 were interacting with the cysteine residue on the opposing face of the barrel at position C342 and, that C52 was interacting with C226 due to their complementary alignment, although C56 and C49 may also be involved, as shown in **Figure 5.21**.
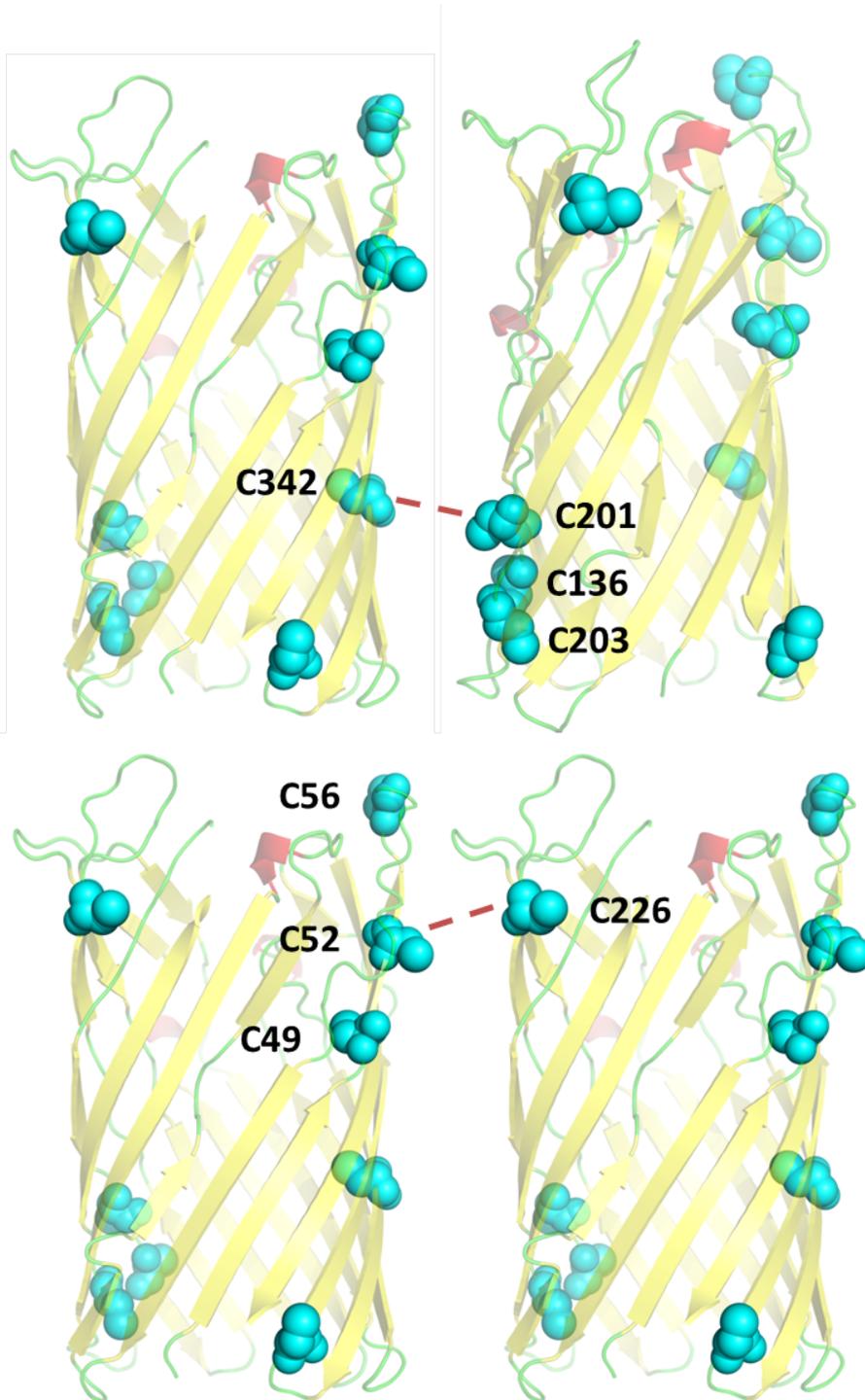
Figure 5.21. **Two MOMP crystal structures indicating the hypothesised intermolecular disulphide bonds between C342 and C201, and C52 and C226.** The nine cysteines residues are shown in cyan, with key residues labelled. Hypothesised intermolecular disulphide bonding is indicated by red dashed lines.

5.5.2   Use of fluorescence microscopy to understand MOMP's OM localisation

5.5.2.1 Cysteine mutant design

In order to determine which of the nine cysteine residues might play a significant role in intermolecular disulphide bonding between MOMP molecules, as a model of the COMC, mutants were carefully designed based on the low resolution crystal structure, the MD simulations, and the novel cysteine rich pocket hypothesis developed during the course of this work. Using the crystal structure, the putative location of all nine cysteine residues can be visualised. Interestingly, some of these residues appear to group together, for example, those residues at positions 136, 201, and 203, and those at positions 49, 52 and 56. This supports the novel hypothesis proposed in this work suggesting that, due to the large number of cysteine residues amongst the 17 proteins of the COMC, intermolecular disulphide bonds may be generated in cysteine rich regions within a pocket, rather than towards specific residues. With 194 documented cysteine residues and a large number of interacting proteins within the COMC, specific interactions appear to be not only unlikely, but also energetically unfavourable in terms of locating the correct protein and organising the correct orientation. Whilst MD simulations did not directly characterise the intermolecular disulphide bonds, it provided an indication of which faces of the β-barrel were interacting. From these regions, a number of cysteines were identified, namely C136, C201, and C203. Therefore, these residues were selected based on their location at the interface of the interacting region as well as the fact that these form a cysteine rich pocket on the outer face of the barrel. Additionally, C226 was also selected for mutation in order to characterise all the available residues within this potential interacting region.

As the only experimental data characterising potential disulphide bonding in MOMP was conducted for *C. trachomatis* MOMP, were obtained from peptide fragments that could

have re-associated in a number of ways, and gave rise to conflicting results across two independent research groups (Yen *et al.*, 2005; Wang *et al.*, 2006), these data could not be considered reliable for use in the cysteine mutant design and subsequent analysis presented herein. Furthermore, there is currently no published data available regarding how *Chlamydia* may form disulphide bonds, which renders predictions based on this new hypothesis entirely plausible and contributes significantly new knowledge to the field.

Once rationalised, all mutants were successfully created in the pETMOMP$_{H6}$ construct using SDM, however, this technique did not yield results with the pOPINHaloMOMP construct, which instead required In-Fusion cloning of the mutated sequence into the linearised vector. This vector suite possesses the genes required for expression in multiple hosts such as *E. coli*, mammalian cell lines, and insect cells, for parallel screening. Due to this, a β-globin polyA signal sequence is included in the vector in order to increase mRNA stability during mammalian cell expression. This repeated sequence of adenines likely disrupted SDM attempts, through non-specific annealing with the designed primers, resulting in the multiple faint bands observed in agarose gel electrophoresis of the PCR products.

All cysteines were mutated to alanine in order to neutralise the effect of the side chain, through truncation, whilst also maintaining the hydrophobicity of the region. Additionally, mutation to serine may have resulted in salt bridge or hydrogen bond formation, due to serine's polarity.

5.5.2.2 Growth and expression testing of initial pETMOMP$_{H6}$ cysteine mutants

Initially, only mutants C201A and C203A, plus their double mutant, were created in the pETMOMP$_{H6}$ construct for preliminary fluorescence microscopy experiments. As expected, in a growth curve experiment the control C41 (DE3) *E. coli* cells not expressing MOMP grew

more quickly and to a higher OD_{600}, with each mutant and the native MOMP growing at rates comparable to one another for the total 10h duration (**Figure 5.11**). This comparable rate of growth suggests that none of the mutants were more toxic on *E. coli* growth than the native MOMP protein itself, which would exhibit as lower OD_{600} measurements at equal time points. Additionally, the comparable growth rates suggested that each mutant MOMP was being over-expressed, as none grew more quickly or reached a higher OD_{600}, which one might expect if the protein was not being synthesised. Following this simple experiment, a medium scale expression test was conducted on the same constructs, which confirmed equivalent MOMP expression through Western blotting (**Figure 5.12**). The additional bands present in the Western blot, at a lower molecular weight, were likely due to freeze-thaw associated degradation. At this stage, it was not possible to determine if MOMP was expressed in the inner or outer membrane. However, as the membrane fraction was isolated through centrifugation, it was unlikely that MOMP was being expressed as inclusion bodies, which are typically discarded along with unlysed cells during the first 12,000 rpm centrifugation after cell lysis. Nevertheless, fluorescence microscopy was used in subsequent experiments to confirm membrane localisation of rMOMP.

5.5.2.3 Sample preparation for epifluorescence and TIRF microscopy

During the course of this work, it was necessary to optimise a number of parameters to achieve the highest image quality for preliminary epifluorescence and TIRF microscopy. These are highlighted herein. Following poly-L-lysine treatment of the microscope slides, the bacteria were plated using a range of dilutions in order to obtain an equilibrium between undesirable confluence, which would make membrane regions hard to distinguish, and a sparse distribution, which would require capture of a larger number of images to reasonably

characterise the population. An OD_{600} value of 0.5, equating to approximately $4x10^8$ cells, measured at the time of harvesting was consistently successful in this regard. Standardising the cell number for each sample also assisted in determining if a particular mutation was resulting in cell death and, therefore, resulting in increased particulate matter, which also scatters light at 600 nm and increases the OD_{600} value, whilst also evident by a decrease in cell number.

When imaging at 488 nm, it was important to account for bacterial autofluorescence, which can prove problematic during the detection of specific but weak fluorescence with green-labels, such as GFP or the novel his-tag binding P3NTA probe. LB and other rich media typically contain yeast extract and tryptophan, both of which contribute to autofluorescence through excitation of their aromatic groups under UV and blue light. Alternatively, cultures can be grown in MM, although it is important to be aware that cell morphology is evolutionarily affected by the decrease in nutrient availability and must be considered when analysing cell features, such as membranes. Due to the associated changes in cell morphology with growth in MM, cells were grown in LB but were washed three times in PBS at low speed centrifugation, before plating, in order to limit the aforementioned autofluorescence. Furthermore, it has been well established that naturally occurring metabolites produced in bacteria, such as flavins and nicotinamide adenine dinucleotide (NADH), contribute to autofluorescence at this wavelength. To account for these effects, unstained control C41 (DE3) *E. coli* cells were imaged with the 488-laser and settings were adjusted so that the autofluorescent signal was no longer perceivable. These baseline settings were then used in all epifluorescence and TIRF imaging of cells labelled with either the P3NTA probe or Alexa-488 antibody.

As discussed previously (**Chapter 1 Section 1.6.5**), due to the suspected inaccessibility of the N-terminus within the β-barrel, MOMP was C-terminally his-tagged. However, as the C-terminus of MOMP likely lies at the periplasmic side of the OM, membrane permeabilisation was mandatory to permit access of the chosen fluorophore to the his-tag. Despite permeabilisation with 0.1% Triton X-100, staining with anti-his primary antibody conjugated to a secondary Alexa fluorophore appeared erratic and weak. It was hypothesised that the peptidoglycan layer that lies between the inner and outer membrane in *E. coli* was restricting access of the primary and/or secondary antibodies. In response to this, an additional treatment step involving incubation with 100 μg/ml lysozyme immediately succeeded cell permeabilization, in order to hydrolyse the linker bonds within the peptidoglycan and improve probe access. The improvement in fluorescent signal observed in lysozyme treated cells was dramatic, with cell morphology now clearly visible in the fluorescent images for every cell identified with brightfield imaging (**Figure 5.13**), as opposed to the weak and patchy fluorescence observed in the absence of lysozyme. Consequently, it is strongly recommended that a lysozyme treatment step be included when fluorescently labelling a tag exposed to the periplasmic region within a Gram negative bacterium, such as *E. coli*, in order to improve the fluorescent signal.

5.5.2.4 Alternative labelling techniques for preliminary experiments

An anti-MOMP antibody obtained from the *Chlamydia* biobank and produced from *C. pneumoniae* was not effective at labelling recombinant MOMP for fluorescence microscopy imaging in whole cells or Western blot analysis of membrane preparations. Being an OM protein in *Chlamydia*, it has been widely accepted that MOMP is externally exposed and, therefore, natively produced antibodies would not require membrane permeabilisation to

gain access to the desired MOMP epitope. Initially, in fluorescence microscopy analysis, it was expected that MOMP was not correctly inserted into the *E. coli* OM, as no fluorescence was observed using a range of dilutions of the *Chlamydia* derived antibody. However, this was investigated further, through Western blotting of a number of different membrane preparations including those from both pETMOMP$_{H6}$ and pBADaMOMP (used for the functional growth assay, **Chapter 4**) constructs in addition to purified MOMP, none of which revealed any observable bands in a Western blot analysis. The method by which this antibody was produced in *Chlamydia* was unclear. However, purified MOMP is notoriously hard to obtain and as it is a major contributor to the *Chlamydial* COMC, it is possible that some additional residual COMC proteins were present and bound by intermolecular disulphide bonds. Therefore, when tested in *Chlamydial* samples antibody binding occurs, whereas in *E. coli*, without the presence of the additional epitope regions of other, possible contaminating COMC protein(s), antibody binding cannot occur. Alternatively, it is entirely possible that the antibody had degraded at some stage, dramatically reducing its efficacy, although both theories are difficult to investigate in the absence of native MOMP controls.

Additionally, the novel P3NTA probe, as used in detergent screening with FSEC to fluorescently label the his-tag, was tested as a potential single step labelling method. However, photobleaching occurred extremely quickly and did not tolerate the capture of both epifluorescence and TIRF images and, therefore, use of this probe would certainly not be amenable to super-resolution imaging with an ultra-high power laser. Redesign of the probe with a more robust fluorophore such as Alexa-647 was considered, however, due to the excessive expense, the time required for synthesis, and the low binding strength of the NTA region to this his-tag, this option was ultimately disregarded. Consequently, primary

anti-his antibody was conjugated to secondary Alexa-488 fluorophore for preliminary epifluorescence and TIRF imaging.

5.5.2.5 Clustering differences observed with epifluoresecence and TIRF between native and mutant MOMP

Upon establishment of a robust and efficient protocol for fluorescence imaging of recombinant MOMP in *E. coli*, the three cysteine mutants, C201A, C203A, C201/203A were imaged along with OmpA, a natively expressed *E. coli* outer membrane protein control. From these images it was apparent that native MOMP exhibited a speckled (uneven) membrane distribution, in both epifluorescence and TIRF, hypothesised to result from intermolecular disulphide bonding, resulting in unevenly distributed protein clusters in the membrane. As MOMP is the most abundant COMC protein, it is likely to be forming these intermolecular disulphide bonds with other MOMPs and, from these images, it was apparent that MOMP displayed the same behaviour as hypothesised when recombinantly expressed in *E. coli*. OmpA served as an excellent control to demonstrate homogeneous membrane dispersion, which was indisputably distinct from that of native MOMP. It was hypothesised that mutation of particular cysteine residues considered to be important for intermolecular disulphide bonding would result in a more homogenous membrane localisation pattern, more akin to that of OmpA. From the epifluorescence images of both C201A and C201/3A it appeared that, whilst not as intense as OmpA (**Figure 5.14A**), these mutants had somewhat increased the homogeneous dispersion (**Figures 5.14C and 5.14E**, respectively). For the C201/203A sample, TIRF microscopy supported this observation (**Figure 5.14E**). However, for the single C201A mutant, the TIRF microscopy image still exhibited a reasonable level of clustering (**Figure 5.14C**). Additionally, the C203A single mutant appeared to have had a negligible effect upon MOMP clustering, observed in both

techniques (**Figure 5.14D**). As these data are solely qualitative, it was difficult to draw robust conclusions regarding the mutants' effect on MOMP clustering, although this preliminary experiment was vital in indicating an observable difference that warranted additional investigation using a higher resolution technique that could provide quantitative data.

5.5.2.6 Additional cysteine mutant design and expression

As discussed (**Section 5.5.2.5**), preliminary fluorescence microscopy of recombinant MOMP in *E. coli* indicated a decrease in clustering for mutants C201A and C201/203A. To expand the breadth of this investigation, additional mutants were designed. As residues C201 and C203 exist in a hypothesised cysteine rich pocket, their accompanying residue C136 on the neighbouring β-strand was mutated singly and also coupled with C201 and C203 to create two new double mutants, namely C136/201 and C136/203. Residue C226 at the top of the barrel was also mutated to alanine, due to its alignment with this interacting face as also indicated by MD simulations of the two MOMP models within the lipid bilayer. Additional medium scale expression tests were conducted on each of the new mutants, with membranes prepared for analysis with Western blotting, as before. From this Western blot it was clear that all four of the novel MOMP mutants expressed protein at comparable levels to that of the native protein and the original three mutants (**Figure 5.15**), again suggesting that these mutations did not adversely affect protein folding and export to the membrane, as inclusion bodies, where misfolded proteins accumulate, and are largely eliminated at the first centrifugation step following cell lysis.

5.5.2.7 Advancement to super-resolution STORM

In order to obtain both higher resolution images and quantitative data for an analysis of clustering, super-resolution STORM imaging was conducted. The established protocol at the CLF Harwell, where the STORM experiments were conducted, was to use inverted superfrost microscope slides with the sample immersed in buffer and sealed with a 0.17 mm coverslide. In contrast to the Ibidi slide method adopted in TIRF microscopy, the samples were treated and labelled in microcentrifuge tubes, as opposed to being plated first. In order to wash the cells after each treatment, cells were pelleted by low-speed centrifugation and resuspended three times. Accounting for the total number of steps involved in the preparation procedure, samples were subjected to centrifugation and resuspension around seventeen times and, whilst care was taken to gently resuspend each pellet, the cells prepared this way exhibited very poor labelling. It was concluded that the shearing forces involved in both pelleting and resuspension resulted in damage to the cell architecture, especially the exposed outer membrane, which led to suboptimal labelling. Fortunately, as Ibidi slides consist of a 0.17 mm glass bottom, their use was considered appropriate for STORM imaging and sample processing using this technique was considerably more efficient, as the cells were plated first prior to treatment and labelling. As each slide has eight separated wells, samples could be incubated with the same volume over identical areas, as opposed to traditional microscope slides where aqueous solutions are difficult to confine to one region and tend to spread across the slide. Additionally, with eight wells per slide, the efficiency of data collection was dramatically improved since, to exchange samples, the chamber location can simply be adjusted without complete removal of the slide, thus eliminating the need to relocate the plane of focus with each sample change. As the wells could accommodate a total volume of 300 μl, the risk of imaging buffer

evaporation also was eradicated, as opposed to the coverslide method whereby it was only feasible to add 10 µl of buffer and, therefore each sample had to be incubated directly with buffer prior to imaging to limit evaporation.

5.5.2.8 Investigation of HaloTag-SiR as a labelling method

A new construct known as pOPINHaloMOMP was created in order to tag the C-terminus of MOMP with the Halo-tag protein, to allow irreversible covalent labelling with the cell permeable fluorophore, HaloTag-SiR. Initially, only C201A, C203A and C201/3A mutant pOPINHalo constructs were created in order to test the robustness of this labelling method. As SiR is cell permeable, the Triton X-100 detergent treatment step was omitted during preparation of this construct with the Ibidi slide method. Upon imaging of labelled cells, an unusual 'criss-cross' pattern was observed for each of the constructs instead of a distinct fluorescent signal at the membrane region, as in epifluorescence and TIRF microscopy. Initially, it was suspected that this unfamiliar signal was a result of many cells clumping together and overlapping. The importance of brightfield imaging (used to observe cells with conventional light microscopy) was emphasised at this stage, which revealed that the fluorescent signal originated from a single cell. The pattern appeared to be angular, almost forming a helix around the outside of the cells, although it wasn't clear if the fluorescence detected in the central region of the cells was from inclusion bodies or was connected in a link to the fluorescence spots nearer to the membrane region. Subsequently, 3D STORM was used to investigate this in two independent samples, both of which demonstrated a continuous band of fluorescence looping and twisting around the cell, as opposed to a shell-like layer of fluorescence surrounding the outside of the membrane, as expected and previously observed. As 3D STORM was considerably more time consuming than standard

2D STORM, both in terms of data collection and calibration with predefined probes, only two pOPINHaloMOMP samples were tested before this method of labelling was ultimately retired.

It is hypothesised that, due to the large size of the Halo-tag at 33 kDa, particularly in comparison to the size of MOMP (40 kDa), protein expression, folding, and/or export was disrupted. Previous attempts to tag MOMP with GFP, a protein of 27 kDa, were unsuccessful due to poor expression (Atanu, 2014) and, therefore, it is not entirely unexpected that this effect has also been exhibited here with the HaloTag. As a result, an anti-his Alexa-647 conjugated antibody was used with the pETMOMP$_{H6}$ construct for all future STORM experiments, with the re-addition of the permeabilisation treatment step. Since the native *E. coli* protein OmpA was, of course, not his-tagged, the primary anti-OmpA antibody was conjugated to a secondary Alexa-647 antibody so that the same grouping parameters, to account for trails, could be applied to both control and test samples.

5.5.2.9 STORM imaging and results

Alexa-647 is regarded as one of the best fluorophores for STORM, mainly due to its effective blinking and high photostability (Dempsey *et al*., 2011). These characteristics were observed in the STORM data collection experiments for the pETMOMP$_{H6}$ mutants using the Ibidi slide method, where 20,000 frames were collected before any decrease in blinking efficiency or photobleaching occurred. In total, at least ten STORM data sets were collected for each sample, under the same conditions, in order to obtain at least fifty regions of interest and to provide substantial data for analysis. Additionally, C41 (DE3) control cells were imaged with and without labelling with Alexa-647 antibody, as before, so as to identify non-specific binding and autofluorescence, respectively. Non-specific binding was detected to a small

233

degree in control samples incubated with the standard 1/100 concentration of antibody in 1% (w/v) BSA and, therefore, to completely eliminate this effect a higher concentration of 3% (w/v) BSA was certified and used for all test samples to ensure that the fluorescent signals detected arose from specific binding.

In all cases, following collection, data were processed to account for camera drift and trails. In comparison to the images captured in the preliminary epifluorescence and TIRF experiments, the STORM images acquired were of considerably higher resolution, predicted between 20-30 nm from the localisation precision histogram. However, with over ten images for each sample, and a number of cells in each image, visual comparison of clustering behaviour again was complex and subjective, due to the qualitative nature of the images.

5.5.2.10   Clustering analysis using ClusDoC software

However, with STORM the user acquires more than just a high resolution image. Each fluorescent spot visible within the STORM images actually has localisation data associated with it, all of which can be extracted as a simple text file using the Zen Black software (Zeiss). This information can then be uploaded into the ClusDoC software (Pageon *et al*., 2016), which reconstitutes the original image within the GUI. As bacteria are so small, this feature was particularly useful in that within each image there were varying numbers of cells present and, therefore it was more appropriate to analyse clustering within each cell as opposed to across all cells in combination within one image. In doing so, each cell was accounted for and individually contributed to the data set, which then could collectively be statistically processed in the same manner.

Following clustering analysis, a measure referred to as the 'relative density in clusters', an arbitrary value which accounts for the number of fluorophores and the size of the ROI, was used to compare clustering between samples. Due to the arbitrary nature of this measurement, it was imperative to include control OmpA data in the study, to act as a reference point for low clustered samples. Data from each ROI was collated and averaged, before the standard error was calculated to represent the accuracy of the data. A smaller SE reflects a more representative sample and mutants were deemed to be significantly different to the native MOMP if their SEs did not overlap. From the results (**Figure 5.20** and **Table 5.1**), it is evident that native MOMP was the most clustered sample with OmpA (representing homogeneous dispersion and low clustering) demonstrating a considerably lower score, both of which supports the observations made in the preliminary epifluorescence and TIRF microscopy experiments. Statistical analysis with a one-way ANOVA revealed a p-value <0.0001, permitting rejection of the null hypothesis, thus indicating a statistically significant difference between groups. In order to determine where these differences arose, a Tukey post-hoc test was conducted.

From comparison of the SEs in **Figure 5.20**, the mutants with the most significant decrease in clustering were the double mutants C201/203A and C136/201A, as well as the single mutant C226A, although all other mutants also displayed significant decreases compared to native MOMP. This was supported by the Tukey post-hoc test which indicated significant differences between all seven cysteine mutants in comparison to native MOMP, with the aforementioned mutants showing no significant difference to OmpA, suggestive of comparable membrane homogeneity (**Table 5.3**). With regard to the double mutants, notably, their respective single mutants did not exhibit the same decrease in clustering, indicating that two mutations in combination were required to produce a more dramatic

response. Contrary to this, double mutant C136/203A did not decrease clustering to the same extent as that observed for the other two double mutants. For the C136/203A double mutant, the data are suggestive of something paradoxical; it suggests that residue C201 has a more critical role in disulphide bond formation than C203, but contrarily, the effect of the single C201A mutant was less significant than that of the C203A single mutant.

Analysis of the putative location of the cysteines in the low resolution MOMP crystal structure, together with the interaction assessment data from the MD simulations, suggests that residue C342, on the opposite side of the barrel, is most likely interacting with the cysteine rich pocket (C136, C201, and C203) investigated here (**Figure 5.21**). With regard to the novel hypothesis, that the COMC proteins target cysteine rich pockets during disulphide bond formation, as opposed to specific residues, these results suggest that when only one of the cysteine residues within the C136, C201, and C203 pocket is mutated, the two remaining cysteine residues can compensate and continue to form the disulphide bond. However, with reference to the paradox previously mentioned, it is hypothesised that residue C201 is much more effective at forming the disulphide bond alone when C136 and C203 are mutated, in comparison to either C136 or C203 when singularly available. This may be because residue C201 is situated higher within the β-barrel and more in alignment with C342 and, therefore, sterically is able to form a more favourable interaction, although the neighbouring residues are still available to form a compensatory disulphide bond if required. From these results, it is evident that some compensatory mechanism is occurring, otherwise it would be expected that mutation of one particular residue would cause a decrease in both single and double mutants in equal (or greater) measure.

Residue C226A also had a significant effect on clustering, despite being a single mutation that is not arranged in a cluster. Again with respect to the MD simulation interaction data

and the cysteine residue locations within the β-barrel, it is hypothesised that this residue is key in forming disulphide bonds with an additional cysteine rich region at residues C49, C52 and C56. Due to its alignment with C226, it is possible that residue C52 would form the most favourable interaction, with its flanking residues C49 and C56 available for compensation (**Figure 5.21**).

5.6 Conclusions

MD is a powerful computational technique that allows us to model the behaviour of proteins that may be difficult to characterise experimentally. In this research, CG MD simulations modelled the insertion and behaviour of two low resolution MOMP models within an asymmetric lipid bilayer. From four initial orientations, only one resulted in correct insertion within the bilayer, which also resulted in a plausible protein-protein interaction. The resulting simulation was considered biologically relevant, as published experimental data to date indicates that MOMP is expressed within the outer membrane (Liu *et al.*, 2010; Sun *et al.*, 2007; Wolf *et al.*, 2001) and, therefore, any deviation from this is less likely to be biologically relevant. Subsequent assessment of the interacting residues revealed that none were explicitly a result of disulphide bonds. Investigation of the original script revealed that only predefined intramolecular disulphide bonds could be accounted for and, therefore, mapping of potential intermolecular disulphide bonds between two MOMP models would require the reclassification of the two models as one protein and manual assignment of the interacting cysteines. This process would, of course, not only be extremely time consuming but also result in a high degree of bias, the results of which would be difficult to assess. However, upon closer inspection, the interacting faces of the barrel were located in close proximity to a number of cysteine residues, indicating that

these residues may have a role in assisting the alignment of MOMP into the correct orientation for disulphide bond formation.

Thus a novel hypothesis emerged, during the course of this research, which suggests that, due to the high number of COMC proteins and consequently the numerous cysteine residues available for disulphide formation, the COMC proteins target cysteine rich pockets or regions as opposed to specific residues. Therefore, the cysteines in the region of C136, C201, and C203 of MOMP were mutated to alanine in order to neutralise their disulphide bonding ability, the effect of which was observed with fluorescence microscopy. During the preliminary stages, only C201 and C203 were mutated and analysed, along with native MOMP and control OmpA, with lower resolution epifluorescence and TIRF microscopy. From these images, it was apparent that MOMP, as hypothesised, was forming clusters within the *E. coli* membrane, characterised as speckled fluorescence using both techniques. Two of the mutants indicated more homogenous membrane dispersion with epifluorescence, although the single C201 mutant still appeared speckled with TIRF. Due to the qualitative nature of epifluorescence and TIRF microscopy imaging, a higher resolution technique was pursued to enable further detailed characterisation of these differences.

For super-resolution STORM, the C136 residue was mutated both singularly and in pairing with its neighbouring residues C201 and C203. Additionally, residue C226 near the top of the barrel and at the interacting interface, as predicted by MD simulations, was also mutated. Initially, for this higher resolution technique, it was intended that the membrane permeable HaloTag-SiR fluorophore would be used for labelling of MOMP in order to preserve the membrane integrity. However, use of this fluorophore required expression of MOMP within the pOPINHalo construct, which upon imaging exhibited an unusual labelling pattern most likely resulting from disrupted folding and export of MOMP to the membrane.

Alternatively, conjugated anti-his Alexa-647 fluorophore was applied with the pETMOMP$_{H6}$ construct which was used in preliminary experiments and was also known to express MOMP within the *E. coli* membrane due to previously observed sarkosyl insolubility, as discussed in **Chapter 3**. The subsequent efficiency of the Alexa-647 fluorophore permitted superior STORM data collection, resulting in a series of high resolution images. The localisation data from each image was analysed using clustering software, which revealed MOMP to be highly clustered and OmpA as more homogeneous, as hypothesised. Double mutants C201/203A and C136/201A reduced the clustering of MOMP most significantly, suggesting that within cysteine rich regions a compensatory mechanism is occurring, whereby neighbouring cysteine residues can continue to form intermolecular disulphide bonds in the absence of the most important residue. Notably, the single mutant C226A produced a significant effect upon disulphide bonding. In relation to the interacting regions predicted by MD simulations, it is likely that this residue is interacting with an additional cysteine rich region, likely formed by residues C49, C52, and C56. These results support the hypothesis that during COMC formation a general cysteine rich region is targeted, as opposed to specific residues, although it is possible that some residues within these regions, such as C201, have more substantial roles, perhaps even depending on the specific target protein partner in the COMC. However, until high resolution structural data emerges detailing the location of the cysteine residues within these other partner proteins of the COMC, then predictions of these interactions remain extremely challenging.

Chapter 6

General discussion and recommendations for future research

6   General discussion and recommendations for future research

Chlamydia is a widespread bacterial pathogen most infamous for its sexually transmitted infection caused by *C. trachomatis. C. pneumoniae*, another less well-known strain of this genus, most often implicated in human respiratory diseases such as pneumonia, sinusitis, and bronchitis, also has a more controversial role in chronic inflammatory diseases such as atherosclerosis (Sorrentino *et al*., 2015), asthma (Webley and Hahn, 2017), and arthritis (Zeidler and Hudson, 2016), following infection during the early stages of life. Common to each strain of *Chlamydia* is the major outer membrane protein, an abundant surface exposed cysteine rich protein, which is highly conserved amongst the species with demonstrable immunodominance in both humans and mice (Wolf *et al*., 2001; Molina *et al*., 2010). Despite the huge global economic and health impact arising from *Chlamydia* infections, there are currently no suitable preventative vaccines. Currently, infections are treated with antibiotics such as azithromycin and doxycycline, however the increasing threat of antimicrobial resistance has the potential to soon render these methods nonviable, increasing the drive towards novel vaccine therapeutics. Being surface exposed, conserved, and highly abundant, accounting for 60% of the outer membrane protein, MOMP has long been the focus of vaccine studies (Kari *et al*., 2009; Olsen *et al*., 2015; O'Meara *et al*., 2017; Xia *et al*., 2013). Despite this, there is currently no high resolution structural data for MOMP, which would provide vital information regarding both the surface exposed domains and epitopes, as well as provide valuable insight into MOMP's function, which could then be targeted for novel antimicrobials. CD spectroscopy has been used by multiple researchers to confirm MOMP's β-sheet content (Wyllie *et al*., 1998; Sun *et al*., 2007; Cai *et al*., 2009; Wen *et al*., 2016), with homology modelling used to make predictions on more specific features

and side chain orientations, although due to the paucity of highly homologous proteins with structural data, results are conflicting (Atanu *et al*., 2013; Feher *et al*., 2013).

Within the past three years, low resolution structural data for MOMP was obtained to a resolution of 4 Å using molecular replacement, by this research group (Atanu, 2014). Although the barrel structure was evident, it was not possible to accurately determine individual side chain orientations or to trace the extracellular loop regions with convincing electron density. As a result, the initial focus of this research was toward producing stable, active, and correctly folded recombinant MOMP to be used in structural studies, namely in optimising crystallisation conditions for analysis with X-ray diffraction, in an effort to produce quality higher resolution structural data. Not only would this provide valuable insight into MOMP's structure, but also its potential functional role, information which could be transferrable to the other *Chlamydial* strains due to MOMP's highly conserved nature in this organism. Due to the presence of an occluded region within the barrel of the density map, as modelled by the 14-stranded homology model (Atanu *et al*., 2013), this homology model was primarily used in cohort with the low resolution structural model in this research.

The protocol for rMOMP expression and purification that previously resulted in diffraction quality crystals was not replicable, despite extensive and multiple attempts. It was hypothesised that increased MOMP expression through improved aeration resulted in protein that was more difficult to extract with the mild detergent DDM, likely due to the increase in intermolecular disulphide bonding and, therefore, substantial efforts were placed on solubility screening in order to improve the yield. The two techniques adopted, Western blotting and FSEC, identified zwitterionic SB3-14, also used in the solubilisation of

C. jejuni major outer membrane porin (Ferrara *et al*., 2013) as well as native and recombinant *Chlamydial* MOMP (Findlay *et al*., 2005; Pal *et al*., 2005), as the most promising lead. Further comprehensive optimisation of the IMAC purification procedure successfully enhanced the MOMP yield, which was then thwarted by the concentration of excess detergent micelles and proved detrimental to rMOMP stability. Instead, during the purification procedure MOMP was buffer exchanged into DDM, as used in the initial protocol, which due to its larger micellar size permitted MOMP concentration without the excessive build-up of smaller detergents SB3-14, OG, and sarkosyl. Under these conditions, MOMP was concentrated and used in crystallisation trials with two commercial screening kits explicitly designed for membrane proteins. Following a period of incubation, small crystalline products were observed in a number of wells, which were then analysed with a microfocus beamline *in situ*. The diffraction patterns observed from the majority of these crystals were likely due to detergent, although a small subset of crystals produced a regular pattern of reflections, indicative of protein. Data collection with a microfocus beamline often requires thousands of crystals to obtain a complete dataset, and therefore in this research this method was solely used for screening. Whilst the regular pattern of reflections observed was of low resolution, it is recommended that future crystallisation screening include the optimisation of this condition (D11, MemGold II). As MOMP buffer exchanged into detergent OG produced diffraction quality crystals in previous research (Atanu, 2014), MOMP concentrated in DDM was then buffer exchanged, using a desalting column, in order to concurrently decrease the salt concentration whilst also buffer exchanging into OG, also ensuring the removal excess detergent micelles. Whilst success was observed following this procedure during small scale experiments, using a larger culture volume and thus purification of a greater quantity of protein likely resulted in aggregation during

concentration over 1 mg/ml. Therefore, it is recommended that following overnight storage at 4 °C, samples be supplemented with additional TCEP in order to limit intermolecular disulphide bonding. Additionally, the flow rate used in buffer exchange should be maintained at the slowest rate possible to ensure complete buffer exchange for soluble and stable protein. Whilst high resolution quality crystals were not obtained, the robust and rigorous optimisation of each stage of rMOMP purification has resulted in higher yields than previously reported, with half the culture volume previously required. It must be noted that membrane proteins are notoriously difficult to purify to quantities suitable for crystallography whilst also maintaining a stable and active state, due to the requirement to extract them from their native lipid environment, which then must be detergent exchanged into a more crystallisation amenable detergent. There are a huge range of detergents available with a multitude of different properties for membrane protein solubilisation, and thus a suitable amount of time must be invested in detergent screening and optimisation. Additionally, as the detergent envelopes the hydrophobic regions of the protein, much of the protein is masked thus reducing the number of contacts that are able to form during the crystallisation process, and therefore a range of different detergents must also be tested to optimise crystallisation.

Not only is MOMP a membrane protein, but it is also cysteine rich, with one of its known functions being a major contributor to the COMC through the formation of intermolecular disulphide bonds. Therefore, it is to be expected that many if not all of the nine cysteine residues within MOMP are externally exposed and available for disulphide bonding, as opposed to the more common intramolecular disulphide bond used for protein secondary structure stability. As a result, many of the difficulties observed in solubility and stability

during the purification procedure were likely exacerbated by disulphide bonding. Even though reducing agents were utilised, the nature and potential number of disulphide bonds in MOMP suggests that these residues will unlikely be readily accessible to such agents, likely being protected by the inward barrel orientations. Despite the difficulties associated with membrane protein research, the method presented herein not only produces a greater yield but is more time efficient, removing the requirement to culture and process such large quantities, and is also more cost effective through reduction of the quantity of reagents required. The final adjustment and optimisation of MOMP concentration will no doubt lead to the production of sufficient protein for crystallisation trials, potentially leading to more robust crystals suitable for high quality data collection.

In the absence of structural data, the MOMP homology model, based on the fatty acid transporter FadL of *E. coli* developed by this research group (Atanu *et al*., 2013), was used in research investigating MOMP's functional role. This model indicated an external α-helical loop domain with an internal N-terminal plug occluding the β-barrel at the periplasmic side. The low resolution structure for *C. pneumoniae* MOMP recombinantly expressed in *E. coli* revealed density within the lumen of the barrel, concurrent with the FadL homology model plug domain. Typically, β-barrel proteins function as porins, which non-specifically transport small solutes (Fairman *et al*., 2011), with some reports also indicating porin function for MOMP (Bavoil *et al*., 1984; Jones *et al*., 2000; Sun *et al*., 2007). However, in this instance due to the occluded barrel it was hypothesised that MOMP may have a more specific and active role requiring remodelling of the tertiary structure to permit ligand passage, possibly mediated by the NPA and NPS motifs (Guan *et al*., 2010). *Chlamydia* also has an extremely small genome, and has been known to scavenge various components from the infected

host, such as long chain fatty acids palmitate and oleate for incorporation into the bacterial membrane (Yao, Dodson *et al.*, 2015). MOMP's thermal stability also increased in the presence of these two ligands during SRCD experiments (Atanu, 2014), which was suggestive of ligand binding. Taken together, homology with FadL, experimental evidence demonstrating host fatty acid uptake, and increased thermal stability upon fatty acid binding, provided the basis for a functional growth assay assessing the ability of MOMP to act as a transporter for fatty acids. Whilst MOMP exhibits high expression in the circulating and infectious EBs, within RBs expression is downregulated although not completed eliminated (Roulis *et al.*, 2013), which is suggestive of an additional role for MOMP in the metabolically active stage of the *Chlamydial* lifecycle, similar to PorB, which also functions as both a porin and a contributor to the COMC (Kubo and Stephens, 2002; Liu *et al.*, 2010). It is well known that *E. coli* is able to use LCFAs as a sole carbon source, with FadL the recognised fatty acid transporter (Kumar and Black, 1993; van den Berg, 2010) and, therefore, a number of *E. coli fadL* knockout strains were utilised in the growth assay development. In order to create a robust assay, it was essential that the knockout mutant be complemented with the pBAD construct containing the *fadL* gene, the same vector to be used for MOMP expression. It was also imperative that knockout strains grew as expected under various carbon sources, for example with no growth on LCFAs. Double knockout mutant, LS6164, was sourced from researchers who investigated FadL LCFA transport (van den Berg, 2004), which grew as expected. Rhamnose is a tuneable inducer that is able to control expression with the pBAD vector, which is largely dependent on the quantity of cellular permease for activity (Rosano and Ceccarelli, 2014). As FadL is a protein native to *E. coli*, over expression through rhamnose induction was in fact detrimental to growth, most likely due to overloading of the membrane. Instead, basal expression of FadL was sufficient

to reinstate LCFA transport in the knockout strain. Being a *Chlamydial* protein, it was uncertain how the level of MOMP expression would affect transport, and therefore a range of conditions were examined whereby rhamnose was included in plates, in the liquid medium for a variety of durations, or not at all. Interestingly, rhamnose included within plates (but not in liquid cultures) was highly toxic to both FadL and MOMP expressing cells, completely arresting growth in palmitate plates. This effect was also observed in glycerol plates containing rhamnose (**Figure 4.9A**), whereby LS6164 transformed with pBADaFadL showed individual colonies at a dilution of 10^{-2}, as opposed to confluent growth down to 10^{-4} in plates containing just glycerol (**Figure 4.10A**). In fact, the optimum growth conditions for FadL were achieved with basal expression, without rhamnose induction in plates or liquid media. Despite the multitude of conditions to encourage differing levels of MOMP expression, no growth of MOMP expressing cells was observed on LCFA plates, although MOMP exhibited comparable growth to FadL expressing strains on glycerol plates (**Figure 4.10B**), which suggested that MOMP expression was not toxic. Another hypothesis, as to why growth did not occur, was that MOMP was expressed as inclusion bodies, as opposed to within the outer membrane and thus not in the correct location for ligand transport. STORM, a super resolution microscopy technique, was used to identify MOMP's cellular location through comparison with control membrane protein OmpA, whereby MOMP appeared to be localised to the membrane. Therefore, a number of hypotheses were suggested as to why growth did not occur on a LCFA carbon source. Primarily, it is of course possible that, despite the homology to FadL and the increased thermal stability with oleate and palmitate, MOMP is not the fatty acid transporter of *Chlamydia*. Instead, either another outer membrane protein could be responsible, which at this stage is difficult to predict due to the lack of structural data for any OMP of *Chlamydia*, or that due to the truncated form

of LPS (also known as LOS) LCFAs are able to flip-flop across the outer membrane via a diffusion gradient (Hamilton, 2007). Alternatively, it is possible that MOMP requires a co-activator protein to assist in the active rearrangement of the tertiary structure to enable transport. Since the low resolution structure for MOMP does not indicate a lateral opening, similar to that observed with FadL (Hearn *et al.*, 2009), it is not expected that MOMP transports fatty acids in the same manner, although higher resolution data will confirm this. Additionally, it is currently unknown how FadL communicates with the inner membrane transporter FadD and, therefore, it is possible that LCFAs are transported into the periplasmic space by MOMP, but then are not able to access FadD for transport to the oxidation enzymes. Again, higher resolution structural data for MOMP will provide greater insight into the potential function of MOMP and the mechanisms that may be employed in ligand transport, which can then be targeted in *Chlamydia* studies with radiolabelled substrates, such as in the study conducted in the characterisation of FadL (Lepore *et al.,* 2011).

A more established role for MOMP is in the COMC, a network of cysteine rich outer membrane proteins linked via intermolecular disulphide bonding (Bavoil *et al.*, 1984; Hatch, 1996; Liu *et al.*, 2010; Newhall *et al.*, 1983). MOMP, being both cysteine rich and the most abundant outer membrane protein, naturally has an important role in this mechanism. The COMC proteins are upregulated during the EB stage of the *Chlamydial* lifecycle and act to provide structural stability to the circulating infectious cells in the absence of peptidoglycan, the sugar polymer found in virtually all bacteria. Despite playing such a vital role in EB stability, there is a paucity of data on the interactions formed between the seventeen contributing proteins, most likely due to the difficulties associated with the expression and

purification of cysteine rich outer membrane proteins. A number of studies have however been conducted on MOMP to analyse the disulphide bond interactions formed, although each report provides conflicting results whereby some cysteine pairs were purported to form intramolecular disulphide bonds whilst others suggested these residues were freely available to form intermolecular disulphide bonds (Wang *et al*., 2006; Yen *et al*., 2005). In an effort to help resolve these discrepancies and shed light on the role of specific cysteine residues, in this work, the low resolution crystal structure was used to rationally target specific cysteine residues for both site-directed mutagenesis studies and to perform preliminary CG molecular dynamics simulations. Whilst intermolecular disulphide bonding was not an interaction specifically accounted for during the MD simulation of two MOMP crystal structure models in an asymmetric bilayer, the resulting interaction interface aligned with many of the cysteine residues. It is possible that MOMP's cysteine rich structure encourages alignment of specific residues, through the location of particular hydrophobic regions. A novel hypothesis is proposed, arising from this research, suggesting that intermolecular disulphide bonds are formed between cysteine rich regions, as opposed to specific residues, in order to increase the likelihood of interaction and decrease the energy required to obtain the correct orientation in the COMC. Such a cysteine rich pocket (C136, C201, C203) was located in MOMP, using the low-resolution structure and coincident with the interacting plane identified in the MD simulations, and a number of mutant combinations subsequently were created. In preliminary experiments, TIRF and epifluorescence microscopy were adopted to visualise the clustering behaviour of MOMP within the *E. coli* outer membrane. Whilst it was possible to visually discern differences between the control protein OmpA, which exhibited homogenous membrane dispersion, to native MOMP, which appeared more clustered, and a number of the mutants, which

appeared to decrease clustering, it was not possible to quantify these variations. As a result, STORM was implemented to obtain both higher resolution images as well as quantitative data for comprehensive comparison. Initially, MOMP was to be labelled with a membrane permeable fluorophore known as HaloTag-SiR in order to limit disruption to the membrane architecture during the staining process. In order to do so, MOMP was expressed within the pOPINHalo vector, which introduces the HaloTag to the C-terminus of the POI, which is then able to form an irreversible covalent bond with the fluorophore (England *et al*., 2015). However, the size of the tag (33 kDa) likely disrupted MOMP folding and insertion into the outer membrane, as STORM imaging revealed an unusual geometric pattern. To further investigate this, 3D STORM was implemented in order to clearly discern the fluorophore localisation, which appeared to be looping around the cell as opposed to within the membrane region. As a result of this, Alexa-647 conjugated to an anti-his antibody was used for all future STORM experiments, selected as only one labelling step was required, and more importantly because Alexa-647 is reportedly the strongest fluorophore, both in terms of blinking and high photostability (Dempsey *et al*., 2011). Data were collected for mutants in the cysteine rich area, in addition to C226A, a residue situated at the top of the barrel in the same plane of orientation as the cysteine pocket. Whilst the images collected were clearly of a higher resolution, it was still difficult to draw qualitative comparisons between samples, and as a result the quantitative data arising from the localisations captured in each frame were analysed with clustering analysis software ClusDoC (Pageon *et al*., 2016). Comparison of the relative density in clusters revealed that the mutants with the most significant impact upon MOMP clustering were double mutants C201/203A and C126/201A within the cysteine rich pocket, in addition to the single mutant C226A located at the top of the barrel. Interestingly, the respective single mutants C136A, C201A, and C203A did not

decrease clustering to the same degree, suggesting that two mutants were required to observe an effect. Additionally, double mutant C136/203A did not exhibit the same effect, with its degree of clustering similar to that of the single mutants. This evidence strongly suggested that two mutants within the pocket were required to disrupt disulphide bonding, with residue C201 having a more important role than its neighbouring residues. In C201 containing mutants, the remaining residue was less able to compensate for the neighbouring mutants, as opposed to C201 which was in a position to maintain the interaction alone. This research is highly supportive of the cysteine rich pocket hypothesis proposed herein, as it appears that cysteine residues are able to compensate for neighbouring mutants, with some residues such as C201 having a more vital role. It may seem that the similar decrease in clustering observed with mutant C226A contradicts this theory, however, on the opposing face of the barrel is an additional cysteine rich pocket (C49, C52, C56), which is likely interacting with C226 due to their orientations. Therefore, it is proposed that C226 is a vital residue in this interaction, but that similarly any of the cysteines in the C49-C56 pocket can form the disulphide bond, with residue C52 likely being favoured due to its location. Similarly with the C136, C201 and C203 pocket, it is likely that residue C342 on the opposing face of the barrel is also vital to this interaction. To investigate these hypotheses, mutants should be created of the C49-56 pocket together with residue C342, as well as an additional triple mutant of the cysteine rich pocket analysed in this research.

Due to the difficulties associated with the extraction of rMOMP from the outer membrane, the cysteine mutants designed here, which exhibited the least clustering should be implemented in further purification and crystallisation studies. It is recommended that

mutants C201/203A, C136/201A, and C226A be detergent screened using FSEC, as previous solubilisation with SB3-14 revealed unusual migration patterns with Western blotting. As each mutant expressed comparable levels of MOMP to the wild type, it is anticipated that optimisation of mutant solubilisation will greatly assist in the purification procedure, namely during solubilisation and extraction but also in concentration, where it is expected that disulphide bonding is resulting in aggregation.

Despite MOMP accounting for 60% of the *Chlamydial* outer membrane protein, as well as displaying immunodominant features, little is known regarding its true structure and function. As MOMP is downregulated in the RB (Roulis *et al*., 2013), but not completely eliminated, it is possible that MOMP has a role in metabolic functions, such as the transport of important substrates. The transport of host derived substrates is highly likely, due to the small size of *Chlamydia*'s genome and the truncated or absent metabolic systems (Käding *et al*., 2014). Despite MOMP's homology to the fatty acid transporter FadL (Atanu *et al*., 2013), this research concluded MOMP's inability to transport fatty acids, although an improved structural model would provide insight into the plug domain and the potential mechanism employed in its possible reorganisation. Previous studies have reported the diffusion of small solutes, such as arabinose (Sun *et al*., 2007) and all twenty amino acids (Jones *et al*., 2000), through MOMP, albeit at a considerably slower rate compared to more well characterised porins, such as OmpF. It is therefore possible that small solutes are able to bypass the plug domain. Such a channel would be visualised more clearly with improved structural data. Although it has been established that MOMP contributes toward formation of the COMC (Liu *et al*., 2010), the exact interactions formed and with which of the other sixteen cysteine rich proteins, along with the paucity of knowledge on how disulphide

bonding is mediated within *Chlamydia*, has left many questions unanswered. As the COMC encompasses the whole of the EBs as these circulate the host searching for new cells to infect, it seems to be a vital target for antimicrobials and vaccines, not only due to its surface exposure but also due to its essential role in cell structural stability. Only by enhancing our understanding of both the interactions that stabilise this network and the structures of the contributing proteins, can therapeutics be rationally designed to disrupt this mechanism.

This research details original data pertaining to disulphide bonding in MOMP, proposing a novel hypothesis whereby cysteine rich pockets within the *Chlamydial* cysteine rich outer membrane proteins are responsible for clustering in the outer membrane, thus stabilising the COMC and contributing to structural stability of *Chlamydial* cells during infection. Additionally, the novel cysteine mutants designed herein are likely to facilitate improved purification of MOMP for structural studies, through reduced aggregation and improved membrane extraction, which will reveal not only important structural details but may also shed light on MOMP's functional role.

References

AbdelRahman YM., and Belland RJ. (2005) "The *Chlamydial* developmental cycle." *FEMS Microbiology Reviews*. **29**(5): 949–959.

Abraham MJ., Murtola T., Schulz R., Páll S., Smith JC., Hess B., and Lindahl E. (2015) "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers." *SoftwareX*. **1-2**: 19–25.

Atanu FO. (2014) *Structural studies of Chlamydophila pneumoniae Major Outer Membrane Protein (MOMP) and MOMP-derived peptides: candidate vaccine targets for human inflammatory diseases*. PhD thesis. University of Reading.

Atanu FO., Oviedo-Orta E., and Watson KA. (2013) "A novel transport mechanism for MOMP in *Chlamydophila pneumoniae* and its putative role in immune-therapy." *PloS One* **8**(4): e61139.

Backmark AE., Olivier N., Snijder A., Gordon E., Dekker N., and Ferguson AD. (2013) "Fluorescent probe for high-throughput screening of membrane protein expression." *Protein Science: A Publication of the Protein Society* **22**(8): 1124–32. doi:10.1002/pro.2297.

Baehr W., Zhang YX., Joseph T., Su H., Nano FE., Everett KD., and Caldwell HD. (1988) "Mapping antigenic domains expressed by *Chlamydia trachomatis* major outer membrane protein genes." *Proceedings of the National Academy of Sciences.* **85**(11): 4000–4004.

Bates M., Jones SA., and Zhuang X. (2013) "Stochastic Optical Reconstruction Microscopy (STORM): A method for superresolution fluorescence imaging." *Cold Spring Harbor Protocols.* **6**

Bavoil P., Ohlin A., and Schachter J. (1984) "Role of disulfide bonding in outer membrane structure and permeability in *Chlamydia trachomatis*." *Infection and Immunity* **44**(2): 479–85.

Beis K., Whitfield C., Booth I., and Naismith JH. (2006) "Two-step purification of outer membrane proteins." *International journal of biological macromolecules*. 39(**1–3**): 10–14.

Belland RJ., Ouellette SP., Gieffers J., and Byrne GI. (2004) "*Chlamydia pneumoniae* and atherosclerosis." *Cellular Microbiology*. **6**(2): 117–127.

Berg JM., Tymoczko JL., and Stryer L. (2002) "Three-Dimensional protein structure can be determined by NMR Spectroscopy and X-ray crystallography." *Biochemistry. 5th edition*.

Betts-Hampikian HJ., and Fields KA. (2010) "The *Chlamydial* type III secretion mechanism: revealing cracks in a tough nut." *Frontiers in Microbiology*. **1**: 114.

Betzig E., Patterson GH., Sougrat R., Lindwasser OW., Olenych S., Bonifacino JS., Davidson MW., Lippincott-Schwartz J., and Hess, HF. (2006) "imaging intracellular fluorescent proteins at nanometer resolution." *Science.* **313** (5793): 1642–45.

Black PN. (1990) "Characterization of FadL-specific fatty acid binding in *Escherichia coli*." *Biochimica et Biophysica Acta*. **1046**(1): 97–105.

Black PN., Said B., Ghosn CR., Beach JV., Nunn WD. (1987) "Purification and characterization of an outer membrane-bound protein involved in long-chain fatty acid transport in *Escherichia coli*." *Journal of Biological Chemistry.* **262**: 1412-1419.

Block H., Maertens B., Spriestersbach A., Brinker N., Kubicek J., Fabis R., Labahn J., and Schäfer F. (2009) "Immobilized-metal affinity chromatography (IMAC): a review." *Methods in Enzymology*. **463:** 439–473.

Bolanos-Garcia VM., and Davies OR. (2006) "Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography." *Biochimica et Biophysica Acta.* **1760**(9): 1304–13.

Bond PJ., Holyoake J., Ivetac A., Khalid S., and Sansom MSP. (2007) "Coarse-grained molecular dynamics simulations of membrane proteins and peptides." *Journal of Structural Biology*. **157**(3): 593–605.

Bornhorst JA., and Falke JJ. (2000) "Purification of proteins using polyhistidine affinity tags." *Methods in enzymology*. **326**: 245–254.

Brunham RC., and Rey-Ladino J. (2005) "Immunology of *Chlamydia* infection: Implications for a *Chlamydia trachomatis* vaccine." *Nature Reviews Immunology* **5**(2): 149–61.

Cai M., Huang Y., Sakaguchi K., Clore GM., Gronenborn AM., and Craigie R. (1998) "An efficient and cost-effective isotope labelling protocol for proteins expressed in *Escherichia coli*." *Journal of Biomolecular NMR*. **11:** 97–102.

Cai S., He F., Samra HS., de la Maza LM., Bottazzi ME., Joshi SB., and Middaugh CR. (2009) "Biophysical and stabilization studies of the *Chlamydia trachomatis* mouse pneumonitis major outer membrane protein." *Molecular Pharmaceutics* **6**(5): 1553–61.

Caldwell HD., and Schachter J. (1982) "Antigenic analysis of the Major Outer Membrane Protein of *Chlamydia* Spp." *Infection and Immunity.* **35**(3): 1024–31.

Caldwell HD., Kromhout J., and Schachter J. (1981) "Purification and partial characterization of the major outer membrane protein of *Chlamydia trachomatis*." *Infection and Immunity* **31**(3): 1161–76.

Carabeo RA., Grieshaber SS., Fischer E., and Hackstadt T. (2002) "*Chlamydia trachomatis* induces remodelling of the actin cytoskeleton during attachment and entry into HeLa cells." *Infection and Immunity.* **70**(7): 3793–3803.

Cardona ST., and Valvano MA. (2005) "An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*." *Plasmid*. **54**(3): 219–28.

Carter JD., Inman RD., Whittum-Hudson J., and Hudson AP. (2012) "*Chlamydia* and chronic arthritis." *Annals of Medicine*. **44**(8): 784–792.

Cowan SW., Schirmer T., Rummel G., Steiert M., Ghosh R., Pauptit RA., Jansonius JN., and Rosenbusch JP. (1992) "Crystal structures explain functional properties of two *E. coli* porins." *Nature* **358**(6389): 727–33.

Dempsey GT., Vaughan JC., Chen KH., Bates M., and Zhuang X. (2011) "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging." *Nature Methods.* **8**(12): 1027–36.

Dessau MA., and Modis Y. (2011) "Protein crystallization for X-ray crystallography." *Journal of Visualized Experiments : JoVE*. (**47**): 2285.

Domínguez-Escobar J., Chastanet A., Crevenna AH., Fromion V., Wedlich-Söldner R., and Carballido-López R. (2011) "Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria." *Science.* **333** (6039): 225–28.

England CG., Luo H., and Cai W. (2015) "HaloTag technology: A versatile platform for biomedical applications." *Bioconjugate Chemistry*. **26**(6): 975–986.

Ester M, Kriegel HP, Sander J, and Xu X. (1996) *A density-based algorithm for discovering clusters in large spatial databases with noise.* In: Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Menlo Park, CA: AAAI, 226–231

Fairman JW., Noinaj N., and Buchanan SK. (2011) "The structural biology of β-barrel membrane proteins: a summary of recent reports." *Current opinion in structural biology*. **21**(4): 523–531.

Feher VA., Randall A., Baldi P., Bush RM., de la Maza LM., and Amaro RE. (2013) "A 3-Dimensional trimeric β-barrel model for *Chlamydia* MOMP contains conserved and novel elements of Gram-negative bacterial porins." *PLOS ONE* **8**(7): e68934.

Feng Y., and Cronan JE. (2009) "*Escherichia coli* unsaturated fatty acid synthesis." *The Journal of Biological Chemistry.* **284**(43): 29526–35.

Ferrara LGM., Wallat GD., Moynié L., Dhanasekar NN., Aliouane S., Acosta-Gutiérrez S., Pagès J-M., *et al*. (2016) "MOMP from *Campylobacter jejuni* is a trimer of 18 stranded β–barrel monomers with a Ca^{2+} ion bound at the constriction zone." *Journal of Molecular Biology*. **428** (22): 4528-4543.

Filip C., Fletcher G., Wulff JL., and Earhart CF. (1973) "Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate." *Journal of Bacteriology.* **115**: 6.

Findlay HE., McClafferty H., and Ashley RH. (2005) "Surface expression, single-channel analysis and membrane topology of recombinant *Chlamydia trachomatis* major outer membrane protein." *BMC Microbiology.* **5**: 5.

Fisher DJ., Fernández RE., and Maurelli AT. (2013) "*Chlamydia trachomatis* transports NAD via the Npt1 ATP/ADP translocase." *Journal of Bacteriology*. **195**(15): 3381–3386.

Foadi J., Aller P., Alguel Y., Cameron A., Axford D., Owen RL., Armour W., Waterman DG., Iwata S., and Evans G. (2013) "Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography." *Acta Crystallographica Section D: Biological Crystallography*. **69**(8): 1617–1632.

Ford DC., Ireland PM., Bullifent HL., Saint RJ., McAlister EV., Sarkar-Tyson M., and Oyston PCF. (2014) "Construction of an inducible system for the analysis of essential genes in *Yersinia pestis*." *Journal of Microbiological Methods*. **100**: 1–7.

Freeman TC., Landry SJ., and Wimley WC. (2011) "The prediction and characterization of YshA, an unknown outer membrane protein from *Salmonella typhimurium*." *Biochimica et biophysica acta*. **1808**(1): 287–297.

Fujita Y., Matsuoka H., and Hirooka K. (2007) "Regulation of fatty acid metabolism in bacteria." *Molecular Microbiology.* **66**(4): 829–39.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., and Bairoch A. (2005) *Protein Identification and Analysis Tools on the ExPASy Server;* John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press: 571-607.

Ghuysen JM., and Goffin C. (1999) "Lack of cell wall peptidoglycan versus penicillin sensitivity: new insights into the *Chlamydial* anomaly." *Antimicrobial Agents and Chemotherapy.* **43**(10): 2339–44.

Ginsburgh CL., Black PN., and Nunn WD. (1984) "Transport of long chain fatty acids in *Escherichia coli*. Identification of a membrane protein associated with the *fadL* gene." *The Journal of Biological Chemistry.* **259**(13): 8437–8443.

Godin AG., Lounis B., and Cognet L. (2014) "Super-resolution microscopy approaches for live cell imaging." *Biophysical Journal.* **107** (8): 1777–84.

Grayston JT. (2000) "Background and current knowledge of *Chlamydia pneumoniae* and atherosclerosis." *Journal of Infectious Disease*. **181**(3): S402-10.

Grayston JT., Kuo CC., Campbell LA., and Wang SP. (1989) "*Chlamydia pneumoniae* sp. nov. for *Chlamydia* sp. strain TWAR." *International Journal of Systematic Bacteriology*. **39**: 88-90.

Guan X-G., Su W-H., Yi F., Zhang D., Hao F., Zhang H-G., Liu Y-J., Feng X-C., and Ma, T-H. (2010) "NPA motifs play a key role in plasma membrane targeting of aquaporin-4." *IUBMB life*. **62**(3): 222–226.

Hafner L., Beagley K., and Timms P. (2008) "*Chlamydia trachomatis* infection: Host immune responses and potential vaccines." *Mucosal Immunology.* **1**(2): 116–30.

Hahn DL., and McDonald R. (1998) "Can acute *Chlamydia pneumoniae* respiratory tract infection initiate chronic asthma?" *Annals of Allergy, Asthma & Immunology: Official Publication of the American College of Allergy, Asthma, & Immunology* **81**(4): 339–44.

Hamilton, JA. (2007) "New insights into the roles of proteins and lipids in membrane transport of fatty acids." *Prostaglandins, Leukotrienes, and Essential Fatty Acids.* **77**(5–6): 355–61.

Hatch TP. (1996) "Disulfide cross-linked envelope proteins: The functional equivalent of peptidoglycan in *Chlamydiae*?" *Journal of Bacteriology* **178**(1): 1–5.

Hatch TP., Allan I., and Pearce JH. (1984) "Structural and polypeptide differences between envelopes of infective and reproductive life cycle forms of *Chlamydia* Spp." *Journal of Bacteriology* **157**(1): 13–20.

He W., Felderman M., Evans AC., Geng J., Homan D., Bourguet F., Fischer NO., *et al*. (2017) "Cell-free production of a functional oligomeric form of a *Chlamydia* major outer-membrane protein (MOMP) for vaccine development." *Journal of Biological Chemistry.* **292**(36): 15121–32.

Hearn EM., Patel DR., and van den Berg B. (2008) "Outer-membrane transport of aromatic hydrocarbons as a first step in biodegradation." *Proceedings of the National Academy of Sciences.* **105**(25): 8601–8606.

Hearn EM., Patel DR., Lepore BW., Indic M., and van den Berg B. (2009) "Transmembrane passage of hydrophobic compounds through a protein channel wall." *Nature.* **458**(7236): 367–70.

Hess ST., Girirajan TPK., and Mason MD. (2006) "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy." *Biophysical Journal.* **91** (11): 4258–72.

Hospital A., Goñi JR., Orozco M., and Gelpí JL. (2015) "Molecular dynamics simulations: Advances and applications." *Advances and Applications in Bioinformatics and Chemistry.* **8**: 37–47.

How SJ., Hobson D., and Hart CA. (1984) "Studies in vitro of the nature and synthesis of the cell wall of *Chlamydia trachomatis*." *Current Microbiology*. **10**:269–274.

Huang B., Bates M., and Zhuang X. (2009) "Super resolution fluorescence microscopy." *Annual review of biochemistry.* **78**: 993–1016.

Hybiske K., and Stephens RS. (2007) "Mechanisms of *Chlamydia trachomatis* entry into nonphagocytic cells." *Infection and Immunity.* **75**(8): 3925–34.

Jacob-Dubuisson F., Villeret V., Clantin B., Delattre A-S., and Saint N. (2009) "First structural insights into the TpsB/Omp85 superfamily." *Biological Chemistry.* **390**(8): 675–84.

Janßen HJ., and Steinbüchel A. (2014) "Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid-based biofuels." *Biotechnology for Biofuels.* **7**: 7.

Jones HM., Kubo A., and Stephens RS. (2000) "Design, expression and functional characterization of a synthetic gene encoding the *Chlamydia trachomatis* major outer membrane protein." *Gene*. **258**(1–2): 173–181.

Jovanovic G., Mehta P., McDonald C., Davidson AC., Uzdavinys P., Ying L., and Buck M. (2014) "The N-terminal amphipathic helices determine regulatory and effector functions of Phage shock protein A (PspA) in *Escherichia coli*." *Journal of Molecular Biology.* **426** (7): 1498–1511.

Käding N., Szaszák M., and Rupp J. (2014) "Imaging of *Chlamydia* and host cell metabolism." *Future Microbiology*. **9**(4): 509–521.

Kari L., Whitmire WM., Crane DC., Reveneau N., Carlson JH., Goheen MM., Peterson EM., Pal S., de la Maza LM., and Caldwell HD. (2009) "*Chlamydia trachomatis* native major outer membrane protein induces partial protection in nonhuman primates: Implication for a trachoma transmission-blocking vaccine." *Journal of Immunology* **182**(12): 8063–70.

Kawate T., and Gouaux E. (2006) "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." *Structure.* **14** (4): 673–81.

Kefala G., Ahn C., Krupa M., Esquivies L., Maslennikov I., Kwiatkowski W., and Choe S. (2010) "Structures of the OmpF porin crystallized in the presence of foscholine-12." *Protein Science : A Publication of the Protein Society*. 19(**5**): 1117–1125.

Khoshouei M., Radjainia M., Baumeister W., and Danev R. (2017) "Cryo-EM structure of haemoglobin at 3.2Å determined with the Volta phase plate." *Nature Communications*. **8**: 16099.

Kosma, P. (1999) "*Chlamydial* lipopolysaccharide." *Biochimica et Biophysica Acta - Molecular Basis of Disease*. **1455**(2): 387–402.

Krogh A., Larsson B., von Heijne G., and Sonnhammer EL. (2001) "Predicting transmembrane protein topology with a hidden markov model: application to complete genomes." *Journal of Molecular Biology.* **305**(3): 567–80.

Kubo A., and Stephens RS. (2000) "Characterization and functional analysis of PorB, a *Chlamydia* porin and neutralizing target." *Molecular Microbiology*. **38**(4): 772–780.

Kumar GB., and Black PN. (1993) "Bacterial long-chain fatty acid transport. Identification of amino acid residues within the outer membrane protein FadL required for activity." *The Journal of Biological Chemistry.* **268**(21): 15469–15476.

Ladizhansky V. (2017) "Applications of solid-state NMR to membrane proteins." *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1865**(11b): 1577–1586.

Laemmli, UK. (1970) "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature*. **227**(5259): 680-685.

Lepore BW., Indic M., Pham H., Hearn EM., Patel DR., and van den Berg B. (2011) "Ligand-gated diffusion across the bacterial outer membrane." *Proceedings of the National Academy of Sciences.* **108**(25): 10121–26.

Liechti G., Kuru E., Packiam M., Hsu Y-P., Tekkam S., Hall E., Rittichier JT., van Nieuwenhze M., Brun YV., and Maurelli AT. (2016) "Pathogenic *Chlamydia* lack a classical sacculus but synthesize a narrow, mid-cell peptidoglycan ring, regulated by MreB, for cell division." *PLOS Pathog* **12**(5): e1005590.

Liu X., Afrane M., Clemmer DE., Zhong G., and Nelson DE. (2010) "Identification of *Chlamydia trachomatis* outer membrane complex proteins by differential proteomics." *Journal of Bacteriology* **192**(11): 2852–60.

Los GV., Encell LP., McDougall MG., Hartzell DD., Karassina N., Zimprich C., Wood MG., *et al*. (2008) "HaloTag: A novel protein labeling technology for cell imaging and protein analysis." *ACS Chemical Biology.* **3** (6): 373–82.

Mac T-T., von Hacht A., Hung K-C., Dutton RJ., Boyd D., Bardwell JCA., and Ulmer TS. (2008) "Insight into disulfide bond catalysis in *Chlamydia* from the structure and function of DsbH, a novel oxidoreductase." *The Journal of Biological Chemistry.* **283** (2): 824–32.

Marrink SJ., Risselada HJ., Yefimov S., Tieleman DP., and de Vries AH. (2007) "The MARTINI Force Field: Coarse grained model for biomolecular simulations." *The Journal of Physical Chemistry*. **111** (27): 7812–7824.

Materlik G., Rayment T., and Stuart DI. (2015) "Diamond Light Source: status and perspectives." *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*. 373(**2036**)

Matsumoto A., and GP Manire. (1970) "Electron microscopic observations on the effects of penicillin on the morphology of *Chlamydia psittaci*." *Journal of Bacteriology* **101**(1): 278–85.

McCoy AJ., and Maurelli AT. (2006) "Building the invisible wall: updating the *chlamydial* peptidoglycan anomaly." *Trends in Microbiology*. **14**(2): 70–77.

Miroux B., and Walker JE. (1996) "Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." *Journal of Molecular Biology*. 260(**3**): 289–298.

Mirza SP., Halligan BD., Greene AS., and Olivier M. (2007) "Improved method for the analysis of membrane proteins by mass spectrometry." *Physiological genomics*. **30**(1): 89.

Molina DM., Pal S., Kayala MA., Teng A., Kim PJ., Baldi P., Felgner PL., Liang, X., and de la Maza LM. (2010) "Identification of immunodominant antigens of *Chlamydia trachomatis* using proteome microarrays." *Vaccine*. **28**(17): 3014–3024.

Mölleken K., and Hegemann JH. (2008) "The *Chlamydia* outer membrane protein OmcB is required for adhesion and exhibits biovar-specific differences in glycosaminoglycan binding." *Molecular Microbiology* **67**(2): 403–19.

Mölleken K., Becker E., and Hegemann JH. (2013) "The *Chlamydia pneumoniae* invasin protein Pmp21 recruits the EGF receptor for host cell entry." *PLoS Pathogens* **9**(4).

Molloy MP., Herbert BR., Slade MB., Rabilloud T., Nouwens AS., Williams KL., and Gooley AA. (2000) "Proteomic analysis of the *Escherichia coli* outer membrane." *European journal of biochemistry.* **267**(10): 2871–2881.

Moraes I., Evans G., Sanchez-Weatherby J., Newstead S., and Stewart PDS. (2014) "Membrane protein structure determination — The next generation." *Biochimica et Biophysica Acta*. 1838(**1**): 78–87.

Murata K., and Wolf M. (2018) "Cryo-electron microscopy for structural analysis of dynamic biological macromolecules." *Biochimica et Biophysica Acta (BBA) - General Subjects.* 1862(**2**): 324–334.

Myers-Morales T., Cowan C., Gray ME., Wulff CR., Parker CE., Borchers CH., and Straley SC. (2007) "A surface-focused biotinylation procedure identifies the *Yersinia pestis* catalase KatY as a membrane-associated but non-surface-located protein." *Applied Environmental Microbiology*. **73**(18): 5750–5759.

Newby ZER., O'Connell JD., Gruswitz F., Hays FA., Harries WEC., Harwood IM., Ho JD., Lee JK., Savage DF., Miercke LJW., and Stroud RM. (2009) "A general protocol for the crystallization of membrane proteins for X-ray structural investigation." *Nature Protocols*. 4(**5**): 619–637.

Newhall V., Wilbert J., and Jones RB. (1983) "Disulfide-linked oligomers of the major outer membrane protein of *Chlamydiae*." *Journal of Bacteriology* **154**(2): 998–1001.

Newman J., Fazio VJ., Lawson B., and Peat TS. (2010) "The C6 Web Tool: A resource for the rational selection of crystallization conditions." *Crystal Growth & Design*. 10(**6**): 2785–2792.

Nicholson TL., Olinger L., Chong K., Schoolnik G., and Stephens RS. (2003) "Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*." *Journal of Bacteriology* **185**(10): 3179–89.

Noinaj N., Kuszak AJ., Gumbart JC., Lukacik P., Chang H., Easley NC., Lithgow T., and Buchanan SK. (2013) "Structural insight into the biogenesis of β-barrel membrane proteins." *Nature*. 501(**7467**): 385–390.

Norell D., Heuck A., Tran-Thi T-A., Götzke H., Jacob-Dubuisson F., Clausen T., Daley DO., Braun V., Müller M., and Fan E. (2014) "Versatile *in vitro* system to study translocation and functional integration of bacterial outer membrane proteins." *Nature Communications*. 5: 5396.

Novy R., and Morris B. (2001) "Use of glucose to control basal expression in the pET system." *Innovations BioTechniques*. **12**: 1–3.

Nozaki Y., Schechter NM., Reynolds JA., and Tanford C. (1976) "Use of gel chromatography for the determination of the Stokes radii of proteins in the presence and absence of detergents. A reexamination." *Biochemistry*. 15(**17**): 3884–3890.

Nunes A., Borrego MJ., Nunes B., Florindo C., and Gomes JP. (2009) "Evolutionary dynamics of ompA, the gene encoding the *Chlamydia trachomatis* key antigen." *Journal of Bacteriology* **191**(23): 7182–92.

Nunn WD., and Simons RW. (1978) "Transport of long-chain fatty acids by *Escherichia coli*: mapping and characterization of mutants in the *fadL* gene." *PNaS.* **75** (7): 3377–81.

O'Meara CP., Armitage CW., Andrew DW., Kollipara A., Lycke NY., Potter AA., Gerdts V., Petrovsky N., and Beagley KW. (2017) "Multistage vaccines containing outer membrane, type III secretion system and inclusion membrane proteins protects against a *Chlamydia* genital tract infection and pathology." *Vaccine* **35**(31): 3883–88.

Olsen AW., Follmann F., Erneholm K., Rosenkrands I., and Andersen P. (2015) "Protection against *Chlamydia trachomatis* infection and upper genital tract pathological changes by vaccine-

promoted neutralizing antibodies directed to the VD4 of the major outer membrane protein." *The Journal of Infectious Diseases.* **212**(6): 978–89.

Pageon SV., Nicovich PR., Mollazade M., Tabarin T., and Gaus K. (2016) "Clus-DoC: A combined cluster detection and colocalization analysis for single-molecule localization microscopy data." *Molecular Biology of the Cell.* **27**(22): 3627–36.

Pal S., Peterson EM., and de la Maza LM. (2005) "Vaccination with the *Chlamydia trachomatis* major outer membrane protein can elicit an immune response as protective as that resulting from inoculation with live bacteria." *Infection and Immunity.* **73**(12): 8153–60.

Parker JL., and Newstead S. (2016) "Membrane protein crystallisation: Current trends and future perspectives." *Advances in Experimental Medicine and Biology*. 922: 61–72.

Perez-Melgosa M., Kuo CC., and Campbell LA. (1991) "Sequence analysis of the major outer membrane protein gene of *Chlamydia pneumoniae*." *Infection and Immunology.* **59**: 2195–2199.

Pourhajibagher M., and Bahador A. (2016) "Designing and *in Silico* analysis of PorB protein from *Chlamydia trachomatis* for developing a vaccine candidate." *Drug Research.* **66** (9): 479–83.

Puolakkainen M., Kuo CC., and Campbell LA. (2005) "*Chlamydia pneumoniae* uses the mannose 6-phosphate/insulin-like growth factor 2 receptor for infection of endothelial cells." *Infection and Immunity* **73**(8): 4620–25.

Rahman MU., Hudson AP., and Schumacher HR Jr. (1992) "*Chlamydia* and Reiter's syndrome (reactive arthritis)." *Rheumatic Disease Clinics of North America*. **18**: 67-79.

Redgrove KA., and McLaughlin EA. (2014) "the role of the immune response in *Chlamydia trachomatis* infection of the male genital tract: a double-edged sword." *Frontiers in Immunology.* **5**: 534.

Rodi PM., Bocco Gianello MD., Corregido MC., and Gennaro AM. (2014) "Comparative study of the interaction of CHAPS and Triton X-100 with the erythrocyte membrane." *Biochimica et Biophysica Acta (BBA) – Biomembranes*. 1838(**3**): 859–866.

Rodríguez-Marañón MJ., Bush RB., Peterson EM., Schirmer T., and de la Maza LM. (2002) "Prediction of the membrane-spanning beta-strands of the major outer membrane protein of *Chlamydia*." *Protein Science: A Publication of the Protein Society* **11**(7): 1854–61.

Rosano GL. and Ceccarelli EA. (2014) "Recombinant protein expression in *Escherichia coli*: advances and challenges." *Frontiers in Microbiology*. **5**.

Roulis E., Polkinghorne A., and Timms P. (2013) "*Chlamydia pneumoniae*: Modern insights into an ancient pathogen." *Trends in Microbiology* **21**(3): 120–28.

Rund S., Lindner B., Brade H., and Holst O. (1999) "Structural analysis of the lipopolysaccharide from *Chlamydia trachomatis* serotype L2." *Journal of Biological Chemistry.* **274**(24): 16819–24.

Rust MJ., Bates M., and Zhuang X. (2006) "Stochastic Optical Reconstruction Microscopy (STORM) provides sub-diffraction-limit image resolution." *Nature Methods.* **3** (10): 793–95.

Saikku P., Leinonen M., Mattila K., Ekman MR., Nieminen MS., Mäkelä PH., Huttunen JK., Valtonen V. (1988) "Serological evidence of an association of a novel *Chlamydia,* TWAR, with chronic coronary heart disease and acute myocardial infarction." *Lancet.* **2**:983-6.

Sanderson MJ., Smith I., Parker I., and Bootman MD. (2014) "Fluorescence microscopy." *Cold Spring Harbor Protocols.* **10**

Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., *et al*. (2012), "Fiji: an open-source platform for biological-image analysis". *Nature methods* **9**(7): 676-682.

Schirmer T., Keller TA., Wang YF., and Rosenbusch JP. (1995) "Structural basis for sugar translocation through maltoporin channels at 3.1 a resolution." *Science* **267**(5197): 512–14.

Schneider CA., Rasband WS., and Eliceiri KW. (2012) "NIH Image to ImageJ: 25 years of image analysis". *Nature Methods* **9**: 671-675.

Schoborg RV. (2011) "*Chlamydia* persistence – a tool to dissect *Chlamydia*-host interactions." *Microbes and infection.* **13**(7): 649–662.

Schulz GE. (2002) "The structure of bacterial outer membrane proteins." *Biochimica et Biophysica Acta (BBA) - Biomembranes*. **1565**(2): 308–17.

Seltmann G., and Holst O. (2002) "Periplasmic Space and Rigid Layer. In: The Bacterial Cell Wall." Springer. Berlin, Heidelberg, pp.103–132.

Shahid SA., Bardiaux B., Franks WT., Krabben L., Habeck M., Rossum B-J. van, and Linke D. (2012) "Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals." *Nature Methods*. **9**(12): 1212–1217.

Slotboom DJ., Duurkens RH., Olieman K., and Erkens GB. (2008) "Static light scattering to characterize membrane proteins in detergent solution." *Methods.* 46(**2**): 73–82.

Sonoda Y., Cameron A., Newstead S., Omote H., Moriyama Y., Kasahara M., Iwata S., and Drew D. (2010) "Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins." *FEBS Letters*. 584(**12**): 2539–2547.

Sorrentino R., Yilmaz A., Schubert K., Crother TR., Pinto A., Shimada K., Arditi M., and Chen S. (2015) "A single infection with *Chlamydia pneumoniae* is sufficient to exacerbate atherosclerosis in ApoE deficient mice." *Cellular immunology*. **294**(1): 25–32.

Steimle A., Autenrieth IB., and Frick J-S. (2016) "Structure and function: Lipid A modifications in commensals and pathogens." *International Journal of Medical Microbiology*. **306**(5): 290–301.

Stephens RS., and Lammel CJ. (2001) "*Chlamydia* outer membrane protein discovery using genomics." *Current Opinions in Microbiology*. **4**(1): 16-20.

Stephens RS., Kalman S., Lammel C., Fan J., Marathe R., Aravind L., Mitchell W., *et al.* (1998) "Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*." *Science.* **282**(5389): 754–59.

Su H., Watkins NG., Zhang YX., and Caldwell HD. (1990) "Chlamydia trachomatis-host cell interactions: Role of the *Chlamydial* major outer membrane protein as an adhesin." *Infection and Immunity* **58**(4): 1017–25.

Sun G., Pal S., Sarcon AK., Kim S., Sugawara E., Nikaido H., Cocco MJ., Peterson EM. and de la Maza LM. (2007) "Structural and functional analyses of the Major Outer Membrane Protein of *Chlamydia trachomatis*." *Journal of Bacteriology*. **189**(17): 6222–6235.

Thomas JA., and Tate CG. (2014) "Quality control in eukaryotic membrane protein overproduction." *Journal of Molecular Biology*. 426(**24**): 4139–4154.

Thorley JA., Pike J., and Rappoport JZ. (2014) *Fluorescence Microscopy: Super-resolution and other novel techniques*. Boston: Academic Press.

Tifrea DF., Sun G., Pal S., Zardeneta G., Cocco MJ., Popot JL., and de la Maza LM. (2011) "Amphipols stabilize the *Chlamydia* major outer membrane protein and enhance its protective ability as a vaccine." *Vaccine* **29**(28): 4623–31.

Van den Berg B., Black PN., Clemons WM., and Rapoport TA. (2004) "Crystal structure of the long-chain fatty acid transporter FadL." *Science* **304**(5676): 1506–9.

van den Berg, B. (2010) "Going forward laterally: Transmembrane passage of hydrophobic molecules through protein channel walls." *Chembiochem: A European Journal of Chemical Biology.* **11** (10): 1339–43.

Verardi R., Shi L., Traaseth NJ., Walsh N., and Veglia G. (2011) "Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method." *Proceedings of the National Academy of Sciences.* **108**(22): 9101–9106.

Voss BJ., and Cover TL. (2015) "Biotinylation and purification of surface-exposed *Helicobacter pylori* proteins." *Bio-protocol*. **5**(8): e1455.

Wallace B., and Janes R. (2010) "Synchrotron radiation circular dichroism (SRCD) spectroscopy: an enhanced method for examining protein conformations and protein interactions." *Biochemical Society Transactions.* **238**(4): 861-73.

Wang S., Munro RA., Shi L., Kawamura I., Okitsu T., Wada A., Kim S-Y., Jung K-H., Brown LS., and Ladizhansky V. (2013) "Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein." *Nature Methods*. **10**(10): 1007–1012.

Wang SP., and Grayston JT. (1990) "Population prevalence antibody to *Chlamydia pneumoniae*, strain TWAR. In: Bowie WR., Caldwell HD., Jones RP., Mardh PA., Ridgway GL., Schachter J., *et al*, eds. *Chlamydial Infections*. Cambridge: Cambridge University Press :402-5.

Wang Y., Berg EA., Feng X., Shen L., Smith T., Costello CE., and Zhang Y-X. (2006) "Identification of surface-exposed components of MOMP of *Chlamydia trachomatis* serovar F." *Protein Science* **15**(1): 122–34.

Wassenaar TA., Ingólfsson HI., Böckmann RA., Tieleman DP., and Marrink SJ. (2015) "Computational lipidomics with Insane: A versatile tool for generating custom membranes for molecular simulations." *Journal of Chemical Theory and Computation*. **11**(5): 2144–2155.

Webley WC., and Hahn DL. (2017) "Infection-mediated asthma: Etiology, mechanisms and treatment options, with focus on *Chlamydia pneumoniae* and macrolides." *Respiratory Research* **18.**

Wegerer A., Sun T., and Altenbuchner J. (2008) "Optimization of an *E. coli* L-rhamnose-inducible expression vector: Test of various genetic module combinations." *BMC Biotechnology.* **8**: 2.

Wen Z., Boddicker MA., Kaufhold RB., Khandelwal P., Durr E., Qiu P., Lucas BJ., *et al*. (2016) "Recombinant expression of *Chlamydia trachomatis* major outer membrane protein in e. coli outer membrane as a substrate for vaccine research." *BMC Microbiology* **16**.

Willson BJ., Kovács K., Wilding-Steele T., Markus R., Winzer K., and Minton, NP. (2016) "Production of a functional cell wall-anchored minicellulosome by recombinant *Clostridium acetobutylicum* ATCC 824." *Biotechnology for Biofuels* **9**: 109.

Wlodawer A., Minor W., Dauter Z., and Jaskolski M. (2013) "Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination." *The FEBS journal*. 280(**22**): 5705–5736.

Wolf K., Fischer E., Mead D., Zhong G., Peeling R., Whitmire B., and Caldwell HD. (2001) "*Chlamydia pneumoniae* major outer membrane protein is a surface-exposed antigen that elicits antibodies primarily directed against conformation-dependent determinants." *Infection and Immunity.* **69**(5): 3082–3091.

Wyllie S., Ashley RH., Longbottom D., and Herring AJ. (1998) "The major outer membrane protein of *Chlamydia psittaci* functions as a porin-like ion channel." *Infection and Immunity* **66**(11): 5202–7.

Wyrick PB. (2010) "*Chlamydia trachomatis* persistence *in Vitro* – An overview." *The Journal of Infectious Diseases.* **201**(2): S88–95.

Xia M., Chen D., Endresz V., Faludi I., Szabo A., Gonczol E., Kakkar V., and Lu X. (2013) "Immunization of *Chlamydia pneumoniae* (Cpn)-infected ApoB^{tm2Sgy}Ldlr^{tm1Her}/J mice with a combined peptide of Cpn significantly reduces atherosclerotic lesion development." *PLoS ONE* **8**(12)

Yao J., Cherian PT., Frank MW., and Rock CO. (2015) "*Chlamydia trachomatis* relies on autonomous phospholipid synthesis for membrane biogenesis." *The Journal of Biological Chemistry*. **290**: 18874–18888.

Yao J., Dodson VJ., Frank MW., and Rock CO. (2015) "*Chlamydia trachomatis* scavenges host fatty acids for phospholipid synthesis via an acyl-acyl carrier protein synthetase." *Journal of Biological Chemistry.* **290**(36): 22163–22173.

Yao Z., and Carballido-López R. (2014) "Fluorescence imaging for bacterial cell biology: From localization to dynamics, from ensembles to single molecules." *Annual Review of Microbiology* **68**(1): 459–76.

Yen T-Y., Pal S., and de la Maza LM. (2005) "Characterisation of the disulfide bonds and free cysteine residues of the *Chlamydia trachomatis* mouse pneumonitis major outer membrane protein." *Biochemistry*. **44**(16): 6250–6256.

Zabara A., Chong JTY., Martiel I., Stark L., Cromer BA., Speziale C., Drummond CJ., and Mezzenga R. (2018) "Design of ultra-swollen lipidic mesophases for the crystallization of membrane proteins with large extracellular domains." *Nature Communications*. 9(**1**): 544.

Zeidler H., and Hudson AP. (2016) "Causality of *Chlamydiae* in arthritis and spondyloarthritis: a plea for increased translational research." *Current Rheumatology Reports*. **18**(2): 9.

Appendices

Appendix A1

```
sp|P23732|MOMP_CT_A          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|P08780|MOMP_CT_C          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|P13467|MOMP_CT_H          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCATWCDAIS    59
sp|P23114|MOMP_CT_L3         MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|P16155|MOMP_CT_F          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|P23421|MOMP_CT_B          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWVDAIS    59
sp|P06597|MOMP_CT_L2         MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|Q46409|MOMP_CT_D          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCATWCDAIS    59
sp|P19542|MOMP_CT_L1         MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|P17451|MOMP_CT_E          MKKLLKSVLVFAA-LSSASSLQALPVGNPAEPSLMIDGILWEGFGGDPCDPCTTWCDAIS    59
sp|Q07430|MOMP_CP_IOL-207    MKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATWCDAIS    60
sp|P27455|MOMP_CP_AR39       MKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATWCDAIS    60
sp|Q9XBF4|MOMP_CP_TWAR       MKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATWCDAIS    60
sp|Q00087|MOMP_CPS_F         MKKLLKSALLFAA-AGSALSLQALPVGNPAEPSLLIDGTMWEGASGDPCDPCATWCDAIS    59
sp|Q46203|MOMP_CPS_VR-125    MKKLLKSALLFAA-TGSALSLQALPVGNPAEPSLLIDGTMWEGASGDPCDPCATWCDAIS    59
sp|P10332|MOMP_CPS_A22/M     MKKLLKSALLFAA-TGSALSLQALPVGNPAEPSLLIDGTMWEGASGDPCDPCATWCDAIS    59
                            *******.*: **   .*. **********:.***:*** :*** .********:** ****


sp|P23732|MOMP_CT_A          MRMGYYGDFVFDRVLKTDVNKEFQMGAA-PTTRDVAGLEKDPVVNVARPNPAYGKHMQDA   118
sp|P08780|MOMP_CT_C          MRVGYYGDFVFDRVLKTDVNKEFQMGAA-PTTSDVAGLQNDPTINVARPNPAYGKHMQDA   118
sp|P13467|MOMP_CT_H          MRVGYYGDFVFDRVLKTDVNKEFQMGAA-PTTNDAADLQNDPKTNVARPNPAYGKHMQDA   118
sp|P23114|MOMP_CT_L3         MRVGYYGDFVFDRVLKTDVNKEFQMGAE-PTTSDTAGLSNDPTTNVARPNPAYGKHMQDA   118
sp|P16155|MOMP_CT_F          MRMGYYGDFVFDRVLKTDVNKEFEMGEA-LAGASGN-TTSTLSKLVERTNPAYGKHMQDA   117
sp|P23421|MOMP_CT_B          MRMGYYGDFVFDRVLKTDVNKEFQMGAK-PTTTTGN-AV-APSTLTARENPAYGRHMQDA   116
sp|P06597|MOMP_CT_L2         MRMGYYGDFVFDRVLQTDVNKEFQMGAK-PTTATGN-AA-APSTCTARENPAYGRHMQDA   116
sp|Q46409|MOMP_CT_D          MRVGYYGDFVFDRVLKTDVNKEFQMGAK-PTTDTGN-SA-APSTLTARENPAYGRHMQDA   116
sp|P19542|MOMP_CT_L1         MRMGYYGDFVFDRVLQTDVNKEFQMGAK-PTATTGN-AA-APSTCTARENPAYGRHMQDA   116
sp|P17451|MOMP_CT_E          MRMGYYGDFVFDRVLKTDVNKEFQMGDK-PTSTTGN-AT-APTTLTARENPAYGRHMQDA   116
sp|Q07430|MOMP_CP_IOL-207    LRAGFYGDYVFDRILKIDAPKTFSMGAK-PTGS----ATANYTTAVDRPNPAYNKHLYDA   115
sp|P27455|MOMP_CP_AR39       LRAGFYGDYVFDRILKVDAPKTFSMGAK-PTGS----AAANYTTAVDRPNPAYNKHLHDA   115
sp|Q9XBF4|MOMP_CP_TWAR       LRAGFYGDYVFDRILKVDAPKTFSMGAK-PTGS----ATANYTTAVDRPNPAYNKHLHDA   115
sp|Q00087|MOMP_CPS_F         IRAGFYGDYVFDRILKVDVNKTISGMAAAPTAASGTASNTT--VAADRSNFAYGKHLQDA   117
sp|Q46203|MOMP_CPS_VR-125    IRAGYYGDYVFDRVLKVDVNKTFSGMAATPTQATGNASNTNQPEANGRPNIAYGRHMQDA   119
sp|P10332|MOMP_CPS_A22/M     IRAGYYGDYVFDRVLKVDVNKTFSGMAATPTQATGNASNTNQPEANGRPNIAYGRHMQDA   119
                            :* *:***:****:*: *. * :.       :                * * **.:*: **


sp|P23732|MOMP_CT_A          EMFTNAAYMALNIWDRFDVFCTLGATTGYLKGNSASFNLVGLFGTKTQSSG-FDTANIVP   177
sp|P08780|MOMP_CT_C          EMFTNAAYMALNIWDRFDVFCTLGATTGYLKGNSASFNLVGLFGTKTQSSS-FNTAKLIP   177
sp|P13467|MOMP_CT_H          EMFTNAAYMALNIWDRFDVFCTLGATTGYLKGNSASFNLVGLFGTKTKSSD-FNTAKLVP   177
sp|P23114|MOMP_CT_L3         EMFTNAAYMALNIWDRFDVFCTLGATTGYLKGNSASFNLVGLFGTKTQSTN-FNTAKLVP   177
sp|P16155|MOMP_CT_F          EMFTNAACMTLNIWDRFDVFCTLGATSGYLKGNSASFNLVGLFGDGVNATK-PA-ADSIP   175
sp|P23421|MOMP_CT_B          EMFTNAACMALNIWDRFDVFCTLGASSGYLKGNSASFNLVGLFGNNENQTK-VSNGAFVP   175
sp|P06597|MOMP_CT_L2         EMFTNAAYMALNIWDRFDVFCTLGATSGYLKGNSASFNLVGLFGDNENHAT-VSDSKLVP   175
sp|Q46409|MOMP_CT_D          EMFTNAACMALNIWDRFDVFCTLGATSGYLKGNSASFNLVGLFGDNENQKT-VK-AESVP   174
sp|P19542|MOMP_CT_L1         EMFTNAAYMALNIWDRFDVFCTLGATSGYLKGNSASFNLVGLFGDNENQST-VK-KDAVP   174
```

```
sp|P17451|MOMP_CT_E          EMFTNAACMALNIWDRFDVFCTLGASSGYLKGNSASFNLVGLFGDNENQST-VK-TNSVP          174
sp|Q07430|MOMP_CP_IOL-207    EWFTNAGFIALNIWDRFDVFCTLGASNGYVKGNSAAFNLVGLFGVKGTSVN----ANELP          171
sp|P27455|MOMP_CP_AR39       EWFTNAGFIALNIWDRFDVFCTLGASNGYIRGNSTAFNLVGLFGVKGTTVN----ANELP          171
sp|Q9XBF4|MOMP_CP_TWAR       EWFTNAGFIALNIWDRFDVFCTLGASNGYIKGNSTAFNLVGLFGVKGTSVA----ANELP          171
sp|Q00087|MOMP_CPS_F         EWCTNAAYLALNIWDRFDVFCTLGASNGYFKASSDAFNLVGLIGLAG-----TDFANQRP          172
sp|Q46203|MOMP_CPS_VR-125    EWFSNAAFLALNIWDRFDIFCTLGASNGYFKASSAAFNLVGLIGFSAASSISTDLPMQLP          179
sp|P10332|MOMP_CPS_A22/M     EWFSNAAFLALNIWDRFDIFCTLGASNGYFKSSSAAFNLVGLIGFSATSSTSTELPMQLP          179
                             *  :**. ::*********:******:.**.:..* :*******:*            *


sp|P23732|MOMP_CT_A          NTALNQAVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNASEF          237
sp|P08780|MOMP_CT_C          NTALNEAVVELYINTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNASEF          237
sp|P13467|MOMP_CT_H          NIALNRAVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNASEF          237
sp|P23114|MOMP_CT_L3         NTALNQAVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCDASEF          237
sp|P16155|MOMP_CT_F          NVQLNQSVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKIEELNVLCNAAEF          235
sp|P23421|MOMP_CT_B          NMSLDQSVVELYTDTAFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNAAEF          235
sp|P06597|MOMP_CT_L2         NMSLDQSVVELYTDTTFAWSAGARAALWECGCATLGASFQYAQSKPKVEELNVLCNAAEF          235
sp|Q46409|MOMP_CT_D          NMSFDQSVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNAAEF          234
sp|P19542|MOMP_CT_L1         NMSFDQSVVELYTDTTFAWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNAAEF          234
sp|P17451|MOMP_CT_E          NMSLDQSVVELYTDTAFSWSVGARAALWECGCATLGASFQYAQSKPKVEELNVLCNAAEF          234
sp|Q07430|MOMP_CP_IOL-207    NVSLSNGVIELYTDTTFAWSVGARGALWECGCATLGAEFQYAQSKPKVEELNVICNVSQF          231
sp|P27455|MOMP_CP_AR39       NVSLSNGVVELYTDTSFSWSVGARGALWECGCATLGAEFQYAQSKPKVEELNVICNVSQF          231
sp|Q9XBF4|MOMP_CP_TWAR       NVSLSNGVVELYTDTSFSWSVGARGALWECGCATLGAEFQYAQSKPKVEELNVICNVAQF          231
sp|Q00087|MOMP_CPS_F         NVEISQGIVELYTDTAFSWSVGARGALWECGCATLGAEFQYAQSNPKIEMLNVTSSPAQF          232
sp|Q46203|MOMP_CPS_VR-125    NVGITQGVVEFYTDTSFSWSVGARGALWECGCATLGAEFQYAQSNPKIEMLNVTSSPAQF          239
sp|P10332|MOMP_CPS_A22/M     NVGITQGVVEFYTDTSFSWSVGARGALWECGCATLGAEFQYAQSNPKIEVLNVTSSPAQF          239
                             *  :  ..::*:* :*:*:**.***.************.******:**:* *** .. ::*


sp|P23732|MOMP_CT_A          TINKPKGYVGA--EFPLDITAGTEAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          295
sp|P08780|MOMP_CT_C          TINKPKGYVGA--EFPLNITAGTEAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          295
sp|P13467|MOMP_CT_H          TINKPKGYVGA--EFPLDITAGTEAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          295
sp|P23114|MOMP_CT_L3         TINKPKGYVGA--EFPLDITAGTEAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          295
sp|P16155|MOMP_CT_F          TINKPKGYVGK--EFPLDLTAGTDAATGTKDASIDYHEWQASLSLSYRLNMFTPYIGVKW          293
sp|P23421|MOMP_CT_B          TINKPKGYVGK--ELPLDLTAGTDAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          293
sp|P06597|MOMP_CT_L2         TINKPKGYVGQ--EFPLDLKAGTDGVTGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          293
sp|Q46409|MOMP_CT_D          TINKPKGYVGK--EFPLDLTAGTDAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          292
sp|P19542|MOMP_CT_L1         TINKPKGYVGK--EFPLDLTAGTDAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          292
sp|P17451|MOMP_CT_E          TINKPKGYVGQ--EFPLALIAGTDAATGTKDASIDYHEWQASLALSYRLNMFTPYIGVKW          292
sp|Q07430|MOMP_CP_IOL-207    SLNKPKGYKGV--AFPLPTDAGVVTAAGTKSATINYHEWQVGASLSYRLNSLVPYIGVQW          289
sp|P27455|MOMP_CP_AR39       SVNKPKGYKGV--AFPLPTDAGVATATGTKSATINYHEWQVGASLSYRLNSLVPYIGVQW          289
sp|Q9XBF4|MOMP_CP_TWAR       SVNKPKGYKGV--AFPLPTDAGVATATGTKSATINYHEWQVGASLSYRLNSLVPYIGVQW          289
sp|Q00087|MOMP_CPS_F         MIHKPRGYKGTAANFPLPVAAGTATATDTKSATVKYHEWQVGLALSYRLNMLVPYIGVNW          292
sp|Q46203|MOMP_CPS_VR-125    VIHKPRGYKGASSNFPLPITAGTTEATDTKSATIKYHEWQVGLALSYRLNMLVPYIGVNW          299
sp|P10332|MOMP_CPS_A22/M     VIHKPRGYKGASSNFPLPITAGTTEATDTKSATIKYHEWQVGLALSYRLNMLVPYIGVNW          299
                             ::**:** *    :**   **. .:.**.*::.*****.. :****** :.*****:*


sp|P23732|MOMP_CT_A          SRVSFDADTIRIAQPKLAKPVLDTTTLNPTIAGKGTVVS-S-AENELADTMQIVSLQLNK          353
sp|P08780|MOMP_CT_C          SRVSFDADTIRIAQPKLAEAILDVTTLNRTTAGKGSVVS-AGTDNELADTMQIVSLQLNK          354
sp|P13467|MOMP_CT_H          SRVSFDADTIRIAQPKLAEAILDVTTLNPTIAGKGTVVA-SGSDNDLADTMQIVSLQLNK          354
sp|P23114|MOMP_CT_L3         SRVSFDADTIRIAQPKLAEAVLDVTTLNPTIAGKGSVVA-SGSENELADTMQIVSLQLNK          354
```

271

```
sp|P16155|MOMP_CT_F         SRASFDSDTIRIAQPRLVTPVVDITTLNPTIAGCGSVAG-ANTEGQISDTMQIVSLQLNK    352
sp|P23421|MOMP_CT_B         SRASFDADTIRIAQPKSAETIFDVTTLNPTIAGAGDVKT--SAEGQLGDTMQIVSLQLNK    351
sp|P06597|MOMP_CT_L2        SRASFDADTIRIAQPKSATTVFDVTTLNPTIAGAGDVKA--SAEGQLGDTMQIVSLQLNK    351
sp|Q46409|MOMP_CT_D         SRASFDADTIRIAQPKSATAIFDTTTLNPTIAGAGDVKT--GAEGQLGDTMQIVSLQLNK    350
sp|P19542|MOMP_CT_L1        SRASFDADTIRIAQPKLATAIFDTTTLNPTIAGAGEVKA--NAEGQLGDTMQIVSLQLNK    350
sp|P17451|MOMP_CT_E         SRASFDADTIRIAQPKSATAIFDTTTLNPTIAGAGDVKA--SAEGQLGDTMQIVSLQLNK    350
sp|Q07430|MOMP_CP_IOL-207   SRATFDADNIRIAQPKLPTAILNLTAWNPSLLGSATAVSSS---DQFSDFMQIVSCQINK   346
sp|P27455|MOMP_CP_AR39      SRATFDADNIRIAQPKLPTAVLNLTAWNPSLLGNATALSTT---DSFSDFMQIVSCQINK   346
sp|Q9XBF4|MOMP_CP_TWAR      SRATFDADNIRIAQPKLPTAVLNLTAWNPSLLGNTTTLATS---DSFSDFMQIVSCQINK   346
sp|Q00087|MOMP_CPS_F        SRATFDADTIRIAQPKLASAILNLTTWNPTLLGVATTLDTS---NKYADFMQIVSMQINK   349
sp|Q46203|MOMP_CPS_VR-125   SRATFDADTIRIAQPKLKSEILNITTWNPSLIGSTTALPNNSGKDVLSDVLQIASIQINK   359
sp|P10332|MOMP_CPS_A22/M    SRATFDADTIRIAQPKLKSEILNITTWNPSLLGSTTTLPNNGGKDVLSDVLQIASIQINK   359
                           **.:**:*.******:     :.: *: * :   *    .      .  .* :**.* *:**


sp|P23732|MOMP_CT_A         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   396
sp|P08780|MOMP_CT_C         MKSRKSCGIAVGTTIVDADKYAVTVEARLIDERAAHVNAQFRF   397
sp|P13467|MOMP_CT_H         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   397
sp|P23114|MOMP_CT_L3        MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   397
sp|P16155|MOMP_CT_F         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   395
sp|P23421|MOMP_CT_B         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   394
sp|P06597|MOMP_CT_L2        MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   394
sp|Q46409|MOMP_CT_D         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   393
sp|P19542|MOMP_CT_L1        MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   393
sp|P17451|MOMP_CT_E         MKSRKSCGIAVGTTIVDADKYAVTVETRLIDERAAHVNAQFRF   393
sp|Q07430|MOMP_CP_IOL-207   FKSRKACGVTVGATLVDADKWSLTAEARLINERAAHISGQFRF   389
sp|P27455|MOMP_CP_AR39      FKSRKACGVTVGATLVDADKWSLTAEARLINERAAHVSGQFRF   389
sp|Q9XBF4|MOMP_CP_TWAR      FKSRKACGVTVGATLVDADKWSLTAEARLINERAAHVSGQFRF   389
sp|Q00087|MOMP_CPS_F        MKSRKACGIAVGATLIDADKWSITGEARLIDERAAHINAQFRF   392
sp|Q46203|MOMP_CPS_VR-125   MKSRKACGVAVGATLIDADKWSITGEARLINERAAHMNAQFRF   402
sp|P10332|MOMP_CPS_A22/M    MKSRKACGVAVGATLIDADKWSITGEARLINERAAHMNAQFRF   402
                           :****:**::**:*::****:::* *:***:*****:..****
```

Figure A1.1. Multiple sequence alignment of MOMP from *C. pneumoniae, C. trachomatis and C. psittaci* human pathogenic strains. Conducted in Clustal Omega. CP, *C. pneumoniae*; CT, *C. trachomatis*; CPS, *C. psittaci*.

```
sp|P27455|MOMP_CP_AR39    --MKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATWCDA    58
sp|P10384|FADL_EC_K12     MSQKTLFTKSALAV------------------AVALISTQAWSAGFQLNE-----F-SS    35
                            *.*:..: *:.                     **.   *...         : .:


sp|P27455|MOMP_CP_AR39    ISLRAGFYGDYVF--DRILKVDAPKTFSMGAKPTGSAAANYTTA---VDR----PNPAYN   109
sp|P10384|FADL_EC_K12     SGLGRAYSGEGAIADDAGNVSRNPALITMFDRPTFSAGAVYIDPDVNISGTSPSGRSLKA    95
                           .*   .: *: .: *        * ::*  :** **.* *       :.        .


sp|P27455|MOMP_CP_AR39    KHLHDAEWFTNAGFIALNIWDRFDVFCTLGA----------------SNGYIRGNSTAFN   153
sp|P10384|FADL_EC_K12     DNIAPTAWVPNMHFVAP-INDQFGWGASITSNYGLATEFNDTYAGGSVGGTTDLETMNLN   154
                          .::  : *. *  *:*  * *:*.  .:: :                 .*    ::  :*


sp|P27455|MOMP_CP_AR39    LVGLFGVK----------------------GTTVNANEL--PN---------VSLSN    177
sp|P10384|FADL_EC_K12     LSGAYRLNNAWSFGLGFNAVYARAKIERFAGDLGQLVAGQIMQSPAGQTQQGQALAATAN   214
                          * * : ::                      * * .: : *          .: :*


sp|P27455|MOMP_CP_AR39    GVV---ELYTDTSFSWSVGARGALWECGCA-TLGAEFQYAQSKPKVEELNVICNVSQFSV   233
sp|P10384|FADL_EC_K12     GIDSNTKIAHLNGNQWGFG-----WNAGILYELDKNNRYALTYRSEV---KIDFKGNYSS   266
                          *:     ::    .. .*..*      *:.*    *. : :** : .      *   .::*


sp|P27455|MOMP_CP_AR39    NKPKGYKGVAFPLPTDAGVATATGTKSATINYHEWQVGASLSYRLNSLVPYIGVQWSRAT   293
sp|P10384|FADL_EC_K12     DLNRAFNNYGLPIPTATGGATQSGYLTLNLP-EMWEVSGYN-----R----VDPQWAI--   314
                          :  :.::. .:*:** :* ** :*   : .:  . *:*..         :. **:


sp|P27455|MOMP_CP_AR39    FDADNIRIAQPKLPTAVLNLTAWNPSLLGNATALSTTDSFSDFMQIVSCQINKFKSRKAC   353
sp|P10384|FADL_EC_K12     --------------HYSLAYTSWSQFQQLKATSTSGDTLFQK--------HE--GFKDAY   350
                                        *   *:*.    :**: *    *..        :    :.*


sp|P27455|MOMP_CP_AR39    GVTVGATLVDADKWSLTAEARLIN------ERAAHVSGQFRF------------------   389
sp|P10384|FADL_EC_K12     RIALGTTYYYDDNWTFRTGIAFDDSPVPAQNRSISIPDQDRFWLSAGTTYAFNKDASVDV   410
                           :::*:*    *:*:: :   : :      :*:  : .* **


sp|P27455|MOMP_CP_AR39    -----------------------------------    389
sp|P10384|FADL_EC_K12     GVSYMHGQSVKINEGPYQFESEGKAWLFGTNFNYAF    446
```

Figure A1.2. Sequence alignment of *C. pneumoniae* AR39 MOMP with *E. coli* K12 FadL. Conducted in Clustal Omega. CP, *C. pneumoniae*; EC, *E. coli*.
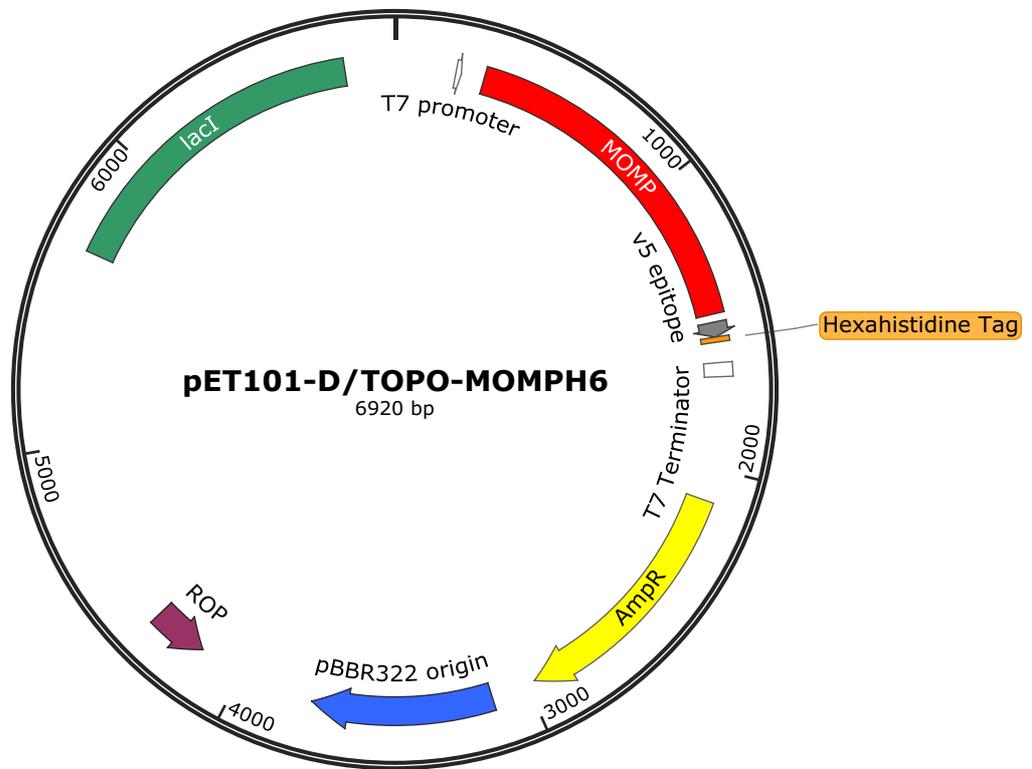
Figure A2.1. Plasmid map of pET101-D/TOPO MOMP$_{H6}$. Created in SnapGene Viewer.
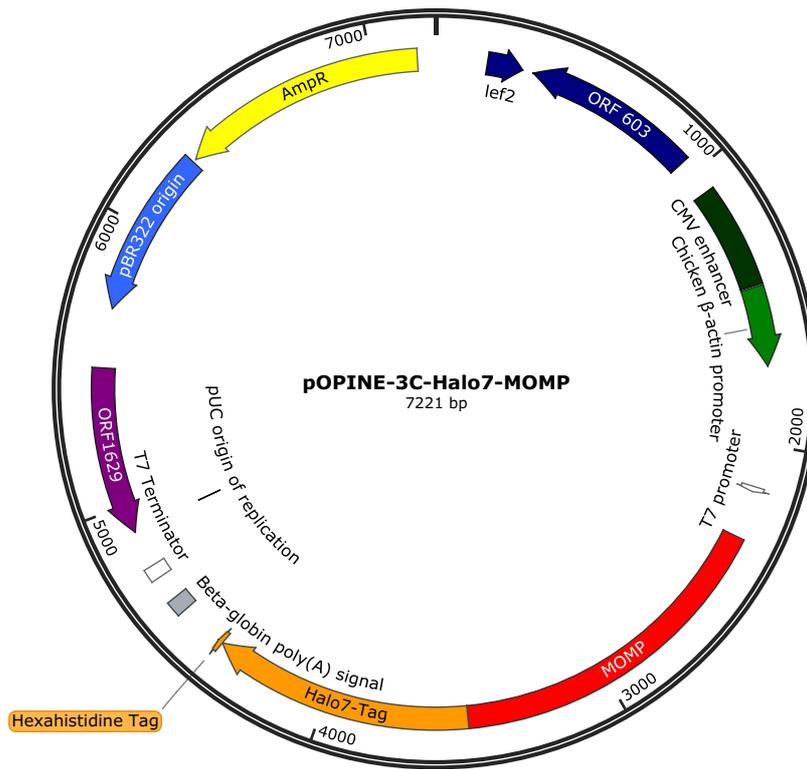
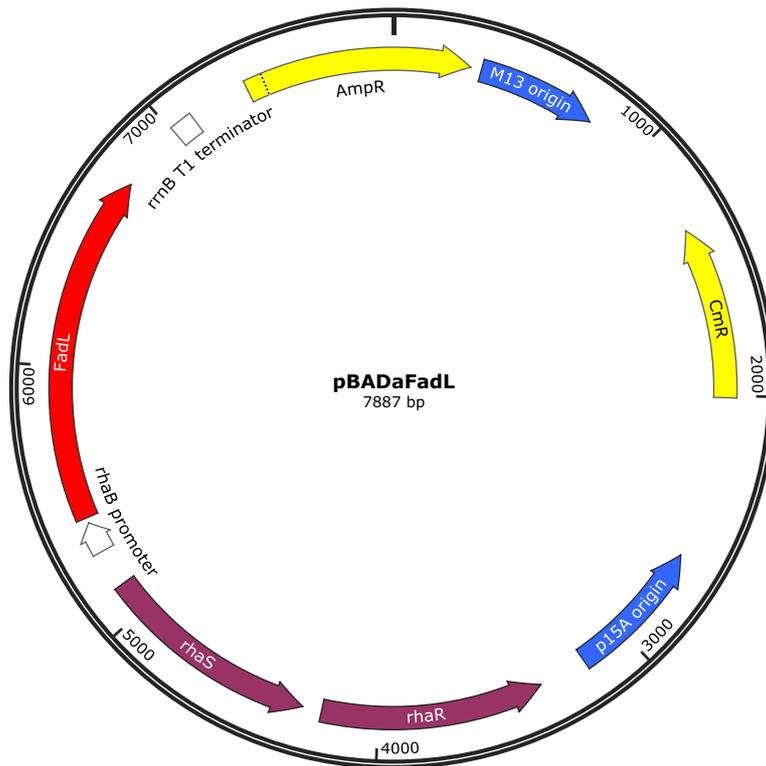**Figure A2.2. Plasmid map of pOPINE-3C-Halo7 MOMP.** Created in SnapGene Viewer.

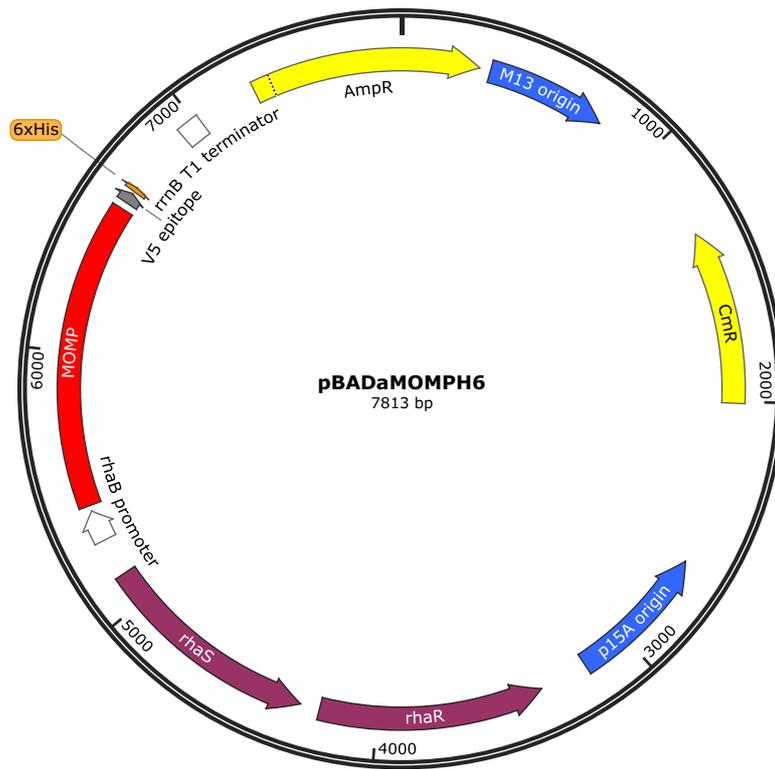**Figure A2.3. Plasmid map of pBADaFadL.** Created in SnapGene Viewer.

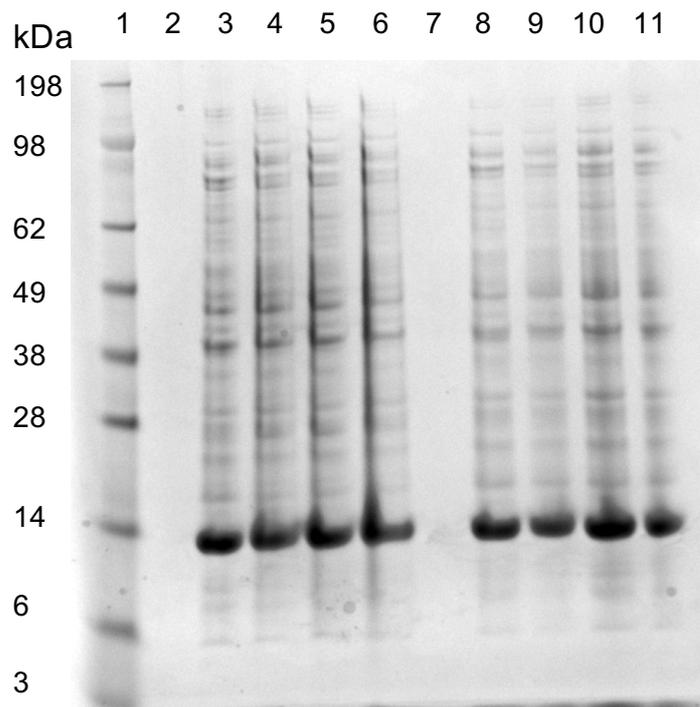**Figure A2.4. Plasmid map of pBADaMOMP_{H6}.** Created in SnapGene Viewer.

Figure A3.1. **Coomassie blue stained SDS-PAGE gel as a loading control for Western blotting detected expression of his-tag mutants.** 1, His-tagged protein standard; 2, Blank; 3, C41 (DE3) control; 4, C41 (DE3) His-6; 5, C41 (DE3) His-8; 6, C41 (DE3) His-10; 7, Blank; 8, BL21 (DE3) control; 9, BL21 (DE3) His-6; 10, BL21 (DE3) His-8; 11, BL21 (DE3) His-10.
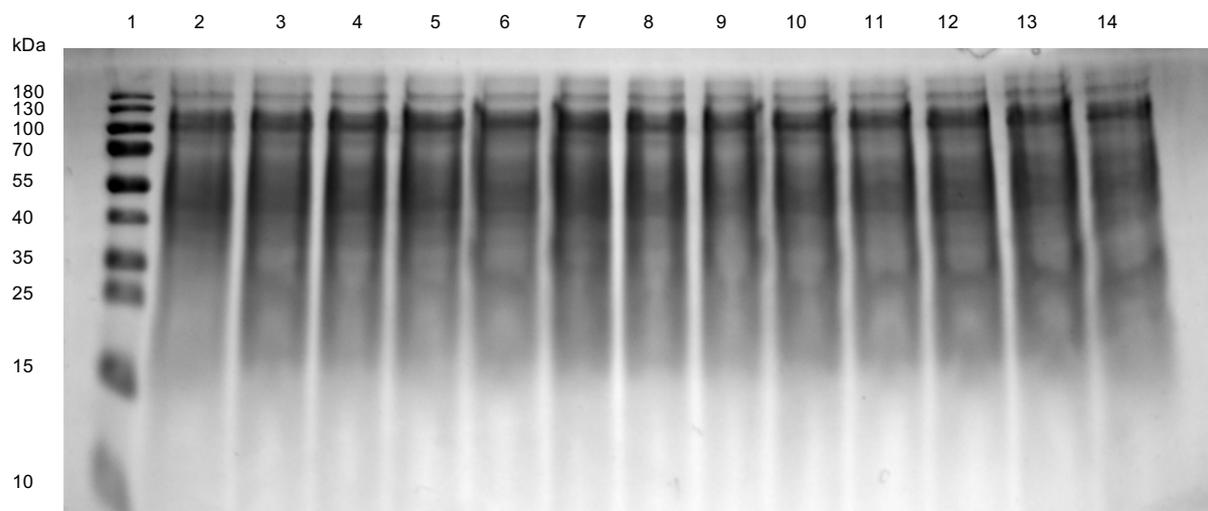
Figure A4.1. **Coomassie blue stained SDS-PAGE gel as a loading control for Western blotting detected expression of whole LS6164 knockout cells expressing pBADaMOMP$_{H6}$.** 1, MW marker; 2, LS6164 control cells; 3, 0% 1h; 4, 0% 3h; 5, 0% 5h; 6, 0.005% 1h; 7, 0.005% 3h; 8, 0.005% 5h; 9, 0.01% 1h; 10, 0.01% 3h; 11, 0.01% 5h; 12, 0.02% 1h; 13, 0.02% 3h; 14, 0.02% 5h. All percentages refer to rhamnose concentration (w/v) and hours to time of harvesting.

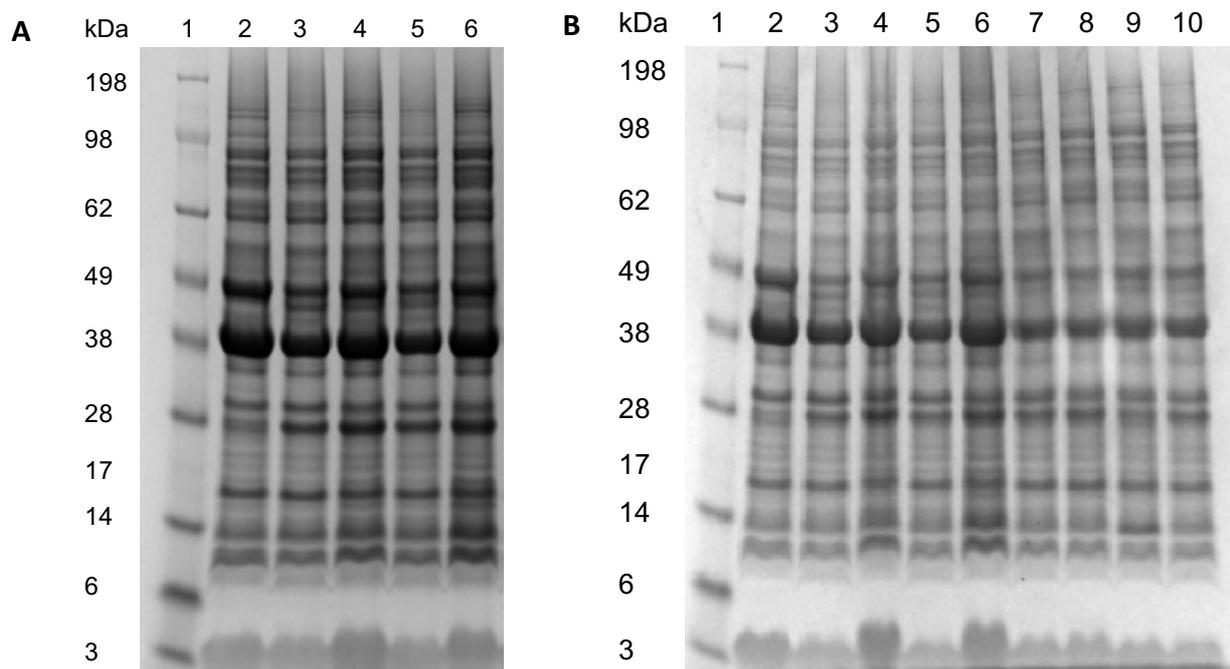**Figure A5.1**. **Coomassie blue stained SDS-PAGE gels as loading controls for Western blotting detected expression of cysteine mutants. A,** 1, MW ladder; 2, C41 (DE3) control; 3, native MOMP; 4, MOMP C201A; 5, MOMP C203A; and 6, MOMP C201/3A. **B**, 1, MW ladder; 2, C41 (DE3) control; 3, native MOMP; 4, MOMP C201A; 5, MOMP C203A; 6, MOMP C201/3A; 7 MOMP C136A; 8, MOMP C136/201A; 9, MOMP C136/203A; and 10, MOMP C226A. A, **Figure 5.12** control; B, **Figure 5.15** control.