# Sheffield Hallam University

## A Sheffield Hallam University thesis

ProQuest Number: 10701125

ProQuest 10701125

# Heterologous expression of *Chlamydia trachomatis* Polymorphic Membrane Proteins for *in vitro* studies

Louise Vickers

A thesis submitted in partial fulfilment of the requirements of Sheffield Hallam University for the degree of Doctor of Philosophy

**January 2013**

# Acknowledgements

I am thankful to my supervisor Dr. Dawn Hadden who has given me valuable help, advice and support throughout my studies. I am also grateful to Dr. Ben Abell for the advice he has given to me.

I am also thankful to my friends and colleagues in the Biomedical Research Centre for their practical support and emotional support; in particular I owe a big thank you to Helenne and Claire for their help in the final stretch. I have undoubtedly met some great friends for life.

More than anything, I am grateful to my family and close friends at home. To my friends, especially Hayley, Jamie and Linzi whom have all helped by looking after my children in times of need over the past few years. I will always be grateful that I could rely on you. To my oldest daughter Megan, who has put up with a busy mum yet has continued to make me immensely proud with all her achievements. Also to my youngest daughter Ella, who was born during my PhD studies and who has grown into a beautiful and joyful little toddler despite being incredibly hard work. And finally a massive thank you to my partner Simon, for his endless patience, support and encouragement. Without him I would not have finished my studies.

# Abstract

*Chlamydia trachomatis* is the most common bacterial sexually transmitted disease worldwide. Not only does *Chlamydia* target urogenital tissue that can lead to host infertility, it is also responsible for a high occurence of blinding trachoma in the developing world.

Complete sequencing analysis of the *C. trachomatis* genome revealed a unique family of Polymorphic membrane proteins (Pmps) that had previously remained undiscovered. Nine to twenty-one *pmp* genes were found across the varied *Chlamdia* species, encoding for large proteins all phylogetically related to one of six basic subtypes. *C. tractomatis* encodes nine of these proteins termed PmpA – PmpI. At present, information about the Pmps is limited with the vast majority of data derived from genomic analysis. Such studies have suggested that the Pmps act as virulence factors possibly offering antigenic variation to evade the host defenses. Although this has been displayed genotypically it is yet to be confirmed phenotypically. One of the major problems for studying such proteins in this bacterium, is that *Chlamydia* reside within a host cell and differentiate between two very different morphological forms during infection and the Pmps appear to be expressed at different stages within this biphasic life cycle. Additional theories have suggested that the Pmps act as adhesions, possibly displaying tissue specific tropism in attachment to host cells, an essential step in Chlamydial growth and survival. This has previously been shown for specific regions of Pmp D and Pmp G homologues belonging to *C.pneumoniae* but has not been repeated with other Pmps. Furthermore, bioinformatic data has proposed that the Pmp proteins are autotransporter proteins, a secretory protein that forms a beta-barrel pore within the outer membrane to allow translocation of the functional protein to the bacterial outer surface.

The aim of this study was to gain a better understanding of the Pmps proteins. Specifically to clone, express and purify full length or truncated recombinant Pmps (rPmp) for secondary structural and functional analysis to test the theory that Pmps are putative adhesins. Expressing membrane proteins in a heterologous system poses many obstacles and the expression of rPmp proteins was hindered by the formation of insoluble aggregates and low yields. Several rPmps were produced and three truncated rPmps were expressed as soluble proteins with the use of nanodiscs to mimic the lipid bilayer. Purification of these domains resulted in a substantial loss of protein and subsequently secondary structural analysis could not be carried out on these proteins. However protein interactions were tested with urogenital cell lines using chemical cross-linking, immunocytochemistry, surface plasmon resonance and ellipsometry. Truncated rPmp domains of Pmp D, Pmp G and Pmp I were shown to interact with endometrial cells, supporting the theory that Pmps are involved in attachment. Also, putative proteolytic cleavage was observed, possibly indicating PmpI is cleaved and released from the Chlamydial surface.

# Table of Contents

X

# List of Figures

## Chapter 3

## Chapter 4

## Chapter 5

**Chapter 6**

# List of Tables

# Abbreviations

| | |
|---|---|
| Ab | Antibody |
| ABC | Adenosine triphosphate binding cassette |
| Ag | Antigen |
| Amp$^R$ | Ampicillin resistant |
| ATP | Adenosine triphosphate |
| BLAST | Basic Local Alignment Search Tool |
| b-ME | Beta-mercaptoethanol |
| bp | Base pairs |
| BSA | Bovine serum albumin |
| cAMP | Cyclic adenosine monophosphate |
| CD | Circular dichroism |
| CF | Cell-free |
| CIP | Calf intestine alkaline phophatase |
| CMC | Critical Micelle Concentration |
| COMC | Chlamydial outer membrane complex |
| CTD | Cytolethal Distending Toxin |
| DAPI | 4', 6-diamidin-2-phenylindole |
| DDM | n-Dodecyl β-D-Maltopyranoside |
| DMSO | Dimethyl sulphoxide |
| DNA | Deoxynucleic acid |
| dNTPs | Deoxynucleotide triphosphate |
| DPBS | Dulbeccos phosphate buffered saline |
| EB | Elementary body |
| EDTA | Ethylenediaminetetraacetic acid |
| ER | Endoplasmic reticulum |
| Fc1/Fc2 | Flow cell 1 or 2 |
| FITC | Flurescein isothiocyanate |
| GAGs | Glycosaminoglycans |
| GPCRs | G-protein coupled receptors |
| His | Histidine |
| HIV | Human immunodeficiency virus |
| HRP | Horseradish peroxidase |
| IBs | Inclusion bodies |
| ICC | Immunocytochemistry |

| Ig | Immunoglobulin |
| IMAC | Immobilised metal affinity chromatography |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| Kan$^R$ | Kanamycin resistant |
| kDa | Kilodalton |
| LB | Luria broth |
| LGV | Lymphogranuloma venereum |
| LPS | Lipopolysaccharide |
| MCS | Multiple cloning site |
| MFP | Membrane fusion protein |
| MOMP | Major outer membrane protein |
| mRNA | Messenger ribonuceic acid |
| MWCO | Molecular weight cut off |
| NADPH | Nicotinamide adenine dinucleotide phosphate |
| NEB | New England Biolabs |
| NHS | *N*-hydroxysulfosuccinimide |
| Ni-NTA | Nickel-nitrilotriacetic acid |
| NLPs | Nanolipoproteins |
| OD | Optical density |
| OMP | Outer membrane protein |
| P | Pellet |
| PAH | polyallylamine hydrochloride |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PID | Pelvic inflammatory disease |
| Pmps | Polymorphic membrane proteins |
| RBS | Ribosome binding site |
| RBS | Reticulate body |
| rPmps | Recombinant Polymorphic membrane proteins |
| RU | Response units |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SN | Supernatant |
| SPR | Surface plasmon resonance |
| SRP | Signal recognition particle |
| STD | Sexually transmitted disease |
| T1S | Type I secretion system |

| | |
|---|---|
| T2S | Type II secretion system |
| T3S | Type III secretion system |
| T4S | Type IV secretion system |
| TAE | Tris-acetate-EDTA |
| TARP | Translocated actin-recruiting phosphoprotein |
| TBS(T) | Tris buffered saline (Tween) |
| TCA | Trichloroacetic acid |
| TCP | Total cell-free protein |
| TIRE | Total Internal Reflection Ellipsometry |
| tRNA | Transfer ribonucleic acid |
| UV | Ultra violet |

## 1.1 Membrane proteins

Membrane proteins are essential components of the most basic cellular processes. Despite the extraordinary differences amongst organisms nearly a third of all gene products in both prokaryotes and eukaryotes encode proteins destined to be inserted into a lipid membrane (Wagner *et al.,* 2006). These proteins often form self-assembled complexes, providing a diverse range of essential functions necessary for the uptake of nutrients, waste removal, energy generation, cell division and cell-signalling. A large proportion of membrane proteins are therefore transporters enabling the movement of ions, small molecules and macromolecules across the membrane and thus have key roles in many diseases. Consequently they are of immense economical importance as many current pharmaceuticals and emerging drug candidates target membrane proteins therefore studies to determine the structure and function of such proteins are vital for drug discovery. Despite this, there are very few membrane proteins with known three-dimensional structural information. As of December 2012, 376 unique structures in the Protein Data Bank (PDB) are of independent integral membrane proteins, these account for only ~0.4% of the total entries. This representation is extremely low due to the tremendous challenges associated with studying membrane proteins.

## 1.2 The Chlamydiales

The order of Chlamydiales consists of a large group of non-motile, coccoid, Gram-negative bacteria characterised by their obligatory growth in host animal cells (Cevenini *et al.,* 2002). This order holds one family, the Chlamydiaceae that can be further subdivided into two genera, although *Chlamydia* and *Chlamydophila* taxonomy at present remains controversial due to the rapid advances in this field of research (AbdelRahman and Belland 2006). *Chlamydia trachomatis* is a species belonging to the genus *Chlamydia* as are *Chlamydia pneumoniae, Chlamydia psittica* and *Chlamydia pecorum* (Cevenini *et al.,* 2002). Collectively this family are the etiological mediators of many important human and animal diseases with *C. trachomatis, C. pneumoniae* and *C. psittica* all being pathogenic to humans. However *C. psittaci* is primarily a threat for animals, mainly birds, only infecting humans accidentally. In contrast, *C. pneumoniae* is a widespread human pathogen causing chronic respiratory diseases accounting for 10-15% of community-acquired pneumonia, bronchitis and sinusitis. More recently this organism has been closely associated with a heightened risk of developing atherosclerotic and cerebrovascular problems amongst sufferers (AbdelRahman and Belland 2006). Some strains of non-human origin have also been

2

isolated from marsupials, frogs and horses (Cevenini *et al.,* 2002), indicating that *C. pneumoniae* could be adapting to new environments. Also advantageous to this strain is the ability to infect and replicate within a diverse range of cells including epithelial, endothelial, smooth muscle cells and monocytes/macrophages that would all add to the pathogenesis of this species (Kuo *et al.,* 1995).

### 1.2.1 *Chlamydia trachomatis*

*C. trachomatis* is strictly a human pathogen that primarily infects columnar epithelial cells of the genital and conjunctival regions and thus has a smaller host range than that of *C.pneumoniae* (Cevenini *et al.,* 2002). To date this species has been shown responsible for the most common bacterial sexually transmitted disease (STD) in both the UK (Health Protection Agency, 2012) and worldwide, (Jotblad *et al.,* 2012, Katz *et al.,* 2012) particularly in young adults. Genital infections with this species also carry other identified risks; these include a higher chance of the transmission of HIV and an increased likelihood of developing human papilloma virus-induced cervical neoplasia (Crane *et al.,* 2006). However there is a vast spectrum of other complications associated with a *Chlamydia* STD that are far more common. Amongst these are pelvic inflammatory disease (PID) leading to possible ectopic pregnancy and/or infertility, epididimitis, urethritis, cervicitis and infant pneumonitis that can be passed from mother to child during birth (Nunes *et al.,* 2007). Infant pneumonitis often occurs because the infection is frequently asymptomatic and therefore remains untreated usually being detected during other routine tests. Alternatively when *C. trachomatis* infects the ocular regions it causes trachoma, a chronic follicular conjunctivitis (Crane *et al.,* 2006) that when left untreated results in scarring of the conjunctiva tissues, visual impairment and eventually complete loss of sight. The World Health Organisation (2012) estimates that 300-500 million people are afflicted by trachoma worldwide making it the most prevalent form of infectious preventable blindness. The majority of cases are found in the developing world, where good hygiene and medication are scarce and the disease can easily be transmitted by flies carrying the *C. trachomatis* organsism. The type of disease inflicted by *C. trachomatis* is dependent upon the serovar of the organism that causes the infection, as certain serovars display tropism for specific tissues (Jeffrey *et al.,* 2010).

### 1.2.1.1 Serovars

Within *C.trachomatis* there are 2 subcategories arranged by pathobiotypes, the trachoma biovar and the lymphogranuloma venereum (LGV) biovar (Brunelle and

Sensabaugh 2006). Collectively these are made up of at least 15 serologically distinct serovars grouped according to their disease type, severity and tissue tropism (Carlson et al., 2005). The trachoma biovariants A,B,Ba,C,D,E,F,G,H,I,J and K are responsible for infecting the columnar epithelial cells of genital and ocular tissues. Specifically, serovars A-C are associated with blinding trachoma, while D-K are involved in STDs. LGV biovariants consist of serovars $L_1$, $L_2$ and $L_3$ that penetrate the submucosa of the urogenital epithelial surfaces infecting the primary lymph nodes (Brunelle and Sensabaugh 2006) and subsequently cause more invasive urogenital diseases often involving genital ulceration, lymphadentitis and proctitis (Nunes et al., 2007). Although Chlamydiae prove to be a diverse group of bacteria in terms of disease, tissue tropism and pathogenesis they share one common unifying feature, their biphasic life cycle that bears much resemblance to that of the viral life cycle (AbdelRahman and Belland 2006).

## 1.3 The Chlamydial Life Cycle

In order to reproduce, all Chlamydia alternate between 2 morphological states, the elementary body and the reticulate body with an intermediate inclusion body phase (Nunes et al., 2007). A full replication cycle takes between 40-72 hours (Cevenini et al., 2002) and is finalised by the release of numerous elementary bodies to further spread infection.

## 1.3.1 Elementary body

Chlamydiae in the form of the elementary body (EB) are in an infectious but metabolically inert state adapted for survival in an extracellular environment (Nunes et al., 2007). At approximately 0.3 μm, the nucleoid is highly compact due to chromosomal condensation by the bacterial histone-like proteins HctA and HctB (AbdelRahman and Belland 2006). Their cell wall is rigid (**section 1.5**), ensuring osmotic stability by the cross-linking of the cysteine rich outer membrane proteins major outer membrane protein (MOMP), OmcB and OmcA, collectively forming the Chlamydial outer membrane complex (COMC). Also thought to contribute is a hexagonal arrayed protein layer made predominantly of OmcB situated on the inner surface of the COMC.

The EB initiates the primary stage of replication by attachment and subsequent entry to the host cell shown in **Figure 1.1** (Cevenini et al., 2002). The ability of Chlamydia to efficiently infect a diverse range of non-phagocytic cell types indicates that the bacterial adhesin mechanism involved must recognise a conserved cellular receptor

4

amongst the hosts (AbdelRahman and Belland 2006). So far neither of these adhesins or receptors has been positively identified. Some suggestions for adhesions point towards the bacterial OmpA, OmcB and Hsp70 proteins, but the more recent discovery of the polymorphic membrane proteins (Pmps) has led to the proposal that these large outer membrane proteins could play an important role in attachment.

The interaction between the EBs and the host cell surface is thought to be a two-step process with initial attachment being reversible. This occurs through electrostatic interactions of the bacteria via Chlamydial OmpA, OmcB and/or Pmps possibly with host cell heparan-sulfate containing glycosaminoglycans (GAGs). The second interaction is irreversible and the receptor involved is yet to be identified. Clifton *et al.* (2004) found that immediately following the irreversible binding step, a Type III Secretion Sytem (T3S) exported protein is transported into the host cell. This protein is rapidly phosphoylated at the tyrosine residues with actin recruitment likely to act as a signalling molecule, subsequently getting its name TARP (Translocated actin-recruiting phosphoprotein). TARP is likely to be involved in triggering the uptake of *Chlamydia* by phagocytosis or as an early effector molecule involved in the primary differentiation of EBs into RBs. Once internalised the EB sits protected from the host's defences in a non-acidified phagolysosome/inclusion where differentiation into the RB begins (Cevenini *et al.,* 2002).

### 1.3.2 Inclusion membrane and energy acquisition

Replication within a unique non-acidified membrane bound vacuole, termed the inclusion, occupies the next stage in the life cycle. Very little is known about the biogenesis and composition of this membrane and how it functions. As the chlamydial micro-colony grows and chlamydiae advance through their developmental cycle, the enveloping inclusion membrane must enlarge to accompany this (**Figure 1.1**). Acting as a barrier, impermeable to molecules as small as 520 Daltons it allows the passage of very low molecular weight compounds to diffuse freely across it (Grieshaber *et al.,* 2002). Moreover this barrier must also be crossed by chlamydial proteins delivered into the host cell cytosol therefore it is likely to contain proteins that serve this purpose (Zhong *et al.,* 2001). One group of proteins that have been discovered in the inclusion membrane are termed the Inc proteins (Rockey *et al.,* 2002). They share minimal sequence identity and are considered to provide a diverse range of functions acting as effector molecules involved in host cell signalling pathways and interacting with the host microtubule network and centrosomes (Mital *et al.,* 2010) possibly secreted via

Type III secretion (**section 1.4.4**), yet largely the function of these proteins remain uncharacterised.

Unable to synthesise ATP for replication and intracellular survival, *Chlamydia* acquire nutrients and energy from host cell membrane lipids, such as sterols, sphingolipids, glycerophospholipids and neutral lipids from as early as two hours post infection (Scidmore *et al.,* 2003, Valdivia 2008). This process is characterised by an increased metabolic function of the mitochondria and an enhanced uptake of long-chain fatty acids using both vesicle-dependent and vesicle-independent methods (Cocchiaro *et al.,* 2008, Valdivia 2008).

Vesicle dependent uptake is achieved when the inclusion interrupts the passage of Golgi-derived exocytic vacuoles and multi-vesicular bodies which move from the Golgi to the plasma membrane. Via inclusion-vacuole fusion the chlamydiae seize the vesicle contents (Valdivia 2008). Vesicle-independent uptake involves interactions of the inclusion with endoplasmic-derived lipid droplets. These organelles are filled with energy rich substrates (sterol esters and triaglycerides) encapsulated by a phospholipid monolayer and utilised as an energy source during replication (Cocchiaro *et al.,* 2008).

### 1.3.3 Reticulate body

The reticulate body (RB) is the metabolically active non-infectious form more specialised for replication and is subsequently found after infection upon entry into the host cell (AbdelRahman and Belland 2006) (**Figure 1.1**). Structurally it has a more flexible, thus fragile cell wall adapted for intracellular survival where it resides in the inclusion protected from the host's immune defences. Its size ranges from 0.5-1.0 µm making them much larger than EBs with less dense nuclear material and more ribosomes. The RBs surface is covered with T3S needle-like projections that extend and penetrate the inner surface of the inclusion membrane (Matsumoto 1982, Peters *et al.,* 2007).

The exact triggering mechanism of primary differentiation from EB to RB has not been clearly identified. However, the process can be blocked by transcription and translation inhibiting antibiotics suggesting *de novo* protein expression is necessary to begin intracellular growth (Scidmore *et al.,* 1996). Carry over mRNA from the EBs is rapidly degraded as transcription of new mRNA needed for growth begins (AbdelRahman and Belland 2006). These genes encode proteins involved in the translocation of host-derived metabolites into the bacterial cell such as ADP/ATP

translocase, nucleotide phospate transporters and oligopeptide permease (Shaw *et al.,* 2002). All these proteins are essential for the parasitic growth mechanism of the RB, which feeds from the host's energy supply to undergo binary fission. Following the period of rapid cell division, 10-12 hours post-infection, the RBs begin to re-differentiate back into EBs (AbdelRahman and Belland 2006). Again the signal for this remains unclear. Possibilities include quorum sensing, although genomic analyses do not support this. Other proposals involve the dissociation of RBs from the inclusion membrane via the removal of the T3S needle apparatus from the inclusion's inner membrane surface (Bavoil *et al.,* 2000). However no definitive conclusions have been made.

Secondary differentiation involves the expression of numerous late-cycle genes encoding components of the COMC (OmcA and OmcB) and proteins involved in condensation of the chromosome (HctA and HctB) (AbdelRahman and Belland 2006). Also at this stage it is thought that some of the early proteins associated with the EB infection are expressed, such as T3S needle like projection proteins (**section 1.4.4**) for entry and Pmps for possible attachment. This stage is quickly followed by host cell lysis with the release of newly formed EBs back into the extracellular environment to start the process over again. Given the marked differences in the outer membrane organisation between the EBs and the RBs, knowledge of the composition of the outer membrane is important for understanding the life cycle.
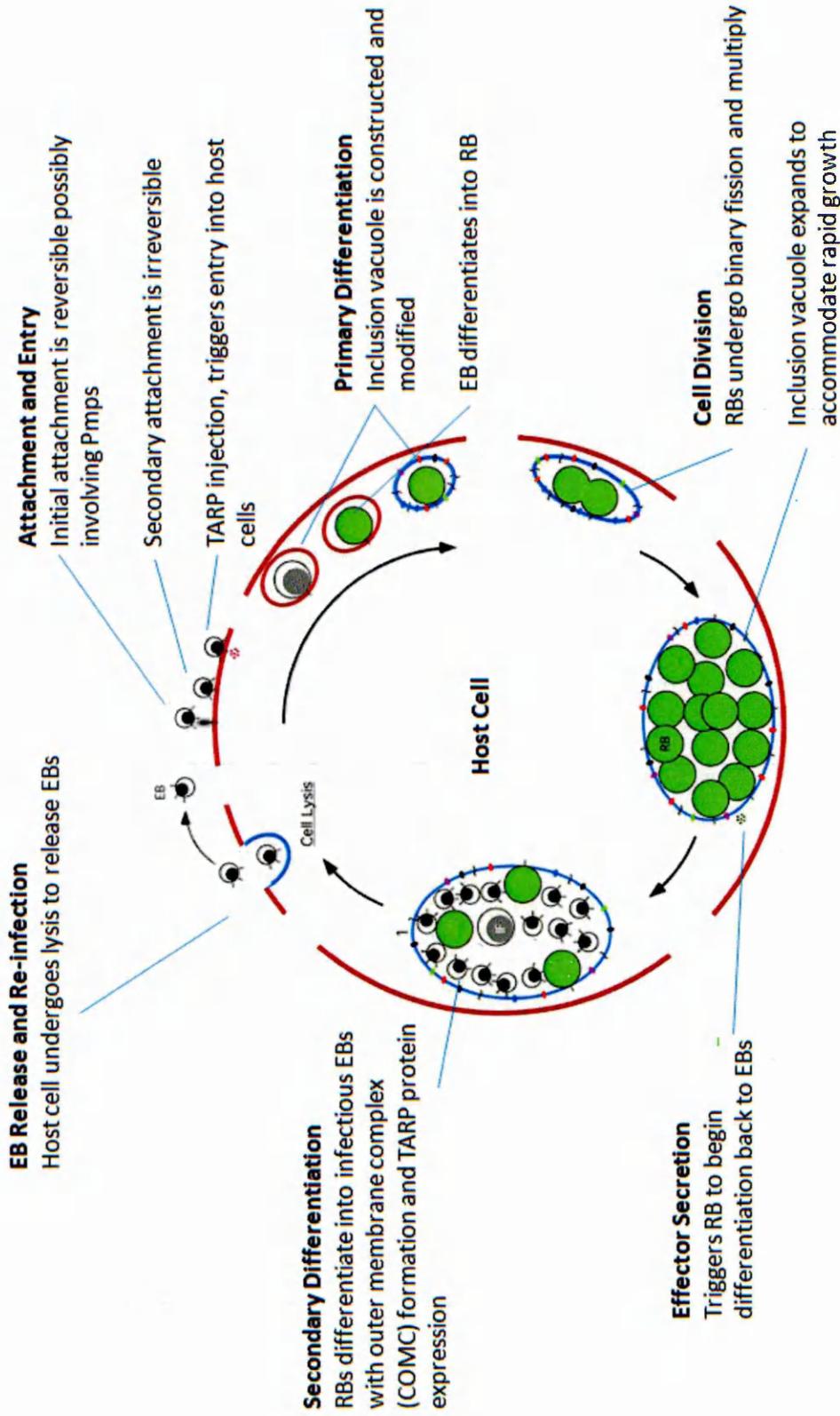
**Figure 1.1 - Diagrammatical representation of the Chlamydial life cycle.** The host cell cytoplasmic membrane is shown in red. TARP (Translocated Actin-Recruiting Phosphoprotein). Adapted from Abdel-Rahman and Belland (2005).

8

## 1.4 The Gram-negative Envelope

In contrast to Gram-positive bacteria, which only possess one biological membrane, the cell envelope of Gram-negative bacteria is composed of two asymmetrical membranes (**Figure 1.2**) (Desvaux *et al.,* 2004a). An inner and outer membrane system sandwiches the periplasmic space and the thin peptidoglycan layer characteristic of Gram-negative organisms (Beveridge 1999). In the context of protein transport, Gram-negative organisms can also be referred to a 'diderm' bacteria (two lipid bilayers) as oppose to Gram-positive which are 'monoderm' due to the single bilayer (Desvaux *et al.,* 2009). This double bilayer structure produces a robust cell envelope, which can endure extreme pH and temperature changes whilst being elastic in order to expand under atmospheric pressure changes. The outer membrane comprises phospholipids, proteins and a complex lipopolysaccharide (LPS). The lipid region of the LPS acts as an endotoxin causing illness upon entering the circulatory system. Although the cell envelope of Gram-negative bacteria is a more complex structure it is still capable of transporting essential substances both in and out of the cell. This takes place via membrane spanning proteins that form pores or via more complicated transport systems. To date, five major secretion systems have been described among Gram-negative bacteria (types I to V **sections 1.4.2** to **1.4.6**), (Desvaux *et al.,* 2004a), some of these mechanisms are shown to be important in *Chlamydia*.
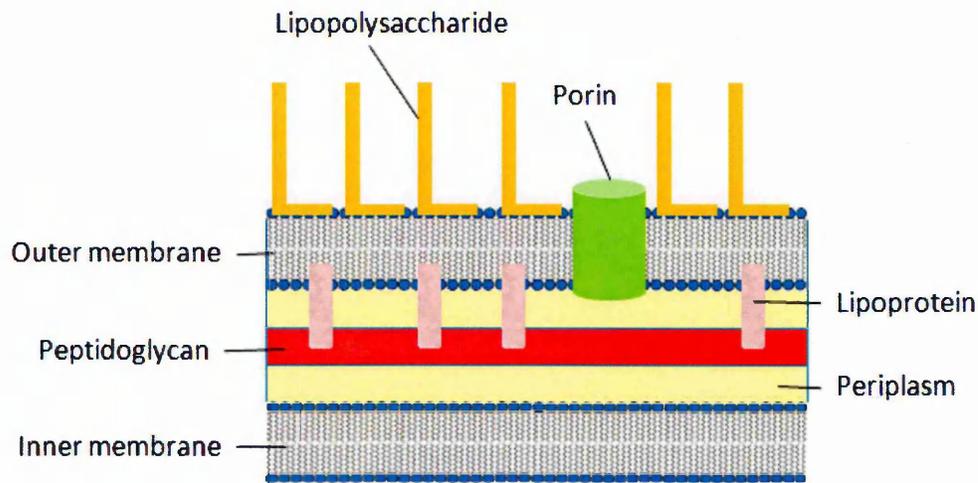
**Figure 1.2 - Schematic diagram depicts the typical structure of the Gram-negative cell wall.** Shown is the outer membrane linked by lipoproteins to a single-layered peptidoglycan. The peptidoglycan is situated within the periplasmic space sandwiched between the outer and inner membranes. The outer membrane includes porins, which allow the passage of small hydrophilic molecules across the membrane, and lipopolysaccharide molecules that extend into extracellular space. The Chlamydial cell envelope differs by the loss of the peptidoglycan layer, instead a highly cysteine enriched protein network is in place in the outer membrane (described in section 1.5). Adapted from Cabeen and Jacobs-Wagner (2005).

## 1.4.1 The Sec Translocon

Protein secretion and protein insertion into the membranes is essential in outer membrane protein biogenesis and relies on specific transportation systems within the cell. For this there exists a ubiquitous membrane protein complex positioned in the plasma membrane of bacteria and archaea (SecY), and in the ER of eukaryotic cells (Sec61) (Gold *et al.,* 2007). The bacterial core translocase consists of the integral inner membrane proteins SecY, SecE, and SecG, which make up an oligomeric complex (Sijbrandi *et al.,* 2003) **(Figure 1.3)**. Two modes of translocation, co- and post-translational are carried out at the Sec complex. Some of the proteins synthesised in the cytosol carry N-terminal signal sequences, which are recognised for Sec targeting. In the event of co-translational translocation, the ribosome-nascent chain complex is targeted to the membrane via the signal recognition particle (SRP) and its receptor. In principle, the SRP-targeting pathway in *E. coli* is homologous to the SRP-targeting pathway that targets both secretory and membrane proteins to the endoplasmic reticulum (ER) membrane in eukaryotes. The eukaryotic SRP binds to the N-terminal signal sequence of secretory and membrane proteins as soon as it is exposed outside the ribosome (De Gier and Luirink 2001). Further translation is inhibited until the SRP contacts its receptor at the ER membrane and then dissociates from the nascent chain, occurring in close proximity to the Sec complex, enabling protein translocation to be driven by the associated chain elongation (Gold *et al.,* 2007). During post-translational translocation, bacteria have a mechanism that pushes proteins across the channel by use of a motor protein ATPase SecA. In addition, the reaction requires a cytosolic component SecB to aid targeting to the membrane and to prevent premature protein folding and aggregation of the pre-protein. Upon translocation the protein is delivered into the periplasm and the signal peptide is cleaved.
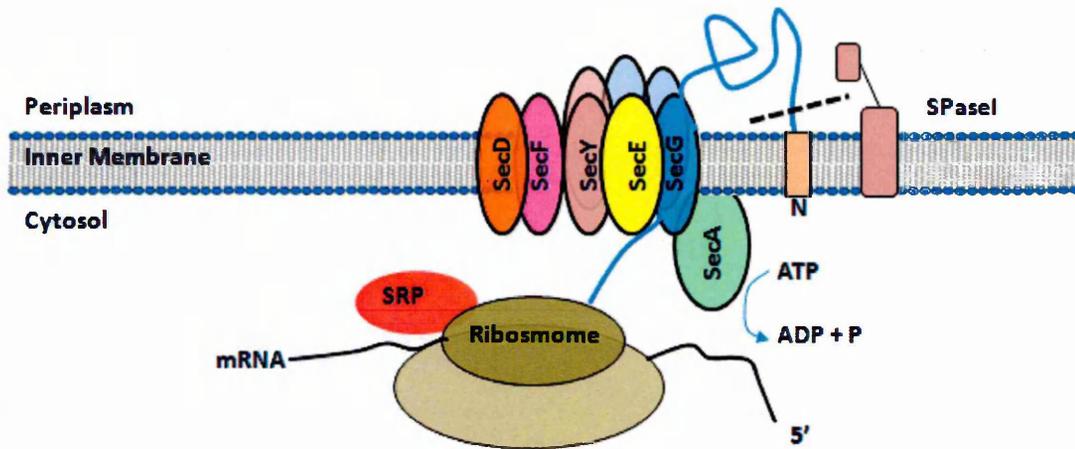
**Figure 1.3 - Diagram of the SRP mediated pathway for Sec co-translationanal pathway**. The ribosome-nascent chain complex is targeted to the membrane via the signal recognition particle (SRP) and its receptor. Translation is inhibited until the SRP contacts its receptor at the membrane. Protein translocation is driven by the associated chain elongation into the periplasm. During post-translational translocation, bacteria have a mechanism that pushes proteins across the channel by use of a motor protein ATPase SecA. The cytosolic component SecB aids targeting to the membrane and to prevent premature protein folding and aggregation of the pre-protein in the cytoplasm. Upon translocation the protein is delivered into the periplasm and the signal peptide is cleaved.

## 1.4.2 Type I Secretion

T1S is associated with an oligomeric complex consisting of three functional subunits: the pore-forming outer membrane protein, a membrane fusion protein (MFP), and within the inner membrane sits the ATP binding cassette (ABC) protein (Binet *et al.,* 1997). It is largely believed that secretion is Sec independent and is a continuous procedure through both membranes. Continuous secretion is initiated when an effector molecule binds to the inner membrane ATP binding cassette. The effector molecule itself contains a secretion signal located within its C-terminal region, this signal sequence is thought to be specific, only interacting with its dedicated ABC transporter protein (Jarchau *et al.,* 1994). Therefore the molecules secreted by T1S are not processed during secretion and do not form distinct periplasmic intermediates (Henderson *et al.,* 2004). The binding of the effector molecule to the ABC protein causes an interaction between the MFP and the outer membrane protein allowing secretion. One example of this is the secretion of alpha hemolysin (HlyA) a large exotoxin from *E.coli*, believed to interact with host cells leading to insertion in to the plasma membrane to of eukaryotic cells (Gentschev *et al.,* 2002). This T1S is made up of TolC (outer membrane pore protein), HlyD (MFP) and HlyB (ABC protein). Binding of the effector molecule triggers the interaction between HlyD and TolC to allow passage of the effector molecule. However, there are various suggestions for the mechanisms of the complex detailing small disparities between different species of bacteria. To date, TS1 has not been identified amongst the Chlamydia species.

## 1.4.3 Type II Secretion

All proteins secreted via the T2S have an N-terminal signal peptide (Henderson *et al.,* 2004). Secretion via this pathway is a two step Sec-dependent process in which the protein or molecule is first exported across the inner membrane under the direction of SecB, into the periplasmic space (Desvaux *et al.,* 2004b). Sec B is a cytoplasmic chaperone that recognises the mature part of the unfolded protein targeted for secretion and subsequently directs it to the Sec translocon (**section 1.4.1**). Upon transportation into the periplasm, studies have shown that the protein adopts a quasi-native state which is aided by certain chaperones such as disulfide bond isomerise, DsbA (Thanassi 2002). These steps are necessary for further transportation across the outer membrane complex termed the secreton. The secreton is a multimer of up to 16 protein subunits (Desvaux *et al.,* 2004b, Desvaux *et al.,* 2004c) that form a channel to allow translocation to the external milieu. The

best studied T2S is for the pullulanase enzyme PulA from *Klebsiella oxytoca*. This starch hydrolysing lipoprotein forms micelles once secreted from the bacteria (Takizawa and Murooka 1985). This pathway has been identified in a variety of other bacterial secretory proteins including cholera toxin of *Vibrio cholera*, exotoxin A of *Pseudomonas aeruginosa* and several other cell-wall degrading enzymes (Sandkvist 2001).

It is also proposed that this process uses the Tat (twin-arginine translocation) pathway (Desvaux *et al.*, 2004b). In contrast to the Sec pathway (which transports proteins in an unfolded manner), the Tat pathway serves to actively translocate folded proteins across the lipid bilayer. Like TS1, this secretion system has also not yet been identified as a method of protein secretion amid the *Chlamydia* species.

## 1.4.4 Type III Secretion

The type III secretion system was first identified in pathogenic *Yersinia* spp. for the secretion of Yop proteins (Henderson *et al.*, 2004). Since then more T3S proteins have been found in several plant and mammalian pathogens including *Salmonella enterica, Shigella flexeri, E.coli* and *C.trachomatis*. Type III secretion systems are largely associated with virulence amongst Gram-negative pathogens so their presence in *C. trachomatis* has raised much interest.

T3S translocate their effector molecule across both the inner and outer membranes in a sec-independent manner (Henderson *et al.*, 2004). The mechanism in which the effector molecule itself is targeted to, or recognises the T3S apparatus remains unclear. However, once the effector molecule reaches the cytoplasmic side of the T3S, secretion proceeds through the needle-like structure of the apparatus (**Figure 1.4**)

The T3S apparatus is composed of a complex array of about 20 different proteins, conserved among T3S systems that span both the inner and outer membranes (Stephens and Lammel 2001). The needle-like structure (CopN) protrudes from the surface of the bacterium and is able to penetrate eukaryotic host cells to inject effector molecules (CopD/CopB) (Peters *et al.*, 2007). The components of the chlamydial T3S are clustered in three operons. In terms of protein structure and content the core apparatus bears a strong resemblance to a Gram-negative flagellar basal body. The secretin proteins in the outer membrane oligomerise to form the outer membrane pore CdsC at the distal portion of the complex (Fields *et al.*, 2006). Among the T3S systems this region generally has surface-exposed domains, normally these do not directly interact with host components, however, this is

different in *C. trachomatis* where they are thought to be host-interactive. The distal portion is linked to the inner membrane components by the CdsJ region that spans the periplasm (**Figure 1.4**). The anchored region in the inner membrane is made up of integral proteins CdsQ-V and CdsD. Portions of CdsD are also surface-exposed and may also be host-interactive (Tanzer and Hatch 2001a). These T3S systems are energy dependent, possessing an ATPase thought to be the CdsN protein in *Chlamydia* (Hueck 1998).

To date there has been no definitive mechanism identified as to what triggers the T3S function in chlamydiae (Fields *et al.,* 2006). Evidence suggests that they are polarised and are induced upon contact with host cell membranes (Hueck 1998). T3S needle-like protrusions can be seen on the surface of RBs and may be involved in the differentiation of RBs into EBs as well as having some role in the attachment and entry into host cells (**section 1.3**).
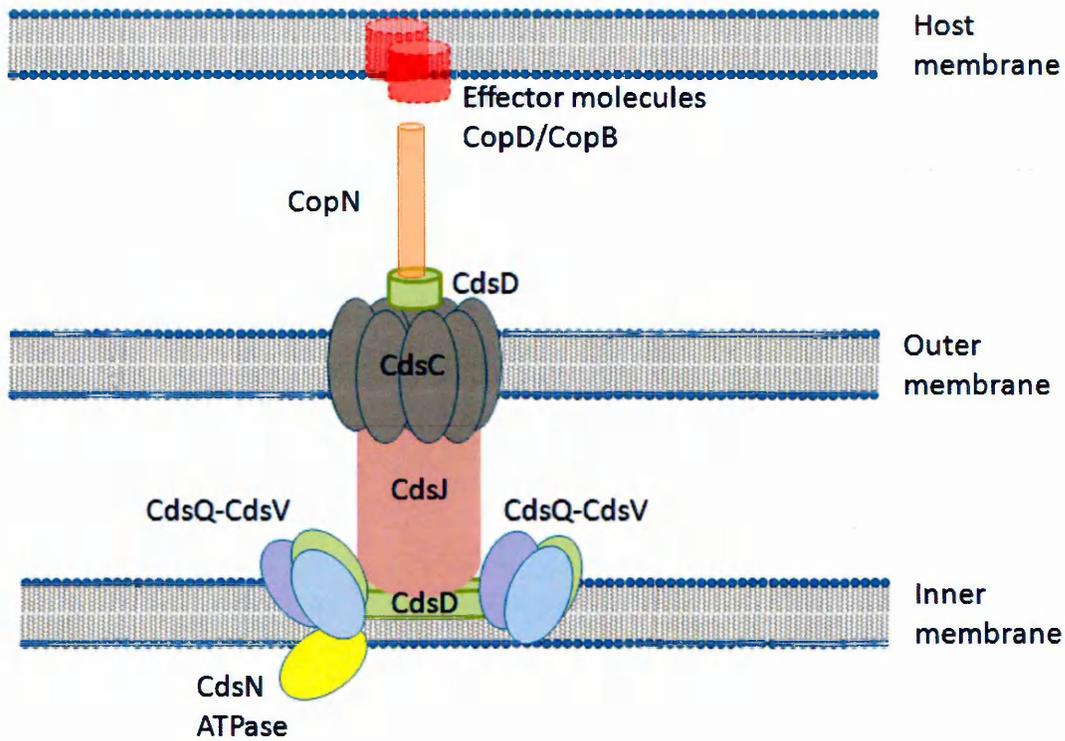
**Figure 1.4 - Diagrammatical representation of the chlamydial T3S apparatus.** The secretin-like protein CdsC sits exposed in the outer membrane and is linked by the periplasmic lipoprotein CdsJ to the multiple inner membrane proteins. CopN signifies the surface exposed needle-like projection. Adapted from Peters *et al.,* ( 2007).

## 1.4.5 Type IV Secretion

The least characterised of the secretion systems, T4S is mainly associated with the transfer of DNA by conjugation (Desvaux *et al.*, 2004b). As with other pathways, the effector molecules also have a wide variety of other functions and thus secretion includes multi-subunit toxins, monomeric proteins as well as nucleoprotein DNA conjugation intermediates (Henderson *et al.*, 2004). One well studied mechanism is of secretion of the pertussis toxin from *Bordetella pertussis,* the causative agent of whooping cough. This toxin belongs to the A-B$_5$ family of toxins (Tamura *et al.*, 1982). Unlike other T4S, this multi-subunit protein is secreted to the extracellular milieu rather than directly into the host cell. Subunits S2-S5, also known as the B domain, interact with host cell glycoprotein receptors mediating translocation of the A domain (subunit S1) into the host cytosol to interfere with signalling pathways (Malaisse *et al.*, 1984). In contrast, the *Agrobacterium tumefaciens* T-DNA transfer system involves the secretion of several effectors, VirD2, VirE2 and VirF directly into the host cell cytosol (Citovsky *et al.*, 1992, Howard *et al.*, 1992). VirD2 is secreted as a nucleoprotein complex, covalently associated with a single-stranded copy of T-DNA for transfer. Delivery of the T-DNA into the host cell is mediated upon interaction of the complex with VirE2 resulting in crown gall tumour formation. The role of VirF is not yet fully understood. The T2S system can be either Sec dependent or Sec independent (Desvaux *et al.*, 2004b). The transfer of T-DNA by *A.tumefaciens* takes place as a single continuous step whereas the export of pertussis toxin resembles that of the T2S pathway where secretion occurs in two steps. First the toxin subunits are translocated across the inner membrane via Sec machinery and then are subsequently targeted for export across the outer membrane (Henderson *et al.*, 2004).

## 1.4.6 Type V Secretion System

Autotransporters are a distinct family of secreted proteins employing the type V secretion pathway (Va or AT-1 system) in which all the necessary elements for transportation of the protein are contained within the polypeptide sequence itself (Veiga *et al.*, 2004). Other proteins categorised under the type V secretion umbrella are those secreted via the the two-partner secretion pathway (type Vb), and the type Vc system (also termed AT-2) (Henderson *et al.*, 2004). Proteins secreted via these pathways have similarities in their primary structures as well as prominent similarities in their modes of biogenesis. Members of this protein family include

17

virulence factors such as the IgA1 proteases from *Neisseria gonorrhoeae* and *Haemophilus influenzae*.

### 1.4.6.1 Va pathway - Autotransporters

Generally autotransporters are large secretory proteins made up of three distinct domains shown in **Figure 1.5a**. The N-terminal region contains the signal peptide leader sequence that targets the protein to the Sec apparatus in the inner membrane (Desvaux *et al.*, 2004a) and this region itself is composed of three domains, a positively charged amino terminus (*n*-domain), a hydrophobic core region (*h*-domain) and a consensus signal peptidase recognition site, SPase1 (*c*-domain) that is cleaved by components of the Sec machinery to release the pro-protein into the periplasm. Adjacent to the signal domain is the functional passenger domain. This region offers the most sequence diversity among the proteins as this performs the extracellular function. To date this function has always been linked to a role in bacterial virulence, either by exhibiting enzymatic activity (protease, peptidase or lipase) or by acting as an adhesin, toxin or cytotoxin. Finally a 25-30 kDa C-terminal region builds the translocation unit needed to export the passenger domain. This can be further divided into two subunits, a short linker region with a α-helical structure and a β-domain which folds into a β-barrel secondary structure forming a pore embedded in the outer membrane (**Figure 1.5b**) (Maurer *et al.*, 1999). The mechanism which triggers the folding of the C-terminal domain is still not recognised. Periplasmic chaperones could be involved or the process may just be spontaneous yet the β-domains of all autotransporters are thought to have some homology. Importantly there are several proposals regarding the translocation mechanism of the passenger domain, including a monomeric form (**Figure 1.6a**) and an oligomeric form whereby autotranslocator β-domains form oligomeric rings in the membrane to produce a large pore for translocation of several passenger domains concomitantly shown in **Figure 1.6b** (Henderson *et al.*, 2004). Furthermore, these passenger domains may sometimes be cleaved to produce a mature active protein.

It is proposed that autotransporters have and continue to evolve by domain shuffling such that novel members arise by the linking a new passenger domain to a generic β-barrel domain (Henderson *et al.*, 2004). Phylogenetic analysis has shown that horizontal transfer of autotranslocator β-domain encoding regions between distant organisms is a rare event yet functional passenger domains have spread by horizontal transfer (Henderson and Nataro 2001b) suggesting that most

autotransporters have arisen through fusion events between passenger domains and β-domains.

### 1.4.6.2 Vb pathway

The bacterial two-partner Vb secretion pathway shares many similarities to the conventional autotansporter Va secretion pathway whereby the passenger domain possesses a signal sequence that directs translocation across the inner membrane and following its export into the periplasm, the passenger domain inserts into an outer membrane pore formed by a β-barrel (Henderson *et al.,* 2000). Once translocation to the surface of the bacterium has occurred, the passenger domain may also undergo further proteolytic processing however, in contrast to the autotransporter pathway, the passenger domain and the pore-forming β-domain (the transporter domain) are translated as two separate proteins referred to as TpsA and TpsB. In addition compared to the Va autotransporter, the β-barrel topology appears different, by containing more amphipathic β-strands compared to the β-domain of type Va autotransporters (Yen *et al.,* 2002).

### 1.4.6.3 Vc pathway

Members of the Oca (Oligomeric Coiled-coil Adhesins) family have been described as a subfamily of oligomeric autotransporters. *Y. pestis* YadA possesses six different domains: an N-terminal signal sequence, head, neck, stalk, linker, and a C-terminal region consisting of only four β-strands (Hoiczyk *et al.,* 2000). This type of trimeric conformation and mode of secretion has been proposed as an alternative model for autotransporter secretion system and has been designated the type Vc or AT-2 pathway (**Figure 1.6c**).
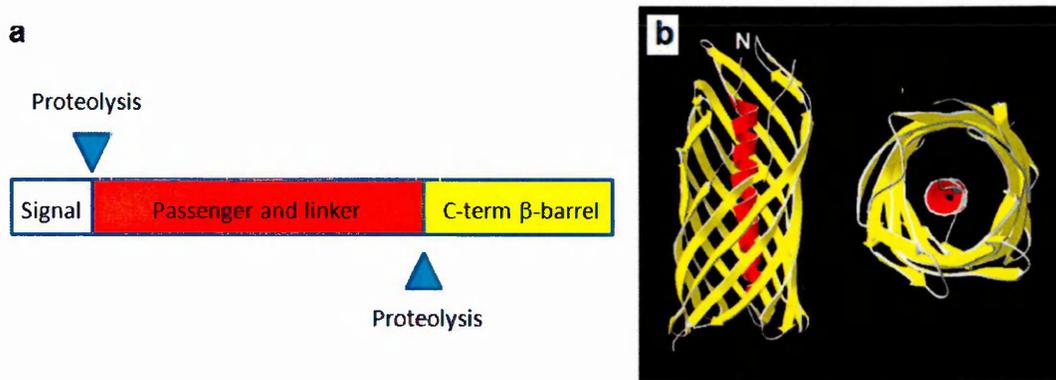
**Figure 1.5 - Schematic representation of the three functional domains and folding of monomeric autotransporter protein.** (a) A short N- terminal leader is necessary for localisation to the inner membrane of Gram-negative bacteria, adjacent to the functional passenger domain (red) linked to the C-terminal autotranslocator β-barrel pore (yellow) adapted from (Henderson and Lam 2001a). (b) Crystal structure of a typical β-barrel autotanslocator membrane domain, the functional passenger/linker domain (red) traverses the translocating unit (yellow). Typically the barrel interior is highly hydrophilic due to the presence of charged amino acids. Taken from Henderson *et al.,* (2004).
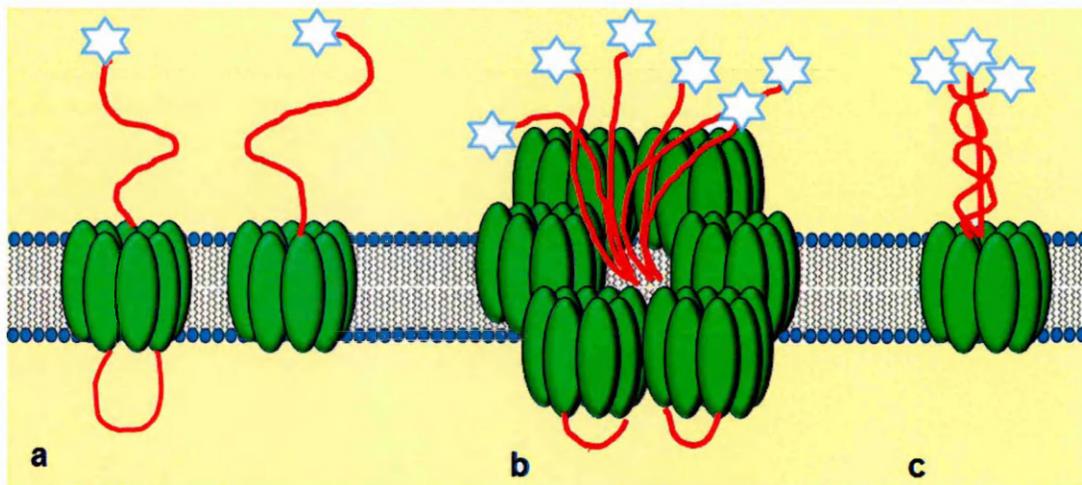


**Figure 1.6 - Diagram of the various proposed strategies for the translocation of the passenger domain to the external surface of the Gram-negative outer membrane.** (a) The autotransporter domain (green) exports the passenger domain (star) outside the cell in typical Va autotransporter pathway, passenger domain may be cleaved from the linker region (red) to release the functional domain or may remain intact with the membrane. (b) The autotansporters form oligomers at the membrane to produce large pores for translocation of the passenger domains. (c) An oligomeric coiled passenger domain forms as tetramers exposed to the external milieu.

20

## 1.5 The Chlamydial Envelope

During all the stages of the developmental cycle (**section 1.3**), *Chlamydia* cells appear to be surrounded by the double membrane system characteristic of Gram-negative cells (**Figure 1.2**) (Bavoil *et al.,* 1984). However, their cell envelope differs by the lack of a peptidoglycan layer within the periplasm (Kosma 1999). This anomaly remains unclear as genomic analyses have shown a complete set of genes that would enable its synthesis. Instead the chlamydial envelope contains a considerable amount of disulfide bond crosslinked proteins in the outer membrane and periplasm, a feature which is thought to stabilize the cell envelope (Hatch 1996) and so the peptidoglycan layer could have been lost by evolution where the requirement for a rigid cell wall in the RBs form was lost due to its intracellular location (Kosma 1999). Like other gram-negative bacteria, Chlamydiae have a vast array of proteins spanning the outer membrane. Many of these are unique to this organism, thus raising much interest in their function. There are difficulties in studying these proteins due to the intracellular residence of the Chlamydiae species, so the majority of data to date is derived from genomic analysis.

Typically the Chlamydial outer membrane proteins (OMPs) boast several functions: firstly they are involved in specific nutrient and metabolite attainment (Rzomp *et al.,* 2006), secondly, they act as adhesions to, or invaders of, the host cells and tissues (Wehrl *et al.,* 2004, Fadel and Eley 2008), thirdly, some are thought to be type III secretion proteins in place to deliver pathogen effector kinases or phosphotases (i.e. TARP) (Jewett *et al.,* 2008); and finally others act as host immune evaders by employing antigenic shift variations (Stephens and Lammel 2001). The main Chlamydial OMPs are crucial components in the attachment and entry of EBs during the life cycle and include OmpA (MOMP), T3S and the Pmps.

## 1.5.1 The Major Outer Membrane Protein

The Major Outer Membrane Protein is a major constituent of the Chlamydial outer membrane with a similar structure and molecular mass of 40kDa among species (Kosma 1999). It functions as a porin bearing structural similarities to porins found in other Gram-negative organisms (Bavoil *et al.,* 1984). This cysteine-rich protein comprises of four domains with seven cysteine residues (Caldwell *et al.,* 1981) that are essential in the formation of the disulfide-linked COMC of the EB (McCafferty *et al.,* 1995). MOMPs are surface exposed, closely associated with LPS and have species-specific epitopes (Batteiger *et al.,* 1993). Su et al, (1996) , found strong evidence to suggest that MOMP functions as a chlamydial cytoadhesin to

glycosaminoglycans (GAGs) of host cells. This protein has therefore been the main focus for an anti-chlamydia vaccine. Partial protection against infection has been observed in animals immunised with preparations of MOMP although protection against re-infection appears to be serovar-specific (Batteiger *et al.,* 1993).

## 1.5.2 The Polymorphic Membrane Proteins (Pmp)

The polymorphic membrane proteins remained largely undiscovered until complete sequence analysis of the *C .trachomatis* genome by Stephens et al, (1998). A super-family of genes were revealed with a nine-member family termed *pmpA - pmpI* present in *C. trachomatis* (Stothard *et al.,* 2003), which has expanded to 21 homologs in *C. pneumoniae* (Mygind *et al.,* 2004). Importantly, these coding regions represent 13.6% and 17.5% of the *Chlamydia*-specific coding regions within these two genomes (Kalman *et al.,* 1999, Read *et al.,* 2000), since *Chlamydia* only retain genomic material critical for survival, these proteins are considered to have a key function in *Chlamydia* biology (Hatch 1998).

## 1.5.2.1 Phylogeny and classification

The Pmp proteins are all phylogenetically related to one of 6 basic subtypes, PmpA, B, D, E, G and H (Grimwood *et al.,* 2001). Pmp C is adapted from Pmp B, Pmp F is derived from Pmp E and Pmp I is a variation of Pmp G. The *pmp* genes vary in size, ranging from approximately 2,600 to 5,300 base pairs, thus encode large proteins between 90 –187kDa in mass (Stothard *et al.,* 2003). They are unique to *Chlamydia* and account for 3.15% of the organism's genetic coding capacity (Gomes *et al.,* 2006). The genes are diverse in nucleotide sequence, sharing between 9-42% sequence homology, however the sequence diversity of individual *pmps* between different serovars has not been explored extensively. Diversification and expansion is most prominent with *pmpG* of *C. pneumoniae, C. caviae* and *C. abortus* where four subtypes have evolved suggesting this Pmp is under selective pressure possibly responding to specific host traits or defences requiring a higher level of antigenic variation (Tan, et al. 2006). Interestingly, some of the *C. trachomatis pmp* genes are located in clusters suggesting that they regulate or may work as one entity, *pmp A, B* and *C* are in close association with each other, as are *E/F* and *G/H* although on opposing DNA strands with *pmp I* in close proximity (**Figure 1.7**) (Gomes *et al.,* 2007). However there is no evidence to suggest they are polycistronic. Instead this may allow for a high mutation rate to provide antigenic variation among species. Conversely *pmp D* sits alone on the chromosome suggesting some important preserved function for this protein.
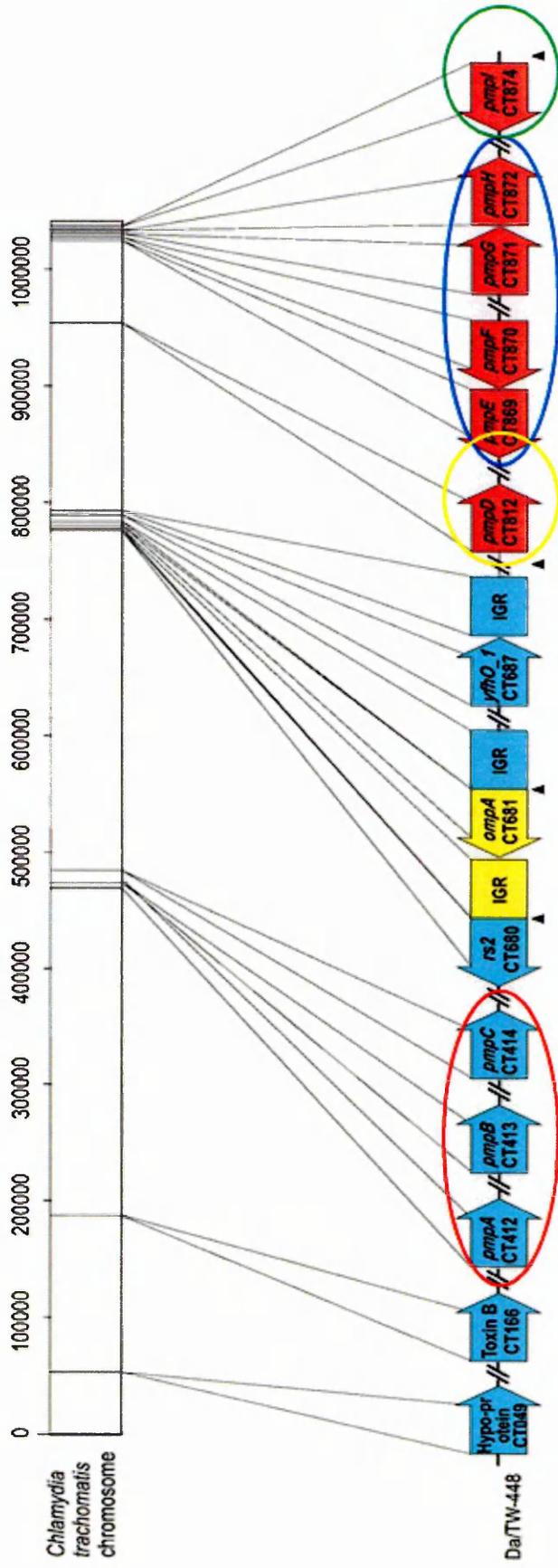
**Figure 1.7 - The linear representation of the ~1.05 Mb chromosome and nucleotide location of the 9 *pmp* gene loci of *Chlamydia trachomatis*.** Adapted from Gomes *et al.*, (2007).

23

## 1.5.2.2 Characteristics of Pmp proteins

The *pmp* family encode several carboxy tryptophan residues and a carboxy terminal phenylalanine typical of proteins that localise to the outer membrane (Stephens *et al.,* 1998). In addition an amino terminal leader peptide sequence consistent with localisation to the inner membrane is also present on most of the Pmp proteins with the exception of Pmp A that differs in this respect (Skipp *et al.,* 2005). Protein sequence analysis has indicated that these Pmp carboxy terminal domains are predicted to form a β-barrel pore structure embedded within the outer membrane and are likely to function as an autotransport secretory protein (type V secretion), transferring the adjacent functional passenger domain through the pore to the extracellular surface as described in **section 1.4.6** (Henderson and Lam 2001a, Henderson and Nataro 2001b). Additionally P-loop motifs associated with pore regulatory gating mechanisms via GTP/ATP binding have been identified within the Pmp I autotanslocator domain and Pmp G passenger domain of *C. trachomatis* (Henderson and Lam 2001a) supporting the formation of a β-barrel pore although a P-loop motif in the passenger domain of Pmp G may not be consistent with regulatory closing of a pore region. In contrast, generally Pmps have relatively high levels of cysteine residues (Tanzer and Hatch 2001a, Tanzer *et al.,* 2001b) inconsistent with the ability to secrete proteins via the autotransporter pathway as studies have shown intra-peptide disulfide bonds inhibit translocation of the passenger domains (Jose *et al.,* 1996). However, studies have since emerged showing that limited disulfide bond formation within the passenger domain of *Shigella* protein IcsA does not prevent translocation of this domain through the β-barrel pore (Brandon and Goldberg 2001). The folding mechanisms of Pmps prior to translocation are unclear. Whether or not these cysteine residues form extensive disulfide bonds such that translocation via an autotransporter β-barrel is impractical has yet to be determined. Alternatively these cysteines may form inter-peptide disulfide bonds with other cysteine-rich proteins such as MOMP, OmcA and OmcB maybe offering added structure and rigidity associated with the EB COMC (Hatch 1996, Henderson and Lam 2001a).

The autotransporter theory is further supported where proteolytic processing of Pmps into amino passenger domain and carboxy β-barrel domains has been positively identified *in vitro* (Grimwood *et al.,* 2001, Vandahl *et al.,* 2002, Wehrl *et al.,* 2004, Vandahl *et al.,* 2005) however these studies have not extended beyond investigations with Pmp D or Pmp D homologues. Also consistent with autotransport is that the putative passenger domains of some of the Pmps are most immunogenic suggesting

this domain is more accessible to the host's immune mechanisms, playing a key role in *Chlamydia* pathogenesis (Crane *et al.,* 2006, Byrne 2010).

The most common features shared by all the Pmps is the repeated tetrapeptide motifs GGAI and FXXN (where X represents any amino acid) found in the putative passenger domains of the proteins (Stephens *et al.,* 1998). These motifs have been associated with attachment to host cells in many other organisms (Grimwood and Stephens 1999). Another feature that indicates a role as an adhesin is the presence of RGD motifs (within the Pmp D and Pmp F passenger domains) of the type implicated in cell-cell adhesion, although such motifs are not recognised in the other *C. trachomatis* Pmps. Furthermore, studies have demonstrated that for some of the Pmps, the passenger domains associates with host cell surfaces including Pmp D of *C. trachomatis* and Pmp D and Pmp G homologues of *C. pneumoniae* (Wehrl *et al.,* 2004, Mölleken *et al.,* 2010a). Speculation that Pmps function as adhesins in Chlamydial attachment is prevalent yet rationally, studies to support this are still limited and lacking in published data. Furthermore, studies have not investigated the other Pmp proteins to determine if they display the same functionality. Additionally there are no repeated studies focussed on the expression and processing of all the Pmps to determine if they all carry the autotransport mechanism.

### 1.5.2.3 Expression, localisation and the role of Pmps

Studies have shown that all nine *C.trachomatis* Pmps are transcribed in EBs, but stable translation and presence in the outer membrane has not been demonstrated for them all. Pmp E, Pmp G and Pmp H have been identified as major constituents of this *C. trachomatis* LGV serovar L2 outer membrane late in the developmental cycle whereas none of the seven other predicted Pmp proteins were detected (Tanzer and Hatch 2001a, Mygind *et al.,* 2006). More recently Pmp B, Pmp C, Pmp D, Pmp E, Pmp F, Pmp G, and Pmp H were found in the COMC of the same serovar (Liu *et al.,* 2010). Pmp A has not been detected in the EB throughout the studies but had previously been detected in the RB (Skipp *et al.,* 2005b) so it is speculated that Pmp A does not reside in the outer membrane of EBs, also supported by bioinformatic analysis that shows a lack of signal peptide in this Pmp (Stephens *et al.,* 1998). However, the Pmps are not abundant proteins in the chlamydial membranes therefore expression levels of certain Pmps not identified in these studies may have been below the limits of detection or expressed in a tissue tropism manner by some serovars but not others.

Evidence shows that Pmps stimulate the expression of cytokines (Johnson 2004, Wang *et al.,* 2005) and serum samples from *C. trachomatis*-infected patients have also shown anti-Pmp antibodies to all nine rPmps (Tan *et al.,* 2009b) indicating Pmps are possibly targeted by the host immune response. Pmp D has already been shown to act as an immune decoy mechanism (Crane *et al.,* 2006). However, there is still no clear picture of how the immune system fully responds to *Chlamydia* infections except that the organism has a dynamic ability to evade the hosts system in the cell inclusion.

## 1.6 The challenges of studying membrane proteins

In order to obtain structural and functional data it is often necessary to produce tens, sometimes hundreds, of milligrams of highly purified material (Bill *et al.,* 2011). However, the majority of membrane proteins are endogenously expressed at levels too low to purify and it is rare that a single protein species is a major component of the membrane. Where this is the case it has been exploited. Bacteriorhodopsin, is the only protein constituent of the purple membrane of the bacterium *Halobacteria salinaria* where straightforward preparation of the membrane was sufficient to obtain high yields of pure protein *(Belrhali et al., 1999)*. The requirement for large amounts of protein has therefore resulted in a bias of structure determination efforts towards membrane proteins which are naturally abundant, such as photosynthetic reaction centres and aquaporins. This is not possible for studies that aim to investigate the structure and function of Pmps as these proteins are not endogenously expressed in high abundance (Tan *et al.,* 2009a). However, membrane proteins can often be over-expressed in an ectopic expression system as recombinant proteins. Although recombinant expression of membrane proteins is technically challenging it is not impossible whilst the inability to perform basic genetic manipulations with *Chlamydia* is well acknowledged and subsequently studies lag behind those of other infective organisms. Since the *C. trachomatis* genome was published less than two decades ago leading to the discovery of the *pmp* gene family, the majority of information with reference to the Pmps is derived from genomic analyses. Studies that have explored the Pmps at the protein level have employed recombinant Pmps in the form of misfolded inclusion bodies and have yet to investigate the structure and function of each of these proteins. An advantage with using recombinant expressed protein is that the facility to manipulate or modify the protein to improve stability, solubility, enable detection and purification is possible (Wagner *et al.,* 2006).

## 1.7 Aims of the Study

The overall aim of this study is to gain an understanding of the role of *Chlamydia trachomatis* Polymorphic Membrane Proteins as potential autotransporters that act as adhesins to initiate infection of host cells. Specifically the aims are to clone, express and purify recombinant Pmps or the predicted Pmp functional domains for structural and functional analyses. Purified recombinant Pmps and recombinant Pmp functional domains can be used in secondary structure analysis to verify that the predicted Pmp C-terminal autotranslocator domains are consistent with the $\beta$-barrel structure of autotransporters as bioinformatic analyses predicts. Binding interactions with female urogenital cells may show that Pmps specifically interact with cells as putative adhesins. Currently there are no effective vaccines to tackle *Chlamydia* infections due to the lack of immunogenic target proteins. Treatment relies upon the administration of antibiotics, which can often be problematic due to the bacteria's ability to evolve and acquire resistance. Data generated from these studies will hopefully contribute to the understanding of the role Pmps play in the attachment to and invasion of host cells. As a crucial step in the chlamydial infective cycle, these proteins could become a suitable target for future vaccine developments in an effort to prevent the spread of chlamydial infections. This is of both fundamental and topical importance given that *C. trachomatis* is the most prevalent STD worldwide.

Specific objectives of this study:

- To produce recombinant expression constructs containing full length *pmpA-pmpI* genes and/or the smaller predicted *pmp* passenger or autotranslocator domain sequences. This is to be achieved by the amplification of each gene and/or domain sequence from genomic DNA using custom designed primers with PCR. Each construct will be designed to incorporate a hexahistidine tag on the C-terminus of the recombinant Pmp for detection with an anti-His antibody during expression trials.

- To attempt the overexpression of recombinant full length Pmps and/or recombinant truncated Pmp domains using *E. coli* as an experimentally amenable system. The expression of recombinant Pmps as inclusion bodies will be avoided since the requirement for protein suitable for structural and functional analyses is of importance. Instead efforts will focus on targeting

expression to the membranes or to a membrane-like environment. This is expected to involve several different expression screening trials and where expression is observed, focus will begin to optimise these methods.

- Purification of the recombinant Pmp proteins using immobilised metal ion affinity chromatography facilitated by the hexahistidine tag on the C-terminus coupled with other conventional purification techniques such ion exchange chromatography will be employed where necessary. The secondary structure of the purified recombinant Pmps will then be examined by circular dichroism.

- Recombinant Pmps expressed and targeted to *E. coli* or artificial membranes (to mimic a native membrane protein environment) will be used in functional adherence assays to investigate Pmps as putative adhesins to female urogenital cells. Several methods will be explored, including methods to exploit the C-terminal poly-histidine tag where generic anti-His antibodies can be used to detect and/or immobilise recombinant Pmps for interaction studies. Specifically, recombinant Pmps will be incubated with whole cells followed by chemical cross-linking to permanently fix recombinant Pmp interactions. Putative interactions will be detected using Western blotting and Immunocytochemical staining techniques.

- Optical biosensor techniques will be employed where appropriate to also investigate putative interactions between the recombinant Pmps and urogenital cells. Methods including Surface Plasmon Resonance and Ellipsometry will be explored, whereby interactions are monitored in real time by changes in reflected light from an experimental surface.

A summary for each the aims and outcomes is displayed in **Figure 6.2**, **Figure 6.3** and **Figure 6.4**.

# Chapter 2 - Recombinant Cloning Strategies for *C.trachomatis* Pmp Expression Constructs

## 2.1 Introduction

This chapter describes the preliminary work undertaken to lay the groundwork necessary for studying the structure and function of *C. trachomatis* Polymorphic Membrane Proteins (Pmps). It describes the development of cloning strategies to produce suitable *E. coli* expression constructs for the production of both full length and truncated recombinant Pmps.

Cloning strategies were developed using the published genome of *C. trachomatis,* serovar D (std.gen.northwestern.edu) for primer design, sequencing alignment and verification. It was thought feasible to use serovar D for this purpose as the genome of serovar E was not yet published and a high sequence homology between some *pmps* amongst the serovars of *C. trachomatis* was reported (Stothard *et al.,* 2003). Using restriction fragment length polymorphism (RFLP), little variation among the 15 serovars for *pmpA* and *pmpD* was shown as both were found to be highly conserved. Interestingly a significant variation in sequences between serovars was evident for some of the other genes, in particular *pmp F, pmpH, pmpI* and *pmpE.* These Pmps are thought to have a role in tissue specific targeting. However serovars D and E are of the same tissue specific tropism both infecting urogenital epithelial cells so any variation within the gene sequence was expected to be modest. Because of repeated similarities, it was expected that *pmpA* and *pmpD* would be the least challenging to clone and interestingly most of the published information that exists to date is based on data obtained from studies of Pmp D where function as an adhesion and as an immune decoy protein has been suggested (Wehrl *et al.,* 2004, Crane *et al.,* 2006). Pmp A also stimulates interest being the only Pmp that does not appear to have a recognised signal leader peptide sequence, suggesting an alternative functional role or localisation (Gomes *et al.,* 2005). Highly variable Pmp F, Pmp H, Pmp I and Pmp E predicted to have a role in tissue specific targeting, are of particular interest as microbial virulence factors (Stothard *et al.,* 2003).

It is widely accepted that expression of recombinant membrane proteins can be a difficult process due to the high attrition rate associated with the stages involved during cloning, expression and purification. Therefore, few recombinant proteins are expressed in native, full-length. One approach to reduce the high attrition is to begin with a larger number of clones using a high-throughput approach (Cowieson *et al.,* 2005). The method used by the Northeast Structural Genomics Consortium Project examines certain characteristics within protein sequences to determine possible structural domain boundaries and subsequently produce nested sets of constructs for

full-length proteins, multi-domain and single domain constructs. These alternative constructs often led to better expression and increased the likelihood for success in further investigations (Gräslund *et al.,* 2008). A similar approach was used to begin this project as it was difficult to determine which particular Pmps would be expressed.

Therefore, as the starting point for this research, one particular Pmp was not singled out for analysis; instead attempts were made to construct expression plasmids for each of the *pmps*. Consequently the targets for cloning were all nine of the full length *pmps*, and truncated *pmps* based on the predicted autotansporter domains by creating expression constructs containing the N-terminal passenger domain predicted to exhibit the functional part of the protein, and the C-terminal autotranslocator region that forms the pore within the membrane (**section 1.4.6**).

The full length sequences were termed *pmpx-FL*, N-terminal passenger domain sequences are termed *pmpx-N* and the C-terminal autotranslocator domains termed *pmpx-C* where x represents the particular *pmpA - pmpI*. All putative *pmp* constructs underwent expression screening, optimisation and ultimately purification for subsequent functional studies (chapters 3, 4 and 5). This was deemed the most favourable approach to begin this project in an effort to increase the outcome of successful expression targets.

## 2.2 Materials/Methods

### 2.2.1 Suppliers used in this chapter

Analytical grade chemicals were used throughout this study with all solutions being prepared using sterile water. Materials were obtained from the following suppliers.

**Bio-Rad Laboratories Ltd.**, Hemel Hemsted, Hertfordshire; **GE Healthcare Ltd.**, Amersham, Buckinghamshire; **Invitrogen Life Technologies.**, Paisley; **Merck Chemicals Ltd.**, Beeston, Nottingham; **New England Biolabs (UK) Ltd.**, Hitchin, Hertfordshire; **Novagen/ Merck Chemicals Ltd.**, Beeston, Nottingham; **Sigma Aldrich Company Ltd.**, Gillingham, Dorset; **Starlab UK.**, Blakelands, Milton Keynes; **Stratagene Ltd.**, Cambridge; **Thermo-Fisher Scientific UK.**, Loughborough, Leicestershire; **Qiagen Ltd.**, Crawley, West Sussex.

Molecular biology methods used were essentially those described in Sambrook *et al.* (2001), with some modifications. Glassware and media were sterilised by autoclaving at 121°C, 15 p.s.i. for 20 minutes. Solutions were made from molecular biology grade reagents and sterilised by filtration or autoclaving before use. Custom DNA primers were synthesised by Sigma Aldrich.

### 2.2.2 Primer Design

Commercial primers were not available for the amplification of *pmp* genes therefore primers incorporating restriction sites were designed to flank the *C. trachomatis pmp* genes and/or the sequences of predicted functional domains. As previously described the nucleotide sequence of the *pmps* of serovar E was not available, therefore serovar D *pmp* sequences were used as a reference source, obtained from the STD sequence database of Los Alamos National Laboratory (std.gen.northwestern.edu). The primer sequences are given in **Table 2.1**.

To allow directional cloning into the multiple cloning site (MCS) of the desired pET vector (**section 3.2.2**) the restriction sites chosen to be incorporated into the primers, were those that did not appear to be present within the gene so each *pmp* sequence was examined using NEBcutter V2.0 (New England Biolabs) to identify restrictions sites within the sequence. In addition the cloning procedure was designed to incorporate a poly-histidine tag at the C-terminus. The restriction sequences were manually aligned with the 5' and 3' regions of the reference gene sequence. This led to some mismatching and minor changes to the N and C terminal sequences as the restriction sequence and gene sequence were not complementary. Adjacent

complementary nucleotide bases were extended either side of the restriction sequence to produce a primer of between 15-30 bases in length with approximately 40-60% GC content. The primer pairs were checked for compatible melting temperature $T_m$ using the DNA calculator (Sigmaaldrich.com). Regions within primers that could result in primer dimers or primer hairpin loops were avoided.

To ensure the resultant protein coding sequence was in the correct open reading frame and bearing the polyhistidine tag, the recombinant theoretical sequence was translated using the Expasy Translate Tool (Expasy http://web.expasy.org/translate/). Bases were added or subtracted accordingly in a way to minimise alteration to the amino acid composition at the N and C terminals. Primers were manufactured by Sigma Aldrich Ltd and are displayed in **Table 2.1.**

### 2.2.3 PCR Amplification of Target DNA from Genomic DNA

A PCR mixture was made up of the primers shown in **Table 2.1** with the components listed in **Table 2.2** according to the KOD Hotstart kit instructions (Novagen). The tested cycle conditions are displayed in **Table 2.3** and the equipment used was MWG Biotech Primus 96 plus thermal cycler. The genomic DNA from a serovar E clinical strain *C.trachomatis* was provided by Dr Adrian Eley, (University of Sheffield). Each of the target DNAs optimal set up and cycle conditions were optimised and are displayed in **Table 2.4**. The PCR products were examined by agarose gel electrophoresis (**section 2.2.4**). Initial positive identification was made by comparison of the size of the PCR product in relation to serovar D predicted *pmp* gene size. PCR products were purified and digested with appropriate enzymes before ligation into pET vectors.

| Gene/Region | Primers | Restriction sites | Vector | Fusion Tag | Selective Marker |
|---|---|---|---|---|---|
| *pmpA-FL* | Forward 5' T**GCTAGC**AATCGAGTTATAGAAATCC 3'<br>Reverse 3' ACGCCTCCGATACAACCGAATC**GCCGGGCG**A 5' | NheI<br>NotI | pET23b | C-term hexaHis | Amp[R] |
| *PmpB-FL* | Forward 5' A**CCATGG**AATGGCTGTGTCAGCTACTGC 3'<br>Reverse 3' GACGCCACGAGCGTACTAGAAG**GAGCTC**T 5' | NcoI<br>XhoI | pET28b | C-term hexaHis | Kan[R] |
| *PmpC-FL* | Forward 5' AT**CCATGG**AATTTATGTCAGCTACTGC 3'<br>Reverse 3' GCCACGAGCATACTGTAA**GAGCTC**TT 5' | NcoI<br>XhoI | pET28b | C-term hexaHis | Kan[R] |
| *PmpD-FL* | Forward 5' TC**CCATGG**GTTCCGAGAAAGATATAAAAAGC 3'<br>Reverse 3' GACCTAACGCTAACTAGAAA**CGCCGGCG**A 5' | NcoI<br>NotI | pET28b | C-term hexaHis | Kan[R] |
| *PmpE-FL* | Forward 5' T**CCATGG**AAAAAGCGTTTTTCTTTTTCC 3'<br>Reverse 3' AGTTACCCCTTTAACGAGAC**GCCGGGCG** 5' | NcoI<br>NotI | pET28b | C-term hexaHis | Kan[R] |
| *PmpF-FL* | Forward 5' CA**CATATG**ATTAAAAGAACTTCTCTATCC 3'<br>Reverse 3' CGTCCTCCTCGACACCAGAAA**GAGCTC**T 5' | NdeI<br>XhoI | pET22b | C-term hexaHis | Amp[R] |
| *PmpG-FL*<br>Set 1 | Forward 5' AACA**GTCGAC**GTTTGCAG 3'<br>Reverse 3' TTTCTAGG**GCGCGGGCG**TTTATAACC 5' | SalI<br>NotI | pET22b | N-term pelB<br>C-term hexa-His | Amp[R] |
| *PmpG-FL*<br>Set 2 | Forward 5' CGTTTG**CAGTCGAC**TATTCAAG 3'<br>Reverse 3' ACC**CGCCGGCG**TTTTTATAACC 5' | SalI<br>NotI | pET22b | N-term pelB<br>C-term hexa-His | Amp[R] |
| *PmpH-FL* | Forward 5' AT**CATATG**CCCTTTTTTCTTTGAGATCTAC 3'<br>Reverse 3' CGAACTTATCTTAGAAAG**GAGCTC**AG 5' | NdeI<br>XhoI | pET22b | C-term hexaHis | Amp[R] |

**Table 2.1 Primers for the PCR amplification of serovar E pmp genes and predicted functional gene regions (continued)**

| Gene | Primer | Site | Vector | Fusion tag | Marker |
|---|---|---|---|---|---|
| **Pmpl-FL** | Forward 5' TA**GCTAGC**CGACCTGATCATATGAACTTC 3' | NheI | pET23b | C-term hexaHis | Amp^R |
| | Reverse 3' CCTGGTGAATGTCCAAT**CGCCGGCG**TA 5' | NotI | | | |
| **pmpA-N** | Forward 5' T**GCTAGC**AATCGAGTTATAGAAATCC 3' | NheI | pET23b | C-term hexaHis | Amp^R |
| | Reverse 3' CGTTTATCTTG**GAGCTC**AA 5' | XhoI | | | |
| **pmpD-N** | Forward 5' TC**CCATGG**GTTCCGAGAAAGATATAAAAAGC 3' | NcoI | pET28b | C-term hexHis | Kan^R |
| | Reverse 3'CGCATACCTT**CAGCT*A***ATAAGAT 5'<br>(*base error A, intended G* lost Sall site) | *Sall * | | | |
| **pmpH-N** | Forward 5' AT**CATATG**CCTTTTTCTTTGAGATCTAC 3' | NdeI | pET22b | C-term hexaHis | Amp^R |
| | Reverse 3' CGAGATCTGACACA**GAGCTC**CCCTC 5' | XhoI | | | |
| **pmpl-N** | Forward 5' TA**GCTAGC**CGACCTGATCATATGAACTTC 3' | NheI | pET23b | C-term hexaHis | Amp^R |
| | Reverse 3' GGGTGATGT**CGCCGGCG**AGCTTG 5' | NotI | | | |
| **pmpD-C** | Forward 5' **GCTAGC**CCATGGAATTCG 3' | NheI | pET28b | C-term hexaHis | Kan^R |
| | Reverse 3' GACCTAACGCTAACTAGAAA**CGCCGGCG**A 5' | NotI | | | |
| **pmpG-C** | Forward 5' GAAGCTTGC**GTCGAC**CTAATACAGCAAATATG 3' | Sall | pET22b | N-term pelB | Amp^R |
| | Reverse 3' CATAGCTCTATGA**GAGCTC**CAATACC 5' | XhoI | | C-term hexaHis | |

**Table 2.1 - Primers for the PCR amplification of serovar E pmp genes and predicted functional gene regions (continued).** A set of primers is displayed for each *pmp* in order to amplify the full-length gene, with the exception of full-length *pmpG-FL*. Also displayed are the primer sets designed to amplify truncated *pmps*, functional C and N terminal domains of *pmpA*, *pmpD*, *pmpG*, *pmpH* and *pmpl*. An error within the reverse primer sequence of the *pmpD-N* was later identified and is highlighted by the asterix. Incorporated restriction sites are highlighted in yellow. The commercial vectors, selection markers and intended fusion tags are also shown for each construct.

| PCR set-up conditions | | |
| --- | --- | --- |
| Reagent | Set-up 1 | Set-up 2 |
| KOD x10 Buffer | 5.0 µl | 5.0 µl |
| MgSO$_4$ | 1.0 mM | 1.5 mM |
| dNTPs | 0.2 mM | 0.2 mM |
| Forward primer | 0.2 µM | 0.3 µM |
| Reverse primer | 0.2 µM | 0.3 µM |
| Template DNA | ~1 ng | ~1 ng |
| KOD polymerase | 1.5 µl | 1.5 µl |
| Total volume | 50 µl | 50 µl |

**Table 2.2 - The two optimised PCR component set-up conditions used to amplify full-length and truncated *pmp* genes**. The set up condition used to amplify each particular gene is displayed in Table 2.4 denoted by 1 or 2

| Thermocycler Conditions | | | |
| --- | --- | --- | --- |
| | a | b | c |
| Initial Denaturation | 94°C - 2 min | 94°C - 2 min | 94°C - 2 min | |
| Denaturation | 94°C - 15 sec | 94°C - 15 sec | 94°C - 15 sec | |
| Annealing | 57°C - 30 sec | 54°C - 30 sec | 50°C - 30 sec | 30 cycles |
| Extension | 72°C - 4 min | 72°C - 4 min | 72°C – 2.5 min | |
| Final Extension | 72°C - 10 min | 72°C - 10 min | 72°C - 10 min | |
| | 4°C – hold | 4°C – hold | 4°C – hold | |

**Table 2.3 - The three optimised PCR thermocycler conditions used to amplify full-length and truncated *pmp* genes.** The set-up condition used to amplify each particular gene is displayed in Table 2.4 denoted by a or b or c using an MWG Biotech Primus 96 plus thermal cycler.

| Target DNA | PCR set-up conditions | PCR cycle conditions |
| --- | --- | --- |
| *pmpA* | 1 | a |
| *pmpB* | 2 | a |
| *pmpC* | 2 | a |
| *pmpD* | 1 or 2 | a |
| *pmpE* | 2 | c |
| *pmpF* | 1 | a |
| *pmpH* | 1 | a |
| *pmpI* | 1 | a |
| *pmpA-N* | 1 | b |
| *pmpD-N* | 1 or 2 | b |
| *pmpH-N* | 2 | b |
| *pmpI-N* | 1 | b |
| *pmpD-C* | 2 | a |
| *pmpG-C* | 1 | b |

**Table 2.4 - Shows optimum PCR reaction set-up conditions and PCR cycle conditions for the amplification of each target DNA**. See Table 2.2 and Table 2.3 for individual parameters.

## 2.2.4 Agarose gel electrophoresis

DNA samples were prepared with 6X DNA sample buffer (40% (w/v) sucrose in sterile water containing 0.1% bromophenol blue) and separated in a 1% agarose slab gel with TAE buffer (40 mM Tris-acetate, pH 8.0, 1 mM EDTA) and ethidium bromide (0.5 µg/ml). The molecular size and quantification of DNA was carried out by visual comparison of the bands under UV illumination (images captured using a UVP epi II dark room with Labworks software version 4.0) to DNA markers (2-log DNA ladder, 1kb DNA ladder or 100 bp DNA ladder, New England Biolabs) which were separated alongside the samples.

## 2.2.5 Purification of PCR products

PCR reactions were purified using the QIAGEN gel extraction purification kit, where specific amplification of *pmps* was observed. This was done according to the manufacturer's instructions with the exception of the gel excision steps 1-4 which were omitted (Qiagen).

## 2.2.6 Restriction digestion of PCR products and plasmid DNA

Purified PCR products and commercial pET plasmid DNA (100 – 500 ng) were digested with the appropriate restriction endonuclease (shown in **Table 2.1**). The reactions were made with (0.5 – 2 U) endonuclease in the manufacturer's buffer according to their instructions, supplied by New England Biolabs. Endonuclease digestion of plasmid DNA was confirmed by agarose gel electrophoresis.

## 2.2.7 Purification of restriction digested DNA fragments

Digested DNA was mixed with DNA sample buffer and separated by agarose gel electrophoresis. DNA fragments were excised using a sterile scalpel and weighed. The DNA was purified using the QIAGEN gel extraction purification kit according to the manufacturer's instructions (Qiagen). The absorbance of 1 µl of purified DNA product was measured using Thermo Scientific NanoDrop® ND-1000 Spectrophotometer to quantify the DNA.

## 2.2.8 Dephosphorylation of linearised plasmid DNA

To minimise self ligation, linearised vector DNA was routinely dephosphorylated with calf intestine alkaline phosphatase (CIP) before the protein coding sequence was inserted. CIP (1 U) was added to approximately 5 µg DNA in CIP reaction buffer (CIP and CIP reaction buffer supplied by New England Biolabs, and diluted as directed). The mixture was incubated at 37° C for 30 minutes, before arresting

the action of the enzyme by heating to 75° C for 10 minutes or immediately purifying the DNA (**section 2.2.5**).

## 2.2.9 Ligation of DNA fragments

Vector and insert DNA were ligated using T4 DNA ligase which catalyses the formation of a phosphodiester bond between the 5' phosphate and the 3' hydroxyl termini. To 30 ng dephosphorylated vector DNA, a 1 to 3-fold molar excess of insert DNA was mixed with 1 U T4 DNA ligase (New England Biolabs) in the corresponding T4 DNA ligase buffer. Dependent upon the size of the insert DNA the reaction was incubated between 1 and 16 hours at room temperature, 16° C or 8° C. Ligation was assessed by agarose gel electrophoresis. The ligated DNA was transformed into highly efficient NEB 10-beta competent *E.coli* host cells (New England Biolabs) as described **section 2.2.10**.

## 2.2.10 Transformation of cells

A 20 µl aliquot of competent NEB 10-beta *E.coli* cells (**section 2.2.11**) was thawed on ice, and 20 – 200 ng plasmid DNA added and gently stirred. The cells were incubated for 30 minutes on the ice and then heat shocked at 42° C for 45 seconds. 500 µl SOC outgrowth media (2 % Vegetable Peptone, 0.5 % Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM Glucose), was added to the cells and the suspension incubated at 37° C for 1 hour with agitation. All of the cell suspension was spread onto a pre-warmed LB agar plate (Tryptone 10g/L, Yeast extract 5g/L, NaCl 5g/L, agar 15 g/L containing the appropriate antibiotic (ampicillin 100 µg/ml or kanamycin 34 µg/ml)). The plate was incubated overnight at 37° C for colony propagation.

## 2.2.11 Genotype of competent E.coli

**NEB 10-beta:** *araD139* Δ(*ara-leu*)7697 *fhuA lacX74 galK* (φ80 Δ(*lacZ*)M15) *mcrA galU recA1 endA1 nupG rpsL* (Str$^R$) Δ(*mrr-hsdRMS-mcrBC*) (New England Biolabs)

## 2.2.12 Selective colony propagation

Single bacterial colonies were selected and used to inoculate 5 ml Luria-Bertani broth (LB, 1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract, pH 7) containing the appropriate antibiotic in a 50 ml Falcon tube, and incubated overnight at 37° C in an orbital incubator at 200 rpm. The culture was subsequently centrifuged at 5000 g for 10 minutes, and plasmid extracted as in **section 2.2.13.**

## 2.2.13 Isolation of plasmid DNA

The plasmid was transformed into competent *E.coli* as described **section 2.2.10**. The cell pellet was lysed and the plasmid DNA was extracted using QIAGEN mini-prep kit according to the manufacturer's instructions (Qiagen). For extracting high concentrations of plasmid DNA <500 µg/ml, the Purelink™ HiPure plasmid maxiprep kit was used according to manufacturer's instructions (Invitrogen).

## 2.2.14 Restriction Mapping

Restriction sites predicted to be present within the recombinant plasmids were mapped using restriction endonuclease digestion as in **section 2.2.6**. Both single and double restriction digests were carried out. Using the sizes of the DNA fragments after digestion, the positions of these sites could be confirmed. Restriction mapping was used to determine the orientation and size of the insert in the cloning vector. Putative clones identified by mapping were sent for DNA sequencing as in **section 2.2.15** to confirm the insert was correct and in frame.

## 2.2.15 Automated DNA sequencing

DNA sequence analysis was carried out by Eurofins MWG based on *Taq* dyedeoxy™ terminator cycle sequencing with the use of T7 promoter and terminator region primers flanking both 3'-5' and 5'-3' sequences (Eurofins MWG Operon).

## 2.2.16 Site Directed Mutagenesis

Site-directed mutagenesis was performed using PCR and the Quikchange™ Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer instructions. Primers specific for the desired mutation were individually designed (**Table 2.5**). The QuikChange site-directed mutagenesis method used the high fidelity PfuTurbo™ DNA polymerase, the PCR conditions were adapted for each mutagenesis based upon the length of the DNA sequence as in **Table 2.6**.

| Construct | Primers 5' – 3' |
|---|---|
| **pPmpA23b-N** | Forward - GTGGTGGTGGTGGTGCTACGAGGTTCTATTTGC |
| | Reverse - GCAAATAGAACCTCGATGCACCACCACCACCAC |
| **pPmpl23b-N** | Forward – ACAGCGGCCGCAACTCGAGCACCAC |
| | Reverse – GTGGTGCTCGAGTTGCGGCCGCTGT |
| **pPmpG22b-C** | Forward – CCGTCGACCCTAATACAGCAAATAATGGTCCTTATACTC |
| | Reverse - GAGTATAAGGACCATTATTTGCTGTATTAGGGTCGACGG |

**Table 2.5 - Primers for site-directed mutagenesis.** The primers designed to correct the reading frame for each of the truncated *pmps* were used with the thermocycler conditions in Table 2.6.

| Thermocycler Conditions | | | | |
|---|---|---|---|---|
| | **pPmpA23b-N** | **pPmpG22b-C** | **pPmpl23b-N** | |
| Initial Denaturation | 95˚C - 1 min | 95˚C - 1 min | 95˚C - 1 min | |
| Denaturation | 95˚C - 50 sec | 95˚C - 50 sec | 95˚C - 50 sec | 18 cycles |
| Annealing | 60˚C - 50 sec | 60˚C - 50 sec | 60˚C - 50 sec | |
| Extension (1 min/ kb) | 68˚C – 5.5 min | 68˚C - 7 min | 68˚C - 5 min | |
| Final Extension | 68˚C - 7 min | 68˚C - 7 min | 68˚C - 7 min | |
| | 4˚C – hold | 4˚C – hold | 4˚C – hold | |

**Table 2.6 - PCR conditions for Site-Directed Mutagenesis**. Shows the cycle conditions for each of the truncated *pmp* constructs (*pmpA-N, pmpG-C and pmpl-N*) that required site-directed mutagenesis using the primers in Table 2.5 to correct the reading frame of the expression construct.

## 2.3 Results

### 2.3.1 PCR of full length pmp genes (pmp-FL)

Numerous trials altering the components and/or the thermocycler conditions were carried out to obtain optimal amplification for all nine *pmp* genes and truncated *pmp* genes. Changes in annealing temperature based on the $T_m$ of the primers were essential to obtain amplification of certain *pmps*. Variations in the primer concentration and $MgSO_4$ aided the amplification of specific *pmps* and truncated *pmps*. These variables resulted in two different component set-up conditions and three different cycle conditions as shown in **Table 2.2** and **Table 2.3**. PCR amplifications of the *pmp* gene were achieved for 8 of the 9 full length DNA targets. The exception was *pmpG-FL* where specific PCR products of the expected size could not be amplified (data not shown) with either of the two primers sets in combination with several alterations to the PCR conditions and components.

The agarose gel image in **Figure 2.1** shows the amplification of putative *pmpB-FL* and *pmpC-FL*. The predicted size of serovar D *pmpB* is ~5250 bp and *pmp C* is ~5300 bp. The obtained sizes of amplified *pmpB-FL* and *pmpC-FL* were approximately 5300 bp indicated by red arrows. Additionally non-specific amplifications were observed with *pmpB-FL*, *pmpC-FL* visible at approximately 1600 bp, 2500 bp and 3000 bp To minimise the possibility of these being incorporated into the vector the PCR products of the desired size were isolated from the gel prior to purification.

**Figure 2.2** shows the amplification of *pmpD-FL*, the predicted size of serovar D *pmpD is* ~4592 bp. The obtained size of *pmpD-FL* was approximately 4600 bp indicated by red arrow. The non-specific amplification product visible at approximately 3500 bp was omitted from ligation into a vector by excising *pmpD-FL* from the gel for purification.

Finally the remainder of the amplified full length *pmps* are displayed in **Figure 2.3**. The predicted sizes of serovar D *pmps* are: *pmpA* ~2920 bp, *pmp E* ~2980 bp, *pmp F* ~3100, *pmp H* ~ 3000 and *pmp I* ~ 2630 bp. The obtained sizes of *pmpA-FL, pmpE-FL, pmpF-FL* and *pmpH-FL* were all approximately 3000 bp and *pmpI-FL* was approximately 2600 bp indicated by the red arrows.
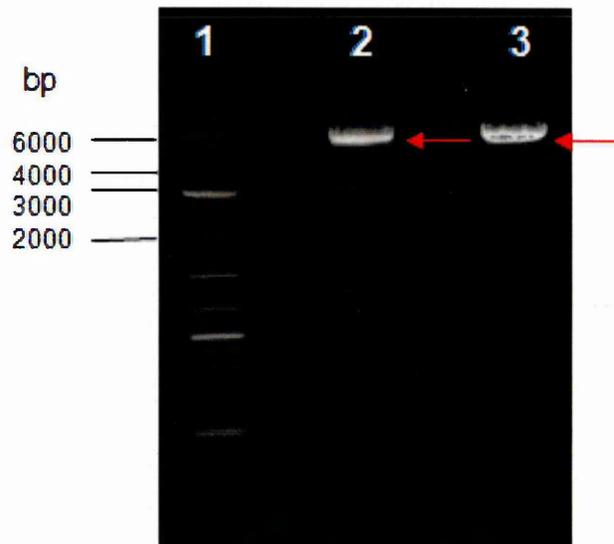
**Figure 2.1 - Agarose gel showing the amplification of *pmpB-FL* and *pmpC-FL* obtained by PCR**. The gel shows the full length *pmpB* and *pmpC* (red arrows) amplified from genomic DNA using primers shown in Table 2.1. Lanes 2 and 3 contained 10μl of the PCR products *pmpB-FL* and *pmpC-FL* respectively, obtained using optimised reaction set-up conditions and thermocycler conditions as shown in Table 2.4. Lane 1 contains 2-log DNA marker (New England Biolabs) with sizes indicated on the left in base pairs.
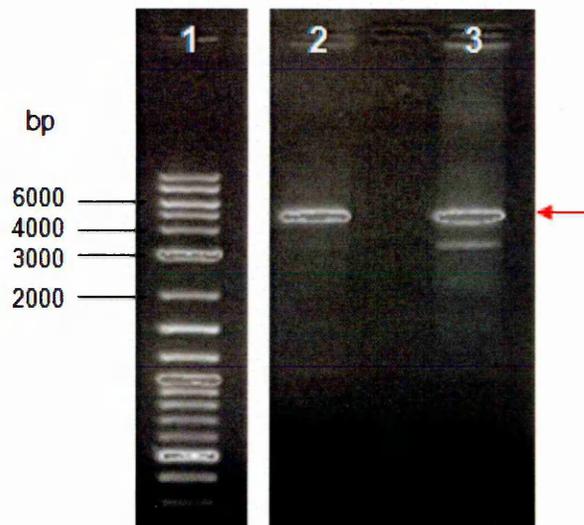


**Figure 2.2 - Agarose gel showing the amplification of *pmpD-FL* obtained by PCR**. The gel shows *the full length pmpD* amplified from genomic DNA using primers shown in Table 2.1 and cycle conditions displayed in Table 2.3. Lanes 2 and 3 contained 10μl of the PCR products obtained using both reaction set-up conditions 1 and 2 respectively, shown in Table 2.2. Lane 1 contains 2-log DNA marker (New England Biolabs) with sizes indicated on the left in base pairs. The cut in the gel is to remove blank lanes.

69
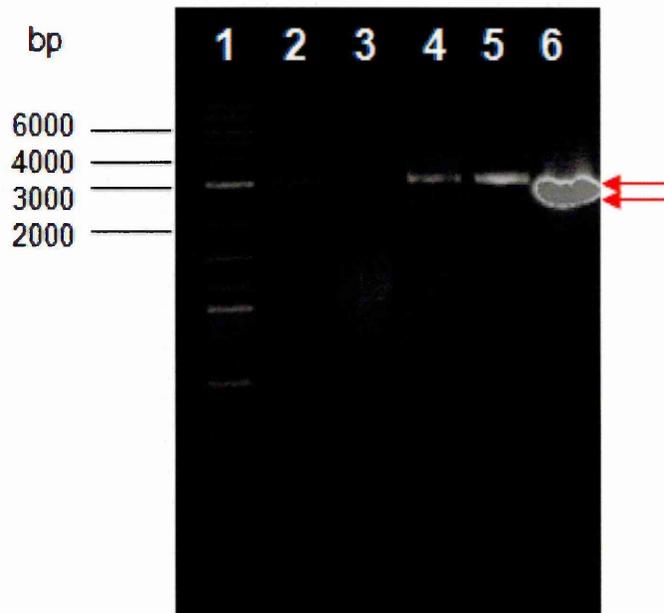
**Figure 2.3 - Agarose gel showing the amplification of *pmpA-FL, pmpE-FL, pmpF-FL* and *pmp I* obtained by PCR.** The gel shows the full length *pmpA, pmpE, pmpF and pmpI* amplified from genomic DNA using primers shown in Table 2.1 and cycle conditions displayed in Table 2.3. All lanes contained 10μl of PCR product obtained using reaction set-up conditions shown in Table 2.2. Lane 2 shows *pmpA-FL*, lane 3 shows *pmpE-FL*, lane 4 *pmpF-FL,* lane 5 *pmpH-FL* and lane 6 contained *pmpI-FL* Lane 1 contains 2-log DNA marker (New England Biolabs) with sizes indicated on the left in base pairs.

70

## 2.3.2 Construction of full length pmps expression vectors with pET vectors pET23b(+), pET22b(+) and pET28b(+)

Purified, digested vectors and amplified PCR products were ligated and transformed into *E.coli* strain NEB 10-beta (New England Biolabs) for plasmid propagation as described in **section 2.2.12**. The restriction sites and pET vector for each construct are listed in **Table 2.1**.

Transformant colonies were propagated and plasmid DNA isolated. These plasmids were analysed by gel electrophoresis alongside empty pET vector controls. Putative clones that were of a size expectant upon uptake of the *pmp* insert were subjected to both single and double restriction digestion mapping at specific recognised site, these restriction sites were determed from serovar D genomic sequences as close similarity was expected, to produce a map of fragments that were distinct to confirm the orientation and presence of the *pmp* insert DNA. The restriction maps of putative clones containing *pmpB-FL, pmpC-FL, pmpE-FL, pmpF-FL* and pmp*H-FL* did not produce the expected insert size or fragment sizes. The fragments produced using restriction digestion were indicative of a truncated insert, an incorrect insert or an insert that was inserted in the wrong direction, sequencing of these inserts confirmed the coding sequences were unsuitable for expression (data not shown). **Figure 2.4a**, **Figure 2.5a** and **Figure 2.6a** show the putative recombinant clones containing *pmpA, pmpD* and *pmpI* respectively. The fragment sizes obtained from restriction mapping of each of these putative clones suggested the coding sequence inserted in each plasmid was correct (**Figure 2.4b**, **Figure 2.5b** and **Figure 2.6b**). These constructs were named pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL, to indicate the *pmp* insert and the pET vector.

Automated DNA sequencing verified that the coding regions for full length PmpA, PmpD and PmpI were present, correct and in-frame with a C-terminal hexahistidine tag. The protein encoded by the DNA sequence within the N- and C- terminal regions showed differences to the published serovar D as a consequence of cloning procedure is displayed in **Table 2.7**. The remainder of the serovar E Pmp sequences within these regions were the same as published serovar D sequences as far as the sequencing reaction was performed. The automated sequencing reaction determines the sequence up to 1000 nucleic acids before it is exhausted, therefore only the N and C terminal regions of these inserts were determined for the purposes of verifying the insert was correct.
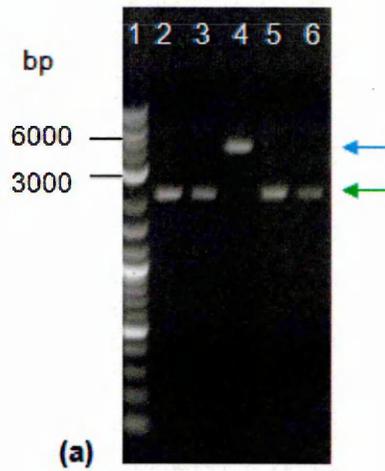
| Restriction Sites | Expected Fragments (bp) |
|---|---|
| NheI/NotI | ~3600, ~3000 |
| NotI | ~6600 |
| XbaI | ~5300, ~1300 |
| XbaI/XhoI | ~3600, ~1700, ~1300 |
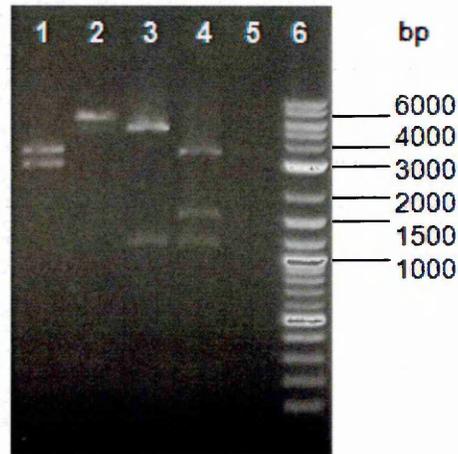
**Figure 2.4 - Agarose gel showing recombinant plasmid pPmpA23b-FL miniprep preparation and restriction mapping characterisation.** (a) The gel shows isolated plasmids from *E.coli* clones transformed with ligated pET23b and *pmpA-FL* extracted as in section 2.2.13. Lane 6 contains undigested empty control plasmid pET23b. Lanes 2, 3 and 5 show putative supercoiled non-recombinant vector, indicated by green arrow. Lane 4 shows putative supercoiled recombinant vector, indicated by the blue arrow. Lane 1 contains a 2-log DNA marker (New England Biolabs) with sizes indicated on the left in base pairs. Comparison of lanes and empty vector gives indication of successful ligation. (b) The single clone from lane 4, panel (a) was subjected to single and double digestion with the restriction enzymes *NheI/NotI* (lane 1), *NotI* (lane 2), *XbaI* (lane 3) and *XbaI/XhoI* (lane 4). Lane 5 contained linear pET23b at 3666 bp, digested with *XhoI* while lane 6 contained DNA ladder.
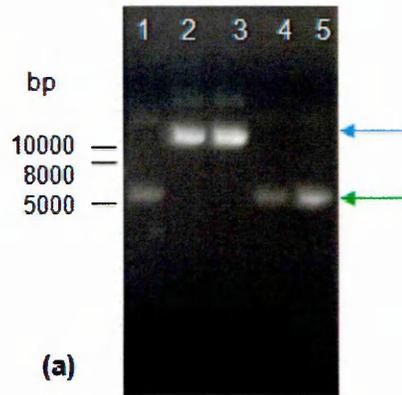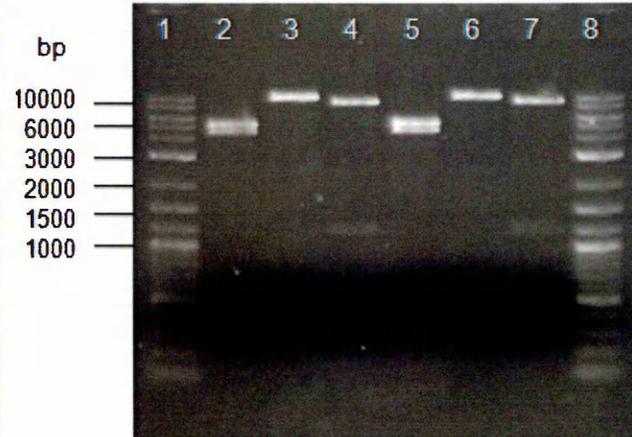
**Figure 2.5 - Agarose gel showing recombinant plasmid pPmpD28b-FL miniprep preparations and restriction mapping characterisation.** (a). The gel shows isolated plasmids from *E.coli* transformed with ligated pET28b and *pmpD-FL* extracted as in section 2.2.13. Lane 5 contains undigested empty control plasmid pET28b. Lanes 1 and 4 show supercoiled non-recombinant vectors, indicated by green arrow. Lanes 2 and 3 show putative supercoiled recombinant vector, indicated in line with the blue arrow. Comparison of lanes and empty vector gives indication of successful ligation. (b) The two clones from lanes 2 and 3 in gel (a), were subjected to single and double digestion with the restriction enzymes *NcoI/NotI* (lanes 2 and 5) or *NcoI* (lanes 3 and 6) or *BamHI* (lanes 4 and 7). Lanes 1 and 8 contained 2-log DNA ladder.
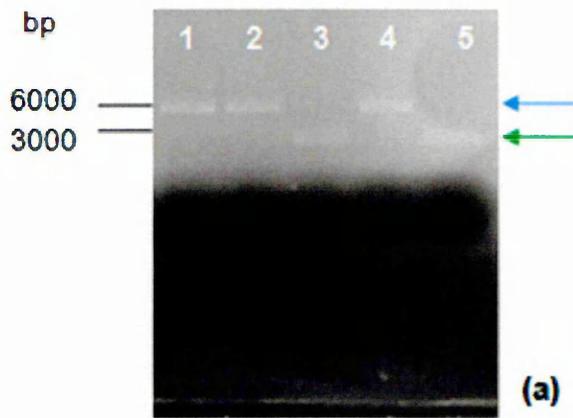
| Restriction Sites | Expected Fragments (bp) |
|---|---|
| XhoI | ~4900, ~1400 |
| NheI/NotI | ~3600, ~2600 |
| NotI/XbaI | ~3500, ~2100, ~600 |
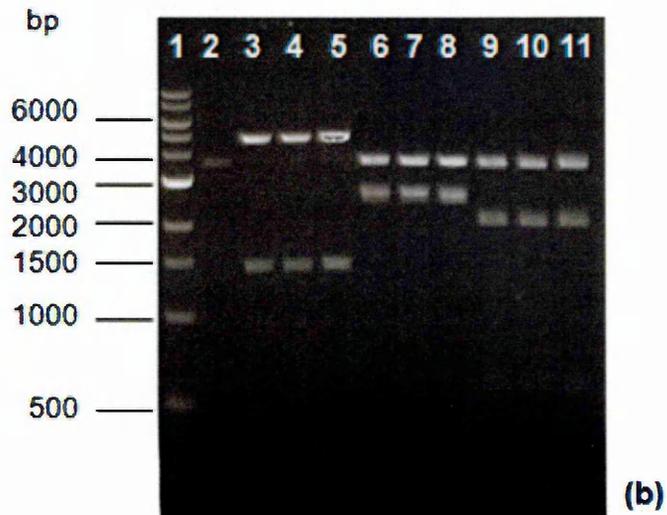
**Figure 2.6 - Agarose gels showing recombinant plasmid pPmpI23b-FL miniprep preparation and restriction mapping characterisation.** (a) The gel shows isolated plasmids from *E.coli* transformed with ligated pET23b and *pmpI*. Lane 5 contains undigested empty control plasmid pET23b at ~3500 bp. Lane 3 shows supercoiled non-recombinant vector at ~3500 bp, indicated by green arrow. Lanes 1, 2 and 4 show putative supercoiled recombinant vectors, indicated by the blue arrow. Comparison of lanes and empty vector gives indication of successful ligation. (b) The three clones from lanes 1, 2 and 4 in gel (a) above were subjected to single and double digestion with the restriction enzymes *XhoI* (lanes 3-5) or *NheI/NotI* (lanes 6-8) or *NotI/XbaI* (lanes 9-11). Lane 2 contained linear pET23b digested with XhoI at 3666pb. Lane 1 contained Quick-load 100 bp DNA ladder (New England Biolabs).

| PmpA | recombinant N-terminus; | `MASNRVIEIHAHYDQRQLS-` |
| | published N-terminus; | `M--NRVIEIHAHYDQRQLS-` |
| | | `*..**************` |
| | recombinant C-terminus; | `-SNSLSCGGYVGLAAALEHHHHHHH` |
| | published C-terminus; | `-SNSLSCGGYVGE` |
| | | `***********...........` |
| PmpD | recombinant N-terminus; | `MGSEKDIKSTCSKFSL-` |
| | published N-terminus; | `MSSEKDIKSTCSKFSL-` |
| | | `* *************` |
| | recombinant C-terminus; | `-KLGYEANAGLRLIFAAALEHHHHHH` |
| | published C-terminus; | `-KLGYEANTGLRLIF` |
| | | `******.******..........` |
| PmpI | recombinant N-terminus; | `MASRPDHMNFCCLCAA-` |
| | published N-terminus; | `M--RPDHMNFCCLCAA-` |
| | | `*..*************` |
| | recombinant C-terminus; | `-SYSLDLGTTYRLAAALEHHHHHH` |
| | published C-terminus; | `-SYSLDLGTTYRF` |
| | | `***********..........` |

**Table 2.7 - Sequence variation of the C and N terminals of the serovar E clinical recombinant full length Pmps compared to the published serovar D sequences.** Sequnce matches are highlighted by (`*`) and recombinant changes as a result of the cloning procedure are highlighted by (`.`).

### 2.3.3 PCR of truncated pmp genes

The predicted functional domains of some of the Pmp protein sequences (Pmp A, Pmp D, Pmp G, Pmp H and Pmp I) were determined using the protein BLAST analysis tool (http://blast.ncbi.nlm.nih.gov), based upon this the primers were designed to flank each end of the domain as in **section 2.2.2.** The agarose gel images in **Figure 2.7** show the amplification of putative truncated *pmpA-N* (Pmp A passenger domain), *PmpH-N* (Pmp H passenger domain) and *PmpI-N* (Pmp I passenger domain). The predicted approximate sizes of these passenger domain sequences are ~1450 bp for *pmp A*, ~1500 bp for *pmp H* and ~1220 bp for *pmp I*. The fragments obtained were of the expected sizes. **Figure 2.8** shows the amplification of putative *PmpD-N* (Pmp D passenger domain) and *pmpG-C* (Pmp G autotranslocator domain). The predicted approximate size of these coding regions are ~3500 bp and ~1100 bp respectively, the obtained size for *pmpD-N* was between 3000 and 4000 bp as indicated by the red arrow and the obtained product for *pmpG-C* was ~1100 bp, indicated by the purple arrow. Finally, **Figure 2.9** shows amplification of *PmpD-C* (Pmp D autotranslocator domain) the predicted size of the autotranslocator region of *pmpD* is ~1000 bp. The fragment obtained for *pmpD-C* is approximately 1000 bp indicated by red arrow
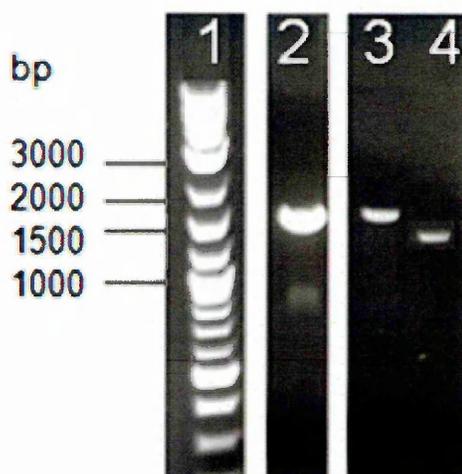
**Figure 2.7 - Agarose gel showing the amplification of** *pmpA-N,* *PmpH-N* **and** *PmpI-N* **obtained by PCR.** The gel shows truncated N-terminal predicted passenger regions of *pmpA, pmpH and pmpI* amplified from genomic DNA using primers shown in Table 2.1 and cycle conditions displayed in Table 2.3. All lanes contained 10μl of PCR product obtained using reaction set up conditions shown in Table 2.2. Lane 2 shows *pmpA-N*, lane 3 shows *pmpH-N*, and lane 4 shows *pmpI-N*. Lane 1 contains a 2-log DNA marker.
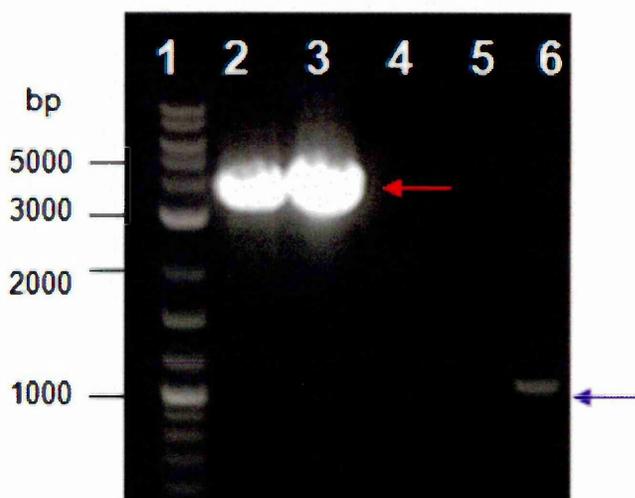


**Figure 2.8 - Agarose gel showing the amplification of** *PmpD-N* **and** *PmpG-C* **obtained by PCR.** The gel shows the truncated N-terminal predicted passenger region of *pmp D* and predicted C-terminal autotransporter coding region of *pmp G* amplified from genomic DNA using primers shown in Table 2.1 and cycle conditions displayed in Table 2.3. All lanes contained 10μl of PCR product obtained using reaction set up conditions shown in Table 2.2. Lanes 2 and 3 show N-term *pmpD* amplified using set up conditions 1 and 2 respectively. Lanes 4 and 5 were empty. Lane 6 shows C-term *pmpG* indicated by the purple arrow. Lane 1 contains 2-log DNA marker with sizes indicated on the left in base pairs.
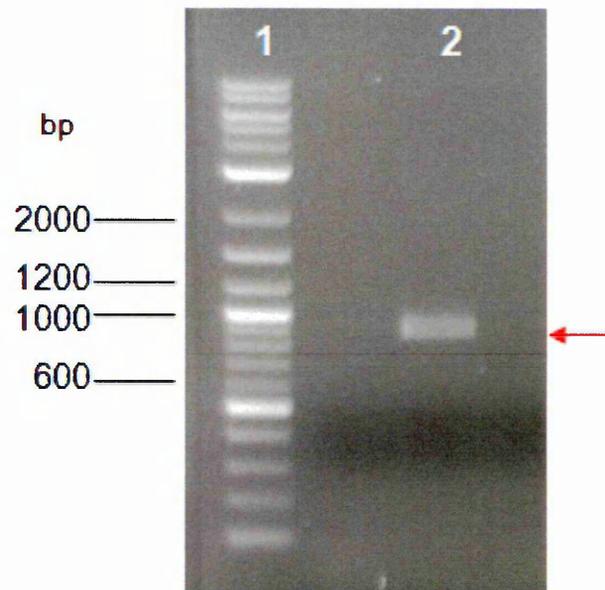
**Figure 2.9 - Agarose gel showing the amplification of *pmpD-C* obtained by PCR.** The gel shows the predicted C-terminal autotransporter region of *pmpD* (red arrow) amplified from genomic DNA using primers shown in Table 2.1 and cycle conditions displayed in Table 2.3. Lane 2 contained 10μl of the PCR product obtained using reaction set up conditions 2, shown in Table 2.2. Lane 1 contains 2-log DNA marker with sizes indicated on the left in base pairs.

## 2.3.4 Construction of truncated pmp expression vectors with pET vectors pET23b(+), pET22b(+) and pET28b(+)

Purified, digested vectors and amplified PCR products were ligated and transformed into *E.coli* strain NEB 10-beta (New England Biolabs) for plasmid propagation as described earlier for full length *pmp* expression constructs. Transformant colonies were propagated and the plasmid DNA isolated was analysed by gel electrophoresis alongside empty pET vector controls. Putative clones were subjected to both single and double restriction digestions at specific recognised sites derived from serovar D genomic sequences as close similarity was expected. The map of fragments (**Figure 2.10**, **Figure 2.11**, **Figure 2.12** and **Figure 2.13**) confirmed the orientation and presence of the truncated *pmp* insert DNA for PmpA-N, PmpD-C, PmpG-C and PmpI-N these constructs were named pPmpA23b-N, pPmpD28b-C, pPmpG22b-C and pPmpI23b-N to indicate the truncated *pmp* insert and the pET vector. However the sequencing of three of these inserts (*pmpA-N*, *pmpG-C* and *PmpI-N*) were not in the correct open reading frame and required further manipulation using site-directed mutagenesis to correct the frameshifts. The sequencing of the PmpD autotranslocator domain was identical to serovar D autotranslocator domain (**Figure 2.16**).

The restriction maps of putative clones containing the *pmp* passenger domain sequences *pmpD-N* and pmp*H-N* did not produce the expected insert size or fragment sizes. Restriction mapping analysis showed some of the expected restriction sites were not present within the construct containing the passenger domain sequence of *pmpD*, in particular some sites that were expected to be located within the last 600 - 700 bases of sequence were absent. Sequencing analysis revealed this region to be missing from the construct (data not shown) resulting in the loss of approximately 200 amino acids from the Pmp D passenger domain. The cloning procedure was repeated a further five times with new amplifications of *pmpD-N* and newly digested pET vector however the outcome remained the same. Numerous efforts were made to ligate the passenger domain sequence of *pmpH* (*pmpH-N*) into digested pET22b but only a small percentage of clones appeared to contain the vector with an insert. Restriction mapping analysis showed the expected restriction sites were not present within the construct (data not shown) therefore the insert was unlikely to contain *pmpH-N*.
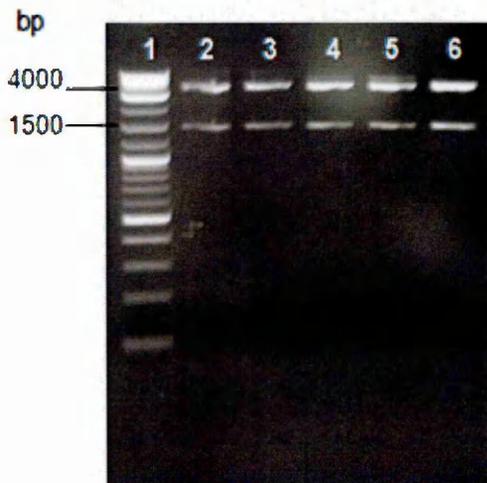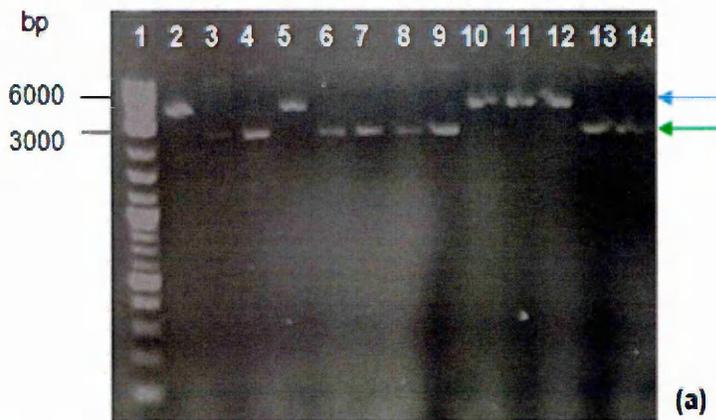
**Figure 2.10 - Agarose gel showing putative recombinant plasmid pPmpA23b-N miniprep preparation and restriction mapping characterisation**. (a) The gel shows isolated plasmids from *E.coli* transformed with ligated pET23b and *pmpA-N* extracted as in section 2.2.13. Lane 14 contains undigested empty control plasmid pET23b. Lanes 3, 4, 6, 7, 8, 9 and 13 show supercoiled non-recombinant vector, indicated by blue arrow. Lanes 2, 5, 10, 11 and 12 show putative supercoiled recombinant vector, indicated by the green arrow. Lane 1 contains 2-log DNA marker with sizes indicated on the left in base pairs. Comparison of lanes and empty vector gives indication of successful ligation. (b) The 5 clones from lanes 2, 5, 10, 11 and 12 were subjected to a double digestion with the restriction enzymes *Nhel/Xhol* to check the size of the inserts. All fragments were of expected size. The clone in lane 6 was sent for DNA sequencing for verification.

| Restriction Sites | Expected Fragments (bp) |
|---|---|
| XhoI | ~6000, ~500 |
| BamHI/ EcoRV | ~5000, ~1500 |
| NcoI | ~6500 |
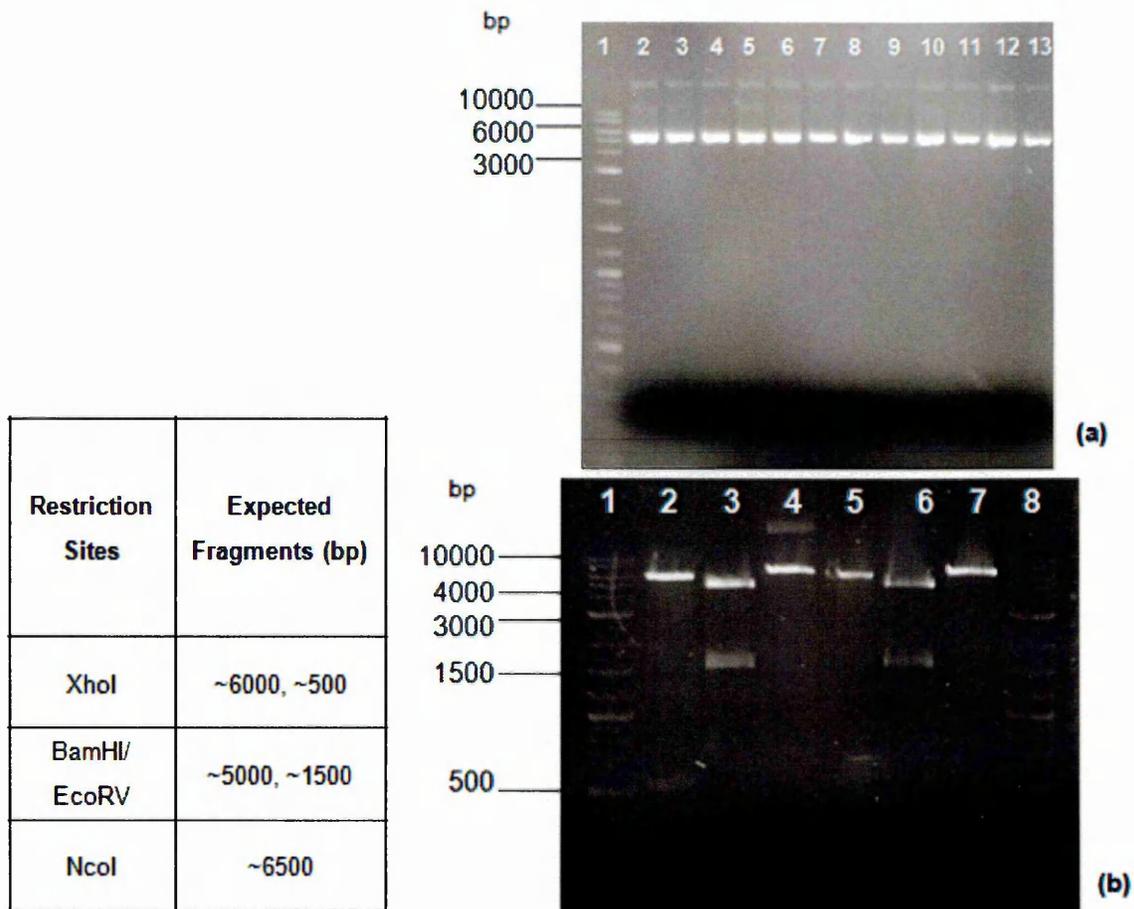
**Figure 2.11 - Agarose gel showing recombinant plasmid pPmpD28b-C miniprep preparations and restriction mapping characterisation**. (a). The gel shows isolated plasmids from *E.coli* transformed with ligated pET28b and *pmp D* autotranslocator domain sequence, extracted as in section 2.2.13. Lanes 2-13 show putative successful supercoiled recombinant plasmids. The expected size is approximately 6500 bp. Lane 1 contains 2-log DNA ladder. (b) Two clones, lanes 2 and 8 from gel (a) were subjected to single and double digestion with the restriction enzymes *XhoI* (lanes 2 and 5), *BamHI/EcoRV* (lanes 3 and 6) and *NcoI* (lanes 4 and 7). Lane 8 contained 2-log DNA ladder.

81

| Restriction Sites | Expected Fragment Sizes (bp) |
|---|---|
| XhoI/SalI | ~5500, ~1100 |
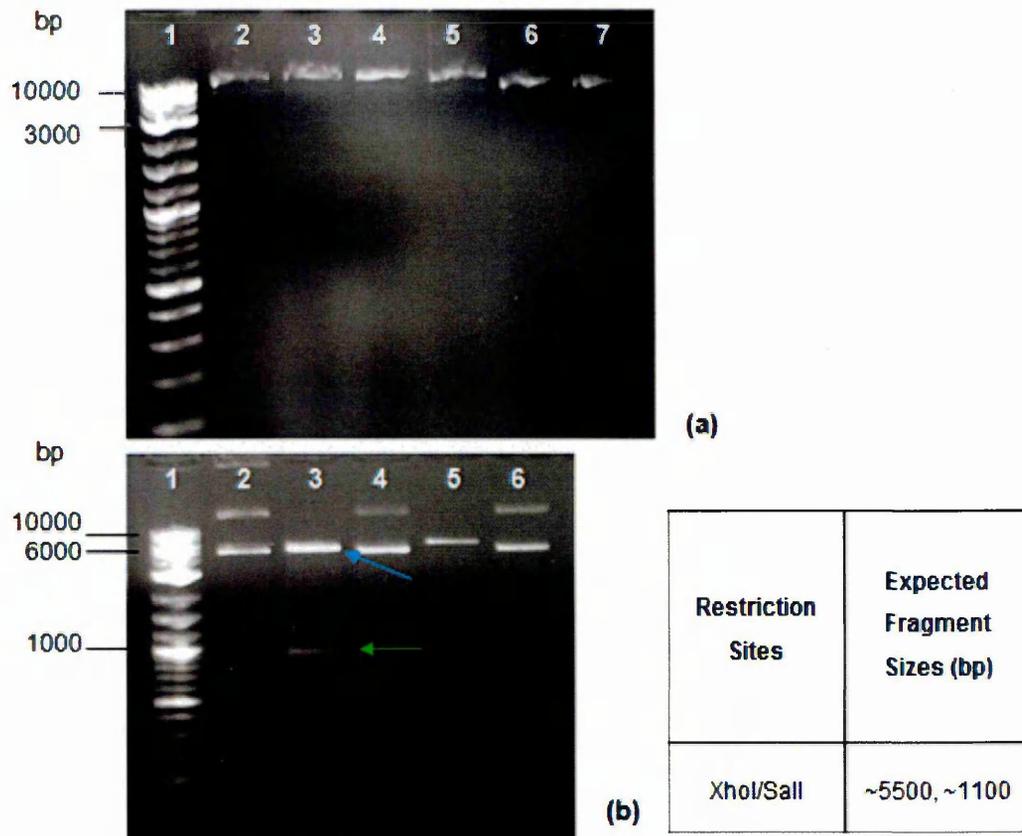
**Figure 2.12 - Agarose gel showing recombinant plasmid pPmpG22b-C miniprep preparations and restriction mapping characterisation.** (a) The gel shows isolated plasmids from *E.coli* transformed with ligated pET22b and *pmp G* autotranslocator domain extracted as in section 2.2.13. Lanes 2-6 show putative successful supercoiled recombinant vector. Lane 7 contains supercoiled pET22b. Lane 1 contains 2-log DNA ladder. (b) All clones from gel (a) were subjected to double digestion with the restriction enzymes *XhoI* and *SalI* to examine the size of the inserts. The fragments of the clone in lane 3 were of expected size with an insert of approximately 1100 bp (green arrow) and linearised pET22b at ~5700 bp (blue arrow), this clone was sent for sequencing. Lane 1 contained 2-log DNA ladder.

| Restriction Sites | Expected Fragments (bp) |
|---|---|
| NheI/notI | ~3600, ~1100 |
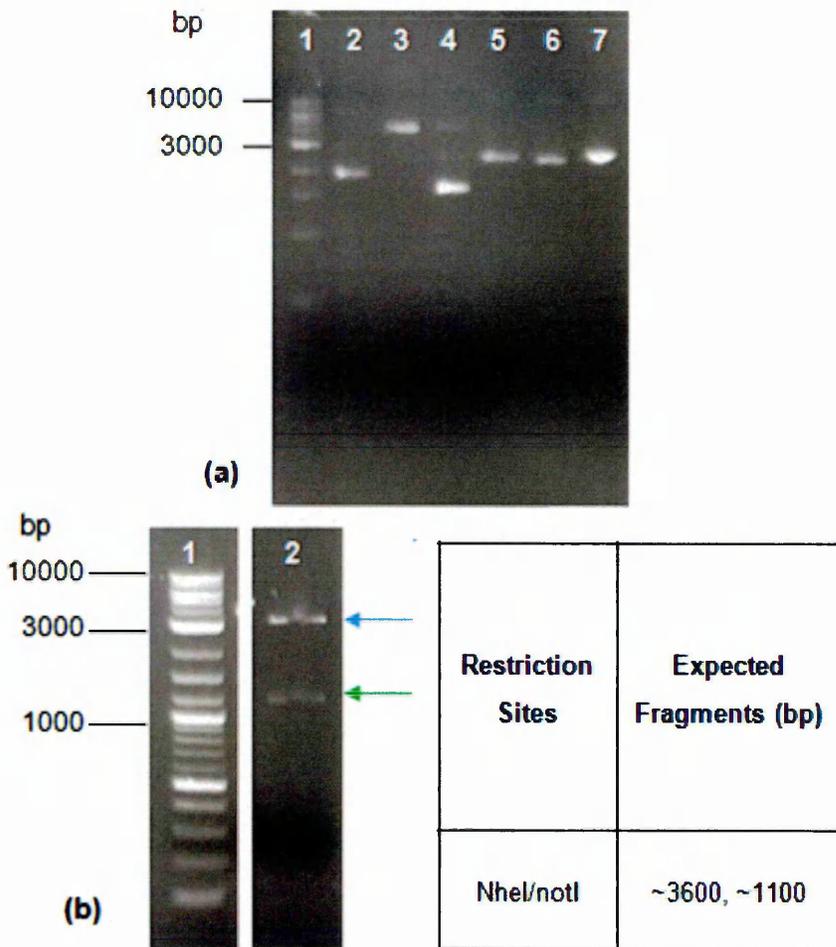
**Figure 2.13 - Agarose gel showing recombinant plasmid pPmpl23b-N miniprep preparations and restriction mapping characterisation**. (a) The gel shows isolated plasmids from *E.coli* transformed with ligated pET23b and *the pmp I* passenger domain sequence (*pmpl-N*), extracted as in section 2.2.13. Lane 3 shows putative successful supercoiled recombinant vector. Lanes 2, 4, 5 and 6 show supercoiled non-recombinant vector. Lane 7 contains supercoiled pET23b. Comparison of lanes and empty vector gives indication of successful ligation. Lane 1 contains 100bp DNA ladder (New England Biolabs). (b) The putative clone in lane 3 from gel (a) was subjected to a double digestion with the restriction enzymes *NheI* and *NotI* to cut either side of the size of the insert. The fragments were of expected size with an insert of approximately 1100 bp (green arrow) linearised pET23b (blue arrow). Lane 1 contained 2-log DNA ladder.

```
M E F D Y S T N V W G F A F G G F R T L S A E N L V A I D G Y K G
M E F D Y S T N V W G F A F G G F R T L S A E N L V A I D G Y K G


A Y G G A S A G V D I Q L M E D F V L G V S G A A F L G K M D S Q
A Y G G A S A G V D I Q L M E D F V L G V S G A A F L G K M D S Q


K F D A E V S R K G V V G S V Y T G F L A G S W F F K G Q Y S L G
K F D A E V S R K G V V G S V Y T G F L A G S W F F K G Q Y S L G


E T Q N D M K T R Y G V L G E S S A S W T S R G V L A D A L V E Y
E T Q N D M K T R Y G V L G E S S A S W T S R G V L A D A L V E Y


R S L V G P V R P T F Y A L H F N P Y V E V S Y A S M K F P G F T
R S L V G P V R P T F Y A L H F N P Y V E V S Y A S M K F P G F T


E Q G R E A R S F E D A S L T N I T I P L G M K F E L A F I K G Q
E Q G R E A R S F E D A S L T N I T I P L G M K F E L A F I K G Q


F S E V N S L G I S Y A W E A Y R K V E G G A V Q L L E A G F D W
F S E V N S L G I S Y A W E A Y R K V E G G A V Q L L E A G F D W


E G A P M D L P R Q E L R V A L E N N T E W S S Y F S T V L G L T
E G A P M D L P R Q E L R V A L E N N T E W S S Y F S T V L G L T


A F C G G F T S T D S K L G Y E A N A G L R L I F A A A L E H H H
A F C G G F T S T D S K L G Y E A N T G L R L I F Stop
```

H H H Stop

**Figure 2.14 - Sequence of *C.trachomatis* serovar E recombinant C-terminal Pmp D autotranslocator domain compared to published serovar D.** Recombinant Pmp D autotranslocator sequence is shown in black, published serovar D Pmp D sequence is aligned in red with the recombinant changes underlined.

## 2.3.4.1 Site-Directed mutagenesis of pPmpA23b-N, pPmpG22b-C and pPmpI23b-N expression vectors

The Expasy translated sequencing results for the passenger domain of construct (pPmpA23b-N) showed the C-terminal region of this domain was incorrect (Expasy Translate Tools). The translated sequence of the *pmpA-N* insert within pPmpA23b-N was found to be consistent with serovar D sequence up to the penultimate serine, after which an isoleucine and glutamic acid were replaced by a single serine resulting in a frameshift and the loss of the Histidine affinity tag and the stop codon, underlined in **Figure 2.15**. Primers were designed for site-directed mutagenesis to correct the frameshift shown in **Table 2.5**, the addition of one nucleic acid upstream of the affinity tag sequence produced a penta-his tag at the C-terminus in addition to five amino acids preceding the stop codon as shown in **Figure 2.16**.

DNA sequencing and Expasy translation of the insert within pPmpG22b-C (Pmp G autodomain) showed the construct was out of frame within its N-terminus region **Figure 2.17**. The pelB sequence of pPmpG22b-C was correct up to the highlighted valine, this was the region where the *pmpG-C* was inserted into the vector, a frameshift within this region resulted in the incorrect Pmp G autotranslocator protein sequence and multiple stop codons. Primers were designed and used for site-directed mutagenesis to correct the frameshift, the addition of one nucleic acid modified the N-terminus resulting in reinstatement of the predicted Pmp G autotranslocator sequence as determined by the published genomic sequence of serovar D, **Figure 2.18**.

Finally Expasy translation of the passenger domain of the *pmp I* construct (pPmpI23b-N) showed the C-terminal region was incorrect (Expasy Translate Tools). The expected pET23b hexahistidine tag sequence and the stop codon were not present as shown in **Figure 2.19.** The sequence of the insert in pPmpI23b-N was found to be correct up to the highlighted serine where primers were subsequently designed for site-directed mutagenesis to correct the frameshift, the addition of one nucleic acid upstream of the affinity tag sequence regained the hexa-histidine tag at the C-terminus in addition to the stop codon **Figure 2.20**.

```
--G V I Q L Q G D M K G S V S F V D Q R G A I I F T N N Q A V T S
--G V I Q L Q G D M K G S V S F V D Q R G A I I F T N N Q A V T S

S S M K H S G R G G A I S G D F A G S R I L F L N N Q Q I T F E G
S S M K H S G R G G A I S G D F A G S R I L F L N N Q Q I T F E G

N S A V H G G A I Y N K N G L V E F L G N A G P L A F K E N T T I
N S A V H G G A I Y N K N G L V E F L G N A G P L A F K E N T T I

A N G G A I Y T S N F K A N Q Q T S P I L F S Q N H A N K K G G A
A N G G A I Y T S N F K A N Q Q T S P I L F S Q N H A N K K G G A

I Y A Q Y V N L E Q N Q D T I R F E K N T A K E G G G A I T S S Q
I Y A Q Y V N L E Q N Q D T I R F E K N T A K E G G G A I T S S Q

C S I T A H N T I I F S D N A A G D L G G G A I L L E G K K P S L
C S I T A H N T I T F S D N A A G D L G G G A I L L E G K K P S L

T L I A H S G N I A F S G N T M L H I T K K A S L D R H N S I L I
T L I A H S G N I A F S G N T M L H I T K K A S L D R H N S I L I

K E A P Y K I Q L A A N K N H S I H F F D P V M A L S A S S S P I
K E A P Y K I Q L A A N K N H S I H F F D P V M A L S A S S S P I

Q I N A P E Y E T P F F S P K G M I V F S G A N L L D D A R E D V
Q I N A P E Y E T P F F S P K G M I V F S G A N L L D D A R E D V

A N R T S I E H H H H H H Stop
A N R T S - S T T T T T T E I R L L T K P E R
```

**Figure 2.15 - Depicts the frame shift mutations caused by the loss of nucleic acids in putative pPmpA23b-N**. The expected C-terminal sequence of the recombinant pmp A passenger domain following ligation of digested pmpA-N into digested pET23b is shown in black. The actual sequence after ligation is aligned in blue. The translational changes caused by the frameshift are underlined. The frameshift occurred following the highlighted serine residue. The hexahistine tag is replaced by a poly T tag and the stop codon was lost.

```
M A S N R V I E I H A H Y D Q R Q L S Q S P N T N F L V H H P Y L
M - - N R V I E I H A H Y D Q R Q L S Q S P N T N F L V H H P Y L

T L I P K F L L G A L I V Y A P Y S F A E M E L A I S G H K Q G K
T L I P K F L L G A L I V Y A P Y S F A E M E L A I S G H K Q G K

D R D T F T M I S S C P E G T N Y I I N R K L I L S D F S L L N K
D R D T F T M I S S C P E G T N Y I I N R K L I L S D F S L L N K

V S S G G A F R N L A G K I S F L G K N S S A S I H F K H I N I N
V S S G G A F R N L A G K I S F L G K N S S A S I H F K H I N I N

G F G A G V F S E S S I E F T D L R K L V A F G S E S T G G I F T
G F G A G V F S E S S I E F T D L R K L V A F G S E S T G G I F T

A K E D I S F K N N H H I A F R N N I T K G N G G V I Q L Q G D M
A K E D I S F K N N H H I A F R N N I T K G N G G V I Q L Q G D M

K G S V S F V D Q R G A I I F T N N Q A V T S S S M K H S G R G G
K G S V S F V D Q R G A I I F T N N Q A V T S S S M K H S G R G G

A I S G D F A G S R I L F L N N Q Q I T F E G N S A V H G G A I Y
A I S G D F A G S R I L F L N N Q Q I T F E G N S A V H G G A I Y

N K N G L V E F L G N A G P L A F K E N T T I A N G G A I Y T S N
N K N G L V E F L G N A G P L A F K E N T T I A N G G A I Y T S N

F K A N Q Q T S P I L F S Q N H A N K K G G A I Y A Q Y V N L E Q
F K A N Q Q T S P I L F S Q N H A N K K G G A I Y A Q Y V N L E Q

N Q D T I R F E K N T A K E G G G A I T S S Q C S I T A H N T I T
N Q D T I R F E K N T A K E G G G A I T S S Q C S I T A H N T I I

F S D N A A G D L G G G A I L L E G K K P S L T L I A H S G N I A
F S D N A A G D L G G G A I L L E G K K P S L T L I A H S G N I A

F S G N T M L H I T K K A S L D R H N S I L I K E A P Y K I Q L A
F S G N T M L H I T K K A S L D R H N S I L I K E A P Y K I Q L A

A N K N H S I H F F D P V M A L S A S S S P I Q I N A P E Y E T P
A N K N H S I H F F D P V M A L S A S S S P I Q I N A P E Y E T P

F F S P K G M I V F S G A N L L D D A R E D V A N R T S M H H H H
F F S P K G M I V F S G A N L L D D A R E D V A N R T S I F N Q P

H L R S G C
V H L Y N G
```

**Figure 2.16 - Sequence of *C.trachomatis* serovar E recombinant N-terminal Pmp A passenger domain following site-directed mutagenesis** as in section 2.2.16. Recombinant Pmp A passenger domain sequence is shown in black, published serovar D Pmp A sequence is aligned in red. The recombinant changes are underlined.

```
M K Y L L P T A A A G L L L L A A Q P A M A M D I G I N S D P N S
M K Y L L P T A A A G L L L L A A Q P A M A M D I G I N S D P N S


S S V D P N T A N N G P Y T L K A T W T K T G Y N P G P E R V A S
S S V E S I S I L I S S M A I Y F Y I T C / R S I H R S S N H N S


L V P N S L W G S I L D I R S A H S A I Q A S V D G R S Y C R G L
M S C / M K G I C C P C L L M G / E R C L S T R D G A I S I T D I


W V S G V S N F F Y H D R D A L G Q G Y R Y I S G G Y S L G A N S
S R H K A I F I R M C T R T S I K L H S N R N / I / E M S A L E C
```

**Figure 2.17 - Depicts the changes caused by frame shift mutations in putative pPmpG22b-C.** The N-terminal region of the recombinant protein sequence expected of the putative pPmpG22b-C construct following ligation of digested *pmpG-C* into digested pET22b shown in black. The actual sequence after ligation is aligned in blue. The translational changes caused by the frameshift are underlined, resulting in the transcription of the incorrect Pmp G protein sequence for the entire protein domain with several stop codons. The frameshift occurred after the pelB sequence (bold) at the highlighted valine residue.

M K Y L L P T A A A G L L L L A A Q P A M A M D I G I N S D P N S
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
- - - -

S S V D P N T A N N G P Y T L K A T W T K T G Y N P G P E R V A S
- - W D P N T A N N G P Y T L K A T W T K T G Y N P G P E R V A S

L V P N S L W G S I L D I R S A H S A I Q A S V D G R S Y C R G L
L V P N S L W G S I L D I R S A H S A I Q A S V D G R S Y C R G L

W V S G V S N F F Y H D R D A L G Q G Y R Y I S G G Y S L G A N S
W V S G V S N F F Y H D R D A L G Q G Y R Y I S G G Y S L G A N S

Y F G S S **M** F G L A F T E V F G R S K D Y V V C R S N H H A C I G
Y F G S S **M** F G L A F T E V F G R S K D Y V V C R S N H H A C I G

S V Y L S T K Q A L C G S Y L F G D A F I R A S Y G F G N Q H **M** K
S V Y L S T K Q A L C G S Y L F G D A F I R A S Y G F G N Q H **M** K

T S Y T F A E E S D V R W D N N C L V G E I G V G L P I V I T P S
T S Y T F A E E S D V R W D N N C L V G E I G V G L P I V I T P S

K L Y L N E L R P F V Q A E F S Y A D H E S F T E E G D Q A R A F
K L Y L N E L R P F V Q A E F S Y A D H E S F T E E G D Q A R A F

R S G H L **M** N L S V P V G V K F D R C S S T H P N K Y S F **M** G A Y
R S G H L **M** N L S V P V G V K F D R C S S T H P N K Y S F **M** G A Y

I C D A Y R T I S G T Q T T L L S H Q E T W T T D A F H L A R H G
I C D A Y R T I S G T Q T T L L S H Q E T W T T D A F H L A R H G

V I V R G S **M** Y A S L T S N I E V Y G H G R Y E Y R D T L E H H H
V I V R G S **M** Y A S L T S N I E V Y G H G R Y E Y R D T S R G Y G

H H H **Stop**
L S A G S K V R F

**Figure 2.18 - Sequence of *C.trachomatis* serovar E recombinant Pmp G autotranslocator domain following site-directed mutagenesis aligned with published serovar D**. The sequencing results for Pmp G autotranslocator sequence after site-directed mutagenesis (section 2.2.16) is shown in black. Recombinant Pmp G autotransloctor domain was designed to incorporate the pelB fusion petptide, published serovar D Pmp G autotransloctor sequence is aligned in red and does not contain a fusion pelB peptide (highlighted in yellow). The other recombinant changes are underlined.

```
--N L I C S G N V N P L F F T G N S A T N G G A I C C I S D L N T
--N L I C S G N V N P L F F T G N S A T N G G A I C C I S D L N T

S E K G S L S L A C N Q E T L F A S N S A K E K G G A I Y A K H M
S E K G S L S L A C N Q E T L F A S N S A K E K G G A I Y A K H M

V L R Y N G P V S F I N N S A K I G G A I A I Q S G G S L S I L A
V L R Y N G P V S F I N N S A K I G G A I A I Q S G G S L S I L A

G E G S V L F Q N N S Q R T S D Q G L V R N A I Y L E K D A I L S
G E G S V L F Q N N S Q R T S D Q G L V R N A I Y L E K D A I L S

S L E A R N G D I L F F D P I V Q E S S S K E S P L P S S L Q A S
S L E A R N G D I L F F D P I V Q E S S S K E S P L P S S L Q A S

V T S P T P A T A S P L V I Q T S A N R S V I F S S E R L S E E E
V T S P T P A T A S P L V I Q T S A N R S V I F S S E R L S E E E

K T P D N L T S H Y S G Q L A A A L E H H H H H H Stop
K T P D N L T S P L Q R P H S S T T T T T T E I R L L T K P R
```

**Figure 2.19 - Depicts the frame shift mutations caused by the loss of nucleic acids in putative pPmpl23b-N.** The expected C-terminal sequence of the recombinant pmp I passenger domain following ligation of digested *pmpI-N* into digested pET23b is shown in black. The actual sequence after ligation is aligned in blue. The translational changes caused by the frameshift are underlined. The frameshift occurred following the highlighted serine residue. The hexahistine tag is replaced by a poly threonine tag and the stop codon was lost.

```
M A S R P D H M N F C C L C A A I L S S T A V L F G Q D P L G E T
M - - R P D H M N F C C L C A A I L S S T A V L F G Q D P L G E T

A L L T K N P N H V V C T F F E D C T M E S L F P A L C A H A S Q
A L L T K N P N H V V C T F F E D C T M E S L F P A L C A H A S Q

D D P L Y V L G N S Y C W F V S K L H I T D P K E A L F K E K G D
D D P L Y V L G N S Y C W F V S K L H I T D P K E A L F K E K G D

L S I Q N F R F L S F T D C S S K E S S P S I I H Q K N G Q L S L
L S I Q N F R F L S F T D C S S K E S S P S I I H Q K N G Q L S L

R N N G S M S F C R N H A E G S G G A I S A D A F S L Q H N Y L F
R N N G S M S F C R N H A E G S G G A I S A D A F S L Q H N Y L F

T A F E E N S S K G N G G A I Q A Q T F S L S R N V S P I S F A R
T A F E E N S S K G N G G A I Q A Q T F S L S R N V S P I S F A R

N R A D L N G G A I C C S N L I C S G N V N P L F F T G N S A T N
N R A D L N G G A I C C S N L I C S G N V N P L F F T G N S A T N

G G A I C C I S D L N T S E K G S L S L A C N Q E T L F A S N S A
G G A I C C I S D L N T S E K G S L S L A C N Q E T L F A S N S A

K E K G G A I Y A K H M V L R Y N G P V S F I N N S A K I G G A I
K E K G G A I Y A K H M V L R Y N G P V S F I N N S A K I G G A I

A I Q S G G S L S I L A G E G S V L F Q N N S Q R T S D Q G L V R
A I Q S G G S L S I L A G E G S V L F Q N N S Q R T S D Q G L V R

N A I Y L E K D A I L S S L E A R N G D I L F F D P I V Q E S S S
N A I Y L E K D A I L S S L E A R N G D I L F F D P I V Q E S S S

K E S P L P S S L Q A S V T S P T P A T A S P L V I Q T S A N R S
K E S P L P S S L Q A S V T S P T P A T A S P L V I Q T S A N R S

V I F S S E R L S E E E K T P D N L T S P L Q R P Q L E H H H H H
V I F S S E R L S E E E K T P D N L T S Q L Q Q P I E L - - - - -

H Stop
- - -
```

**Figure 2.20 - Sequence of *C.trachomatis* serovar E recombinant N-terminal Pmp I passenger domain following site-directed mutagenesis aligned with published serovar D**. The sequencing results for recombinant Pmp I passenger domain sequence after site-directed mutagenesis (section 2.2.16) is shown in black. Published Pmp I sequence is aligned in red. The recombinant changes are underlined.

91

## 2.4 Discussion

Initialising studies with a large number of clones is one approach to counteract the high attrition rates associated with the expression of membrane proteins, a screening process often used in high-throughput studies (Cowieson *et al.,* 2005). Of the initial fifteen *pmp* cloning targets, fourteen were successfully amplified using polymerase chain reaction with the exception of full length *pmpG* that proved very difficult to amplify. To address the difficulty with *pmpG-FL* two different sets of primers were designed incorporating alternative restriction sites and various PCR conditions were trialled with each combination resulting in either zero amplification or a high amount of non-specific amplification. This could be an indication that the sequence is uniquely different from that of the published serovar D *pmp G* on which the primers were designed, possibly within the N-terminal region as there were no difficulties in amplifying the C-terminal *pmpG* autotranslocator domain. Furthermore Pmp G has been suggested to display a high antigenic variation where four subtypes have evolved in *C. pneumoniae*, *C, caviae* and *C. abortus* (Tan, et al. 2006).

Of the eight full length genes that were amplified, three were successfully incorporated into a pET vector in a manner suitable for expression trials. These recombinant constructs containing *pmpA-FL*, *pmpD-FL* and *pmpI-FL* were termed pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL respectively. The translated products of these recombinant sequences were designed to possess a C-terminal hexa-histidine fusion tag to allow detection and isolation of the expressed product via chromatographic methods. Unfortunately restriction mapping and sequencing of the recombinant constructs containing full length *pmpB, pmpC, pmpE, pmpF* and *pmpH* showed these to be unsuitable for expression screening. These constructs contained undesirable truncated sequences that had lost the native N or C terminal regions or were deficient in the intended fusion tag due to an alteration in the reading frame. The errors could have occurred during amplification as a result of mismatched priming, during ligation or transformation if the gene was unstable during replication. Many various attempts to adjust the cloning procedure were made. However the constructs were still unsuitable for further analysis.

By an adaptation of cloning strategies used to produce constructs for full length *pmpA*, *pmpD* and *pmpI*, recombinant constructs were made for the predicted functional domains of these Pmps with the use of nested primers. Of particular interest was the predicted passenger domains of Pmp A and Pmp I. Out of the nine

*C. trachomatis* Pmps, Pmp A is the only member with no obvious signal peptide sequence (Gomes *et al.,* 2006) and is highly conserved (Stothard *et al.,* 2003) therefore a putative role as a microbial virulence factor may be unique and differ distinctively from the other Pmps in the family. The Pmp I passenger domain was also of particular interest because of its variable sequence amongst the different serovars indicating a vital function for pathogenesis. It was therefore proposed that these predicted functional domains would be very attractive to study in functional adherence assays and structural analysis. The recombinant constructs for these Pmp domains were named pPmpA23b-N and pPmpI23b-N since the predicted passenger domains reside at the N-terminal region of the gene. The expressed products were also designed to contain a poly-histidine affinity tag at the C-terminus. Efforts to construct a recombinant clone for the Pmp H and Pmp D passenger domains were executed as both Pmp H and Pmp D are documented to be surface expressed in the bacterial cell, with Pmp D having a highly conserved sequence and Pmp H having a variable sequence amongst different strains indicative of tissue tropism (Stothard *et al.,* 2003). However the constructs were not suited for expression screening as the insert within the vectors was incorrect although further analysis has shown an error within the reverse primer used to amplify the *pmpD* passenger domain (primer set: *pmpD-N*) was responsible for a truncated insert. The one base error within the primer sequence resulted in the loss of the restriction site intended to be incorporated into the amplified sequence for subsequent cloning into the pET vector. A new primer with the correct base could easily rectify this as amplification was strong and ligation conditions were successful. Having identified this, successful cloning and expression of the passenger domain of Pmp D would be a consideration for further work.

Although the passenger domains were of interest in identifying Pmp function, the 'autotranslocator' C-terminal regions amongst the Pmps were of particular interest with regards to an understanding of the structure. Current structural data exists based solely on bioinformatic analysis that predicts Pmps to belong to the autotransporter family (Henderson and Lam 2001a, Wehrl *et al.,* 2004). These data suggest that the C-terminal autotranslocator regions have a β-barrel structure, typical of other prokaryotic autotransporters (Henderson *et al.,* 1998) but this has not been substantiated by circular dichroism or any other biophysical technique for any of the predicted Pmp pore-forming regions. In an attempt to confirm β-barrel structure a recombinant construct was made to contain the predicted autotranslocator domain of Pmp D and Pmp G. These recombinant vectors were

termed pPmpD28b-C and pPmpG22b-C. The expressed product of pPmpG22b-C was designed to contain a pelB leader sequence, an aid to help protein localisation to the periplasm as it was unknown at this point if the lack of the native N-terminal signal leader peptide would affect localisation of the expressed products. Additionally the recombinant cloning of PmpG autotranslocator domain had led to the loss of the carboxy terminal phenylalanine, an amino acid considered essential for localisation to the outer membrane (Struyvé *et al.,* 1991).

Site-directed mutagenesis was necessary to adjust the plasmid in some of the constructs. Errors in the sequence were not unexpected because the primers used were not completely complementary to the sequences they were intended to amplify and the custom design of the primers were designed using the published sequence of DNA of serovar D. Furthermore, the primers incorporated synthetic restriction sites for directional cloning which inevitably resulted in a certain amount of base mismatching, possibly introducing frame-shifts and point mutations.

## 2.5 Summary

Overall, the cloning strategies developed led to the successful construction of seven recombinant expression plasmids containing either full length or truncated *pmp* sequences obtained from a clinical strain of a serovar E *C. trachomatis,* all bearing a poly-histidine fusion tag ready for small scale expression trials in *E. coli.*

# Chapter 3 - Heterologous expression of full length and truncated recombinant Pmps in *E.coli* for downstream analysis

## 3.1 Introduction

To date, *C.trachomatis* Pmps have not been expressed as recombinant proteins at high enough levels for biophysical and functional analyses. Several of the homologous *Chlamydia pneumoniae* Pmps have been successfully expressed as recombinant protein yet structural information has not been obtained (Niessner *et al.,* 2003). As discussed in **section 1.5.2** the Pmps are predicted to be autoransporters that are likely to interact with the host cell to initiate infection, for which little analysis has been carried out to either substantiate or reject this notion. In order to investigate this further it is necessary to express high quantities of pure recombinant proteins suitable for biophysical and protein interaction studies.

The recombinant expression system provides opportunities to include modifications such as fusion tags and leader peptides which can direct protein targeting, increase stability and facilitate detection and purification. The most commonly used vehicle to express both prokaryotic and eukaryotic membrane proteins is *E.coli* since it is well characterised, inexpensive and easy to handle. However the over-expression of membrane proteins is rarely a straightforward task and often requires extensive optimisation due to their insolubility in aqueous buffers and a high occurrence of protein misfolding into inclusion bodies is observed (Seddon *et al.,* 2004). The choice of vector and host strain are therefore an important consideration because they have a direct effect on the level and timing of protein expression, which affects the location in the cell where the protein will accumulate, and whether the protein is biologically active (Miroux and Walker 1996). This is of particular importance when expressing membrane proteins, which can also be toxic to the cell, especially when produced at high levels. Therefore many parameters have to be met to ensure successful expression. However high levels of successfully expressed bacterial membrane proteins have been observed using *E.coli*, including numerous membrane proteins (Dutzler *et al.,* 2002, Griffin *et al.,* 2003, Korepanova *et al.,* 2009, Lee *et al.,* 2011). Importantly, it is currently impossible to predict whether a recombinant membrane protein can be over-expressed. If expression is attainable it is not currently possible to predict if it will be over-expressed in the membrane or as misfolded aggregates in inclusion bodies.

This chapter investigates the use of different expression vectors and *E.coli* hosts for the expression screening of rPmp A, rPmp D and rPmp I. As an alternative strategy the expression of predicted Pmp functional domains as truncated rPmps was also examined in an effort to increase the outcome of successful expression products

where changes in solubility and functionality of the proteins may have been affected. The hosts and vectors are discussed in **sections 3.1.1** and **3.1.2**. Several methods to overcome the common bottlenecks have been employed enabling the production of recombinant Pmp proteins suitable for downstream processing.

### 3.1.1 The pET expression system

Of the various expression systems that exist, many are designed for different applications and compatible with a range of hosts. The most widely used system for the recombinant expression of proteins in *E.coli* is the pET expression system commercialized by Novagen for which more than 40 different plasmids are available (Novagen). The system consists of multiple cloning sites for a range of genetic background modifications such as the incorporation of fusion partners, protease cleavage sites and hybrid promoters (Sørensen and Mortensen 2005).

Due to its adaptability, the pET system was used throughout these studies. In particular, vectors pET22b, pET23b and pET28b were chosen as each offered different potential attributes; pET22b allows the incorporation of a pelB signal peptide to encourage periplasmic localisation, pET23b allows incorporation of an N-terminal T7 tag that can be used as an alternative detection technique and pET28b possess the high stringency T7lac promoter to help suppress basal expression. The three vectors all allow for incorporation of a C-terminal histidine tag for purification and detection. Furthermore, the successful expression of several membrane transporters including the *Bacillus subtilis* tetracycline efflux protein TetA and the Na -dependent citrate carrier of *Klebsiella pneumoniae* has been achieved with the pET vectors (Cheng *et al.,* 1996, Kästner *et al.,* 2000). Moreover, the expression of a truncated recombinant Pmp C from *C. trachomatis* has been expressed using these strategies (Gomes *et al.,* 2005).

The system is designed for use with several strains of λDE3 lysogenised *E.coli* (**section 3.1.2**) which carry a T7 RNA polymerase gene under control of the Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible *lacUV5* promoter (**Figure 3.1**). IPTG binds to the *lacI* repressor protein, preventing interaction with the *lac* operator which in turn allows for the transcription of T7 RNA polymerase encoded for on the *E.coli* chromosome (Sørensen and Mortensen 2005). This T7 RNA polymerase binds to the T7 promoter on the pET plasmid to begin transcription of the target gene. Some pET vectors such as pET28b (**section 3.2.2.1**) confer a more stringent regulatory promoter, the T7*lac* promoter that offers a second *lac* operator situated downstream of the T7 promoter (**Figure 3.1**) such that binding of the

repressor molecule inhibits T7 RNA polymerase interacting at this site. This duo of lac operators inhibits both the binding of *E.coli* RNA polymerase at the *lacUV5* promoter reducing transcription of T7 RNA polymerase whilst also inhibiting the interaction of T7 RNA polymerase with the T7 promoter, collectively offering a tighter regulation over basal expression of the target gene (Novagen). The pLysS plasmid, compatible with many of the strains in **section 3.1.2** presents an even tighter regulation by encoding for T7 lysozyme, a natural inhibitor of T7 RNA polymerase thus reducing its capacity to transcribe the target gene in uninduced cells. The addition of IPTG is strong enough to overcome the levels of T7 lysozyme. Under non-stringent conditions without these regulatory mechanisms a small amount of transcription is permitted in the uninduced state and can be useful for the expression of genes that have innocuous effects on the host cell growth. However with membrane proteins a more stringent method is often necessary to avoid deleterious effects on cell growth when the expressed protein is toxic to the cell and/or when protein aggregation levels are high.

## 3.1.2 Strains for E.coli expression screening

BL21 (DE3) is the standard strain used for protein expression and is often the initial choice of host for expression screening trials. Protease deficient strains such as BL21 (DE3) and its derivatives are lacking in the *lon* protease and the *ompT* outer membrane proteases. Strains containing *lacY1* deletion mutants of BL21 (DE3) such as Tuner™ (DE3) cells allow adjustable levels of protein expression throughout all cells in culture. The lac permease (lacY1) mutation prevents active transport of IPTG into the cells, entry is via diffusion producing a concentration dependent homogenous level of induction allowing some repression by *lacI* see **Figure 3.1**. Controlling the expression rate this way can help with the solubility of the target protein by helping minimise the aggregation of protein. Furthermore, such protease deficient strains were chosen for this study as they offer some stability against protein degradation especially when downstream purification is necessary (Grodberg and Dunn 1988).

Strains with glutathione reductase (*gor*) and thioredoxin reductase (*trxB*) mutations greatly enhance the formation of disulfide bonds in the *E.coli* cytoplasm (Prinz *et al.,* 1997). Many proteins require the formation of stable disulfide bonds to ensure proper folding and thus activity. Usually disulfide bonds are formed only upon export into the periplasmic space. Without disulfide bonds, the proteins may accumulate as mis-folded inclusion bodies or instead may be degraded. This was deemed an

98

important consideration for expressing rPmps, which rich in cysteine residues may also fold in the periplasmic compartment (Tanzer and Hatch 2001a) and/or form intra-molecular disulfide bonds like other autotransporters (**section 1.5.2.2**). This can be limited in *E.coli* due to the relatively high reducing potential in the cytoplasmic compartment however the Origami™ strains (Novagen) carry both the *gor* and *trxB* mutations to allow disulphide formation within the cytoplasm.

*E.coli* and other species carry their own bias in the usage of the 61 available amino acid codons. In each cell the tRNA population closely reflects the codon bias of the mRNA population (Kane 1995). When the mRNA of heterologous target genes is expressed in *E.coli*, differences in codon usage can impede translation due to the demand for one or more tRNAs that may be lacking in the population. In particular, codons for arginine (AGG, AGA, CGA), leucine (CTA), isoleucine (ATA) and proline (CCC) can pose problems with expression. The presence of more than ~ 8-9 of the two rarest arginine codons (AGG and AGA), particularly back-to-back as a tandem repeat can lead to translational stalling, early translation termination, translational frame-shifting and amino acid mis-incorporation. Rosetta strains (Novagen) are designed to enhance the expression of eukaryotic proteins by supplying these rare tRNAs providing a 'universal' translation. All the aforementioned were considered carefully in these studies when choosing a suitable host for expression and codon analysis was carried out on the Pmps undergoing expression trials.
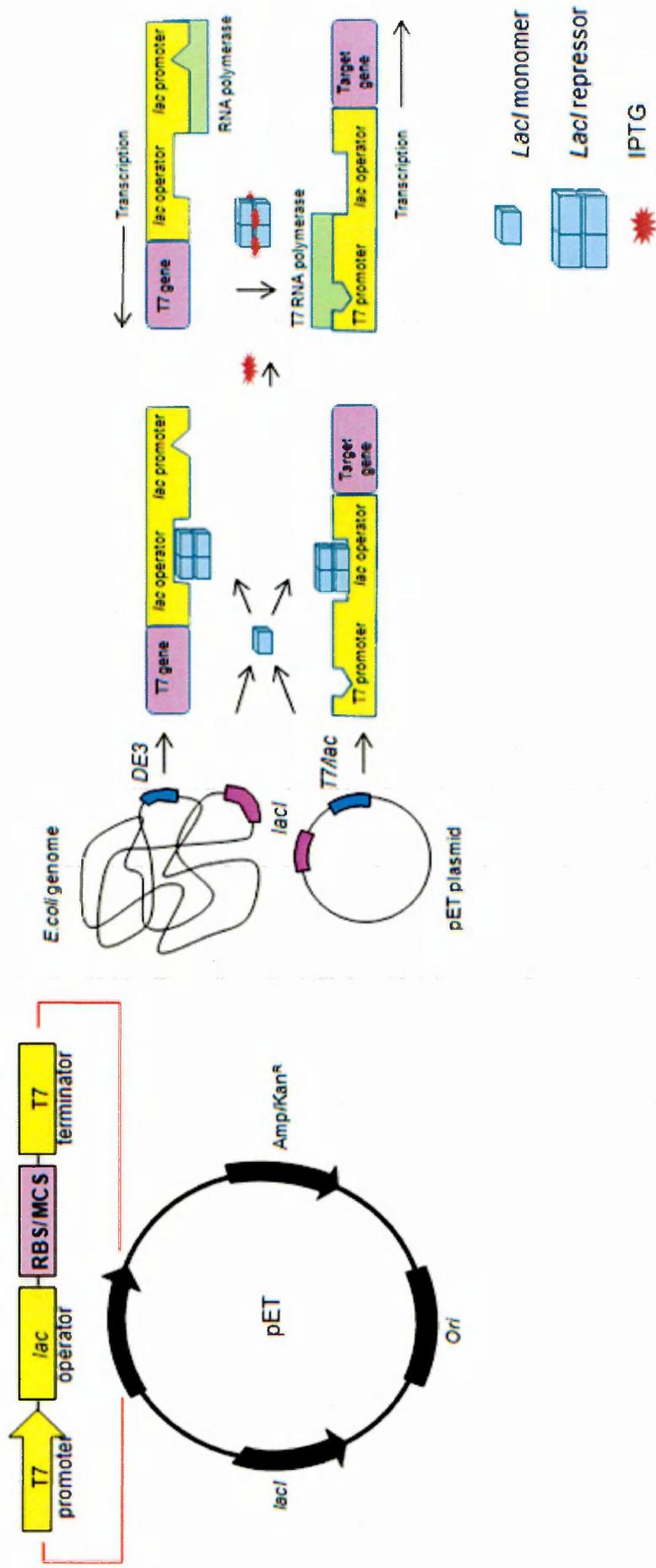
**Figure 3. 1 - Overview of the pET expression system**. Pre-induction, T7*lac* pET vectors confer a more stringent regulatory promoter, the T7*lac* promoter that offers a second *lac* operator situated downstream of the T7 promoter in addition to the lac operator situated downstream of the T7 gene on the chromosomal DNA. Binding of the repressor molecule occurs at both of these sites inhibiting the binding of *E.coli* RNA polymerase at the *lacUV5* promoter reducing transcription of T7 RNA polymerase and inhibiting the interaction of T7 RNA polymerase with the T7 promoter, collectively offering tight regulation over basal expression. Adapted from Sørensen, H.P., and Mortensen, K.K. (2005).

100

## 3.2 Materials/ Methods

### 3.2.1 Suppliers used in this chapter

Analytical grade chemicals were used throughout this study with all solutions being prepared using sterile water. Materials were obtained from the following suppliers.

**Bio-Rad Laboratories Ltd.**, Hemel Hemsted, Hertfordshire, **Cambridge Bioscience.**, Cambridge, **GE Healthcare Ltd**, Amersham, Buckinghamshire, **Invitrogen Life Technologies.**, Paisley, **Merck Chemicals Ltd.**, Beeston, Nottingham, **New England Biolabs (UK) Ltd.**, Hitchin, Hertfordshire, **Novagen, Merck Chemicals Ltd.**, Beeston, Nottingham, **Perbio Science UK Ltd.**, Cramlington, Northumberland, **Qiagen Ltd.**, Crawley, West Sussex, **Sigma Aldrich Company Ltd.**, Gillingham, Dorset, **Starlab UK.**, Blakelands, Milton Keynes, **Stratagene Ltd.**, Cambridge, **Stratech Scientific.**, Newmarket, **Thermo-Fisher Scientific UK.**, Loughborough, Leicestershire.

Glassware and media were sterilised by autoclaving at 121°C, 15 p.s.i. for 20 minutes. Solutions were made from molecular biology grade reagents and sterilised by filtration or autoclaving before use.

### 3.2.2 pET expression vectors

Recombinant vectors were produced using pET vectors shown in **section 3.2.2.1** to **3.2.2.3** and Pmp coding sequences amplified from genomic DNA as described in chapter 2. Plasmids pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL contained the coding regions for full length Pmp A, Pmp D and Pmp I respectively. Pmp N-terminal signal/passenger domains of Pmp A and Pmp I and the C-terminal autotranslocator domains of Pmp D and Pmp G were present within plasmids pPmpA23b-N, pPmpI23b-N, pPmpD28b-C and pPmpG22b-C respectively.

### 3.2.2.1 Plasmid pET28b(+)

In the plasmid pET28b, expression via the T7*lac* promoter is controlled by repression from the *lacI* gene present on the plasmid. A pET28b expression construct can be transformed into many host strains since it operates independently of the host. The plasmid allows incorporation of a hexahistidine tag both at the N-terminus and the C-terminus of the inserted gene to produce an affinity tag to assist in the detection and purification of the recombinant protein. The plasmid offers kanamycin resistance as a selective marker.
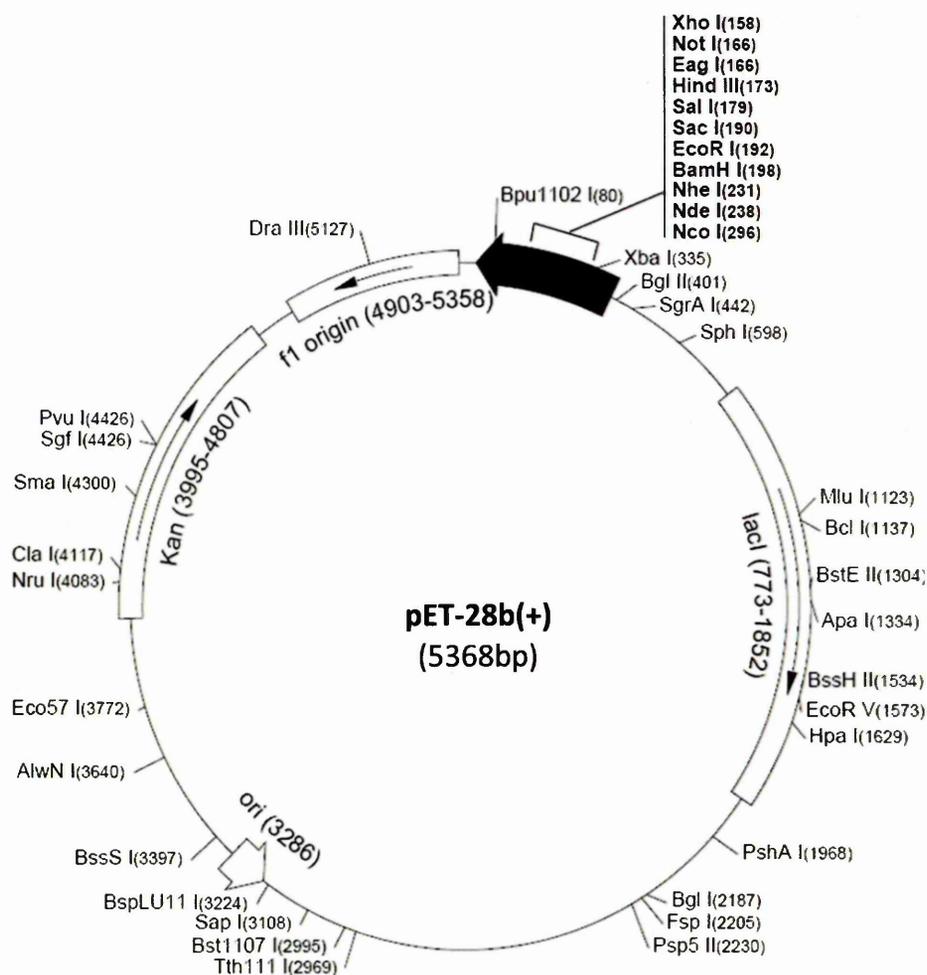


**Figure 3. 2 - The expression vector map, pET28b (+)** (Novagen)

## 3.2.2.2 Plasmid pET23b(+)

Unlike pET28b, the *lacI* gene is not present on the pET23b plasmid. A pET23b expression construct can also be transformed into many host strains. The plasmid allows incorporation of a hexahistidine tag at the C-terminus and a T7 tag at the N-terminus. The plasmid offers ampicillin resistance.
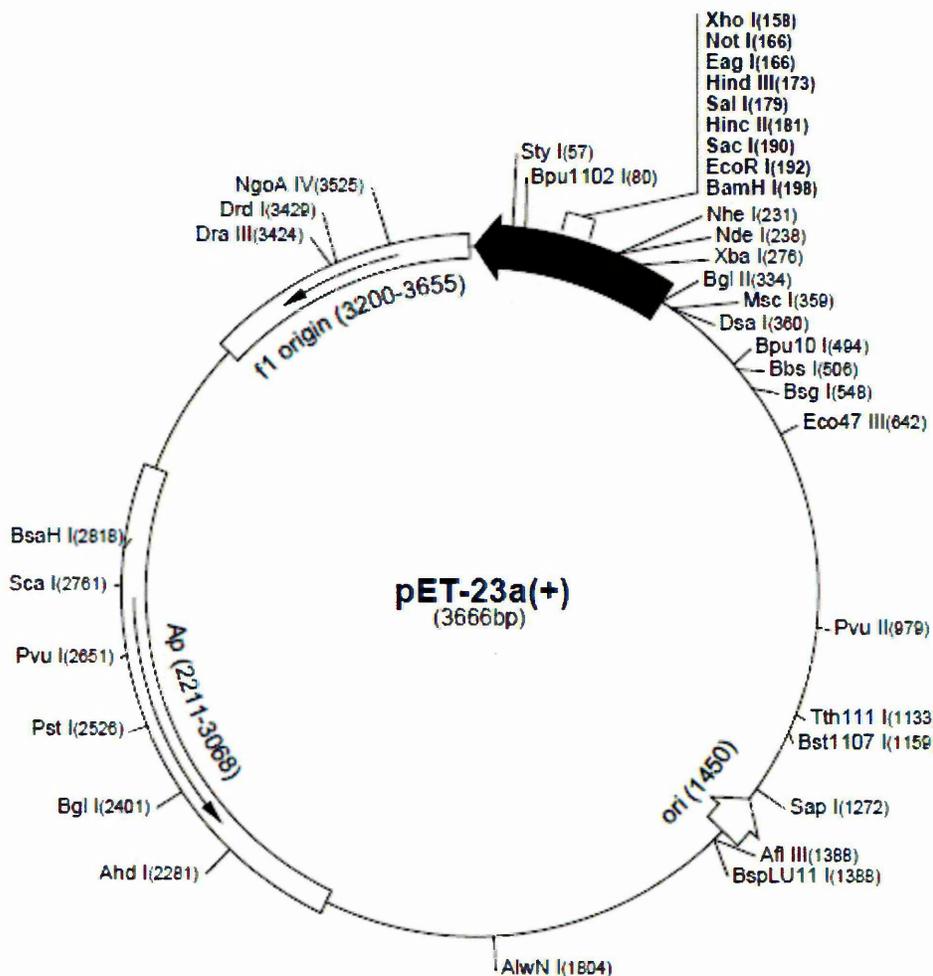


**Figure 3. 3 - The expression vector map, pET23b (+)** (Novagen)

119

### 3.2.2.3 Plasmid pET22b(+)

The pET22b(+) vector has similar properties to pET23b (+) but differs by carrying an N-terminal *pelB* signal sequence for potential periplasmic localisation, plus an optional C-terminal hexahistidine sequence. This can be potentially helpful for sequences that do not have a signal sequence but where periplasmic localisation of the target product would be advantageous.
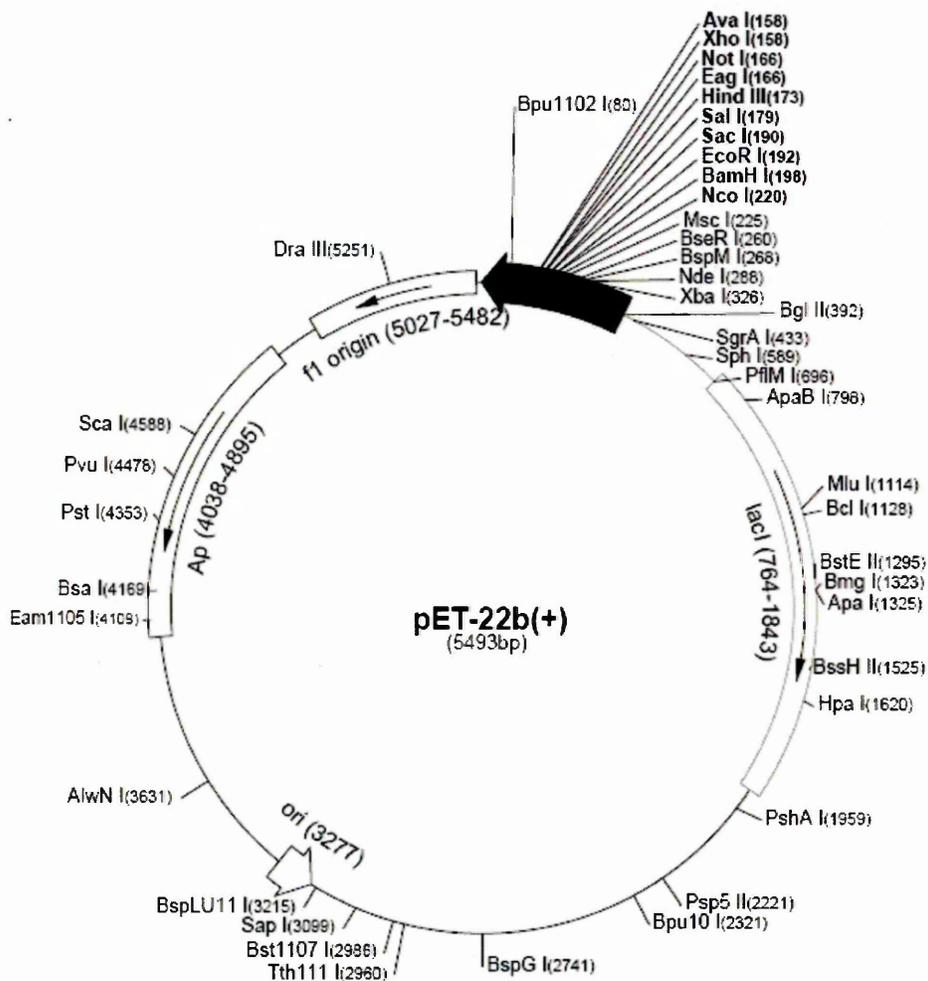


**Figure 3. 4 - The expression vector map, pET22b (+)** (Novagen)

### 3.2.3 E. coli transformation and protein expression

### 3.2.3.1 E. coli transformation

A 20 µl aliquot of competent (DE3) lysogenic strain, see genotypes in **section 3.2.3.7**, was transformed with plasmid DNA as in **section 2.2.9** with appropriate antibiotic (50 µg/ml ampicillin, 15 µg/ml kanamycin and/or 34 µg/ml chloramphenicol). Single bacterial colonies were selected as in **section 2.2.11** to inoculate liquid culture by placing a single colony into 50 ml LB broth with the appropriate antibiotic, and grown at 37°C, 200 rpm.

### 3.2.3.2 Protein expression

A 50 ml overnight culture of *E. coli* in LB (1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract) with the appropriate antibiotic(s) (50 µg/ml ampicillin, 15 µg/ml kanamycin and/or 34 µg/ml chloramphenicol), was used to inoculate 500 ml LB. The culture was incubated at varying temperatures to optimise expression, from 25 ° C - 37° C, 200 rpm until $A_{600}$ = 0.6-0.8. Protein expression was induced by the addition of 0.5 - 1 mM IPTG or alternatively expressed without IPTG, followed by a further incubation with varying time periods to optimise expression. Cells were harvested by centrifugation at 5000*g* for 20 minutes and were stored at -20° C.

### 3.2.3.3 Sec inhibition

Azide is recognized as a potent inhibitor of secA of the inner membrane sec translocon (Oliver *et al.,* 1990). Inhibition of the sec machinery can prevent the translocation and subsequent cleavage of signal peptides from pro-proteins such as autotransporters which are predicted to be processed via the sec pathway. Consequently, unprocessed pro-proteins that are restricted from sec processing accumulate in the cytoplasm as inclusion bodies (Gathmann *et al.,* 2006). During protein expression, where sec inhibition was required, cultures were treated with 0.2 mM, 1 mM or 2 mM sodium azide at the same time as IPTG induction.

### 3.2.3.4 Preparation of total cell protein for SDS-PAGE analysis

A 50 ml aliquot of induced culture **section 3.2.3.2** was centrifuged at 5000*g* for 10 minutes. The pellet was resuspended in 2 ml sonication buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0). This was incubated with 50 µl lysozyme (40 mg/ml) for 15 minutes at 37°C to disrupt the cell wall. Samples were passed through a Hamilton syringe until the viscosity was reduced followed by dilution in SDS sample buffer as described in **section 3.2.5.1**.

### 3.2.3.5 Fractionation of cells under non-denaturing conditions

Fractionation of cells into soluble and insoluble components was carried out. A culture was induced to express protein as described in **section 3.2.3.2**. The cell pellet was freeze/thawed prior to resuspension in ice cold lysis buffer (1XPBS, 20mM $\beta$-ME, 3mg/ml lysozyme, DNase (50 U) and 3mM $MgCl_2$) and incubated on ice for 20 mins. All handling buffers were supplemented with 1X Halt™ EDTA-free protease inhibitors (Thermo-scientific). The suspension was then sonicated on ice 5 X 10 second bursts with a sonication probe (set at 30% amplitude), then incubated on ice for a further 30 minutes before centrifugation at 13,000$g$ for 20 minutes at 4°C. The supernatant was retained (soluble fraction). The pellet was resuspended in solubilisation buffer (1XPBS, 20mM $\beta$-ME and 1-2% detergent (Triton X-100 or $n$-Dodecyl $\beta$-D-maltoside, DDM)). The suspension was incubated with gentle mixing at 4°C for 2 hours and centrifuged at 36,000$g$ for 30 minutes at 4°C, during this step, membrane proteins are released into the buffer. The supernatant was retained (solubilised membrane fraction). The insoluble pellet was washed in solubilisation buffer before centrifuging again at 36,000$g$ at 4°C to pellet inclusion bodies. All samples were diluted in SDS-sample buffer as in **section 3.2.5.1** for analysis by SDS-PAGE and western blotting.

### 3.2.3.6 Fractionation of cells under denaturing conditions

Fractionation of cells into soluble and insoluble components was carried out. A culture was induced to express protein as described in **section 3.2.3.2**. The cell pellet was freeze/thawed prior to resuspension in denaturing lysis buffer (100 mM $NaH_2PO_4$ 10 mM Tris-HCl 8 M urea pH8.0). The suspension was incubated with gentle mixing of the tube, end over end at room temperature for 1 hour and centrifuged at 12,000$g$ for 30 minutes at room temperature. The supernatant was retained (solubilised denatured protein fraction). The insoluble pellet contained urea insoluble debris and proteins. All samples were diluted in SDS-sample buffer as in **section 3.2.5.1** for analysis by SDS-PAGE and western blotting.

### 3.2.3.7 Genotypes of the strains of Escherichia coli used

(**C41 & C43 (DE3) +/- pLysS**: F- ompT hsdSB (rB- mB-) gal dcm (DE3). (Cambridge Bioscience)

**Rosetta-gami 2 (DE3) pLysS**: Δ(ara-leu)7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL (DE3) F'[lac+ lacIq pro] gor522::Tn10 trxB pLysSRARE2 (CamR, StrR, TetR) (Novagen, Merck)

**Tuner (DE3):** F– ompT hsdSB (rB– mB–) gal dcm lacY1(DE3) (Novagen, Merck)

**Tuner (DE3)** pLysS: F- ompT hsdSB (rB-mB-) gal dcm lacY1(DE3) pLysS (CmR) (Novagen, Merck)

## 3.2.4 Protein Purification

### 3.2.4.1 Metal affinity column chromatography with Ni-NTA agarose

Crude membrane fractions from 500 ml of *E. coli* culture were solubilised in 5-10 ml buffer containing the appropriate detergent (1XPBS, 10 mM imidazole, 1-2% (v/v) detergent [Triton X-100 or *n*-dodecyl-β-D-maltopyranoside DDM], 100μl HALT EDTA-free protease inhibitor cocktail). The suspension was incubated with gentle mixing of the tube, end over end at 4°C for 2 hours and centrifuged at 36,000*g* for 30 minutes at 4°C. The supernatant was applied to a 2ml column containing 750μl of pre-washed 50% Ni-NTA agarose slurry (Qiagen). The flow through was collected. The column was washed with 2 ml wash buffer (1XPBS, 20 mM imidazole, 0.1% (v/v) detergent). Protein was eluted in steps with increasing concentrations of imidazole (1XPBS, 0.1% (v/v) detergent with increasing gradients, from 20 mM to 500 mM imidazole), using 4 X 0.5 ml fractions for each concentration. Aliquots of all fractions were examined by SDS-PAGE and western blot as in **sections 3.2.5.1** and **3.2.5.2** for the presence of protein. To examine proteins under denaturing conditions the buffer was replaced by denaturing lysis buffer as **section 3.2.3.6**, and the same procedure was followed in the presence of imidazole. Where 50 ml cultures were used the protocol reagents were scaled down accordingly.

## 3.2.5 General protein techniques

### 3.2.5.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed according to Laemmli (1970). The electrophoresis of resolving gels, 1.5 mm thick, 10-12% acrylamide (w/v), was performed using a Biorad vertical minigel apparatus for approximately 1 hour at 40 mA per gel. Samples were diluted 1:4 in SDS-sample buffer (0.2 M Tris-HCl pH 6.8, 4 mM EDTA, 4% (w/v) SDS, 4% (v/v) glycerol, +/- β-ME). Gels were stained with Coomassie blue (0.25% (w/v) Coomassie brilliant blue in 25% (v/v) methanol/10% (v/v) acetic acid) and destained with 10% (v/v) acetic acid.

### 3.2.5.2 Western blotting

Transfer to nitrocellulose was performed as described by Towbin et al. (1979). Following SDS-PAGE gels were rinsed in transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol, pH 8.3). A nitrocellulose membrane (0.45 µm pore size, GE Healthcare, Amersham) was cut to the same dimension as the gel and soaked in $dH_2O$ for 10 minutes. A Bio-rad mini-gel blotting apparatus was set up according to the manufacturer's instructions, and the proteins were transferred for 1 hour at 100V in transfer buffer. Membranes were washed Tris-buffered saline (TBS, 20 mM Tris, 500 mM NaCl, pH 7.5) for 10 minutes prior to blocking with 1% alkali-soluble casein (Merck) in TBS containing 0.2% Tween-20 (TBST), for 2 hours at room temperature. A penta-His primary antibody (2 µg/ml) (Qiagen) in antibody buffer (TBST containing 1% alkali-soluble casein) was added for 2 hours at room temperature. Membranes were washed three times in TBST. Secondary antibody, a horseradish peroxidase conjugate of goat anti-mouse IgG (Stratech) was added at a dilution of 1:10,000 in antibody buffer for 1 hour at room temperature. Membranes were washed three times in TBST and peroxidase activity detected using the Pierce Supersignal ECL kit (Perbio Scientific UK) according to the manufacturer's instructions.

### 3.2.5.3 Protein precipitation

Protein precipitation of the media was performed using the Trichloroacetic acid (TCA) protein precipitation protocol (Novagen, 2005). 10ml of the media from the cell harvest in **section 3.2.3.2** was incubated with 1ml (1/10 volume) of 100% TCA (w/v) on ice for 15 minutes. The sample was centrifuged at 16,000g for 10 minutes and the supernatant discarded. The pellet was washed twice with 1ml acetone and centrifuged at 16,000g for 5 minutes. The acetone was removed and the pellet was allowed to air dry in a fume hood for 2 hours or until dry. Samples were resuspended in 1X PBS and SDS-sample buffer.

## 3.3 Results

### 3.3.1 Expression trials for full length rPmp A, D and I using the *E.coli* strains Tuner™ (DE3), Rosetta-gami 2™ (DE3), Overexpress C41™ (DE3) and C43™ (DE3)

Membrane proteins that are expressed at a rapid rate have a high tendency to aggregate into inclusion bodies (Rudolph and Lilie 1996). As explored in **section 3.1.2** the lac permease (*lacY1*) mutation present within the Tuner™ (DE3) cells prevents active transport of IPTG. Instead entry is via diffusion producing a concentration dependent homogenous level of induction. Controlling the expression rate this way can help with the solubility of the target protein and so this was the initial strain chosen for early expression trials.

Expression plasmids pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL were transformed separately into Tuner™ (DE3) cells (**section 3.2.3.1**) and expression trials were cultured in LB (**sections 3.2.3.2**). The expression of rPmps was examined under a variety of conditions with cultures induced to express recombinant Pmps for between 2-16 hours, dependent upon the incubation temperature used; at 37 °C the induction period was between 2 and 3 hours whilst cultures induced at room temperature were allowed an extended growth period, typically overnight. IPTG concentrations used for induction were typically between 0.5 mM and 1 mM following the recommended guidelines for the particular pET vector: Coding sequences inserted in to pET28b are under control of the T7*lac* promoter. This promoter allows binding of the *lac* repressor thus reducing transcription by T7 RNA polymerase to suppress basal expression as discussed in **section 3.1.1**. This high stringency promoter requires a higher concentration of IPTG (1 mM) to fully induce compared to the regular T7 promoter of pET23b and pET22b (0.5 mM). Alongside each of the induced cultures was an uninduced culture grown under the same conditions that served as a control. Following expression, harvested cultures were separated into total cell protein, soluble and insoluble fractions under both denaturing and non-denaturing conditions (**sections 3.2.3.5 and 3.2.3.6**) for analysis by SDS-PAGE and Western blotting (**sections 3.2.5.1 and 3.2.5.2**). Any secreted proteins localised to the media were also precipitated using TCA (**section 3.2.5.3**).

Western blot analysis showed rPmp A, rPmp D and rPmp I were not detected in induced Tuner™ (DE3) cells under any of the variant conditions (data not shown)

when compared to uninduced cultures and cultures expressing empty pET vectors. The expected sizes of rPmp A, rPmp D and rPmp I are approximately 105 kDa 160 kDa and 95 kDa, respectively. Such bands were not detected when probed with a range of different anti-His primary antibodies as described in **section 3.3.6**. Growth rates were observed before and after induction with IPTG to monitor the effects of IPTG induction on the cells. The growth of the induced culture slowed notably in relation to the uninduced culture following induction of each of the rPmp expression plasmids (data not shown). This was encouraging as it is likely the resources of the induced cells were being diverted in an effort to produce the rPmps. Induction with IPTG did not appear to be lethal to the cells as growth was maintained albeit at a slower rate (data not shown).

Rosetta strains (Novagen) are designed to enhance the expression of proteins carrying rare codons not recognised in *E. coli* by supplying the rare tRNAs (**section 3.1.2**). Using the Rare Codon Calculator (RaCC) (NIH MBI Laboratory for Structural Genomics and Protemoics) codon analysis of *pmp A*, *pmp D* and *pmp I* revealed a high number of rare codons for these *pmps*. *pmp D* contained 55, 30 of these were rare Arg codons with 1 tandem rare Arg codon. *pmp A* contained 57 rare codons, 17 being rare Arg codons and *pmp I* contained 32 rare codons with only 11 as rare Arg codons. These were all at levels that could affect the expression of the full length Pmps.

Therefore following expression attempts using Tuner™ (DE3) cells, trials were performed using a cell line that could allow for rare tRNAs. The Rosetta-gami 2™ (DE3) pLysS cell line was transformed with pPmpA23b-FL and pPmpI23b-FL and were cultured in LB (**section 3.2.3.2**). Expression was induced with 0.5 mM IPTG for between 2 and 4hrs at both 30 °C and 37 °C alongside uninduced cultures. Expression was also examined for an extended period typically overnight at room temperature for both plasmids. The cells were harvested and prepared into total cell protein, soluble and insoluble fractions under both denaturing and non-denaturing conditions for analysis by SDS-PAGE and Western blotting. Secreted proteins localised to the media were also precipitated using TCA for examination. Expression trials for rPmp D were not carried out with this cell line because the strain is kanamycin sensitive and this expression plasmid carries the kan[R] gene.

Western blot analysis indicated that rPmp A and rPmp I were not expressed in Rosetta-gami 2™ (DE3) pLysS cells under any of the variant conditions (data not shown). Interestingly the addition of IPTG had a marked effect upon the growth of

pPmpI23b-FL transformed Rosetta-gami 2™ (DE3) pLysS cultures. Following induction, the culture growth rate began to plateau and did not recover after three hours (shown in **Figure 3.5**). It is possible that the recombinant Pmp I may have had a deleterious effect on the growth rate. The pLysS plasmid was present to reduce basal expression and minimise putative toxic effects of expressed harmful protein prior to induction. However up to this point pre-induction growth rates were comparable. The change in growth rate led to the assumption that recombinant protein was produced. However rPmp I was not detected via Western blotting or via SDS-PAGE. The addition of IPTG did not have a deleterious effect when inducing the expression of recombinant Pmp A in Rosetta-gami 2™ (DE3) pLysS cells (data not shown). However it was noted that the transformation efficiency of pPmpA23b-FL compared to pPmpI23b-FL was poor using this strain. This could indicate that the rPmpA plasmid or recombinant protein also had deleterious effects. Cultures were always propagated from freshly transformed *E.coli*, however the stability of the plasmid maintenance was not checked. A positive control was not available.

OverExpress™ C41 and C43 (Lucigen) strains have been developed to allow high level expression of a wide variety of toxic proteins proteins previously difficult or impossible to express in bacteria (Miroux and Walker 1996). To this end, trials were carried out with OverExpress™ C43 (DE3) and Overexpress C41™ (DE3) cells transformed separately with pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL. Transformed cells were cultured in LB. Cultures were induced for between 3 and 4 hours both at 30°C and 37°C with both 0.5 mM and 1 mM concentrations of IPTG. Overnight trials at room temperature were also performed. The cultures were prepared into total cell protein, soluble and insoluble fractions under both denaturing and non-denaturing conditions for analysis by SDS-PAGE and Western blotting. Secreted proteins localised to the media were also precipitated for analysis.

Recombinant proteins rPmpA, rPmpD and rPmpI were not detected in induced OverExpress™ C43 (DE3) or Overexpress C41™ (DE3) cells under any of the variant conditions according to Western blot analysis and SDS-PAGE (data not shown). Although transformation efficiency of this cell line was good with all three recombinant plasmids, examination via Western blot and SDS-PAGE did not detect over-expressed His-tagged proteins.

Commonly a very poor yield of expression hampers detection by Western blot or SDS-PAGE. Ni-NTA affinity was therefore used to concentrate putative His-tagged proteins prior to analysis for each of the fractionated cultures expressing rPmps.

This was performed under denaturing conditions (**section 3.2.3.7**) to ensure exposure of the fusion tag. The solubilised cellular fractions were loaded on to the column to collect any putative polyHis tagged protein. The washing steps were omitted to prevent dilution of the proteins for detection. Instead the Ni-NTA associated proteins were eluted from the column with 500 mM imidazole to produce a sample 10 times more concentrated for Western blot analysis. Western blotting did not detect any putative His-tagged rPmps.

Based on the theory that the protein is targeted to the Sec machinery, full length recombinant Pmps were expressed in the presence of a sec inhibitor, sodium azide to inhibit the translocation and subsequent cleavage of the full length Pmp. By attempting to inhibit the sec machinery using methods described by Oliver *et al.,* (1990), expressed recombinant protein could be restricted to the cytoplasm. Although this would most likely cause the protein to aggregate into inclusion bodies unsuitable for this study, it was hoped that it could confirm protein expression. The growth rates of the cultures were monitored during expression trials in the presence of azide. Compared to untreated controls the growth rates slowed but were maintained steadily showing no lethal effect on the cells (data not shown). Subsequently Western blotting and SDS-PAGE did not detect any putative His-tag recombinant Pmps as soluble or inclusion bodies from these trials (data not shown).
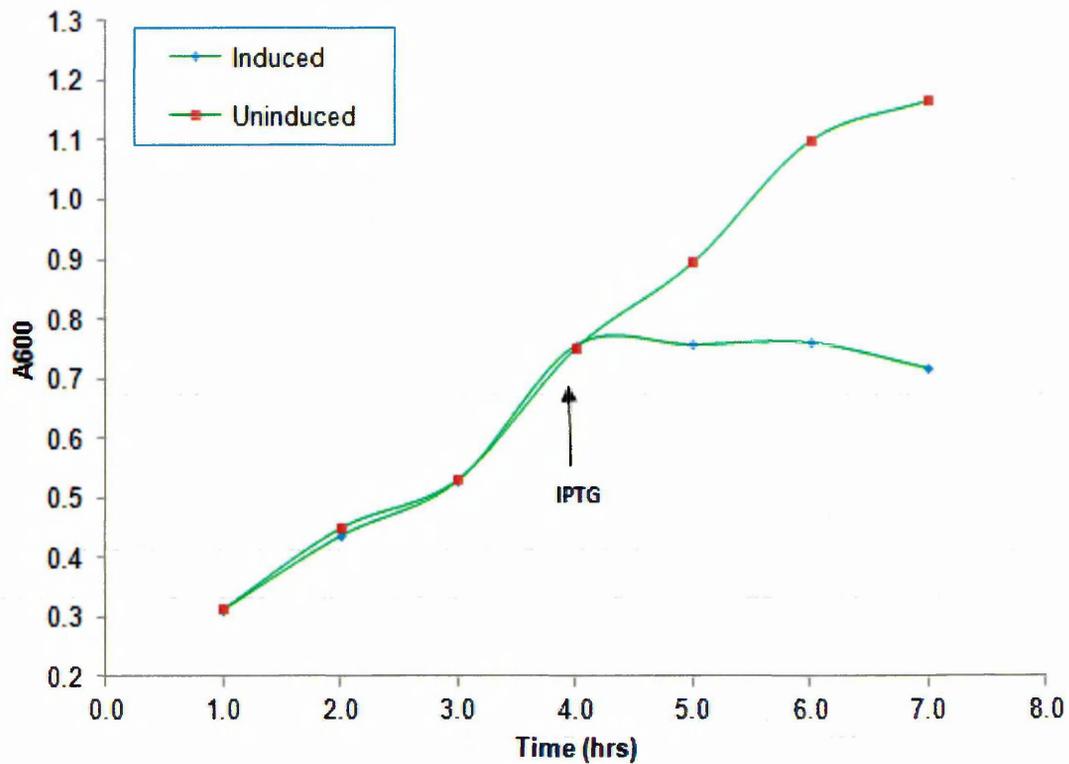
**Figure 3. 5 - Growth curve of induced and uninduced strain Rosetta-gami 2™ (DE3) pLysS transformed with plasmid pPmpl23b-FL**. Cells were cultured as described in **section 3.2.3.2** with or without induction using 0.5 mM IPTG at 30 °C. IPTG was added when $A_{600}$ reached 0.7-0.8, as indicated. The $A_{600}$ was measured every hour.

### 3.3.2 Expression and purification trials of the putative β-barrel autotranslocator domain of Pmp D (rPmpDauto)

Expression trials were performed in order to produce a recombinant autotranslocator domain of Pmp D. Tuner™ (DE3) cells transformed with pPmpD28b-C were induced with 1.0mM IPTG for 3hrs at 30°C as in **section 3.2.3.2**. Harvested cells were lysed and resuspended in 8M urea phosphate buffer, separated into soluble and insoluble, as in **section 3.2.3.6** to denature the proteins and expose the C-terminal His-tag of putative expressed protein. Proteins from the media were concentrated using TCA precipitation to check for protein secretion.

**Figure 3.6** shows the expression of the putative recombinant Pmp D autotranslocator domain (rPmpDauto) at approximately 27kDa observed by Coomassie stained SDS-PAGE and confirmed by Western blot probed with the penta-His primary antibody (red arrows). The penta-His antibody binds with high avidity to rPmpDauto domain. Expression was observed in both the induced and uninduced cultures with bands of similar density on the Coomassie stained SDS-PAGE gel. The protein was observed in the insoluble fractions. Insoluble fractions consist of unsolubilised cellular membrane proteins and inclusion bodies. 8M urea is a strong denaturant but it is not always capable of solubilising all proteins and does not always differentiate between inclusion body aggregates and unsolubilised membrane bound proteins. Therefore it was necessary to determine if the protein detected in the insoluble fraction was situated in the membrane or if it was in the form of inclusion bodies. This would require the use of detergents. There was no indication that rPmpDauto was secreted from the cell into the media.

Harvested cells from identical expression trials were lysed to isolate the inclusion bodies as in **section 3.2.3.6**. Visualisation on a Coomassie SDS-PAGE gel indicated the recombinant PmpDauto had aggregated as inclusion bodies typically found in high level expression shown in **Figure 3.7**. Importantly the high level of basal expression did not have detrimental effects on the *E.coli* growth (data not shown) indicating that expression of this particular rPmp was tolerated. Interestingly high levels of target protein expression were observed in both the induced and uninduced control. The expression of rPmp D autotranslocator domain was under the control of the T7*lac* promoter, a tight regulatory promoter that should prevent basal expression (described in **section 3.1.1**) but the inhibitory effect of this mechanism was not observed here. Furthermore, Tuner™ (DE3) cells permit uniform entry of IPTG to allow control over expression levels within the cell,

however, the expression observed in the uninduced cells was comparable to the expression seen in the IPTG induced cells in both **Figure 3.6** and **Figure 3.7** suggesting that expression was influenced by other factors.

To test whether cyclic AMP (cAMP) levels were responsible for the high basal expression of rPmpDauto trials were repeated in the presence of media supplemented with 1% glucose, a method termed catabolite repression. Expression was performed at a reduced temperature of 25 °C to also reduce basal expression. Samples were fractionated under non-denaturing conditions and analysed by Coomassie stained SDS-PAGE. As it was previously noted that uninduced Tuner (DE3) also expressed rPmpDauto, this time cells were also transformed with the empty vector pET28b to serve as a negative control.

When comparing the differences between induced and uninduced inclusion bodies shown in **Figure 3.7** with the induced and uninduced inclusion bodies shown in **Figure 3.8,** catabolite repression accompanied with a lower incubation temperature of 25 °C reduced basal expression of rPmpDauto prior to IPTG induction. However this was insufficient to prevent basal expression completely as rPmpDauto protein was still expressed as insoluble inclusion bodies in the uninduced culture **Figure 3.8** Further efforts were necessary to minimise protein aggregation to allow rPmpDauto to localise to the membranes.

Subsequently expression trials were performed using heat shock methods to enhance chaperones at 42 °C for 1 minute prior to IPTG induction. An alternative technique that is thought to elicit the same response is to add 1-2% ethanol to the culture prior to induction with IPTG. These methods were utilised in order to facilitate protein folding and thus help direct rPmpDauto to the membrane. However all rPmpDauto protein was found to be in inclusion bodies concluding that these techniques used alone or in combination were not useful in the elimination of insoluble aggregates (data not shown).

To modulate basal expression observed using *E.coli* Tuner™ (DE3) cells, the pLysS plasmid was used to ensure expression was regulated by IPTG levels (**section 3.1.1**) The pLysS encodes for T7 lysozyme that breaks down T7 RNA polymerase necessary for expression. After an initial growth period at 30°C, expression was induced at 25 °C and cultures were incubated overnight. The harvested cells were lysed and fractionated under non-denaturing conditions into soluble proteins, detergent solubilised membranes and inclusion body fractions. **Figure 3.9** shows the addition of the pLysS plasmid significantly reduced basal expression of the

uninduced control where inclusion bodies were minimal. In a further attempt to regulate expression and improve localisation of rPmpDauto to the membranes, sub-culturing phases and the exponential growth phase of the *E.coli* culture were performed at room temperature 25 ºC. This had previously been done at 30 °C or 37 °C before lowering the incubation temperature for the induction period. Cultures were then incubated overnight without IPTG to modulate expression. Monomers of putative rPmpD autodomain at 27 kDa were detected in the detergent solubilised membrane fractions of **Figure 3.10.** Interestingly, it was noted that the majority of rPmpDauto expression was still in the form of inclusion bodies even under these tightly regulated expression rates.
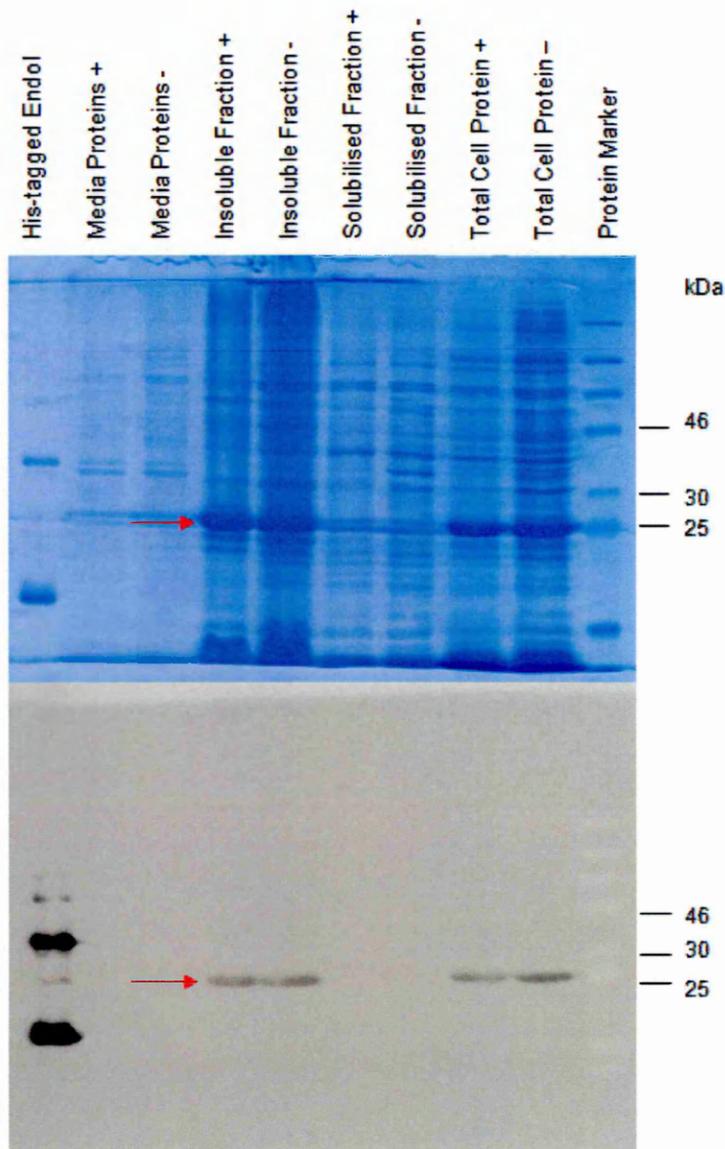
**Figure 3. 6 - Coomassie stained SDS-PAGE gel (top) and Western blot (below) of rPmp D autotranslocator domain expression in Tuner™ (DE3) cell line**. Tuner (DE3) cultures were induced with 1.0mM IPTG (+) or uninduced (-) at 30°C. Molecular weights are shown in kDa. The *E.coli* lysates were separated into soluble and insoluble fractions using 8M urea and centrifugation. Proteins were also precipitated from the media. An equal volume of each fraction was loaded into each lane. Lanes are: His-tagged positive control Endo1 monomers (17 kDa) and dimers (35 kDa), precipitated media proteins, urea insoluble proteins, urea soluble proteins and total cell proteins, followed by a prestained protein marker (NEB Biolabs). The expression of rPmpDauto at approximately 27 kDa was detected in both the induced and uninduced cultures and is shown by the red arrow in the total cell protein fractions and the insoluble fractions.

133
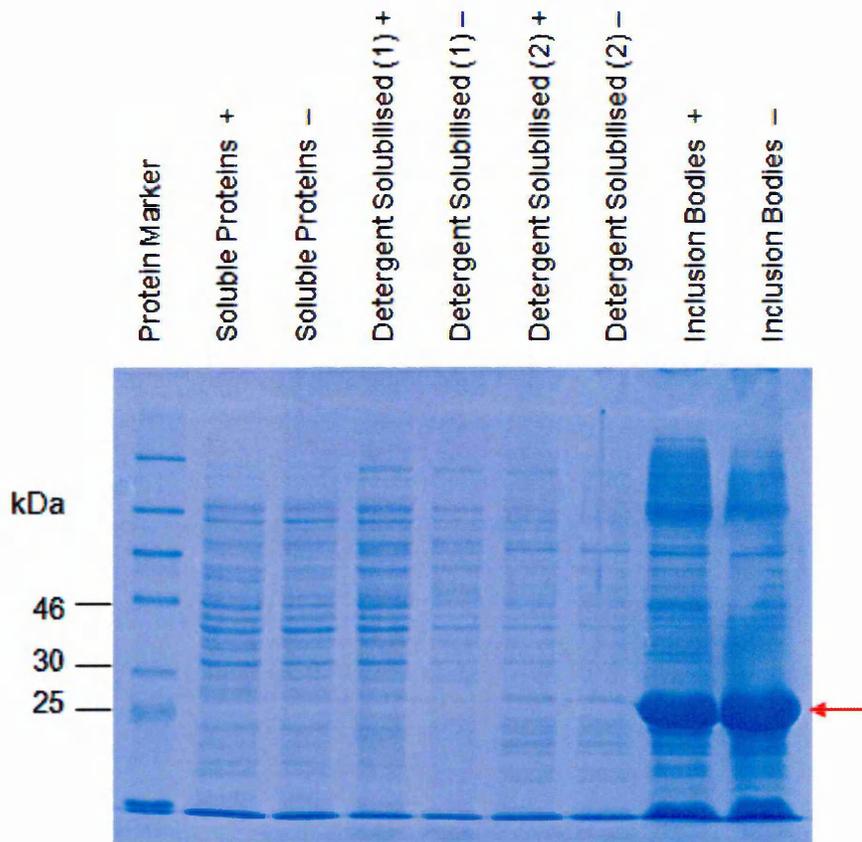
**Figure 3.7 - Coomassie stained SDS-PAGE gel showing rPmp D autotranslocator domain expressed as inclusion bodies in Tuner™ (DE3)**. Tuner (DE3) cultures were induced with 1.0mM IPTG (+) or uninduced (-) at 30 °C. The *E.coli* lysates were separated into soluble and insoluble fractions. The insoluble fractions were suspended in 1% Triton-X100 to release membrane bound proteins. The remaining pellet contained the inclusion body aggregates. An equal volume of each fraction was loaded into each lane. Lanes are as labeled, the protein marker is shown in kDa. rPmpD autodomain was found as inclusion bodies at approximately 27kDa, shown by red arrow.
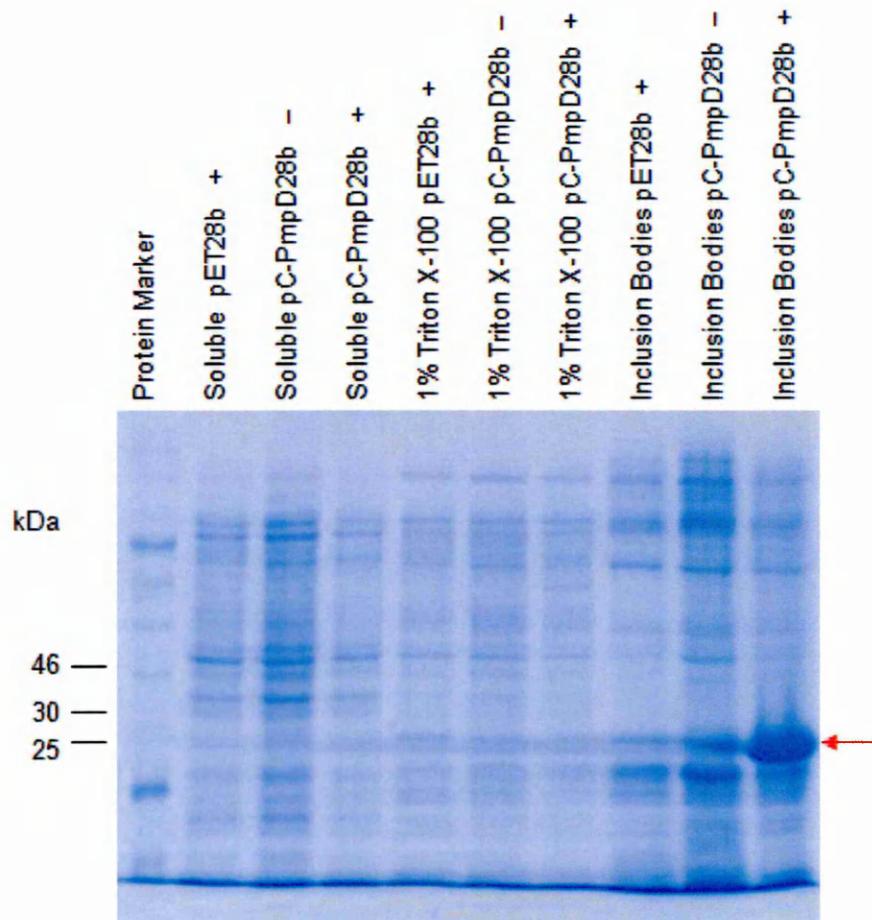
**Figure 3. 8 - Coomassie stained SDS-PAGE gel shows catabolite repression reduces basal expression of rPmp D autodomain**. Tuner (DE3) cultures were induced with 1.0mM IPTG (+) or uninduced (-) at room temperature with 1% glucose supplementation. The *E.coli* lysates were separated into soluble fractions, detergent solubilised proteins using 1% Triton X-100 and inclusion bodies. An equal volume of each fraction was loaded into each lane as labelled. *E.coli* transformed with empty vector pET28b were induced as a negative control. The prestained protein marker (NEB biolabs) depicts MW in kDa.
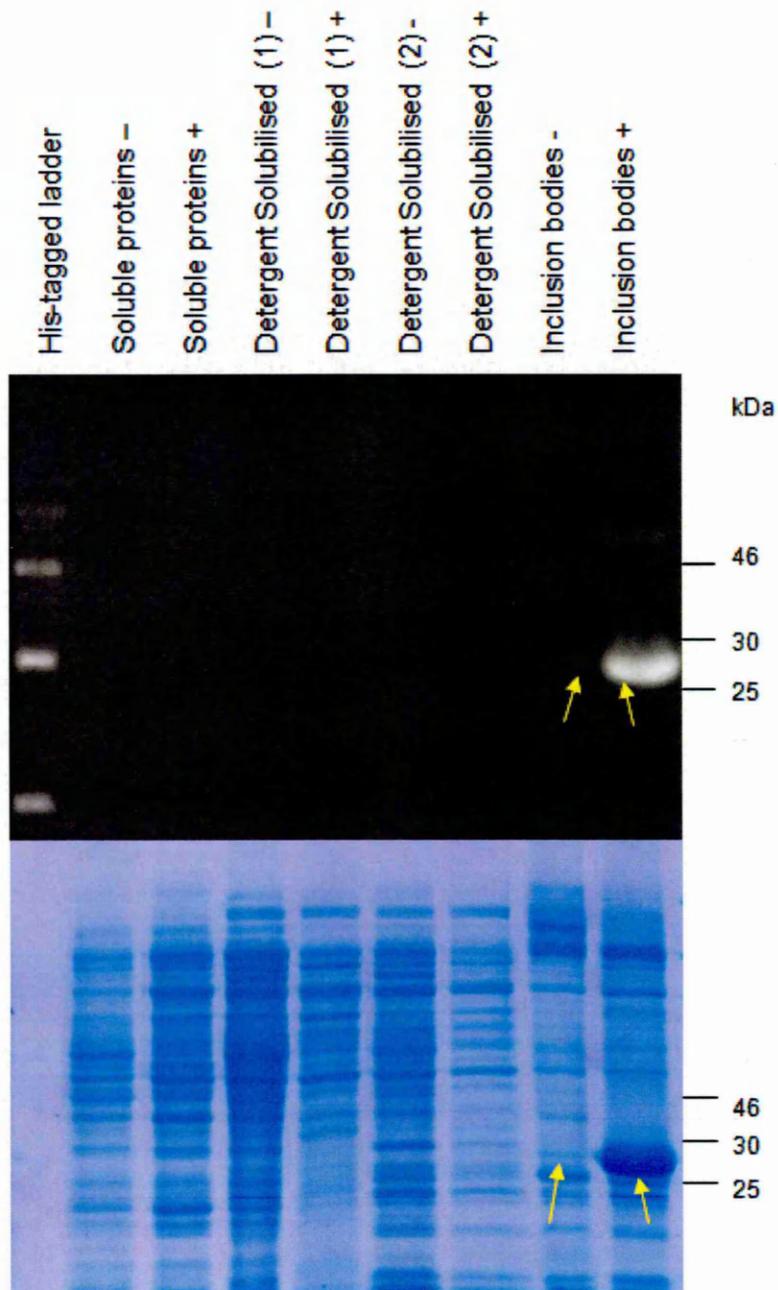
**Figure 3.9 - Western blot (top) and Coomassie SDS-PAGE (below) showing the expression of rPmp D autodomain using the pLysS plasmid to inhibit basal expression.** Tuner™ (DE3) pLysS transformed with pPmpD28b-C were induced (+) or not induced (-) with 1.0mM IPTG at room temperature. The *E.coli* lysates were separated into soluble fractions, detergent solubilised proteins using 1% Triton X-100 and inclusion bodies. An equal volume of each fraction was loaded into each lane as labelled. The basal expression of rPmpD autodomain was reduced by inhibiting T7 RNA polymerase, a reduction in rPmpD autodomain inclusion bodies between the induced and uninduced cultures is shown by yellow arrow.

136
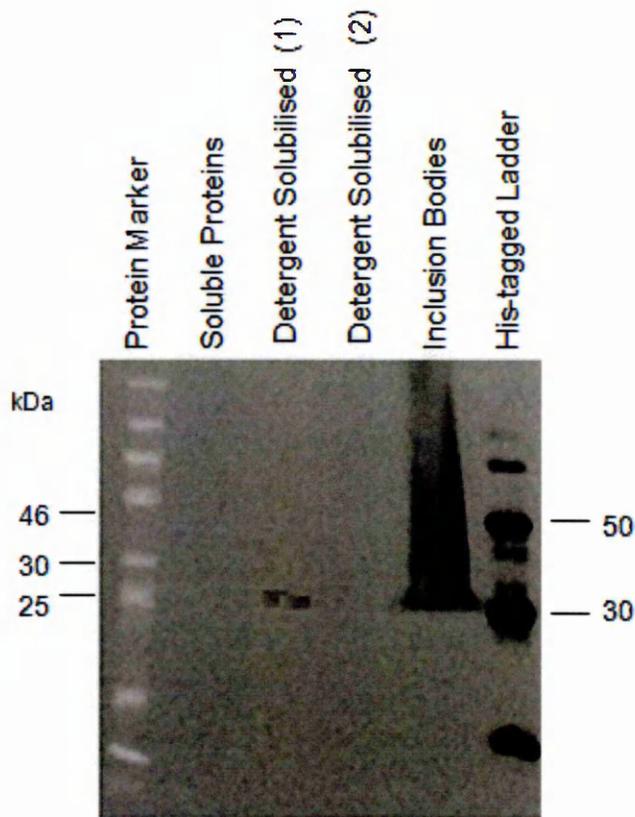
**Figure 3.10 - The natural unregulated expression of rPmp D autotranslocator domain shows expression to the *E.coli* membranes for solubilisation using detergent.** Tuner™ (DE3) pLysS transformed with pPmpD28b-C expressed rPmpD autotranslocator domain without IPTG regulation, relying on the natural unregulated expression of the plasmids and *E.coli*. All growth phases were carried out at room temperature. The *E.coli* lysates were separated into soluble fractions, detergent solubilised membrane fractions and inclusion bodies. An equal volume of each fraction was loaded into each lane. A faint band was detected in the detergent solubilised fraction at approximately 27kDa. This was also detected in the inclusion body fractions. Discrepancies between the pre-stained marker (NEB biolabs) and the 6xHis ladder (Qiagen) between 25kDa and 30kDa bands are evident. rPmpDauto is observed at ~27kDa when compared against the pre-stained marker, but appears to be ~32kDa when compared against the 6xHis ladder. This discrepancy was seen on all Western blots where both markers were used.

### 3.3.2.1 Small-scale purification of E.coli derived rPmpDauto using Ni-NTA chromatography

The detergent solubilised membrane fractions of IPTG independent expression of rPmpD autotranslocator domain from Tuner™ (DE3) pLysS were bound to Ni-NTA agarose as described in **section 3.2.4.1**. To optimize the purification, numerous wash and elution conditions under using a range of imidazole concentrations of between 10 and 500 mM were trialed in the binding and washing steps. This protein was only partially purified using Ni-NTA chromatography. Despite varying the purification parameters it was difficult to obtain a sample with at least 95% homogeneity suitable for circular dichroism. Western blot analysis shows recombinant PmpDauto domain eluted with 500mM imidazole under low stringency conditions where imidazole wash concentrations were kept below 50 mM to allow a sharp elution of rPmpDauto. However the Coomassie gel (**Figure 3.11**) shows the 500 mM imidazole fraction containing rPmpDauto contains several contaminant proteins. A step-by-step gradient of increasing imidazole at 50 mM, 60 mM, 75 mM, 100mM and 250mM in the elution buffers were tested in an attempt to remove the contaminant proteins but resulted in early elution of rPmpDauto at 50 mM imidazole (data not shown).
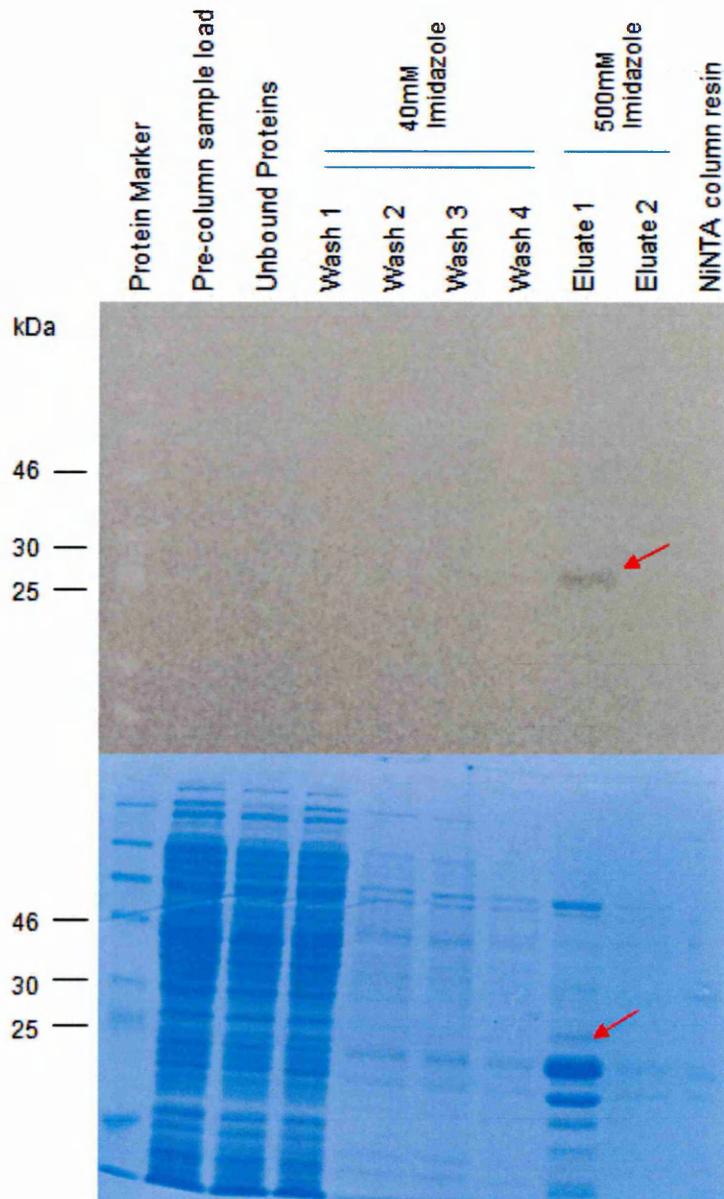
**Figure 3.11 - Western blot and Coomassie gel showing the fractions obtained during Ni-NTA purification of detergent solubilised rPmpD autotranslocator domain.** (Extracted from a 500 ml *E.coli* culture). Western blot analysis (top) indicated that specifically bound recombinant Pmp D autotranslocator of correct monomer size ~27 kDa (red arrow) eluted from the column with 500 mM imidazole. Traces of rPmp D autotranslocator domain, were also found in the fourth 40mM wash due to a low binding affinity. The Coomassie stained SDS gel (below) demonstrated the level of purity where many other non-specific proteins were present in the sample, rPmpDauto is depicted by red arrow. Downstream purification steps are necessary to remove these contaminants.

### 3.3.3 Expression trials of the putative β-barrel rPmp G autotranslocator domain (rPmpGauto) using E.coli strain Tuner™ (DE3) +/- pLysS

Expression trials were performed in order to produce a recombinant autotranslocator domain of Pmp G. Tuner™ (DE3) pLysS cells transformed with pPmpG22b-C were induced with 1.0mM IPTG overnight at 30°C as in **section 3.2.3.2**. Harvested cells were freeze/thawed and lysed in buffer and fractionated into soluble, detergent solubilised and inclusion bodies and examined on Western blot. Expression of the putative recombinant Pmp G autotranslocator domain (rPmpGauto) at approximately 40kDa was detected by Western blot (**Figure 3.12**) using the penta-His primary antibody. Unlike the homologue in Pmp D, the recombinant Pmp G autotranslocator domain was designed as a fusion protein with the pelB leader sequence on the N-terminus to direct the recombinant protein across the inner membrane of *E.coli*, where it may be incorporated into the lipid bilayer. This was not observed; instead insoluble aggregates of rPmpGauto appeared to be present in two bands, indicated by the red arrows at approximately 40 kDa, in the inclusion body fraction. The smaller band was thought to be a truncated fragment of rPmpGauto, possibly as a result of cleavage of the pelB leader peptide. The same results were observed using Tuner™ (DE3) cells without the pLysS plasmid (data not shown).

### 3.3.4 Expression screening of N-terminal rPmpA passenger domains using E.coli strain Tuner™ (DE3) +/- pLysS

Expression trials were performed in order to produce a recombinant passenger domain of Pmp A (rPmpA-N). The expression of this domain was not observed in *E.coli* Tuner™ (DE3) pLysS (data not shown), further expression trials of rPmpA-N was carried out using *E.coli* Tuner™ (DE3) cells lacking the pLysS plasmid to minimise regulation within the cells with no expression detected. The growth of the cultures was monitored and did not show the recombinant expression plasmid to have any deleterious effects on the cells however extensive screening trials were omitted using *E.coli* as an alternative expression system was being explored that could express rPmpA-N (see Chapter 4).
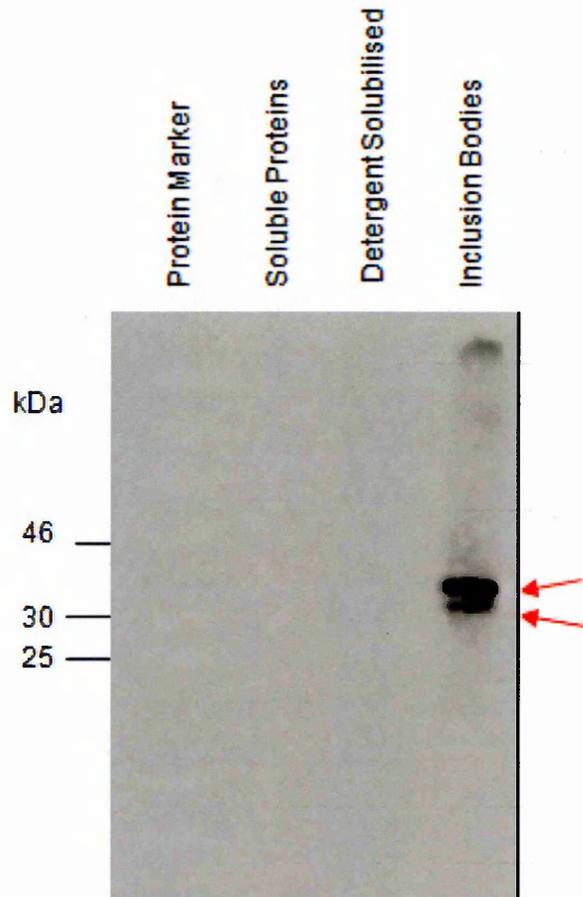
**Figure 3.12 - Western blot showing the expression of rPmp G autotranslocator domain in Tuner™ (DE3) pLysS**. *E.coli* transformed with pPmpG22b-C were induced with 1.0mM IPTG at 30 °C. The *E.coli* lysates were separated into soluble fractions, detergent solubilised proteins using 1% w/v DDM and inclusion bodies. An equal volume of each fraction was loaded into each lane as labelled. Putative rPmpGauto was detected at ~40 kDa, with an additional smaller band thought to be an fragment of rPmpGauto, depicted by red arrows.

### 3.3.5 Expression screening of N-terminal rPmp I domain using E.coli strain Tuner™ (DE3) pLysS

Expression trials were performed in order to produce a recombinant passenger domain of Pmp I. Tuner™ (DE3) pLysS cells transformed with pPmpI23b-N were induced with 0.5 mM IPTG overnight at 30°C (**section 3.2.3.2**). Harvested cells were freeze/thawed in lysis buffer and fractionated into soluble, detergent solubilised and inclusion bodies that were examined by Western blot. Detection using the penta-His primary antibody indicated the expression of the putative recombinant Pmp I passenger domain (rPmpI-N) at approximately 44kDa observed by Western blot (**Figure 3.13**). The protein was observed as inclusion bodies with some traces present in the soluble cytoplasmic fraction, the detergent solubilised membrane fraction and in the culture media. The rPmpI passenger domain has a native amino-terminal signal peptide sequence for inner membrane sec processing. It was expected that this native signal leader peptide would help direct recombinant protein to the *E.coli* membrane to be processed via sec machinery for translocation and possible secretion out of the cell as described in **section 1.4.6**. The predicted size of the passenger domain of Pmp I is approximately 46 kDa, recombinant Pmp I passenger domain was detected in some fractions as two bands at approximately 46 kDa and 40 kDa (**Figure 3.13**) which may have demonstrated proteolytic cleavage of the native signal peptide that is predicted to be approximately 6 kDa. In addition a third band of approximately 25 kDa was detected in the detergent solubilised membrane fraction, indicated by yellow arrow given in **Figure 3.13**. This 20 kDa band was not apparent in the soluble, inclusion body or TCA precipitated media fractions, and so appeared to be associated with the membranes.

### 3.3.5.1 Sodium azide effects on the expression of the N-terminal rPmpI domain with and without native signal peptide

To assess the processing of the passenger domain of rPmpI (rPmpI-N) further, expression cultures were treated with varying concentrations of sodium azide to inhibit sec processing (**section 3.2.3.3**). Inhibition of the sec machinery can prevent the translocation and subsequent cleavage of signal peptides from pro-proteins such as autotransporters which are predicted to be processed via the sec pathway. As a control, an expression construct lacking the signal peptide sequence was produced for rPmpI-N by digesting the recombinant expression plasmid pPmpI23b-FL with *BamHI* and *NotI* (**section 2.2.6**). Digesting at these sites removes the signal sequence and the newly digested DNA sequence was subcloned into a newly

digested pET23b vector (*BamHI* and *NotI*) (**section 2.2.9**). In the absence of the signal leader peptide sequence, the recombinant passenger domain was classed as a non-secretory protein using the bioinformatic prediction programs SignalP 3.0 and SignalP-HMM (Expasy Tools). The sequence with intact signal peptide rPmpI-N is predicted to be a secretory protein. It was expected that the passenger domain lacking the signal leader peptide would be restricted to the cytoplasm and therefore azide would have no effect on the accumulation of the protein.

The inclusion bodies of expressed recombinant Pmp I passenger domains with and without signal peptides were isolated for analysis by Western blot. In the presence of azide, an increase in the formation of inclusion bodies was observed for the expression of the Pmp I passenger domain with its native signal peptide (**Figure 3.14a**). Furthermore, this was in a concentration dependent manner. When the recombinant passenger domain was expressed without a signal peptide the domain accumulated as inclusion bodies without azide treatment as expected (**Figure 3.14b**). The growth rates of the cultures were monitored and showed that the high inclusion body production was not due to a higher cell density. In fact, azide slowed the growth rates in a concentration dependent manner, but did not appear to be directly lethal to the cells as growth was maintained and stable throughout all the cultures (data not shown). Trials were also performed with 2mM azide treatment but the growth rates and the isolated inclusion bodies were almost identical to the cultures treated with 1mM.
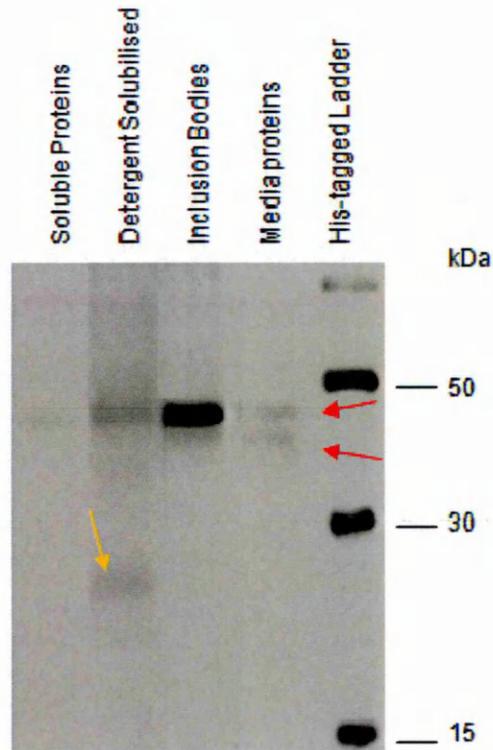
**Figure 3.13 - Western blot showing the expression of rPmp I passenger domain in Tuner™ (DE3) pLysS.** *E.coli* transformed with pPmpI23b-N were induced with 0.5mM IPTG at 30 °C. The *E.coli* lysates were separated into soluble fractions, detergent solubilised proteins using 1% w/v DDM and inclusion bodies and TCA precipitated proteins from the media. An equal volume of each fraction was loaded into each lane as labelled. Putative rPmpI-N was detected at ~44 kDa, with an additional smaller band thought to be a fragment of rPmpI-N, depicted by red arrows. A third band ~20 kDa was detected within the detergent solubilised membrane fraction, depicted by yellow arrow.
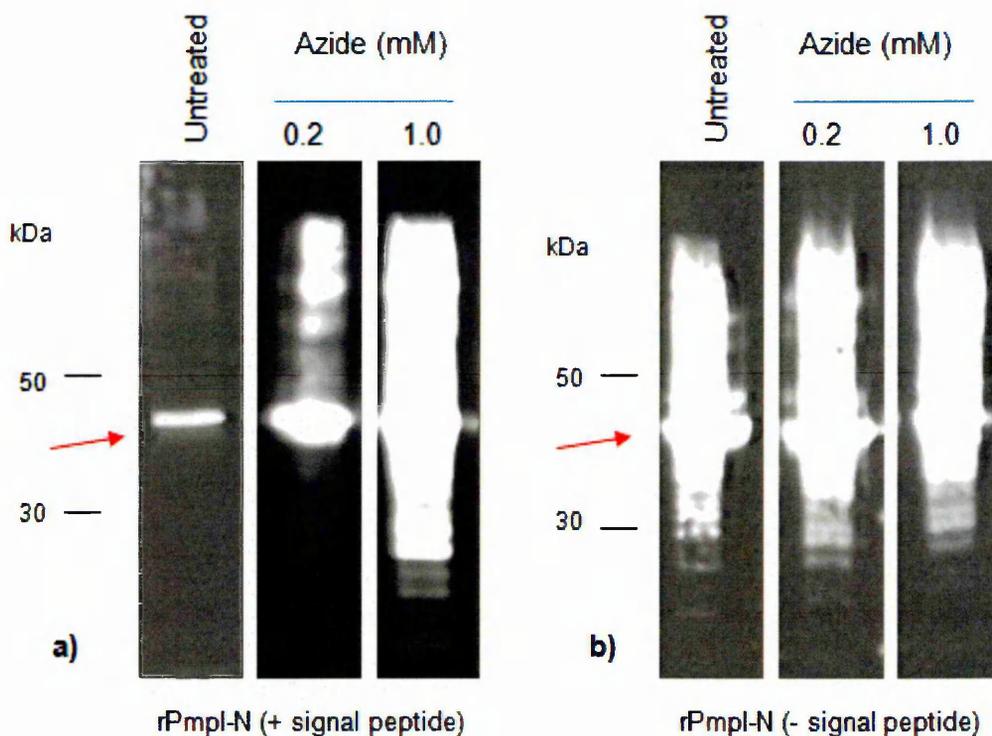
**Figure 3.14 - Western blots showing sodium azide effects on the expression of rPmp I passenger domain with and without native signal peptide sequence, in Tuner™ (DE3) pLysS**. The inclusion body fractions were isolated as in **section 3.2.3.5**. An equal volume of each fraction was loaded into each lane. (a) *E.coli* transformed with pPmpl23b-N was induced with 0.5mM IPTG at 30 °C and treated with sodium azide at varying concentrations (0.2 mM, 1.0 mM (2.0 mM not shown)). (b) *E.coli* transformed with pPmpl23b-N-signal (without signal peptide) were induced with 0.5mM IPTG at 30 °C and treated with sodium azide ((0.2 mM, 1.0 mM (2.0 mM not shown).

### 3.3.6 Western blot optimisation

During expression trials several commercial anti-His Abs were trialled to find a sensitive but specific antibody for use in Western blot to detect expressed recombinant His-tagged protein. Initially a mouse monoclonal anti-polyHistidine antibody used at a 1:3000 dilution (Sigma) was used both in conjunction with a rabbit anti-mouse IgG-HRP (Sigma) for chemiluminescence detection and a goat IRDye 800CW anti-mouse secondary antibody (Licor) for infrared fluorescence using an Odyssey Imaging System. These were tested with a positive control (Endol with an 8 residue poly-Histidine tag, obtained from Dr J.Hadden, University of Leeds) alongside the expression trial samples. The signal via these methods was very weak in comparison to protein load (data not shown) and therefore was deemed an unreliable method to detect putative expressed rPmps. At this point in the study it was unclear whether the absence of expressed full length rPmps in **section 3.3.1** was due to an insensitive detection technique with the monoclonal anti-polyHistidine antibody (Sigma) or due to a lack of expression. An authentic positive control (hexahistidine tagged recombinant Pmp) was not available at that point. Further trials used mouse monoclonal tetra-His and penta-His antibodies (Qiagen) with a commercial 6xHis protein ladder (Qiagen) as a positive control. Subsequently, after the successful expression of a detectable polyHis-tagged truncated Pmp (rPmpDauto **section 3.3.2**) the penta-His antibody was chosen for its high avidity and specificity to both the 6xHis ladder control and rPmpDauto. All rPmp expression trials from **section 3.3.1** were re-examined using this optimised technique with the penta-His antibody as described in **section 3.2.5.2** and full length rPmps were not present.

## 3.4 Discussion

### 3.4.1 Heterologous expression of full length recombinant Pmps

The first aim of this chapter was to express the full-length rPmps in *E.coli* of which several commercially adapted strains are available to optimise expression yields. Being a Gram-negative bacterium and displaying native autotransport proteins it was hoped that *E.coli* would be a suitable vehicle to express these proteins, possibly to the membrane for detergent solubilisation and purification for downstream assays although inclusion bodies were also anticipated. The trials carried out in an attempt to express full length rPmps are displayed in **Table 3.1**.

| Recombinant protein (Expression plasmid) | E. coli strain | Expression yields |
|---|---|---|
| **rPmpA** (pPmpA23b-FL) | Tuner™ (DE3) | Not detected |
| | Rosetta-gami 2™ (DE3) pLsyS | Not detected |
| | OverExpress C41™ (DE3) | Not detected |
| | OverExpress C43 ™(DE3) | Not detected |
| **rPmpD** (pPmpD28b-FL) | Tuner™ (DE3) | Not detected |
| | OverExpress C41™ (DE3) | Not detected |
| | OverExpress C43 ™(DE3) | Not detected |
| **rPmpI** (pPmpI23b-FL) | Tuner™ (DE3) | Not detected |
| | Rosetta-gami 2™ (DE3) pLsyS | Not detected |
| | OverExpress C41™ (DE3) | Not detected |
| | OverExpress C43 ™(DE3) | Not detected |

**Table 3.1 - Summary of *E. coli* strains used in full length rPmp expression trials.** Each construct was expressed at 37 C, 30 C and 25 C for short incubation periods (<3hrs) and prolonged incubation periods (<16hrs). His-tagged proteins were not detected.

Tuner™ (DE3) was the initial strain chosen to begin expression trials as this strain allows a high level of transcriptional control (Novagen). To this end, variable concentrations of IPTG can be employed to allow control over expression rates; an important consideration for expressing membrane proteins due to their low solubility. However expression of rPmpA, rPmpD or rPmpI was not observed using this strain of *E.coli* under many variable conditions including expression at different temperatures, with variable amounts of IPTG and extended expression periods.

Interestingly the growth rates of the cultures undergoing expression trials were indicative of a culture diverting its resources in an effort to transcribe the rPmps yet polyHistidine rPmps were not detected.

Although Pmps are from a Gram negative prokaryotic organism, they differ in that they would naturally reside encapsulated within the eukaryotic host and within a Gram-negative membrane that lacks peptidoglycan so the Tuner™ (DE3) strain may lack the necessary resources, environment and machinery to construct large heterologous membrane proteins such as Pmps. It is possible that the Pmps may have evolved to carry eukaryotic codon sequences like other intracellular bacteria that have evolved to carry eukaryotic-like coding sequences from lateral gene transfer such as *Legionella pneumophila* (Lurie-Weinberger *et al.*, 2010). When the mRNA of heterologous target genes are expressed in *E.coli*, differences in codon usage can impede translation due to the demand for one or more tRNAs that may be lacking leading to translational stalling, early translation termination, translational frame-shifting and amino acid mis-incorporation (Kane 1995). In view of this, codon analysis was performed on the full length *pmp* sequences and revealed several rare codons not commonly used by *E. coli*. One strain of *E.coli* that could potentially overcome this difficulty was the Rosetta™ strain. This strain and its derivatives contain a chloramphenicol-resistant plasmid that supplies the rare tRNAs thus has the ability to express proteins containing rare codons. In addition the Rosetta-gami 2™ (pLysS) strain is adapted to allow disulfide bond formation within the cytoplasm. Many proteins require the formation of stable disulfide bonds to ensure proper folding and thus activity (Prinz *et al.,* 1997) and this can be limited in *E.coli* due to the relatively high reducing potential in the cytoplasmic compartment. Considering the Pmps are relatively rich in cysteine residues, efforts to allow disulfide bond formation within the cytoplasm in an attempt to aid folding or allow rPmps to be chaperoned to prevent mis-folding was another reason to examine expression in this strain. To date, the folding of Pmps has not been explored and so it remains unclear where and what mechanisms are involved, however rPmp were not detected using this strain under all of the variable conditions trialled.

A further consideration for the lack of expression was plasmid instability through segregational loss of the plasmid to daughter cells. Importantly it is known that several autotransporters act as toxins and although it remains uncertain if Pmps function as toxins, it is possible that the rPmps or the *rpmp* genes could have been toxic to the *E.coli*, possibly at expression levels too low to detect by conventional

means. Interestingly the growth curves were not suggestive of rPmps having a toxic effect on the Tuner™ (DE3) cells as growth was maintained upon IPTG induction, however there were differences in the growth rates when expressing rPmpI-FL in the Rosetta-gami 2™ strain where the growth began to plateau upon the addition of IPTG induction. If expression was deleterious to the cell it is possible that the plasmid may have been lost by the Tuner™ (DE3) so the gene and its products could not be expressed. The presence of the pLysS plasmid within the Rosetta-gami 2™ cell line works to prevent basal expression of the target protein by inhibiting T7 RNA polymerase, therefore putative toxic effects should only be demonstrated upon IPTG induction as was observed. Although this was purely speculative, choosing a strain shown to confer resistance to toxic expression was an appropriate choice to overcome such difficulties. Futhermore expression trials with the plasmid pPmpI23b-FL, using an *E.coli* derived cell-free protein expression system that utilises nanolipoproteins to mimic the lipid bilyer (**section 4.1.1**), produced a ~80 kDa His-tagged protein; a size expectant of rPmpI-FL indicating that transcription and translation of the gene was viable (**chapter 4**).

OverExpress™ C43 (DE3) and Overexpress C41™ (DE3) were developed to allow high level expression of a wide variety of toxic proteins previously difficult or impossible to express in bacteria. Both strains are reported to be very effective in expressing complex membrane proteins from all classes of organisms including viruses, bacteria, yeasts, plants, insects and mammals (Masi *et al.,* 2003, Drory *et al.,* 2004, Pos *et al.,* 2004, Thai *et al.,* 2009). In particular, the OverExpress™ C43 (DE3) have been developed to allow high level expression of a wide variety of toxic proteins previously difficult or impossible to express in bacteria (Miroux and Walker 1996).The strain derived from BL21 (DE3) contains genetic mutations phenotypically selected for conferring toxicity tolerance although the mutation responsible is uncharacterised. Dumon-Seignovert et al., (2004) showed that C43 (DE3) presented only 4 per cent expression-induced toxicity compared to 96 per cent observed in BL21 (DE3) cells so became a favourable choice for screening trials. Yet after numerous expression trials with variable conditions, rPmps were not detected using this strain.

Other possible explanations for the lack of expression of full length rPmps could be due to the size of the expression plasmids. The recombinant plasmids transformed into *E.coli* for expression were relatively large constructs, each over 10kb in size. There are correlations between plasmid size and instability; the larger the plasmid

the higher the propensity for intramolecular recombination or genome integration (Oliveira *et al.,* 2009). In addition such plasmids also have an increased metabolic burden on the cell.

In summary, full length recombinant Pmps were not detected when expressed in a variety of *E.coli* expression hosts. Early indications implied that induced expression of certain rPmps may have had a lethal effect on certain *E.coli* hosts however there is no published data to support this and performing trials in a strain that confers high tolerance to toxic proteins did not substantiate this, nor did it improve detectable expression. A major drawback throughout these early trials was the reliability of the detection technique. Until a suitable sensitive and specific primary antibody was determined, re-examination of the expression samples obtained from the full length Pmp expression trials with an optimised detection technique revealed no discrepancies from the original findings. Using the assumption that Pmps were expressed in poor yields below detectable levels, methods to concentrate the protein using IMAC were explored and putative rPmp-FL remained undected. A further consideration was that Pmps, may have undergone proteolytic cleavage during expression into their functional autotransporter domains before, during and/or after inner membrane translocation. If this had occurred it may have hindered detection of the protein since the C-terminus of the rPmps boast a hexahistine tag, both necessary for Ni-NTA chromatography and penta-His probing. Method to inhib processing using sodium azide did not result in the accumulation of rPmps. Finally it was concluded that the over-expression of such large complex membrane proteins was beyond the scope of *E.coli's* capabilities and an alternative approach to produce rPmps was necessary.

### 3.4.2 Heterologous expression of truncated Pmps

As full length rPmp expression was unsuccessful in *E.coli,* focus on expressing the smaller functional domains of these proteins was considered. As autotransporters are large proteins made up of three distinct domains described in **section 1.4.6** (Henderson *et al.,* 1998) other autotransporter proteins have been studied this way with success. In particular the IgAP protease from *Neisseria gonorrhoeae*, an organism that also infects the genital epithelial cells has had its autotranslocator domain structure determined by expressing the C-terminal functional region (Veiga *et al.,* 2002) where the 25-30kDa C-terminal region builds the translocation unit needed to export the passenger domain. In addition the N-terminal passenger domain of the Hia autotransporter from *Haemophilus influenza* has been expressed

for use in adherence assays (Laarmann *et al.,* 2002). Expression screening trials to produce such regions of *C.trachomatis* Pmps in this study have led to the successful expression of several truncated rPmp functional domains in *E.coli* (**Table 3.2**).

| Recombinant protein (Expression plasmid) | E. coli strain | Expression yields |
|---|---|---|
| **rPmpA-N** (pPmpA23b-N) | Tuner™ (DE3) Tuner™ (DE3) pLysS | Not detected Not detected |
| **rPmpDauto** (pPmpD28b-C) | Tuner™ (DE3) Tuner™ (DE3) pLysS | Inclusion bodies Mainly inclusion bodies with some targeted to membranes after optimisation to reduce expression rates significantly (**section 3.3.2**) |
| **rPmpGauto** (pPmpG22b-C) | Tuner™ (DE3) Tuner™ (DE3) pLysS | Inclusion bodies Inclusion bodies |
| **rPmpI-N** (pPmpI23b-N) | Tuner™ (DE3) pLysS | Inclusion bodies and putative soluble cleaved products |

**Table 3.2 - Summary of outcome for the expression trials of recombinant truncated Pmps.** The autotranslocator domain of rPmpDauto was the first to be expressed and be positively identified. Extensive efforts were made to optimise the expression to target protein to the membranes. Due to the difficulties, the other rPmp domains were not followed up with optimisation trials.

However expression of truncated rPmps was mostly in the form of inclusion bodies, a common problem when expressing membrane proteins. Inclusion bodies are dense particles of mis-folded proteins (Lilie *et al.,* 1998) formed intracellularly, although the mechanisms behind the aggregation are not fully understood, it is believed they form as part of a stress response when recombinant protein is expressed at a high rate. Because inclusion bodies are often resistant to proteolysis and contain a high percentage of over-expressed protein they can sometimes be exploited for purification and are often used in favour of proteins that are unstable, toxic or easy to refold (Baneyx and Mujacic, 2004). Still, in order to achieve this they must first be solubilised in strong denaturants to disrupt aggregation. In order to study the structure and function of the protein it is then necessary to refold by renaturing the proteins via a titration method (Vallejo and Rinas, 2004). 8M urea has already been shown incapable of denaturing rPmpDauto inclusion bodies in this study, therefore a stronger alternative would be needed such as 6M guanidine hydrochloride (GuHCl). However, refolding can lead to conformational changes different from the native structure particularly where disulfide bonds are present therefore denaturation and refolding is not a reliable method for production of functionally active protein. Optimisation of such a method also requires a control assay to demonstrate function which is not possible in this study. Therefore a more suitable approach to produce a homogenous, natively folded, active protein suitable for structural and functional analysis was to prevent target protein aggregation to such a level that allows target protein to localise to the membranes. Disruption of the membranes with detergents would allow for solubilisation of the rPmp domain, in a correctly folded and active state, into detergent micelles for purification. The optimisation trials in this study were time consuming and focused on rPmpDauto.

Early trials to express the rPmp D autotranslocator domain in this study were under the control of the T7*lac* promoter only, a tight regulatory promoter that should have prevented basal expression (**section 3.1.1**) however the inhibitory effect of this mechanism was not observed during the expression and the propensity to form inclusion bodies was extremely high. Furthermore, Tuner™ (DE3) cells permit uniform entry of IPTG into the cell to allow control over expression levels however, the expression observed in the uninduced cells was comparable to the expression seen in the IPTG induced cells and indicated expression was affected by other factors. cAMP can directly affect the level of transcription (Grossman *et al.,* 1998). cAMP forms complexes with the cAMP receptor protein (CRP) and this cAMP/CRP complex binds upstream of the *lacUV5* promoter, directly stimulating transcription of

T7 RNA polymerase. Such levels are strongly influenced by the carbon source available to the cells and one approach to counteract this and increase protein solubility is to perform expression trials in media supplemented with glucose. When glucose is absent, the cell is forced to use an alternative carbon source, such as glycerol that causes cAMP levels to rise. The addition of glucose acts as an alternative carbon source, keeping cAMP levels low within the cell which in turn inhibits the production of T7 RNA polymerase and expression of the target gene until IPTG is added, a term referred to as catabolite repression (Zhang *et al.,* 2003). Subsequently the addition of glucose within cultures reduced the basal expression thus reducing the amount of inclusion bodies, but did not avert them.

Often molecular chaperones are necessary to assist with the folding, especially of large multidomain and over-expressed proteins (Baneyx and Mujacic, 2004). In *E.coli*, where transcription and translation are tightly coupled, one polypeptide chain is released from the ribosome every 35 seconds (Lorimer, 1996), which can lead to the saturation of the cellular chaperones and the translocation machineries. Consequently failure to interact with folding modulators has two possible consequences, either degradation or deposition into inclusion bodies (Middelberg, 2002). Chen et al., (2002) described a simple method for enhancing soluble protein yields using heat shock prior to inducing protein expression. By exposing the culture to a temperature of 42 °C for 1 minute prior to induction increased the solubility of the target protein significantly with no deleterious effects upon yield. Although the mechanism of increased solubility is not fully understood, the likelihood is heat-shock proteins aid correct folding. Previous studies have shown that co-expression of heat-shock proteins with target proteins can enhance proper protein folding, since heat-shock proteins, such as GroEL, GroES and DnaK, function as molecular chaperones to prevent misfolding (Gragerov *et al.,* 1992; Goenka and Rao, 2001). In view of this expression trials for rPmpDauto were performed using heat shock methods to enhance folding modulators. However rPmpDauto remained in the form of inclusion bodies.

Efforts to decrease inclusion body formation in conjunction with increasing expression at the membrane can be extremely cumbersome if little is known about how the membrane protein is targeted or the mechanisms involved in folding. Yet it is largely accepted that lowering expression temperatures decreases aggregation and this was fundamental to the expression of truncated rPmp domains in these trials. Previous studies have explored this, Ami et al., (2005) investigated the folding

kinetics of *Pseudomonas fragi* lipase, a cold-adapted enzyme susceptible to aggregation when expressed even at moderate temperatures and found that expression into inclusion bodies was decreased with a lower expression temperature and subsequent soluble protein levels were increased compared to expression at 37 °C. Further support in the form of a kinetic model proposed by (Kiefhaber *et al.,* 1991) showed that the yield of native protein depends only on the rate of folding, the rate of aggregation and the rate of synthesis. Therefore the yield of native protein increases, and inclusion body formation decreases, with a lowered rate of protein expression. This correlates with other studies that have explored expression under reduced growth conditions including decreased growth temperature, using specific culture media and limiting the induction to produce increased yields of native, soluble protein (Schein and Noteborn, 1988; Kopetzki *et al.,* 1989). Although these studies were largely based on soluble proteins, implementing these adaptable parameters were considered to target the rPmpDauto domain to the membranes of the host, possibly by limiting the inclusion body formation.

Consequently the pLysS plasmid was employed, the pLysS encodes for T7 lysozyme that breaks down T7 RNA polymerase necessary for expression. Lowering expression temperatures with the addition of the pLysS plasmid significantly reduced basal expression and prevented early formation of inclusion bodies however upon induction the inclusion bodies were formed. Unsurprising when under control of the strong pET promoters, induction of recombinant protein can exceed 50% of the total cellular protein leading to the titration of other endogenous proteins (Lorimer, 1996). Because IPTG induction previously resulted in a high level of unfavourable inclusion body formation IPTG independent expression was trialled with success. The expression of membrane-localised rPmpDauto appeared to be facilitated by slow (overnight), IPTG independent expression under the tight control of the pLysS plasmid at low incubation temperatures (25 °C) a term that could be referred to as 'restrained expression'. In theory, one would expect to observe an absence of recombinant rPmpDauto with these restrictions in place but these conditions appeared to provide an optimal environment for incorporation of the recombinant membrane protein into the membrane probably due to the folding rate kinetics of the protein, avoiding both chaperone and membrane saturation. This would support Kiefhaber's *et al.*, (1991) kinetic model where the native protein yield depends upon the rate of folding, the rate of aggregation and the rate of synthesis. It is likely that rPmpDauto domain

154

produced in these conditions is in a native thus active conformation suitable for purification and downstream biophysical studies, although the yield of native protein was very low. Further support from a recent study that showed the omission of IPTG induction led to higher yields of toxic membrane proteins (Narayanan *et al.,* 2011). The absence of the inducer IPTG, led to improved expression levels of cation/$Ca^{2+}$ exchangers as the expression levels did not have a deleterious effect on the cells. The low level expression occurs due to the regulatory mechanism whereby the equilibrium between free and bound states of the *lac* operator and *lac* repressor control the expression.

Even when expression of rPmpDauto was directed to the membrane, challenges remained. Membrane proteins embedded within a complex and dynamic lipid bilyer, have to be isolated as the use of many biophysical techniques is limited (Seddon *et al.,* 2004). Therefore following successful membrane protein expression, a suitable detergent had to be found that simultaneously extracts the protein from the membrane while retaining the native fold and function. Detergents can provide the protein with a favourable environment as the hydrophobic portions of the detergent molecules adsorb onto the apolar surface of the protein, forming a protein-detergent complex that is soluble in aqueous solvents (Gordon *et al.,* 2008). However this protein-detergent complex is often heterogeneous itself creating numerous problems with purification and downstream functional assays (Savage *et al.,* 2009). Furthermore poor stability facilitating solubilisation is common and complicated the handling and purification of rPmpDauto. Subsequently the purification of rPmpDauto by nickel chelate affinity chromatography faced new challenges. It was difficult due to the low binding affinity to the Ni-NTA column, the poor yield from expression, the heterogeneity of the sample and the tendency to precipitate during handling in detergents. Alternative strategies to facilitate purification by batch binding and altering the buffer conditions did not improve the outcome. It is possible that rPmpDauto may not have had the polyHistidine tag exposed for binding but denaturing the protein to unfold the His-tag fully was not feasible as a biologically active protein sample was necessary for downstream analysis methods. Consequently the partial purification that was achieved was unsuitable for circular dichroism so secondary structural analysis could not be carried out.

Several further downstream purification methods could have been carried out to improve the purity, such as ion exchange chromatography, however, as rPmpDauto was only present in a low concentrations beyond the reliable limits of detection, this

155

was deemed unfeasible due to the substantial losses that would be associated with further purification methods. A further problem was that the purified rPmpDauto would precipitate out of solution during storage and samples did not recover well from storage at 4 °C, -20 °C or -80 °C, meaning that each purification trial had to be carried out from freshly expressed and processed samples, a long process when expression was carried out slowly. Additional purification steps would prolong the handling time and require buffer-exchange, which would likely exacerbate precipitation and the loss of soluble protein. This problem in combination with poor starting yields that were already difficult to detect, made further handling of the protein very problematic and further purification could not be considered. However trials continued exploring the expression of other truncated Pmp domains in this way. The expression, targeting, recovery and purification of rPmpDauto was a very time consuming process and is an example of the high attrition rates associated with producing recombinant membrane proteins.

The autotranslocator domain from Pmp G was expressed as a fusion with a pelB leader peptide signal sequence to aid localisation to the inner membrane of *E.coli*, an approach also used to express the autotranslocator domain of IgAP by Veiga *et al.,* (2002). Although expression was successful with two His-tagged fragments detected, indicating possible proteolytic cleavage of the signal peptide, this was also in the form of inclusion bodies. Due to problems encountered in trying to produce a purified rPmpDauto domain, it was deemed an uneconomical use of time to pursue the expression of rPmpGauto in the same way.

In addition to expressing the C-terminal autotranslocator domains of the Pmps, expression of the functional N-terminal passenger domains was also explored, in particular those of Pmp A and Pmp I. The N-terminal region contains the signal peptide leader sequence that targets the protein to the Sec apparatus in the inner membrane (Desvaux *et al.,* 2004a). This peptide is cleaved by components of the Sec machinery to release the pro-protein into the periplasm. Adjacent to the 'signal' domain is the functional 'passenger' domain. This region offers the most sequence diversity among the proteins as this performs the extracellular function. This passenger domain may remain intact but exposed from the cell surface or may be cleaved into the extracellular milieu to carry out its function.

Interestingly the expression of the recombinant passenger domain of Pmp I in *E.coli* was detected in the soluble and insoluble fractions including the cytoplasm, the media and solubilised from the membranes. Passenger domains of autotransporters

are often presented externally of the cell surface as well as within the membrane so it was anticipated that these domains may be less hydrophobic in nature compared to their C-terminal membrane-bound counterparts. Purification trials were not attempted with the passenger domains as a novel cell-free expression system was being explored (chapter 4) with a more efficient handling time to express and purify these domains.

However, as the expression of the PmpI-N passenger domain showed possible proteolytic cleavage consistent with cleaved autotransporters, these putative proteolytic cleavage events were investigated further in an attempt to examine the processing of the domain where three His-tagged bands had been consistently detected upon the expression of rPmpI-N. These bands may imply sec processing of the signal peptide and further cleavage within the functional region of the passenger domain of Pmp I was occurring. Autotransporter proteins can undergo proteolytic cleavage within the passenger domain upon localisation to the membrane or external milieu (Henderson *et al.,* 2004). Although the mechanisms behind this cleavage event are not fully understood, interactions with components of the outer membrane bilayer or the extracellular milieu are thought to trigger this and so the smallest band detected in the membrane fraction was suggestive of further processing of the rPmp I passenger domain.

Inhibition of the Sec machinery can prevent this translocation and subsequent cleavage of signal peptides from autotransporters and other pro-proteins. Consequently, unprocessed pro-proteins that are restricted from Sec processing accumulate in the cytoplasm as inclusion bodies. Studies to inhibit the Sec apparatus with sodium azide, a known inhibitor of SecA (Oliver *et al.,* 1990) resulted in an increase of inclusion body formation of the Pmp I passenger domain within the cell. In a similar study Gathmann *et al.,* (2006) found that expression of plastocyanine (a soluble pre-apo-protein with an N-terminal signal sequence targeted to the thykaloid lumen and periplasm of *Cyanobacterium*), also formed highly enriched inclusion bodies when expressed in the presence of azide. Expressing the same domain without a native signal peptide also serves to inhibit targeting to the Sec machinery and also led to the aggregation of this domain as inclusion bodies.

Azide has the same effect as removal of the leader/signal sequence and so these data would indicate that the sec translocon is necessary for the processing of rPmpI, under direction of the signal peptide. However to demonstrate this

conclusively would need considerable further study beyond the scope of this project. One particular consideration is that azide is not specific only to secA but also affects many other proteins and pathways within the cell. To rule out other azide disrupted cellular processes as the cause for increased inclusion body formation, rPmpI-N would have to be expressed in an azide resistant (DE3) lysogen. Consequently this could help elucidate the theory that Pmps are in fact autotransporters as bioinformatic studies predict. The passenger domain of Pmp A, which differs by not possessing a native signal peptide compared to other Pmps within the family, was not expressed in *E.coli*.

## 3.5 Summary

*E. coli* were not able to express the large multidomain full length Pmps or at least not at levels that could be detected. Where successful expression was observed, for three truncated rPmp domains, the expression was very 'leaky' and conditions implemented to modulate this, in particular to minimise basal expression and prevent inclusion body formation were met with some success. The host cells were successful in that the formation of inclusion bodies could be reduced to allow some protein targeting to the membranes however the downstream purification that followed was difficult due the heterogeneity of the expression sample and the poor yield of recombinant protein solubilised from the membranes which became increasingly difficult to detect following each downstream application. This was further hampered by the necessity to express fresh recombinant proteins for each individual optimisation step as the protein became insoluble in detergents.

However this was a positive outcome as it remains impossible to predict whether a recombinant membrane protein can be expressed in useful quantities, if expression is possible it is not known if it will be targeted to the membrane or as misfolded aggregates in inclusion bodies. It remains unsurprising that over-expression of membrane proteins is challenging. Unlike soluble cytoplasmic proteins, membrane proteins do not meet favourable conditions upon displacement from the ribosome unless the protein is targeted and inserted within a lipid environment where secretory machinery and/or chaperones are necessary in order to achieve this correct folding and targeting. Determining the most suitable conditions for expression of membrane proteins can be a laborious task. Currently there is no ideal over-expression system that exists for the production of such proteins and so the parameters for each target protein have to be met individually as the expression of almost identical proteins can vary significantly. Therefore, heterologous over-expression of membrane proteins in *E. coli* is largely accepted as a matter of 'trial and error' often time consuming and futile. As well as inclusion body formation, the production of membrane proteins can often result in low protein yield, insufficient for purification. These challenges were observed with the expression of recombinant full length Pmps and truncated Pmps and *E. coli* as a host had many deficiencies.

Alternatively homology screening can be a powerful approach, and is the centre of many structural genomics initiatives. Homologues provide natural variation that may be less toxic, better expressed and more stable (Mancia and Love 2010). However Pmps are unique to *Chlamydia* so this was not an option.

A change of expression host was considered but ultimately in view of the time involved in establishing these techniques, developing new cloning strategies and repeating the time-consuming expression trials already carried out in *E. coli* for each rPmp and its respective domains, this approach was thought to be an uneconomical use of time. A reasonable alternative to designing and constructing new expression systems was to change the expression host to one also compatible with the T7-based expression system. As cloning strategies had already led to the successful expression of several smaller rPmps domains in *E. coli*, an *E. coli* based cell-free expression system with a composition less complex than viable *E.coli* was chosen for its rapid screening ability, toxicity tolerance, and ease of handling (chapter 4). The fact that successful expression was observed for three Pmp domains meant that such methods could be improved upon to obtain purified rPmps using a novel cell-free expression system.

# Chapter 4 - Expression of *C. trachomatis* Recombinant Pmps in a Cell-free Expression System

## 4.1 Introduction

The heterologous expression of recombinant Pmps and Pmp domains was problematic using *E. coli* as an expression host (chapter 3). To overcome some of the difficulties, an alternative, cell-free (CF) protein expression system was considered.

*In vitro* protein expression systems emerged based on the early demonstration that cell integrity is not a requirement for protein synthesis, such that translation can be accomplished using a crude lysates containing the translational machinery, enzymes and tRNAs in a coupled or uncoupled system (Katzen *et al.,* 2005). The *in vitro* 'coupled' transcription/translation system transcribes a DNA template *in situ*, as opposed to an 'uncoupled' system that requires RNA template to be added for translation. The most popular CF protein expression systems are derived from *E. coli* lysates although alternative systems originate from wheat germ extracts and rabbit reticulocytes. In theory, any organism could potentially be used as a source for the preparation of a CF expression system. Typically *E. coli* based systems are better for bacterial genes and provide higher yields at milligram levels of protein per millilitre of reaction, are more homogenous, and have been directly used for structural analysis of the product (Kigawa *et al.,* 1999).

Since protein expression processes are not dependent on cell viability, cell-free translation systems have become increasingly recognised as viable alternatives to overcome the obstacles observed during *in vivo* over-expression. This method has proved particularly useful for expressing proteins that are cytotoxic or impossible to express within a heterologous cellular environment, for example the Cytolethal Distending Toxin (CTD) of *Helicobacter hepaticus* was expressed as recombinant protein *in vitro*, where previous efforts to examine the activity of the toxin were limited due the difficulties in producing sufficient quantities using *in vivo* expression systems (Avenaud *et al.,* 2004).

Moreover, membrane proteins which are notoriously difficult to express and can also pose a lethal effect on the cell have also been produced at high levels using *E.coli* derived cell-free expression systems (Klammt *et al.,* 2004). Studies by Savage et al, (2009), found the *in vitro* CF system to be more robust than the *E.coli in vivo* system in over-expressing over one hundred *E. coli* membrane proteins. Upon comparison, 63% of the target MPs were expressed *in vitro* compared to 44% expressed *in vivo*. This was interesting considering the CF system has similarities to *in vivo* mechanisms, having the same transcriptional, translational and presumably

162

translocational components. Unsurprisingly over the past decade, due to the open nature of the system in that it offers flexibility to manipulate the reaction conditions, the field has expanded greatly resulting in significant improvements to the yield and solubility of membrane protein expressed this way. These manipulations come in the form of accessory reagents that can be directly added to a CF expression reaction, including lipids, folding catalysts and other amphiphilic molecules potentially useful in providing a more favourable environment for membrane protein expression (Katzen *et al.,* 2005). Such advances in this area have led to the publication of the first crystal structure of a membrane protein, EmrE a small multidrug transporter expressed using an *E.coli*-derived cell-free expression system Chen et al, (2007). It was found that the *in vitro* expressed EmrE, displayed essentially identical biochemical behaviour to *in vivo* expressed EmrE. One commercially available cell-free expression kit has been developed by Invitrogen and has shown a high success rate for producing tens or hundreds of milligrams of active recombinant membrane proteins, this system is termed the MembraneMax™ cell-free expression system (Katzen *et al.,* 2008).

### 4.1.1 MembraneMax™

MembraneMax™ is an *in vitro* coupled cell-free expression system developed by Invitrogen Life Technologies, that utilises nanolipoprotein particles (NLPs) or nanodiscs as a vessel for membrane protein insertion (**Figure 4.1**). Nanodiscs are self-assembled, soluble and stable complexes comprised of lipids encircled by an amphipathic 'scaffold' protein belt to produce a discoidal planar lipid bilayer (Borch *et al.,* 2008). In general, the scaffold proteins of nanodiscs are expressed by synthetic genes designed from a truncated form of the naturally occurring protein Apolipoprotein A-I and the length of the scaffold dictates the diameter of the disc. NLPs have an approximate average height of 5 nm with diameters varying between 10-60 nm (Cappuccio *et al.,* 2008). The NLPs used in the MembraneMax™ system are approximately 10 nm in diameter. The development of the MembraneMax™ system stemmed from several studies (Bayburt *et al.,* 2002a, Bayburt and Sligar 2002b, Civjan *et al.,* 2003, Leitz *et al.,* 2006, Bayburt *et al.,* 2007, Bayburt and Sligar 2009, Bayburt and Sligar 2010) that showed successful incorporation of integral membrane proteins into NLPs as an alternative tool to using liposomes or detergent micelles to 'solubilise' and aid the handling of membrane proteins. Katzen et al., (2008) reported that NLPs present within the cell-free reaction dramatically increased the solubility of a multitude of biologically important membrane proteins,

suggestive that nanodiscs provide a suitable artificial membrane lipid bilayer to mimic the cellular environment. Importantly, analysis on a subset of these membrane proteins indicated that the proteins were biologically active, correctly folded and were embedded within the NLPs.

Since *E.coli* proved to be a problematic host for the expression of serovar E full length recombinant Pmps, with no expression detected and with some suggestions that certain expression plasmids could induce deleterious effect on some of the *E.coli* strains, a cell-free approach appeared to be an optimal choice for overcoming these challenges. Furthermore, the cell-free system offers major benefits in overcoming the challenges in downstream purification by reducing both the heterogeneity of the samples and the handling time, as recombinant protein can be achieved in less than two hours without the need for storage between processes. However these cell-free expression systems are very expensive compared to more cost-effective methods such as *E. coli*.
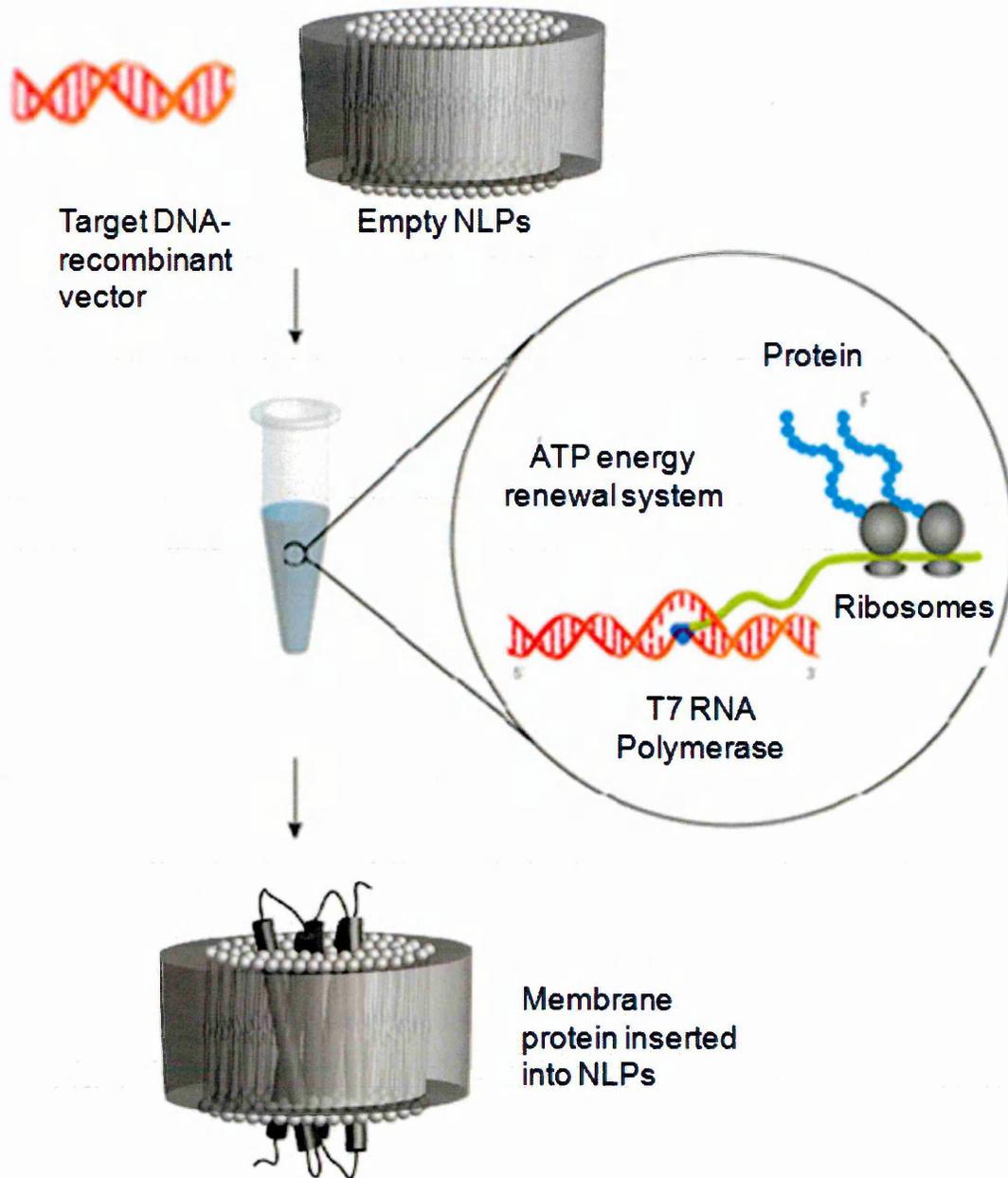
**Figure 4.1 - Diagram depicting the MembraneMax™ cell-free protein expression reaction**. The vector containing the gene of interest is added with the MembraneMax™ Nanolipoprotein (NLP) discs to a reaction containing *E.coli* lysate, an ATP energy regeneration buffer and T7 RNA polymerase. The *E.coli* lysate provides the transcriptional and translational machinery necessary to produce the protein of interest which can be localised into the NLP discs to produce high yields of soluble membrane proteins, using a simple and scalable method. Adapted from Invitrogen Flier (2008) Invitrogen.com.

## 4.2 Materials/Methods

Analytical grade chemicals were used throughout this study with all solutions being prepared using sterile water. Materials were obtained from the following suppliers.

**Bio-Rad Laboratories Ltd.**, Hemel Hemsted, Hertfordshire, **GE Healthcare Ltd**, Amersham, Buckinghamshire, **Generon.**, Maidenhead, Berkshire, **Invitrogen Life Technologies.**, Paisley, **New England Biolabs (UK) Ltd.**, Hitchin, Hertfordshire, **Novagen, Merck Chemicals Ltd.**, Beeston, Nottingham, , **Qiagen Ltd.**, Crawley, West Sussex, **Sigma Aldrich Company Ltd.**, Gillingham, Dorset, **Starlab UK.**, Blakelands, Milton Keynes, **Thermo-Fisher Scientific UK.**, Loughborough, Leicestershire,

## 4.2.1 Protein expression, MembraneMax™ Expression kit (Invitrogen)

100 μl reaction samples were set up according to the manufacturer's instructions (**Table 4.1**) with the appropriate expression plasmid. A negative control reaction was set up using the corresponding empty pET vector. The reactions were incubated at 37 °C, 250rpm in an orbital shaker for 2 hours. ATP stocks were replenished after a 30 minute interval with feed buffer (**Table 4.2**). Samples were placed on ice and separated into soluble supernatant (containing NLPS) and insoluble pellet fractions (containing aggregated proteins) by centrifugation at 17,000$g$ at 4 °C for 15 minutes according to the manufacturer's instructions, in the presence of protease inhibitors (1X HALT™ EDTA-free cocktail, Thermo-scientific). Soluble and insoluble fractions were resuspended in SDS-buffer for analysis by SDS-PAGE (**section 3.2.5.1**) and Western blot (**section 3.2.5.2**). Insoluble proteins were not used in any downstream assays. These reactions were scaled up where necessary.

## 4.2.2 Detergent solubilisation of recombinant proteins from NLPs for purification trials

The soluble fractions containing NLP-recombinant protein complexes from cell-free expression reactions were resuspended in 2X binding buffer (**section 4.2.3**) with gentle agitation, end over end for one hour with 1% (v/v) detergent (Triton-X100, Octyl glucoside or Dodecyl maltoside) to release the recombinant Pmp from the NLP disc. Samples were fractionated into detergent solubilised (supernatant) and detergent insoluble pellet fractions by centrifugation at 40,000$g$ for 20 minutes at 4 °C. Detergent solubilised samples were passed through a 0.22 μM filter (Millipore)

to remove any precipitated protein before undergoing purification as **section 4.2.3** or **section 4.2.4.**

| Reagent | Protein Synthesis Reaction |
|---|---|
| *E. coli* Extract | 20 µl |
| 2.5X IVPS Reaction Buffer | 20 µl |
| 50 mM Amino acids (-Met) | 1.25 µl |
| 75 mM Methionine | 1 µl |
| MembraneMax™ NLPs | 2 µl |
| T7 Enzyme Mix | 1 µl |
| DNA Template | 1 µg |
| DNase/RNase-Free Water | To final volume 50 µl |

**Table 4.1 - MembraneMax™ cell-free expression set-up conditions**. Reactions were set up for protein expression with expression plasmids or empty pET plasmids.

| Reagent | Protein Synthesis Reaction |
|---|---|
| 2.5X IVPS Feed Buffer | 25 µl |
| 50 mM Amino acids (-Met) | 1.25 µl |
| 75 mM Methionine | 1 µl |
| DNase/RNase-Free Water | To final volume 50 µl |

**Table 4.2 - MembraneMax™ cell-free expression buffer feed set-up conditions**. The feed buffer was added to the reaction in Table 4.2, 30 minutes into the reaction to replenish ATP and amino acid stocks to make a total reaction of 100 µl.

### 4.2.3 Affinity column chromatography of recombinant protein with Ni-NTA agarose

Soluble fractions containing NLPs (from 200 µl of cell-free expression reactions) were resuspended in 2X binding buffer to make a final concentration of (1XPBS, 10 mM imidazole, 1X HALT EDTA-free protease inhibitor cocktail) with or without 1% (v/v) detergent (Triton-X100 or DDM). Where detergent extraction was used, the sample was fractionated again into 'detergent solubilised' to release rPmps from NLPs and 'detergent insoluble' fractions as in **section 4.2.2**. The material that was not solubilised by detergents was examined but was not used in purification as this contained aggregated material. Soluble and detergent solubilised samples were passed through a 0.22 µM filter (Millipore) to remove any precipitated protein before being applied to a 2 ml column containing 250 µl of pre-washed 50% Ni-NTA agarose slurry (Qiagen). The flow through was collected and reapplied 3 times (before reserving as unbound fraction). The column was washed with 200 µl wash buffer twice (1XPBS, 20 mM imidazole, 0.1% detergent). Protein was eluted sharply with imidazole (typically 250 mM), using a 200 µl volume to minimise dilution of the eluted protein. Aliquots of all fractions were examined by SDS-PAGE and western blot for the presence of protein. Where batch-binding methods were used the soluble and/or detergent solubilised fraction was incubated with 250 µl pre-washed 50% Ni-NTA agarose slurry by gentle mixing of the tube, end over end at 4 °C for 1 hour prior to loading on a 2 ml column and collecting the unbound fraction. Larger expression reactions were scaled up accordingly.

### 4.2.4 Small Scale Ion Exchange Chromatography

Prior to chromatographic analysis, soluble fractions (containing NLP-rPmp expression from **section 4.2.1)** or Ni-NTA purified samples from **section 4.2.2** were buffer-exchanged into a buffer of appropriate pH and ionic strength (typically 50 mM NaCl) by washing in a Vivaspin 500, 10 MWCO concentrator (Generon). Due to sample precipitation this step was later replaced with dialysis overnight at 4 °C using GeBAflex dialysis tubes (Generon). For anion exchange chromatography using Q sepharose (Amersham Biosciences), a 2X buffer to make a final concentration of 50 mM Tris, 50 mM NaCl, 0.1% (w/v) dodecyl maltoside (DDM), pH 8.0 was initially used. pH 8.0 was chosen to promote the formation of a significant number of negatively charged regions on the protein, allowing it to interact with the positively charged Q sepharose. Similarly a 2X MES buffer to make a final concentration of 25 mM MES, 50 mM NaCl, 0.1% (w/v) DDM, pH 6.0 was selected

for cation exchange chromatography using SP sepharose (Amersham Biosciences). This promotes the formation of positively charged regions on the protein, thus allowing interaction with the negatively charged SP sepharose.

Small columns (150 µl) of Q and SP sepharose were prepared by pre-equilibrating the packing with either the Tris buffer (Q sepharose) or MES buffer (SP sepharose) detailed above. The protein sample was applied to the column, which was then washed with four column volumes of the Tris or MES buffer to remove non bound proteins. Bound proteins were eluted in 200 µl volumes with increasing concentrations of NaCl (in Tris or MES buffer) of; 100 mM, 250 mM, 500 mM and 1M. All fractions were examined by SDS-PAGE and Western blotting for the presence of protein. Larger expression reactions were scaled up accordingly.

## 4.2.5 Estimation of protein concentration (Bradford method)

Protein concentrations were determined with the Bradford assay (Biorad) according to the manufacturer's instructions using Bovine Serum Albumin as a standard.

## 4.3 Results

### 4.3.1 Expression of full length rPmp A, rPmp D and rPmp I using the MembraneMax™ Protein Expression System

Micrograms of recombinant expression plasmids pPmpA23b-FL, pPmpD28b-FL and pPmpI23b-FL were produced using the PureLink™ HiPure Maxiprep kit (Invitrogen) as in **section 2.2.13.** The MembraneMax™ expression trials were set up as in **section 4.2.1.** Aliquots of each expression reaction were resuspended in SDS buffer to examine the total cell-free proteins and the remainder of each reaction was separated into soluble supernatant (containing NLPs) and insoluble material via centrifugation as in **section 4.2.1.**

Analysis by Western Blot with the penta-His antibody (**section 3.3.5.2**) and SDS-PAGE showed the expression of full length putative rPmp I detected in the region of 90 kDa, shown in the insoluble fraction therefore not associated with soluble NLPs, **Figure 4.2.** The predicted molecular mass of Pmp I is ~95 kDa. The two larger bands detected were also recognised by the antibody as bearing a histidine tag so are likely to be dimers and trimers of rPmp I, although the sizes could not be determined using the scale of 6xHis ladder on the Western blot. The band for full length rPmpI was not clearly visible in the total cell-free protein (TCP) as this sample was more dilute. Insoluble proteins pellet upon centrifugation to produce more concentrated samples. The antibody did not recognise any protein in the lanes containing expression products of rPmp A or rPmp D.
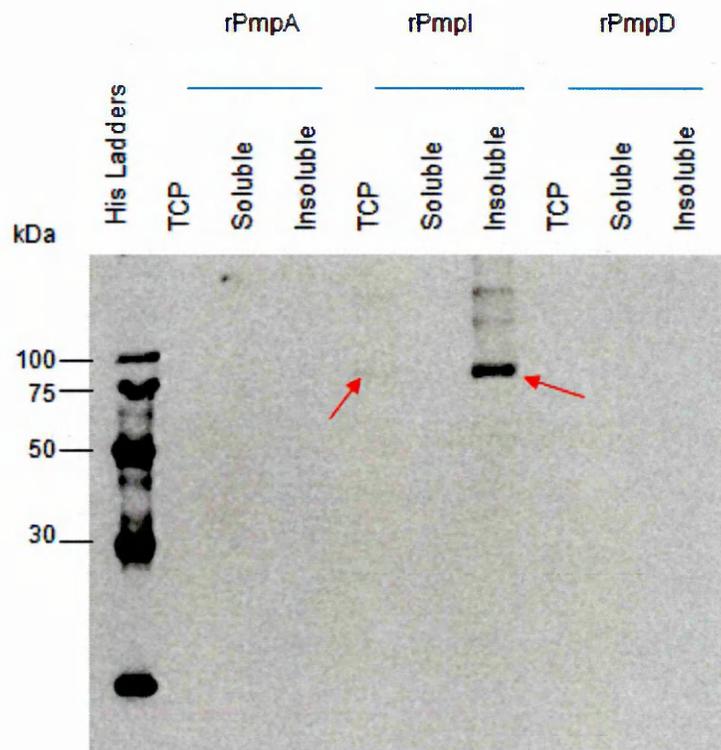
**Figure 4.2 - Western blot showing the expression of full length rPmp A, rPmp D and rPmp I using a cell-free expression system, MembraneMax™**. The proteins were expressed in 100 µl reactions for 2 hours at 37 °C as described in **section 4.2.1**. The total cell-free protein (TCP) was examined, alongside the soluble and insoluble fractions. Putative monomers of ~90 kDa with possible dimers and trimers of rPmp I were detected on the Western blot (top) in the total cell protein and insoluble fraction indicated by the red arrows. There was no detection showing the expression of rPmp A or rPmp D.

### 4.3.2 Expression and purification of the putative β-barrel Pmp D autotanslocator domain (rPmpDauto) in MembraneMax™

Micrograms of recombinant expression plasmid pPmpD28b-C were produced using the PureLink™ HiPure Maxiprep kit (Invitrogen) as in **section 2.2.13**. Cell-free expression reactions were set up as in **section 4.2.1** and were fed after a 30 minute interval. As rPmpDauto was under the control of the lac repressor gene via the T7*lac* promoter (discussed in **section 3.1.1**), trials to establish if IPTG was necessary to induce the promoter were carried out with and without IPTG.

Western blotting and SDS-PAGE confirmed the successful expression of the His-tagged autotranslocator domain of recombinant Pmp D. The bands detected in **Figure 4.3a** show that the *E.coli* derived transcriptional/translational machinery readily expresses the rPmpDauto domain without the need for IPTG. Protein was of the expected molecular weight of this domain at approx 27kDa and is the same molecular weight as previously detected expressing rPmpDauto in *E.coli* (**chapter 3**). Discrepancies between the pre-stained marker and the 6XHis ladder between 25 kDa and 30 kDa bands remain (**section 3.3.2**).

Fractionation of the expression products as in **section 4.2.1** to separate soluble (NLP fractions) and insoluble products showed a small proportion was detected in the soluble fraction presumably incorporated into NLPs **Figure 4.3b** (Katzen *et al.,* 2008). rPmpDauto was also expressed as insoluble protein and more easily detected as this was concentrated into a pellet upon fractionation. The amount in the soluble fraction was within the limits of detection from a 100 μl cell-free expression reaction and could be scaled up. Since downstream applications were dependent upon Western blotting and Coomassie SDS-PAGE it was important that the expressed protein was within detectable limits. This method of expression could have allowed for downstream processing to be monitored more easily in comparison to detergent solubilised rPmpDauto from 0.5 L of *E.coli* culture (**chapter 3**) where difficulties in the detection hampered purification trials. Variations to the incubation set-up conditions, including lowering the incubation to 30 °C in an effort to increase the levels of expression and the yield of soluble rPmpDauto did not have any beneficial effects and showed that the cell-free system and vector were most likely expressing at their maximal rates (data not shown).
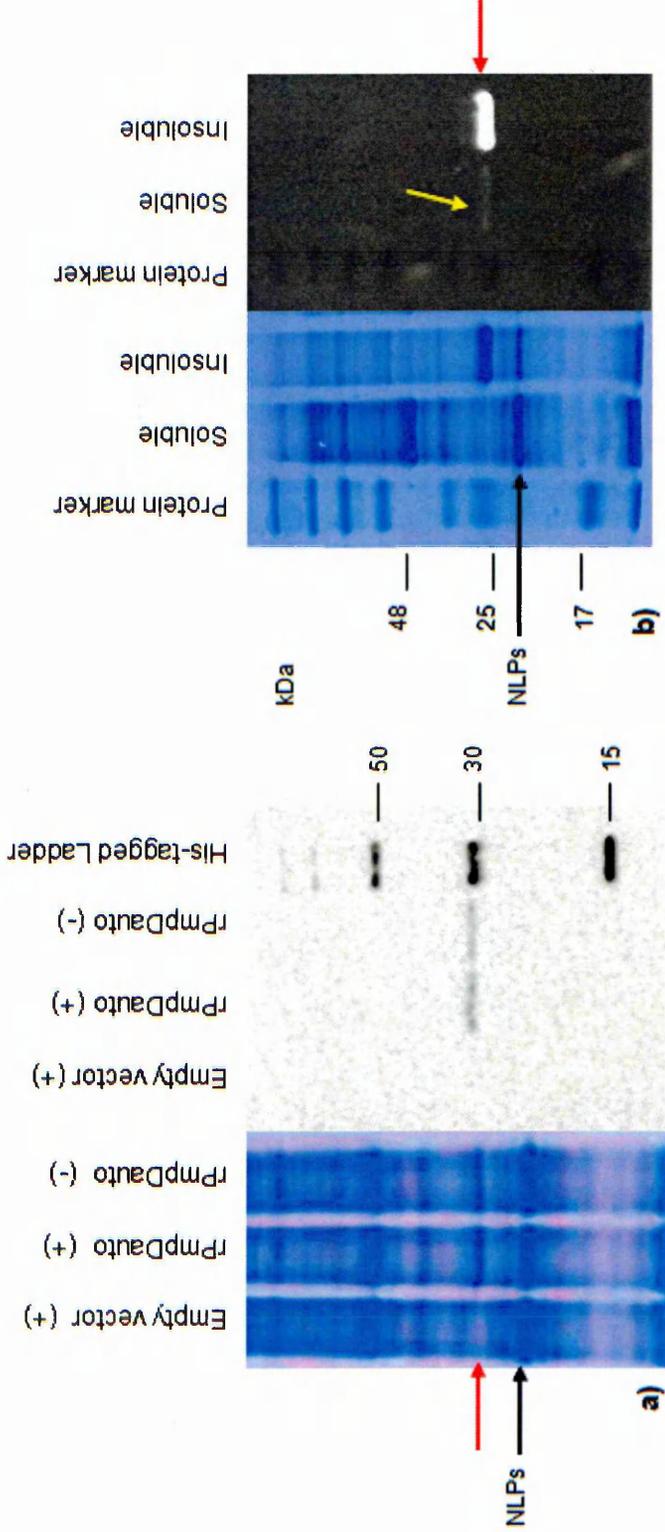
**Figure 4.3 - Western blot and Coomassie stained SDS-PAGE showing the expression of the rPmp D autotransiocator domain using the cell-free expression system MembraneMax™** ( a) The truncated protein domain was expressed alongside an empty pET28b vector as a negative control for 2 hours at 37 °C in one 100 µl reaction for each as described in **section 4.2.1**. To examine the regulation of the vector the expression was examined with (+) and without (-) the presence of IPTG. rPmpDauto was visualised on a Coomassie stained SDS-PAGE gel (left) shown by the red arrows and detected with via Western blotting (right), histidine ladder, (Qiagen). (b) The total cell-free protein from a 100 µl non-IPTG induced expression was then fractionated in to soluble and insoluble pellet fractions to determine if rPmpDauto had been incorporated into the soluble NLPs. rPmpDauto was detected and visualised on a Coomassie gel (left) in line with the red arrow in the insoluble pellet. Western blotting (right) detected small amounts of rPmpDauto in soluble NLP fraction, depicted by the yellow arrow. Pre-stained protein marker in kDa (New England Biolabs). Both SDS-PAGE gels show NLPs 9orange arrow).

173

## 4.3.2.1 Small-scale purification of rPmpDauto-NLPs complexes using Ni-NTA chromatography

To gain a better insight into the nature of the Pmps, purified protein would be required for the future analyses such as circular dichroism. rPmpDauto was to be purified by affinity column chromatography using Ni-NTA agarose (Qiagen) to bind the C-terminal histidine tag. As the NLPs were understood to offer a stable environment for recombinant proteins (Katzen *et al.,* 2009), the purification was first performed using the soluble fraction containing rPmpDauto in NLPs, from a 200 µl MembraneMax™ cell-free expression diluted in 2X binding buffer without detergents as in **section 4.2.3**.

rPmpDauto in NLPs remained in the column flow-through as unbound proteins and therefore the binding efficiency to Ni-NTA agarose was poor (data not shown). Additionally, rPmpDauto began to elute from the column with a relatively low concentration of imidazole (20 mM), showing that the protein had weak binding to the Ni-NTA agarose. Weak binding efficiency could have been due to the hexahistidine tag being only partially accessible. In such circumstances, it can bind as long as more than two histidines are available to interact with the nickel ions, but generally, the smaller the number of accessible histidines, the weaker the binding will be (Qiagen, 1997). Extending the association time between rPmpDauto and Ni-NTA by batch-binding the soluble fraction with Ni-NTA as in **section 4.2.3** did not improve the binding efficiency (data not shown).

To test the possibility that the NLPs surrounding rPmpDauto, were not obstructing the accessibility of the His tag. rPmpDauto was released from the NLPs using detergent. Detergents are necessary reagents in membrane protein biochemistry as they serve to disrupt lipid bilyers and solubilise membrane proteins into micelles. A micelle is a thermodynamically stable colloidal aggregate of detergent monomers in which the hydrophillic polar heads are directed outward in contact with the water and the nonpolar tails are clustered inward, avoiding exposure to water thus act to stabilise membrane proteins providing hydrophobic surfaces for exposed amino acids normally within the bilayer (Seddon *et al.,* 2004). Using this method, detergent micelles could replace the NLPs as the NLPs would themselves be unsuitable for structural analysis by circular dichroism and ultimately would have to be removed during purification.

Many prokaryotic membrane transporters are soluble in Triton-X100, dodecyl-β-D-maltoside (DDM) or octyl-β-D-glucoside (OG) (personal communication, Dr D. Hadden). Octyl glucoside and dodecyl maltoside are similar in structure, based on alkyl chains with a carbohydrate moiety, and both are chemically well defined. One advantage of octyl glucoside is it has a high critical micelle concentration (CMC), which enables the detergent to be easily removed from solution by dialysis. Triton-X100 is an inexpensive non-ionic detergent, commonly used in the laboratory for many applications and is well suited for membrane protein extraction, it contains impurities but is useful as a cost effective method for use during optimisation trials.

Solubilisation trials were therefore initiated using Triton-X100, OG and DDM at a concentration range of 1% (w/v). Soluble expression fractions containing rPmpDauto with NLPs were incubated for one hour, with gentle agitation, end over end in 1% (w/v) detergent as detailed in **section 4.2.2**. Precipitated proteins were pelleted into the detergent insoluble fraction by centrifugation at 4 °C, 40,000g for 20 minutes and the supernatant was passed through a 0.22 μm filter (Millipore). The pellet and supernatant samples were analysed by Western blot.

Western blots showing the results of detergent extraction are shown in **Figure 4.4.** rPmpDauto was not recovered into detergent micelles using OG as all the protein was detected in the detergent insoluble pellet as precipitate. Triton-X100 and DDM were able to partially extract rPmpDauto shown by yellow arrows and therefore were selected as the detergent of choice to begin purifying rPmpDauto extracted from NLPs.

## 4.3.2.2 Small-scale purification of detergent extracted rPmpDauto with Triton-X100 or DDM, using Ni-NTA chromatography

Initial attempts to purify rPmpDauto by Ni-NTA chelation revealed cell-free expressed rPmpDauto in NLP complexes had a very low affinity for the Ni-NTA (**section 4.3.2.1**). On the basis of the detergent solubilisation trials it was decided the purification of rPmpDauto would be attempted using Triton-X100 and DDM extracted protein. The purification was performed on a small scale using the soluble fraction from a 200 μl reaction that had undergone incubation with 1% (w/v) detergent to extract the rPmpDauto from the NLPs. rPmpDauto in detergent was applied to 250 μl Ni-NTA 50% agarose slurry and the purification carried out as described in **section 4.2.3**. Fractions were analysed by SDS-PAGE and Western blot.

The Western blots and Coomassie SDS-gels showing the purification of rPmpDauto in both Triton-X100 and DDM are shown in **Figure 4.5** and **Figure 4.6** respectively. Triton-X100 extracted rPmpDauto bound to the Ni-NTA and was eluted with 250 mM imidazole. However the binding efficiency was poor as this protein also eluted with the 20 mM imidazole wash. A similar outcome was obtained using DDM however the bands remained very faint and difficult to identify despite efforts to load more protein on to the blotting membrane. Several methodologies to improve the binding efficiency included batch binding as in **section 4.2.3** at both 22 °C and 4 °C, were attempted but did not improve the outcome.

Coomassie SDS-PAGE analysis of the same fractions showed few protein bands within the second 20 mM and within the 250 mM imidazole eluted fractions (lanes 'Wash 2' and 'Eluate' in **Figure 4.6** bottom).This indicated that the overall presence of contaminants was low and the stringency of the wash conditions (20 mM imidazole) was adequate. Because of the low quantity of proteins, it was difficult to assess whether rPmpDauto was present as a pure protein in the 250 mM fraction. To examine these fractions more closely, TCA/acetone precipitation was used to concentrate the purified 250 mM fraction from **Figure 4.6**. Coomassie SDS-PAGE revealed the presence of rPmpDauto (red arrow) with few larger unidentified contaminant bands (green arrows) that co-purified with rPmpDauto given in **Figure 4.7** and confirmed by Western blot.

Efforts to separate the contaminant proteins from the sample with an extra higher stringency wash at both 40 mM and 75 mM imdazole did not separate the proteins (data not shown). Additional purification methods such as ion exchange chromatography or size exclusion chromatography could not be implemented to remove the larger contaminant proteins as the yield of rPmpDauto was already at the limit of reliable detection following Ni-NTA purification, further losses as a result of buffer exchange and purification steps made this an impractical step.

Circular dichroism requires a 200 µl sample of at least 0.5 mg/ml and 95% homogeneity. It was difficult to estimate the protein concentration of the partially purified rPmpDauto fraction because it was so low. The Bradford (**section 4.2.5**) protein estimation method was used and revealed the total protein in the 200 µl fraction to be very low at approximately 20 µg/ml, a level far below the concentration required for CD analysis. Although CD analysis could not be carried out for rPmpDauto, efforts to express and purify other rPmp domains were attempted in the hope that these homologues would display a higher affinity for Ni-NTA purification.

176
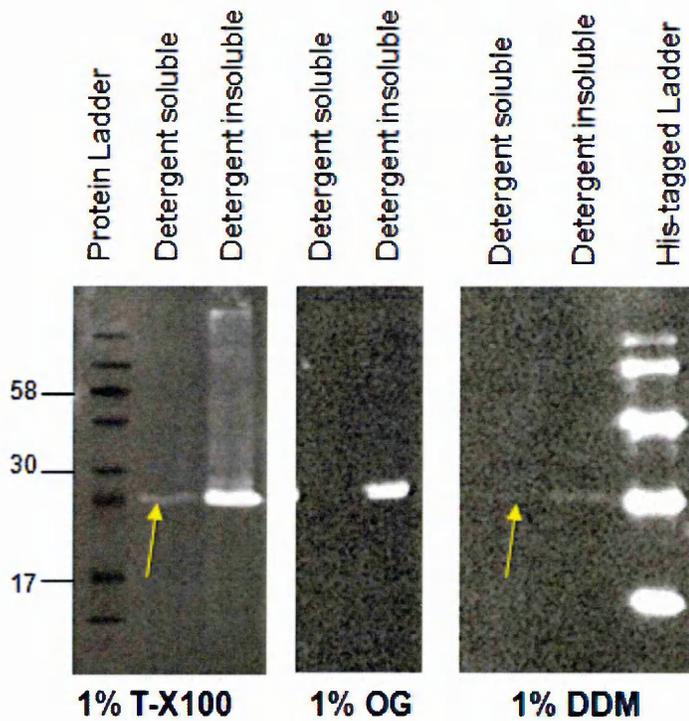
**Figure 4.4 - Western blots showing the detergent solubilisation (extraction) of rPmpDauto from NLPs with Triton-X100, OG and DDM to release rPmpDauto from the MembraneMax™ NLP scaffolds.** The soluble fraction from the cell-free expression reaction of rPmpDauto in NLPs was resuspended in 1XPBS with 1% detergent to disrupt the NLPs and release rPmpDauto into detergent micelles. After detergent extraction, the samples were centrifuged to remove the detergent insoluble proteins i.e. aggregated proteins displaced from the NLPs that did not encounter detergent micelles. Triton-X100 and DDM offered some detergent extracted protein recovery with faint bands detected via Western blot (yellow arrows). There was no recovery into the detergent using OG. Triton-X100 and DDM offered the best recovery from NLPs out of these trials.
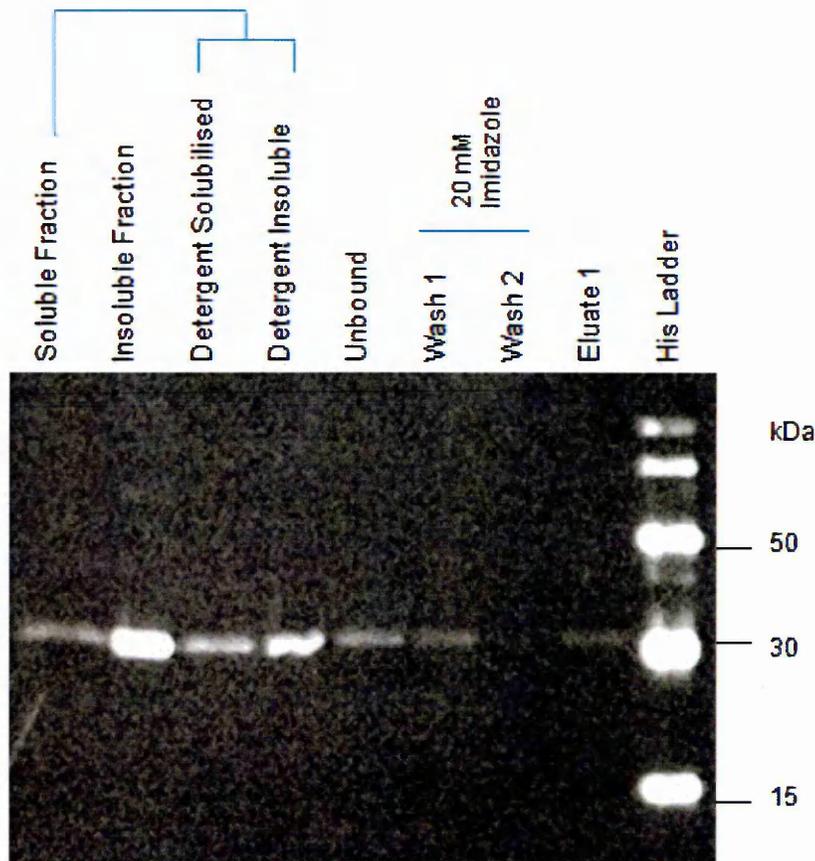
**Figure 4.5 - Western blot showing the initial trials to purify rPmpDauto expressed using MembraneMax™ protein expression system, released from NLPs using 1% Triton-X100.** The soluble rPmpDauto expressed in NLPs (soluble fraction) was extracted from the NLPs using 1% Triton-X100 to produce detergent solubilised rPmpDauto  and detergent insoluble rPmpDauto as indicated above. Subsequently the detergent solubilised sample was applied to Ni-NTA for binding 3 times.  A large proportion of rPmpDauto remained in the unbound fraction after application onto a Ni-NTA column. A small amount of rPmpDauto was eluted with 20 mM (Wash 1) and 250mM imidazole (Eluate 1). The presence of rPmpDauto within these fractions was not visible on a Coomassie SDS-PAGE.
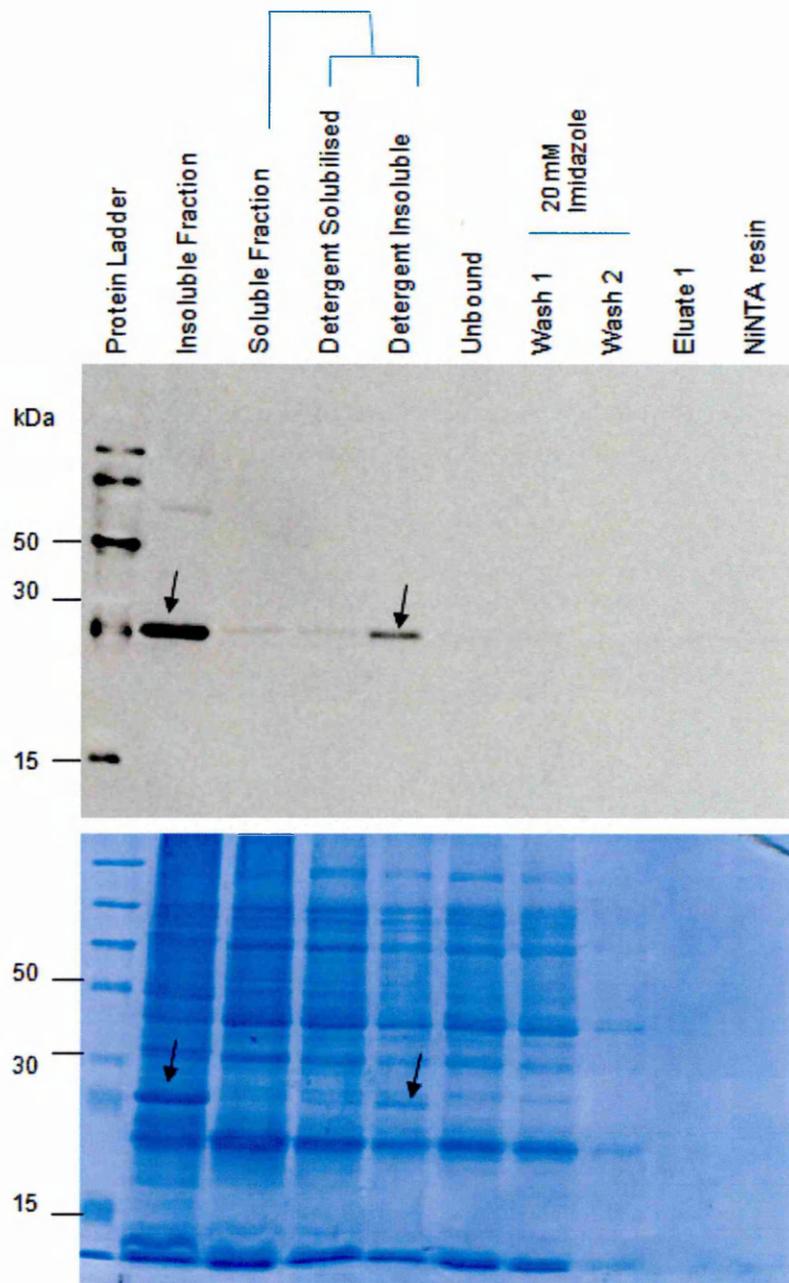
**Figure 4.6 - Western blot (top) and Coomassie gel (below) showing the purification of rPmpDauto expressed using MembraneMax™ released from NLPs using 1% DDM.** The soluble rPmpDauto expressed in NLPs (soluble fraction) was extracted from the NLPs using 1% DDM to produce detergent solubilised rPmpDauto and detergent insoluble rPmpDauto as indicated above. Subsequently the detergent solubilised sample was applied to Ni-NTA for binding 3 times. Although very faint the majority of rPmpDauto applied to the resin remained in the unbound fraction and the first wash. A very faint trace of rPmpDauto was detected in 250mM imidazole eluate fraction via Western blot. However these bands were not visible on the Coomassie SDS-PAGE. The insoluble rPmpDauto (black arrows) was used as a guide to determine putative rPmpDauto in the eluted fractions.
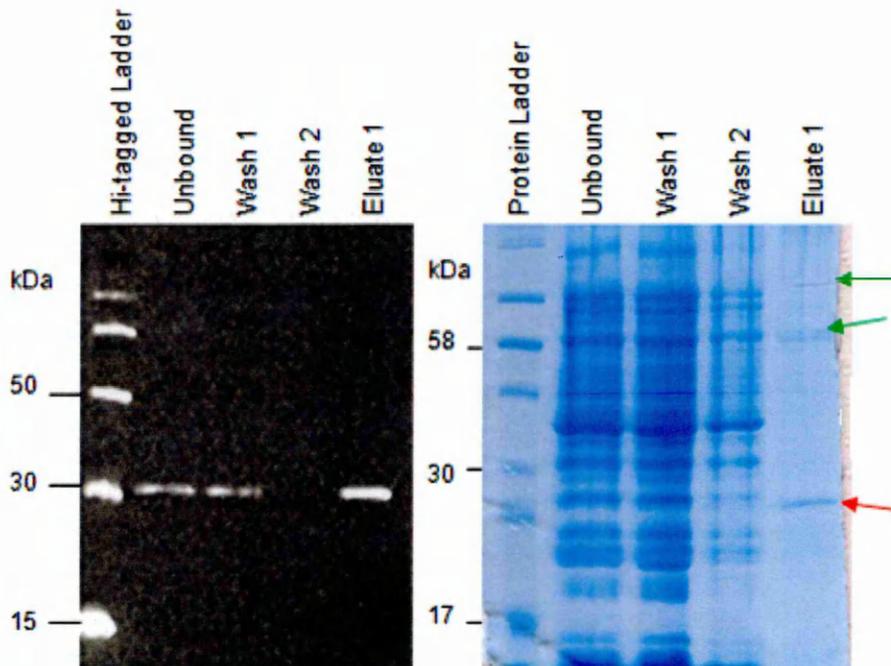
**Figure 4.7 - Western Blot (left) and Coomassie gel (right) of the X10 concentrated fractions of purified rPmpDauto**. The unbound, Wash 1 and 2 and Eluate 1 fractions from the purification trials using DMM in Figure 4.6 were concentrated X10 using TCA/acetone precipitation to pellet the proteins. The pellets were resuspended in SDS loading buffer and examined via Western blot and Coomassie SDS-PAGE. Concentrating the proteins showed the presence of rPmpDauto (red arrow), at levels previously beyond detection in addition to co-purified protein bands at approximately 58 kDa and 80 kDa that were not recognised by the penta-His antibody (green arrow).

### 4.3.3 MembraneMax™ cell-free expression of the predicted β-barrel domain of rPmpGauto

Micrograms of recombinant expression plasmid pPmpG22b-C were produced using the PureLink™ HiPure Maxiprep kit (Invitrogen) as in **section 2.2.13**. Cell-free expression reactions were set up as in **section 4.2.1** and were fed after a 30 minute interval. After expression the complete reaction was separated into soluble proteins and insoluble aggregates by centrifugation prior to examination by SDS-PAGE and by Western blot using a penta-His primary antibody (**section 3.2.5.2**).

Western blotting shows the successful expression of putative rPmpGauto as both soluble and insoluble protein in **Figure 4.8a**. The protein was detected as two bands in the soluble fraction at approximately 40 kDa and 34 kDa and insoluble rPmpGauto was visualised on the Coomassie SDS-PAGE gel. These molecular weights were as predicted for the Pmp G autotranslocator domain and were comparative to the sizes observed when expressing rPmpGauto in *E.coli* (chapter 3). The two bands were an indication that proteolysis was taking place, possibly of the pelB fusion peptide.

### 4.3.3.1 Small-scale purification of MembraneMax™ derived rPmpGauto using Ni-NTA chromatography

Attempts to purify rPmpGauto were by affinity column chromatography using Ni-NTA agarose (Qiagen) to bind the histidine tag. The purification was performed on a small scale using rPmpGauto expressed in the NLP-soluble fraction from a 200 μl reaction, later extracted from the NLPs in 1% (w/v) DDM using similar methods employed to purify rPmpDauto (**section 4.3.2.2**). The detergent solubilised rPmpGauto was applied to 250 μl Ni-NTA 50% agarose slurry and the purification carried out as described in **section 4.2.3**. Fractions were analysed by SDS-PAGE and Western blot with the penta-His antibody.

The Western blot showed very faint putative rPmpGauto bands of approximately 40 kDa present in the detergent solubilised fraction the majority of rPmpGauto was lost in the detergent insoluble pellet (**Figure 4.8b**). However the detergent extracted rPmpGauto was applied to a Ni-NTA column and the domain appeared to bind to the Ni-NTA agarose as it was not detected in the unbound flow-through fraction or any of the wash and eluate fractions. The Ni-NTA resin was probed showing rPmpGauto had not remained bound however as the amount of protein loaded on to the Ni-NTA column from the detergent soluble fraction was already very difficult to detect this hampered efforts to locate rPmpGauto throughout the fractions despite efforts to

181

load more protein onto the gel for immunoblotting. Concentration of these fractions using TCA to precipitate the proteins did not improve the Western blot detection of rPmpGauto and did not show that the protein was widely distributed throughout the fractions (data not shown). Purification trials were repeated using Triton-X100 to extract rPmpGauto from the NLPs to improve the yield of detergent solubilised protein yet the outcome was similar and rPmpGauto could not be detected (data not shown).

### 4.3.4 MembraneMax™ cell-free expression of the predicted passenger domain of PmpA (rPmpA-N)

Micrograms of recombinant expression plasmids pPmpA23b-N were produced using the PureLink™ HiPure Maxiprep kit (Invitrogen) as in **section 2.2.13**. Cell-free expression reactions were set up as in **section 4.2.1** and were fed after a 30 minute interval. The expression reaction was fractionated into soluble and non-soluble samples. The separated fractions were analysed by Western blotting and SDS-PAGE as in **sections 2.2.5.**

The Western blot revealed the successful expression of putative rPmpA-N at the expected molecular weight of 56 kDa (**Figure 4.9**) however this was as insoluble aggregates not associated with NLPs. Efforts to produce some soluble rPmpA-N in NLPs by lowering the incubation temperature, reducing the aeration to slow expression were met without success (data not shown).
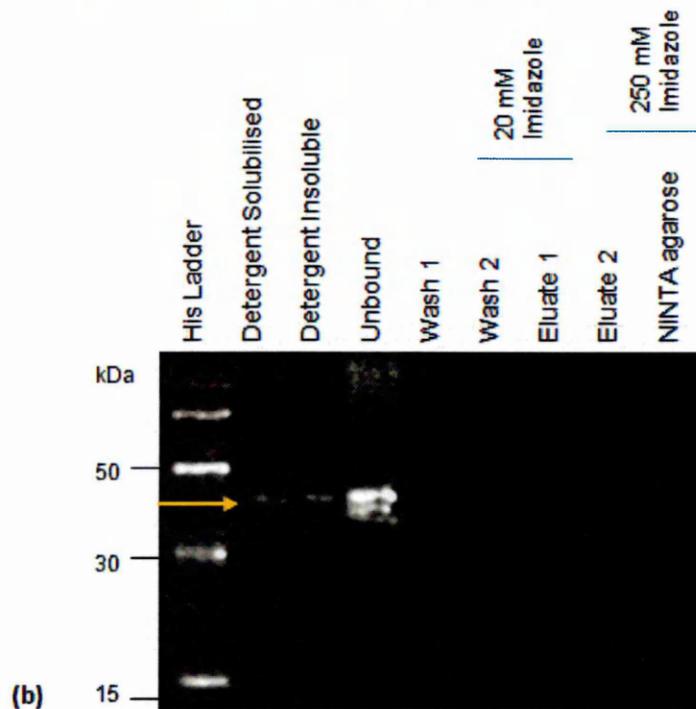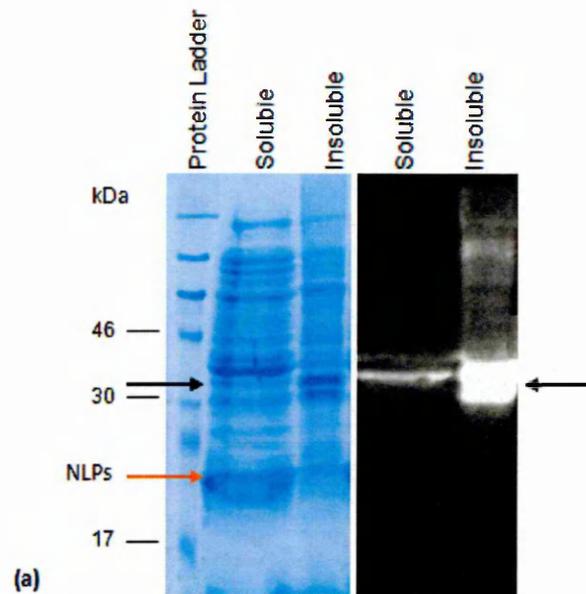
**Figure 4.8 - Coomassie SDS-PAGE and Western blots showing the expression and purification trials of rPmp G autotranslocator domain (rPmpGauto).** (a) The truncated rPmpGauto was expressed for 2 hours at 37 °C in 100 µl MembraneMax™ cell-free reaction. The sample was fractionated into soluble (containing NLPs) and insoluble fractions. rPmpGauto is shown by the black arrows at approximately 34 kDa with a larger 40 kDa band detected in the soluble NLP fraction (right). (b) The rPmpGauto expressed in NLPs (the soluble fraction from above (a)) was extracted from the NLPs using 1% DDM to produce detergent solubilised rPmpGauto and detergent insoluble rPmpGauto (yellow arrow). The detergent solubilised sample was applied to Ni-NTA for binding 3 times. rPmpGauto was not detected in the remainder of the fractions.
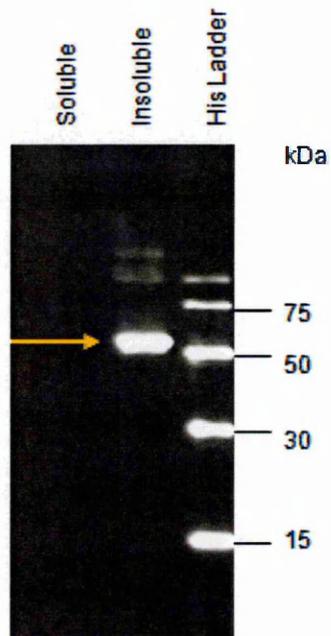
**Figure 4.9 - The expression of rPmp A passenger domain (rPmpA-N) using a cell-free expression system MembraneMax™.** The truncated protein domain was expressed for 2 hours at 37 °C in 100 µl reaction and was fractionated into soluble and insoluble fractions as described in section 4.2.1. rPmpA-N was detected as insoluble aggregates, shown by the yellow arrow at approximately 56 kDa. The protein markers are shown in kDa.

### 4.3.5 MembraneMax™ cell-free expression of the predicted passenger domain of PmpI (rPmpI-N)

Micrograms of recombinant expression plasmids pPmpI23b-N were produced using the PureLink™ HiPure Maxiprep kit (Invitrogen) as in **section 2.2.13**. Cell-free expression reactions were set up as in **section 4.2.1** and were fed after a 30 minute interval. The expression reaction was fractionated into soluble and non-soluble samples.

Probing with a penta-His primary antibody shows the successful expression of putative rPmpI-N as both soluble (with NLPs) and insoluble aggregates confirmed by Western blot and Comassie stained SDS-PAGE (**Figure 4.10a**). The wide band was due to abundant detection by the penta-His antibody. At approximately 44 kDa the molecular size was as expected for this domain and corresponded with the size of the rPmpI-N domain expressed in *E.coli* (**Chapter 3**).

### 4.3.5.1 Small-scale purification of MembraneMax™ derived rPmpI-N using Ni-NTA chromatography

As this passenger domain was intended to undergo secondary structural analysis it was necessary to purify the protein and remove the NLPs. Considering the NLPs construction is based around a scaffold protein co-purification of NLP scaffolds would have interfered with CD structural analysis of rPmpI-N. This exchange from NLPs into detergent micelles was done prior to purification as opposed to during purification when the protein is bound to the column. Disrupting the NLPs whilst bound to the column was not favoured as protein precipitation occured when adding detergents (observed previously when adding detergents to release rPmpDauto and rPmpGauto) that ultimately could have led to the loss of protein conformation. The purified fractions would then have become contaminated with misfolded rPmpI-N unsuitable for structural and functional analysis. Based on the detergent extraction trials of rPmpDauto, where Triton-X100 and DDM were capable of recovering some the Pmp domain from the NLPs, DDM was chosen to extract the passenger domain of Pmp I since Triton X-100 contains aromatic rings that absorb at 260-280 nm and is generally known to contain impurities such as peroxides as contaminants unsuitable for CD analysis (Kelly *et al.,* 2005). Therefore the NLPs were disrupted using 1.0% (w/v) DDM to release the proteins (**section 4.2.3**) and detergent solubilised recovery of rPmpI-N was observed in **Figure 4.10b**.

Subsequently rPmpI-N was purified by affinity column chromatography using Ni-NTA agarose (Qiagen). The purification was performed using the detergent extracted rPmpI-N from the soluble fraction of a 200 µl MembraneMax™ cell-free expression reaction. Protein was eluted in steps with increasing concentrations of imidazole (10 mM, 40 mM, 75 mM and 250 mM imidazole), using 200 µl volumes for each concentration. Aliquots of each fraction were examined by SDS-PAGE and Western blotting. Using this method, rPmpI-N was partially purified, and putative rPmpI-N bands were detected on the Western blot at 38 kDa and 44 kDa (blue arrows) in the 75 mM and 250 mM imidazole eluted fractions (**Figure 4.11a**). These bands were also observed when expressing rPmpI-N in *E.coli* possibly an indication of cleavage of the predicted native signal leader peptide as the smaller 38 kDa band becomes more prominent during downstream processing. The predicted molecular weight of the passenger domain of Pmp I is approximately 44 kDa with a predicted 6 kDa signal peptide. However a third band seen previously in *E.coli* expression trials of rPmpI-N was not detected here. A Coomassie stained SDS-PAGE gel showing the purification profile of rPmpI-N using Ni-NTA agarose (**Figure 4.11a**) could not determine the presence of rPmpI-N in any of the fractions as the yield of rPmpI-N was too low to detect with this staining technique.

To investigate the purity of the fraction containing rPmpI-N, TCA/acetone precipitation of the 250 mM imidazole fraction to produce a sample ten times more concentrated verified the presence of rPmpI-N at approximately 38 kDa (red arrow) and two larger contaminant proteins (green arrows) that would render the sample unsuitable for circular dichroism analysis **Figure 4.11b**. Although rPmpI-N was not fully purified, attempts to remove the co-purified contaminant proteins could be made using additional purification methods such as ion exchange as the yield of purified rPmpI-N by Ni-NTA chelation was still within the limits of detection by Western blot (**section 4.3.5.2**).
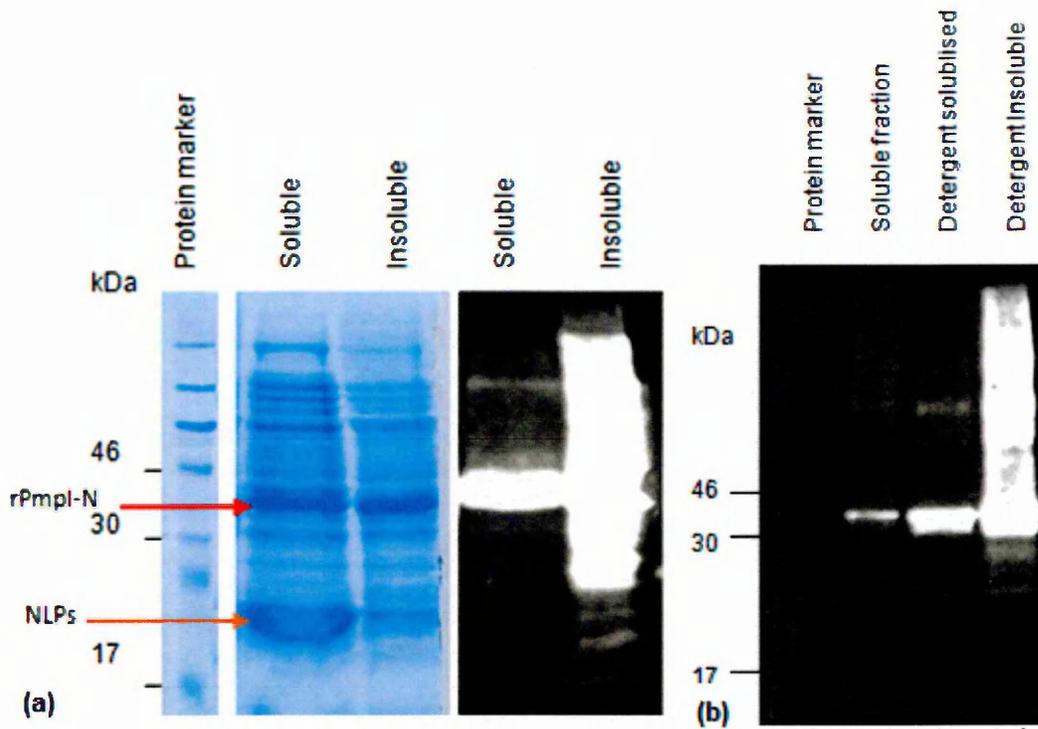
**Figure 4.10 - Coomassie SDS-PAGE gel and Western blots showing the expression of rPmp I passenger domain (rPmpI-N) and detergent solubilisation recovery from MembraneMax™ NLPs**. (a) The truncated protein domain was expressed for 2 hours at 37 °C in 100 µl MembraneMax™ cell-free expression reaction and was fractionated into soluble and insoluble fractions. rPmpI-N was present as both soluble and insoluble aggregates, visualised on a Coomassie stained SDS-PAGE gel (left) shown by the red arrow and detected with via Western blotting (right) at approximately 44 kDa. NLPs are shown by orange arrow. The protein markers are shown in kDa. (b) The soluble fraction from the cell-free expression reaction of rPmpI-N (from part (a)) was resuspended in 1% DDM to disrupt the NLPs and release rPmpI-N into detergent micelles. Precipitated protein was separated into the 'detergent insoluble' fraction by centrifugation and was not used further.
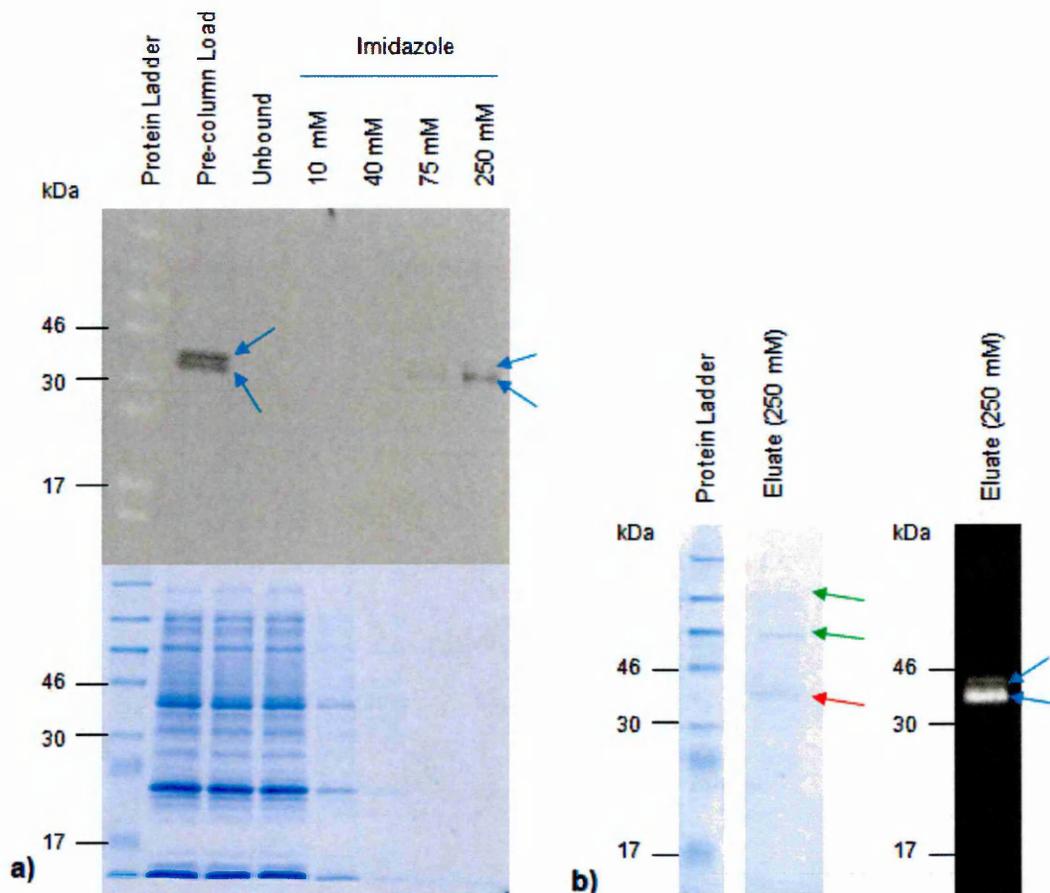
**Figure 4.11 - Western blots and Coomassie gels showing the partially purified rPmpI-N expressed using MembraneMax™ by Ni-NTA affinity chromatography.** (a) Western blot (top) shows the Ni-NTA purification of the two fragments of the Pmp I passenger domain. The detergent extracted rPmpI-N protein prior to application on to the Ni-NTA column is shown in the pre-column load lane. The sample was applied to the column 3 times and the unbound proteins were retained. The bound proteins were washed with 10 mM, 40 mM, 75 mM and 250 mM imidazole concentrations. rPmpI-N started to elute with 75 mM imidazole. And was detected in the 250 mM imidazolefraction (blue arrows). No protein was detected by Coomassie stained SDS-PAGE gel (below). (b) Concentration of the eluate (250 mM fraction) shows the purity of the fraction and rPmpI-N can be visualised on the Coomassie gel (left) indicated by the red arrow with two larger unidentified proteins at ~58 kDa and ~80 kDa (green arrows) previously seen when purifying rPmpDauto (**Figure 4.7**) rPmpI-N is confirmed by Western blot (right).

## 4.3.5.2 Small-scale purification of MembraneMax™ derived rPmpI-N using Ni-NTA chromatography and ion exchange chromatography

To improve the purity of rPmpI-N produced using nickel affinity chromatography, ion exchange chromatography was employed. Ion exchange chromatography separates molecules based on differences in their accessible surface charges. The partially purified reaction (**section 4.3.5.1**) was divided into two aliquots. In order to lower the ionic strength and alter the pH, the aliquots were buffer exchanged by applying the fractions to a concentrator (Vivaspin 500, 10 MWCO, Generon) and washing through with either 50 mM Tris, 25 mM NaCl, 0.1% DDM, pH 8.0, or 25 mM MES, 25 mM NaCl, 0.1% (w/v) DDM, pH 6.0. Samples in the alkaline buffer were applied to an anion exchanger, Q sepharose (Amersham Biosciences), and samples in the acidic buffer applied to a cation exchanger SP sepharose (Amersham Biosciences). Ion exchange chromatography was performed as described in **section 4.2.4.**

The described method failed to produce any visible bands on a Coomassie stained SDS-PAGE gel and rPmpI-N was not detected via Western blot after the buffers were exchange in the concentrators (data not shown). The vivaspin 500 concentrators were chosen for their hydrophilic polyethersulfone (PES) membrane that reduces sample binding however the rapid changes to the pH and ionic strength of the buffer using this method may have caused rPmpI-N to precipitate and bind to the PES membrane. A gentler method to exchange the buffer and minimise protein precipitation was done by dialysing rPmpI-N overnight at 4 °C as in **section 4.2.4** after which the sample was passed through a 0.22 μM Millipore filter to remove any precipitated proteins before application on to the equilibrated anion/cation exchanger as in **section 4.2.4**.

A Western blot showing the results of the purification using SP sepharose is shown in **Figure 4.12.** At pH 6.0, the rPmpI-N bound to SP sepharose tightly and did not elute until the column was washed with 1 M NaCl, where three His-tagged bands were detected on a Western blot (**Figure 4.12a**). Because of the weak detection of the protein, these were not visible on a Coomassie SDS-PAGE gel (data not shown) and so it was necessary to concentrate the 1M fraction to assess the purity of the fraction. Examining this concentrated fraction on a Coomassie stained gel (**Figure 4.12b**) confirmed the larger contaminant proteins that co-purified with rPmpI-N following Ni-NTA purification (**Figure 4.11b** depicted by green arrows) were not present on the Coomassie gel in **Figure 4.12b**. Presumably, these proteins did not bind to the sepharose, or were eluted from the column at a lower NaCl

concentration. However a smaller contaminant band was detected on the Coomassie gel at ~15 kDa that was not recognised by the penta-His when probing the same fraction. Additonally although Western blotting detected rPmpI-N in the 1M fraction, the yields were too low to be identified clearly by Coomassie SDS-PAGE even after concentration with TCA (**Figure 4.12b**). Furthermore it was difficult to estimate the protein concentration within this fraction as this was beyond the reliable lower limits of protein estimation by Bradford assay. Consequently two purification steps produced a protein unsuitable for CD analysis due to poor yields. Purification using Q sepharose was unsuccessful as rPmpI-N could not be detected in any of the fractions after column loading (data not shown) this could be due to dilution of rPmpI-N where the protein elutes in a wide range of NaCl concentrations at pH 8.0. If the protein had eluted sharply in few fractions, detection of the protein by Western blotting may have been easier.
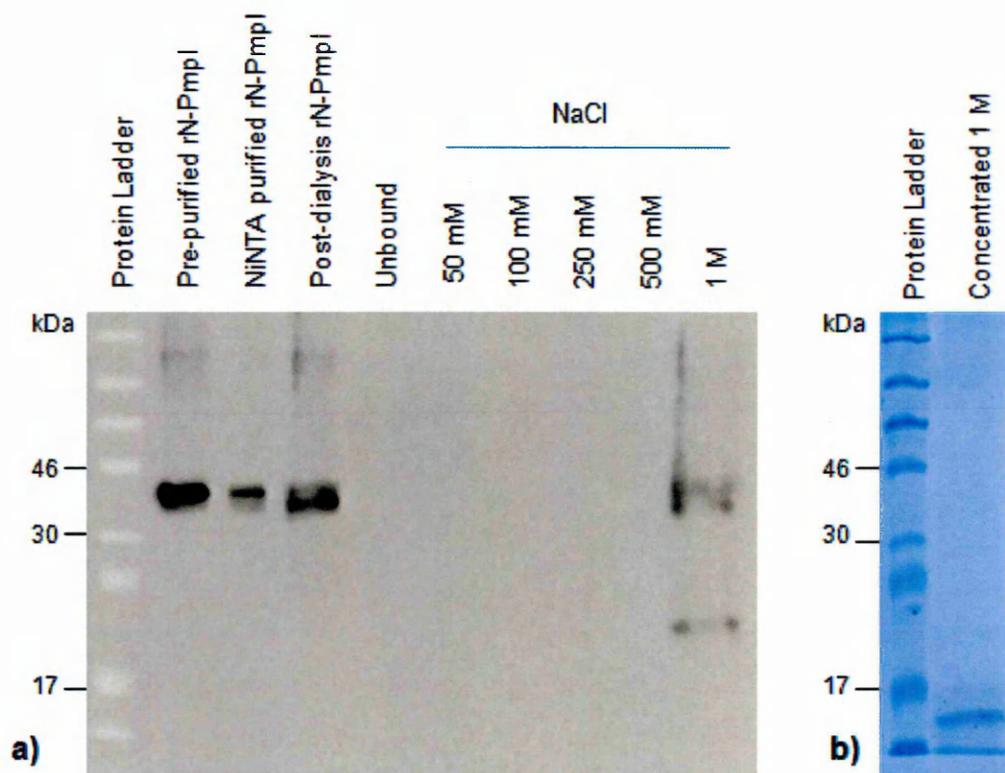
**Figure 4.12 - Western blot and Coomassie gel showing purification of Ni-NTA purified rPmpI-N by ion exchange chromatography.** (a) Western blot shows the second step of rPmpI-N purification using ion exchange chromatography. The detergent solubilised rPmpI-N (Pre-purified) was partially purified by Ni-NTA as in **section 4.3.5.1** (Ni-NTA purified). This partially purified rPmpI-N was dialysed into the MES buffer and remained soluble (post-dialysis). This dialysed rPmpI-N was applied to the SP sepharose column and the unbound proteins were retained (unbound). The column was washed with 50 mM, 100 mM, 250 mM, 500 mM and 1 M NaCl. The rPmpI-N protein was eluted at 1M NaCl as three His-tagged bands at approximately 44 kDa, 38 kDa and 20 kDa. (b) The 1M fraction was concentrated X10 to show the purity of the fraction on a Coomassie SDS-PAGE gel. rPmpI-N could not be clearly identified on the Coomassie gel.

### 4.3.5.3 Small scale purification of MembraneMax™ derived rPmpI-N using ion exchange chromatography alone

Based on the finding that recombinant Pmp I passenger domain had a strong binding affinity for SP sepharose (**section 4.3.5.2**), purification was carried out using cation exchange alone. By omitting the Ni-NTA purification, the losses in yield could be minimised. 200 µl of soluble expression of rPmpI-N was resuspended in 25 mM MES buffer containing 1% DDM to disrupt the NLPs and extract rPmpI-N into micelles. The detergent solubilised fraction was applied to the cation exchanger and subsequent fractions eluted in MES buffer with 0.1% DDM with increasing concentrations of NaCl as in **section 4.2.4**.

A Western blot shows the majority of rPmpI-N eluted within the 700 mM NaCl fraction (**Figure 4.13** (W7)), where the levels were too low for detection on the complementary Coomassie SDS-PAGE. This fraction was concentrated using TCA precipitation to check the purity. rPmpI-N was identified on the Coomassie SDS-PAGE and confirmed by Western blot (**Figure 4.14** red arrows). Several smaller contaminant proteins had co-purified and were more prominent. These contaminant proteins were tightly bound to the sepharose, eluting at and above 600 mM NaCl and were still evident in the 1M NaCl fraction (see W9 fraction on Coomassie gel, **Figure 4.13**). It would be very difficult to separate them from rPmpI-N using this technique. Additionally the yield of rPmpI-N was low making further purification of this eluted fraction impractical.

Following ion exchange chromatography the rPmpI passenger domain was detected as three peptides that were specifically identified by the penta-His antibody indicating the presence of His-tag. These sizes were ~38 kDa, ~44 kDa as previously observed in **Figure 4.11** with a third additional band present at ~22 kDa **Figure 4.14**. These bands are consistent with the peptides observed after expression of rPmpI-N in E.coli (**chapter 3**) where it was speculated the protein underwent auto-proteolysis. This 22 kDa band was not observed prior to ion exchange chromatography. Furthermore the Coomassie SDS-PAGE gel in **Figure 4.12b** shows the accumulation of a prominent non-His tagged peptide at approximately 12 kDa within the purified 1M NaCl rPmpI-N fraction that is not detected on the corresponding Western blot. This peptide was not seen previously in the **Figure 4.11b** where rPmpI-N was purified by Ni-NTA affinity chromatography. This peptide could be a by-product of rPmpI-N as a result of protein cleavage. This was consistently found following ion exchange chromotography purification trials, all in the presence of

protease inhibitors. Despite numerous efforts to produce a pure sample of the passenger domain of Pmp I, the co-elution of contaminant proteins remained a consistent problem and deemed the samples unsuitable for CD analysis.
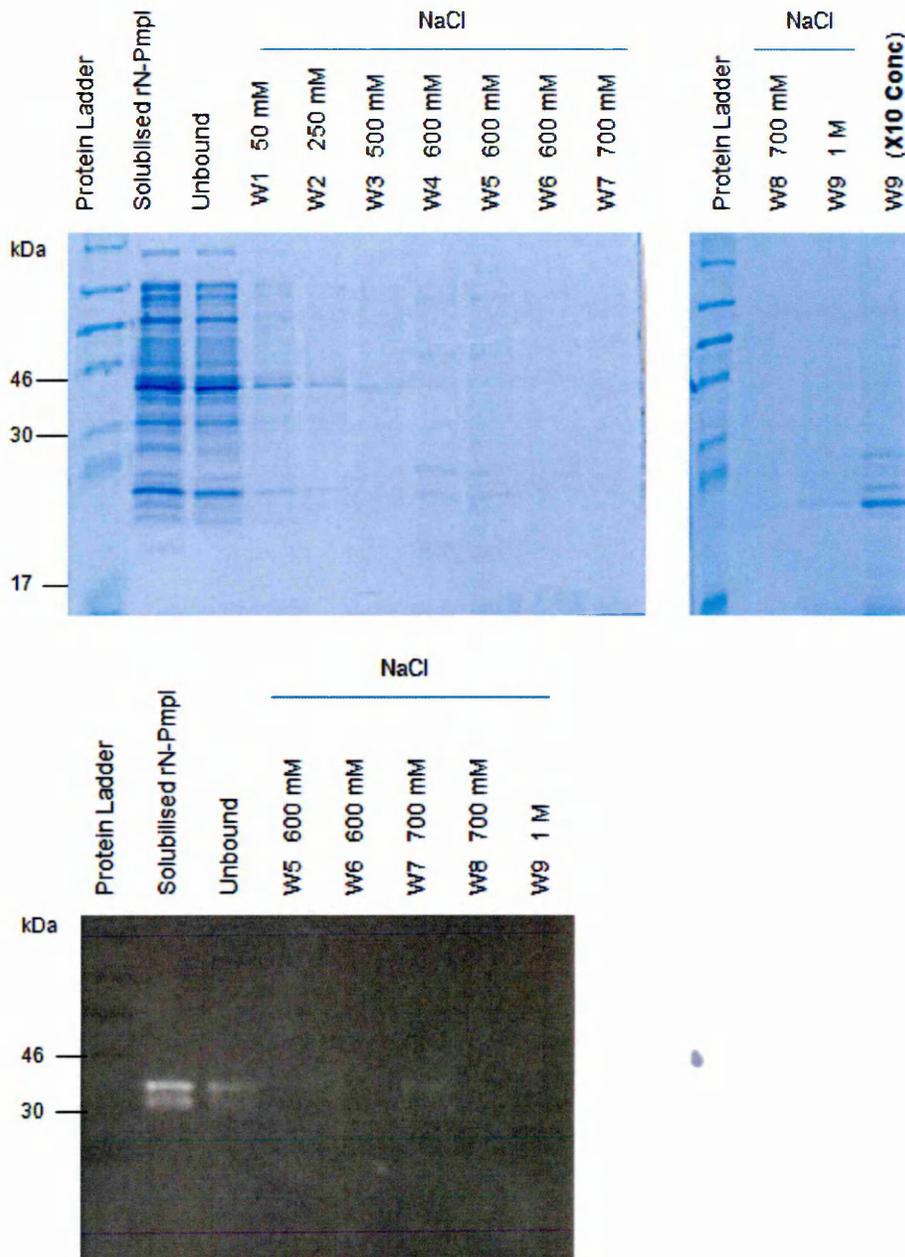
**Figure 4.13 - Coomassie gel (top) and Western blot (below) showing purification rPmpI-N by ion exchange chromatography.** The Western blot shows the rPmpI-N domain extracted from NLPs by detergents before binding to the column (solubilised rPmpI-N). This sample was applied to the SP sepharose three times prior to collecting the unbound fraction. The column was washed with increasing concentrations of NaCl as shown (W1-W9) rPmpI-N eluted at 700 mM NaCl (W7) as two faint His-tagged bands. rPmpI-N could not be clearly identified on the Coomassie gel above so fractions were concentrated (shown in Figure 4.14).
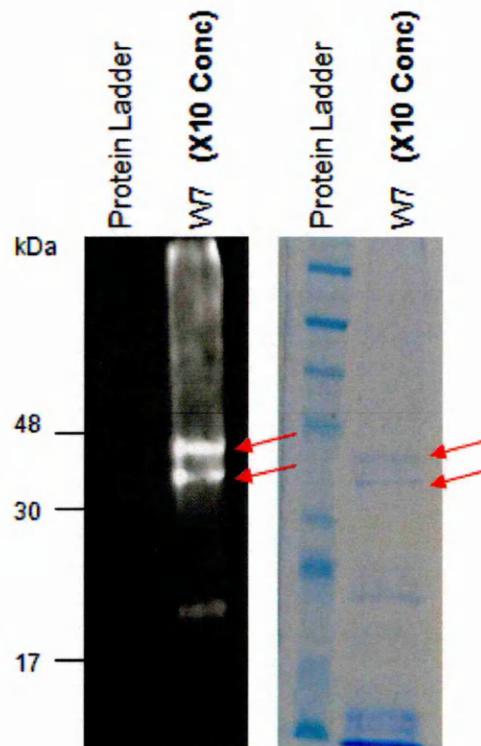
**Figure 4.14 - Western blot (left) and Coomassie SDS-PAGE (right) of the eluted fraction W7 from ion exchange chromatography X10 concentrated.** rPmpI-N eluted with 700 mM NaCl during ion exchange chromatography (Figure 4.13 lane W7) was concentrated X10 using TCA/acetone precipitation to visualise the bands on a Coomassie gel. rPmpI-N is shown as three His-tagged peptides at approximately 44 kDa, 38 kDa (red arrows) and a third 22 kDa on the Western blot. The 44 kDa and 38 kDa bands can be seen as very faint bands on the Coomassie gel, shown by the red arrows. The smaller band at ~20 kDa is not clearly identifiable on the Coomassie gel. Several other proteins also eluted with 700 mM Nacl.

## 4.4 Discussion

The aim of this section of work was to express recombinant Pmps or rPmp domains in a manner suitable for downstream processes including purification for structural and functional analyses. Problems with the yield, handling and purification were problematic using *E.coli* as an expression host and so hampered progress at this stage. It was hoped that the novel MembraneMax™ cell-free expression system would overcome some of the challenges previously met.

The MembraneMax™ cell-free system appeared to express one of the full length Pmps, recombinant Pmp I in an insoluble form. It was hoped that these proteins would be incorporated into the nanolipoprotein discs (NLPs) in the soluble fraction however it is possible that the full length rPmpI was too large at approximately 95 kDa, and also the system may not provide an environment suitable for full length Pmp protein folding. The MembraneMax™ cell-free system has displayed the expression of membrane proteins up to 100 kDa in previous studies however membrane proteins over 100 kDa have not been expressed or are expressed as truncated products (Katzen *et al.,* 2008), Although rPmp I was in an aggregated form, this was an improvement on the expression trials previous using *E.coli* where the expression of full length rPmps was not observed.

The autotranslocator domain of Pmp D, that embeds and acts as a pore within the membrane bilayer was expressed in a soluble state using the novel cell-free expression system that utilises nanolipoprotein scaffolds to mimic the lipid bilyer. Purification of the recombinant Pmp D autotranslocator (rPmpDauto) domain from this cell-free expression using Ni-NTA agarose was partially successful. It was separated from the majority of the *E. coli* endogenous proteins present within the cell-free system. However, rPmpDauto did not bind tightly to the Ni-NTA column, and therefore began to elute with a low concentration of imidazole (20 mM), which limited the purification efficiency. This was unexpected as rPmpDauto produced in *E.coli* bound with a useful efficiency to Ni-NTA during purification (**chapter 3**). It was considered that the NLPs surrounding the cell-free expressed rPmpDauto, being large proteins themselves, could have been obstructing the accessibility of the His tag reducing its affinity to interact with the nickel ions. Disrupting the nanolipoprotein discs to release rPmpDauto into detergent micelles did not improve the binding of rPmpDauto to the nickel ions. It is possible that the use of denaturing conditions to fully expose the His-tag would have provided some success in binding efficiency but this would not be appropriate since the protein was required for secondary structure

196

analysis. Alternatively, additional histidines could be incorporated into the expression vector to lengthen the hexahistidine tag where in some cases the use of longer polyhistidine tags has resulted in increased purity due to the ability to use more stringent washing steps (Bornhorst and Falke 2000). However increasing the length of the Histidine tag up to 10 residues has been shown to decrease the purification yield of Histidine tagged membrane proteins by both metal affinity chromatography and by ion exchange possibly due to overall changes in ionic behaviour in pH solutions, also the position of the tag (N- or C-terminus) has been shown to have little effect upon purification yield (Mohanty and Wiener 2004). It is advised to use the smallest number of histidine residues as required for efficient purification to minimise changes to the protein that could affect protein function. In general, a six histidine tag as used in these studies is widely accepted as an appropriate choice (Bornhorst and Falke 2000). Furthermore, a longer polyHistidine tag may only be useful when the tag is located and accessible on the outer surface of the protein. As rPmpDauto was predicted to form an autotransporter pore region, of β-barrel structure, there were uncertainties over the protein confirmation and tag location. In addition, expressing an alternative autotranslocator domain Pmp G for purification via the Histidine affinity tag resulted in a similar outcome. On the basis that purification was reliant upon detection by Western blot and early purification trials led to undetectable levels of the autotranslocator domain of Pmp G (rPmpGauto) the production of a homogenous sample, sufficient for CD analysis was not realistic for this domain. The cost of pursuing the purification of these domains only had implications on the rest of the study. In order to pursue the expression and structural analysis of Pmp passenger domains, the purification of Pmp D and Pmp G autotranslocator domains were not explored further.

The passenger domain of Pmp I (rPmpI-N) was successfully expressed in a soluble state using nanolipoprotein scaffolds. Using Nickel affinity chromatography, rPmpI-N was partially purified with relative ease and was eluted sharply into a single fraction with very few contaminants that could be identified on a Coomassie SDS-PAGE gel. However for secondary structural analysis by CD a useful sample requires at least 95% homogeneity (Kelly *et al.,* 2005) making it necessary to purify further. A two step purification with Ni-NTA agarose and the cation exchanger SP sepharose was implemented with some success and the rPmpI-N domain was separated firstly from the majority of the endogenous proteins by Ni-NTA and those that co-purified during Ni-NTA affinity chromatography were removed by ion exchange chromatography.

Unfortunately, the yield in the final sample following two purification steps was poor rendering the sample unsuitable for CD analysis.

As rPmpI-N had a strong affinity for the cation exchanger SP sepharose, preliminary nickel purification was omitted to minimise the loss of protein that resulted from using a two step purification technique. However, when purifying such a heterogeneous sample directly using this method, contaminants that would have been eliminated previously by Ni-NTA methods, became bound to the ion exchange column and co-eluted with rPmpI-N therefore purification using this method alone was also unsuccessful.

Interestingly aside from the problems in obtaining sufficient quantities of purified protein, the cell-free expression and purification trials for rPmpI-N resulted in the detection of three bands of His-tagged proteins at ~44 kDa, ~38 kDa and ~22 kDa. Similar products were obtained following the expression of rPmpI-N in *E.coli*. The size of these fragments may have indicated proteolytic cleavage of the predicted native signal peptide producing a fragment ~6 kDa smaller than the expected and observed ~44 kDa band. Additionally the third fragment of ~22 kDa was detected after ion exchange chromatography when rPmpI-N was introduced into higher ionic concentrations. Some autotransporters have a cleavage site within the passenger domain and undergo proteolytic cleavage upon certain environmental triggers (Henderson *et al.,* 2004, Desvaux *et al.,* 2004a, Desvaux *et al.,* 2009). It is possible that autoproteolysis of the Pmp I passenger domain may have been displayed here. One explanation is that Pmp I functions as a secretory autotransporter in which the passenger domain is cleaved upon exposure to the external milieu during host infection as described in **section 1.4.6**. To investigate this further the fragments would need to be purified at levels high enough for proteomic analysis or for N-terminal sequencing of the fragments so that the cleavage sites could be identified. Following identification of these sites, mutagenesis could be carried out to investigate this notion. However this was hampered by the low yields that remained after purification and such methods rely on the ability to identify the bands on a Coomassie SDS-PAGE gel.

The recombinant passenger domain of Pmp A was also produced however this was only present as insoluble aggregates and not incorporated into the soluble NLPs. Bioinformatic data suggest this protein does not contain an N-terminal signal peptide (Gomes *et al.,* 2006), which may indicate that the protein is targeted to the membrane by other mechanisms. Possibly hydrophobic regions within the C-terminal

autotransloctor domain of this Pmp target this protein to the membrane. It is possible that the truncated Pmp A passenger domain cannot be targeted to the membranes, or in this case the NLPs, when lacking its hydrophobic C-terminal autotranslocator domain. However the mechanisms of targeting to the NLPs may differ from native environments and are not yet fully understood.

The purification of the recombinant Pmp domains was fraught with difficulties making structural analysis by circular dichroism unrealistic. Although soluble protein was expressed, the losses in yield that resulted from the purification steps resulted in protein levels below the detection limits of Coomassie SDS-PAGE and Western blotting making it very difficult to locate and isolate the protein of interest. While low yields or dilute samples of proteins can be concentrated using a variety of methods, the levels necessary for CD analysis would require costly scale up. In addition the optimisation steps are time consuming. Furthermore for each optimisation step, a fresh expression sample had to be obtained as the rPmps were highly susceptible to precipitation therefore storing the protein between downstream processes (at 4 °C, -20 °C and -80 °C) was not suitable. With the extremely high costs involved in expressing proteins using a cell-free system, limitations in expressing, optimising purification and scaling up the samples had to be considered.

Furthermore, a common problem with relying on the use of polyHistidine affinity tags which was observed in this study was nonspecific binding of untagged proteins. Although histidine occurs relatively infrequently (2% of all protein residues are histidine) (Schmitt *et al.,* 1993), some cellular proteins contain two or more adjacent histidine residues and so also display an affinity for the Ni-NTA matrix, co-eluting with the protein of interest, resulting in contamination of the final product. This problem is generally more pronounced in systems other than *E.coli*. Mammalian cells, for example, have a higher natural abundance of proteins containing consecutive histidine residues (Crowe *et al.*, 1994). An alternative method metal affinity purification using $Co^{2+}$-carboxylmethylaspartate ($Co^{2+}$-CMA) matrix (Talon™ resin) which displays a lower affinity for polyHistidine affinity tags than the Ni-NTA resin, resulting in elution of the tagged proteins under milder conditions (Chaga *et al.,* 1999a, Chaga *et al.,* 1999b). It has been reported to exhibit less nonspecific protein binding than the Ni-NTA resin, resulting in higher elution product purity and so may have overcome some of the problems associated with NiNTA purification in these studies.

## 4.5 Summary

Practically the expression of rPmp domains was relatively simple using a cell-free expression method specialised for the expression of membrane proteins. Although downstream purification resulted in yields insufficient to allow structural analysis, compared to the expression *E.coli*, the cell-free system provided a sample that was easier to handle with less heterogeneity and rapid results. This proved an advantageous method for the expression for rPmp domains where extended handling times cause protein precipitation. Most importantly, as soluble protein was produced for two Pmp autotranslocator domains and one Pmp passenger domain, indicating insertion into nanolipoprotein discs (Bayburt and Sligar, 2010). It is assumed that the domains would be in a native conformation ideal for functional studies (Katzen *et al.,* 2008, Katzen *et al.,* 2009) where purification was not necessary. Therefore it remains important that these domains can be expressed as soluble protein with the aid of NLPs. As NLPs mimic the lipid bilyer (Katzen *et al.,* 2009, Yang *et al.,* 2011), the investigation of Pmp domains can be explored further to investigate if and how they interact with host cells and host cell membranes.

# Chapter 5 - *In vitro* investigations of recombinant truncated Pmps in interaction assays with HeLa and Hec1B cells

## 5.1 Introduction

Chlamydiae appear to have various mechanisms in which attachment to host cells is established; thus far, the complexity of these mechanisms have not been fully elucidated and several theories remain with regards to which particular ligands and receptors are involved. To date, several ligands belonging to *Chlamydia* have been implicated in attachment to the host including heparan sulfate-like glyosaminoglycan (GAG), the major outer membrane protein (MOMP), OmcB, heat shock protein 70 (Hsp70) and the high-mannose oligosaccharide glycan moiety on such proteins (Campbell and Kuo 2006). In view of this, studies have focused on determining the type of ligand - host cell receptor combinations involved and found that *Chlamydia* may use several different receptors and that dependence was upon host cell type (Moulder 1991, Kuo *et al.,* 2002). Multiple studies have shown the importance of heparan sulfate GAGs in attachment of *Chlamydia*. GAGs bind extracellular matrix components including fibronectin, collagen and laminin (Bernfield *et al.,* 1999). *Neisseria gonorrhoeae, Leishmania donovani* and varicella zoster virus amongst other pathogens have all been shown to use heparan sulfate on the host cell as a receptor (Butcher *et al.,* 1992, Zhu *et al.,* 1995, Freissler *et al.,* 2000, Van Putten *et al.,* 2002). However studies with *C.trachomatis* have been inconclusive in determining if the GAG is displayed on the organism or on the host cell to initiate attachment. MOMP, also implicated in attachment through heparan sulfate on epithelial cells has been shown to interact through hydrophobic interactions and through possessing glycan moieties (Campbell and Kuo, 2006). These glycan moieties consist of mannose and mannose-6-phosphate residues and studies have hypothesised that Chlamydiae also use the mannose receptor and the mannose-6-phosphate receptors for attachment (Swanson and Kuo 1994, Kuo *et al.,* 1996, Kuo *et al.,* 2002, Kuo *et al.,* 2004). In addition, the oestrogen receptor on human endometrial cells has been revealed as an important receptor for the attachment of *C.trachomatis* serovars B and E (Davis *et al.,* 2002). Earlier studies had discovered electrostatic interactions were important for the binding of *Chlamydia* to host cells and demonstrated that positively charged polycations such poly-L-lysine, divalent cations such as $Ca^{2+}$, $Mg^{2+}$ and $Mn^{2+}$ and monovalant cations of $Na^+$ all increased the binding of different *Chlamydia* species to cells, whilst the addition of negatively charged heparan sulfate inhibited the attachment and entry again indicating its importance (Kuo *et al.,* 1973, Hatch *et al.,* 1981, Sneddon and Wenman 1985). Collectively, this information details a complex array of adherence mechanisms displayed by Chlamydiae.

202

Bioinformatic studies have revealed characteristics about the Pmps that would suggest these proteins also act as potential adhesion molecules involved in host cell attachment especially since the Pmp proteins contain repeated signature motifs GGAI and FXXN. These signature motifs have been implicated in adhesion for other microbial virulence factors such as OmpA of *Rickettsia* spp. (Grimwood and Stephens, 1999). Furthermore, Pmps have been implicated in tissue specific tropism (Stothard *et al.,* 2003, Gomes *et al.,* 2006), adding weight to the adhesion theory such that Pmps function to target specific host tissue types. Additionally, truncated portions of Pmp D were expressed as inclusion bodies and identified as a stimulatory antigen, recognised by CD4+ cells (Goodall *et al.,* 2001) and Pmp F was predicted to possess T-cell epitopes that bind human leukocyte antigen (Carlson *et al.,* 2005), both indicating surface interactions with human cells.

Since these discoveries, few studies have proceeded to demonstrate attachment of Pmps *in vitro*. The studies that have explored this notion have shown that the N-terminal passenger domain of the Pmp D homologue, Pmp21 from *C.pneumoniae* translocates to the outer membrane and antibodies raised against this, have blocked chlamydial infectivity (Wehrl *et al.,* 2004). More recently, Pmp G and Pmp D homologues, Pmp6 and Pmp21 respectively, from *C.pneumoniae,* were shown to adhere to human epithelial cells whilst using truncated Pmp domains on yeast cells and latex beads (Mölleken *et al.,* 2010a). However these studies only examined the central passenger domain of the proteins and so lacked data associated with adherence of the C-terminal autotranslocator domains or the signal leader peptides (although it is expected this would not be displayed on mature Pmps). These studies are by no means conclusive, and little attention has been paid to any of the nine Pmps (A-I) of *C.trachomatis* with regards to adherence and interaction with mammalian cell surfaces.

In this study, with the use of nanodiscs to mimic the lipid bilayer, truncated rPmps were expressed for use in protein interaction assays with cervical (HeLa) and endometrial (Hec1B) cells of the female urogenital tract to screen for putative ligand-receptor interactions. Several techniques were explored including cross-binding assays using chemical cross-linkers after incubating rPmp domains with whole cells, immunofluorescence to examine the possibility of rPmp domains in localisation to the host cell surfaces and finally biophysical techniques using surface plasmon resonance (SPR) and ellipsometry as a ligand fishing approach utilising lysed mammalian cells as potential orphan receptors.

The particular rPmps domains used in these screening trials were the (N-term) passenger domain of Pmp I and the (C-term) autotranslocator domains of Pmp D and Pmp G. These truncated rPmps were expressed as soluble proteins with the aid of NLP nanodiscs (**chapter 4**). However, as purification of the rPmp domains was challenging, producing only partially pure protein, which resulted in low protein yields and increased handling time escalating the chances of degradation and precipitation, these rPmp domains were not isolated from other MembraneMax™ endogenous proteins. Instead, as a control the empty vector was applied to the MembraneMax™ expression system and was used in control assays to eliminate potential false positives. The N-terminal passenger domains of Pmp D and Pmp G have already been suggested to carry adherent mechanisms in the homologue *C.pneumoniae* (Mölleken *et al.*, 2010a).

## 5.1.1 Chemical cross-linking

Chemical cross-linking is a process whereby two proteins or protein regions are fixed through covalent bonds formed at the end reaction groups present on the cross-linker with functional groups on the proteins (Tang and Bruce 2009). Cross-linkers are used for identification of closely associated protein relationships and ligand-receptor interactions. The cross-linkers used to identify surface receptors to their ligands often require a cross-linker with a longer space arm compared to a zero or short spacer arm used to fix subunits within a protein. The most common reactive groups on cross-linkers are *N*-hydroxysulfosuccinimide (NHS) esters, malemaides (sulfhydryl-reactive) and non-specific photo-reactive groups.

## 5.1.2 Immunocytochemistry

Immunocytochemistry (ICC) can be a useful tool to investigate the spatial distribution of proteins within cultured cells using antibodies directed against the protein of interest. Immuofluorescence is dependent upon the coupling of fluorophores to antibodies, which when bound to the antigen of interest and excited by a light source, such as UV light, emit light that is visible under a fluorescent microscope to allow identification of antigen/protein localisation. Prior to antibody application the cells are fixed to retain cellular and subcellular structures, interactions and to fix the distributions of proteins. Permeabilisation of the cellular membranes is necessary to allow entry of the antibody and can be accomplished by using an organic solvent such as acetone which also fixes the cells. A blocking step is implemented with a non-specific protein such as bovine serum albumin (BSA) or using serum to which the antibody is raised in order to reduce non-specific binding to endogenous Fc

receptors (FcR) present on the surface of some cells (Buchwalow *et al.,* 2011) . The cells are washed between each step to ensure unbound proteins, antibodies and other reagents are removed. Determination of a suitable staining method is essential for successful amplification and identification of protein-cell and protein-protein interactions.

Two types of immunofluorescence exist: direct and indirect, the method of choice is often determined upon the necessity to amplify the target signal where the target antigen/protein may be present in low concentrations or where the epitope may be obstructed. A direct method of immunofluorescence utilises a primary antibody conjugated with a fluorochrome, this conjugate is directed to the antigen of interest and the method overall involves fewer steps. However, the technique is limited as only a certain amount of primary antibodies can access the antigen of interest before saturation is met. Using an indirect method, a secondary antibody that targets the primary antibody is used to amplify this signal. This secondary antibody is conjugated to a fluorophore and is directed against the species of immunoglobulin (Ig) that the primary antibody was raised. For the purposes of this study, as the recombinant protein was of unknown concentration in a heterogeneous mix and the amount necessary to establish an interaction with the cells was unknown an indirect method is the preferred method, using two antibodies to visualise the protein/antigen and so an indirect method was chosen to ensure that if such an interaction or localisation had taken place, this would increase the likelihood of this being detected.
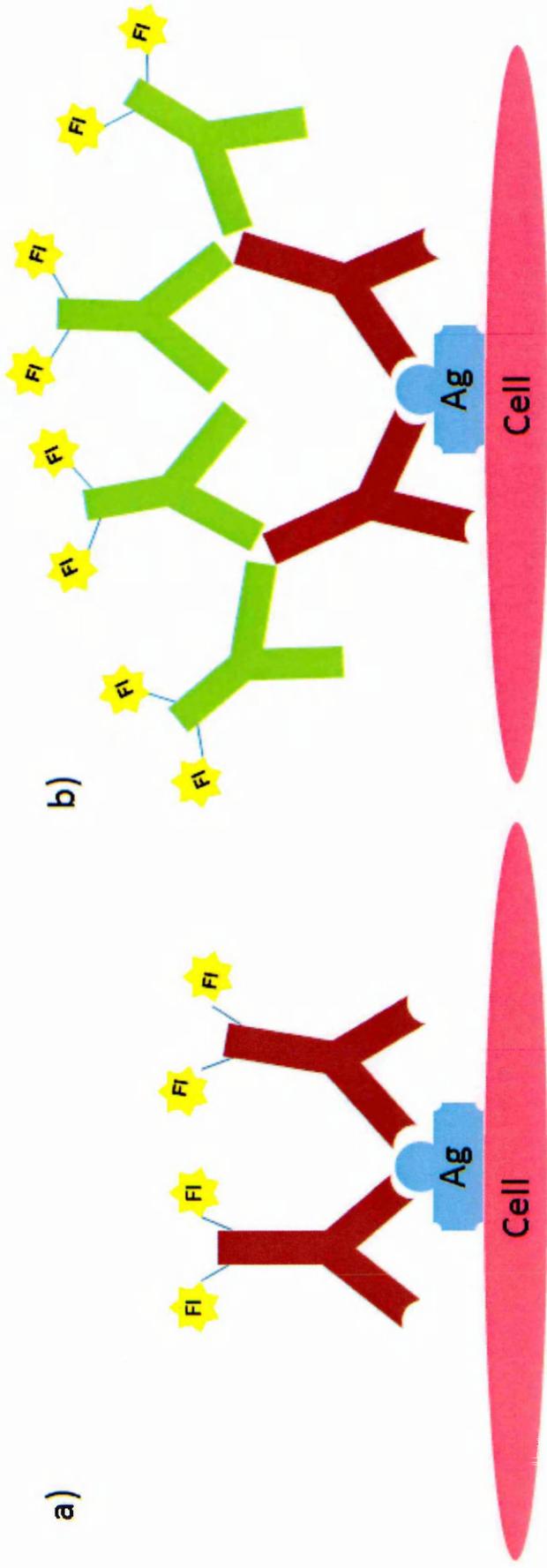
**Figure 5.1 - Direct and indirect immunofluorescence of a cell surface antigen.** (a) A primary antibody, to the antigen (Ag) of interest is conjugated to the fluorochrome (Fl) to give direct immunofluorescence. (b) A secondary antibody, specific to the unlabelled primary antibody, is conjugated to a fluorochrome using an indirect method of immunofluorescence. Upon excitation, the fluorochrome emits light visible by fluorescent microscopy. Indirect immunofluorescence can provide a higher amplification signal.

### 5.1.3 Surface Plasmon Resonance

Over recent years surface plasmon resonance (SPR) biosensors have become an established method of measuring molecular interactions. SPR-based instruments use an optical method to measure the refractive index near a sensor surface to which a ligand is immobilised and the interaction with analytes, passing over in solution are monitored (Cooper, 2002). Binding of molecules in solution to surface-immobilised receptors alters the refractive index of the medium near the surface (Cooper, 2002). Since the development of the first biosensors in the late 1980s, the use of optical biosensors in research and development has expanded. These include ligand fishing (Zhukov *et al.,* 2004, Ravanat *et al.,* 2009), epitope mapping (Jokiranta *et al.,* 2000, Fägerstam *et al.,* 2004), and cell adhesion (Nielsen *et al.,* 2000) amongst many others.

Using Biacore SPR analysis, the ligand is immobilised to a sensor chip surface, the surface is comprised of a 50 nm layer of gold sandwiched between a glass layer and the micro-fluidic cartridge, in which the sample analyte solution flows through and passes over the ligand (**Figure 5.2**). A polarised laser light directed through a prism with a high refractive index, focused on the back of the sensor chip, is reflected back to a diode array detector unit that monitors the intensity of reflected light (Beseničar *et al.,* 2006). At a critical angle of incident light, surface plasmons are generated at the surface of the gold layer. This absorbs the light and is visible as a decrease in the intensity of reflected light. When a change in mass occurs near the sensor chip surface, as a result of a binding event or a change in solution/buffer in the micro-fluidic system, the angle of light at which the SPR occurs shifts due to a change in the refractive index near the sensor chip surface (Jason-Moller., *et al* 2006). This is measured in real-time and plotted on a sensorgram with response measure in resonance units (RU). Each RU is equal to a critical angle angle shift of $10^{-4}$ degrees with a linear relationship between the mass of the components at the surface of the chip and the RU such that 1 RU is equal to 1 pg/mm$^2$ (Stenberg *et al.,* 1991).

One of the most widely used SPR systems is the Biacore, produced by BIAcore AB (Biomolecular Interaction Analysis), (GE Healthcare). Biacore instruments are capable of characterising binding interactions in real-time without the need for a labelled analyte (Rich and Myszka 2000). Identification of binding interactions has been achieved with a variety of samples ranging from proteins, nucleic acids, small molecules to more complex sample mixtures including lipid vesicles, viruses, bacteria and eukaryotic cells (Jason-Moller., *et al* 2006).

## 5.1.3.1 Biacore X apparatus

Biacore X, used in these studies is a highly sensitive, semi-automated system suited for low-sample throughput. To understand and develop optimised methods, a working knowledge of the internal micro-fluidic set up was necessary. Briefly, the sensor chip surface forms the floor of the small flow cell which an aqueous solution (running buffer) passes under continuous flow. As described in **section 5.1.3** to test for an interaction the ligand is fixed to the surface and the analyte is flushed over in sample buffer under continuous flow. A change in mass as a result of a binding event increases the refractive index to form a response curve/peak. Importantly, a background response will also be generated if there is a change in the running and sample buffers and so this change in refractive indices must be subtracted by using a reference flow cell (**Figure 5.3**). The integrated fluidic cartridge in the Biacore X instrument contains only two flow cells allowing the immobilisation of two molecules on one chip or set-up with one flow cell used as a reference and the other immobilised with a ligand of interest (**Figure 5.4**). The system has a dual-channel detection of the SPR signal such that simultaneous monitoring of both flow cells can be used and the instrument software allows in-line reference subtraction of the control surface signal from the sample injection (Biacore, 2002). One advantage of investigating interactions by this method, due to the high sensitivity, was that only small amounts of sample were needed and where strong interactions take place, only nanomolar concentrations may be necessary.

Biacore have a range of sensor chips suited to different applications, all are coated with a thin uniform layer of gold, gold is amenable to covalent attachment of surface matrix layers and in physiological buffers it is mostly inert and so is well suited for SPR applications. The original and most widely used due to its versatility is the CM5 chip coated with a carboxymethylated dextran layer for the coupling of ligands to the surface by covalent attachment via amino, thiol, aldehyde or carboxyl groups (Beseničar *et al.,* 2006). Molecules bound to the dextran matrix are able to behave similarly as in solution due to the high flexibility of the matrix (Day *et al.,* 2009).
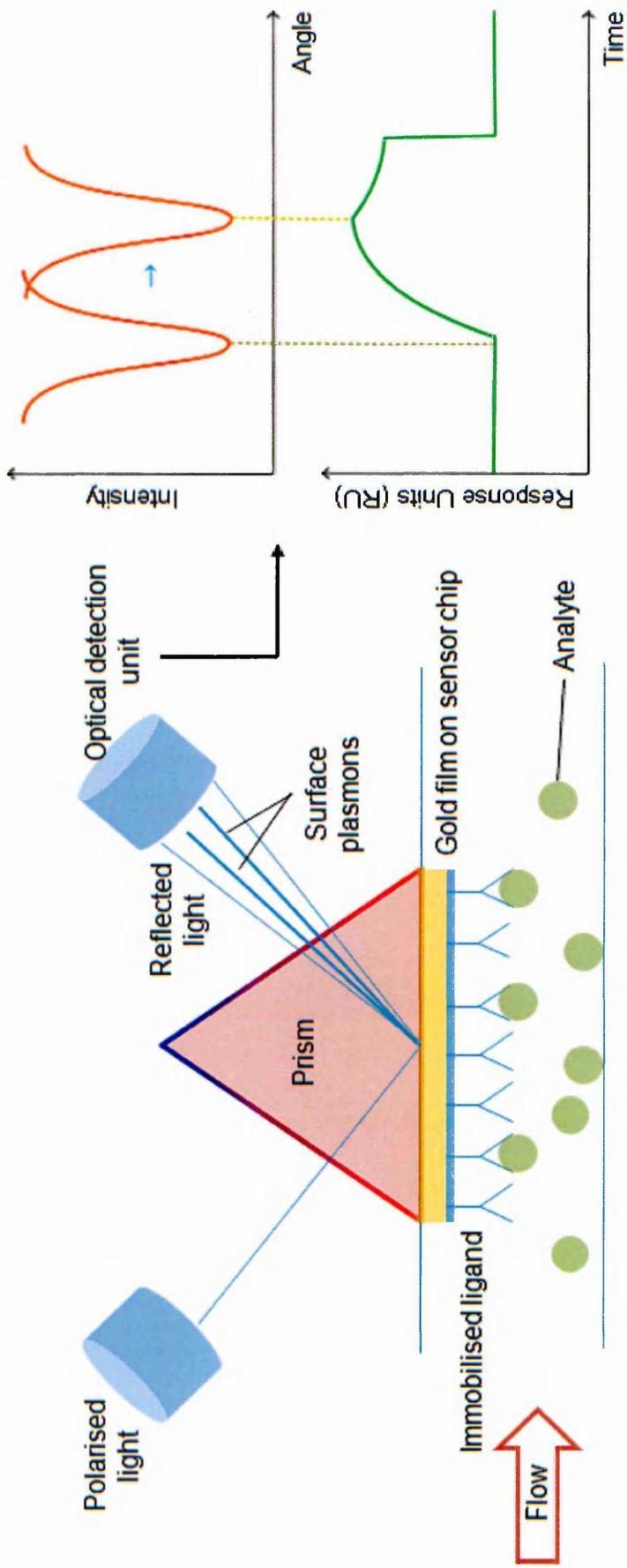
**Figure 5.2 - Schematic diagram showing the principals of surface plasmon resonance.** SPR detects changes in the refractive index in the reflected light, a shift in the refractive index due to binding events is detected as a decrease in intensity. This forms a sensorgram of response units in real-time (bottom right). (Adapted from (Cooper, 2002)).

209

Reference surface      Active surface      Reference subtraction
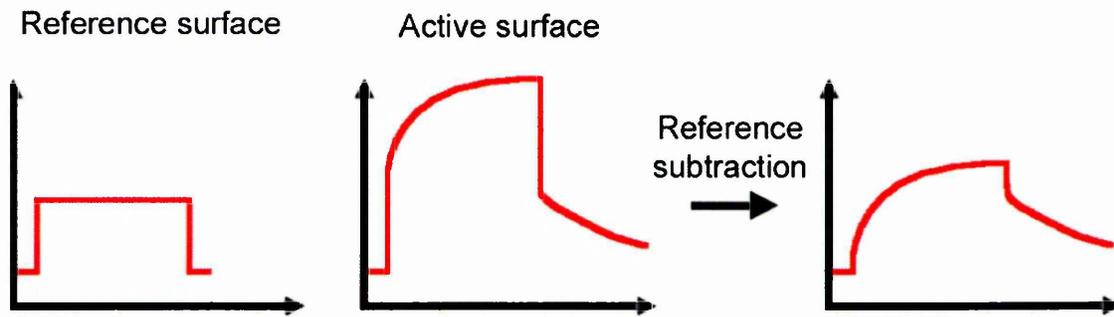
**Figure 5.3 - Schematic diagram to depict the bulk subtraction of reference control from a binding event.** The sensorgram from the reference surface shows a bulk contribution from the sample buffer while the active surface shows association of ligand/analyte with the surface and an increase in baseline as a consequence of binding. This surface also includes the bulk contribution shown on the reference surface, when the reference bulk is subtracted from the active surface, only the binding response is shown. (Biacore X getting started manual, 2002).
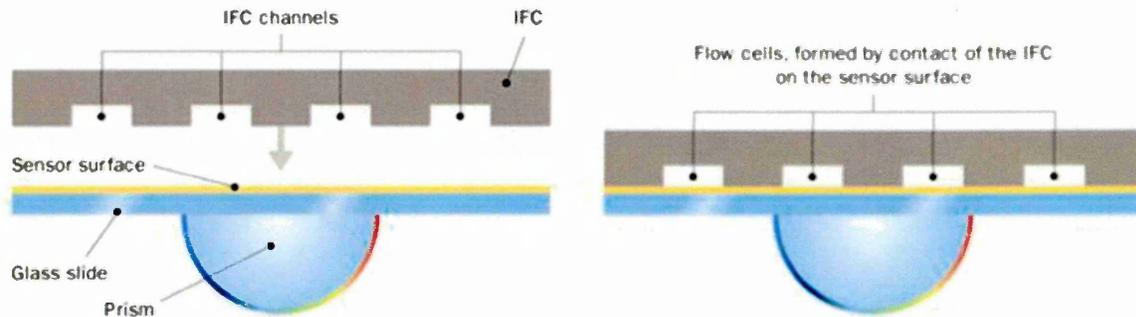


**Figure 5.4 - Schematic diagram depicts the flow cells over the sensor chip surface in Biacore instruments.** Flow cells are formed when the microfluidic cartridge is pressed against a sensor surface. All sample solutions, running buffers, analytes etc. are delivered through each flow cell to the surface of the sensor chip by the integrated micro-fluidics cartridges IFC. The diagram shows 4 flow cells, the Biacore X used in this study contains only 2 flow cells. (www.biacore.com/lifesciences/technology/introduction/Flow_cells/index)

210

## 5.1.4 Ellipsometry

Ellipsometry is an optical biosensor technique used to investigate and characterise film thickness of single layers or complex multi-layers of material by use of polarised light similar to conventional SPR (Nabok *et al.*, 2011). A change in polarised light reflected off a sample yields information about the layers with measurements sensitive to detect changes ranging from a few angstroms to several micrometers with excellent accuracy. Total Internal Reflection Ellipsometry (TIRE) combines the highly sensitive spectroscopic ellipsometry with the experimental conveniences of SPR. Where conventional SPR monitors the intensity of reflected polarised light, TIRE detects two parameters, the amplitude ratio ψ and the phase shift Δ of p- and s-components of reflected light shown in **Figure 5.5**. The method of total internal reflection ellipsometry has shown to be more accurate and sensitive when compared to SPR (Nabok *et al.*, 2005).
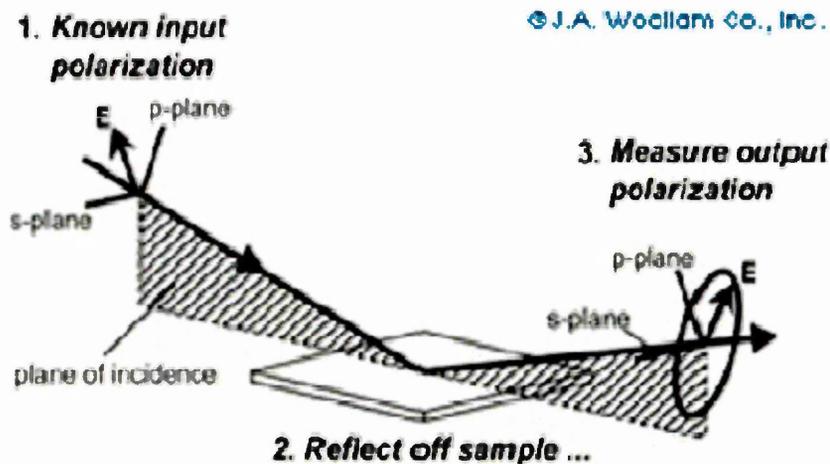


**Figure 5.5 - Principals of TIRE reflected light.** (1) The incident light is linear with both p- and s- components prior to meeting sample material (2). The reflected light has undergone amplitude and phase changes for both p- and s- polarised light (3), and ellipsometry measures these changes. (http://www.jawoollam.com/tutorial_4.html).

## 5.2 Materials/methods

### 5.2.1 Laboratory reagents and equipment:

Analytical grade chemicals were used throughout this study with all solutions being prepared using sterile water. Materials were obtained from the following suppliers.

**Abcam**, Cambridge; **Bio-Rad Laboratories Ltd.**, Hemel Hemsted, Hertfordshire; **GE Healthcare Ltd**, Amersham, Buckinghamshire; **Generon.**, Maidenhead, Berkshire; **Invitrogen Life Technologies.**, Paisley; **New England Biolabs (UK) Ltd.**, Hitchin, Hertfordshire; **Novagen, Merck Chemicals Ltd.**, Beeston, Nottingham; **Qiagen Ltd.**, Crawley, West Sussex; **Sigma Aldrich Company Ltd.**, Gillingham, Dorset; **Starlab UK.**, Blakelands, Milton Keynes; **Thermo-Fisher Scientific UK.**, Loughborough, Leicestershire; Protein expression; **Vector Laboratories Ltd.**, Orton Southgate, Peterborough

### 5.2.2 Protein Expression

Recombinant Pmp domains (rPmpI-N, rPmpDauto and rPmpGauto) were expressed in the cell-free expression system in 100 $\mu$l volumes as in **section 4.2.1**. Insoluble fractions were discarded. HALT protease inhibitors (Thermo Scientific) were applied to the soluble protein fractions prior to interaction assays. Recombinant Pmp domains were not subject to purification for interaction studies.

### 5.2.3 Mammalian cell culture methods

All cell culture reagents, medium and supplements were obtained from Gibco™, Invitrogen UK unless otherwise stated. All cell culture methods were carried out under sterile conditions in a class II laminar flow cabinet. HeLa cervical adenocarcinoma cells and Hec1B endometrial adenocarcinoma cells were both obtained as passage 2 stocks, as a gift from Dr Adrian Eley (University of Sheffield).

### 5.2.3.1 Cell line stocks and recovery

HeLa and Hec1B cells were stored long-term in cryovials immersed in liquid nitrogen until required for experimental work. Cells were preserved in 1 ml of Fetal Bovine Serum containing 10% dimethyl sulphoxide (DMSO) at a density of 1 x 10$^6$ cells per/ml. Upon recovery, cells were placed in 37 °C incubator to thaw for 2 minutes. Upon thawing the cells were transferred into 25 cm$^2$ filter-cap (T25) flasks containing 9 ml of the appropriate pre-warmed, complete cell culture medium and allowed to adhere to the bottom of the flasks overnight at 37 °C in a humidified incubator containing 95% air and 5% $CO_2$. Media was changed twice weekly and upon 80%

confluence, cells were transferred into 75 cm$^2$ filter-cap (T75) flasks to maintain stocks for experimental assays.

### 5.2.3.2 Cell harvesting and serial passage of cells

Thawed stocks of adherent HeLa cervical adenocarcinoma cells and Hec1B endometrial adenocarcinoma cells were routinely grown in a humidified incubator at 37 °C with 5% CO$_2$ in 75 cm$^2$ flasks immersed in specified media with supplements as in **Table 3.1**. When cell monolayers were determined near-confluent at ~80%, cells were passaged into new T75 flasks by trypsinisation. Briefly, the culture medium was removed and the cell monolayer was washed with 10 ml Dulbecco's phosphate buffered saline (DPBS) without calcium or magnesium to remove remaining media. Cells were detached from the flask using 5 ml Trypsin/EDTA solution (0.5% trypsin, 0.2% EDTA) and incubating at 37 °C for 5 minutes followed by gentle tapping of the flask. The enzymatic action of the trypsin was quenched by the addition of 5 ml of complete media to the cell suspension and gentle pipetting before transfer to a sterile Falcon tube where the cell pellet was collected by centrifugation at 650 x g for 5 minutes. The supernatant was discarded and the collected cells were washed in DPBS, pelleted as before and resuspended in a known volume of culture media for either cell counting (**section 5.2.3.4**) for experimental assays, or were seeded at 1:6 ratio for further propagation of cells.

### 5.2.3.3 Preparation of HeLa cell lysates for Biacore and Ellipsometry

HeLa cells suspended in ice cold PBS with protease inhibitors were lysed by sonication with a Sonics Vibra Cell probe sonicator (probe tip diameter 3 mm) on ice with five 10 second pulses at 30% amplitude. Centrifugation at 2900 xg for 20 minutes removed unbroken cells and nuclei (Millipore protein extraction protocols). Supernatant containing cell membranes and cytosol was diluted 1 in 10 in running buffers for application in Biacore SPR and ellipsometry. All cell preparations for these applications were performed fresh and were used immediately.

### 5.2.3.4 Cell counting

Cell counting was carried out using a haemocytometer. Briefly, 10 µl of cell suspension from **section 5.2.3.2** was mixed thoroughly with the same volume of 0.4% Trypan Blue. This viable/non-viable cell staining method is based on the principle that certain cells i.e. with compromised membranes, allow entry of the trypan blue dye to allow calculation of the percentage viability of the cells. Counting the number of cells in 5 of the large outer squares (outer 4 corners and central

square) and multiplying by 50,000 gave the number of cells per ml. To calculate the percentage of viable cells the following equation was used [1.00 − (number of blue cells ÷ number of total cells)] × 100. To seed cells for experimental assays only the viable cell count was used to calculate the dilution factor of the cell suspension to obtain the desired density.

| Cell | Cell Type | Origin | Obtained from | Culture medium composition |
|------|-----------|--------|---------------|----------------------------|
| HeLa | Cervical adenocarcinoma | Human | Dr. A. Eley, Medical School, University of Sheffield, UK | 500 ml Dulbecco's modified eagle's medium (DMEM), high-glucose with GlutaMAX™ <br> 10% (v/v) Heat inactivated foetal bovine serum (HI-FBS) <br> 1% (v/v) Penicillin/streptomycin (100 U/ml/50μg/ml) |
| Hec1B | Endometrial adenocarcinoma | Human | Dr. A. Eley, Medical School, University of Sheffield, UK | 500 ml McCoy's 5A medium <br> 10% (v/v) Heat inactivated foetal bovine serum (HI-FBS) <br> 1% (v/v) Penicillin/streptomycin (100 U/ml/50μg/ml) <br> 1% (v/v) L-Glutamine (2mM) |

Table 5.1 - Summary of mammalian cell lines and growth culture requirements.

## 5.2.4 Chemical cross-linking

### 5.2.4.1 Recombinant protein treatment of cells for chemical cross-linking assays

Cross-linking can be performed on cells in suspension or on adherent cells in culture plates. In the latter, adherent cells have less exposed cell surface area, limiting potential cross-linking therefore assays were carried out with cells in suspension. Hec1B and HeLa cells were seeded into 15 ml falcon tubes at a density of $25 \times 10^6$ cells per ml in 1XPBS, pH 8.0, this calculation had to take into account the amount of protein suspension that would be added to ensure a final volume of 0.5 ml was not exceeded to prevent dilution of the recombinant Pmp protein which limits downstream detection of His-tagged proteins and conjugates. A high concentration of cells was preferred to maximise the chances of identifying an interaction with recombinant protein. Briefly, the medium was removed and the cells were washed three times with DBPS to remove amine containing components that would quench the cross-linking reaction. The total soluble protein fractions from the cell-free protein expression were incubated with the mammalian cells for one hour with gentle agitation. Initially this was done using 200 µl of cell-free expression of rPmp domain (**section 4.2.1**) (300 µg total protein), for each assay so that His-tagged rPmp domains were detectable via Western blot. This interaction was carried out at 37 °C to mimic *in vivo* *C.trachomatis* attachment (Söderlund and Kihlström, 1983; Söderlund and Kihlström, 1988; Fadel and Eley, 2008). Finally any interactions were cross-linked as in **section 5.2.4.2**. Adaptations to optimise these methods were subsequently made and the remainder of the incubation assays were performed at 4 °C to reduce recombinant protein precipitation.

### 5.2.4.2 Chemical cross-linking with BS3

BS3 (Sulfo-DSS) is a homobifunctional amine-amine NHS ester with an 8-carbon spacer arm as seen in **Figure 5.6** (Thermo Scientific). Generally, proteins have several lysine residues rich in amines exposed via side chains in addition to the N-terminus of each polypeptide making ample targets for NHS ester cross-linking. With appropriate controls, cross-linking experiments can allow unambiguous identification of protein interactions when used together with a variety of techniques including Western blotting, immunoprecipitation and mass spectrometry (MS). In this case, the target proteins carried a polyHistidine tag which could be probed using the established Western blotting technique using the penta-His antibody which has been

used to identify rPmps (**section 5.2.6.1**). BS3 was chosen because it is membrane impermeable making it ideal for cell-surface interactions, and most importantly it is soluble in water and can be used in a range of common buffers where the use of organic solvents may disturb protein structure and activity. The cells were seeded as in **section 5.2.3.2** to a density of ~25 x $10^6$ cells/ml in PBS (pH 8.0) following the guidelines set out by Thermo Scientific. Cross-linking with homobifunctional linkers is a shotgun approach for capturing an interaction and according to the manufacturer's recommendations high quantities of cross-linker and targets are often necessary to increase the capture of such an interaction. Without authentic positive controls and following the manufacturer's instructions BS3 is effective when used at a final concentration of between 1 and 5 mM (Thermo Fisher). The higher end of this concentration was used throughout the cross-linking studies to ensure that all interactions were captured. This final concentration of 5 mM BS3 was added to the end of the protein/cell incubation, from a freshly made 100 mM stock dissolved in PBS pH 8.0. Stocks were used immediately as BS3 is rapidly hydrolysed. The reaction was cross-linked for 30 minutes at 4 °C followed by quenching with a final concentration of 20 mM Tris for 15 minutes at room temperature. The cross-linked samples were prepared for analysis as in **section 5.2.4.4.**
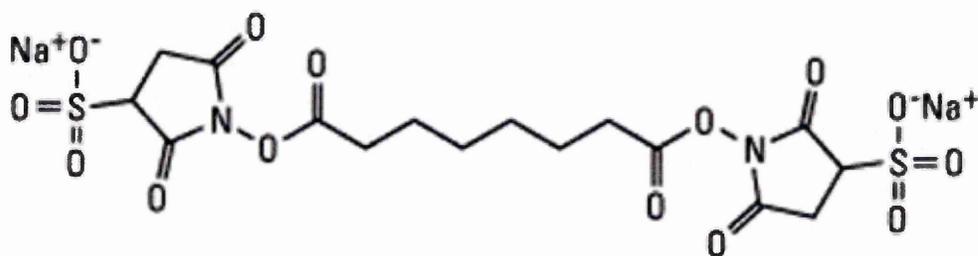


**BS3**
Bis(sulfosuccinimidyl) suberate
MW 572.43
Spacer Arm 11.4 Å

**Figure 5.6 - Chemical structure of BS3**. From http://www.piercenet.com

### 5.2.4.3 Chemical cross-linking with formaldehyde

Formaldehyde is membrane permeable therefore cross-linking occurs both intracellularly and extracellularly. Formaldehyde was used to fix interactions at a final concentration of 1% for 1 hour at room temperature. The reaction was quenched with 300 mM glycine for 30 minutes at room temperature. An aliquot of the crude assay was added to SDS loading buffer for SDS-PAGE and Western blotting and the supernatant and cell pellets were collected as in **section 5.2.4.4**.

### 5.2.4.4 Fractionation of cross-linked assays

Cells were pelleted for 5 minutes at 500 x g at room temperature. The supernatant was retained as the unbound fraction and an aliquot was added to SDS loading buffer. The cell pellets were lysed directly by 5 brief pulses with a Jencons Sonics Vibra Cell probe sonicator (probe tip diameter 3 mm) in SDS loading buffer and passed several times through an Hamilton syringe to reduce viscosity. These fractions were analysed via Western blot.

### 5.2.5 Immunocytochemistry

### 5.2.5.1 Recombinant protein treatment of cells for immunocytochemistry (ICC)

For immunocytochemistry analysis, Hec1B and HeLa cells were seeded into 8-well chamber slides in 400 µl volumes at a density of 1 x $10^5$ cells per ml in appropriate complete media (**section 5.2.3.2**) and allowed to adhere to the slides for 24 hours at 37 °C in a humidified incubator prior to treatment with recombinant protein samples. Following this, the media was removed and the cells were washed three times with DPBS. The total soluble protein fractions from a 100 µl cell-free MembraneMax™ protein expression sample (**section 4.2.1**) were diluted in serum-free media and incubated with the mammalian cells for one hour at 37 °C with 5% $CO_2$ to mimic *in vivo* attachment conditions (Söderlund and Kihlström 1983, Söderlund and Kihlström 1988, Fadel and Eley 2008). A range of protein dilutions were used during optimisation to ensure the recombinant protein was detectable. A 100 µl of diluted expression sample was added to each well. The recombinant protein samples were aspirated and cells washed prior to ICC staining as in **section 5.2.5.3**.

### 5.2.5.2 Sudan black B staining

Sudan black B (SBB) is a fat-soluble dye commonly used to reduce autofluorescence of tissues and cells caused by high levels of lipofuscin, a

fluorescent pigment which accumulates with age in the cytoplasm of cells, which because of its broad excitation and emission spectra, it often interferes with fluorescent microscopy (Viegas et al., 2009). Using 0.3% SBB has previously been shown to be sufficient to quench lipofuscin autofluorescence (Schnell et al., 1999). 0.3% SBB stock is prepared in 70% ethanol, stirred for 2 hours at RT, followed by filtering with Whatman paper and stored wrapped in foil at 4 °C. This was applied to cells for 10 minutes at RT prior to mounting the slides during ICC.

### 5.2.5.3 Indirect immunofluorescence

In order to determine if recombinant truncated Pmps interact with the cell surface of endometrial and cervical cells, immunocytochemistry (ICC) was carried out after incubating truncated Pmps with Hec1B and HeLa cells. Following incubation the cells were washed three times in PBS for 5 minutes to remove non-interacting and non-specifically bound proteins. The cells were fixed with ice cold acetone for 5 minutes and left to air dry for 15 minutes at room temperature The cells were blocked with 1XPBS, 25% goat serum, 1% BSA for 45 minutes at RT. The cells were stained for polyhistidine tagged proteins using rabbit polyclonal 6X His-tag antibody (Abcam) and/or mouse monoclonal Penta-His antibody (Qiagen) as in **Table 5.2**. Preliminary experiments were performed to determine the optimum concentration of Ab for ICC, with each antibody diluted within a range recommended by the supplier. Primary antibodies diluted in DPBS, 1% BSA, were applied to each well and incubated for one hour at RT before three 5 minute washing steps to remove unbound antibody. To detect bound primary antibodies, secondary antibodies of fluorescein isothiocyanate (FITC) conjugated to goat anti-rabbit IgG (for rabbit polyclonal His-tag antibody) and goat anti-mouse IgG (for mouse monoclonal Penta-His antibody) were applied diluted in PBS and incubated for 45 minutes at RT in the dark. Cells were washed three times to remove unbound secondary antibody prior to staining with Sudan black B to reduce autofluorescence in HeLa cells (**Figure 5.7**). Antibody controls were also performed with untreated cells using primary and secondary antibodies independently to check for non-specific background signal, (**Figure 5.7** and **Figure 5.8**). Slides were then mounted using Vectashield™ hard-set mounting medium with 4', 6-diamidino-2-phenylindole (DAPI) (Vector laboratories, UK), which stains the cell nuclei blue. The edges were sealed with clear varnish and slides stored in the dark at 4 °C until analysed.

## 5.2.5.4 Light and fluorescent microscopy

ICC images were captured with equal exposure times using an upright Olympus BX60 fluorescent microscope equipped with a cool-snap pro (Cybernetics) digital camera to acquire images into Labworks™ where they were saved without manipulation.

| Primary antibodies | | | | |
|---|---|---|---|---|
| Immunogen | Species | Clonality | Dilution | Source |
| 6X His-tag -HHHHHH- | Rabbit | Polyclonal | 1:50,1:100,1:250, **1:500**,1:750,1:100, 1:2000,1:5000, | Abcam, UK |
| Penta-His -HHHHH- | Mouse | Monoclonal | 1:50, 1:100, 1:250, 1:500 | Qiagen, UK |
| Secondary antibodies | | | | |
| Immunogen | Species | Fluorochrome | Dilution | Source |
| Rabbit IgG | Goat | FITC | 1:250, **1:350** | Abcam, UK |
| Mouse IgG | Goat | FITC | 1:250, 1:350, 1:500 | Invitrogen, UK |

**Table 5. 2 - Summary table of antibodies used for immunocytochemistry**. Details of primary and secondary antibodies used for immunocytochemical detection of recombinant His-tagged Pmp domains in interactions with Hec1B and HeLa cells. The antibodies used for the final immunocytochemistry assays are highlighted in bold.
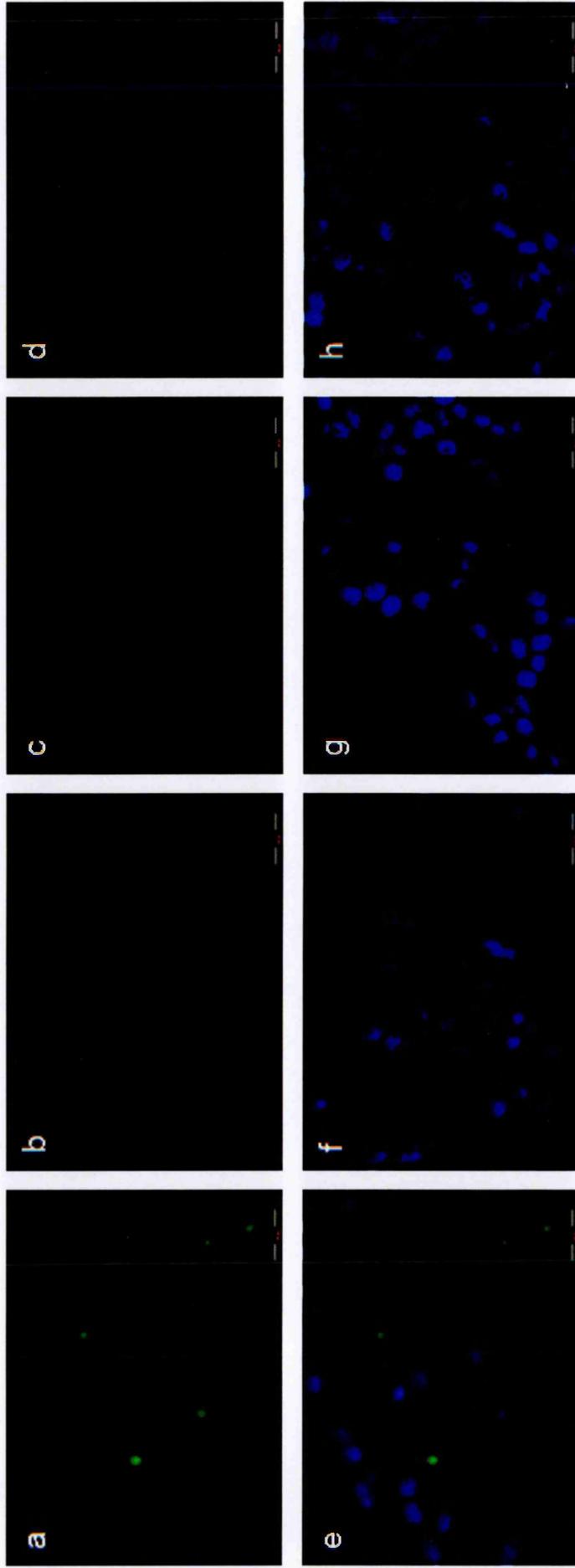
**Figure 5.7 - HeLa controls with FITC (above) and DAPI staining overlay (below).** (a) and (e) antibody staining omitted, cells exhibiting autofluorescence, (b) and (f) antibody staining omitted, cells treated with 0.3% sudan black, (c) and (g) primary anti-His antibody 1:500 (Abcam), (d) and (h) secondary anti-rabbit FITC 1:350 (Abcam). *Scale bar is 50 μm*

220
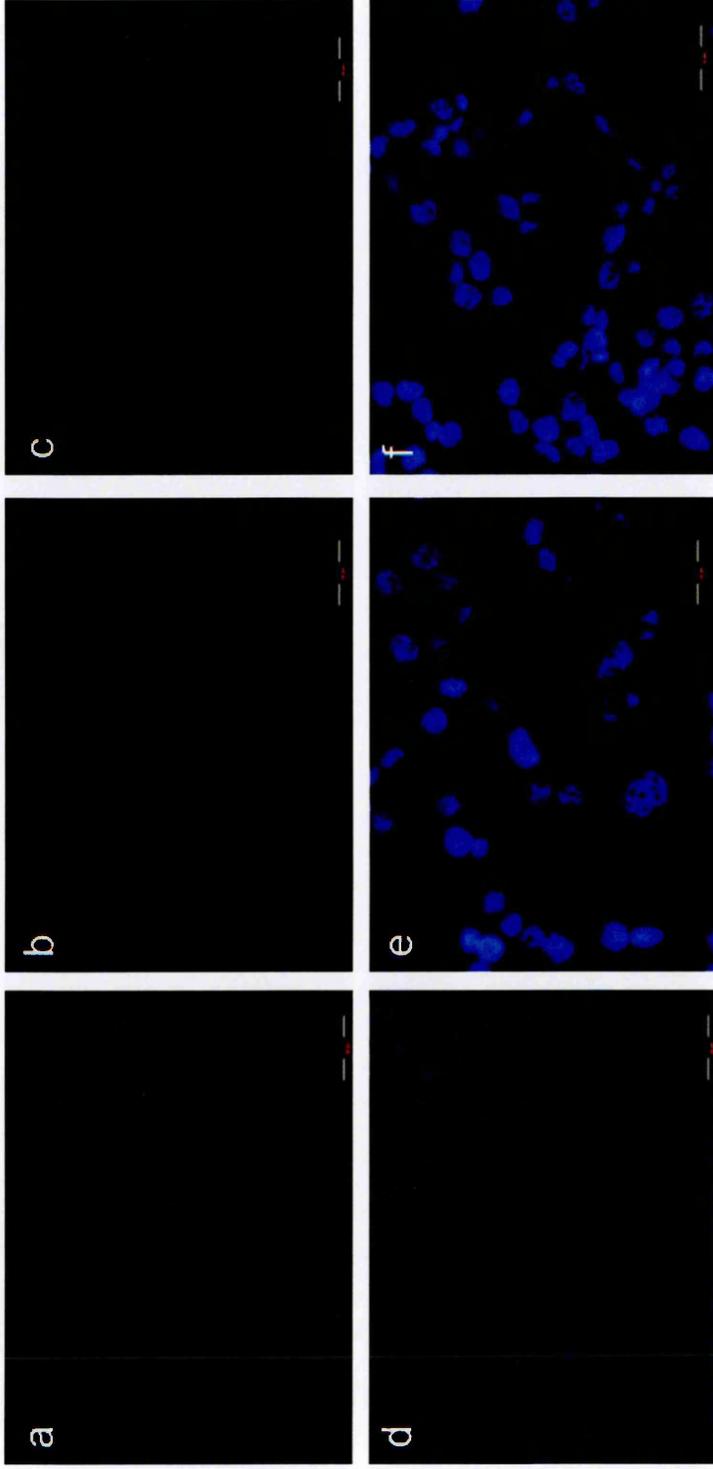
**Figure 5.8 - Hec1B controls with FITC (above) and DAPI staining dual overlay (below).** (a) and (d) antibody staining omitted, cells exhibiting no autofluorescence, (b) and (e) primary anti-His antibody 1:500 (Abcam) (c) and (f) secondary anti-rabbit FITC 1:350 (Abcam). *Scale bar is 50 μm*

221

## 5.2.6 Immunoblotting

### 5.2.6.1 Western blotting with mouse monoclonal penta-His and HRP conjugated secondary antibody

The technique for Western blotting using penta-His primary with HRP secondary is described previously in **section 3.2.5.2**.

### 5.2.6.2 Dot blotting

10 µl of each sample for analysis was applied directly on to a nitrocellulose membrane, allowed to dry for 5 mins before the antibody staining protocol in **section 3.2.5.2** omitting the SDS-PAGE and methanol transfer was used to probe the samples using the penta-His antibody and HRP conjugated secondary.

### 5.2.6.3 Western blotting with rabbit polyclonal 6XHis antibody and anti-rabbit FITC secondary Western blotting

Transfer to nitrocellulose was performed as described by (Towbin *et al.,* 1979). Following SDS-PAGE gels were rinsed in transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol, pH 8.3). A nitrocellulose membrane (0.45 µm pore size, GE Healthcare, Amersham) was cut to the same dimension as the gel and soaked in dH2O for 10 minutes. A Bio-rad mini-gel blotting apparatus was set up according to the manufacturer's instructions, and the proteins were transferred for 1 hour at 100V in transfer buffer. Membranes were washed Tris-buffered saline (TBS, 20 mM Tris, 500 mM NaCl, pH 7.5) for 10 minutes prior to blocking with 1% alkali-soluble casein (Merck) in TBS containing 0.2% Tween-20 (TBST), for 2 hours at room temperature. A 6XHis primary antibody (1 µg/ml) (Abcam) in antibody buffer (TBST containing 1% alkali-soluble casein) was added for 2 hours at room temperature. Membranes were washed three times in TBST. Secondary antibody, a FITC conjugate of goat anti-rabbit IgG (Abcam) was added at a dilution of 1:10,000 in antibody buffer for 1 hour at room temperature. Membranes were washed three times in TBST and the fluorescent signal was detected using the Odyssey Infrared Imaging System (LI-COR Biosciences, UK).

## 5.2.7 Surface plasmon resonance

### 5.2.7.1 Docking sensor chips and maintenance

Before analysis using Biacore apparatus, all reagents and sensor chips must be equilibrated to room temperature. All buffers were filtered (0.22 µm) using a Millipore

filter and degassed under vacuum filtration. The Biacore micro-fluidic system was primed with running buffer after each docking of a sensor chip. There are variations in the reflectance characteristics of the gold dextran surface both between chips and between the same chip at different temperatures. To normalise the signal response between docking sensor chips, the normalise maintenance procedure was routinely carried out using 100 µl 70% BIAnormalising solution (70% glycerol), (Biacore). All assays were carried out at room temperature (22 ±1 °C).

### 5.2.7.2 Sample loading

Sample loading is a critical step in producing unambiguous sensorgrams, as a semi-automated method the sample injection is by manual load and a 'double air bubble' technique, displayed in **Figure 5.9** is necessary to prevent sample dispersal at the beginning and after injection, in particular it is useful for viscous samples that tend to drag in the sample loop walls during injection.

### 5.2.7.3 Amine coupling of penta-His antibody to CM5 sensor chip

Firstly, prior to preparing the sensor chip surface for amine coupling, the intended ligand (penta-His Ab) for immobilisation was checked for sufficient electrostatic interactions with the carboxymethylated dextran in just one flow cell. This pH scouting is normally carried out at a range of pH in sodium acetate buffers. An aliquot of the penta-His antibody was subjected to a short-term dialysis for 1 hour at 4 °C into 10 mM sodium acetate buffer pH 5.0 where it has previously been reported to associate and couple to the dextran surface (Nordin *et al.,* 2005). Upon dissociation of the penta-His using the running buffer (10mM Hepes, 0.15 M NaCl, 0.0005% P20, pH 7.4) the dextran surface was activated using the Biacore amine coupling kit (GE Healthcare) with a freshly made 1:1 mixture of 0.1 M N-Hydroxysuccinimide (NHS) and 0.4 M 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) at 30 µl/min to reach 180-250 RU, to give reactive succinide esters. 30µg/ml of penta-His antibody diluted in 10 mM sodium acetate (pH 5.0) was injected,at 20 µl/min and reapplied until the response units were resolute. The remaining activated groups were blocked with 60 µl 1 M Ethanolamine hydrochloride-NaOH pH 8.5, to quench unreacted esters and remove unbound antibody. The immobilised sensor chip was ready for binding His-tagged proteins. Activation and deactivation was also carried out on the sensor chip surface in the second flow cell to produce a reference surface for subtraction of non-specific protein interactions with the dextran surface.

### 5.2.7.4 Capture of analyte using a penta-His immobilised sensor surface

The unpurified soluble fraction of the cell-free expressed recombinant His-tagged truncated Pmps or purified Endo1-8-His control were injected over the active and reference surfaces for capture by the immobilised penta-His antibody on the CM5 sensor chip. These samples were either directly applied as concentrated proteins or were diluted in running buffer (10mM Hepes, 0.15 M NaCl, 0.0005% P20, pH 7.4) to minimise the changes in refractive indices caused by the change in sample buffer and running buffer. The changes in response units were measured and the reference surface was subtracted from the active surface (BIAevaluation software version 3.2).

### 5.2.7.5 Regeneration of penta-His immobilised surfaces

Regeneration buffer (Glycine-HCl pH 2.5) was injected at 7 $\mu$l/min for 10 minutes. The baseline was monitored after regeneration, where the baseline was not restored, further efforts to regenerate were carried out with harsher regeneration buffers (0.05% SDS in appropriate running buffer).

### 5.2.7.6 NTA sensor chip activation

Based on the optimised activation method by (Nieba *et al.,* 1997) 500 mM $NiCl_2$ was injected at a flow rate of 7 $\mu$l/min for 1 - 4 minutes. $NiCl_2$ was made up using the relevant running buffer used for the experimental analyses although activation has been shown to be stable against variations in pH, NaCl concentration and contact time (Nieba *et al.,* 1997).

### 5.2.7.7 Binding of ligand to NTA sensor chip surface

The protein samples as in **Table 5.3** were injected over the active and reference surfaces for binding of the His-tag to the immobilised Ni-NTA surface. As a His-tagged control to check NiNTA binding specificity, Endo1-8-His was used, and a non His-tagged control, 1 mg/ml BSA was employed to test for non-specific interactions. These samples were either directly applied as concentrated proteins from direct cell-free expression reactions or were diluted in running buffer 1 (10mM Hepes, 0.15 M NaCl, 0.0005% P20, pH 7.4) to minimise the changes in refractive indices caused by the differences in sample buffer and running buffer. Buffer 1 was later replaced by running buffer 2 (PBS, 300 mM NaCl, Halt EDTA-free protease inhibitor cocktail, pH 8.0) to minimise electrostatic interactions. In addition the use of detergents was avoided to prevent the disruption of the nanodiscs in the expressed rPmp samples. Changes in response units were measured and the reference surface was subtracted from the active surface (BIAevaluation software version 3.2).

224

## 5.2.7.8 Regenration of NTA sensor surface

Imidazole, normally used to elute proteins during chromatography has been shown to disassociate Histidine-tagged proteins at 50 mM (Nieba *et al.,* 1997) whilst not removing the nickel ions from activated sensor chips. Regeneration of NTA surfaces carried out with imidazole was unsuitable for the removal of Histidine-tagged proteins in this study. This step was replaced with 350 mM EDTA, pH 8.3 at a flow rate of 7 µl/min for at least 3 minutes. EDTA also removes the nickel ions from the NTA surface and so reactivation of surfaces was necessary following each regeneration step.
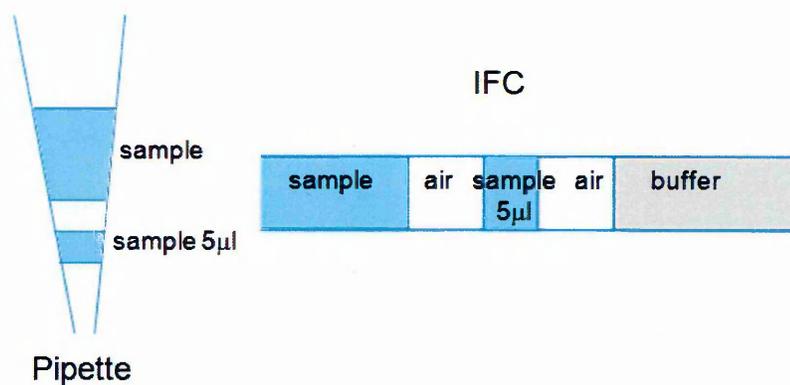
**Figure 5.9 - Air bubble technique for loading sample into the integrated micro-fluidic cartridge.** Shows the sample pipette technique (left) and the sample loop (right). Adapted from Biacore X Getting Started Handbook, 2002

| Ligand | Total protein concentration |
|---|---|
| **rPmpl-N** (C-terminal 6 His-tag) | 30 µg/ml |
| **Expressed empty pET23b control** (no His-tagged proteins) | 30 µg/ml |
| **Endo1-8-His** (C-terminal 8 His-tag) | 2 µg/ml, 20 µg/ml, 200 µg/ml |
| **BSA** (no His-tag) | 1 mg/ml |

**Table 5.3 - Ligands used to test specificity of binding to activated and non-activated NTA sensor surfaces.**

## 5.2.8 Ellipsometry

## 5.2.9 Immobilization of HeLa cell 'orphan receptors' on solid surfaces

The substrates for TIRE study were prepared by thermal evaporation of gold onto microscopic glass slides (kindly provided by Dr Anna Tsargorodskaya, MERI, Sheffield Hallam University) The typical thickness of gold layer was 25–27 nm. The gold-coated glass slides had an enhanced negative surface charge from prior treatment overnight in 0.1 M solution of mercapto-ethylsulfonate sodium salt SH-$(CH_2)_2$-$SO_3^-$ $Na^+$ in methanol. The immobilization of proteins on the gold surface was achieved by positively charging the surface with polyallylamine hydrochloride (PAH, 2 mg/ml). An incubation time of 15 min was sufficient to achieve saturation of PAH adsorption. The sample was then rinsed by purging the cell with 10-times the cell volumes of deionised water. After that, the surface is ready for electrostatic adsorption of proteins, which are typically negatively charged in slightly alkaline (pH 8.0) Tris-HCl buffer solution (Nabok *et al.*, 2005).

## 5.2.10 TIRE experiments

The TIRE experimental set-up in built on the basis of an automatic spectroscopic J. A. Woollam ellipsometer M2000 operating in the 370 – 1000 nm spectral range and exploiting the rotating compensator principle was used for ellipsometry experiments (Nabok and Tsargorodskaya, 2008), (Nabok *et al.*, 2011) and (Kriechbaumer *et al.*, 2011).. The light is coupled to a thin gold film deposited on the glass slide via a 68° trapezoidal prism at the conditions close to total internal reflection (**Figure 5.10**). A 0.2 ml cell is attached to the gold film via rubber O-ring; the inlet and outlet tubes enable injection of required solutions into the cell to perform molecular adsorption and different biochemical reactions on the gold surface. Briefly, the adsorption steps used were 1), PAH (polyallylamine, 2 mg/ml aqueous solution), 2), crude HeLa cell proteins and membranes (in 1X PBS with protease inhibitors, pH 8.0), and 3), 100 µl unpurified soluble rPmp domains (rPmp domains with nanodiscs and endogenous proteins at approximately 300 µg/ml in 1X PBS with protease inhibitors, pH 8.0) with 10-15 min incubation times. Rinsing the cell was carried out after every deposition step using 10 times the cell volume with 100 mM Tris-HCl, pH 8.0. The parameters of the adsorbed layers such as thickness and refractive index were evaluated by fitting the TIRE spectra to the model system by Dr Anna Tsargorodskaya, MERI, Sheffield Hallam University) using the J.A. Woollam software (Nabok *et al.*, 2005).
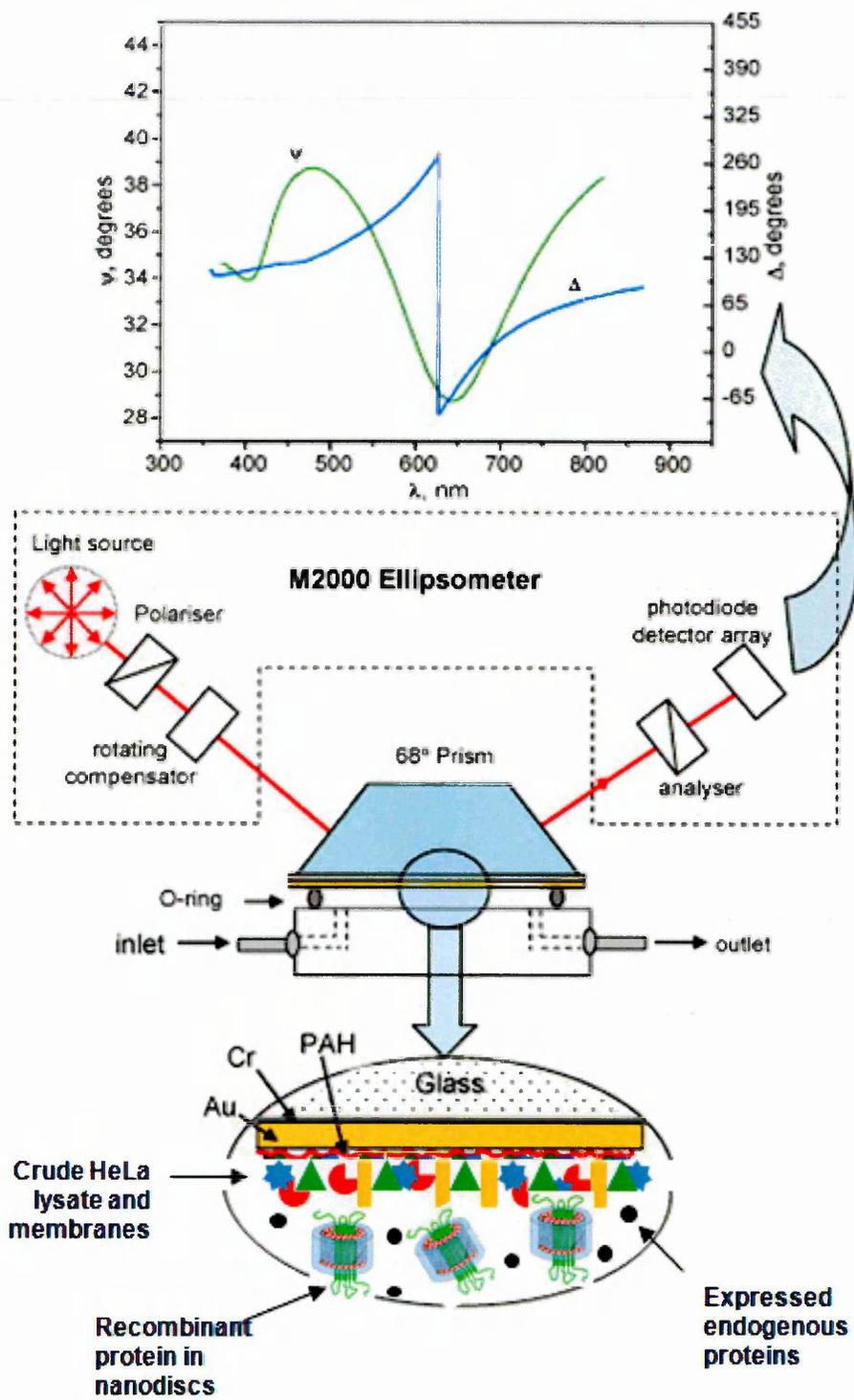
**Figure 5.10 - Schematic diagram of TIRE experimental set-up.** The 'orphan' receptor (crude mix of HeLa cell proteins and membranes) are layered onto the PAH/gold surface by electrostatic interactions, the recombinant protein samples (rPmp domains expressed using MembraneMax™ cell-free expression kit) are passed over in solution, the reflected light is monitored as ψ and Δ as described in **section 5.1.4**. Adapted from (Mustafa *et al.,* 2010).

## 5.3 Results

### 5.3.1 Recombinant Pmp domains in cell surface cross-linking binding studies with Hec1B and/or HeLa cells

To screen for potential binding of Pmps to the cell surface of urogenital cells, rPmp domains were incubated with Hec1B and HeLa cells as in **section 5.2.4.1.** Interactions were then fixed permanently with chemical cross-linkers. Western blots using the penta-His primary antibody as in **section 5.2.6.1** were used throughout the cross-linking assays to detect a shift in molecular weight indicative of protein-protein interactions. Trials to investigate putative interactions between truncated Pmp proteins and the human cells were first performed using rPmpI-N since this truncated His-tagged recombinant protein provided the best signal for detection via western blot. Subsequently when a suitable method with controls was established, these trials were repeated to investigate interactions of the autotranslocator domains of Pmp D and Pmp G with Hec1B and HeLa cells.

Formaldehyde is commonly used for cross-linking protein interactions, with one of the shortest space arms available (2.3–2.7 Å) it has been used for a long time in histology and pathology to fix the native state of tissues and cells (Sutherland *et al.,* 2008). Therefore one disadvantage of using formaldehyde is that only closely associated proteins can be cross-linked due to the small chemical structure yet formaldehyde is fast acting at the right concentration and so can fix potentially transient interactions (Klockenbusch and Kast 2010). Formaldehyde penetrates the cell membrane and so fixes both extracellular and intracellular protein interactions. As formaldehyde was readily available at a fraction of the cost of other cross-linkers, preliminary trials utilised this cross-linker. The majority of rPmpI-N was detected as non-interacting monomers showing no putative interactions with the cells (data not shown).

BS3 a non-cleavable, membrane-impermeable, water-soluble cross-linking agent was used to fix potential cell surface interactions. For all assays recombinant truncated Pmp domains were expressed using MembraneMax™ cell-free expression system as fresh stocks for each cross-linking assay. The rPmp domains were not purified. To control against binding of other MembraneMax™ endogenous proteins present within the samples the total soluble protein from the expression of the empty pET vectors were used in chemical cross-linking assays. As the control expression reaction harbours the expression plasmid without the rPmp domain encoding sequence, the two samples differed only by the expression of the recombinant rPmp domain.

Early trials to investigate the interaction of rPmpI-N to urogenital epithelial cells were carried out at 37 °C to mimic the *in vivo* temperature in which *Chlamydia trachomatis* establishes infection (Su *et al.,* 1996). An aliquot of the whole cell assay (containing cells, recombinant expression proteins and cross-linker) and the supernatant containing rPmpI-N not bound to cells and mammalian cell secretory proteins were examined.

Western blot in **Figure 5.11** revealed a large smear of over 250 kDa, detected in the whole assay fraction containing cross-linked rPmpI-N and Hec1B cells indicating an interaction had occurred (blue arrow). The large cross-linked conjugate remained at the interface of the resolving and stacking gel with monomers detected in the supernatant unbound fraction (green arrow). From this assay alone it could not be determined if the conjugate was as a result of; 1) rPmpI-N fixed to the Hec1B cells or, 2) multimers of rPmpI-N cross-linked to nanodiscs or, 3) rPmpI-N cross-linked as aggregates of itself or indeed, 4) a mix of all these outcomes. Further controls were used to help rule out endogenous Hec1B proteins and help determine where the conjugated rPmp domains were located. Subsequently assays were set up with rPmpI-N with both Hec1B cells and HeLa cells with further controls:

assay 1) cells + recombinant proteins + cross-linker,

assay 2) cells + protein,

assay 3) recombinant protein + cross-linker,

assay 4) cells + empty expression vector proteins + cross-linker,

These controls were implemented in **Figure 5.12** and **Figure 5.13**. The cells were collected to examine if the cross-linked complex was directly associated with the cells. The supernatant containing unbound rPmpI-N, cell-free expression proteins and HeLa secretory proteins was retained. The described samples were examined to show that rPmpI-N was not detected at the HeLa cell surface (**Figure 5.12b**) only a faint band was observed in the assay 1 pellet. This band was an endogenous HeLa protein. The rPmpI-N protein was detected in the supernatant fractions as both monomers in **Figure 5.12a** (red arrow) and large conjuages greater than 175 kDa in assay 2 that did not contain cross-linker. This was only useful in showing that rPmpI-N was not interacting directly and permanently binding with the HeLa cells in the absence of cross-linker as would be expected. Where permanent interactions had been fixed, rPmpI-N could not be detected as strongly.

A similar outcome was observed for Hec1B cells incubated with rPmpI-N. Monomers and higher mass conjugates containing rPmpI-N were detected in the supernatant

fraction where reactions were not permanently fixed with BS3, (**Figure 5.13a** red arrow) however a faint band in the region of 40 kDa, a size indicative of rPmpl-N, was detected in the Hec1B cell pellet, yet cross-linker had not been applied to this assay (highlighted by the yellow arrow **Figure 5.13b**). However since interactions in this assay were not cross-linked, monomers rather than higher molecular weight conjugates would be expected if rPmpl-N was interacting with cell surface proteins as this interaction would most likely be disrupted by SDS.

Most apparent was that the signal of rPmpl-N in these cross-binding assays was weaker in all the fractions fixed where BS3 cross-linker (assays 1 and 3, assay 4 did not contain rPmpl-N so no His-tag signal was detected) (**Figure 5.12** and **Figure 5.13**) compared to the assays without cross-linker (assays 2) (**Figure 5.12** and **Figure 5.13**) indicating that BS3 was responsible for this effect. In addition traces of His-tagged protein were found in the stacking gel wells of all these assays where proteins had not entered the polyacrylamide gels. The samples were run on an 8% resolving SDS-PAGE without a stacking gel for Western blotting, this did not permit the conjugates to be separated for analysis and conjugates were restricted to the wells of the gel (data not shown).

### 5.3.1.1 Dot blot analysis

It was unclear whether the BS3 cross-linking had affected the accessibility of the His-tag epitope of rPmpl-N conjugates to the penta-His antibody. The samples from **Figure 5.13** were re-examined using dot blot analysis, which confirmed the His-tag epitopes were not affected by cross-linking with BS3. **Figure 5.14** shows the anti-His signal of rPmpl-N was strong in assays with and without cross-linker confirming that the reason for the weaker signals via Western blotting was due to the size of the cross-linked rPmpl-N conjugates. The dot blots in **Figure 5.14** also confirmed that whilst the majority of rPmpl-N was detected in the supernatant, a weak signal was observed in the Hec1B cell pellets incubated with rPmpl-N. Penta-His mAb has previously shown some cross-reactivity with Hec1B cell yet the signal was not observed in the negative control cell pellets with pET23b vector proteins. rPmpl-N does appear to have an interaction with Hec1B cells that suggests binding may occur.
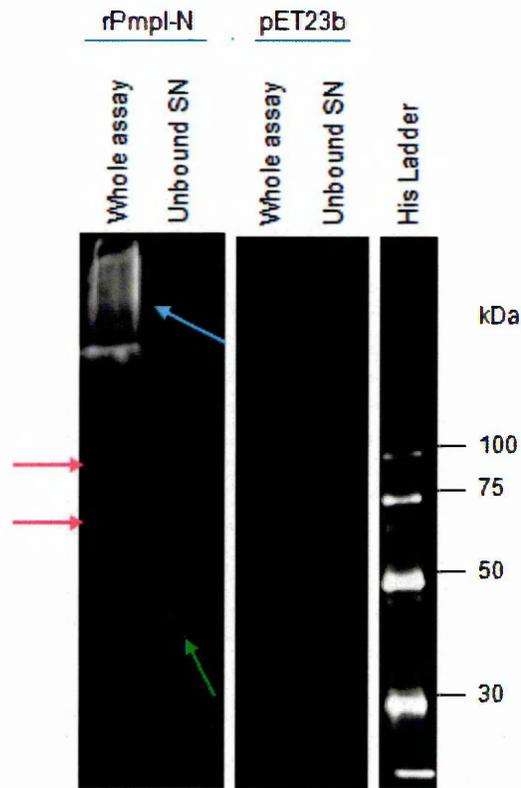
**Figure 5.11 - Western blot showing large smeared conjugate of rPmpI-N indicates interaction of rPmpI-N with Hec1B using cell surface cross-linker BS3**. Hec1B cells were incubated with rPmpI-N (the recombinant passenger domain of Pmp I) and expressed products of empty pET23b for 1 hour before fixing all cell surface interactions with BS3. An aliquot of the whole cell assay includes all the components within the reaction. Following centrifugation the supernatant (SN) was collected and examined for unbound rPmpI-N (unbound fractions) where monomers of rPmpI-N were detected at ~42 kDa (green arrow). Antibody interactions with endogenous Hec1B proteins are shown by pink arrows. His-tagged protein marker (Qiagen) is shown in kDa.
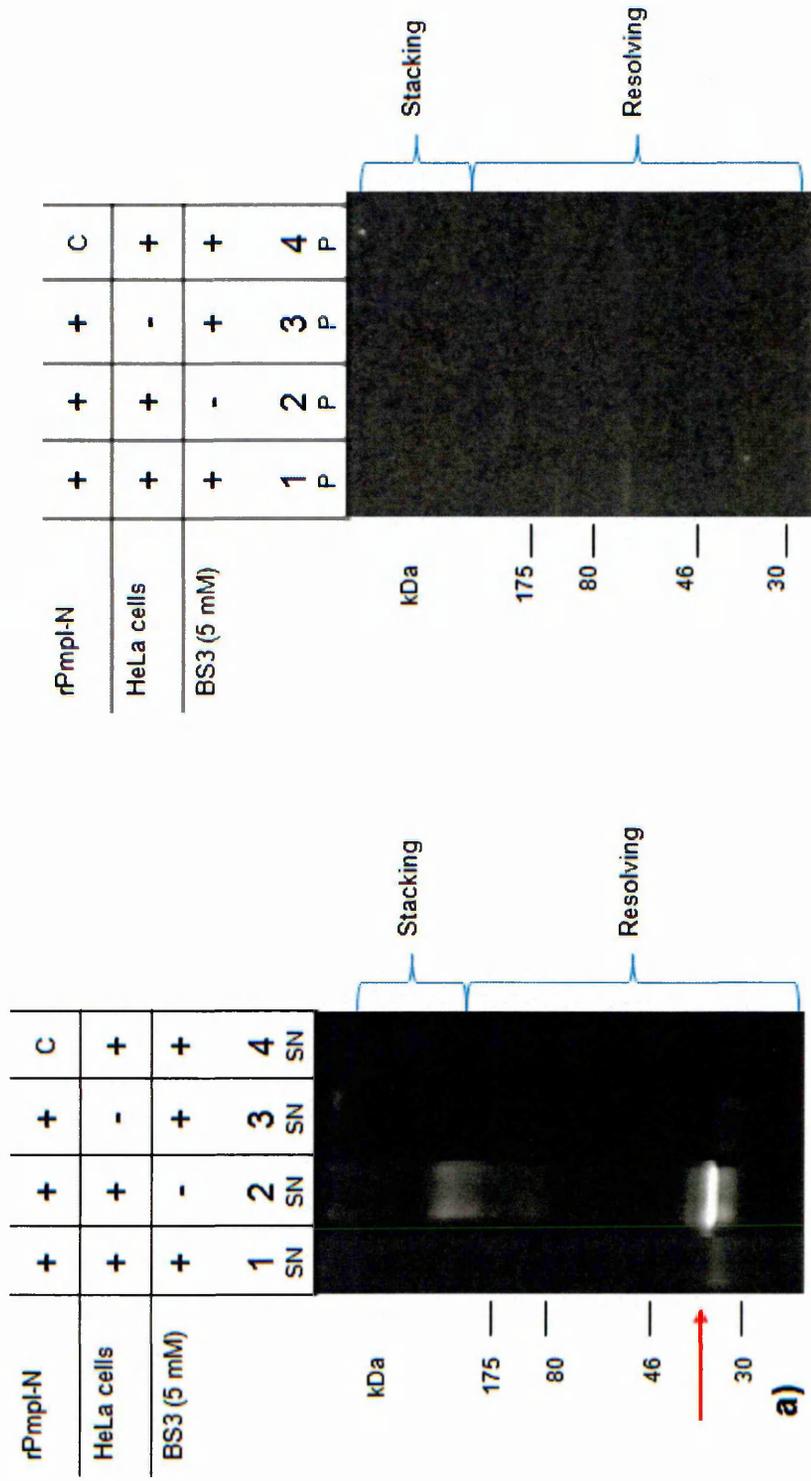
**Figure 5.12 - Western blots of chemical cross-linking: rPmpI-N with HeLa cells using cell surface cross-linker BS3 shows no binding to HeLa cells.** HeLa cells were incubated with (rPmpI-N) or expressed products of empty pET23b (C=control proteins) before fixing all cell surface interactions with BS3. The cells were pelleted (P) and the supernatant (SN) was collected to test for interactions with HeLa secretory proteins (a) An aliquot of the supernatant was examined for each assay 1 to 4. (b) The HeLa cells pellets were examined for cross-linked interacting rPmpI-N at the cell-surface. Assays 1 to 4 were carried out using the components indicated as controls.

**b)**

| rPmpI-N | + | + | + | C |
| --- | --- | --- | --- | --- |
| Hec1B cells | + | + | - | + |
| BS3 (5 mM) | + | - | + | + |
| | 1 P | 2 P | 3 P | 4 P |

kDa — 175 — 80 — 46 — 30

M

Stacking / Resolving

**a)**

| rPmpI-N | + | + | + | C |
| --- | --- | --- | --- | --- |
| Hec1B cells | + | + | - | + |
| BS3 (5 mM) | + | - | + | + |
| | 1 SN | 2 SN | 3 SN | 4 SN |

kDa — 175 — 80 — 46 — 30

Stacking / Resolving
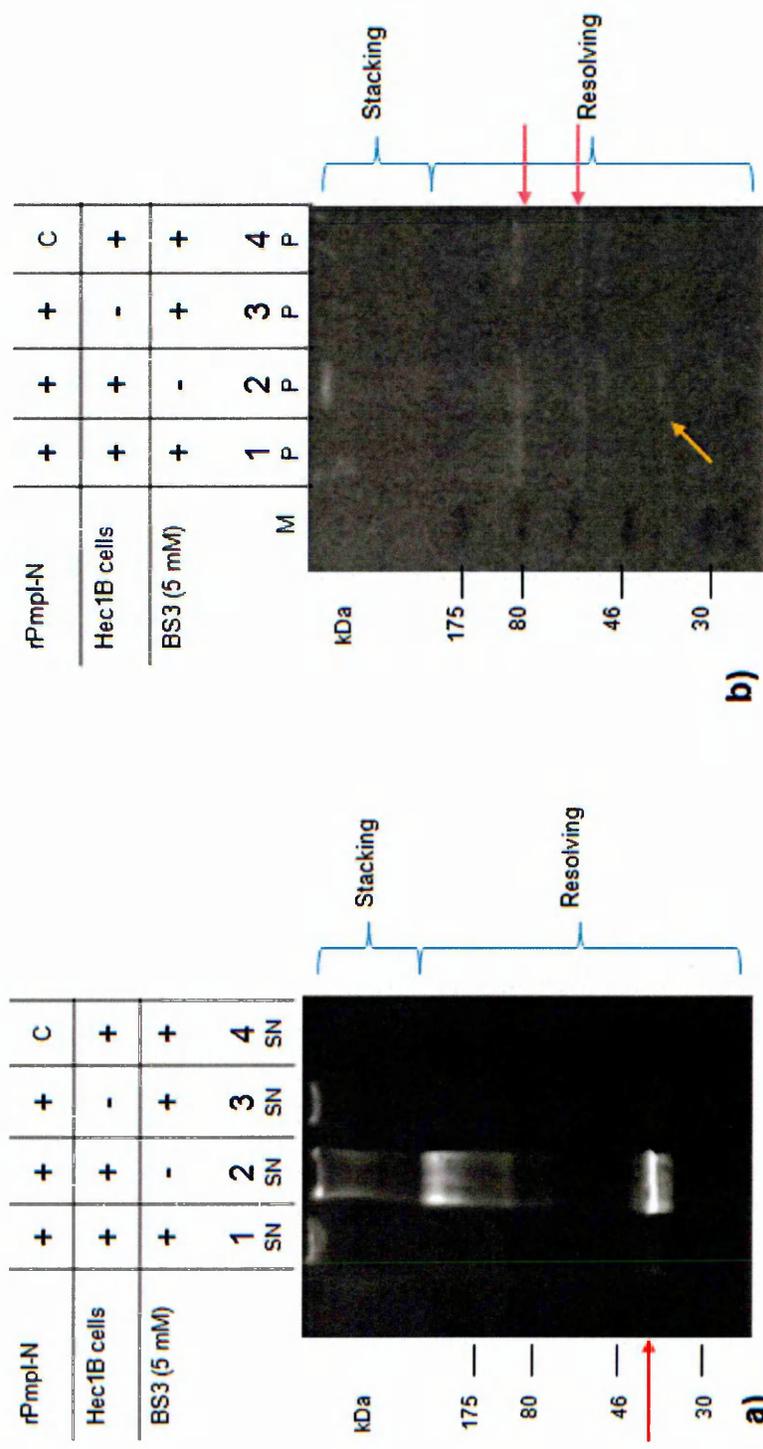
**Figure 5.13 - Chemical cross-linking of rPmpI-N with Hec1B cells using cell surface cross-linker BS3 shows putative interaction with Hec1B cells.** Hec1B cells were incubated with (rPmpI-N) or expressed empty pET23b (C=control proteins) before fixing cell surface interactions with BS3 . The cells were pelleted (P) and the supernatant (SN) was collected to test for interactions with Hec1B secretory proteins (a) An aliquot of the supernatant was examined for each assay 1 to 4. (b) The Hec1B cells pellets were examined for cross-linked interacting rPmpI-N at the cell-surface. Endogenous Hec1B proteins are indicated by pink arrows. Assays 1 to 4 were carried out using the components indicated as controls. M is protein marker in kDa.

234

|  | Assay 1 | Assay 2 | Assay 3 | Assay 4 |
|---|---|---|---|---|
| rPmpI-N | + | + | + | Empty pET |
| Hec1B cells | + | + | - | + |
| BS3 (5 mM) | + | - | + | + |



**Figure 5.14 - Dot blot shows BS3 cross-linker does not affect the epitope of cross-linked rPmpI-N.** Dot blot analysis of the samples shown in **Figure 5.13** was carried out to examine the accessibility of the His-tag epitope after cross-linking with BS3. Assays 1 to 4 were carried out using the components indicated as controls.

## 5.3.1.2 Recombinant Pmp I passenger domain in cell surface cross-linking binding studies at 4 °C

It was difficult to determine whether the large complexes formed during cross-binding studies with rPmpI-N were genuine protein–protein interactions with large putative receptor proteins/molecules or if the rPmpI-N that could not migrate through polyacrylamide gels was as a result of aggregation. As previous studies have shown that adherence of *C.trachomatis* is not inhibited at 4 °C (Zaretzky *et al.,* 1995, Davis and Wyrick 1997) further assays were carried out with a lower incubation temperature at 4 °C to protect the integrity of rPmpI-N that is susceptible to precipitating out of solution into insoluble aggregates, especially at higher temperatures (**chapter 4**).

With Hec1B cells rPmpI-N was detected as monomers in the unbound fraction as expected in assay 2 (without cross-linker) indicated by the blue arrow (**Figure 5.15**) but it was not detected in the other fractions, with the exception of a faint band over 250 kDa in the cross-linked assay 1. The cross-linking assays for HeLa cells resulted in a similar outcome with monomers of rPmpI-N detected in the unbound fractions of the assay 2 (**Figure 5.16**, assay 2, blue arrows). In contrast to the results for Hec1B cells at 4 °C, the large complexes retained in the stacking gel were found in assay 2 without cross-linker and in assay 4 that contained cross-linked rPmpI-N only (no HeLa cells) and a very faint band in assay 1 containing all components. This possibly indicates large cross-linked oligomers or aggregates of rPmpI-N. To conclude, examination of the cell pellets in **Figure 5.15** and **Figure 5.16** do not provide any solid evidence that rPmp-N interacts directly with the cell surfaces of Hec1B or HeLa cells at 4 °C. However, earlier assays at 37 °C observed a faint band associated with the cell pellets, which was confirmed by dot blot analysis therefore if interactions occurring they may have been too weak to detect.

| | Assay 1 | | Assay 2 | | Assay 3 | | Assay 4 | Assay 5 | |
|---|---|---|---|---|---|---|---|---|---|
| rPmpI-N | + | | + | | - | | + | Empty pET | |
| Hec1B cells | + | | + | | + | | - | + | |
| BS3 (5 mM) | + | | - | | + | | + | + | |
| M | SN | P | SN | P | SN | P | SN | SN | P |
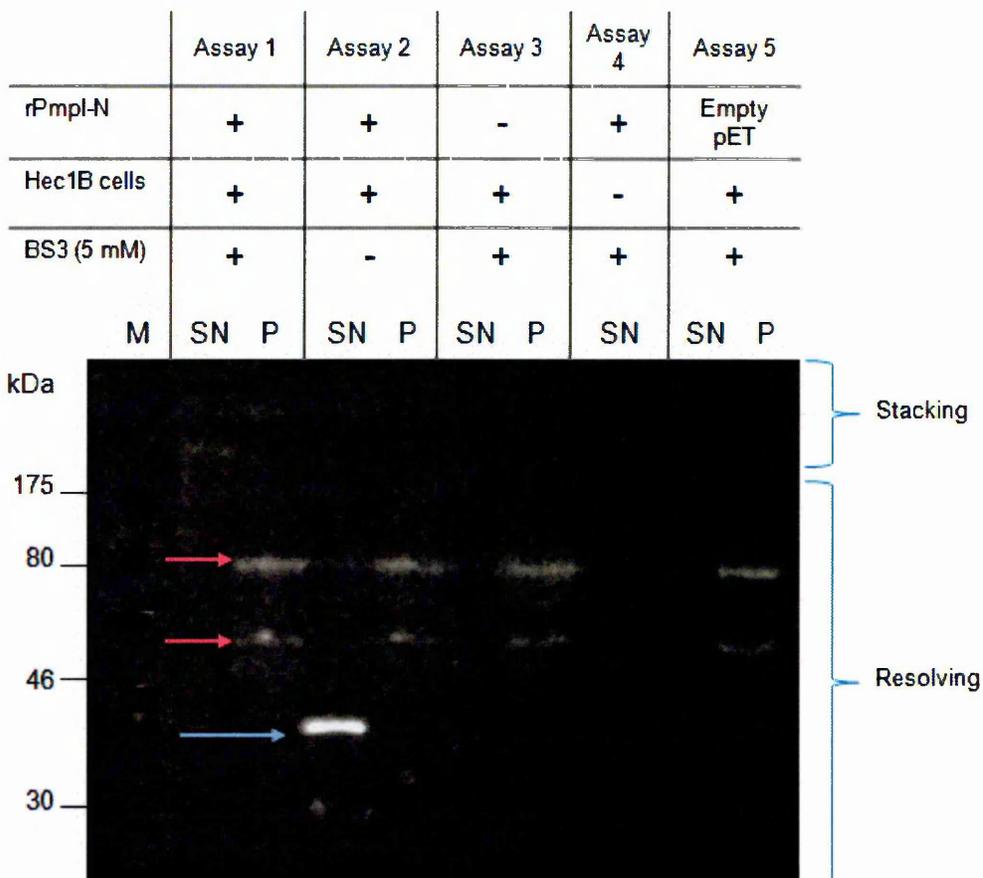


**Figure 5.15 - Western blot showing chemical cross-linking of rPmpl-N with Hec1B cells at 4 °C.** Hec1B cells were incubated with rPmpI-N with controls (incubated without rproteins and with the expressed products of empty pET23b) at 4 °C before fixing cell surface interactions with BS3 . Assay 2 was not cross-linked. The Hec1B cells were collected (P) and the supernatant (SN) containing the MembraneMax™ expression proteins and Hec1B secretory proteins were examined with the exception of control assay 4 that did not contain Hec1B cells. Endogenous Hec1B proteins were identified from control assay 3 that was not treated with heterologous proteins, indicated by pink arrows. Monomers of rPmpl-N are shown by blue arrows. Pre-stained protein marker (New England Biolabs) is shown in kDa.

| | Assay 1 | Assay 2 | Assa 3 | Assay 4 | Assay 5 |
|---|---|---|---|---|---|
| rPmpl-N | + | + | - | + | Empty pET |
| HeLa cells | + | + | + | - | + |
| BS3 (5 mM) | + | - | + | + | + |
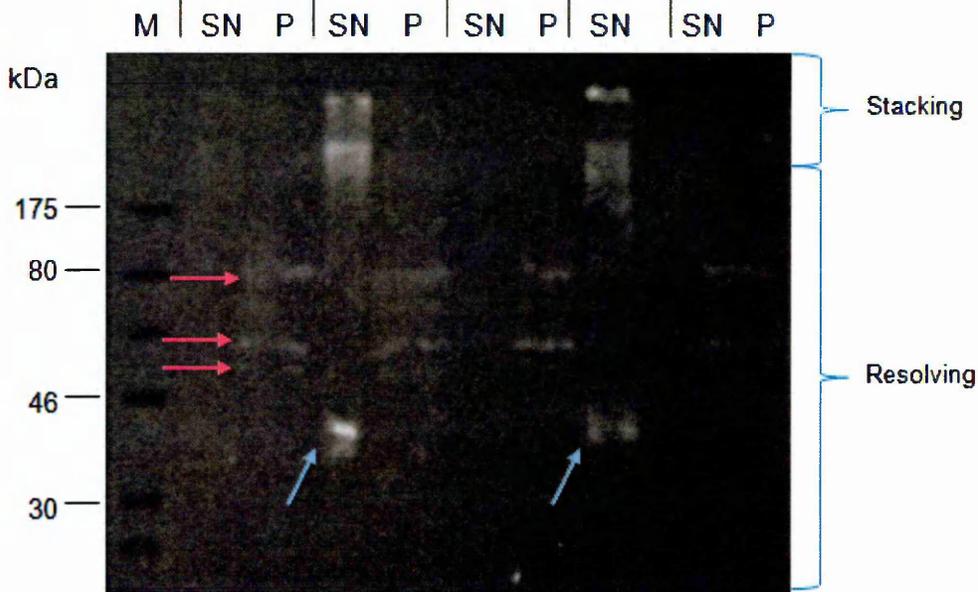
**Figure 5.16 - Western blot shows the chemical cross-linking of rPmpl-N with HeLa cells at 4 °C.** HeLa cells were incubated with rPmpl-N with controls (incubated without proteins and with expressed empty pET23b) at 4 °C before fixing cell surface interactions with BS3. Assay 2 was not cross-linked. The HeLa cells were collected (P) and the supernatant (SN) containing the MembraneMax™ expression proteins and HeLa secretory proteins were examined with the exception of control assay 4 that did not contain HeLa cells. Monomers of rPmpl-N are shown by blue arrows. Endogenous HeLa proteins were identified from control assay 3 that was not treated with heterologous proteins, indicated by pink arrows. Monomers of rPmpl-N are shown by blue arrows. Pre-stained protein marker (M) (New England Biolabs) is shown in kDa.

238

## 5.3.1.3 Recombinant Pmp D and Pmp G autotranslocator domain in cell surface cross-linking binding studies at 4 ºC

Assays were repeated with rPmpGauto and rPmpDauto with Hec1B and HeLa cells incubated at 4 ºC. rPmpGauto was not detected in the cross-linked reactions whilst as expected, monomers were observed in unfixed assays (**Figure 5.17** and **Figure 5.18**). In addition a very faint band of rPmpGauto was detected in the Hec1B cell pellets in assay 4 without cross-linking (**Figure 5.17**) rPmpDauto was below the limits of detection by Western blot during assays after incubation with HeLa cells and so was not repeated with Hec1B cells (data not shown). In view of the data generated from cross-linking assays, alternative methods to investigate these putative interactions were considered.

|  | Assay 1 | Ass-ay 2 | Assay 3 | Assay 4 | Assay 5 |
|---|---|---|---|---|---|
| rPmpGauto | Empty pET | + | - | + | + |
| Hec1B cells | + | - | + | + | + |
| BS3 (5 mM) | + | - | + | - | + |
| | P   SN | SN | P   SN | P   SN | P   SN | M |
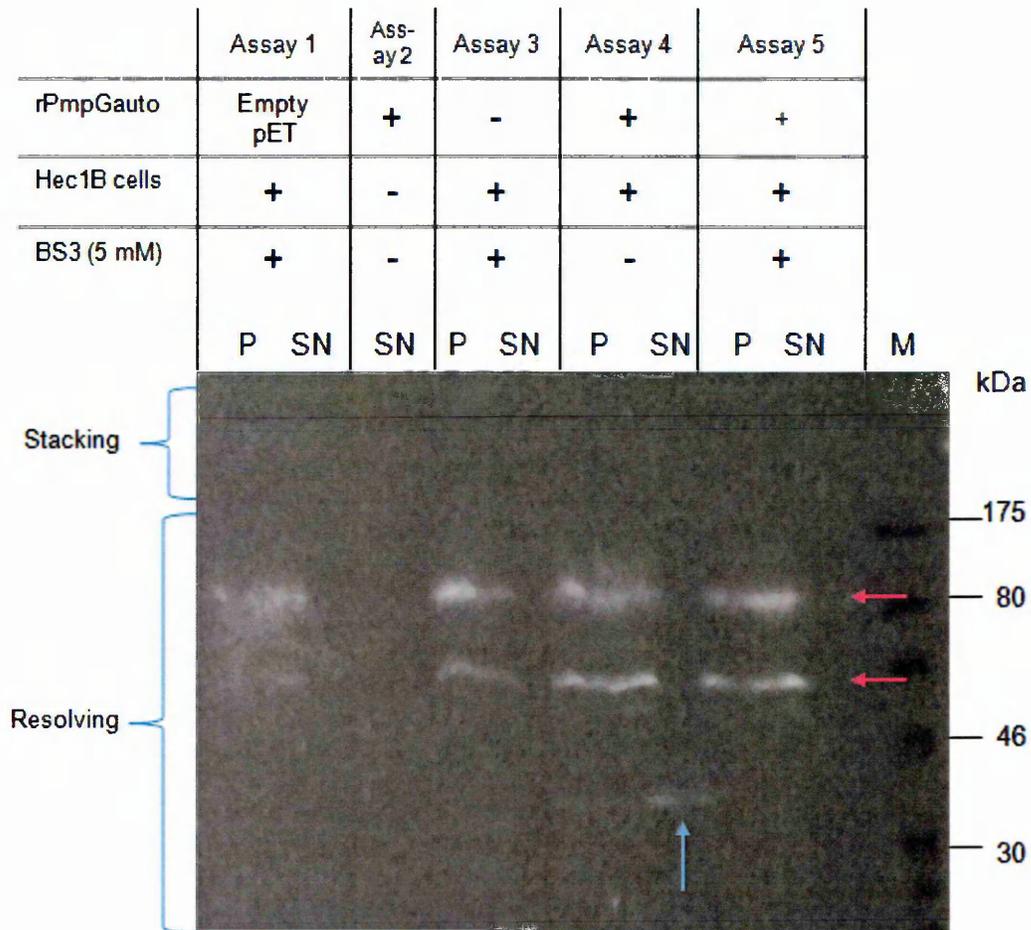
**Figure 5.17 - Western blot showing chemical cross-linking of rPmpGauto with Hec1B cells at 4 °C.** Hec1B cells were incubated with rPmpGauto with controls as above incubated without proteins and with expressed empty pET22b at 4 °C before fixing cell surface interactions with BS3. The Hec1B cells were collected (P) and the supernatant (SN) containing the MembraneMax™ expression proteins and Hec1B secretory proteins were examined with the exception of control assay 2 that did not contain Hec1B cells. Monomers of rPmpGauto are indicated by the blue arrow with a very faint band in the cell pellet of the same assay (yellow arrow). The two bands detected in the cell pellets are endogenous Hec1B cell proteins (pink arrows). Pre-stained protein marker (New England Biolabs) is shown in kDa.

|  | Assay 1 | Assay 2 | Assay 3 | Assay 4 | Assay 5 |
|---|---|---|---|---|---|
| rPmpGauto | + | + | - | + | Empty pET |
| HeLa cells | + | + | + | - | + |
| BS3 (5 mM) | + | - | + | + | + |
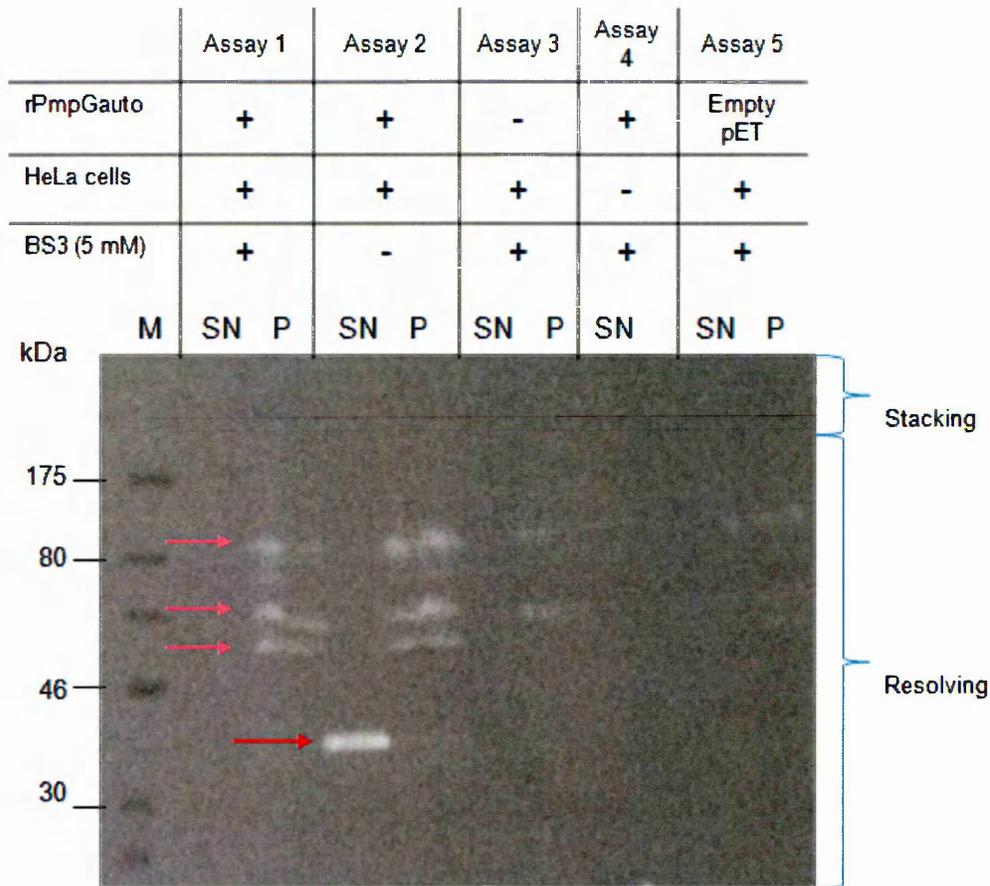
**Figure 5.18 - Western blot of chemical cross-linking of rPmpGauto with HeLa cells at 4 ºC.** HeLa cells were incubated with rPmpGauto with controls incubated without proteins and with expressed empty pET22b at 4 ºC before fixing cell surface interactions with BS3. The HeLa cells were collected (P) and the supernatant (SN) containing the MembraneMax™ expression proteins and HeLa secretory proteins were examined with the exception of control assay 4 that did not contain HeLa cells. rPmpGauto is detected as monomers in the unbound fraction in the unfixed assay 2, indicated by the red arrow. The three bands detected in the cell pellets are endogenous HeLa cell proteins (pink arrows). Pre-stained protein marker (New England Biolabs) is shown in kDa.

## 5.3.2 Methods to investigate recombinant Pmp domains in interactions with Hec1B cell and HeLa cells using immunocytochemistry (ICC)

Methods relying on the cross-binding of truncated rPmp domains with both Hec1B and HeLa cells with chemical cross-linkers were hampered by the dependence upon Western blotting to detect conjugates which were often too large to migrate into an SDS-PAGE gel or were below the limits of detection once diluted in cell suspensions. An alternative method using ICC was proposed to overcome these problems using an indirect probing method. ICC can be a useful tool to probe for the localisation of proteins within cultured cells using antibodies directed against the protein of interest. Determination of a suitable staining method is crucial for successful intensification and identification of protein-cell and protein-protein interactions. Using the penta-His Ab to probe for the localisation of His-tagged rPmp domains upon incubation with mammalian cells was deemed an appropriate staining technique since the penta-His Ab had been shown to identify all the His-tagged rPmp domains. Using an indirect method the signal can be amplified with a suitable fluorescently labelled secondary antibody to reveal whether rPmp domains interact directly with the cells, intracellularly or at the cell surface.

Truncated rPmp domains were expressed using the MembraneMax™ cell-free expression system and were applied to the cells prior to antibody staining as in **section 5.2.5.** As a control to compare the levels of background fluorescence and observe differences in localisation, the same MembraneMax™ protein sample minus the rPmp domain was added to cells by expression of the empty pET vector. During ICC optimisation trials cells were also cross-linked with BS3 prior to staining to develop a suitable ICC method. However as BS3 is a cell surface cross-linker and does not permeabilise the cells, for early ICC trials it was considered necessary to allow the antibodies access within the cell as it was unknown if truncated Pmps would enter the cells. Triton X-100 is the most popular detergent for improving the penetration of the antibody. However, it is not appropriate for the use of membrane-associated antigens since it destroys membranes and therefore was avoided. Fixing of the cells in these studies was therefore carried out with acetone.

### 5.3.2.1 ICC of rPmp domain interactions using penta-His primary antibody and anti-mouse FITC secondary

Numerous attempts of ICC were carried out using the penta-His primary antibody (Qiagen, UK) that had proved useful throughout Western blotting methods in chapter 3 and used for detection of cross-linked recombinant proteins in **section 5.3.1.** The antibody was used in combination with the FITC conjugated secondary antibody as shown in **Table 5.2**. This indirect antibody method was carried out using a range of concentrations but this technique displayed high levels of back-ground signal that could not discriminate between cells that had been treated or untreated with rPmp domains (data not shown). Some cross-reactivity had been observed in the Western blots in **section 5.3.1**, where Hec1B and HeLa protein bands were detected by the penta-His Ab. Titration of the primary antibody to very low concentrations was avoided as low concentrations may not have detected low levels of recombinant His-tagged protein localisation. An alternative primary antibody was chosen.

### 5.3.2.2 Western blot of rPmp domains using 6XHis primary

The specificity of the rabbit polyclonal primary 6XHis tag antibody (Abcam, UK) to His-tagged recombinant proteins showed that this antibody had strong avidity for the rPmp domains with no obvious cross-reactivity with HeLa and Hec1B cells or endogenous MembraneMax™ proteins and so was considered suitable for ICC trials (**Figure 5.19**).

### 5.3.2.3 ICC of rPmpI-N, rPmpDauto and rPmpGauto domain interactions with HeLa cells

A series of dilutions of rPmpI-N (1 in 2, 1 in 5, 1 in 10 and 1 in 20) and a range of antibody concentrations (**Table 5.2**) were carried out to try and optimise ICC methods with the 6XHis primary and anti-rabbit FITC secondary (data not shown). These ICC trials were performed on cervical cells (HeLa) incubated with soluble rPmpI-N. In the preliminary assays the cells were not washed after removal of rPmpN-I in order to confirm the level of recombinant protein added to the cells was in the detection range of the titrated 6XHis-tag primary antibody. In comparison to the untreated cells in **Figure 5.20** (g/h), a 1 in 5 dilution of rPmpI-N was detected with strong staining in the unwashed assay (images a/b, c/d and e/f). Demonstrated within **Figure 5.20** was localisation to some of the cell surfaces (a/b), the cell nuclei of other cells (c/d), with a meshwork of non-interacting rPmpI-N in other areas (e/f). Collectively these showed a diverse range of results as these images were all

captured from the same assay. However the 'meshwork' of recombinant protein observed in (e and f) provided a high background signal that would possibly interfere with the identification of putative protein-cell interactions and so a washing step was included in the experimental assays to remove the excess non-interacting rPmp samples.

Consequently ICC assays with rPmpI-N (1 in 5 dilution) and HeLa cells were repeated with the washing step to remove unbound rPmpI-N after the incubation period in which localisation to the cell surface was not observed (**Figure 5.21**). Cell nuclei were stained brightly (**Figure 5.21a**) and this was comparable to the negative control incubated with the expressed products of empty pET23b (**Figure 5.21b**). This staining was brighter than the staining of untreated cells (**Figure 5.20g**) where moderate cross reactivity/non-specific binding with the HeLa cells was observed.

ICC was also performed on HeLa cells to identify possible cell surface interactions with recombinant Pmp autotranslocator domains of Pmp G and Pmp D. Assays were carried out with a more concentrated application (1 in 2 dilution) of expressed rPmpGauto or rPmpDauto as the expression yield appeared less than the rPmpI-N domain (chapter 3) and the cells were washed to remove unbound protein. Negative controls using the soluble cell-free expression of the empty pET vectors demonstrated a similar staining profile to cells treated with rPmpGauto (**Figure 5.22**) and rPmpDauto (**Figure 5.23**) and so localisation could not be determined.
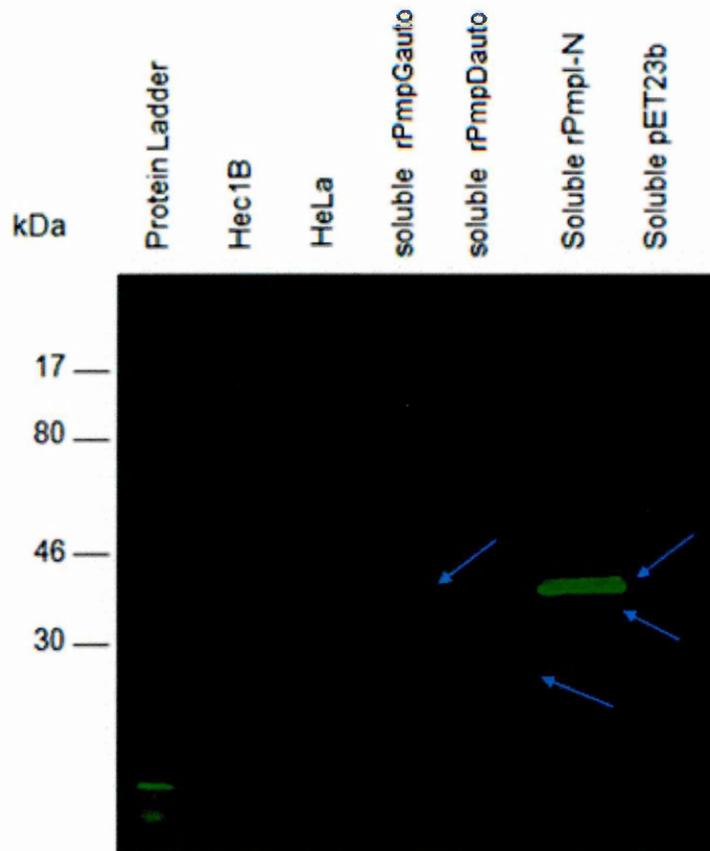
**Figure 5.19 - Rabbit polyclonal 6xHis tag antibody shows avidity for His tagged recombinant Pmp domains.** Western blot carried out using method in **section 5.2.6.3**. Antibody combination shows no cross-reactivity with Hec1B or HeLa cells. Detection of rPmpI-N monomers, (+/- signal peptide) at ~42 kDa, ~36 kDa with high molecular weight multimers, rPmpDauto monomers ~27 kDa and rPmpGauto monomers ~ 40 kDa are shown by blue arrows.
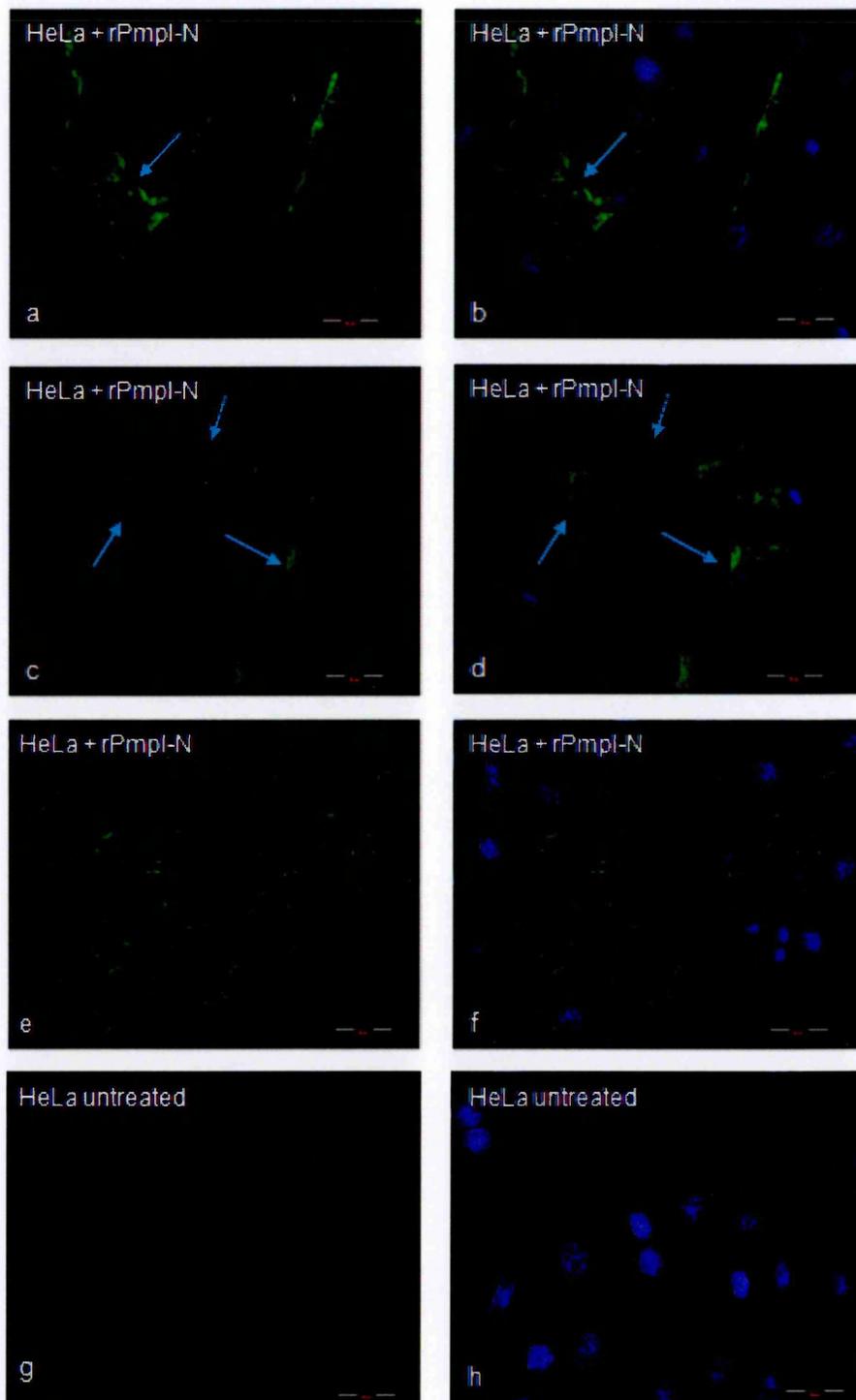
**Figure 5.20 - Immunocytochemistry shows putative rPmpI-N localisation to HeLa cells when the cells are unwashed after incubation with rPmpI-N. FITC (left) and DAPI staining overlay (right).** Images a-f were all captured from the same interaction assay of HeLa cells incubated with rPmpN-I and stained as in **section 5.2.5.3** with 6XHis antibody and ant-rabbit FITC to show the variation in staining and localisation. (a/b and c/d) show varying degrees of rPmpI-N localisation to HeLa cell surface (blue arrows), (e/f) shows no interaction and (g/h) negative cells were treated with PBS only. *Scale bar is 50 μm*

246
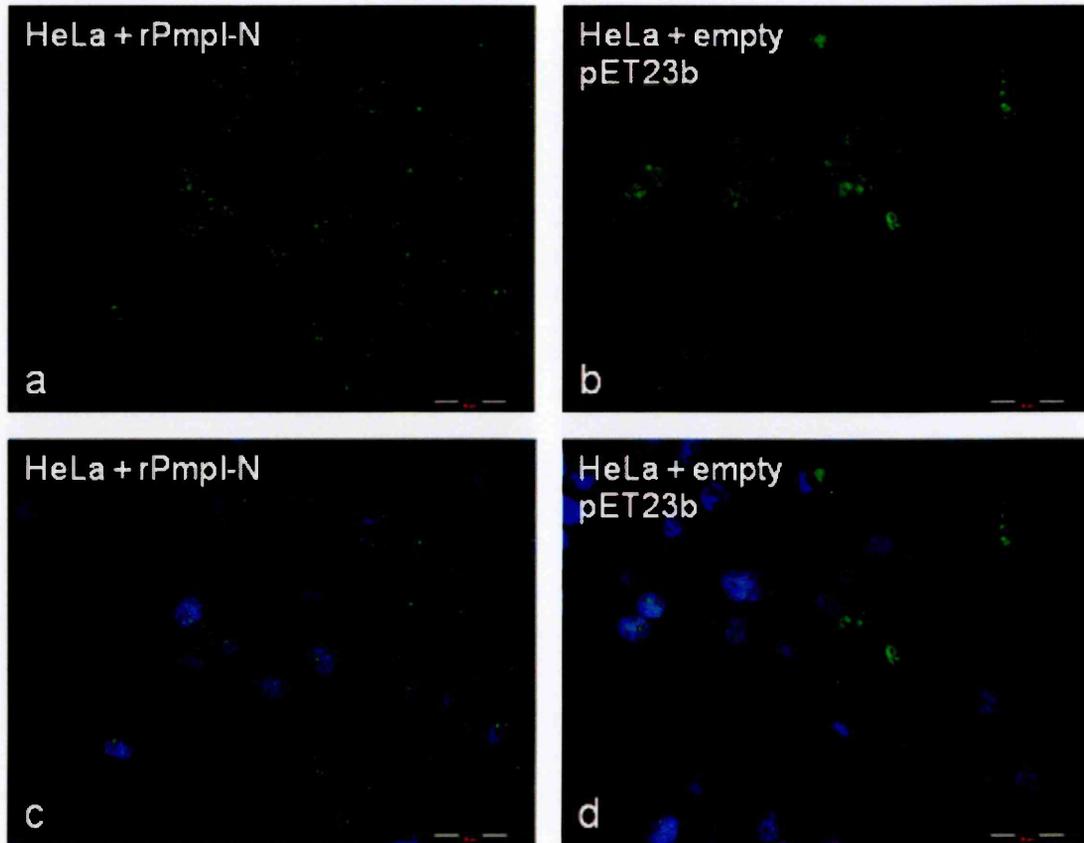
**Figure 5.21 - Immunocytochemistry of HeLa cells post-treatment with rPmpI-N. FITC (top) DAPI staining overlay (bottom)**, (a/c) HeLa cells after treatment with rPmpI-N (b/d) HeLa cells after treatment with pET23b proteins. *Scale bar is 50 μm*
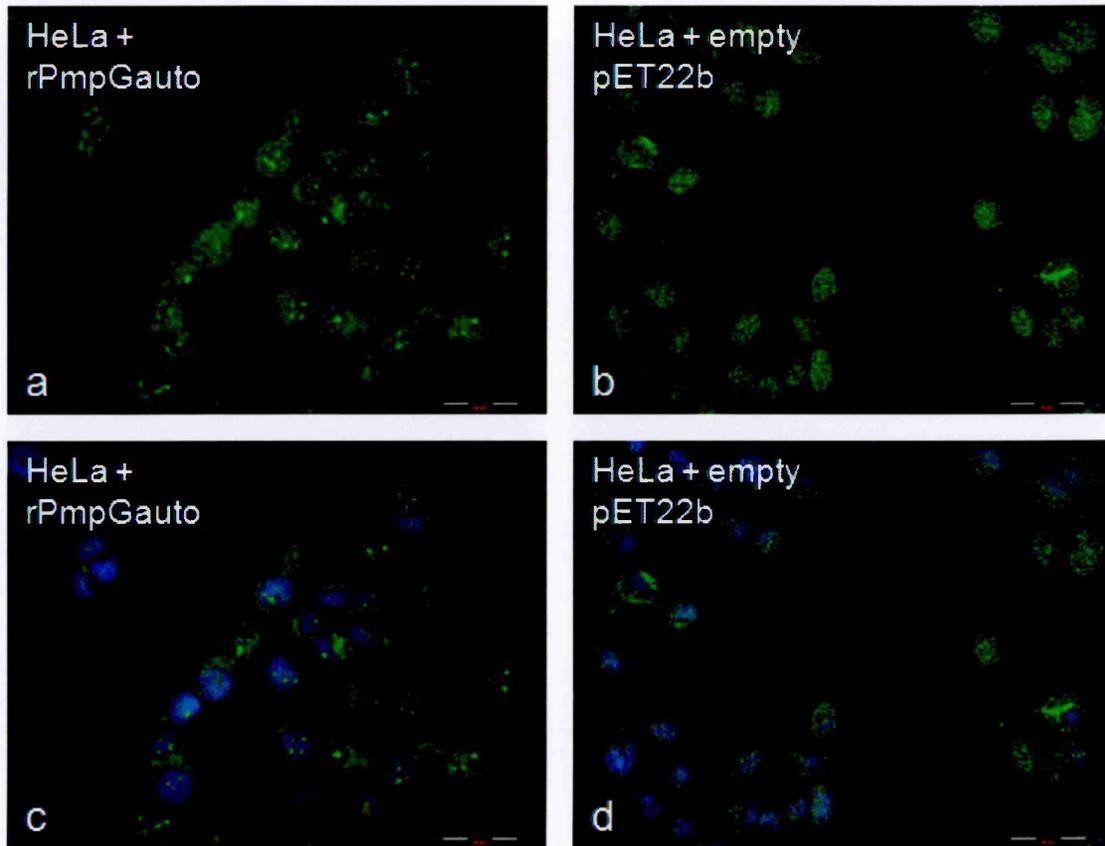
**Figure 5.22 - Immunocytochemistry of HeLa cells post-treatment with rPmpGauto.**
**FITC (top) DAPI staining overlay (bottom)**, (a/c) HeLa cells after treatment with
rPmpGauto, (b/d) HeLa cells after treatment with pET22b proteins. *Scale bar is 50 μm.*
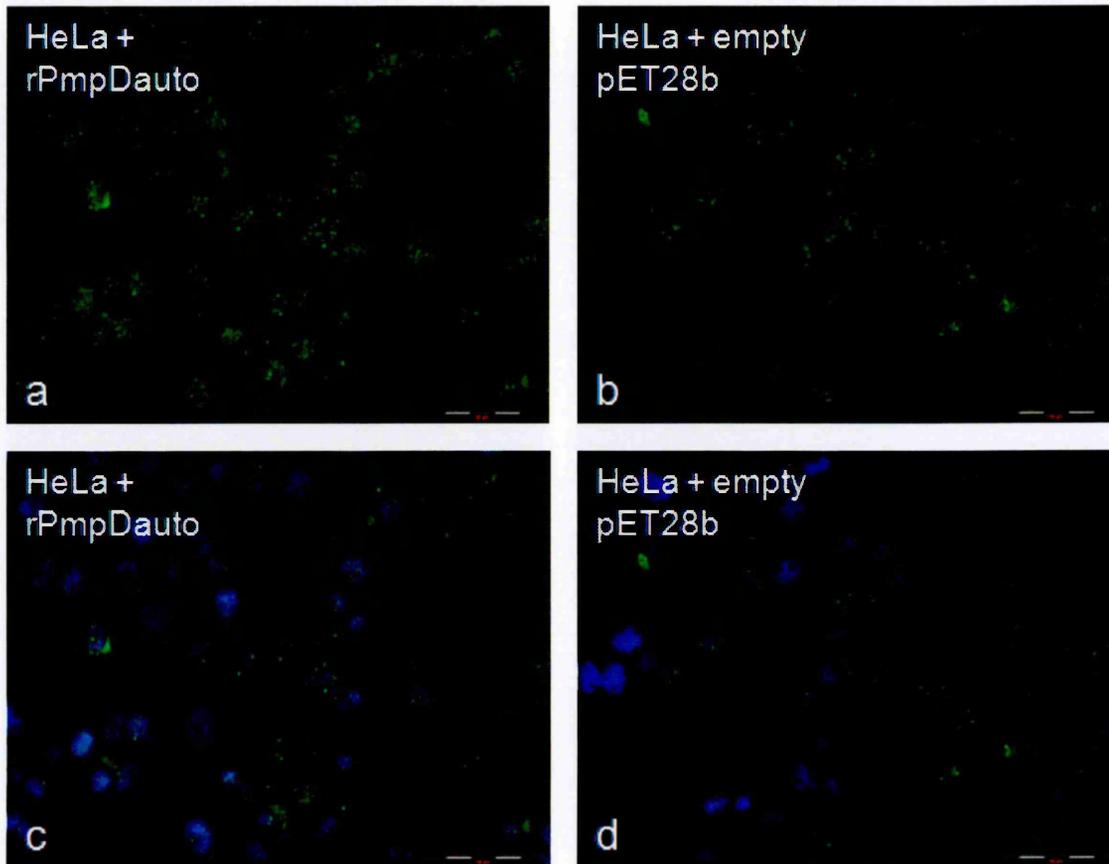
**Figure 5.23 - Immunocytochemistry of HeLa cells post-treatment with rPmpDauto.**
**FITC (top) DAPI staining overlay (bottom)**, (a/c) HeLa cells after treatment with
rPmpDauto, (b/d) HeLa cells after treatment with pET28b proteins. *Scale bar is 50 μm.*

## 5.3.2.4 ICC of rPmpl-N, rPmpDauto and rPmpGauto domain interactions with Hec1B cells

ICC was performed on Hec1B cells to identify possible cell surface or intracellular interactions with recombinant Pmps. Assays were carried out with a 1 in 5 dilution of expressed soluble rPmpl-N and a more concentrated 1 in 2 dilution of expressed rPmpGauto or rPmpDauto applied to the cells (**section 5.2.5.1**). The unbound proteins were washed off prior to fixing the cells with acetone. Negative controls were incubated with the cell-free expressed products of empty pET vectors for comparative analysis.

Differences in staining profiles were observed between cells incubated with truncated rPmps and the cells incubated with the control proteins. In particular rPmpGauto treated cells (**Figure 5.24a**) did not display the granular staining observed in the control cells (**Figure 5.24b**). Cells that had been incubated with rPmpGauto appeared to be stained at the cell surface. Putative cell surface staining was also demonstrated for Hec1B cells after treatment with rPmpDauto shown in **Figure 5.25a** when compared to negative controls shown in **Figure 5.25b**. Finally, differences were also observed in the staining of Hec1B cells after incubation with rPmpl-N where the negative controls had granular staining of the nuclei (**Figure 5.26 a and b**). In this case it was difficult to determine if this staining was localisation to the cell surface or if it was intracellular staining of the nuclear membranes.

In summary, the ICC of Hec1B cells incubated with rPmp domains (rPmpDauto, rPmdGauto and rPmpl-N) all displayed some staining at the cell surface that was indicative of these His-tagged domains interacting with the cells. This signal was stronger and more localised than the signal displayed by the control cells treated with the empty vector expression products.
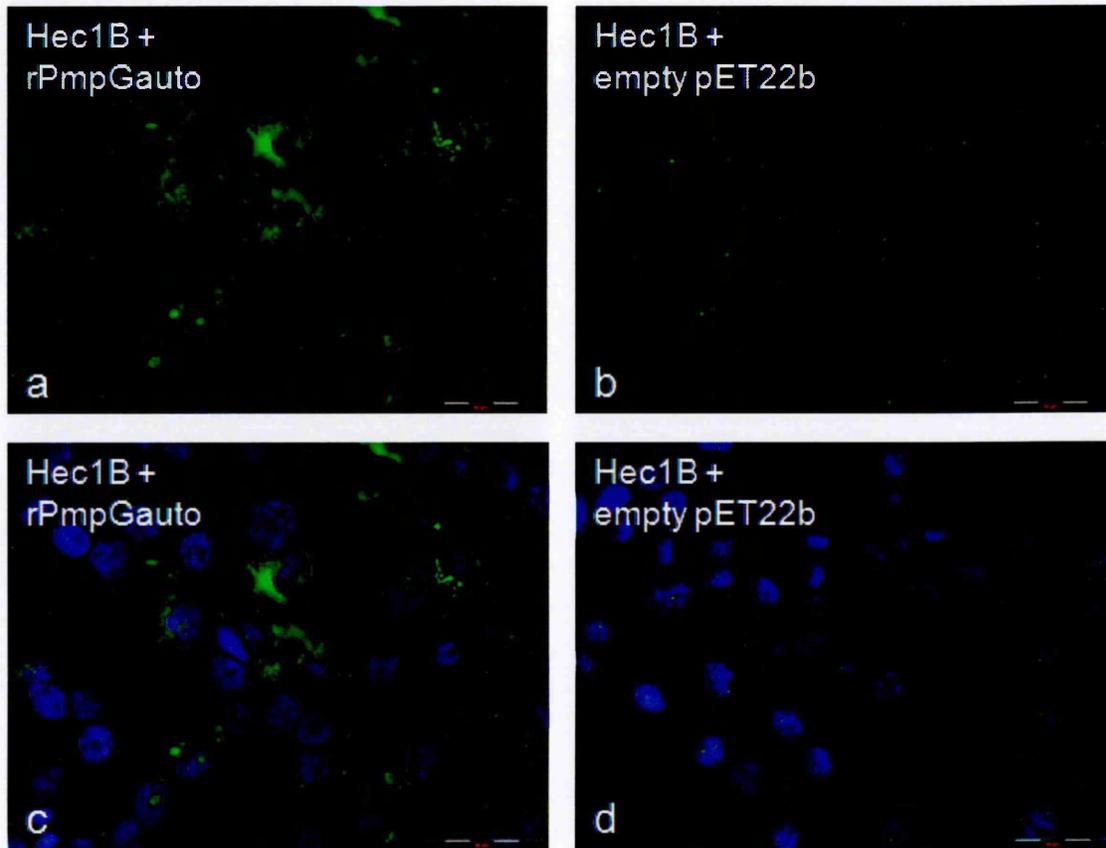
**Figure 5.24 - Immunocytochemistry shows putative rPmpGauto localisation to Hec1B cells. FITC staining (top) DAPI staining overlay (bottom)**, (a/c) Hec1B cells after treatment with rPmpGauto, (b/d) Hec1B cells after treatment with pET22b proteins. *Scale bar is 50 µm.*
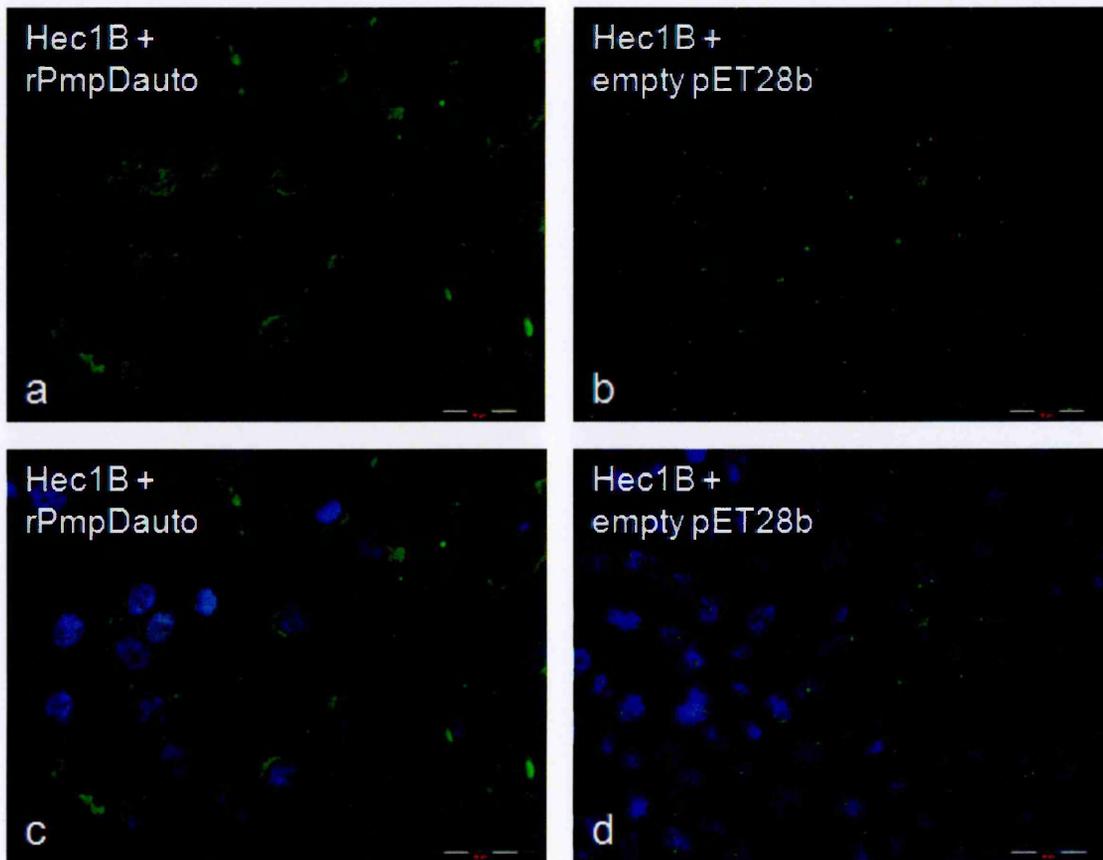
**Figure 5.25 - Immunocytochemistry shows putative rPmpDauto localisation to Hec1B cells. FITC staining (top) DAPI staining overlay (bottom),** (a/c) Hec1B cells after treatment with rPmpDauto (b/d) Hec1B cells after treatment with pET28b proteins. *Scale bar is 50 μm.*
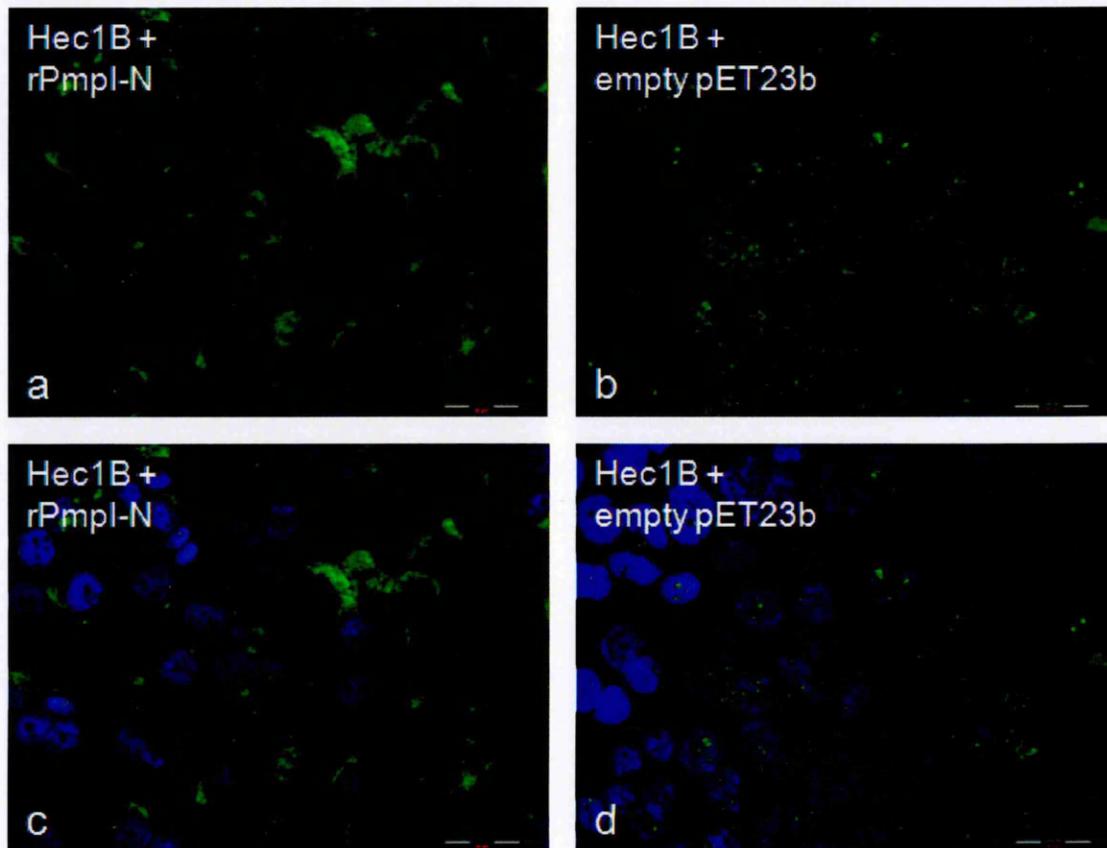
**Figure 5.26 - Immunocytochemistry shows putative rPmpI-N localisation to Hec1B cells. FITC staining (top) DAPI staining overlay (bottom)**, (a/c) Hec1B cells after treatment with rPmpI-N (b/d) Hec1B cells after treatment with pET23b proteins. *Scale bar is 50 μm.*

### 5.3.3 Methods to investigate recombinant Pmp domains in interactions with HeLa and Hec1B cells using surface plasmon resonance

To screen for potential interactions between Pmps with endometrial and/or cervical cells, analysis by surface plasmon resonance was performed with truncated recombinant Pmps expressed in nanodiscs (chapter 4). A variety of biacore sensor surfaces are available, giving a wide range of experimental approaches. For this study the development of a suitable strategy was based on the indirect immobilisation technique. As the previously used penta-His antibody was readily available and showed a high avidity for recombinant His-tagged proteins within MembraneMax™ reactions, this antibody was chosen as the ligand to be immobilised on a CM5 sensor surface. Following successful immobilisation, the antibody was used to capture recombinant His-tagged Pmp domains. Upon capture of rPmp domains, whole or lysed HeLa and Hec1B cells were passed over the captured ligand (rPmp domain) to screen for protein-protein, protein-membrane interactions in a 'ligand fishing' technique. The advantage of using this method was that purified samples were not a requirement, instead a suitable reference surface was used and subtracted from the active surface sensorgram.

### 5.3.3.1 Biacore – CM5 indirect immobilisation technique with penta-His antibody to capture recombinant His-tagged proteins

To eliminate capture of any non-specific proteins present within the samples, empty expression vector was used as a reference to allow for the subtraction of non-specific binding to the penta-His antibody. Given that rPmpI-N produces the strongest signal when probed with penta-His antibody on a Western blot, this truncated Pmp was used in the first screening trials with SPR.

Electrostatic interactions between the carboxymethylated dextran surface and the penta-His antibody were observed using 10 mM sodium acetate buffer at pH 5.0 (**Figure 5.27**). Subsequently the active surface (flow cell 2) was activated and ~2000 RU of penta-His antibody were immobilised to the dextran surface over 5 applications where saturation was reached (**Figure 5.28**).

Initial attempts to capture His-tagged rPmpI-N saw no binding to the active surface or the reference surface. The peaks in **Figure 5.29** show slight association with rapid dissociation back to baseline for both injections of proteins. Subtraction of the reference chip (Fc 2- Fc 1) shows that the immobilised chip did not capture His-tagged rPmpI-N but were refractive bulk indices due to the change in sample buffer

from the running buffer. Subsequently the rPmpI-N sample was diluted with running buffer 1 which lowered the bulk effect and an association curve was seen in the second peaks of Fc 1 and Fc 2 but did not result in any protein binding as the baseline returned to original for both the reference and active surfaces.

A decrease in flow rate from 30 μl/min to 7 μl/min to increase the association time between the antibody and the diluted rPmpI-N protein sample did not result in the capture of proteins. Association curves on the injection peak were not recorded with either sensor surface, only the bulk refractive index is seen (**Figure 5.30**), subtraction of the reference chip (Fc 2- Fc 1) confirmed the penta-His mAb chip had not captured any His-tagged rPmpI-N.

A highly concentrated purified recombinant 8 residue histidine-tagged protein, Endo1 (termed Endo1-8-His in this study for clarity of Histidine tag (supplied by Dr J. Hadden)) was used to test the immobilised penta-His binding capability. Initially, sensorgrams showed binding of Endo1-8-His (diluted in running buffer 1 to 2 μg/ml) with an increase of approximately 850 RU to the active surface, 7 minutes after the injection the protein was still disassociating indicating a strong affinity. However binding was not specific to the Ab coupled surface as approximately 1100 RU of Endo1-8-His also bound to the reference surface, again with a slow dissociation rate. Both surfaces required regeneration to disassociate the Endo1-8-His (**Figure 5.31**). Regeneration with glycine-HCl regeneration buffer pH 2.5 (Biacore Amine Coupling Kit, GE Healthcare) was too mild to regenerate the reference surface, but was too harsh for the active surface where the baseline went into negative RU. The Endo1-8-His was reapplied to test for a reduction in the binding potential on the active surface. Less binding was seen with this application as the RU increased by approximately 600 RU indicating removal or denaturation of some of the penta-His antibody as a result of the regeneration procedure. With this second application, a higher association was observed with the reference surface than the active surface possibly showing Endo1-8-His binds non-specifically to the dextran surface as well as the penta-His. Further applications, with a change in buffer from Hepes to PBS with increased ionic strength (running buffer 2), to reduce electrostatic interactions between Endo1-8-His and the dextran matrix surfaces, led to an unstable baseline with high and low drifts (data not shown). 0.05% SDS a strong regeneration buffer, was also insufficient in removing the Endo1-8-His from the reference surface (**Figure 5.31**). Inevitably, this had led to an accumulation of bulk on the reference surface that would result in ambiguous results and irreversible damage to the active

255

surface. Without a suitable positive control, the ability to determine the ligand activity was compromised therefore trials to capture rPmp domains were not pursued using this approach.
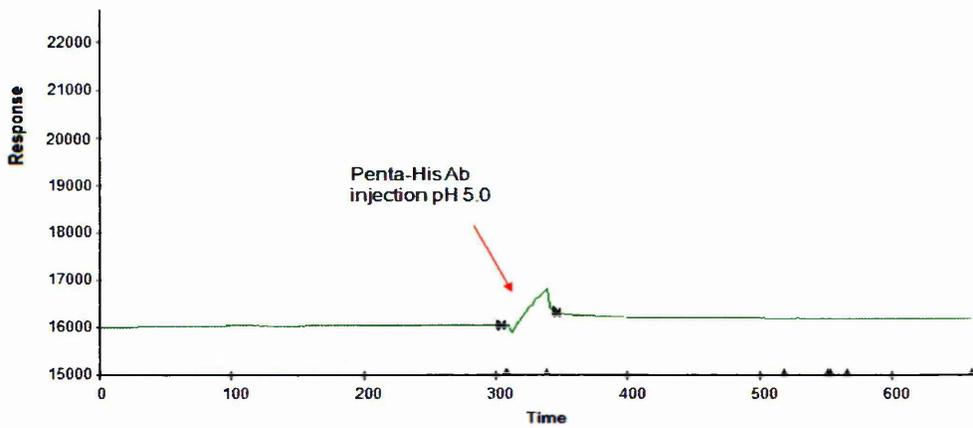
**Figure 5.27 - Sensogram shows the mouse monoclonal Penta-His antibody associates with dextran matrix at pH 5.0.** $30\mu g/ml$ of antibody was injected over flow cell 2 in 10mM sodium acetate to test for association with the unactivated carboxymethylated dextran (CM5) sensor chip. Time is in seconds (s) and response is response units (RU).
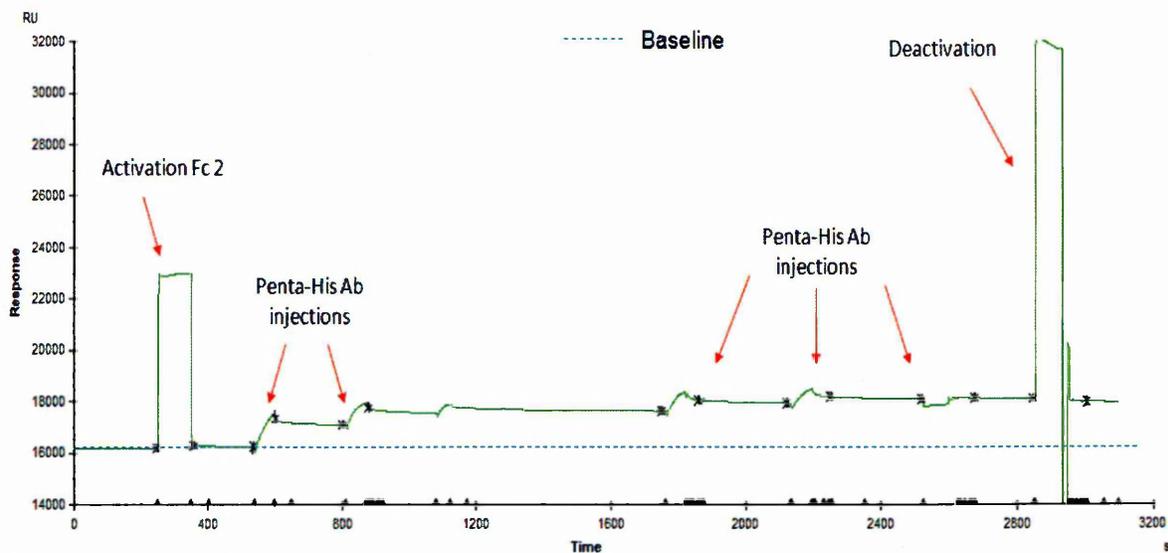


**Figure 5.28 - Immobilisation of the mouse monoclonal penta-His antibody to CM5 sensor chip using amine coupling**. The chip was activated using 1:1 EDC and NHS activation buffer prior to the coupling of $30\mu g/ml$ of antibody over 5 injections (flow cell 2) in 10mM sodium acetate, with a flow rate of $20\mu l/min$. Approximately 2000 RU of penta-His Ab was coupled to the sensor chip surface. The uncoupled reactive groups were quenched with ethanolamine during the deactivation phase.Time is in seconds (s) and response is response units (RU).

257
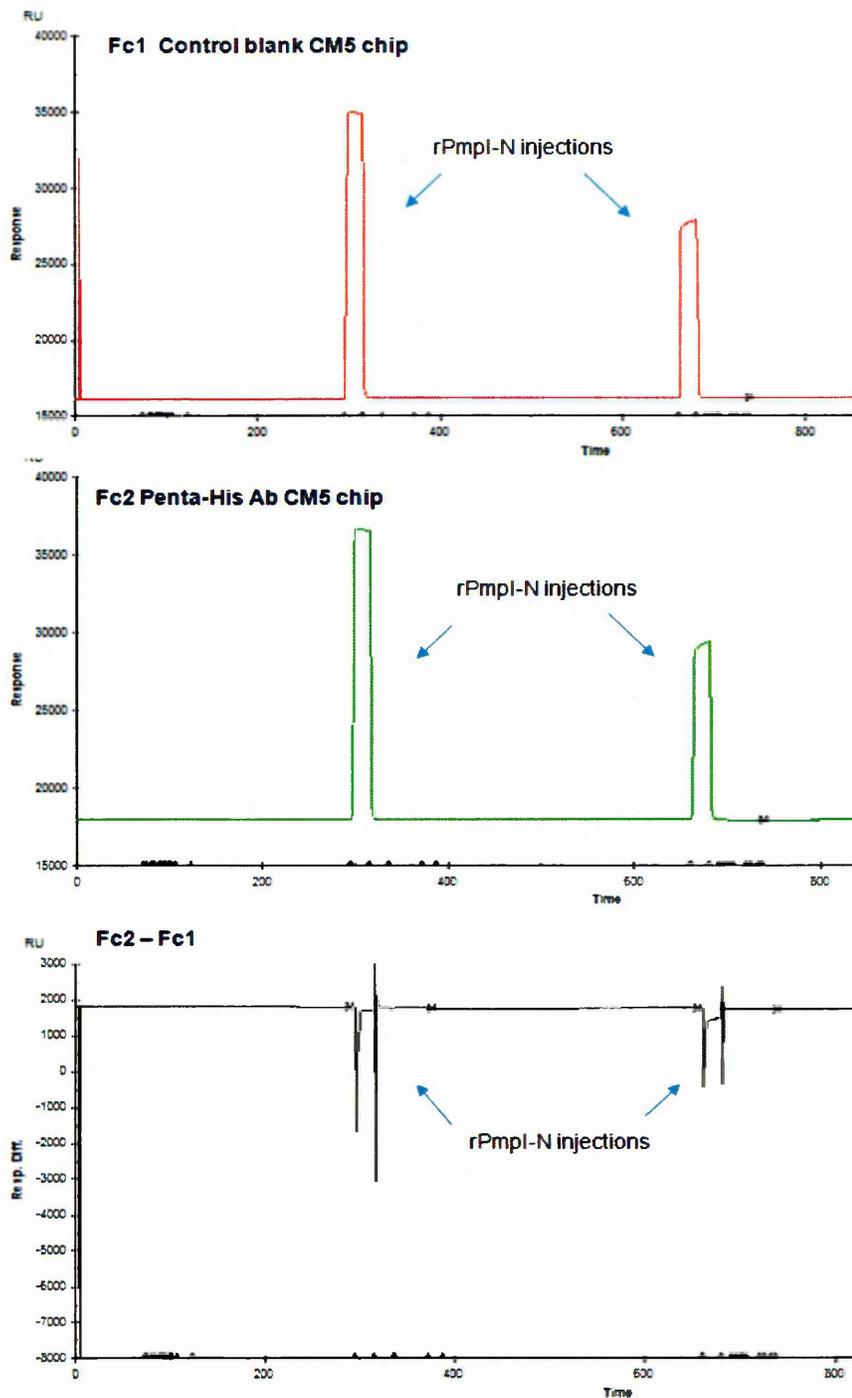
**Figure 5.29 - Capture of concentrated recombinant Pmp I passenger domain on penta-His CM5 chip**. An aliquot of the soluble fraction of the undiluted cell-free expression of rPmpI-N was injected at a flow rate of 30μl/min over both the immobilised penta-His CM5 chip (Fc 2) and the blank reference chip (Fc 1). A second application of rPmpI-N was injected, diluted x2 in running buffer. Sharp spikes in the reference-subtracted sensorgram derive from the small time difference between flow cells 1 and 2. Time is in seconds (s) and response is response units (RU).
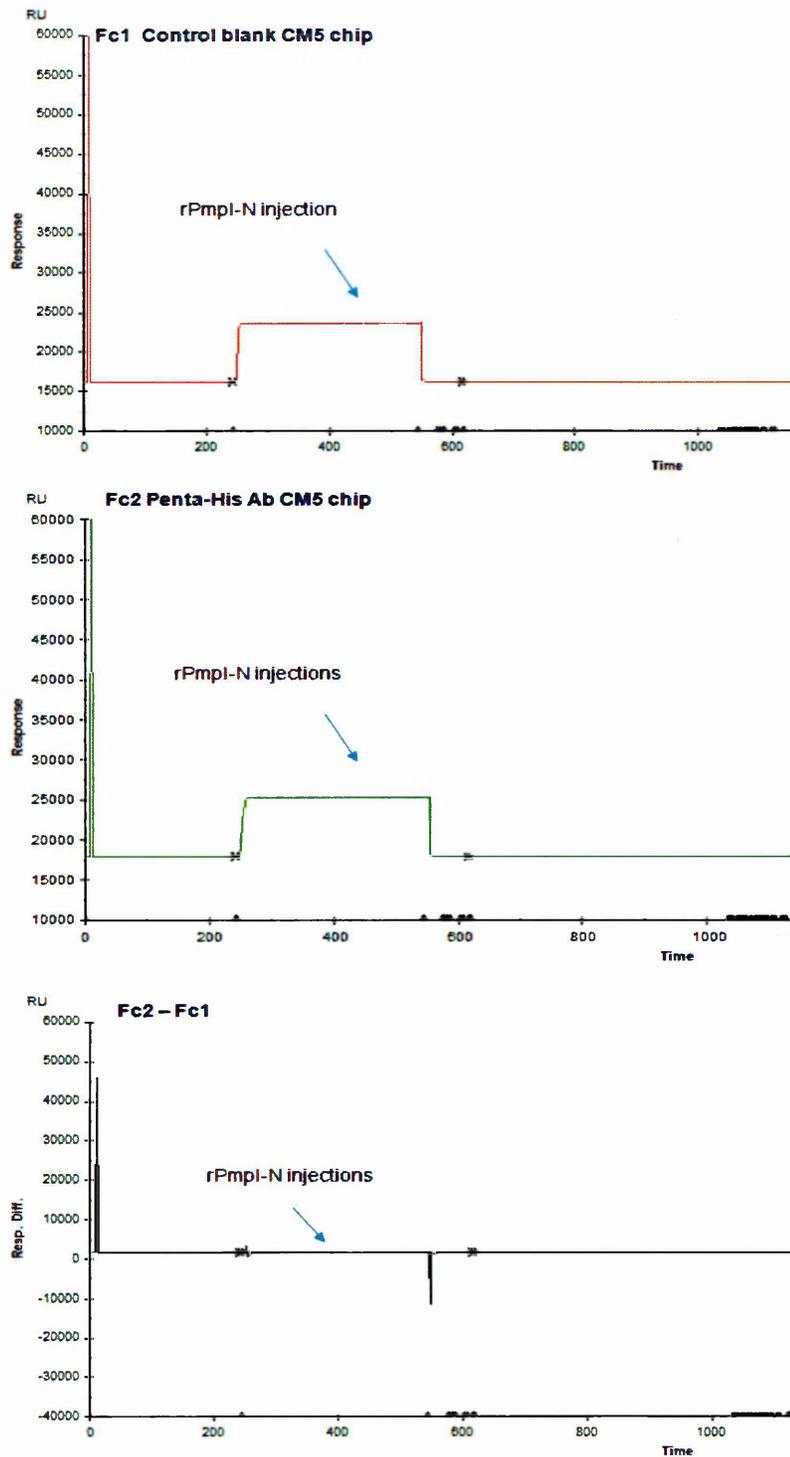
**Figure 5.30 - Capture of diluted recombinant Pmp I passenger domain on penta-His CM5 chip using a slower flow rate to increase capture association**. An aliquot of the soluble fraction of a cell-free expression of rPmpI-N (x2 dilution factor) was injected at a flow rate of 7μl/min over both the immobilised penta-His CM5 chip (Fc 2) and the blank reference chip (Fc 1). Time is in seconds (s) and response is response units (RU).
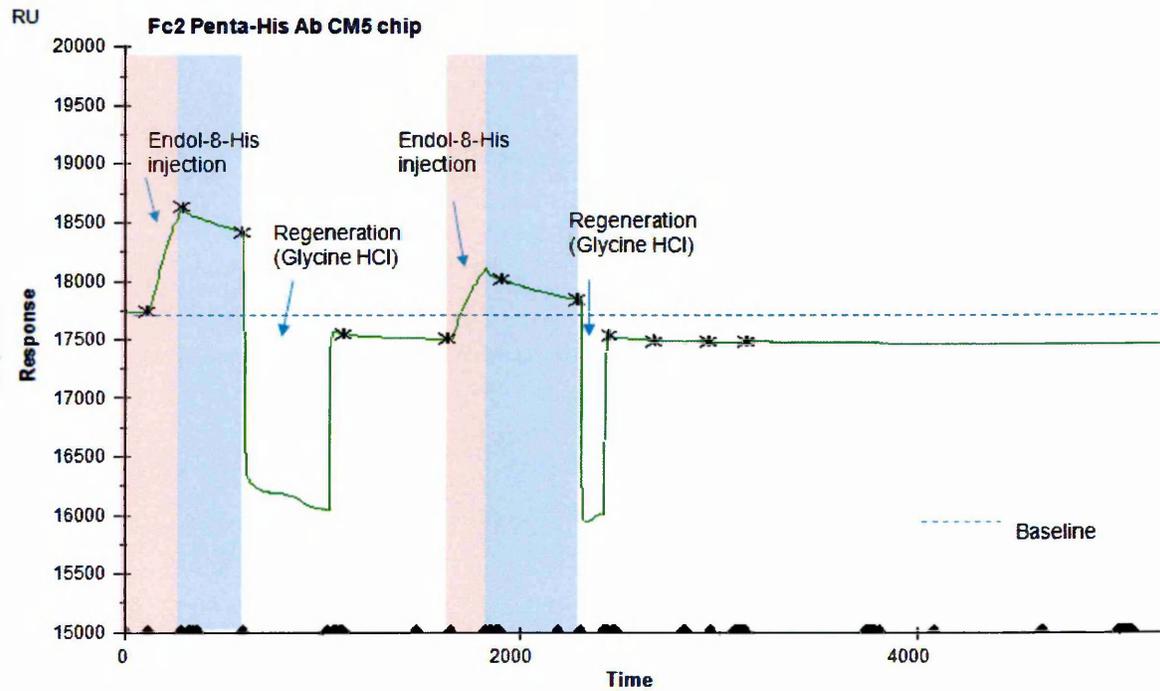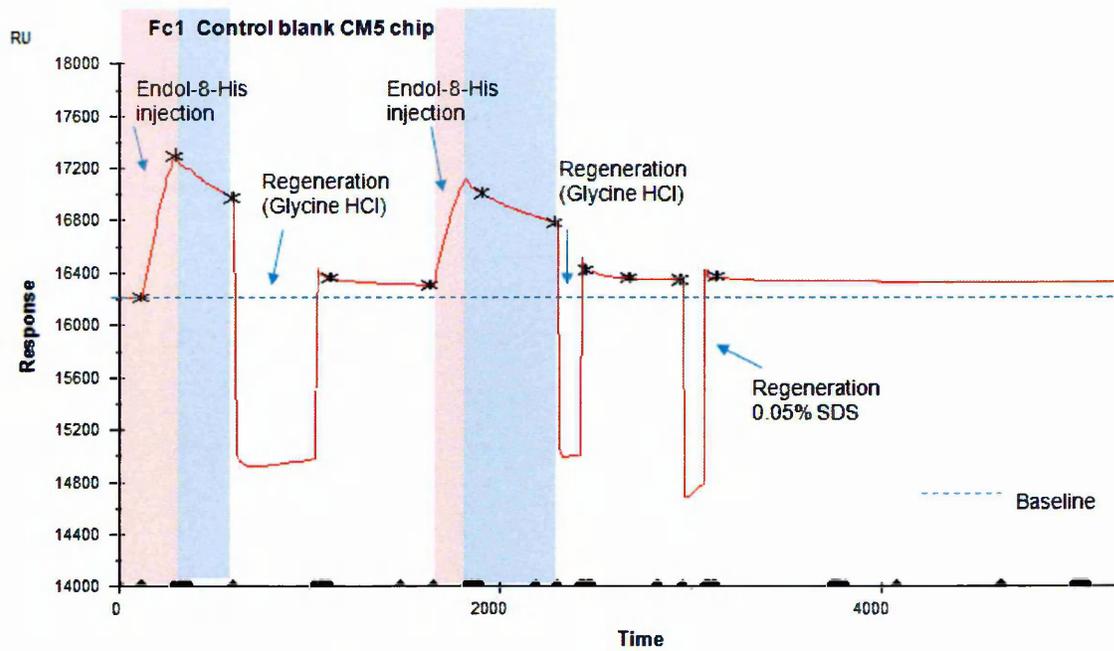
**Figure 5.31 - Capture of purified Endo1-8-His control on penta-His CM5 chip**. Endo1-8-His (2μg/ml) was injected at a flow rate of 7μl/min over both the immobilised penta-His CM5 chip (Fc 2) and the blank reference chip (Fc 1). Association is observed with both chips (pink shading), the dissociation rate was not calculated (blue shading). Regeneration with glycine/HCl did not remove all Endo1-8-His from the reference chip (Fc 1), and regenerated the active chip (Fc 2) stripping protein below the baseline. Endo1-8-His was reapplied and regenerated, baseline was not recovered. Fc 1 was further regenerated with 0.05% SDS, baseline was not reached and Endo1-8-His remained bound to the reference chip. Time is in seconds (s) and response is response units (RU).

260

## 5.3.3.2 Biacore – NTA indirect method to capture recombinant His-tagged proteins for interaction analysis

High levels of non-specific binding were problematic using the CM5 chip immobilised with penta-His antibody. In addition, the regeneration of the surface with mild regeneration buffers appeared to cause irreversible damage to the expensive immobilised surface. Therefore an alternative strategy to capture the recombinant Pmp domains was by use of a NTA sensor chip where NTA is pre-immobilised to the dextran matrix. The chip was activated with $NiCl_2$ to capture Histidine-tagged proteins by metal chelation affinity. Consequently, an NTA method was used to capture histidine-tagged Pmp domains with the empty vector expression products as a control on the reference surface to eliminate non-specific interactions. Upon capture of rPmp domain, interactions with cervical cells were examined in a ligand fishing approach to screen for putative Pmp targets.

Endo1-8-His was readily available and less expensive to use than expressing fresh samples of recombinant Pmp domains for preliminary trials to test the binding specificity of the nickel activated chip (**Figure 5.32**). Endo1-8-His had bound non-specifically to the carboxymethylated dextran of the CM5 sensor chip in previous trials but had not been tested on carboxymethylayed-NTA sensor chips. Upon exposure to both active and reference surfaces the Endo1-8-His protein ($2\mu g/ml$ diluted in running buffer 1) had a higher binding response to the reference surface (without nickel ions) with 1350 RU compared to 970 RU (**Figure 5.33**). Mild regeneration conditions to remove the ligand from the nickel by competitive binding using 50 mM imidazole further increased the RU on the active surface by 150 RU indicating that $Ni^{2+}$ ions on the NTA surface were still available for binding of imidazole, instead of competing and removing bound Endo1-8-His. A harsher regeneration with EDTA, which strips the nickel ions from the NTA surface, resulted in incomplete regeneration; almost 700 RU units remained bound to the sensor surfaces indicating that the protein was non-specifically associated with exposed NTA-dextran matrix as opposed to specific interactions with the nickel ions. To test this theory, a non-specific, non Histidine-tagged protein, BSA (diluted running buffer 1) was applied to the non activated surface and was found to interact directly with the NTA-dextran surface with an increase in 1921 RU (**Figure 5.34**).

Since PBS is a compatible buffer with NiNTA metal affinity chromatography (Qiagen Expressionist, 2003) and Ni-NTA purification trials with truncated rPmps were carried out in phosphate buffers with some success (**chapter 4**), the sample and running

buffers were changed to PBS (running buffer 2), with increased ionic strength, 300 mM NaCl and higher pH 8.0 to reduce non-specific electrostatic interactions. **Figure 5.35** shows the change in buffer conditions prevented non-specific interaction of BSA with the reference surface. Low concentrations of Endo1-8-His in higher ionic strength buffer also did not bind to reference surface, indicating that the non-specific interactions with the NTA chip were most likely through undesirable electrostatic interactions. Since non-specific interactions with the reference surface had been reduced significantly, the activated surface was tested again for specificity with varying concentrations of Endo1-8-His. Subtraction from the reference surface showed the binding of the His-tagged protein was specific to the nickel activated surface with a difference of 3211 RU bound to the activated surface, although high concentrations of Endo1-8-His (200 μg/ml) resulted in some bulk binding to the reference surface (**Figure 5.35**). Furthermore, regeneration with EDTA was sufficient to remove all bound ligands from the surfaces and baseline was restored. Therefore investigations with rPmps and HeLa/Hec1B cells could begin.
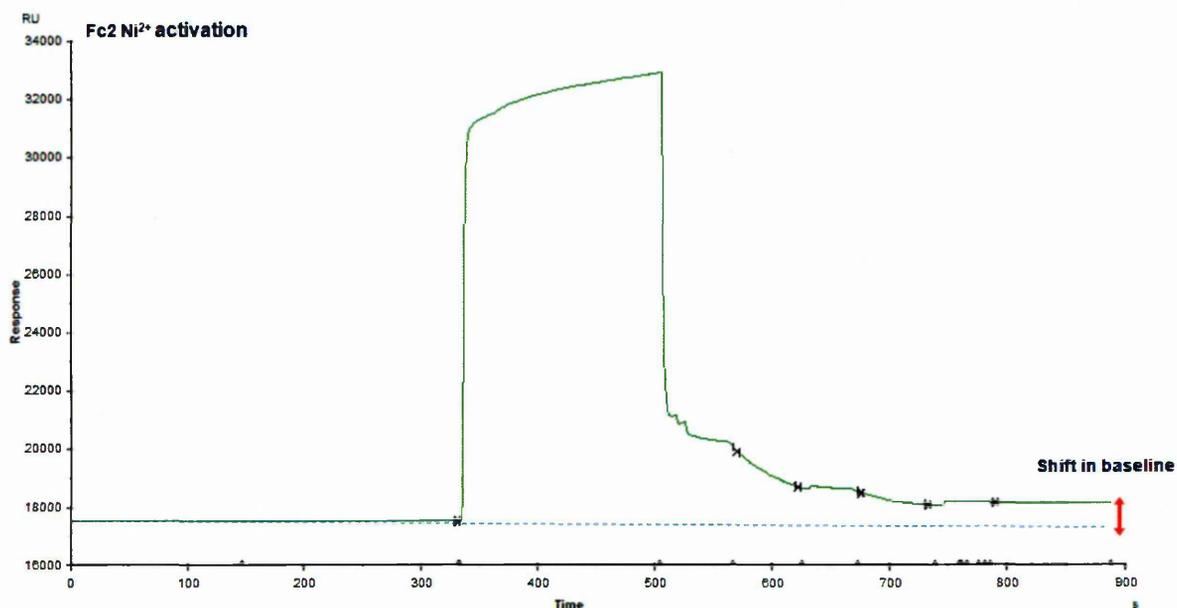


**Figure 5.32 - Nickel activation of pre-immobilised NTA chip, 500 mM NiCl₂ was injected at a flow rate of 7 μl/min, followed by washing with running buffer to remove unbound nickel ions.** The injection produced 629 RU of active surface. Time displayed in seconds.

262

Fc 1 Reference surface

Fc 2 Active surface
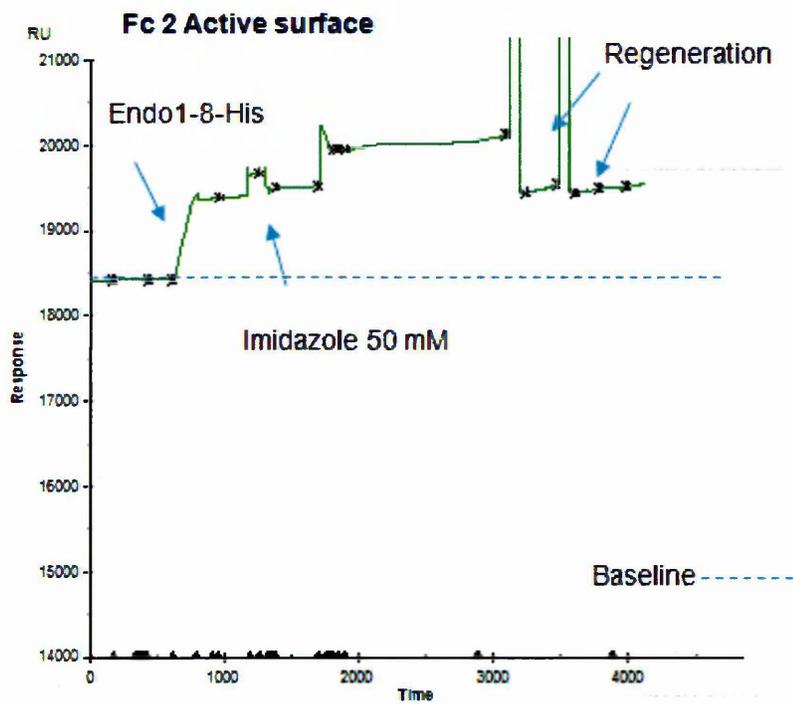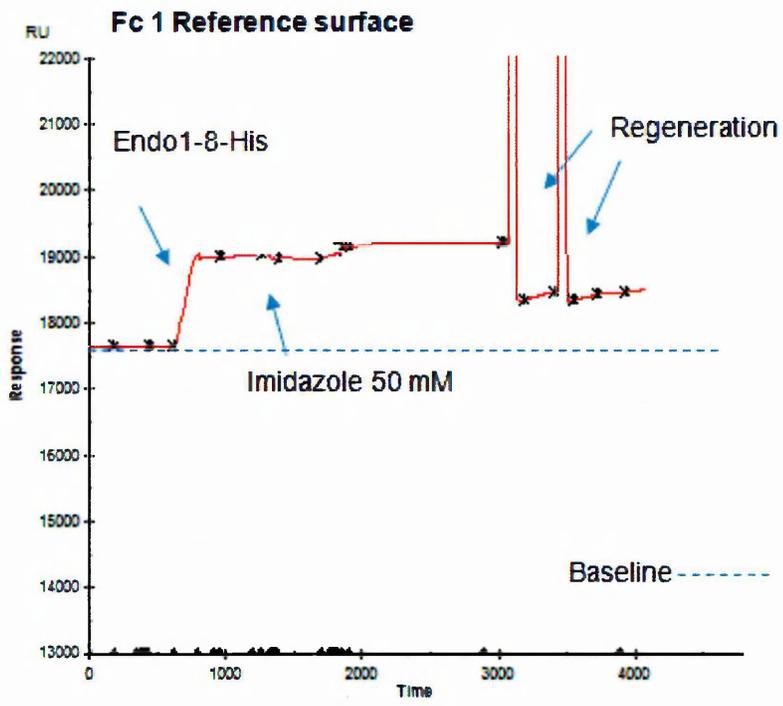
Figure continued on next page...
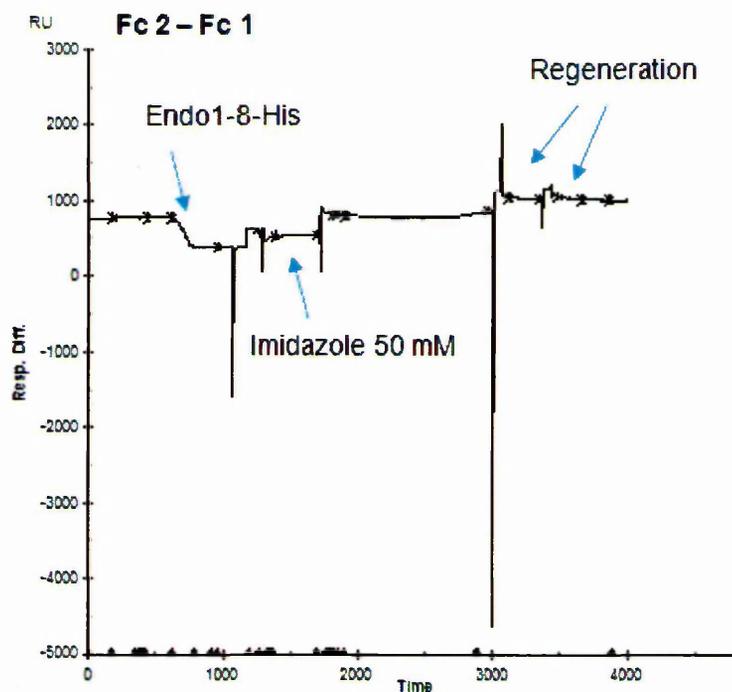
**Figure 5.33 - Endo1-8-His binds non-specifically to NTA dextran surfaces**. Endo1-8-His protein (2µg/ml diluted in running buffer 1, pH 7.4) had a higher binding response to the Fc 1 reference surface (previous page) with 1350 RU compared to 970 RU with FC 2 nickel actived surface (above). Fc 2- Fc 1 (below) shows the negative response difference upon binding with Endo1-8-His. Mild regeneration conditions to remove the ligand from the nickel by competitive binding using 50 mM imidazole further increased the RU on the Fc 2 active surface by 150 RU. Regeneration with EDTA to remove the nickel ions from the NTA surface, resulted in incomplete regeneration; almost 700 RU units remained bound to the sensor surfaces. Time is in seconds (s) and response is response units (RU).
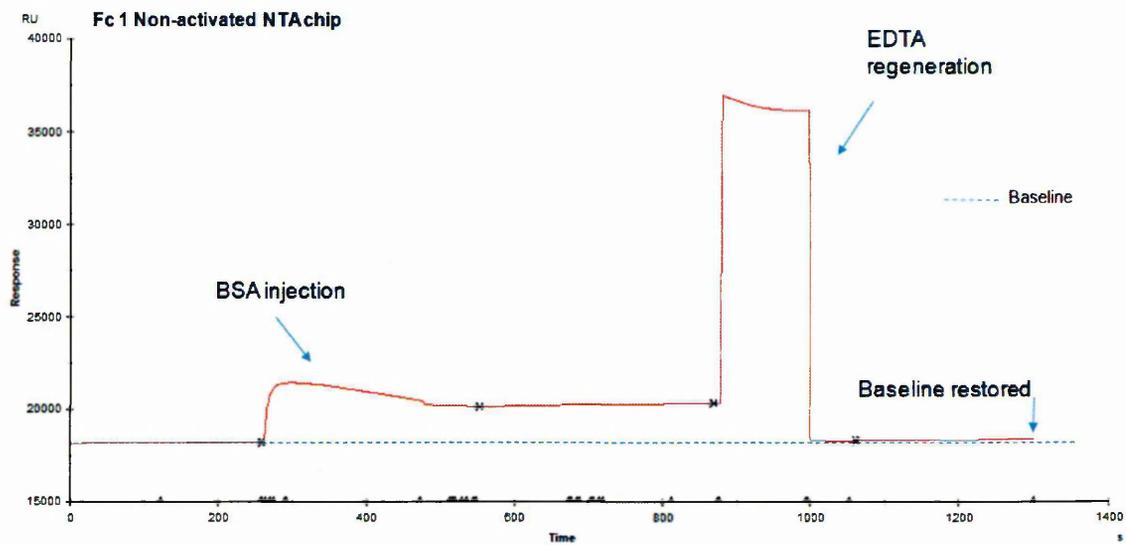
**Figure 5.34 - Non-specific binding of BSA to NTA dextran surface**. Non-Histidine tagged protein (BSA, 1 mg/ml) binds to unactivated reference NTA dextran surface in low ionic running buffer 1. BSA was injected at 7 μl/min and bound with 1921 RU. Regeneration with 350 mM EDTA restored the baseline. Time is in seconds (s) and response is response units (RU).
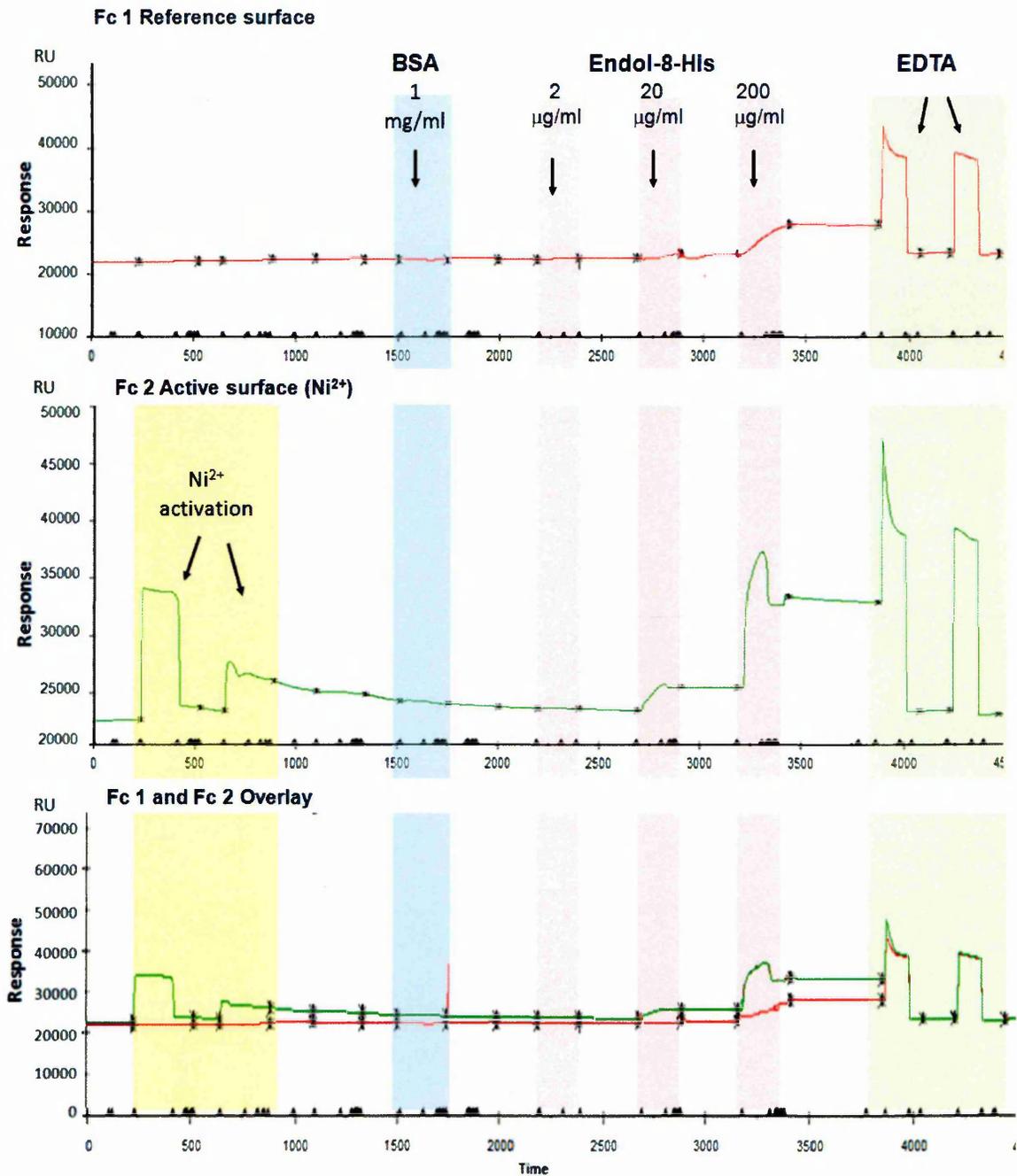
**Figure 5.35 - Sensorgrams show Ni-NTA specific binding of His-tagged Endo1-8-His. Fc 2 was activated with 2 applications of Ni$^{2+}$ (yellow shading).** Non-specific protein BSA (1 mg/ml) was injected over both active and reference flow cells (blue shading). Increasing concentrations of Endo1-8-His were injected over both cells (pink shading) and the 2 step regeneration by 350 mM EDTA is shown in green. Flow rate used was 7 μl/min. Time is shown in seconds.

### 5.3.3.3 Identification of rPmpI-N protein interactions with HeLa cells

The empty vector pET23b was expressed as in **section 4.2.1**. The binding affinity was tested with the reference and active surface to check for non-specific interactions with the NTA dextran surface and nickel activated surfaces. The proteins present within the cell-free expression samples bind with the NTA surface (**Figure 5.36**) but could be used as a control for the subtraction of endogenous binding interactions of the MembraneMax™ proteins present within the rPmpI-N domain cell-free expressed sample.

Proteins from expression of the empty pET vector (30 μg/ml) and the expressed rPmpI passenger domain (rPmpI-N) (30 μg/ml) were dialysed into running buffer 2 and each applied to nickel activated surfaces in **Figure 5.37** and **Figure 5.38**. The results displayed in **Table 5.4** summarise the data showing that the rPmpI-N sample binds to the active surface with a higher response compared to the control sample, indicating putative capture of His-tagged rPmpI-N. At the end of the injection phase of the control sample, the sensorgram showed no binding of the control sample to the surface which was promising except after a lapse of time the response units increased to form a biphasic curve (**Figure 5.38**). The HeLa cell crude lysate was applied to both the control and test flow cell. The response units were subtracted and show the HeLa cell proteins had a stronger response for the control surface over the ligand surface bound with rPmpI-N. The sensorgram shows the binding curves in **Figure 5.37** and **Figure 5.38** and the response units are in **Table 5.4**. The large highlighted peaks in **Figure 5.38** are typical of an accumulation of air bubbles in the integrated fluidic cartridge of the flow cells, this happened regularly despite using freshly made and degassed buffers throughout.

The data generated from these studies did not show any significant binding events that would indicate rPmpI-N interacts with HeLa cells or HeLa cell components as the similar levels of interaction were observed with the control surfaces. Overall the methods were not working with reproducibility or reliability and therefore studies were not extended to include the other rPmp domains or Hec1B cells.
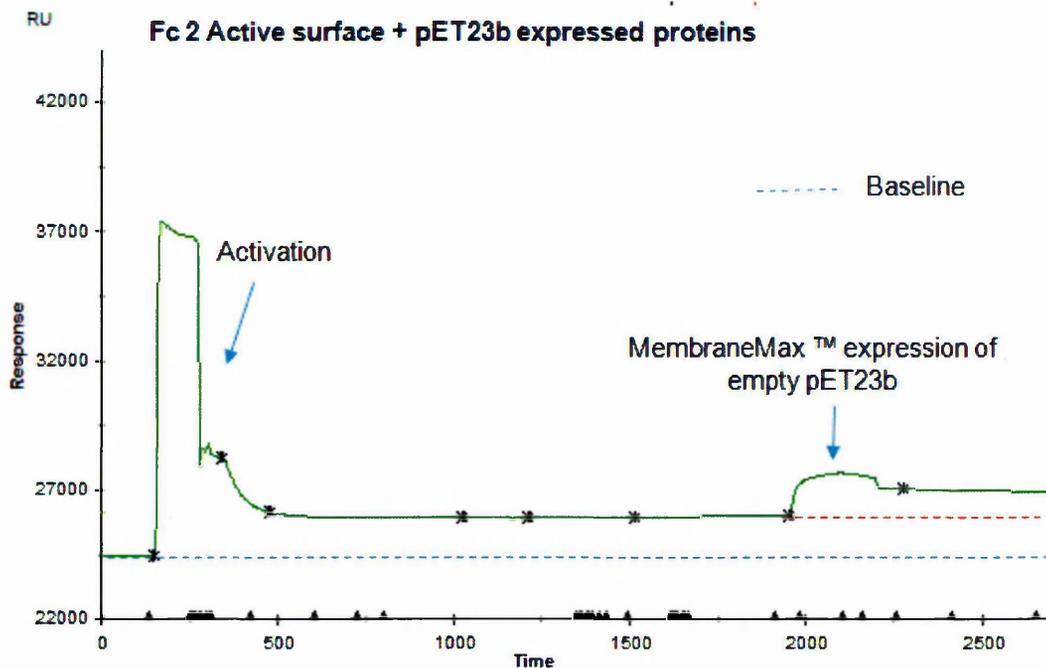
**Figure 5.36 - Sensorgram shows that endogenous proteins in the MembraneMax™ cell-free expression system bind to the active NiNTA surface**. The empty pET23b vector was expressed in the MembraneMax™ system as in section, diluted to 30 µg/ml in running buffer 2 and applied to the activated surface. The surface was activated with nickel ions and the RU was increased from the baseline. Binding can be seen with a further change in RU from the activated surface RU (red dotted line). Time is in seconds (s) and response is response units (RU).

| | Control Ligand (pET23b) | Active Ligand (rPmpl-N) | Response difference (Active – Control) |
|---|---|---|---|
| **Activation** | 526 RU | 400 RU | **-126 RU** |
| **Ligand bound** | 672 RU | 817 RU | **145 RU** |
| **Analyte (HeLa)** | 6950 RU | 4694 RU | **- 2256 RU** |

**Table 5.4 - Summary of response units after activation, binding of the ligand and the analyte, lysed HeLa cells**. The difference in response units show the HeLa cells have a higher binding to the control surface (pET23b). The ligand injection containing rPmpl-N produced a higher response than the control upon association with the nickel activated surfaces.

268
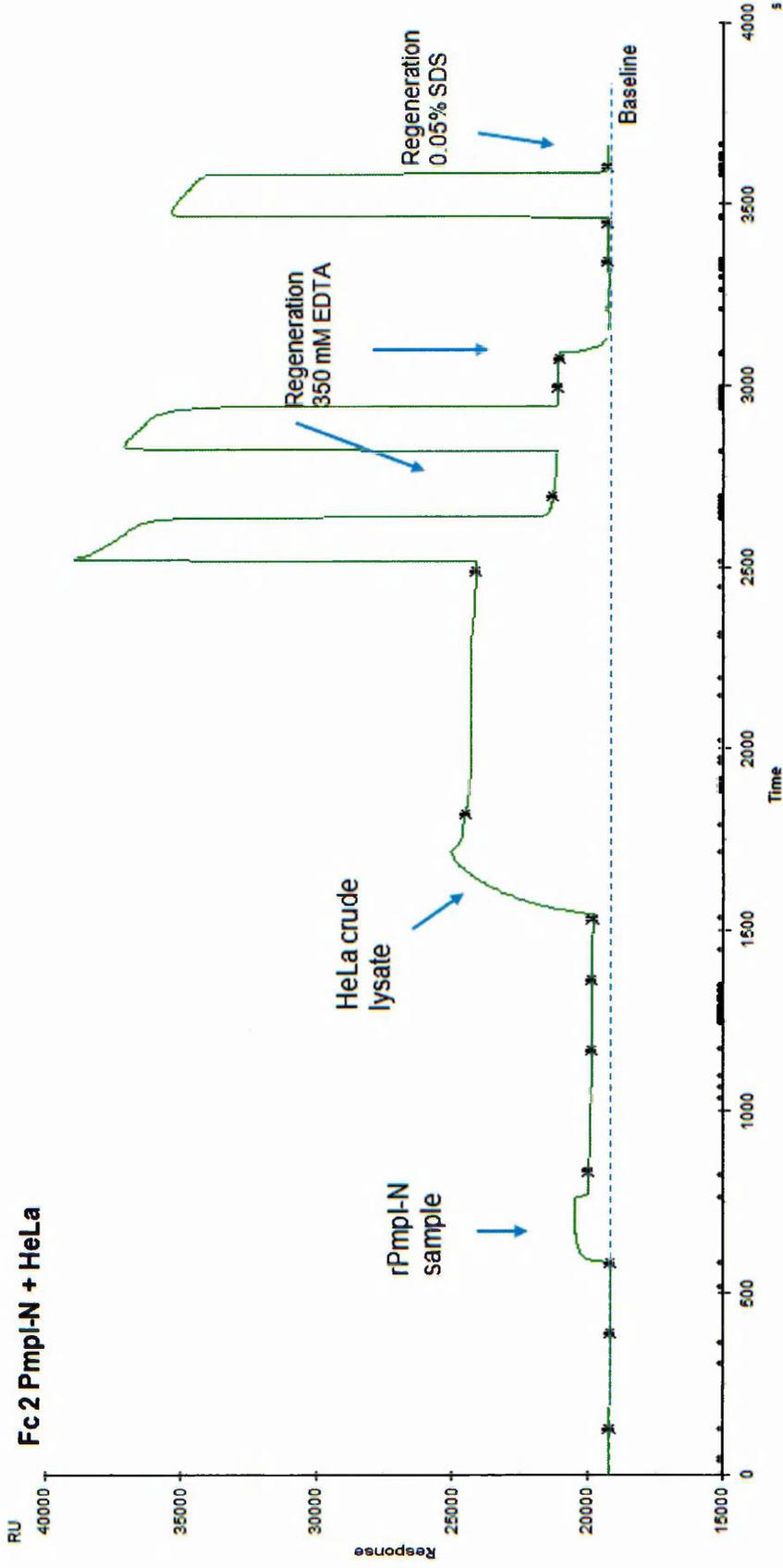
**Figure 5.37 - Sensorgram shows rPmpI passenger domain (rPmpI-N) and endogenous MembraneMax™ cell-free expressed proteins binding with lysed cervical cells.** The ligand (cell-free expression of rPmpI-N) was injected over the NiNTA activated surface of flow cell 2. The HeLa cell crude lysate was passed over the captured proteins to show a large binding event. The surface was regenerated back to baseline with EDTA. The difference in response units compared to the control is summarised in Table 5.4. The control surface is displayed in **Figure 5.38** and the overlay in shown in **Figure 5.39**.

269
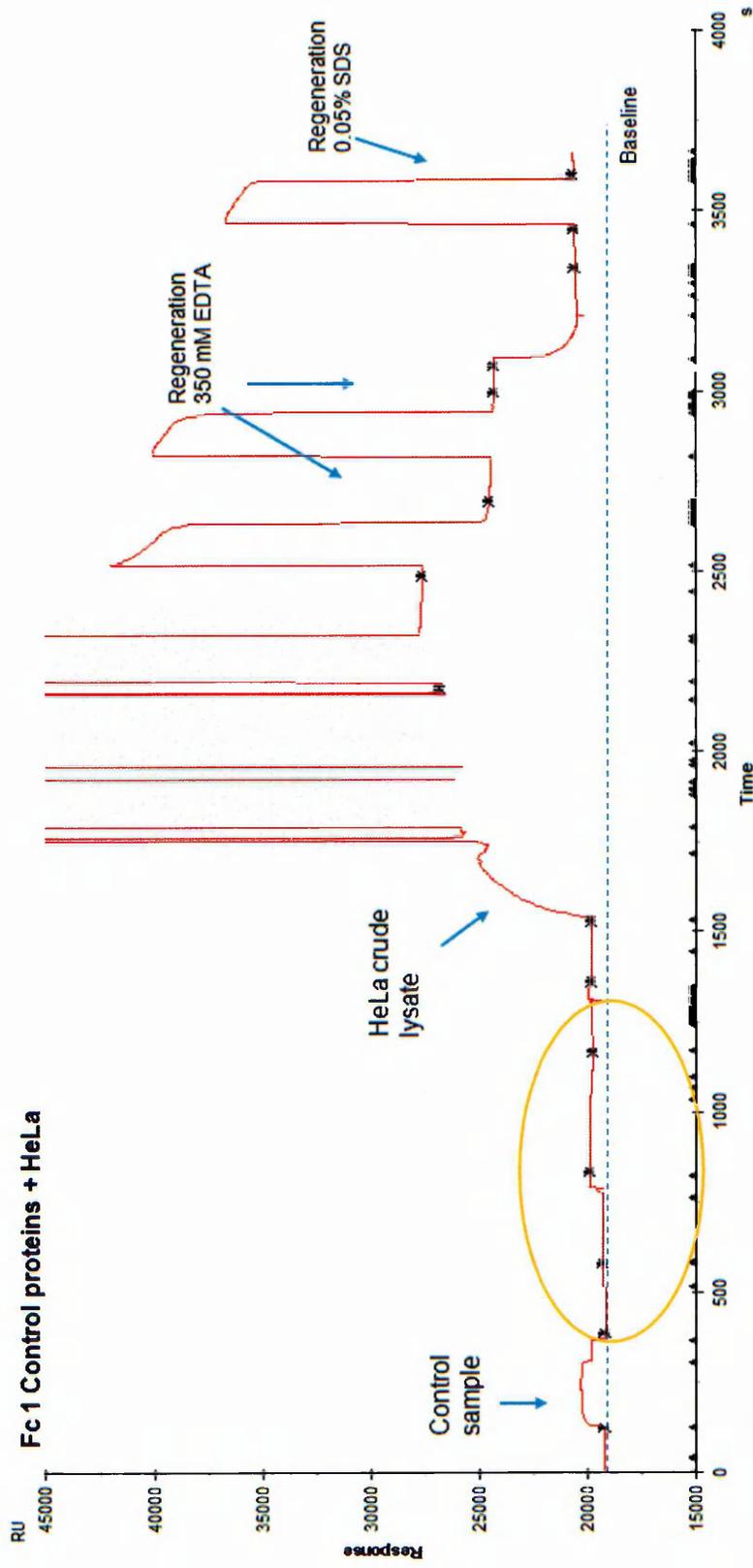
**Figure 5.38 - Sensorgram shows endogenous MembraneMax™ cell-free expressed proteins binding with lysed cervical cells.** The ligand (cell-free expression of empty pET23b vector) was injected over the NiNTA activated surface of flow cell 1 and produced a biphasic curve with response units decreasing then increasing (circled in yellow). The HeLa cell crude lysate was passed over the captured proteins to show a large binding event. Off the scale peaks highlighted in grey area are derived from the problems with IFC in the reference flow cell 1. The surface could not be regenerated fully to baseline with EDTA or SDS. Response units are summarised in Table 5.4. The experimental ligand surface with rPmpI-N is displayed in **Figure 5.37** and the overlay in shown in **Figure 5.39**.
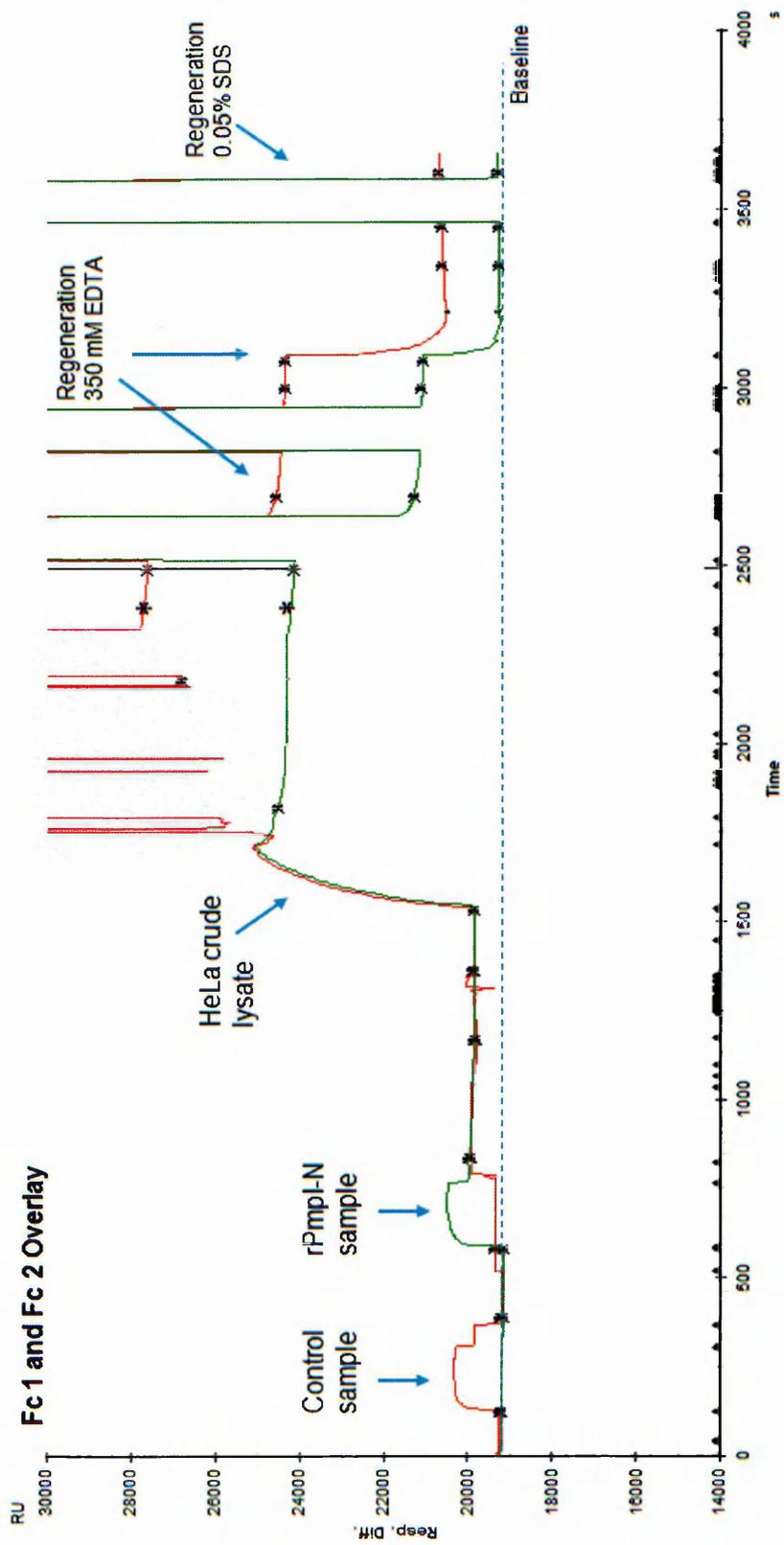
**Figure 5.39 - Sensorgram overlay shows the interaction of both rPmpI-N (Figure 5.37) and control proteins (Figure 5.38) with HeLa cell crude lysate.** Off the scale peaks highlighted in grey area are derived from the problems with IFC in the control reference flow cell 1 (red). The reference surface was not regenerated back to baseline with EDTA or SDS, the active surface (green) was regenerated with EDTA. Response units for each binding event are summarised in Table 5.4.

271

## 5.3.4. Methods to investigate recombinant Pmp domains in interactions with HeLa cells using ellipsometry

Ellipsometry was used to detect potential protein-protein and protein-membrane interactions whereby HeLa cell lysate (with potential orphan receptors) was deposited on the gold surface and recombinant Pmp domains were incubated with the immobilised lysate as in **section 5.2.9**. The rPmp domains were expressed using MembraneMax™ cell-free system and remained unpurified for the assays. Variations in the incidence of reflected light would be indicative of changes to the layers of the thin film surface representative of putative binding events with the lysate.

The spectra in **Figure 5.40** to **Figure 5.43** illustrate the changes in amplitude ratio ($\psi$) and phase shift ($\Delta$) of reflected light from the buffer and the PAH deposition followed by the addition of HeLa lysate and rPmp domains (rPmpI-N, rPmpDauto and rPmpGauto). The HeLa lysate deposition produced a layering of approximately 30-34 nm at the gold surface (**Figure 5.40** to **Figure 5.43**) The addition of each of the rPmp domains with endogenous MembraneMax™ system proteins did not lead to an increase in layering in association with the HeLa lysate, instead a small shift backwards was observed showing that no protein interacted with the HeLa cell crude lysate.
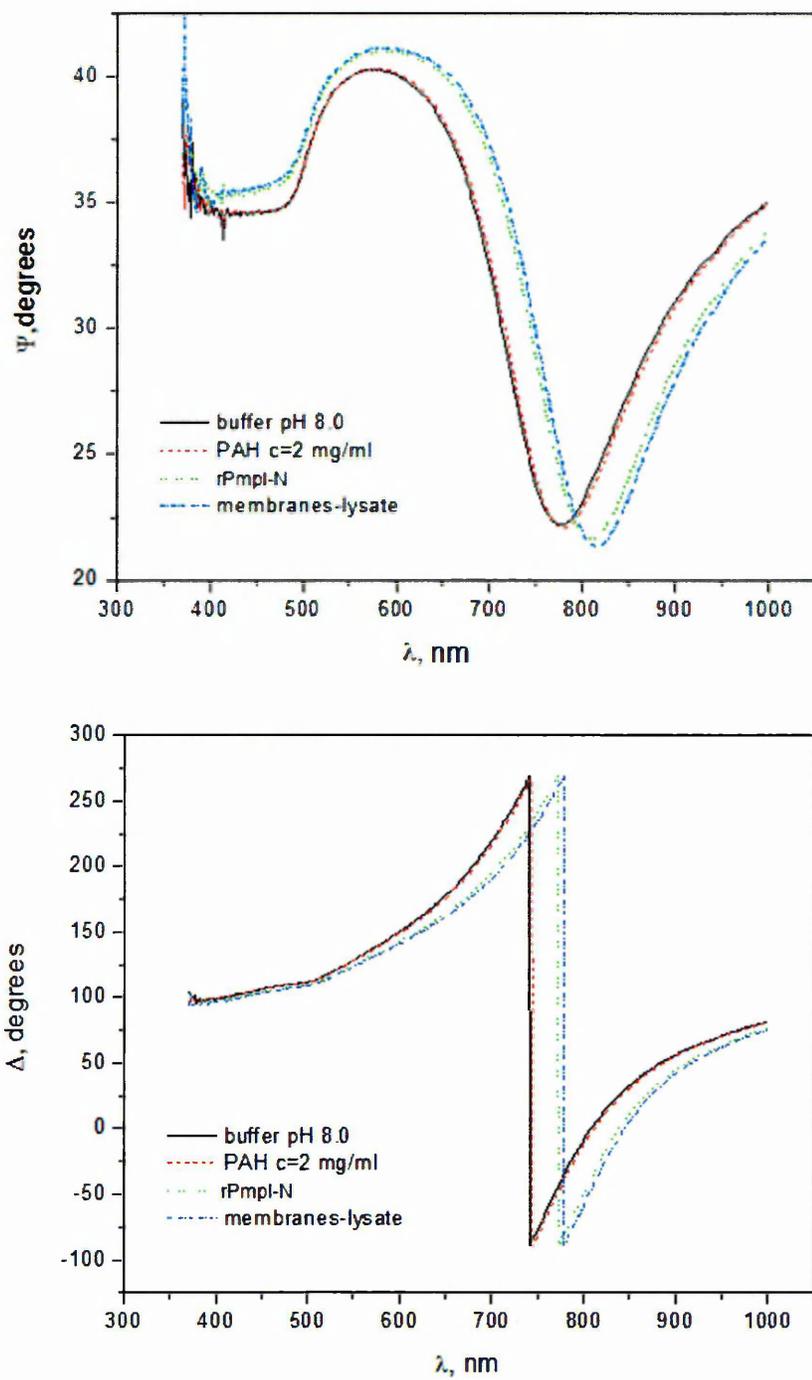
**Figure 5.40 - The amplitude ratio ψ spectra (top) and phase shift Δ spectra (below) from TIRE measurement of the recombinant Pmp I passenger domain (rPmpI-N) in interaction studies with HeLa cell proteins and membranes**. The change in effective layering is shown by the shift upon addition of membranes/lysate to the PAH layer.
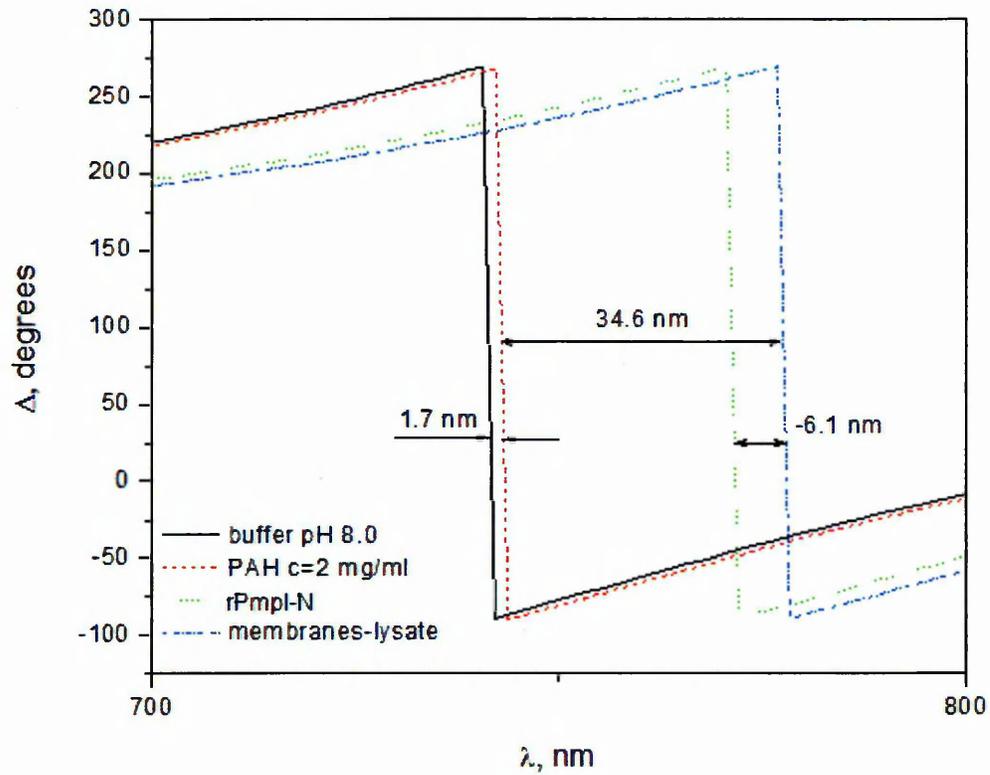
**Figure 5.41 - The phase shift Δ spectra up close showing the layering of rPmpI-N with HeLa cell lysate**. The addition of each application is shown as a difference in layering at the surface in nm, the rPmpI-N domain did not associate with the layer of bound lysed HeLa cells, shown as negative values.
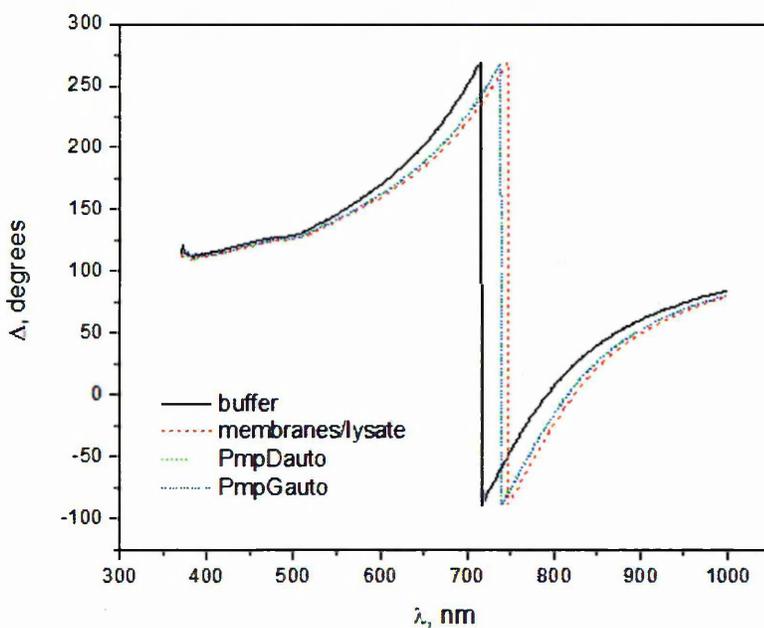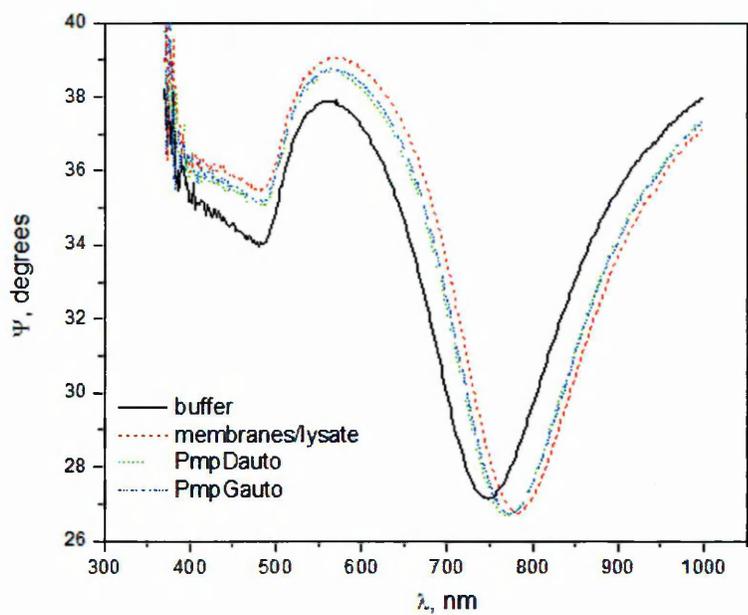
**Figure 5.42** - The amplitude ratio ψ spectra (top) and phase shift Δ spectra (below) from TIRE measurement of recombinant Pmp D and Pmp G autotranslocator domains (rPmpDauto and rPmpGauto) in interaction studies with HeLa cell proteins and membranes. The change in effective layering is shown by the shift upon addition of membranes/lysate to the PAH layer
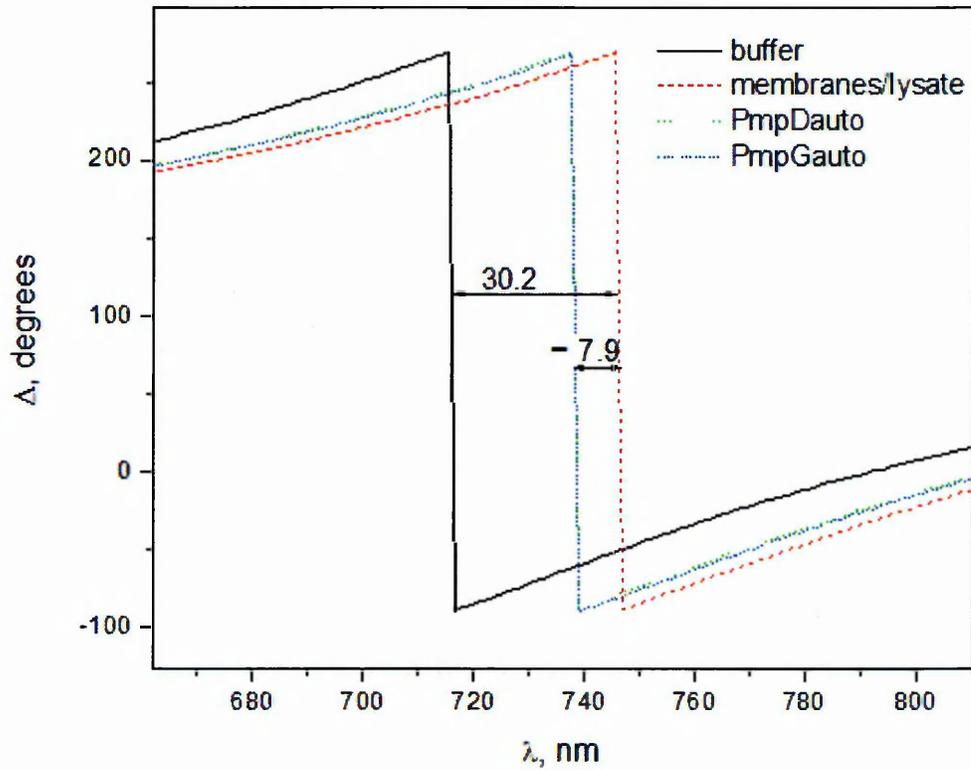
**Figure 5.43 - The phase shift Δ spectra up close showing the layering of rPmpDauto and rPmpGauto with HeLa cell lysate.** The addition of each application is shown as a difference in layering at the surface in nm, the rPmpI-N domain did not associate with the layer of bound lysed HeLa cells, shown as negative values.

## 5.4 Discussion

### 5.4.1 Cross-binding assays using chemical cross-linking

*In vitro* cross-binding assays were unable to determine conclusively that rPmp domains had binding affinity with cervical and/or endometrial cells. Although cross-linking of truncated rPmps was obversed, identifying the components of these cross-linked conjugates was difficult.

Initially, assays investigating interactions between the recombinant Pmp I passenger domain (rPmpI-N) and endometrial cell line (Hec1B) using the cell surface cross-linker BS3 found high mass conjugates difficult to separate by SDS-PAGE. These high molecular weight conjugates detected by Western blotting were retained within the stacking gel and resolving gel interface of the SDS-PAGE gel producing a smear when detected by anti-His antibody. This is typical of protein complexes that are too large to migrate into the SDS polyacrylamide resolving gel using electrophoresis. A possible explanation for this is that Pmps may interact with host cells via a group of known cell surface molecules called heparan sulfate glycosaminoglycans (HSGAGs) (Kobayashi *et al.,* 2011)(Chen and Stephens, 1997). Due to the large size and the glycan moieties, such proteoglycans often migrate between the stacking and running gel interface during SDS-PAGE to also produce the characteristic smears, one example is Agrin, involved in the development of the neuromuscular junction during embryogenesis, (Tsen *et al.,* 1995, Groffen *et al.,* 1998). The large 'smearing' conjugates in this study were in the supernatant fractions (after collecting to cells via centrifugation), extracellular domains of proteoglycans can commonly be shed from the cell surface, generating soluble heparan sulfate proteoglycans (Bernfield *et al.,* 1999). However, rPmpI-N was not applied to the cells as a homogenous protein and therefore it would be premature to speculate cross-linking with endometrial cell proteins had definitely occurred. It is a possibility that these high molecular weight conjugates also involved endogenous *E.coli* proteins (from the MembraneMax™ cell-free *E.coli* extract), nanodiscs and/or oligomers/aggregates of rPmpI-N.

In an effort to optimise the conditions of the cross-binding assays to reduce putative rPmpI-N aggregation, assays were performed at 4 °C and reduced the high molecular weight conjugates, suggesting that the higher temperature of 37 °C forced rPmpI-N to precipitate out of solution into the insoluble complexes that were too large to resolve. Therefore assays with rPmpI-N, rPmpDauto and rPmpGauto were

carried out under these altered conditions but did not reveal any convincing evidence of putative interactions with Hec1B or HeLa cells. In particular, levels of rPmpDauto were so low that detection was difficult despite efforts to incubate the cells with higher levels of protein. It could be disputed that at 4 °C, putative interactions between Pmps and receptors would not take place, it is quite likely that low temperatures were not conducive to binding studies however studies have shown that *Chlamydia* attachment can still occur at these low temperatures in some circumstances (Zaretzky *et al.,* 1995, Davis and Wyrick 1997).

Using the cross-linker BS3 is a shotgun approach to capturing interaction complexes; as a high efficiency NHS ester linker, BS3 will cross-link any and all interacting molecules whose respective lysine residues come within the 11.4 Å spacer arm of the linker (Sinz, 2010). Because lysine residues are normally in plentiful supply on most proteins the formation of conjugates consisting of several components is sometimes an inevitable outcome. Paradoxically, in assays with rPmpl-N and Hec1B/HeLa cells these conjugates were also found when cross-linker was not used. Additionally, the same high mass conjugates were also detected in assays where Hec1B and HeLa cells were omitted to test the cross-linking of rPmpl-N under the same buffer and incubation conditions and led to the conclusion that these conjugates were cross-linked aggregates or oligomers of rPmpl-N. It has been reported that Pmp D forms flower-like oligomer structures (Swanson *et al.,* 2009) and so it is possible that other Pmps could also form oligomers, although this theory would need further investigation.

An alternative approach was to use a cross-linker with higher specificity. With the exception of zero-length cross-linkers, formaldehyde has the shortest cross-linking span (~2-3 Å) of any cross-linking reagent (Sinz, 2010). Cross-linking with formaldehyde did not show any traces of size shifts suggestive of interactions of rPmpl-N with Hec1B cell proteins. However, having such a short arm span formaldehyde requires the reactive groups of interacting proteins to be in very close association and so is only ideal for fixing specific protein-protein interactions.

Heterobifunctional cross-linkers which possess two different reactive groups as opposed to the identical reactive group present on homobifunctional cross-linkers were considered as an alternative approach. Heterobifunctional cross-linkers allow for a two-step conjugation method that can minimise unwanted self-conjugation. In sequential procedures, heterobifunctional reagents are applied to the target protein

278

first where one reactive group reacts with the target protein. In this case, the modified rPmp domain could have then been applied to the uro-gential cell lines where reaction through the second reactive group of the crosslinker would occur where interaction took place. However, this would not be as straightforward for these studies as the rPmp domains were not purified samples therefore other endogenous proteins in the MembraneMax™ system would also become modified with the ability to cross-link leading to false positives. Furthermore, modification by addition of cross-linker directly on the protein may affect the functionality of the rPmp domain and without feasible tests for activity would not be conclusive. Purifying the rPmp domains to reduce the heterogeneity prior to cross-binding assays may have minimised non-specific binding events however, the levels would have been beyond the limits of detection so this was not an option.

In summary one of the major problems throughout these assays was the inability to examine the cross-linked samples thoroughly. The signal from rPmpI-N was weakened significantly when cell surface cross-linking with BS3 was carried out which indicated three possible events; i) degradation/loss of rPmpI-N had occurred upon cross-linking, ii) His-tag epitopes were inaccessible after cross-linking or iii) the majority of the protein was cross-linked into conjugates too large to migrate into the SDS-PAGE gel. Since dot blot analysis found that that the epitope of cross-linked recombinant PmpI-N protein was not affected and the strong signal in all the assays did not indicate significant loss of protein, it was considered that the main problem was the size of cross-linked conjugates and the transfer of these onto nitrocellulose membranes. Using dot blot analysis to overcome these problems was not a solution as successful cross-binding identification relied upon identifying a size shift from monomers to higher mass conjugates. These studies did infer that cross-linking was taking place between rPmpI-N and Hec1B cells, however in other trials high molecular weight complexes were found in the supernatant separated from the cells, with similar bands detected when cross-linking rPmps without cells. Therefore in some circumstances it was concluded that the rPmps had self-conjugated or to extracellular domains of proteoglycans.

## 5.4.2 Immunofluorescence

This technique was used to overcome the problems experienced with aggregated protein that produced ambiguous results in the cross-binding assays as this method does not rely on SDS-PAGE techniques to separate proteins prior to

immunoprobing. Instead, aggregated non-specifically bound protein is removed during washing of the cells to allow fixing of only bound truncated rPmps.

This is the first study of its kind to incubate recombinant truncated Pmp domains with endometrial and cervical cells for immunostaining methods. This method was chosen as a preliminary screening assay to detect fluorescent signals that could indicate localisation of potential rPmp binding to the cell surface. ICC may not be a fitting model to test recombinant proteins in adherence. Instead expressing the rPmps to the membranes of *E.coli* and applying recombinant *E.coli* to interact with the cells is a more practiced method. However, expressing the truncated rPmps to the membranes was very challenging due to the high amounts of inclusion bodies formed during the expression trials and the expression levels of truncated rPmps were low in *E. coli*.

Staining using anti-His antibodies directed against the His-tagged recombinant trunctated Pmps indicated possible localisation of the rPmp domains to the cell surface of Hec1B cells. Such putative interactions were observed between each rPmpGauto, rPmpDauto and rPmpI-N with the endometrial Hec1B cell line but not with the HeLa (cervical) cell line. This could support the theory that Pmps function as adhesins with distinctive organotropisms for different tissue sites (Gomes *et al.,* 2006) to the extent that certain Pmps of serovar E *C.trachomatis* target the endothelium and not the cervix. There is as yet no published data to support this or to illustrate which, if any, Pmps target definite tissue types. However studies with human Hec1B endometrial cells have shown that serovar E is closely associated with the protein disulfide isomerase (PDI), a component of the oestrogen receptor complex (Davis *et al.,* 2002). This is further supported by earlier studies which showed *Chlamydia* attachment is dramatically enhanced in oestrogen dominant primary human endothelial Hec1B epithelial cells, (Maslow *et al.,* 1988). The data from cross-binding truncated rPmps with Hec1B cells in this study were inconclusive and so although it was tempting to suggest such a hypothesis, this was highly speculative and requires further study as immunocytochemistry in this context was not a very specific or reliable technique.

In some instances despite efforts to optimise the amount of antibody used, a strong fluorescent signal was noted in the negative control assays that were untreated with His-tagged protein. Although the staining patterns between the Hec1B cells treated with rPmp domains displayed obvious spatial differences compared to the staining

of the control assays treated with proteins from the same empty expression vector, this could have been as a result of problems with the antibody specificity which was a concern throughout these experiments. Even a trivial event such as insufficient permeabilisation of the cells prior to antibody staining could have prevented access of the antibody to within the cell, resulting in accumulation of the antibody to the cell-surface by hydrophobic or electrostatic interactions.

Therefore one of the problems throughout optimisation of the ICC investigations was the apparent cross-reactivity of the anti-His antibodies with the HeLa cells and the Hec1B cells and efforts were made to minimise these background signals. Initially the penta-His primary antibody was used in combination with a fluorescently tagged secondary, anti-rabbit conjugated FITC, however this combination displayed high cross-reactivity with both HeLa and Hec1B endogenous proteins in spite of titration to low concentrations, this cross-reactivity was also evident during the Western blot analyses of the cross-binding assays. An alternative approach was to use an anti-His primary that was specific to a 6XHis epitope as opposed to a 5XHis epitope (penta-His) to minimise cross-reactivity that can result from the presence of histidine rich proteins residing within the HeLa and Hec1B cells. Concerted efforts were made to titrate the antibodies for each staining method, however the risk with using antibodies at very dilute concentrations was a consideration as this may have compromised the detection of the rPmp domain. The 6XHis antibody displayed no cross-reactivity with HeLa or Hec1B cells when tested on a Western blot, on the basis of this it was considered suitable for ICC however a high background fluorescent signal was still evident in ICC experiments particularly with the HeLa cells. This high background was mainly demonstrated as granular staining within the nuclei. Interestingly in a previous study, upon fractionation of HeLa cytoplasmic or nuclear extracts on NTA columns, enrichment of a cellular protein was observed from HeLa nuclear extracts (Schmitt *et al.,* 1993). This protein bound was identified as Oct-6 a transcriptional factor, previously believed to be expressed only in undifferentiated cells (Suzuki *et al.,* 1990) but has since been found in other cell types including fibroblasts and macrophages (Hofmann *et al.,* 2010). In agreement with its function as a transcriptional regulator during neurogenesis and myelination, Oct-6 is predominantly found in the nucleus, in addition Oct-6 is thought to shuttle constantly between nucleus and cytoplasm by use of a nuclear export signal (Baranek *et al.,* 2005). Interestingly Oct-6 contains a naturally occurring stretch of 6 histidines near the C-terminus (amino acids 436-441) as well as another His rich

281

sequence (amino acids 208-216) (Suzuki *et al.,* 1990). These histidine enriched sites were thought to mediate the high affinity binding of Oct-6 to the NTA matrix in the studies by Schmitt *et al.,* (1993) . The granular staining within the nuclei of the HeLa and Hec1B cells in the ICC assays and the Histidine bands detected in the cross-linking assay Western blots could have been anti-His binding to the 6Xhistidine stretches present within Oct-6 transcriptional factor and so using an anti-His antibody may not have been an appropriate antibody for these trials. However, anti-His antibodies are commonly used in ICC to identify eukaryotic cells expressing Histidine-tagged recombinant proteins (Hao *et al.,* 2006).

Addtionally, autofluorescence of mammalian cells is often a common problem during immunofluorescent techniques therefore Sudan black B that quenches lipofuscin autofluorescence was used throughout the assays to minimise background signal (Schnell *et al.,* 1999). However, many other cellular metabolites exhibit autofluorescence, including flavins and NADPH whilst generally, proteins rich in tryptophan, tyrosine and phenylalanine show some level of autofluorescence (Billinton and Knight 2001) Interestingly, a study was able to quantify *E.coli* and other bacterial cells dependent upon the autofluorescence of bacterial cell lysate (Bao *et al.,* 2008) and so high background signals may have resulted from the natural fluorescence emitted from the MembraneMax™ proteins derived from *E.coli* lysate rather than non-specific binding of the antibodies.

Although localisation was observed with hec1B cells, similar findings were not observed with HeLa cells. ICC methods that involve washing away excess proteins from adhered cells could serve to disrupt 'weak' non-covalent interactions. Interestingly, during a preliminary optimisation trial, localisation may have been demonstrated in assays where the washing step was omitted, this procedure was only intended to verify that the concentration of rPmpI-N applied to the cells was within the detection range of the antibody dilution used but the findings were diverse in showing localisation of staining for rPmpI-N. Notably macromolecular interactions can often be of unexpectedly low affinity even when they are of clear physiological importance. Studies investigating the affinity and kinetics of such interactions using a technique based on surface plasmon resonance have shown that cellular adhesion molecule interactions that mediate transient cell adhesion have surprisingly low affinities and extremely fast dissociation rate constants (Anton van der Merwe, P. and Neil Barclay 1994). Other examples of low affinity complexes are those formed transiently by proteins involved in electron transfer (Pelletier and Kraut

282

1992), and weakly bound multi-enzyme complexes, such as those formed by enzymes of the tricarboxylic acid cycle (Srere *et al.,* 1997). The low stability of these non-covalent associations is such that, *in vitro*, the complexes usually dissociate, making non-covalent interactions challenging to study. However to investigate the putative interactions observed using chemical cross-linking and antibody staining which displayed some promising results, surface plasmon resonance was applied.

### 5.4.3 Ligand fishing for potential binding interactions between passenger domain of rPmp I and cervical cell lines in vitro using Surface Plasmon resonance

SPR was used to investigate putative interactions between truncated rPmp domains observed with crossbinding and ICC that could potentially be weak and transient interactions difficult to confirm within those detection limits, for example rPmpI-N binding weakly to HeLa cells where interactions may have been observed when the protein was not displaced by washing. Two indirect capture methodologies were trialled during these studies as this was a novel method for Pmps and developing an appropriate procedure and suitable conditions was a priority based on the properties of the ligand to be immobilised, including factors such as purity and availability of affinity tags. These indirect methods were based on the capture of the His-tag on the C-terminus of the rPmp proteins by either immobilised anti-His antibodies or nickel activated NTA surfaces. Lysed mammalian cells were then passed over immobilised truncated rPmps in an effort to capture a binding event. To eliminate non-specific binding of MembraneMax™ expression proteins within the rPmpI-N sample, the empty expression vector products were expressed and captured on a reference surface. Using the indirect penta-His capture techniques was problematic due to non-specific binding to the penta-His antibody. Studies using the indirect Ni-NTA immobilisation methods for SPR were only tested for the recombinant Pmp I passenger domain (rPmpI-N) with HeLa cell lines and were inconclusive. Although interactions occurred between the un-purified recombinant PmpI-N domain and the HeLa cells for this method, interactions of a higher response occurred with the negative controls containing the same un-purified proteins minus the rPmpI-N domain.

Initially, as the avidity of the penta-His Ab was good for detecting rPmp domains within MembraneMax™ cell-free heterogeneous samples, this indirect technique based on the capture of the recombinant His-tagged was intended for SPR studies.

However, upon coupling of the penta-His antibody to the sensor surface, un-purified rPmpI-N was not captured by the antibody. Subsequently function and specificity of the coupled penta-His was tested using a purified His-tagged control protein Endo1-8-His. Binding was not specific to the penta-His surface. Furthermore, mild regeneration buffers were sufficient to cause irreversible damage to the coupled surface by stripping or denaturing the penta-His Ab, yet harsh conditions with SDS did not remove Endo1-8-His from the control surface (lacking antibody), indicating an undesirably strong binding affinity to carboxymethylated dextran. Increased pH and ionic strength of the samples did not reduce the non-specific associations. Consequently sensor chip surfaces were unsuitable for experimental assays and optimisation of these conditions were not pursued as such methods are extremely time consuming and expensive with the use of multiple sensor chips and milligram quantities of antibody. However, other studies have successfully used penta-His as a capture molecule in biacore studies (Giannetti and Björkman 2004, Nordin *et al.,* 2005, Glück *et al.,* 2011).

Instead an alternative indirect capture method was used employing a NTA surface based on Nickel metal affinity capture. An advantage of using NTA chips is that generic regeneration conditions can be employed by stripping the nickel ions from the surface to remove bound analyte thus minimise assay development. Importantly, the surfaces can be reactivated with nickel ions between applications unlike the CM5 chips that once regeneration has taken place under conditions too harsh, a new sensor chip is necessary. Before protein-protein interactions could be tested, further assay development was carried out using control ligands (His-tagged and non His-tagged proteins) in an attempt to produce optimal binding conditions as Endo1-8-His, an 8 His-tagged purified protein was found to bind strongly with both nickel activated and non-activated surfaces, displaying little to no specificity for the nickel ions. This non-specific binding was also confirmed with a non His-tagged protein bovine serum albumin (BSA). This non-specific binding was most likely attributed to electrostatic interactions with the highly negatively charged dextran surface. A change in buffer to PBS with increased pH and ionic strength resulted then in the capture of Endo1-8-His whilst preventing the binding of BSA. This also significantly reduced the binding of Endo1-8-His to the reference surface (minus nickel ions).

The revised buffer conditions were used for all subsequent experiments. The rPmpI-N domain was applied to the nickel activated surface alongside the empty

expression vector products lacking rPmpI-N to provide a control as similar as possible for reference subtraction. Subsequently, the rPmpI-N sample was shown to give a higher binding response than the control that would indicate the capture of His-tagged rPmpI-N from the sample. This was observed on each occasion where the control surface binding interactions were subtracted, indicating a higher specificity for His-tagged proteins. However, some endogenous protein binding (from the MembraneMax™ expression system) was expected since several endogenous proteins were seen to bind to the Ni-NTA agarose during the purification of rPmp domains (**chapter 4**). Since the MembraneMax™ expression system is derived from *E.coli* lysate, Tween 20 or urea can be used as additives to reduce *E,coli* components adsorbing non-specifically onto sensor surfaces (Oshannessy *et al.,* 1995), however these were considered unsuitable for use in rPmp binding assays as urea is a denaturant and detergents may disrupt the lipid based components of the NLP nanodiscs and the analyte (HeLa) cell membranes possibly compromising potential binding sites. Therefore with such a sensitive method and closely matched controls it was proposed that subtraction of the binding response would be sufficient to indicate specific capture of rPmpI-N. Furthermore, other studies have used this methodology to capture and purify affinity-tagged proteins using Biacore sensor chips, including the cytoplasmic C-terminal domain of the receptor of the *Xenopus* calcium channel, $IP_3$, (inositol 1,4,5-trisphosphate) using *E.coli* lysate as a reference blank (Natsume *et al.,* 2002). Although the experimental design appeared to be working between active surface and capture of rPmpI-N, the application of the HeLa cell lysate did not bind with a higher affinity to the experimental surface containing rPmpI-N. Instead, the components of the HeLa lysate produced a higher binding response with the control surface containing the expressed empty vector products. It was likely that the high occurrence of non-specific binding was observed due to the heterogenous ligand and analyte samples and ultimately, flaws in the experimental design of these assays were prominent and could not provide any convincing data.

Some of the difficulties with background signal using the Ni-NTA capture method were as a result of using complex, heterogenous samples. This can be overcome by using a direct immobilisation technique in which the ligand (rPmpI-N domain) is directly coupled to the CM5 sensor chip surface and the analyte (HeLa cells) is passed over in solution omitting the use of antibodies and other capture techniques. However purified ligand is usually necessary when using this approach. In this study

this mechanism was not utilised due to several limitations. Firstly, rPmpI-N could be only partially purified, therefore contaminants that are not removed during purification would also be directly coupled to the surface limiting the quality of the procedure. Furthermore, ligands that are directly coupled to the surface can display a random orientation and so potential binding sites within the protein may become inaccessible. This was observed in studies by Wear et al., (2005) where the surface charge potential of CypA a recombinant cyclophilin, a potential drug target for a several diseases including HIV and malaria infection, was altered in low pH immobilisation buffers. Directly immobilising CypA to the sensor chip surface resulted in the molecule being orientated with its binding surface downward toward the activated chip surface. A further consideration using this approach is the regeneration step necessary to remove bound analyte from the ligand in order to perform further applications, critical for successful assays. Regeneration conditions are dictated by the stability and nature of the ligand, often harsh regeneration conditions cause irreversible damage to the bound ligand, yet mild conditions can allow analytes to accumulate on the surface decreasing the available binding sites. As the rPmps have displayed a tendency to precipitate during buffer exchange even in physiological buffers, it was likely that these proteins would not withstand even mild regeneration when bound to a sensor chip surface. In fact it is quite likely that the process of amine coupling the rPmp to the chip surface would denature the protein in the first instance. Moreover, immobilised ligands by indirect capture methods, such as the Ni-NTA method used, are reported to exhibit higher binding activity often as a result of a homogenous arrangement of the ligand and accessibility of the binding sites and so can be advantageous over direct immobilisation (Baumann et al., 1998).

Alternative approaches for Biacore analysis would be by immobilisation of the orphan receptors (HeLa cell lysate or membranes) to the surface of a hydrophobic sensor surface such as the HPA chip, this facilitates adsorption of lipid monolayers for analysis of interactions involving lipid components. However, this was considered unsuitable as in similar procedures, the use of detergents is recommended to prevent aggregation within the microfluidic flow system. The possibility that even small concentrations (0.01%) of detergent would either release potential receptors from cellular membranes or shield targets/ligands in micelles or disrupt the nanodiscs were considered. Ellipsometry, an alternative optical biosensor technique,

lacking a microfluidic system thus allowing the immobilisation of complex materials, was used instead.

Finally, according to an extensive review by (Rich and Myszka 2008) the vast majority of biosensor data published is of poor quality, displaying poor experimental design with over interpreted results and application of complex models to fit data sets. Despite this the data are accepted by peer reviewers. The data obtained from the Biacore study indicated binding events but the sensorgrams produced were sometimes presented with spikes, jumps and drifts, occurrences mainly caused by instrument artifacts and sample aggregation. Biphasic binding curves were also observed when applying the control analyte to the surface which are often described as rapid and slow interaction of molecules with the surface yet this is difficult to substantiate with confidence as such curves detail more complex responses than an exponential binding curve. Additionally, binding responses are not always applicable to true binding events, mismatched sample and running buffers, non specific binding, analyte aggregation/precipitations as well as poor instrument performance are often misinterpreted as interacting molecular events (Rich and Myszka 2008). However in this study, alterations in conditions and repetition of experiments were undertaken to minimise artefacts

## 5.4.4 Elipsometry-TIRE to measure truncated rPmp interactions with HeLa cells

Total internal reflection ellipsometry measures the thickness of surface layers or the stacking of materials on a surface by reflected light, similar to SPR analysis but measures two components of the oscillating reflected light as opposed to the change in intensity and therefore has been described as being more accurate and sensitive than conventional SPR (Nabok *et al.*, 2005). Recently several publications have emerged validating this technique for the measurement of protein-protein interactions (Nabok *et al.*, 2007, Kriechbaumer *et al.*, 2011, Kriechbaumer *et al.*, 2012a, Kriechbaumer *et al.*, 2012b) where protein interactions with membranes and putative membrane receptors have been explored with positive results. This technique has been used to measure quantitative interactions between the novel plastidial chaperone receptor OEP61 and isoforms of the chaperones Hsp70 and Hsp90 (Kriechbaumer *et al.*, 2011). In addition the behaviour of such membrane protein receptors in their native membrane environment was explored (Kriechbaumer *et al.*, 2012a). G-protein-coupled receptors (GPCRs) have since

287

been used in ellipsometry to quantify specific interactions of receptors within native mammalian cell membranes, where the GPCR-ligand CXCL12α and its receptor CXCR4 were used. Human-derived Ishikawa cells were deposited onto gold coated slides and interactions between the receptor CXCR4 on these cells and its ligand CXCL12α were detected. This interaction was inhibited by application of the CXCR4-binding drugs (Kriechbaumer *et al.*, 2012b).

Although this is a relatively new approach for studying membrane protein interactions, the published data shows it to be a reliable method in overcoming some of the challenges associated with studying membrane protein interactions in a lipid environment. Based on the publisehed methods by (Kriechbaumer *et al.*, 2012a), experiments were carried out using HeLa cell lysates as orphan receptors on to the gold surface and the unpurified soluble rPmp domains ( in nanodiscs) were laid over the orphan receptors using the TIRE method to monitor any changes in molecular layering that was indicative of ligand-receptor binding. This method found no interaction between truncated rPmps and HeLa cell proteins. Neither was non-specific binding observed despite the use of two heterogeous samples. A difficulty in this study was screening for interactions between two potentially membrane bound proteins, the rPmp domain and putative urogenital cell surface receptors. On reflection, the experimental conditions, such as insufficient incubation time or buffer conditions may not have been suitable. (Kriechbaumer *et al.*, 2012a) demonstrated that native membrane protein activity was retained after deposition onto gold film, the diversity of these macromolecules means that conditions favoured by one species may not extend to another. In addition, this method would not detect interactions with secreted or cell surface shedding molecules that would possibly have been lost during preparation of the lysates, such as the interactions seen in the crossbinding assays. Immobilising whole cells using the Langmuir-Schaefer deposition technique (Kriechbaumer *et al.*, 2012b) may have improved the experimental design however, these methods were not available at the time. Endometrial Hec1B cells were not tested due to limited access to the equipment but it would be very useful to carry out these methods using the cells and truncated rPmps to confirm interactions found with ICC. Also a consideration for further work would be test all the truncated rPmps with both cell lines using whole cells to investigate interactions with native membranes now these methods are established.

## 5.5 Summary

In summary chemical cross-linking, immunocytochemistry, surface plasmon resonance and ellipsometry were all tested to some degree to screen or 'fish' for potential binding event using rPmp domains as ligands and urogenital cells as 'orphan receptors'. From the data generated from these studies it cannot be determined with confidence that rPmpI-N, rPmpDauto and rPmpGauto act as adhesins to host cells, neither can the data reject this theory. Interestingly, some binding interactions were observed between Hec1B and rPmp domains in ICC studies and for the crosslinking assays. These findings were not tested further with other methods. Similar interactions were not observed when testing HeLa cells using chemical cross-linking, ICC, SPR and ellipsometry. However, this was difficult to conclude as the methods applied in this study were not always working efficiently and the ICC data for the HeLa cells displayed a high background signal. Overall the methodologies were challenging due the heterogeneous nature of both the recombinant proteins and the putative receptors, which resulted in high background and non-specific binding events. Therefore within these heterogeneous binding interactions, the signal from one binding event may have been obstructed by the large non-specific interactions observed such as for the SPR study. Although Biacore was implemented as a very sensitive method to detect even small levels of binding interactions, it does not differentiate between specific and non-specific interactions (Nedelkov and Nelson, 2001). Consequently, relying exclusively on the subtraction of controls to eliminate background signals resulted in ambiguous interaction data. Furthermore, the NLP nanodiscs used in this study may have prohibited interactions if the rPmp binding sites were not accessible. However the incorporation of membrane proteins within nanodiscs is growing and has been successfully implemented in protein binding interactions, in particular for SPR analysis examining the binding of soluble proteins to immobilised membrane receptors in nanodisc complexes (Shaw *et al.,* 2007, Borch *et al.,* 2008) and for detection of drug binding to an enzyme reconstituted in immobilised nanodiscs (Das *et al.,* 2009). Ultimately as this was the first study of its kind in exploring the use of predicted functional rPmp domains as adhesins with the potential to interact with several putative receptors.

## 6.1 Polymorphic membrane proteins (Pmps) from clinical strains of serovar E C.trachomatis are challenging to express as full length recombinant proteins

To date, *C.trachomatis* Pmps have not been expressed as recombinant proteins at high enough levels for structural and functional investigations. The Pmps are predicted to be autoransporters composed of three major domains, the N-terminal signal peptide, the passenger domain and the C-terminal autotranslocator domain. The amino signal peptide region of autotransporters in many other species contain a consensus signal peptidase recognition site, SPase1 (c-domain) that is cleaved by components of the Sec machinery to release the pro-protein into the periplasm (Henderson *et al.,* 2004; Desvaux *et al.,* 2004a; Desvaux *et al.,* 2004c). The Chlamydial Pmps are thought to be processed this way but this is yet to be confirmed. Adjacent to the putative 'signal' domain is the functional 'passenger' domain. This region offers the most sequence diversity among autotransporter proteins as this performs the extracellular function. This passenger domain may remain intact in association with the cell surface or may be cleaved into the extracellular milieu to carry out its function (Henderson *et al.,* 2004, Desvaux *et al.,* 2004b). The β-barrel C-terminal domain forms the translocator pore embedded with the membrane to transport the functional passenger domain to external membrane surface. In addition to the speculative data that shows Pmps to be autotransport proteins with a β-barrel membrane domain, studies have suggested Pmps interact with the host cell to initiate infection, although little analysis has been carried out to determine which, if any Pmps function in this way with the exception of Pmp D (Wehrl *et al.,* 2004p; Crane *et al.,* 2006).

In order to investigate the function of the *C. trachomatis* Pmps it was necessary to express high quantities of pure recombinant proteins. Thus far, very few studies have expressed rPmps, instead the focus has been largely based on purifying native Pmps from elementary bodies isolated from infected mammalian cells yet these have not been used for structural or functional studies, probably due to the low abundance of these proteins within the membranes (Swanson *et al.,* 2009b). *E. coli* was the first choice as an expression host as it is a Gram-negative prokaryote like *Chlamydia* and was considered suitable due the presence of both inner and outer membranes. However, heterologous recombinant expression of membrane proteins poses unique challenges for several reasons. First, their over-expression can often be deleterious to the membrane and consequently toxic to the host organism

(Miroux and Walker, 1996). Second, correct folding of proteins within the membrane is rate-limiting for any expression system (Loll, 2003) and if this capacity is exceeded; inclusion bodies are likely to accumulate within the cell which was a common problem in this study. Finally, many membrane proteins, particularly eukaryotic proteins, undergo functionally important posttranslational modifications that cannot be replicated in heterologous expression systems and so can affect protein function (Sarramegna *et al.,* 2003). As anticipated the expression of recombinant Pmps was extremely challenging using *E. coli* as a heterologous expression host. Of the expression vectors constructed containing *rpmp* genes, full length rPmps were not expressed, yet a variety of *E. coli* strains and a range of optimisation techniques were employed.

It is possible that *E.coli* is insufficiently equipped to express full length Pmps due to their large size and complexity. The lack of published data that has explored the use of recombinant Pmps would support this theory. Where recombinant Pmps have been produced, they have been expressed in the form of inclusion bodies for the production of Pmp specific antibodies mainly raised against *C.pneumoniae* truncated rPmps or peptide sequences (Vandahl *et al.,* 2002; Niessner *et al.,* 2003; Pedersen *et al.,* 2006; Mölleken *et al.,* 2010a). A study that was successful in the production of *C.trachomatis* rPmps in *E.coli,* expressed them as inclusion bodies for the use of antibody generation. Crane et al., 2006 were able to produce rPmps but only in truncated form and lacking a signal peptide (**Figure 6.1**). Previous to this, a study had expressed a recombinant Pmp C as a fusion again without the hydrophobic amino-terminal signal leader sequence However, expression of rPmps without signal peptides were not attempted in this study as the signal peptide was considered important for targeting Pmps to the *E.coli* membranes where expression was desired for the extraction of native rPmps. It was thought that expression would most likely result in inclusion bodies without the signal peptide as seen in these published studies and was unsuitable for the purposes of structural and functional investigations.

Although not reported, collectively the difficulties in expressing full length rPmps compared to the success of expression of rPmps lacking signal peptides in other studies (Gomes *et al.,* 2005; Crane *et al.,* 2006) (although as inclusion bodies) may suggest that the Pmp amino-terminal signal peptides have a detrimental role upon the production of these proteins in *E.coli*. Generally signal peptides consist of a short, positively charged N-terminal region (*n*-region), a 7-15 residue hydrophobic

core (*h*-region), and a polar 3-7 residue C-terminal region (*c*-region) leading up to the signal peptidase cleavage site between the signal peptide and the mature protein (von Heijne and Abrahmsen 1989; Claros *et al.,* 1997). As demonstrated in early studies, there are defined differences between signal peptides between some species. Differences in *c*-regions may be as a result of different signal peptidase specificities and differences in the *n*- and *h*-regions could be due to differences in the lipid composition of membranes between species. For example *E. coli* membranes contain more of the zwitterionic phosphatidylethanolamine and less of the negatively charged cardiolipin than *Bacillus* and *Streptomyces* however (von Heijne and Abrahmsen, 1989). Therefore in the early part of this study it was thought the native Pmp signal peptides may be specific for expressing and targeting Pmps to constituents of the Chlamydiae membranes only and so could possess unrecognised codons resulting in early termination during transcription or translation. However, later findings in this study reject this notion as *E.coli* was used to express the amino-terminal passenger domain of Pmp I with native signal peptide and putative cleavage of the signal peptide was observed.

In some species autotransporter proteins function as toxins containing a toxic passenger domain. It is possible that some of the *C.trachomatis* Pmps may also serve this function such that the N-terminal transcriptional and/or translational products may have deleterious effects on *E.coli* leading to early termination of transcriptional machinery. This was a possibilty as the expression of C-terminal autotranslocator domains of Pmp D and Pmp G were successful. A similar observation was not reported by Crane *et al.*, (2006) during the expression of *C. trachomatis* rPmps.

| Recombinant Pmp (length) | Residues expressed | N-sig Passenger Translocator |
|---|---|---|
| | |  |
| rPmpA (975) | 47–979 |  |
| rPmpB (1751) | 949-1747 |  |
| rPmpC (1770) | 994-1766 |  |
| rPmpD1, rPmpD2, rPmpD3 (1531) | 921-1528, 45-1079, 45-1528 |  |
| rPmpE (964) | 27-962 |  |
| rPmpF (1034) | 28-449 |  |
| rPmpG (1013) | 139-502 |  |
| rPmpH (1016) | 337-1009 |  |
| rPmpI (876) | 30-879 |  |
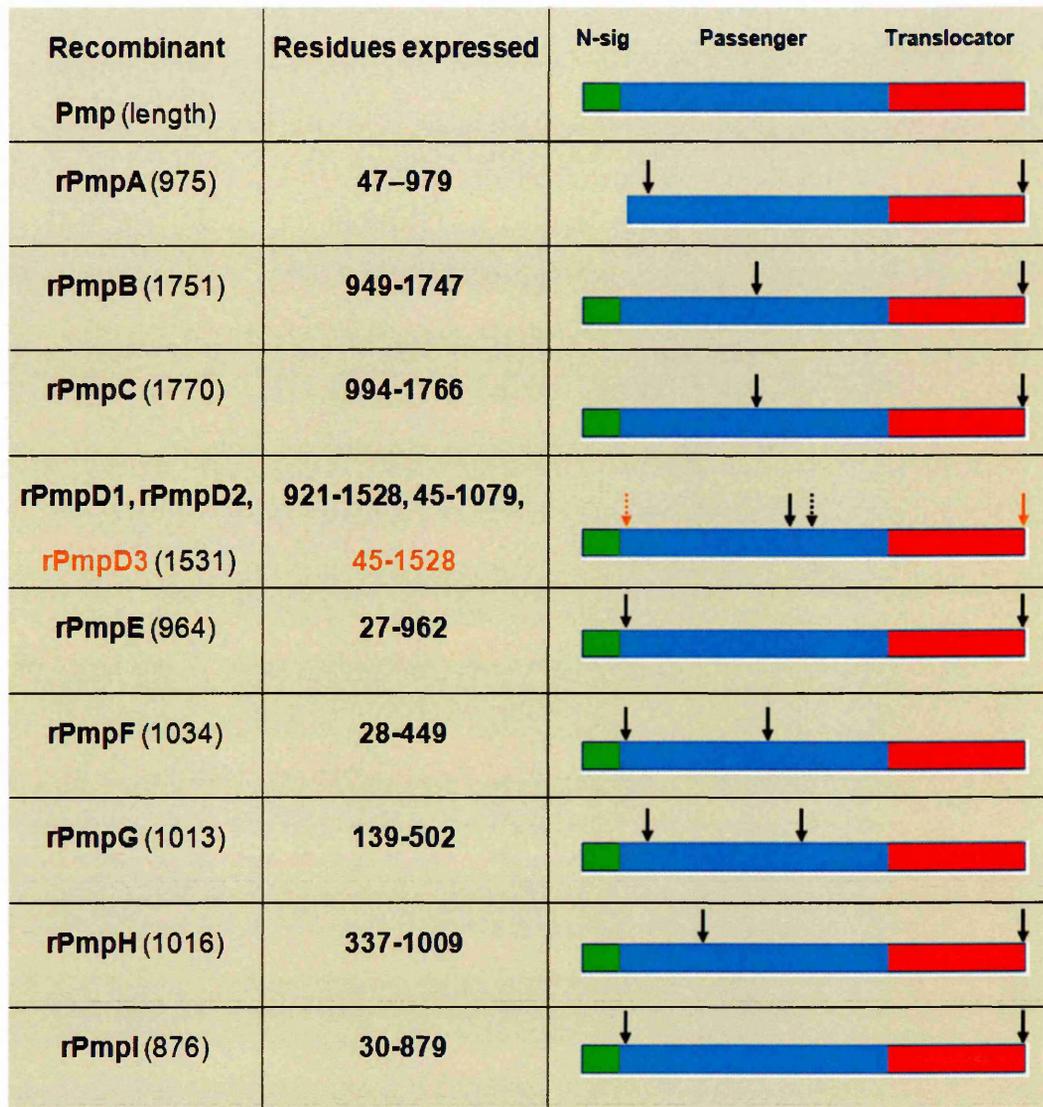
Figure 6.1 - Schematic diagram showing the recombinant Pmp regions derived from *C.trachomatis* expressed as inclusion bodies in other studies. Arrows depict the regions expressed by Crane *et al.*, 2006. All constructs were truncated and lacked the N-terminal signal peptide

Another consideration for the lack of Pmp-FL expression could result from discrepancies within the translation initiation regions of the recombinant expression vectors. Initiation of protein synthesis in *E.coli* requires several features of the mRNA to form a ribosome binding site (RBS) and direct ribosomes to the correct translational start site (Sprengart *et al.,* 1996). The elements of the RBS include the initiation codon (usually AUG or GUG) and upstream of this the Shine-Dalgarno

(SD) region (GAAGGAGA provided by the pET vector). The SD sequence is complementary to the 3' end of 16S rRNA and so the distance between the start codon and the SD is vital to allow interaction with the ribosome (Kozak 1999). The distance between a functional SD sequence and the start codon usually varies between three and 12 nucleotides, with an optimal spacing of 7-9 nucleotides (Ringquist *et al.,* 1992). The distance between the initiation codon of the inserted full length *pmp* genes and the SD of the pET vector were all within this range so it is unlikely ribosomal binding was affected. Furthermore trials using a cell-free expression host later expressed full length Pmp I showing that ribosome binding activity was functional for this expression construct.

Finally, there is evidence to suggest that the folding of autotransporters is both an autonomous process and/or requires chaperones. BrkA is a *Bordetella pertussis* virulence factor that is one of many *B. pertussis* adhesins. BrkA and is expressed as a preproprotein of 103 kDa that is subsequently processed into an N-terminal signal peptide, passenger domain and C-terminal autotranslocator domain (Oliver *et al.,* 2003a). Studies have shown a region in the C-terminus of the BrkA passenger domain is required for folding of this same domain by acting as a intramolecular chaperone (Oliver *et al.,* 2003b). This region is conserved in a large group of autotransporters signifying that it serves a key function but thus far has not been identified in the Pmp proteins. In contrast, periplasmic chaperones such as DsbA and FkpA were shown to influence the folding of heterologous N-terminal passenger domains used in a study to test the tolerance of the C-terminal pore of the autotransporter IgA1 protease (IgAP) from *Neisseria gonorrhoeae* (Veiga et al., 2004). Here it was demonstrated that passenger domains are transported through the C-terminal pore in a pre-folded state. However the passengers domains were made of single or multiple folded Ig domains as opposed to the natural IgAP N-terminal secretory protease and therefore may not be representative of all autotransporter periplasmic folding mechanisms. Overall, studies are contradictory with regards to autotransporter folding mechanisms and it seems one rule does not apply to all and this area of protein folding needs further investigation. Essentially the folding mechanisms of Pmps remain unclear, it is possible that either cytoplasmic and/or periplasmic chaperones have important roles during the synthesis and folding of these proteins *in vivo.* The co-expression of full length rPmps with chaperones in *E. coli* may have been necessary in order to produce full

length Pmps, however there remains no published data to support which, if any chaperones are involved.

One recent study has determined the first ever crystal structure of a complete autotransporter expressed in *E. coli*, the esterase EstA from the opportunistic human pathogen *Pseudomonas aeruginosa* (van den Berg 2010). This is remarkable as there has been a void of information prior to this highlighting the difficulties associated with producing such large complex proteins. With such information emerging, successful expression of full length Pmps could pave the way for successful structural and functional Pmp studies.

## 6.2 Truncated Polymorphic membrane proteins (Pmps) from clinical strains of serovar E C.trachomatis are less challenging to express using a novel cell-free expression system with nanolipoprotein discs

Although not reported, collectively the lack of full length rPmp expression in these investigations and evidently the lack of native rPmp expression in previous studies (Gomes *et al.*, 2005; Crane *et al.*, 2006) would suggest that *E.coli* is an unsuitable host for the production of *C.trachomatis* full length Pmps possibly due to a combination of factors previously discussed. Alternative methods were therefore considered to express Pmps for secondary and functional analysis.

As it is widely accepted that Pmps are autotransporters, made up of three distinct domains, expression of these smaller functional domains of some of the *C.trachomatis* Pmps was explored. Other autotransporter proteins have been studied this way with success; in particular the IgAP protease from *Neisseria gonorrhoeae*, an organism that also infects the genital epithelial cells has had its autotranslocator domain structure determined by expressing the C-terminal functional region to the outer membranes of *E.coli* (Veiga *et al.*, 2002). Oomen *et al.*, (2004) also presented the crystal structure of the translocator domain of the autotransporter NalP from Neisseria meningitidis, again from over-expression in *E.coli*. In addition the N-terminal passenger domain of the Hia autotransporter from *Haemophilus influenza* has been expressed for the use in adherence assays (Laarmann *et al.*, 2002). Furthermore, previous studies have expressed truncated rPmps.

In this study, expression trials to produce such regions of *C.trachomatis* Pmps led to the successful expression of several truncated rPmp functional domains in *E.coli* although mostly in the form of inclusion bodies. This is unsurprising as inclusion

296

bodies are a common occurrence when expressing proteins which are hydrophobic or membrane bound (Wagner *et al.*, 2006). Inclusion bodies can be purified with relative ease for *in-vitro* refolding but can lead to conformational changes different from the native structure. This is especially problematic where disulfide bonds are present and since Pmps are reported to be rich in cysteine residues (Swanson *et al.*, 2009; Liu *et al.*, 2010) this was deemed impractical. Furthermore, refolding often requires extensive optimisation with the use of a control assay to demonstrate that the folded protein is functional, which was not possible in this study. Instead an alternative approach pursued the expression of natively folded, active protein suitable for structural and functional analysis by reducing protein aggregation to such a level that allowed truncated rPmp to localise to the membranes.

By reducing the expression rates significantly (Sørensen and Mortensen, 2005), a reduction in aggregation of the rPmp domains, in particular the translocator domain of Pmp D (rPmpDauto) were met with some success. Tight regulation using the pLysS plasmid, (that encodes T7 lysozyme), IPTG independent expression and low incubation temperature allowed small amounts of rPmpD autotranslocator domain to localise to the membranes of *E.coli* for detergent solubilisation. However the purification of this domain by nickel chelate affinity chromatography was met with great difficulty due to the low binding to the column, a poor yield of expression and a high tendency for the protein to precipitate and aggregate during handling and storage. It is possible that the C-terminal polyhistidine tag was not fully exposed for binding (Bornhorst and Falke, 2000) but denaturing the protein to unfold the His-tag fully was not viable as a biologically active protein sample was necessary. Collectively the purity, yield and tendency to aggregate made this protein domain unsuitable for structural studies. However this is the first study of its kind to express the *C.trachomatis* Pmp D translocator domain both as inclusion bodies and as membrane localised protein with a 6XHistidine affinity tag.

Other methods described to aid localisation to the membrane were by expressing target protein as a fusion with a protein known to localise to the membranes for example the maltose-binding protein (Fox *et al.*, 2008) or as a fusion with an N-terminal signal peptide leader sequence. This was considered a useful exercise as the autotranslocator domains of the Pmps lack a native signal peptide as this is normally located upstream at the amino-terminal so is lost in these truncated C-terminal Pmp constructs. Pmp G was expressed as a fusion with a pelB leader peptide to aid localisation to the inner membrane of *E.coli*, an approach also used to

express the C-terminal autotranslocator domain of IgAP (Veiga *et al.,* 2002b). Although expression was successful this was not localised to the membranes and was in the form of inclusion bodies.

In addition to the production of the C-terminal autotranslocator domains of the Pmp D and Pmp G, expression of the functional N-terminal passenger domains was also explored, in particular those of Pmp A and Pmp I. The N-terminal region of Pmp I contains a native signal peptide leader sequence predicted to target proteins to the Sec apparatus in the inner membrane for processing (Desvaux *et al.,* 2004a). The expression of the recombinant passenger domain of Pmp I in *E.coli* was detected as both soluble and insoluble protein. This fits with other data that have observed non membrane bound passenger domains of autotransporters (Ieva and Bernstein 2009). Some passenger domains, including that of Pmp D (Kiselev *et al.,* 2007), are often presented externally of the cell surface possibly necessitating less hydrophobic regions compared to their C-terminal membrane-bound counterparts. At present there are no repeated studies that comment on the processing and localisation of Pmp I or the Pmp I passenger domain to corroborate the finding that the N-term of this protein is soluble in nature. Interestingly, the passenger domain of Pmp A, which differs from the other Pmps by absence of a native signal peptide (Gomes *et al.,* 2006), was not expressed by *E.coli*. This protein has previously been expressed as inclusion bodies in *E.coli* by Crane *et al.,* (2006) however their recombinant construct lacked the first 46 residues of the amino terminus which may be a significant deletion to ensure successful expression.

In summary, the expression of recombinant full length Pmps and truncated Pmps in *E.coli* was extremely challenging. *E.coli* could not be used to express the large multidomain full length Pmps and where expression was observed, for the truncated domains, the expression was very 'leaky' resulting in inclusion bodies. Conditions implemented to modulate this, in particular to minimise basal expression and prevent inclusion body formation were met with varying degrees of success. They were successful in that the formation of inclusion bodies could be reduced to allow some protein targeting to the membranes where potentially, folding could take place. However, this process of optimisation was very time consuming and so only extended to the expression of Pmp D autotranslocator domain (rPmpDauto). The downstream purification that followed was difficult due the heterogeneity of the sample and a low affinity between the Histidine-tag and purification media. Moreover the poor yield of recombinant protein solubilised from the membranes was

increasingly difficult to detect following each downstream application. This was further hampered by the necessity to express fresh recombinant proteins in large culture volumes for each individual optimisation step as the rPmpDauto aggregated upon storage.

While *pmp* genes are prokaryotic, Chlamydiae do reside within a eukaryotic host so it was considered that mammalian expression systems may be suitable to overcome the problems associated with expression of rPmps in *E.coli*. Furthermore, prokaryotic cells are not equipped to carry out posttranslational modification where eukaryotic systems can. This is particularly important for proteins that are glycosylated as this can affect protein activity or stability. Although there is no evidence to suggest *C.trachomatis* Pmps are glycosylated, studies have shown Pmp homologues of *C.abortus* are glycosylated (Vretou *et al.,* 2001). Few studies have used yeast as a host for the expression of Pmps. One study that did, expressed passenger domains of some of the *C.pnuemoniae* Pmps as fusion proteins to the yeast surface protein Aga2p to direct the Pmp domain to the surface for use in adherence assays and showed the passenger domains of Pmp D, B/C and G homologues act as adhesins (Mölleken *et al.,* 2010a) yet this study did not express full length Pmps. Ultimately considering the time involved in establishing these techniques, developing new cloning strategies and repeating the time-consuming expression trials already carried out in *E.coli* for seven Pmp targets, cloning for an alternative expression system was considered to be an uneconomical use of time. A reasonable alternative to designing and constructing new expression systems was to change the expression host to one also compatible with the T7-based expression system. As cloning strategies had already led to the successful expression of several smaller rPmps domains in *E.coli*, an *E.coli* based cell-free expression system with a composition less complex than viable *E.coli* was chosen for its rapid screening ability, toxicity tolerance, and ease of handling.

The autotranslocator domains of Pmp D and Pmp G (rPmpDauto, rPmpGauto) plus the passenger domain of Pmp I (rPmpI-N) were successfully expressed as soluble protein using nanolipoprotein scaffolds (NLPs or nanodiscs) as part of a novel cell-free expression system (MembraneMax™, Invitrogen). Whilst the production of recombinant protein was consistent using this method, purification was difficult as the yield of recombinant protein was low and not sufficient for any biophysical analysis. Moreover this was a very expensive protein expression technique where again the proteins were highly susceptible to aggregation during handling and so

required newly expressed proteins for every optimisation step. This was possibly due to the use of detergents necessary during purification. Although nanodiscs are reputed to offer a more stable environment than detergent micelles (Nath *et al.,* 2007) the original aim of this study was to obtain biophysical data by circular dichroism and the scaffold protein of the nanodiscs would have interfered with this analysis. Interestingly the rPmp A passenger domain, the only Pmp reputed to lack a native signal peptide sequence was expressed in the cell-free system but not in *E. coli*. However this passenger domain was not incorporated into NLPs and was only present as insoluble aggregates, a possible indication that Pmp A is targeted to the membrane by another mechanism, possibly by use of a C-terminal anchor region, also suggested by Gomes *et al.*, (2006) which was not present in this expression construct as it lacked the C-terminus region.

The cell-free system was also shown to express one of the full length Pmps, recombinant Pmp I (rPmpl-FL), in an insoluble form. It is likely that rPmpl-FL is too large at approximately 95 kDa, to be incorporated in to the nanolipoproteins (NLPs) (according to the manufacturer). It is likely that the system may not provide an environment suitable for full length Pmp protein folding, for example if chaperones are essential. The MembraneMax™ cell-free system used has been used successfully for membrane proteins up to 100 kDa. Very large membrane proteins over 100 kDa have not been reported, or have been expressed as truncated products, for example the human potassium voltage-gated channel ERG protein (~155 kDa) could not be expressed as a full length protein (Katzen *et al.,* 2008). Although full length rPmp I was insoluble, this was an improvement on the expression trials previous using *E.coli* where the expression of full length rPmps was not observed, and was also an indication that the expression vector containing full length *pmp I* was indeed functional. The larger rPmps (rPmpA-FL and rPmpD-FL) that are over 100 kDa at 105 kDa and 156 kDa were not expressed by the cell-free system concurring with previous findings. This is the first study to explore the expression of rPmps and predicted rPmp domains using nanodiscs and a cell-free expression system. Importantly, this is the first study to report the expression of soluble rPmps.

In summary, it is widely accepted that expression of recombinant membrane proteins is a difficult process. Even in the most successful cases, the over-expression and/or purification protocols are often unique for each target, being both technically challenging and extremely expensive (Dobrovetsky *et al.,* 2005).

Furthermore, few recombinant membrane proteins are expressed in native, full-length and this is reflected in the number of membrane proteins that have had their structure determined successfully compared to their soluble counterparts. Similar difficulties were observed in this study which set out to express recombinant Pmps for structural and functional purposes and are an example of the high attrition rates common with membrane protein expression and purification. This is summarised in **Figure 6.2**, where fourteen expression constructs resulted in only five expressed proteins and was narrowed down further to just three proteins suitable for purification. At this final stage, the purified products were then not suited to a basic biophysical analysis such as circular dichroism, therefore secondary structure of these Pmp domains could not be obtained. High attrition rates are evident in other studies. Dobrovetsky *et al.*, (2005) investigated the expression and purification of numerous membrane proteins from *E.coli* and *T. Maritime,* a thermophilic eubacterium, of the total number of genes cloned, only ~8% (*n* 17/203) and ~6% (*n* 5/77) respectively, of the represented proteins were obtained in quantities sufficient for structural studies.

Although expensive, the cell-free system proved to be an advantageous method over *E.coli* for the expression for rPmp domains. Firstly, handling times were reduced significantly; an important factor where protein precipitation is a problem, expression was achieved within two hours and solubilised protein was acquired with the use of nanolipoprotein discs. Although structural analysis was unattainable for the recombinant Pmps domains due to low yields after purification, functional screening assays where purified protein is not a requirement were still explored. Most importantly, as soluble protein was produced for two Pmp autotranslocator domains (Pmp D and Pmp G) and one Pmp passenger domain (Pmp I), it was considered that the domains would be in a native confirmation ideal for protein interaction studies (Katzen*, et al.* 2008; Bayburt and Sligar, 2010), therefore investigation of these Pmp domains with the use of nanodisc technology was explored further to screen for interactions with host cells and host cell membranes.
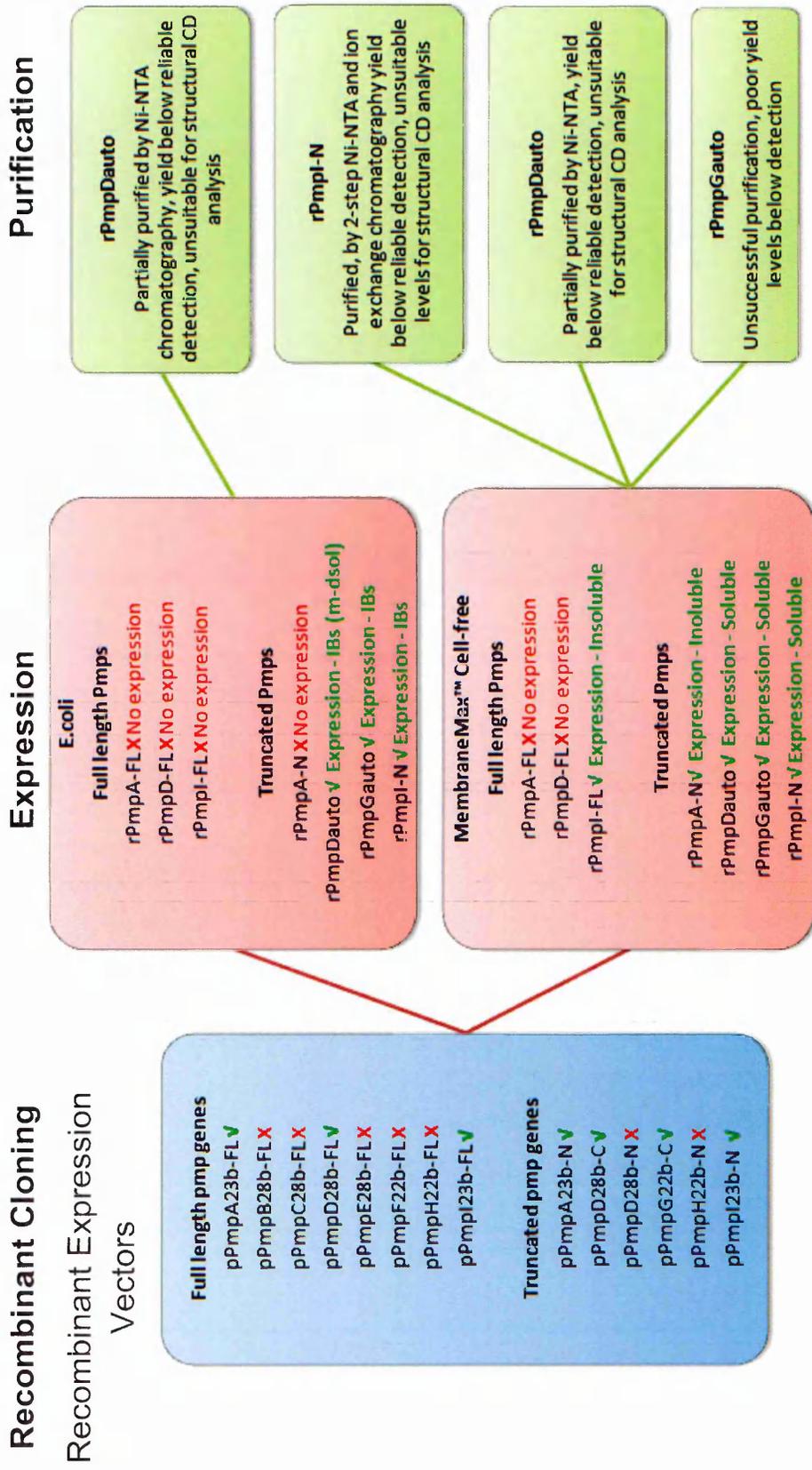
**Figure 6.2 - Summary of cloning, expression and purification of rPmps produced and used in these studies.** IBs, Inclusion bodies, m-dsol, membrane bound detergent solubilised, CD, Circular Dichroism

## 6.3 C-terminal autotranslocator domains of Pmp D, Pmp G and the N-terminal passenger domain of Pmp I as adhesins for Chlamydia trachomatis infection

The importance of studying membrane proteins is highlighted by the fact that almost half of the major pharmaceutical drug targets are designed to target membrane receptors (Cooper, 2004). However, monitoring protein interactions in lipid environments is also technically challenging. Existing biochemical and biophysical methods to study membrane protein interactions require either sufficient native protein or recombinant protein into a lipid or detergent environment to maintain biological activity. In one study, recombinant OmcB from *C. trachomatis* was expressed and targeted to the membranes of *E. coli* and *E.coli* displaying the protein on its surface were used in adherence assays and showed Chlamydial OmcB acts as an adhesin (Fadel and Eley, 2008). However, as demonstrated for the expression of rPmpD autotranslocator domain in this study, directing the expression of membrane proteins to a heterologous membrane is not always a straightforward process, especially where a high tendency to form thermodynamically stable inclusion bodies is also favourable (Foit *et al.*, 2009). Other methods require the membrane protein to be reconstituted into detergent micelles or liposomes yet these techniques are not always suitable as liposomes are large, unstable and difficult to prepare at a controlled size and stoichiometry and detergents can decrease protein stability and interfere with assays. Hence few techniques are available to study specific membrane protein interactions *in situ*.

Over the past decade, reconstituted high-density lipoprotein particles (rHDL) have emerged for the incorporation of membrane proteins in a native-like membrane environment, allowing solubilisation and delivery of hydrophobic drugs (Civjan *et al.*, 2003; Leitz *et al.*, 2006; Whorton *et al.*, 2007; Mi *et al.*, 2008; Bayburt and Sligar, 2009). In this system, membrane proteins in detergent micelles are reconstituted into lipid bilayers of apolipoprotein particles. These particles are homogenous and monodispersed with membrane proteins incorporated in monomeric form. This system has been tested in studies with membrane proteins such as bacteriorhodopsin (Bayburt and Sligar, 2009), cholera toxin and receptor binding (Borch *et al.*, 2008) the pore of the anthrax toxin (Katayama *et al.*, 2010) amongst many other applications making nanodiscs, similar to those used in this study (MembraneMax™ NLPs) suitable to investigate potential receptor-ligand interactions between rPmps and urogenital cells.

Prominent traits of the Pmps shared among the species are the multiple sequence repeats GGAI and FXXN (Stephens *et al.,* 1998), similar motifs are found in several proteins believed to be involved in attachment with eukaryotic cells including OmpA of *Rickettsia* spp. (Grimwood and Stephens, 1999) and uncharacterised surface proteins of *Trichomonas Vaginalis*, another bacterium associated with infection of the urogenital system (Hirt *et al.,* 2007). Chlamydiae have been shown to demonstrate various mechanisms in which attachment to host cells is established but so far the complexity of these mechanisms have not been fully elucidated however the Pmps have been identified as a possible ligand, functioning as an adhesin to bind and interact with host cells for this purpose (Wehrl *et al.,* 2004; Mölleken *et al.,* 2010b). The possibility that Pmps act as adhesins were tested in this study with the use of recombinant truncated Pmps: Pmp I passenger domain (rPmpI-N), Pmp D and Pmp G autotranslocator domains (rPmpDauto and rPmpGauto), expressed using a novel cell-free expression system with nanodiscs.

Truncated rPmp protein-cell interaction assays were problematic and encountered several difficulties due to the heterogeneity of the both the truncated rPmp samples and the orphan receptors (whole or lysed urogenital cells). It is possible that rPmp interactions have been shrouded by some of the high background signals observed using the experimental designs applied in this study. However several techniques were employed to try and overcome such problems with methodologies including cross-linking assays, immunocytochemistry, surface plasmon resonance and ellipsometry to screen for interactions with HeLa cervical and Hec1B endometrial cell lines, the assays and findings are summarised in **Figure 6.3** and **Figure 6.4**.

Collectively the observed protein interactions were often contradictory between methods but a concerted effort was made to try and rule out false positive results from background signals. Cross-linking rPmpI-N with HeLa or Hec1B cells resulted in high mass conjugates containing rPmpI-N suggestive of protein interactions, yet these were too large to be examined by Western blotting techniques and would not migrate into a polyacrylamide gel making it difficult to determine the size or identify the components of the interaction where methods such as mass spectrometry could have been applied. Additionally small traces of rPmpI-N were associated with the cell surface of Hec1B cells, potentially this interaction could have involved large molecules such as heparan sulphate, also a target of Chlamydial MOMP attachment (Su *et al.,* 1996), Heparan sulphate can also be shed from the cell surface (Bernfield *et al.,* 1999) large conjugates that could not be characterised but were separated

from the urogenital cells, could have contained such secretory/shedding molecules. However studies have shown that serovar E attachment is not affected by heparin or heparan sulphate (GAGs) and so this may not be involved in the attachment mechanisms of this particular serovar, instead these GAGs dramatically reduced the attachment of more invasive serovar L2 (Davis and Wyrick, 1997, Vretou et al., 2001). There was little evidence to show rPmpD and rPmpG autotranslocator domains had interacted with the HeLa or Hec1B cell surfaces, however if interaction had occurred these levels may have been below the limits of detection by Western blot as rPmpGauto and rPmpDauto were difficult due to the low expression levels of these domains.

An alternative method to 'fix' and detect cell surface interactions was carried out using immunocytochemistry (ICC). ICC methods showed putative localisation of the passenger domain of rPmpI to the HeLa cell surfaces when washing steps to remove unbound rPmpI-N were omitted, however localisation was not observed for any of the rPmp domains when the HeLa cells were washed indicating possible weak interactions and so more sensitive methods to capture weak, transient interactions were employed. SPR detected binding events between rPmpI-N and lysed HeLa cells, but this could not distinguish between specific binding of the rPmpI-N domain and non-specific binding where endogenous proteins within the expression sample also bound to HeLa cell lysate and so was not extended to test rPmpDauto and rPmpGauto. Finally, Ellipsometry, a method that monitors changes in molecular thickness caused by binding events was used to capture potential weak interactions (Kriechbaumer et al., 2012b) yet this technique revealed no association between lysed HeLa cells and either rPmpI-N, rPmpDauto or rPmpGauto domains.

Differences between the standard protocols for methodologies may have resulted in the contradictory findings. For example the cross-linking and ICC methods were carried out using whole HeLa cells, at 37 °C to mimic in vivo chlamydial attachment (Söderlund and Kihlström, 1983; Söderlund and Kihlström, 1988; Su et al., 1996; Su et al., 1996; Fadel and Eley 2008), with an incubation time of at least one hour. In contrast, SPR and ellipsometry methods required lower incubation temperatures at 25 °C, lysed HeLa cells and shorter incubation times of around ten minutes which may not have been optimal conditions to allow interactions to take place.It is reported that temperatures as low as 4 °C have not prevented adherence of C.trachomatis in previous studies (Zaretzky et al., 1995; Davis and Wyrick, 1997)

and so it could be argued that lower temperatures should not have prevented interactions,. The passenger domains of Pmp D and Pmp G homologues from *C.pneumoniae* have been suggested to carry adherent mechanisms (Mölleken *et al.,* 2010a) at 4 °C and 37 °C, although these were presented at the surface of yeast cells or latex beads and the study did not examine interactions of the C-terminal autotranslocator regions of these Pmps.

Interestingly, immunocytochemistry indicated cell surface localisation for all three rPmp domains when incubated with Hec1B endometrial cells. Furthermore, traces of rPmpI-N were also found associated with the Hec1B cells in the cross-linking assays. These results were promising to indicate *C.trachomatis* serovar E Pmps may have a role in tissue tropism, having a higher affinity for endometrial cells. Conversely, previous studies have shown that the binding of serovar E *C.trachomatis* elementary bodies (EBs) was comparable among Hec1B and HeLa cells (Davis and Wyrick, 1997). Yet those studies were not examining the particular binding partners responsible and so using EBs with an array of outer membrane proteins, could have resulted in binding from Pmps and other outer membrane proteins present on the EB surface such as MOMP and OmcB, also involved in attachment (Campbell and Kuo, 2006; Fadel and Eley, 2008; Moelleken and Hegemann, 2008). Furthermore, this may also show that the autotranslocator domains (rPmpDauto and rPmpGauto) may not only function as translocator proteins but could also interact directly with host cells.

### 6.3.1 Pmps as cleaved autotransport proteins

The methodology and experimental designs in this study were based around the notion that Pmps function with an autotransporter mechanism, however any study is yet to confirm Pmps are in fact autotransporters. Several of the *pmp* alleles encode a lipoprotein signal sequence which is not compatible with autotransport yet it remains unclear whether the lipoprotein sequences are ever functional (Henderson and Lam, 2001a). The findings in this study may suggest the predicted signal peptide of Pmp I is functionally active and the putative Pmp I passenger domain undergoes post-translational processing when expressed in *E. coli.* Similar investigations have shown the cleavage of the signal peptide of autotransporter adhesion: BrkA, a *Bordetella pertussis* virulence factor which is expressed as a preproprotein (Oliver *et al.,* 2003a). Recombinant expression of the predicted amino passenger domain of Pmp I with its native signal peptide observed putative

cleavage events within the passenger region. Autotransporter passenger regions are directed to the membrane under direction of the signal peptide where they can remain surface exposed or undergo proteolytic cleavage to release peptide fragments into the extracellular environment, via cell surface proteases or autoproteolysis (Henderson *et al.,* 1998; Henderson *et al.,* 2004; Desvaux *et al.,* 2004a). Three His-tagged peptides were detected during expression of the amino-passenger domain of Pmp I in *E.coli* that were suggestive of protein processing possibly upon interaction with sec machinery. The sizes of these fragments were indicative of this passenger domain with and without the N-terminal signal peptide and of cleavage within the passenger domain itself. Furthermore, trials inhibiting secA of the sec translocon in *E. coli* showed an increase in the formation of inclusion bodies during the expression of the Pmp I passenger domain possibly showing restriction to within the cytoplasm. Without a native signal sequence so that rPmpI-N would not be targeted to the sec translocon, the expression of the passenger domain of Pmp I appeared to be restricted and confined within the cytoplasm forming large inclusion bodies. These studies may be an early indication to show the importance of the signal sequence for the targeting of Pmp I and processing of the protein to produce mature Pmp I. Expression of the same domain using the *E.coli* derived cell-free expression system resulted in only two His-tagged peptides during expression, again sizes expected for the domain with and without the signal sequence. Exposing this domain to higher ionic strength buffers and increased pH during ion exchange purification, a third peptide of the same molecular weight previously observed during *E.coli* expression, was detected. Published work investigating the processing of Pmps is substantially lacking, however a study by Swanson *et al.,* (2009) also found similar results after probing purified native Pmp D from *C.trachomatis* extracts, several processed fragments derived from both the carboxy-terminal autotranslocator domain and the amino-passenger domain of Pmp D were found. Therefore it is tempting to speculate that the mature form of the Pmp I passenger domain (without signal peptide) undergoes autoproteolytic cleavage, possibly as a result of environmental factors or changes in confirmation due to surface charge.
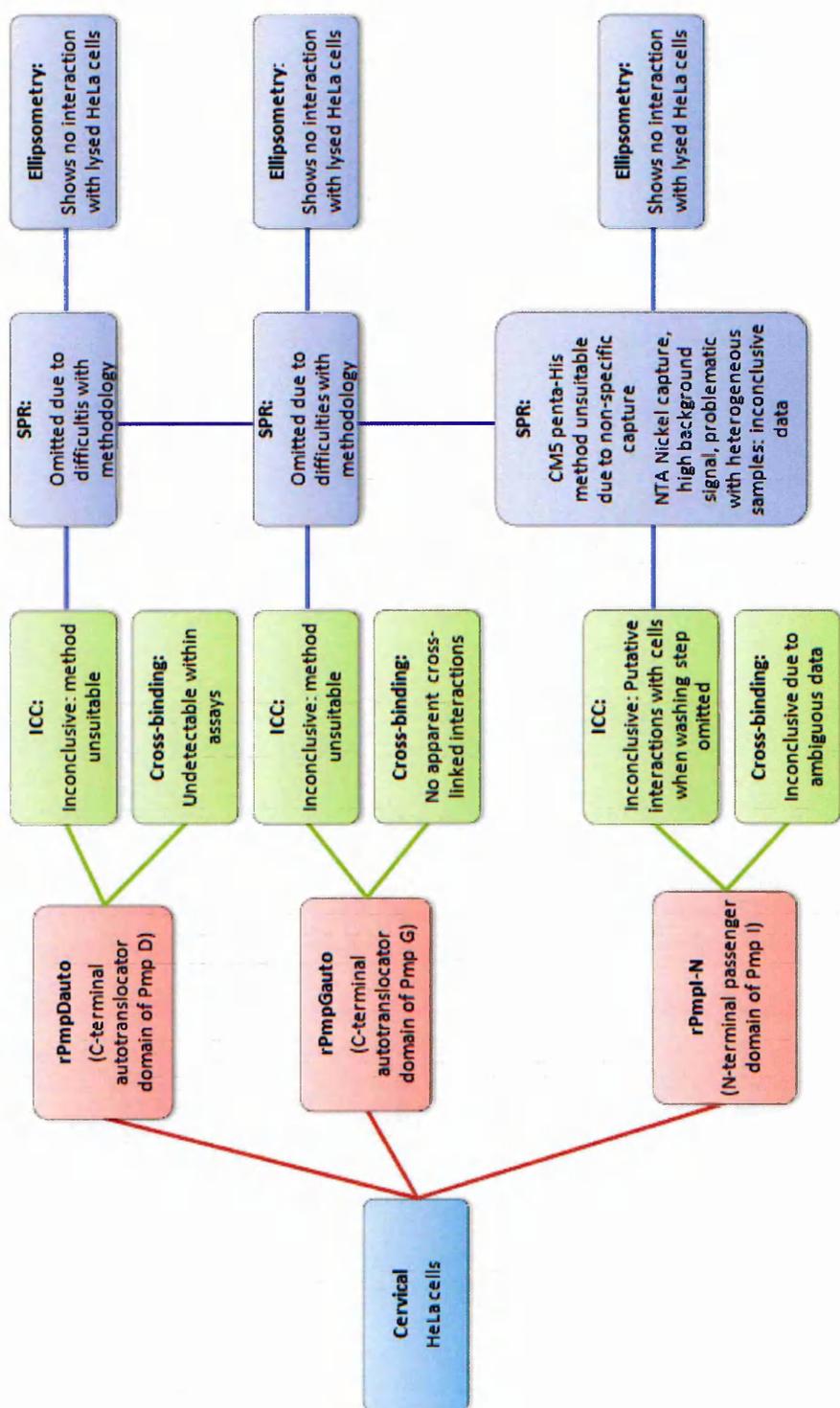
**Figure 6.3 - Summary of findings investigating truncated rPmps (pink) in interaction studies with cervical cells by immunocytochemistry and chemical cross-binding assays.** Assays were followed up with Surface plasmon resonance and ellipsometry. ICC, Immunocytochemistry; SPR, Surface plasmon resonance.

308
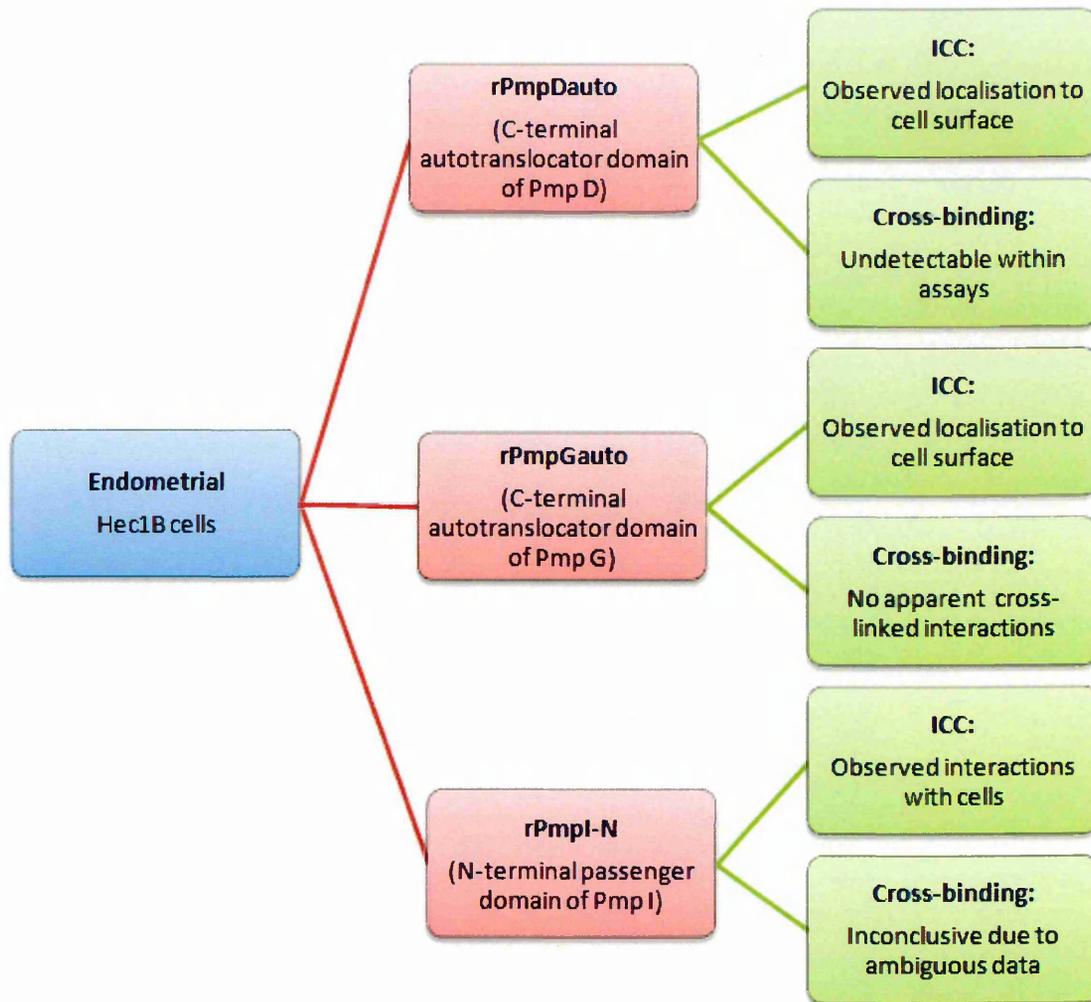
**Figure 6.4 - Summary of findings investigating truncated rPmps (pink) in interaction studies with endometrial cells by immunocytochemistry and chemical cross-binding assays**. ICC, Immunocytochemistry

## 6.4 Future work

Some of the major challenges in this study could possibly have been alleviated by reducing the dependency upon the Histidine affinity tag. For example, capture of the recombinant Pmp domains was reliant upon exposure of the Histidine tag for Ni-NTA purification and SPR Ni-NTA interaction studies. Additionally the expression and purification, cross-binding interactions, ICC investigations and alternative SPR capture techniques were all dependent upon anti-His antibodies binding to the rPmp polyhistidine affinity tags. Moreover, the Histidine tag could have had some affect upon the folding of rPmps. As Pmps are unique to Chlamydia, the production of anti-Pmp antibodies to specific rPmp peptides may have overcome some of these problems by minimising non-specific binding to endogenous proteins, within the

expression samples and eukaryotic cells and by overcoming problems with accessibility of the His-tags. Specific anti-Pmp Abs raised to the rPmp inclusion bodies could replace generic anti-His Abs used throughout these studies for detection of interactions with cells by ICC, for immunoprecipitation with rPmps and orphan receptors, flow cytometry and with direct coupling to a sensor surface for capture of rPmps in SPR. This is potentially an area for future work to proceed with the studies investigating the function and processing of the rPmps produced in this study.

Investigating the processing of Pmps to determine if Pmp I is processed via the inner membrane Sec translocon into a mature protein from its amino-terminal signal peptide would offer clues as to the function of Pmp I as a secretory molecule. If sufficient protein could be obtained for either method, identification of the cleavage site could be achieved by Edman sequencing of the mature passenger domain or using N-terminal peptide sequencing by mass spectrometry following separation of cleavage products. In addition mutagenesis studies could be carried out on the newly identified signal peptidase cleavage site to determine if the targeting of this protein is under direction of the native leader sequence. Preliminary work using sodium azide, which inhibits SecA, an ATPase involved in the translocation of proteins via the Sec machinery (Oliver *et al.,* 1990; Oliver, 1993), has suggested Pmp I is directed to the Sec translocon for N-terminal processing, however this work is by no means conclusive as azide can interrupt other cellular pathways. However, studies carried out expressing the passenger domain of rPmpI without its native peptide, resulted in increased accumulation within the cells, as expected indicating confinement of this expressed domain when not under directional processing. Azide resistant mutants have already been obtained from the Coli Genetic Stock Center (Yale University) but are not suited to expressing target proteins from pET expression vectors. However, integration of λDE3 prophage into this mutant *E. coli* host cell chromosome can be carried out, so that the azide resistant host containing the secA protein resistant to inhibition by azide can be used alongside an azide sensitive host to express rPmpI passenger domain. The effects of azide on protein accumulation could be directly related to secA inhibition or not and would show if Pmp I is processed via Sec processing as predicted.

In addition to the putative cleavage of the signal peptide, other cleaved peptides were observed during the expression of Pmp I passenger domain, potentially from autoproteolytic cleavage within the passenger domain itself (Henderson *et al.,* 1998,

310

Henderson and Nataro 2001b, Desvaux *et al.,* 2004a). Confirmation of this cleavage site is of interest as it may show Pmp I is cleaved and released into the external milieu whilst other autotransporters remain intact at the cell surface (Henderson *et al.,* 1998; Desvaux *et al.,* 2004a; Dé *et al.,* 2008; Benz and Schmidt, 2011). Problems in identifying these cleaved products on a Coomassie SDS gel were as a result of low yields after purification and hampered efforts to proceed with peptide sequencing. Future work to improve these yields could possibly be achieved by using cobalt ions in metal affinity chromatography. The $Co^{2+}$-carboxylmethylaspartate (Co2+-CMA) matrix (Talon™ resin) has been reported to exhibit less nonspecific protein binding than the Ni-NTA resin, resulting in higher elution product purity and so may have overcome some of the problems associated with contaminants difficult to remove via Ni-NTA and ion exchange purification (Chaga *et al.,* 1999a; Chaga *et al.,* 1999b). Subsequently, these peptides could be gel excised for N-terminal peptide sequencing by mass spectrometry to identify the putative cleavage sites and mutagesis studies on these sites could be carried out.

To corroborate the ICC findings with Hec1B cells and rPmpI-N, rPmpDauto and rPmpGauto, where putative localisation at the cell surface was observed, indicating tissue tropism for endometrial cells, ellipsometry should be carried out. This sensitive method requires no labelling of samples, instead changes in the reflected light indicate a change to the molecular layering. As the ICC showed cell surface staining, ellipsometry should be done using whole Hec1B cells via Langmuir-Schaefer film deposition (Kriechbaumer *et al.,* 2012b) as oppose to lysed cells to ensure the membrane outer surface is exposed for binding. Based on these findings, assays could be repeated with whole HeLa cells as previously used methods may have resulted in an immobilised surface where HeLa membranes/membrane proteins were unavailable for binding.

AbdelRahman, Y. M. and Belland, R. J. (2006). The chlamydial developmental cycle. *FEMS Microbiology Reviews,* 29 (5), 949-959.

Ami, D., Natalello, A., Gatti-Lafranconi, P., Lotti, M. and Doglia, S. M. (2005). Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Letters,* 579 (16), 3433-3436.

Anton van der Merwe, P. and Neil Barclay, A. (1994). Transient intercellular adhesion: The importance of weak protein-protein interactions. *Trends in Biochemical Sciences,* 19 (9), 354-358.

Avenaud, P., Castroviejo, M., Claret, S., Rosenbaum, J., Mégraud, F. and Ménard, A. (2004). Expression and activity of the cytolethal distending toxin of< i> helicobacter hepaticus</i>. *Biochemical and Biophysical Research Communications,* 318 (3), 739-745.

Baneyx, F. and Mujacic, M. (2004). Recombinant protein folding and misfolding in escherichia coli. *Nature Biotechnology,* 22 (11), 1399-1408.

Bao, N., Jagadeesan, B., Bhunia, A. K., Yao, Y. and Lu, C. (2008). Quantification of bacterial cells based on autofluorescence on a microfluidic platform. *Journal of Chromatography A,* 1181 (1), 153-158.

Baranek, C., Sock, E. and Wegner, M. (2005). The POU protein oct-6 is a nucleocytoplasmic shuttling protein. *Nucleic Acids Research,* 33 (19), 6277-6286.

Batteiger, B. E., Rank, R. G., Bavoil, P. M. and Soderberg, L. S. F. (1993). Partial protection against genital reinfection by immunization of guinea-pigs with isolated outer-membrane proteins of the chlamydial agent of guinea-pig inclusion conjunctivitis. *Journal of General Microbiology,* 139 (12), 2965-2972.

Baumann, S., Grob, P., Stuart, F., Pertlik, D., Ackermann, M. and Suter, M. (1998). Indirect immobilization of recombinant proteins to a solid phase using the albumin binding domain of streptococcal protein G and immobilized albumin. *Journal of Immunological Methods,* 221 (1), 95-106.

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), 479-485.

Bavoil, P. M., Hsia, R. and Ojcius, D. M. (2000). Closing in on chlamydia and its intracellular bag of tricks. *Microbiology,* 146 (11), 2723-2731.

Bayburt, T. H., Grinkova, Y. V. and Sligar, S. G. (2002a). Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters,* 2 (8), 853-856.

Bayburt, T. H., Leitz, A. J., Xie, G., Oprian, D. D. and Sligar, S. G. (2007). Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *Journal of Biological Chemistry,* 282 (20), 14875-14881.

Bayburt, T. H. and Sligar, S. G. (2002b). Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. *Proceedings of the National Academy of Sciences,* 99 (10), 6725-6730.

Bayburt, T. H. and Sligar, S. G. (2009). Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Science,* 12 (11), 2476-2481.

Bayburt, T. H. and Sligar, S. G. (2010). Membrane protein assembly into nanodiscs. *FEBS Letters,* 584 (9), 1721-1727.

Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J. P., Landau, E. M. and Pebay-Peyroula, E. (1999). Protein, lipid and water organization in bacteriorhodopsin crystals: A molecular view of the purple membrane at 1.9 Å resolution. *Structure,* 7 (8), 909-917.

Benz, I. and Schmidt, M. A. (2011). Structures and functions of autotransporter proteins in microbial pathogens. *International Journal of Medical Microbiology,* 301 (6), 461-468.

Bernfield, M., Götte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J. and Zako, M. (1999). Functions of cell surface heparan sulfate proteoglycans. *Annual Review of Biochemistry,* 68 (1), 729-777.

Beseničar, M., Maček, P., Lakey, J. H. and Anderluh, G. (2006). Surface plasmon resonance in protein–membrane interactions. *Chemistry and Physics of Lipids,* 141 (1), 169-178.

Beveridge, T. J. (1999). Structures of gram-negative cell walls and their derived membrane vesicles. *Journal of Bacteriology,* 181 (16), 4725-4733.

Bill, R. M., Henderson, P. J. F., Iwata, S., Kunji, E. R. S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C. G. and Vogel, H. (2011). Overcoming barriers to membrane protein structure determination. *Nature Biotechnology,* 29 (4), 335-340.

Billinton, N. and Knight, A. W. (2001). Seeing the wood through the trees: A review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence. *Analytical Biochemistry,* 291 (2), 175-197.

Binet, R., Letoffe, S., Ghigo, J. M., Delepelaire, P. and Wandersman, C. (1997). Protein secretion by gram-negative bacterial ABC exporters--a review. *Gene,* 192 (1), 7-11.

Borch, J., Torta, F., Sligar, S. G. and Roepstorff, P. (2008). Nanodiscs for immobilization of lipid bilayers and membrane receptors: Kinetic analysis of cholera toxin binding to a glycolipid receptor. *Analytical Chemistry,* 80 (16), 6245-6252.

Bornhorst, J. A. and Falke, J. J. (2000). [16] purification of proteins using polyhistidine affinity tags. *Methods in Enzymology,* 326 , 245-254.

Brandon, L. D. and Goldberg, M. B. (2001). Periplasmic transit and disulfide bond formation of the autotransported shigella protein IcsA. *Journal of Bacteriology,* 183 (3), 951-958.

Brunelle, B. W. and Sensabaugh, G. F. (2006). The ompA gene in chlamydia trachomatis differs in phylogeny and rate of evolution from other regions of the genome. *Infection and Immunity,* 74 (1), 578-585.

Buchwalow, I., Samoilova, V., Boecker, W. and Tiemann, M. (2011). Non-specific binding of antibodies in immunohistochemistry: Fallacies and facts. *Scientific Reports,* 1 , 1-6.

Butcher, B. A., Sklar, L., Seamer, L. and Glew, R. (1992). Heparin enhances the interaction of infective leishmania donovani promastigotes with mouse peritoneal macrophages. A fluorescence flow cytometric analysis. *The Journal of Immunology,* 148 (9), 2879-2886.

Byrne, G. I. (2010). Chlamydia trachomatis strains and virulence: Rethinking links to infection prevalence and disease severity. *Journal of Infectious Diseases,* 201 (Supplement 2), S126-S133.

Cabeen, M. T. and Jacobs-Wagner, C. (2005). Bacterial cell shape. *Nature Reviews Microbiology,* 3 (8), 601-610.

Caldwell, H., Kromhout, J. and Schachter, J. (1981). Purification and partial characterization of the major outer membrane protein of chlamydia trachomatis. *Infection and Immunity,* 31 (3), 1161-1176.

Cambell, LA and Kuo, CC (2006). Interactions of chlamydia with the host cells that mediate attachment and uptake. In: Bavoil, P. M. and Wyrick, P. B. (eds.). Wymondham: Horizon Bioscience, 505-522.

Cappuccio, J. A., Blanchette, C. D., Sulchek, T. A., Arroyo, E. S., Kralj, J. M., Hinz, A. K., Kuhn, E. A., Chromy, B. A., Segelke, B. W. and Rothschild, K. J. (2008). Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Molecular & Cellular Proteomics,* 7 (11), 2246-2253.

Carlson, J. H., Porcella, S. F., McClarty, G. and Caldwell, H. D. (2005). Comparative genomic analysis of chlamydia trachomatis oculotropic and genitotropic strains. *Infection and Immunity,* 73 (10), 6407-6418.

Cevenini, R., Donati, M. and Sambri, V. (2002). Chlamydia trachomatis–the agent. *Best Practice & Research Clinical Obstetrics & Gynaecology,* 16 (6), 761-773.

Chaga, G., Bochkariov, D. E., Jokhadze, G. G., Hopp, J. and Nelson, P. (1999a). Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt (II)-carboxymethylaspartate crosslinked agarose. *Journal of Chromatography A,* 864 (2), 247-256.

Chaga, G., Hopp, J. and Nelson, P. (1999b). Immobilized metal ion affinity chromatography on Co2 -carboxymethylaspartate–agarose superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. *Biotechnology and Applied Biochemistry,* 29 (1), 19-24.

Chen, J., Acton, T. B., Basu, S. K., Montelione, G. T. and Inouye, M. (2002). Enhancement of the solubility of proteins overexpressed in escherichia coli by heat shock. *Journal of Molecular Microbiology and Biotechnology,* 4 (6), 519-524.

Chen, J. C. R. and Stephens, R. S. (1997). < i> chlamydia trachomatis</i> glycosaminoglycan-dependent and independent attachment to eukaryotic cells. *Microbial Pathogenesis,* 22 (1), 23-30.

Chen, Y. J., Pornillos, O., Lieu, S., Ma, C., Chen, A. P. and Chang, G. (2007). X-ray structure of EmrE supports dual topology model. *Proceedings of the National Academy of Sciences,* 104 (48), 18999-19004.

Cheng, J., Hicks, D. B. and Krulwich, T. A. (1996). The purified bacillus subtilis tetracycline efflux protein TetA (L) reconstitutes both tetracycline–cobalt/H and na (K )/H exchange. *Proceedings of the National Academy of Sciences,* 93 (25), 14446-14451.

Citovsky, V., Zupan, J., Warnick, D. and Zambryski, P. (1992). Nuclear localization of agrobacterium VirE2 protein in plant cells. *Science (New York, NY),* 256 (5065), 1802-1805.

Civjan, N. R., Bayburt, T. H., Schuler, M. A. and Sligar, S. G. (2003). Direct solubilization of heterologously expressed membrane proteins by incorporation into nanoscale lipid bilayers. *BioTechniques, 35* (3), 556-563.

Claros, M. G., Brunak, S. and Heijne, G. (1997). Prediction of N-terminal protein sorting signals. *Current Opinion in Structural Biology, 7* (3), 394-398.

Clifton, D., Fields, K., Grieshaber, S., Dooley, C., Fischer, E., Mead, D., Carabeo, R. and Hackstadt, T. (2004). A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. *Proceedings of the National Academy of Sciences of the United States of America, 101* (27), 10166-10171.

Cocchiaro, J. L., Kumar, Y., Fischer, E. R., Hackstadt, T. and Valdivia, R. H. (2008). Cytoplasmic lipid droplets are translocated into the lumen of the chlamydia trachomatis parasitophorous vacuole. *Proceedings of the National Academy of Sciences, 105* (27), 9379-9384.

Cooper, M. A. (2002). Optical biosensors in drug discovery. *Nature Reviews Drug Discovery, 1* (7), 515-528.

Cooper, M. A. (2004). Advances in membrane receptor screening and analysis. *Journal of Molecular Recognition, 17* (4), 286-315.

Cowieson, N. P., Listwan, P., Kurz, M., Aagaard, A., Ravasi, T., Wells, C., Huber, T., Hume, D. A., Kobe, B. and Martin, J. L. (2005). Pilot studies on the parallel production of soluble mouse proteins in a bacterial expression system. *Journal of Structural and Functional Genomics, 6* (1), 13-20.

Crane, D. D., Carlson, J. H., Fischer, E. R., Bavoil, P., Hsia, R., Tan, C., Kuo, C. and Caldwell, H. D. (2006). Chlamydia trachomatis polymorphic membrane protein D is a species-common pan-neutralizing antigen. *Proceedings of the National Academy of Sciences of the United States of America, 103* (6), 1894-1899.

Crowe, J., Dobeli, H., Gentz, R., Hochuli, E., Stuber, D. and Henco, K. (1994). 6xHis-ni-NTA chromatography as a superior technique in recombinant protein expression/purification. *Methods in Molecular Biology (Clifton, N.J.), 31* , 371-387.

Das, A., Zhao, J., Schatz, G. C., Sligar, S. G. and Van Duyne, R. P. (2009). Screening of type I and II drug binding to human cytochrome P450-3A4 in nanodiscs by localized surface plasmon resonance spectroscopy. *Analytical Chemistry, 81* (10), 3754-3759.

Davis, C. H. and Wyrick, P. B. (1997). Differences in the association of chlamydia trachomatis serovar E and serovar L2 with epithelial cells in vitro may reflect biological differences in vivo. *Infection and Immunity,* 65 (7), 2914-2924.

Davis, C., Raulston, J. and Wyrick, P. (2002). Protein disulfide isomerase, a component of the estrogen receptor complex, is associated with chlamydia trachomatis serovar E attached to human endometrial epithelial cells. *Infection and Immunity,* 70 (7), 3413-3418.

Day, Y. S. N., Baird, C. L., Rich, R. L. and Myszka, D. G. (2009). Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface-and solution-based biophysical methods. *Protein Science,* 11 (5), 1017-1025.

De Gier, J. W. and Luirink, J. (2001). Biogenesis of inner membrane proteins in escherichia coli. *Molecular Microbiology,* 40 (2), 314-322.

Dé, E., Saint, N., Glinel, K., Meli, A. C., Lévy, D. and Jacob-Dubuisson, F. (2008). Influence of the passenger domain of a model autotransporter on the properties of its translocator domain. *Molecular Membrane Biology,* 25 (3), 192-202.

Desvaux, M., Hébraud, M., Talon, R. and Henderson, I. R. (2009). Secretion and subcellular localizations of bacterial proteins: A semantic awareness issue. *Trends in Microbiology,* 17 (4), 139-145.

Desvaux, M., Parham, N. J. and Henderson, I. R. (2004a). The autotransporter secretion system. *Research in Microbiology,* 155 (2), 53-60.

Desvaux, M., Parham, N. J. and Henderson, I. R. (2004b). Type V protein secretion: Simplicity gone awry? *Current Issues in Molecular Biology,* 6 , 111-124.

Desvaux, M., Parham, N. J., Scott-Tucker, A. and Henderson, I. R. (2004c). The general secretory pathway: A general misnomer? *Trends in Microbiology,* 12 (7), 306-309.

Dobrovetsky, E., Lu, M. L., Andorn-Broza, R., Khutoreskaya, G., Bray, J. E., Savchenko, A., Arrowsmith, C. H., Edwards, A. M. and Koth, C. M. (2005). High-throughput production of prokaryotic membrane proteins. *Journal of Structural and Functional Genomics,* 6 (1), 33-50.

Drory, O., Frolow, F. and Nelson, N. (2004). Crystal structure of yeast V-ATPase subunit C reveals its stator function. *EMBO Reports,* 5 (12), 1148-1152.

Dumon-Seignovert, L., Cariot, G. and Vuillard, L. (2004). The toxicity of recombinant proteins in< i> escherichia coli</i>: A comparison of overexpression in BL21 (DE3), C41 (DE3), and C43 (DE3). *Protein Expression and Purification,* 37 (1), 203-206.

Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. and MacKinnon, R. (2002). X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature,* 415 (6869), 287-294.

Fadel, S. and Eley, A. (2008). Differential glycosaminoglycan binding of chlamydia trachomatis OmcB protein from serovars E and LGV. *Journal of Medical Microbiology,* 57 (9), 1058-1061.

Fägerstam, L. G., Frostell, Å., Karlsson, R., Kullman, M., Larsson, A., Malmqvist, M. and Butt, H. (2004). Detection of antigen—antibody interactions by surface plasmon resonance. application to epitope mapping. *Journal of Molecular Recognition,* 3 (5-6), 208-214.

Fields, K. A., Hackstadt, T. and Wyrick, P. (2006). The chlamydia type III secretion system: Structure and implications for pathogenesis. *Chlamydia: Genomics and Pathogenesis,* , 220-233.

Foit, L., Morgan, G. J., Kern, M. J., Steimer, L. R., von Hacht, A. A., Titchmarsh, J., Warriner, S. L., Radford, S. E. and Bardwell, J. C. A. (2009). Optimizing protein stability in vivo. *Molecular Cell,* 36 (5), 861-871.

Fox, J. D., Kapust, R. B. and Waugh, D. S. (2008). Single amino acid substitutions on the surface of escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Science,* 10 (3), 622-630.

Freissler, E., Meyer auf der Heyde, A., David, G., Meyer, T. F. and Dehio, C. (2000). Syndecan-1 and syndecan-4 can mediate the invasion of OpaHSPG-expressing neisseria gonorrhoeae into epithelial cells. *Cellular Microbiology,* 2 (1), 69-82.

Gathmann, S., Rupprecht, E. and Schneider, D. (2006). High level expression of a protein precursor for functional studies. *Journal of Biochemistry and Molecular Biology,* 39 (6), 717-721.

Gentschev, I., Dietrich, G. and Goebel, W. (2002). The< i> E. coli</i> α-hemolysin secretion system and its use in vaccine development. *Trends in Microbiology,* 10 (1), 39-45.

Giannetti, A. M. and Björkman, P. J. (2004). HFE and transferrin directly compete for transferrin receptor in solution and at the cell surface. *Journal of Biological Chemistry,* 279 (24), 25866-25875.

Glück, J. M., Koenig, B. W. and Willbold, D. (2011). Nanodiscs allow the use of integral membrane proteins as analytes in surface plasmon resonance studies. *Analytical Biochemistry,* 408 (1), 46-52.

Goenka, S. and Rao, C. M. (2001). Expression of recombinant ζ-crystallin in< i> escherichia coli</i> with the help of GroEL/ES and its purification. *Protein Expression and Purification,* 21 (2), 260-267.

Gold, V. A. M., Duong, F. and Collinson, I. (2007). Structure and function of the bacterial sec translocon (review). *Molecular Membrane Biology,* 24 (5-6), 387-394.

Gomes, J. P., Bruno, W. J., Nunes, A., Santos, N., Florindo, C., Borrego, M. J. and Dean, D. (2007). Evolution of chlamydia trachomatis diversity occurs by widespread interstrain recombination involving hotspots. *Genome Research,* 17 (1), 50-60.

Gomes, J. P., Hsia, R., Mead, S., Borrego, M. J. and Dean, D. (2005). Immunoreactivity and differential developmental expression of known and putative chlamydia trachomatis membrane proteins for biologically variant serovars representing distinct disease groups. *Microbes and Infection,* 7 (3), 410-420.

Gomes, J. P., Nunes, A., Bruno, W. J., Borrego, M. J., Florindo, C. and Dean, D. (2006). Polymorphisms in the nine polymorphic membrane proteins of chlamydia trachomatis across all serovars: Evidence for serovar da recombination and correlation with tissue tropism. *Journal of Bacteriology,* 188 (1), 275-286.

Goodall, J. C., Yeo, G., Huang, M., Raggiaschi, R. and Gaston, J. S. H. (2001). Identification of chlamydia trachomatis antigens recognized by human CD4 T lymphocytes by screening an expression library. *European Journal of Immunology,* 31 (5), 1513-1522.

Gordon, E., Horsefield, R., Swarts, H. G. P., de Pont, J. J. H. H. M., Neutze, R. and Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in< i> escherichia coli</i>. *Protein Expression and Purification,* 62 (1), 1-8.

Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G. A., Gottesman, M. E. and Nikiforov, V. (1992). Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in escherichia coli. *Proceedings of the National Academy of Sciences,* 89 (21), 10341-10344.

Gräslund, S., Sagemark, J., Berglund, H., Dahlgren, L. G., Flores, A., Hammarström, M., Johansson, I., Kotenyova, T., Nilsson, M. and Nordlund, P. (2008). The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. *Protein Expression and Purification,* 58 (2), 210-221.

Grieshaber, S., Swanson, J. A. and Hackstadt, T. (2002). Determination of the physical environment within the chlamydia trachomatis inclusion using ion-selective ratiometric probes. *Cellular Microbiology,* 4 (5), 273-283.

Griffin, S. D. C., Beales, L. P., Clarke, D. S., Worsfold, O., Evans, S. D., Jaeger, J., Harris, M. P. G. and Rowlands, D. J. (2003). The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, amantadine. *FEBS Letters,* 535 (1), 34-38.

Grimwood, J., Olinger, L. and Stephens, R. S. (2001). Expression of chlamydia pneumoniaePolymorphic membrane protein family genes. *Infection and Immunity,* 69 (4), 2383-2389.

Grimwood, J. and Stephens, R. S. (1999). Computational analysis of the polymorphic membrane protein superfamily of chlamydia trachomatis and chlamydia pneumoniae. *Microbial & Comparative Genomics,* 4 (3), 187-201.

Grodberg, J. and Dunn, J. J. (1988). ompT encodes the escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. *Journal of Bacteriology,* 170 (3), 1245-1253.

Groffen, A. J., Ruegg, M. A., Dijkman, H., van de Velden, T. J., Buskens, C. A., van den Born, J., Assmann, K. J., Monnens, L. A., Veerkamp, J. H. and van den Heuvel, L. P. (1998). Agrin is a major heparan sulfate proteoglycan in the human glomerular basement membrane. *Journal of Histochemistry & Cytochemistry,* 46 (1), 19-27.

Grossman, T. H., Kawasaki, E. S., Punreddy, S. R. and Osburne, M. S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene,* 209 (1), 95-103.

Hao, Z., Li, X., Qiao, T., Du, R., Zhang, G. and Fan, D. (2006). Subcellular localization of CIAPIN1. *Journal of Histochemistry & Cytochemistry,* 54 (12), 1437-1444.

Hatch, T. (1998). Chlamydia: Old ideas crushed, new mysteries bared. *Science,* 282 (5389), 638-639.

Hatch, T. P. (1996). Disulfide cross-linked envelope proteins: The functional equivalent of peptidoglycan in chlamydiae? *Journal of Bacteriology,* 178 (1), 1-5.

Hatch, T. P., Vance, D. W. and Al-Hossainy, E. (1981). Attachment of chlamydia psittaci to formaldehyde-fixed and unfixed L cells. *Journal of General Microbiology,* 125 (2), 273-283.

321

Henderson, I. R., Cappello, R. and Nataro, J. P. (2000). Autotransporter proteins, evolution and redefining protein secretion. *Trends in Microbiology,* 8 (12), 529-532.

Henderson, I. R. and Lam, A. C. (2001a). Polymorphic proteins of chlamydia spp.– autotransporters beyond the proteobacteria. *Trends in Microbiology,* 9 (12), 573-578.

Henderson, I. R. and Nataro, J. P. (2001b). Virulence functions of autotransporter proteins. *Infection and Immunity,* 69 (3), 1231-1243.

Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. and Ala'Aldeen, D. (2004). Type V protein secretion pathway: The autotransporter story. *Microbiology and Molecular Biology Reviews,* 68 (4), 692-744.

Henderson, I. R., Navarro-Garcia, F. and Nataro, J. P. (1998). The great escape: Structure and function of the autotransporter proteins. *Trends in Microbiology,* 6 (9), 370-378.

Hirt, R. P., Noel, C. J., Sicheritz-Pontén, T., Tachezy, J. and Fiori, P. L. (2007). < i> trichomonas vaginalis</i> surface proteins: A view from the genome. *Trends in Parasitology,* 23 (11), 540-547.

Hofmann, E., Reichart, U., Gausterer, C., Guelly, C., Meijer, D., Müller, M. and Strobl, B. (2010). Octamer-binding factor 6 (oct-6/Pou3f1) is induced by interferon and contributes to dsRNA-mediated transcriptional responses. *BMC Cell Biology,* 11 (1), 61-77.

Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A. and Heesemann, J. (2000). Structure and sequence analysis of yersinia YadA and moraxella UspAs reveal a novel class of adhesins. *The EMBO Journal,* 19 (22), 5989-5999.

Howard, E. A., Zupan, J. R., Citovsky, V. and Zambryski, P. C. (1992). The VirD2 protein of A. tumefaciens contains a C-terminal bipartite nuclear localization signal: Implications for nuclear uptake of DNA in plant cells. *Cell,* 68 (1), 109-118.

Hueck, C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and Molecular Biology Reviews,* 62 (2), 379-433.

Ieva, R. and Bernstein, H. D. (2009). Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proceedings of the National Academy of Sciences,* 106 (45), 19120-19125.

Jarchau, T., Chakraborty, T., Garcia, F. and Goebel, W. (1994). Selection for transport competence of C-terminal polypeptides derived from escherichia coli hemolysin: The

shortest peptide capable of autonomous HlyB/HlyD-dependent secretion comprises the C-terminal 62 amino acids of HlyA. *Molecular and General Genetics MGG,* 245 (1), 53-60.

Jason-Mlooer, L., Murphy, M. and Bruno, J. A. (2006). Overview of biacore systems and their applications. In: Wiley Online Library, 19.13. 1-19.13. 14.

Jeffrey, B. M., Suchland, R. J., Quinn, K. L., Davidson, J. R., Stamm, W. E. and Rockey, D. D. (2010). Genome sequencing of recent clinical chlamydia trachomatis strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infection and Immunity,* 78 (6), 2544-2553.

Jewett, T. J., Dooley, C. A., Mead, D. J. and Hackstadt, T. (2008). < i> chlamydia trachomatis</i> tarp is phosphorylated by src family tyrosine kinases. *Biochemical and Biophysical Research Communications,* 371 (2), 339-344.

Johnson, R. M. (2004). Murine oviduct epithelial cell cytokine responses to chlamydia muridarum infection include interleukin-12-p70 secretion. *Infection and Immunity,* 72 (7), 3951-3960.

Jokiranta, T. S., Hellwage, J., Koistinen, V., Zipfel, P. F. and Meri, S. (2000). Each of the three binding sites on complement factor H interacts with a distinct site on C3b. *Journal of Biological Chemistry,* 275 (36), 27657-27662.

Jose, J., Krämer, J., Klauser, T., Pohlner, J. and Meyer, T. F. (1996). Absence of periplasmic DsbA oxidoreductase facilitates export of cysteine-containing passenger proteins to the< i> escherichia coli</i> cell surface via the iga< sub> β</sub> autotransporter pathway. *Gene,* 178 (1), 107-110.

Jotblad, S., Park, I. U., Bauer, H. M., Barandas, A., Deal, M. and Amey, A. (2012). Patient-delivered partner therapy for chlamydial infections: Practices, attitudes, and knowledge of california family planning providers. *Sexually Transmitted Diseases,* 39 (2), 122-127.

Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R. W., Olinger, L., Grimwood, J., Davis, R. W. and Stephens, R. S. (1999). Comparative genomes of chlamydia pneumoniae and C. trachomatis. *Nature Genetics,* 21 (4), 385-389.

Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in< i> escherichia coli</i>. *Current Opinion in Biotechnology,* 6 (5), 494-500.

Kästner, C. N., Dimroth, P. and Pos, K. M. (2000). The na -dependent citrate carrier of klebsiella pneumoniae: High-level expression and site-directed mutagenesis of asparagine-185 and glutamate-194. *Archives of Microbiology,* 174 (1), 67-73.

Katayama, H., Wang, J., Tama, F., Chollet, L., Gogol, E., Collier, R. and Fisher, M. (2010). Three-dimensional structure of the anthrax toxin pore inserted into lipid nanodiscs and lipid vesicles. *Proceedings of the National Academy of Sciences,* 107 (8), 3453-3457.

Katz, A. R., Lee, M. V. C. and Wasserman, G. M. (2012). Sexually transmitted disease (STD) update: A review of the CDC 2010 STD treatment guidelines and epidemiologic trends of common STDs in hawai ʻi. *Hawaiʻi Journal of Medicine & Public Health,* 71 (3), 68-73.

Katzen, F., Chang, G. and Kudlicki, W. (2005). The past, present and future of cell-free protein synthesis. *Trends in Biotechnology,* 23 (3), 150-156.

Katzen, F., Fletcher, J. E., Yang, J. P., Kang, D., Peterson, T. C., Cappuccio, J. A., Blanchette, C. D., Sulchek, T., Chromy, B. A. and Hoeprich, P. D. (2008). Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *Journal of Proteome Research,* 7 (8), 3535-3542.

Katzen, F., Peterson, T. C. and Kudlicki, W. (2009). Membrane protein expression: No cells required. *Trends in Biotechnology,* 27 (8), 455-460.

Katzen, F., Fletcher, J. E., Yang, J. P., Vasu, S., Peterson, T. and Kudlicki, W. (2008). *Cell-free protein expression of membrane proteins using nanolipoprotein particles.* England, BioTechniques, 45 (4), 469 .

Kelly, S. M., Jess, T. J. and Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica Et Biophysica Acta (BBA)-Proteins & Proteomics,* 1751 (2), 119-139.

Kiefhaber, T., Rudolph, R., Kohler, H. H. and Buchner, J. (1991). Protein aggregation in vitro and in vivo: A quantitative model of the kinetic competition between folding and aggregation. *Nature Biotechnology,* 9 (9), 825-829.

Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. and Yokoyama, S. (1999). Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Letters,* 442 (1), 15-19.

Kiselev, A. O., Stamm, W. E., Yates, J. R. and Lampe, M. F. (2007). Expression, processing, and localization of PmpD of chlamydia trachomatis serovar L2 during the chlamydial developmental cycle. *PLoS One,* 2 (6), e568-e575.

Klammt, C., Löhr, F., Schäfer, B., Haase, W., Dötsch, V., Rüterjans, H., Glaubitz, C. and Bernhard, F. (2004). High level cell-free expression and specific labeling of integral membrane proteins. *European Journal of Biochemistry,* 271 (3), 568-580.

Klockenbusch, C. and Kast, J. (2010). Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin $\beta$ 1. *Journal of Biomedicine and Biotechnology,* 2010 , 1-14.

Kobayashi, M., Ishida, K., Matsuo, J., Nakamura, S., Nagasawa, A., Motohashi, K., Yao, T., Hirai, I., Yamamoto, Y. and Suzuki, H. (2011). < i> chlamydophila pneumoniae</i> attachment and infection in low proteoglycan expressing human lymphoid jurkat cells. *Microbial Pathogenesis,* 51 (3), 209-216.

Kopetzki, E., Schumacher, G. and Buckel, P. (1989). Control of formation of active soluble or inactive insoluble baker's yeast α-glucosidase PI in escherichia coli by induction and growth conditions. *Molecular and General Genetics MGG,* 216 (1), 149-155.

Korepanova, A., Gao, F. P., Hua, Y., Qin, H., Nakamoto, R. K. and Cross, T. A. (2009). Cloning and expression of multiple integral membrane proteins from mycobacterium tuberculosis in escherichia coli. *Protein Science,* 14 (1), 148-158.

Kosma, P. (1999). Chlamydial lipopolysaccharide. *Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1455 (2), 387-402.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene,* 234 (2), 187-208.

Kriechbaumer, V., Nabok, A., Mustafa, M. K., Al-Ammar, R., Tsargorodskaya, A., Smith, D. P. and Abell, B. M. (2012a). Analysis of protein interactions at native chloroplast membranes by ellipsometry. *PloS One,* 7 (3), e34455-e34462.

Kriechbaumer, V., Nabok, A., Widdowson, R., Smith, D. P. and Abell, B. M. (2012b). Quantification of ligand binding to G-protein coupled receptors on cell membranes by ellipsometry. *PLOS ONE,* 7 (9), e46221-e46230.

Kriechbaumer, V., Tsargorodskaya, A., Mustafa, M. K., Vinogradova, T., Lacey, J., Smith, D. P., Abell, B. M. and Nabok, A. (2011). Study of receptor-chaperone interactions using the optical technique of spectroscopic ellipsometry. *Biophysical Journal,* 101 (2), 504-511.

Kuo, C., Lee, A. and Campbell, L. A. (2004). Cleavage of the N-linked oligosaccharide from the surfaces of chlamydia species affects attachment and infectivity of the organisms in human epithelial and endothelial cells. *Infection and Immunity,* 72 (11), 6699-6701.

Kuo, C., Puolakkainen, M., Lin, T. M., Witte, M. and Campbell, L. A. (2002). Mannose-receptor positive and negative mouse macrophages differ in their susceptibility to infection by< i> Chlamydia</i> species. *Microbial Pathogenesis,* 32 (1), 43-48.

Kuo, C., Takahashi, N., Swanson, A. F., Ozeki, Y. and Hakomori, S. (1996). An N-linked high-mannose type oligosaccharide, expressed at the major outer membrane protein of chlamydia trachomatis, mediates attachment and infectivity of the microorganism to HeLa cells. *Journal of Clinical Investigation,* 98 (12), 2813-2818.

Kuo, C. C., Wang, S. P. and Grayston, J. T. (1973). Effect of polycations, polyanions, and neuraminidase on the infectivity of trachoma-inclusion conjunctivitis and lymphogranuloma venereum organisms in HeLa cells: Sialic acid residues as possible receptors for trachoma-inclusion conjunctivitis. *Infection and Immunity,* 8 (1), 74-79.

Kuo, C., Grayston, J., Campbell, L., Goo, Y., Wissler, R. and Benditt, E. (1995). Chlamydia pneumoniae (TWAR) in coronary arteries of young adults (15-34 years old). *Proceedings of the National Academy of Sciences,* 92 (15), 6911-6914.

Laarmann, S., Cutter, D., Juehne, T., Barenkamp, S. J. and St Geme, J. W. (2002). The haemophilus influenzae hia autotransporter harbours two adhesive pockets that reside in the passenger domain and recognize the same host cell receptor. *Molecular Microbiology,* 46 (3), 731-743.

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

Lee, C. W., Eu, Y. J., Min, H. J., Cho, E. M., Lee, J. H., Kim, H. H., Nah, S. Y., Swartz, K. J. and Kim, J. I. (2011). Expression and characterization of recombinant kurtoxin, an inhibitor of T-type voltage-gated calcium channels. *Biochemical and Biophysical Research Communications,* 416 (3-4), 277-282.

Leitz, A. J., Bayburt, T. H., Barnakov, A. N., Springer, B. A. and Sligar, S. G. (2006). Functional reconstitution of beta~ 2-adrenergic receptors utilizing self-assembling nanodisc technology. *BioTechniques,* 40 (5), 601-607.

Lilie, H., Schwarz, E. and Rudolph, R. (1998). Advances in refolding of proteins produced in< i> E. coli</i>. *Current Opinion in Biotechnology,* 9 (5), 497-501.

Liu, X., Afrane, M., Clemmer, D. E., Zhong, G. and Nelson, D. E. (2010). Identification of chlamydia trachomatis outer membrane complex proteins by differential proteomics. *Journal of Bacteriology,* 192 (11), 2852-2860.

Loll, P. J. (2003). Membrane protein structural biology: The high throughput challenge. *Journal of Structural Biology,* 142 (1), 144-153.

Lorimer, G. (1996). A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo. *The FASEB Journal,* 10 (1), 5-9.

Lurie-Weinberger, M. N., Gomez-Valero, L., Merault, N., Glöckner, G., Buchrieser, C. and Gophna, U. (2010). The origins of eukaryotic-like proteins in< i> legionella pneumophila</i>. *International Journal of Medical Microbiology,* 300 (7), 470-481.

Malaisse, W., Svoboda, M., Dufrane, S., Malaisse-Lagae, F. and Christophe, J. (1984). Effect of< i> bordetella pertussis</i> toxin on ADP-ribosylation of membrane proteins, adenylate cyclase activity and insulin release in rat pancreatic islets. *Biochemical and Biophysical Research Communications,* 124 (1), 190-196.

Mancia, F. and Love, J. (2010). High-throughput expression and purification of membrane proteins. *Journal of Structural Biology,* 172 (1), 85-93.

Masi, M., Pagès, J. M. and Pradel, E. (2003). Overexpression and purification of the three components of the< i> enterobacter aerogenes</i> AcrA–AcrB–TolC multidrug efflux pump. *Journal of Chromatography B,* 786 (1), 197-205.

Matsumoto, A. (1982). Electron microscopic observations of surface projections on chlamydia psittaci reticulate bodies. *Journal of Bacteriology,* 150 (1), 358-364.

Maurer, J., Jose, J. and Meyer, T. F. (1999). Characterization of the essential transport function of the AIDA-I autotransporter and evidence supporting structural predictions. *Journal of Bacteriology,* 181 (22), 7014-7020.

McCafferty, M. C., Herring, A. J., Andersen, A. A. and Jones, G. E. (1995). Electrophoretic analysis of the major outer membrane protein of chlamydia psittaci reveals multimers which are recognized by protective monoclonal antibodies. *Infection and Immunity,* 63 (6), 2387-2389.

Mi, L. Z., Grey, M. J., Nishida, N., Walz, T., Lu, C. and Springer, T. A. (2008). Functional and structural stability of the epidermal growth factor receptor in detergent micelles and phospholipid nanodiscs†. *Biochemistry,* 47 (39), 10314-10323.

Middelberg, A. P. J. (2002). Preparative protein refolding. *Trends in Biotechnology,* 20 (10), 437-443.

Miroux, B. and Walker, J. E. (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.

Mital, J., Miller, N. J., Fischer, E. R. and Hackstadt, T. (2010). Specific chlamydial inclusion membrane proteins associate with active src family kinases in microdomains that interact with the host microtubule network. *Cellular Microbiology,* 12 (9), 1235-1249.

Moelleken, K. and Hegemann, J. H. (2008). The chlamydia outer membrane protein OmcB is required for adhesion and exhibits biovar-specific differences in glycosaminoglycan binding. *Molecular Microbiology,* 67 (2), 403-419.

Mohanty, A. K. and Wiener, M. C. (2004). Membrane protein expression and production: Effects of polyhistidine tag length and position. *Protein Expression and Purification,* 33 (2), 311-325.

Mölleken, K., Schmidt, E. and Hegemann, J. H. (2010a). Members of the pmp protein family of chlamydia pneumoniae mediate adhesion to human cells via short repetitive peptide motifs. *Molecular Microbiology,* 78 (4), 1004-1017.

Mölleken, K., Schmidt, E. and Hegemann, J. H. (2010b). Members of the pmp protein family of chlamydia pneumoniae mediate adhesion to human cells via short repetitive peptide motifs. *Molecular Microbiology,* 78 (4), 1004-1017.

Moulder, J. W. (1991). Interaction of chlamydiae and host cells in vitro. *Microbiological Reviews,* 55 (1), 143-190.

Mustafa, M., Nabok, A., Parkinson, D., Tothill, I. E., Salam, F. and Tsargorodskaya, A. (2010). Detection of β-amyloid peptide (1–16) and amyloid precursor protein (APP$_{770}$) using spectroscopic ellipsometry and QCM techniques: A step forward towards alzheimers disease diagnostics. *Biosensors and Bioelectronics,* 26 (4), 1332-1336.

Mygind, P. H., Christiansen, G., Roepstorff, P. and Birkelund, S. (2006). Membrane proteins PmpG and PmpH are major constituents of chlamydia trachomatis L2 outer membrane complex. *FEMS Microbiology Letters,* 186 (2), 163-169.

Mygind, T., Vandahl, B., Pedersen, A. S., Christiansen, G., Höllsberg, P. and Birkelund, S. (2004). Identification of an in vivo CD4 T cell-mediated response to polymorphic membrane proteins of chlamydia pneumoniae during experimental infection. *FEMS Immunology & Medical Microbiology,* 40 (2), 129-137.

Nabok, A. and Tsargorodskaya, A. (2008). The method of total internal reflection ellipsometry for thin film characterisation and sensing. *Thin Solid Films,* 516 (24), 8993-9001.

Nabok, A., Tsargorodskaya, A., Holloway, A., Starodub, N. F. and Demchenko, A. (2007). Specific binding of large aggregates of amphiphilic molecules to the respective antibodies. *Langmuir,* 23 (16), 8485-8490.

Nabok, A., Tsargorodskaya, A., Mustafa, M., Szekacs, I., Starodub, N. and Szekacs, A. (2011). Detection of low molecular weight toxins using an optical phase method of ellipsometry. *Sensors and Actuators B: Chemical,* 154 (2), 232-237.

Nabok, A., Tsargorodskaya, A., Hassan, A. and Starodub, N. (2005). Total internal reflection ellipsometry and SPR detection of low molecular weight environmental toxins. *Applied Surface Science,* 246 (4), 381-386.

Narayanan, A., Ridilla, M. and Yernool, D. A. (2011). Restrained expression, a method to overproduce toxic membrane proteins by exploiting operator–repressor interactions. *Protein Science,* 20 (1), 51-61.

Nath, A., Atkins, W. M. and Sligar, S. G. (2007). Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry,* 46 (8), 2059-2069.

Natsume, T., Taoka, M., Manki, H., Kume, S., Isobe, T. and Mikoshiba, K. (2002). Rapid analysis of protein interactions: On-chip micropurification of recombinant protein expressed in esherichia coli. *Proteomics,* 2 (9), 1247-1253.

Nedelkov, D. and Nelson, R. W. (2001). Analysis of native proteins from biological fluids by biomolecular interaction analysis mass spectrometry (BIA/MS): Exploring the limit of detection, identification of non-specific binding and detection of multi-protein complexes. *Biosensors and Bioelectronics,* 16 (9), 1071-1078.

Nieba, L., Nieba-Axmann, S. E., Persson, A., Hämäläinen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, Å. F. and Plückthun, A. (1997). BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Analytical Biochemistry,* 252 (2), 217-228.

Nielsen, P. K., Gho, Y. S., Hoffman, M. P., Watanabe, H., Makino, M., Nomizu, M. and Yamada, Y. (2000). Identification of a major heparin and cell binding site in the LG4 module of the laminin α5 chain. *Journal of Biological Chemistry,* 275 (19), 14517-14523.

Niessner, A., Kaun, C., Zorn, G., Speidl, W., Türel, Z., Christiansen, G., Pedersen, A. S., Birkelund, S., Simon, S. and Georgopoulos, A. (2003). Polymorphic membrane protein (PMP) 20 and PMP 21 of chlamydia pneumoniae induce proinflammatory mediators in

human endothelial cells in vitro by activation of the nuclear factor-κB pathway. *Journal of Infectious Diseases,* 188 (1), 108-113.

Nordin, H., Jungnelius, M., Karlsson, R. and Karlsson, O. P. (2005). Kinetic studies of small molecule interactions with protein kinases using biosensor technology. *Analytical Biochemistry,* 340 (2), 359-368.

Nunes, A., Gomes, J. P., Mead, S., Florindo, C., Correia, H., Borrego, M. J. and Dean, D. (2007). Comparative expression profiling of the chlamydia trachomatis pmp gene family for clinical and reference strains. *PLoS One,* 2 (9), e878.

Oliveira, P. H., Prather, K. J., Prazeres, D. M. F. and Monteiro, G. A. (2009). Structural instability of plasmid biopharmaceuticals: Challenges and implications. *Trends in Biotechnology,* 27 (9), 503-511.

Oliver, D. B. (1993). SecA protein: Autoregulated ATPase catalysing preprotein insertion and translocation across the escherichia coli inner membrane. *Molecular Microbiology,* 7 (2), 159-165.

Oliver, D. B., Cabelli, R. J., Dolan, K. M. and Jarosik, G. P. (1990). Azide-resistant mutants of escherichia coli alter the SecA protein, an azide-sensitive component of the protein export machinery. *Proceedings of the National Academy of Sciences,* 87 (21), 8227-8231.

Oliver, D. C., Huang, G. and Fernandez, R. C. (2003a). Identification of secretion determinants of the bordetella pertussis BrkA autotransporter. *Journal of Bacteriology,* 185 (2), 489-495.

Oliver, D. C., Huang, G., Nodel, E., Pleasance, S. and Fernandez, R. C. (2003b). A conserved region within the bordetella pertussis autotransporter BrkA is necessary for folding of its passenger domain. *Molecular Microbiology,* 47 (5), 1367-1383.

Oomen, C. J., Van Ulsen, P., Van Gelder, P., Feijen, M., Tommassen, J. and Gros, P. (2004). Structure of the translocator domain of a bacterial autotransporter. *The EMBO Journal,* 23 (6), 1257-1266.

Oshannessy, D. J., Odonnell, K. C., Martin, J. and Brighamburke, M. (1995). Detection and quantitation of hexa-histidine-tagged recombinant proteins on western blots and by a surface plasmon resonance biosensor technique. *Analytical Biochemistry,* 229 (1), 119-124.

Pedersen, A. S., Christiansen, G. and Birkelund, S. (2006). Differential expression of Pmp10 in cell culture infected with chlamydia pneumoniae CWL029. *FEMS Microbiology Letters,* 203 (2), 153-159.

Pelletier, H. and Kraut, J. (1992). Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. *Science (New York, NY)*, 258 (5089), 1748-1755.

Peters, J., Wilson, D. P., Myers, G., Timms, P. and Bavoil, P. M. (2007). Type III secretion à la< i> Chlamydia</i>. *Trends in Microbiology*, 15 (6), 241-251.

Pos, K. M., Schiefner, A., Seeger, M. A. and Diederichs, K. (2004). Crystallographic analysis of AcrB. *FEBS Letters*, 564 (3), 333-339.

Prinz, W. A., Åslund, F., Holmgren, A. and Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the escherichia colicytoplasm. *Journal of Biological Chemistry*, 272 (25), 15661-15667.

Ravanat, C., Wurtz, V., Ohlmann, P., Fichter, M., Cazenave, J. P., VanDorsselaer, A., Lanza, F. and Gachet, C. (2009). Use of tandem Biacore–mass spectrometry to identify platelet membrane targets of novel monoclonal antibodies. *Analytical Biochemistry*, 386 (2), 237-243.

Read, T. D., Brunham, R., Shen, C., Gill, S., Heidelberg, J., White, O., Hickey, E., Peterson, J., Utterback, T. and Berry, K. (2000). Genome sequences of chlamydia trachomatis MoPn and chlamydia pneumoniae AR39. *Nucleic Acids Research*, 28 (6), 1397-1406.

Rich, R. L. and Myszka, D. G. (2000). Advances in surface plasmon resonance biosensor analysis. *Current Opinion in Biotechnology*, 11 (1), 54-61.

Rich, R. L. and Myszka, D. G. (2008). Survey of the year 2007 commercial optical biosensor literature. *Journal of Molecular Recognition*, 21 (6), 355-400.

Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G. D. and Gold, L. (1992). Translation initiation in escherichia coli: Sequences within the ribosome-binding site. *Molecular Microbiology*, 6 (9), 1219-1229.

Rockey, D. D., Scidmore, M. A., Bannantine, J. P. and Brown, W. J. (2002). Proteins in the chlamydial inclusion membrane. *Microbes and Infection*, 4 (3), 333-340.

Rudolph, R. and Lilie, H. (1996). In vitro folding of inclusion body proteins. *The FASEB Journal*, 10 (1), 49-56.

Rzomp, K., Moorhead, A. and Scidmore, M. (2006). The GTPase Rab4 interacts with chlamydia trachomatis inclusion membrane protein CT229. *Infection and Immunity*, 74 (9), 5362-5373.

SAMBROOK, J. and RUSSELL, D. W. (2001). *Molecular cloning: A laboratory manual.* CSHL press. 1.

Sandkvist, M. (2001). Type II secretion and pathogenesis. *Infection and Immunity,* 69 (6), 3523-3535.

Sarramegna, V., Talmont, F., Demange, P. and Milon, A. (2003). Heterologous expression of G-protein-coupled receptors: Comparison of expression systems from the standpoint of large-scale production and purification. *Cellular and Molecular Life Sciences,* 60 (8), 1529-1546.

Savage, D. F., Anderson, C. L., Robles-Colmenares, Y., Newby, Z. E. and Stroud, R. M. (2009). Cell-free complements in vivo expression of the E. coli membrane proteome. *Protein Science,* 16 (5), 966-976.

Schein, C. H. and Noteborn, M. H. M. (1988). Formation of soluble recombinant proteins in escherichia coli is favored by lower growth temperature. *Nature Biotechnology,* 6 (3), 291-294.

Schmitt, J., Hess, H. and Stunnenberg, H. G. (1993). Affinity purification of histidine-tagged proteins. *Molecular Biology Reports,* 18 (3), 223-230.

Schnell, S. A., Staines, W. A. and Wessendorf, M. W. (1999). Reduction of lipofuscin-like autofluorescence in fluorescently labeled tissue. *Journal of Histochemistry & Cytochemistry,* 47 (6), 719-730.

Scidmore, M. A., Fischer, E. R. and Hackstadt, T. (2003). Restricted fusion of chlamydia trachomatis vesicles with endocytic compartments during the initial stages of infection. *Infection and Immunity,* 71 (2), 973-984.

Scidmore, M. A., Rockey, D. D., Fischer, E. R., Heinzen, R. A. and Hackstadt, T. (1996). Vesicular interactions of the chlamydia trachomatis inclusion are determined by chlamydial early protein synthesis rather than route of entry. *Infection and Immunity,* 64 (12), 5366-5372.

Seddon, A. M., Curnow, P. and Booth, P. J. (2004). Membrane proteins, lipids and detergents: Not just a soap opera. *Biochimica Et Biophysica Acta (BBA)-Biomembranes,* 1666 (1), 105-117.

Shaw, A. W., Pureza, V. S., Sligar, S. G. and Morrissey, J. H. (2007). The local phospholipid environment modulates the activation of blood clotting. *Journal of Biological Chemistry,* 282 (9), 6556-6563.

Shaw, E., Dooley, C., Fischer, E., Scidmore, M., Fields, K. and Hackstadt, T. (2002). Three temporal classes of gene expression during the chlamydia trachomatis developmental cycle. *Molecular Microbiology,* 37 (4), 913-925.

Sijbrandi, R., Urbanus, M. L., Corinne, M., Bernstein, H. D., Oudega, B., Otto, B. R. and Luirink, J. (2003). Signal recognition particle (SRP)-mediated targeting and sec-dependent translocation of an extracellular escherichia coli protein. *Journal of Biological Chemistry,* 278 (7), 4654-4659.

Sinz, A. (2010). Investigation of protein–protein interactions in living cells by chemical crosslinking and mass spectrometry. *Analytical and Bioanalytical Chemistry,* 397 (8), 3433-3440.

Skipp, P., Robinson, J., O'Connor, C. D. and Clarke, I. N. (2005). Shotgun proteomic analysis of chlamydia trachomatis. *Proteomics,* 5 (6), 1558-1573.

Sneddon, J. M. and Wenman, W. M. (1985). The effect of ions on the adhesion and internalization of chlamydia trachomatis by HeLa cells. *Canadian Journal of Microbiology,* 31 (4), 371-374.

Söderlund, G. and Kihlström, E. (1983). Attachment and internalization of a chlamydia trachomatis lymphogranuloma venereum strain by McCoy cells: Kinetics of infectivity and effect of lectins and carbohydrates. *Infection and Immunity,* 42 (3), 930-935.

Söderlund, G. and Kihlström, E. (1988). Factors affecting the adhesion and uptake of an oculo-genital strain of< i> chlamydia trachomatis</i> to McCoy cells. *FEMS Microbiology Letters,* 49 (3), 467-471.

Sørensen, H. P. and Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in< i> escherichia coli</i>. *Journal of Biotechnology,* 115 (2), 113-128.

Sprengart, M. L., Fuchs, E. and Porter, A. (1996). The downstream box: An efficient and independent translation initiation signal in escherichia coli. *The EMBO Journal,* 15 (3), 665-674.

Srere, P., Sherry, A., Malloy, C. and Sumegi, B. (1997). Channelling in the krebs tricarboxylic acid cycle. *Portland Press Research Monograph,* , 201-218.

Stenberg, E., Persson, B., Roos, H. and Urbaniczky, C. (1991). Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. *Journal of Colloid and Interface Science,* 143 (2), 513-526.

Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L. and Zhao, Q. (1998). Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis. *Science,* 282 (5389), 754-759.

Stephens, R. S. and Lammel, C. J. (2001). Chlamydia outer membrane protein discovery using genomics. *Current Opinion in Microbiology,* 4 (1), 16-20.

Stothard, D. R., Toth, G. A. and Batteiger, B. E. (2003). Polymorphic membrane protein H has evolved in parallel with the three disease-causing groups of chlamydia trachomatis. *Infection and Immunity,* 71 (3), 1200-1208.

Struyvé, M., Moons, M. and Tommassen, J. (1991). Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *Journal of Molecular Biology,* 218 (1), 141-148.

Su, H., Raymond, L., Rockey, D. D., Fischer, E., Hackstadt, T. and Caldwell, H. D. (1996). A recombinant chlamydia trachomatis major outer membrane protein binds to heparan sulfate receptors on epithelial cells. *Proceedings of the National Academy of Sciences,* 93 (20), 11143-11148.

Sutherland, B. W., Toews, J. and Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of Mass Spectrometry,* 43 (6), 699-715.

Suzuki, N., Rohdewohld, H., Neuman, T., Gruss, P. and Schöler, H. (1990). Oct-6: A POU transcription factor expressed in embryonal stem cells and in the developing brain. *The EMBO Journal,* 9 (11), 3723.

Swanson, A. F. and Kuo, C. C. (1994). Binding of the glycan of the major outer membrane protein of chlamydia trachomatis to HeLa cells. *Infection and Immunity,* 62 (1), 24-28.

Swanson, K. A., Taylor, L. D., Frank, S. D., Sturdevant, G. L., Fischer, E. R., Carlson, J. H., Whitmire, W. M. and Caldwell, H. D. (2009). Chlamydia trachomatis polymorphic membrane protein D is an oligomeric autotransporter with a higher-order structure. *Infection and Immunity,* 77 (1), 508-516.

Takizawa, N. and Murooka, Y. (1985). Cloning of the pullulanase gene and overproduction of pullulanase in escherichia coli and klebsiella aerogenes. *Applied and Environmental Microbiology,* 49 (2), 294-298.

Tamura, M., Nogimori, K., Murai, S., Yajima, M., Ito, K., Katada, T., Ui, M. and Ishii, S. (1982). Subunit structure of islet-activating protein, pertussis toxin, in conformity with the AB model. *Biochemistry,* 21 (22), 5516-5522.

Tan, C., Hsia, R., Shou, H., Carrasco, J. A., Rank, R. G. and Bavoil, P. M. (2009a). Variable expression of surface-exposed polymorphic membrane proteins in in vitro-grown chlamydia trachomatis. *Cellular Microbiology,* 12 (2), 174-187.

Tan, C., Hsia, R., Shou, H., Haggerty, C. L., Ness, R. B., Gaydos, C. A., Dean, D., Scurlock, A. M., Wilson, D. P. and Bavoil, P. M. (2009b). Chlamydia trachomatis-infected patients display variable antibody profiles against the nine-member polymorphic membrane protein family. *Infection and Immunity,* 77 (8), 3218-3226.

Tan, C., et al. (2006). The polymorphic membrane protein gene family of the chlamydiaceae. In: Bavoil, P. M. and Wyrick, P. B. (eds.). *Chlamydia, genomics and pathogenesis.* 195-218.

Tang, X. and Bruce, J. E. (2009). Chemical cross-linking for protein-protein interaction studies. *Methods Mol.Biol,* 492 , 283-293.

Tanzer, R. J. and Hatch, T. P. (2001a). Characterization of outer membrane proteins in chlamydia trachomatis LGV serovar L2. *Journal of Bacteriology,* 183 (8), 2686-2690.

Tanzer, R. J., Longbottom, D. and Hatch, T. P. (2001b). Identification of polymorphic outer membrane proteins of chlamydia psittaci 6BC. *Infection and Immunity,* 69 (4), 2428-2434.

Thai, K., Choi, J., Franzin, C. M. and Marassi, F. M. (2009). Bcl-XL as a fusion protein for the high-level expression of membrane-associated proteins. *Protein Science,* 14 (4), 948-955.

Thanassi, D. G. (2002). Ushers and secretins: Channels for the secretion of folded proteins across the bacterial outer membrane. *Journal of Molecular Microbiology and Biotechnology,* 4 (1), 11-20.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences,* 76 (9), 4350-4354.

Tsen, G., Halfter, W., Kröger, S. and Cole, G. J. (1995). Agrin is a heparan sulfate proteoglycan. *Journal of Biological Chemistry,* 270 (7), 3392-3399.

Valdivia, R. H. (2008). < i> Chlamydia</i> effector proteins and new insights into chlamydial cellular microbiology. *Current Opinion in Microbiology,* 11 (1), 53-59.

Vallejo, L. F. and Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microbial Cell Factories,* 3 (1), 11-22.

van den Berg, B. (2010). Crystal structure of a full-length autotransporter. *Journal of Molecular Biology,* 396 (3), 627-633.

Van Putten, J. P. M., Duensing, T. D. and Cole, R. L. (2002). Entry of OpaA gonococci into HEp-2 cells requires concerted action of glycosaminoglycans, fibronectin and integrin receptors. *Molecular Microbiology,* 29 (1), 369-379.

Vandahl, B., Stensballe, A., Roepstorff, P., Christiansen, G. and Birkelund, S. (2005). Secretion of Cpn0796 from chlamydia pneumoniae into the host cell cytoplasm by an autotransporter mechanism. *Cellular Microbiology,* 7 (6), 825-836.

Vandahl, B. B., Pedersen, A. S., Gevaert, K., Holm, A., Vandekerckhove, J., Christiansen, G. and Birkelund, S. (2002). The expression, processing and localization of polymorphic membrane proteins in chlamydia pneumoniae strain CWL029. *BMC Microbiology,* 2 (1), 36-48.

Veiga, E., De Lorenzo, V. and Fernández, L. A. (2004). Structural tolerance of bacterial autotransporters for folded passenger protein domains. *Molecular Microbiology,* 52 (4), 1069-1080.

Veiga, E., Sugawara, E., Nikaido, H., De Lorenzo, V. and Fernández, L. A. (2002). Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. *The EMBO Journal,* 21 (9), 2122-2131.

Viegas, M., Martins, T., Seco, F. and Do Carmo, A. (2009). An improved and cost-effective methodology for the reduction of autofluorescence in direct immunofluorescence studies on formalin-fixed paraffin-embedded tissues. *European Journal of Histochemistry,* 51 (1), 59-66.

von Heijne, G. and Abrahmsen, L. (1989). Species-specific variation in signal peptide design. *FEBS Letter,* 244 , 439-446.

Vretou, E., Giannikopoulou, P. and Psarrou, E. (2001). Polymorphic outer-membrane proteins of chlamydophila abortus are glycosylated. *Microbiology,* 147 (12), 3303-3310.

Wagner, S., Bader, M. L., Drew, D. and de Gier, J. W. (2006). Rationalizing membrane protein overexpression. *Trends in Biotechnology,* 24 (8), 364-371.

Wang, C., Tang, J., Crowley-Nowick, P., Wilson, C., Kaslow, R. and Geisler, W. (2005). Interleukin (IL)-2 and IL-12 responses to chlamydia trachomatis infection in adolescents. *Clinical & Experimental Immunology,* 142 (3), 548-554.

Wear, M. A., Patterson, A., Malone, K., Dunsmore, C., Turner, N. J. and Walkinshaw, M. D. (2005). A surface plasmon resonance-based assay for small molecule inhibitors of human cyclophilin A. *Analytical Biochemistry,* 345 (2), 214-226.

Wehrl, W., Brinkmann, V., Jungblut, P. R., Meyer, T. F. and Szczepek, A. J. (2004). From the inside out–processing of the chlamydial autotransporter PmpD and its role in bacterial adhesion and activation of human host cells. *Molecular Microbiology,* 51 (2), 319-334.

Whorton, M. R., Bokoch, M. P., Rasmussen, S. G. F., Huang, B., Zare, R. N., Kobilka, B. and Sunahara, R. K. (2007). A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proceedings of the National Academy of Sciences,* 104 (18), 7682-7687.

Yang, J. P., Cirico, T., Katzen, F., Peterson, T. and Kudlicki, W. (2011). Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnology,* 11 (1), 57-65.

Yen, M. R., Peabody, C. R., Partovi, S. M., Zhai, Y., Tseng, Y. H. and Saier, M. H. (2002). Protein-translocating outer membrane porins of gram-negative bacteria. *Biochimica Et Biophysica Acta (BBA)-Biomembranes,* 1562 (1), 6-31.

Zaretzky, F. R., Pearce-Pratt, R. and Phillips, D. M. (1995). Sulfated polyanions block chlamydia trachomatis infection of cervix-derived human epithelia. *Infection and Immunity,* 63 (9), 3520-3526.

Zhang, Y., Taiming, L. and Liu, J. (2003). Low temperature and glucose enhanced T7 RNA polymerase-based plasmid stability for increasing expression of glucagon-like peptide-2 in< i> escherichia coli</i>. *Protein Expression and Purification,* 29 (1), 132-139.

Zhong, G., Fan, P., Ji, H., Dong, F. and Huang, Y. (2001). Identification of a chlamydial protease–like activity factor responsible for the degradation of host transcription factors. *The Journal of Experimental Medicine,* 193 (8), 935-942.

Zhu, Z., Gershon, M. D., Ambron, R., Gabel, C. and Gershon, A. A. (1995). Infection of cells by varicella zoster virus: Inhibition of viral entry by mannose 6-phosphate and heparin. *Proceedings of the National Academy of Sciences,* 92 (8), 3546-3550.

Zhukov, A., Schürenberg, M., Jansson, Ö., Areskoug, D. and Buijs, J. (2004). Integration of surface plasmon resonance with mass spectrometry: Automated ligand fishing and sample preparation for MALDI MS using a biacore 3000 biosensor. *Journal of Biomolecular Techniques: JBT*, 15 (2), 112-119.