

Characterisation of a Novel  
Molecular Tension Probe to  
Investigate Translocation of Colicin  
E9 across the Outer Membrane

Daniel George Louis Jones

MSc by Research

University of York

Biology

December 2016

## Abstract

Previous studies have hinted at the presence of a mechanical force that unfolds Colicin E9 (ColE9) at the cell surface, enabling it to be translocated into a target bacterial cell. This hypothesis has yet to be proven, though the Tol system is speculated to be the origin of this force. Numerous Förster resonance energy transfer (FRET) studies have been carried out to elucidate the mechanism of colicin translocation, a protein with implications for medicine. However, none of these single-molecule studies have been able to directly probe mechanical unfolding of ColE9 at the cell surface.

Here, a novel molecular tension probe (MTP) has been proposed, utilising a dsDNA duplex, labelled with a FRET acceptor-donor fluorophore pair, inserted between structural domains of ColE9. The force required to unzip the dsDNA duplex (10-15pN) has been measured by force spectroscopy to be sufficient to unfold colicin (Cocco, Monasson and Marko, 2002; Robinson, 2013; Vankemmelbeke *et al.*, 2013). Should an unfolding force be present, the MTP would observe it as a decrease in FRET efficiency. The proposed MTP model requires the synthesis of two ssDNA-protein chimeras. Described here are the results of purification and modification protocols, designed to engineer ColE9 domains into suitable components for the MTP. Progress has been made towards the generation of a final MTP, with additional modifications and revisions to the thiol coupling strategy planned.

Also described is a structural bioinformatics project exploring readily available protein data bank (PDB) database files belonging to the OMP families OmpF, BtuB and VDAC. Symmetry analysis of these families has discovered preferred patches of intermolecular interaction on the outer surface of their  $\beta$ -barrels. Investigations into these patches have suggested they may be conserved through evolution across these divergent OMP families, first arising in a distant common ancestor.

## List of Contents

Abstract .....	2
List of Contents .....	3
List of Figures .....	8
List of Equations .....	11
List of Tables.....	12
Acknowledgements .....	14
Declaration .....	15
Chapter 1: Introduction .....	16
1.1 Bacterial Membranes are a Selectively Permeable Barrier .....	16
1.2 Translocation across the Outer Membrane of Gram-Negative Bacteria .....	17
1.3 Bacteriocins: Agents of Bacterial Warfare.....	19
1.4 Colicin Translocation across the Bacterial Outer Membrane .....	21
1.5 Investigating the Mechanics of Colicin E9 Translocation .....	23
1.6 Aims and Objectives .....	26
Chapter 2: Materials and Methods .....	28
2.1 Materials.....	28
2.2 Methodology .....	28
2.2.1 Separation of Protein and DNA by Electrophoresis.....	28
2.2.2 Preparation of ssDNA Oligonucleotides .....	29
2.2.3 Protein-Encoding Plasmids and Transformation into Competent <i>E. coli</i> Cells ...	30
2.2.4 Over-Production of Recombinant Proteins in Bacteria.....	32
2.2.5 Extraction and Purification of Recombinant Proteins.....	32
2.2.6 Removal of His <sub>6</sub> -tag from R-Domain using Thrombin.....	34
2.2.7 Characterisation of ColE9 T <sup>1-83-TEV-H103A</sup> DNase and R-domain by ESI-MS.....	35
2.2.8 Production of pUC-ϕ600 Plasmid DNA to Test DNase Activity .....	35

2.2.9 Restriction Digest of pUC- $\phi$ 600 by EcoRI to Linearise DNA.....	36
2.2.10 Assays of DNA Degradation by ColE9 DNase Using Linear and Supercoiled pUC- $\phi$ 600 Plasmid.....	36
2.2.11 Quenching ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Activity with EDTA.....	37
2.2.12 Quantification of Free Thiols using Ellman's Reagent.....	37
2.2.13 Self-Coupling of Colicin E9 R-domain.....	39
2.2.14 Testing Bromomaleimide Labelling of ColE9 R-domain with Ellman's Reagent .....	40
2.2.15 Use of MALDI-MS to test for Bromomaleimide labelling of ColE9 R-domain.....	42
2.2.16 Labelling ColE9 R-domain with N-MBM and Cysteine .....	43
2.2.17 Coupling Fluorescent Tetrapeptide to ColE9 R-domain.....	44
2.2.18 Coupling ColE9 T-domain Linked 3' Thiol Oligonucleotide to ColE9 R-domain .....	45
2.2.19 Annealing of ColE9 R-domain linked 5' Thiol Oligonucleotide to ColE9 T-domain linked 3' Thiol Oligonucleotide .....	46
2.3 Methods for Structural Bioinformatics Investigation of Protein-Protein, Protein-Lipopolysaccharide and Protein-Detergent Interactions in Crystal Structures of Bacterial Outer Membrane Proteins .....	47
2.3.1 Mining protein data bank (PDB) data for monomeric OMP crystal contacts in WinCOOT .....	47
2.3.2 Mining the Crystal Contacts in Other Outer Membrane Protein Structures .....	50
2.3.3 Mining the Crystal Contacts of Mitochondrial Voltage-Dependent Anion Channels .....	50
2.3.4 Hydrophobicity Analysis.....	51
2.3.5 Amino Acid Composition and Location Analysis .....	54
Chapter 3: Production and Analysis of ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Construct .....	56
3.1 Introduction .....	56

3.1.1 Background to ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Construct and Incorporation in MTP Design .....	56
3.2 Results .....	58
3.2.1 Over-Production of ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Construct.....	58
3.2.2 Purification and Characterisation of ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> .....	60
3.2.3 Production and Purification of pUC- $\phi$ 600 Plasmid DNA.....	63
3.2.4 Testing Endonuclease Activity of ColE9 T1-83-TEV-DNaseH103A compared to wild-type ColE9 DNase .....	64
3.2.5 Single-Strand DNA Nicking Assay of ColE9 DNase Variants utilising supercoiled pUC- $\phi$ 600 plasmid DNA .....	67
3.2.6 Inhibiting Wild-Type ColE9 DNase and T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Endonuclease Activity with EDTA .....	69
3.3 Discussion .....	70
3.3.1 Purification and Characterisation of T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Construct.....	70
3.3.2 Assaying the DNase Inactivating H103A Mutation in the T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Construct .....	70
3.3.3 Future Work .....	71
Chapter 4: Developing a Thiol Coupling Protocol for the Molecular Tension Probe using Colicin E9 R-Domain .....	73
4.1 Introduction .....	73
4.2 Results .....	76
4.2.1 Transformation of <i>E. coli</i> cells with plasmid pREN1 .....	76
4.2.2 Purification and Characterisation of ColE9 R-domain.....	78
4.2.3 Removing the Purification Tag and Dimer population of R-domain.....	80
4.2.4 Initial Attempts to Self-Couple R-domain using Bromomaleimide.....	83
4.2.4.1 Analysing Self-Coupling of R-Domain.....	83
4.2.4.2 Labelling R-Domain with N-MBM using different label:protein molar ratios.	86
4.2.5 Testing N-MBM Labelling via Ellman's Test .....	87

4.2.5.1 Generation of a Standard Curve for Ellman's Reagent.....	87
4.2.5.2 Ellman's Test of N-MBM-labelled Uncleaved R-Domain .....	87
4.2.5.3 Labelling of SasG with N-MBM and subsequent Ellman's Tests .....	88
4.2.5.4 Returning to N-MBM Labelling of Uncleaved R-domain .....	89
4.2.6 Use of MALDI and LC-MS to confirm N-MBM labelling of R-domain .....	90
4.2.7 Visualising Thiol Coupling of R-domain using Fluorescent Tetrapeptide .....	92
4.2.8 Thiol Coupling R-domain to Oligo Components of the MTP .....	98
4.3 Discussion .....	102
4.3.1 Initial Attempts of N-MBM-Labelling and Self-Coupling of R-domain .....	102
4.3.2 Ellman's Tests and MALDI LC-MS to confirm N-MBM Labelling.....	104
4.3.3 Fluorescent Peptide Coupling of R-domain .....	105
4.3.4 Coupling of ssDNA Oligonucleotide Handles to R-domain.....	106
4.3.5 Future Work .....	106
Chapter 5: Investigating Protein-Protein, Protein-Lipopolysaccharide and Protein-Detergent Interactions in Crystal Structures of Bacterial Outer Membrane Proteins .	108
5.1 Research Aims.....	108
5.2 Introduction .....	108
5.3 Results .....	110
5.3.1 Interactions in Monomeric OMPs .....	111
5.3.2 Intermolecular Interactions in OmpA and OmpC .....	122
5.3.3 Intermolecular Interactions in Voltage-Dependent Anion Channels .....	129
5.4 Discussion .....	133
5.4.1 Overview of Symmetry Analysis .....	133
5.4.2 Hydrophobicity of PPI Patches .....	134
5.4.3 Future Work .....	135
Chapter 6: Conclusions and Future Work .....	138
6.1 Progress Towards a Functional Molecular Tension Probe.....	138

6.2 Identification of Putative Intermolecular Interaction Patches in Outer Membrane Proteins.....	139
6.3 Future Work .....	139
Appendix .....	141
A – Buffers and Reagents.....	141
B - Protein Construct Coding Sequences .....	143
I. ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> : Im9 His <sub>6</sub> .....	143
I. R-domain (Y324C, L447C), N-Cys (318C), C-Thrombin Cl.S-His <sub>6</sub> .....	143
List of Abbreviations.....	145
References .....	147

## List of Figures

Figure 1-1 – Colicin E9 3D Model and Domain Organisation .....	21
Figure 1-2 – Formation of the Translocon upon Colicin E9 binding to the Cell Surface .....	23
Figure 1-3 - Molecular Tension Probe for Detecting Mechanical Unfolding During ColE9 Translocation Across Outer Membrane .....	25
Figure 1-4 - DNA Duplex of Molecular Tension Probe .....	26
Figure 2-1 – Plasmid map of pREN1 encoding Colicin E9 R-domain .....	31
Figure 2-2 – Standard Curve of Ellman’s Test with DTT.....	39
Figure 2-3 – Reference residues for distance and angle measurements.....	49
Figure 2-4 – Hydrophobicity/side chain volume ( $\text{kcal mol}^{-1} \text{ \AA}^{-3}$ ) for each amino acid	54
Figure 2-5 – Diagram showing the protocol used to determine Z distances of interface residues.....	55
Figure 3-1 – 1 % (w/v) Agarose-TAE gel of pREN1/pPR1 Plasmid Purification.....	58
Figure 3-2 – Growth Curve of BL21*(DE3) bacterial cells transformed with pPR1 plasmid during induction with IPTG.....	59
Figure 3-3 – 12 % (w/v) SDS-PAGE gel analysis of T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> over-production in BL21*(DE3) bacterial cells .....	59
Figure 3-4 – Analysis of Fractions eluted from HisTrap Column .....	61
Figure 3-5 – Mobility Shift Analysis of T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> .....	62
Figure 3-6 – ESI-MS mass spectrum of purified ColE9 T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> .....	63
Figure 3-7 – Linearisation of pUC- $\phi$ 600 by EcoRI and 1 % (w/v) Agarose-TAE Gel Analysis .....	64
Figure 3-8 – 1 % (w/v) Agarose-TAE Gel of DNA Digestion Assay analysing ColE9 DNase Variants .....	65
Figure 3-9 – 1 % (w/v) Agarose-TAE Gel Analysis of Digestion Assay with 2 mM Mg <sup>2+</sup> .....	66
Figure 3-10 – 1 % (w/v) Agarose-TAE Gel Analysis of Digestion Assay with and without added Mg <sup>2+</sup> .....	67
Figure 3-11 – 1 % (w/v) Agarose-TAE Gel Analysis of Single-Strand DNA Nicking Assay for DNase variants.....	68

Figure 3-12 – 1 % (w/v) Agarose-TAE Gel Analysis of Endonuclease Activity for DNase variants with and without added EDTA .....	69
Figure 4-1 – Structure of R-Domain .....	75
Figure 4-2 – Diagram of R Domain-Oligo Conjugation utilising N-MBM.....	75
Figure 4-3 – Growth curve for <i>E. coli</i> BL21(DE3) bacterial cells transformed with pREN1 plasmid during induction with IPTG.....	77
Figure 4-4 – 15 % (w/v) SDS-PAGE analysis of R-domain over-production in <i>E. coli</i> BL21(DE3) bacterial cells.....	77
Figure 4-5 –Analysis of Eluted FPLC Fractions by OD <sub>280</sub> and SDS-PAGE .....	79
Figure 4-6 – ESI-MS mass spectrum of purified Cole9 R-domain .....	80
Figure 4-7 – 15% (w/v) SDS gel of R Domain pre- and post-Thrombin Cleavage .....	82
Figure 4-8 – ESI-MS mass spectrum of Thrombin-cleaved Cole9 R-Domain .....	82
Figure 4-9 – Analysis of Fractions eluted from Size Exclusion column .....	83
Figure 4-10 – 15 % (w/v) SDS-PAGE Analysis of Labelled R-domain Self-Coupling at Room Temperature.....	85
Figure 4-11 – SDS-PAGE Analysis of Labelled R-domain Self-coupling at 30 °C and 37 °C.....	85
Figure 4-12 – 15% (w/v) SDS-PAGE analysis of R-Domain Labelling with Bromomaleimide .....	86
Figure 4-13 –ESI-MS spectra of 0:1, 2:1 and 20:1 N-MBM labelling reactions.....	87
Figure 4-14 – Chemical structure of synthesised fluorescent tetrapeptide .....	92
Figure 4-15 – LC-MS analysis of synthesised fluorescent tetrapeptide.....	93
Figure 4-16 – 15 % (w/v) SDS Gel of Peptide and R-Domain Coupling Reaction.....	95
Figure 4-17 – 15% (w/v) SDS Gel of Coupling Reactions treated with β-ME.....	96
Figure 4-18 – 15% (w/v) SDS Gel of Coupling Reactions with Cysteine Blocker .....	96
Figure 4-19 – 15% (w/v) SDS gel of Peptide Coupled Samples Post Buffer Exchange .....	97
Figure 4-20 15 % (w/v) SDS gel of R-domain-Peptide Post-Cut-off with β-ME.....	97
Figure 4-21- ESI-MS of Uncleaved R-Domain Peptide Coupling Samples .....	100
Figure 4-22– 15% (w/v) SDS gel of R-Domain 3'Oligo ssDNA Coupling.....	101
Figure 4-23 – 15 % (w/v) SDS gel of R-Domain-Oligo mix Reacted with DNase I..	101
Figure 4-24- 15 % (w/v) SDS gel of R-Domain and 5'Oligo ssDNA Coupling .....	102
Figure 5-1 –Distance vs Angle of Mapped Crystal Contacts Relative to BtuB.....	111

Figure 5-2 – Histogram of distances between protein-protein interaction residues and the closest symmetry residue.....	112
Figure 5-3 – LPS binding site of FhuA structures 1QFF and 4CU4.....	113
Figure 5-4 – Intermolecular interaction patches of Monomeric OMPs .....	113
Figure 5-5 – Aromatic girdle residues of BtuB in relation to PPI patches.....	114
Figure 5-6 – Symmetry analysis for FrpB iron transporter .....	115
Figure 5-7 – Scatter plots showing the Z Distance of Patch Residues and their Residue Number.....	121
Figure 5-8 – Symmetry simulation of OmpA (1QJP).....	122
Figure 5-9 – Trimer-Trimer Interactions in OmpC (1OSM).....	123
Figure 5-10 – Plot of OmpC trimer contacts as distance vs angle relative to contacts in BtuB .....	124
Figure 5-11 – Interactions at OmpC Trimer Interface .....	124
Figure 5-12 – Interaction Patches in OmpC and BtuB.....	125
Figure 5-13 – Distance versus angle for crystal contacts in VDACS.....	130
Figure 5-14 – Relative PPI patches of BtuB (3RGM) and VDACS (2JK4) .....	131
Figure 5-15 – Location of aromatic girdle residues and intermolecular interaction patches in VDACS .....	131
Figure 5-16 – Interaction patches of VDACS superimposed onto OmpC .....	133
Figure 5-17– BtuB mutations to mimic Patch 2.....	136
Figure 6-1 – Location of Proposed Cysteine Mutation in Colicin E9 DNaseH103A .	140

**List of Equations**

Equation 2-1 – Calculation of the thiol concentration of a sample using the Ellman’s reagent test.....38

## List of Tables

Table 2-1 – List of ssDNA Oligonucleotides used in Thiol-Coupling Experiments ....	30
Table 2-2 – Terms and definitions for the $\beta$ -barrel surface hydrophobicity calculations .....	52
Table 2-3 – Calculated hydrophobicity contributions for each amino acid .....	53
Table 3-1 - PCR Primers for site-directed mutagenesis of the pRP1 plasmid to create the T <sup>1-83</sup> -TEV-DNase <sup>H103A/H127A</sup> construct .....	72
Table 4-1 – Ellman’s Test Results of R-Domain-N-MBM labelling samples .....	88
Table 4-2 – Ellman’s Test Results of SasG-N-MBM labelling samples .....	89
Table 4-3 – Ellman’s Test Results of R-Domain-N-MBM labelling samples .....	90
Table 4-4 - NCBI BLAST Search Results of R-Domain samples from MALDI LC-MS .....	91
Table 4-5 – Mascot Search Results of MALDI LC-MS run of R-Domain.....	91
Table 5-1 – Final calculated hydrophobicity values for the full $\beta$ -barrel surface and PPI patches.....	116
Table 5-2 – Amino Acid Composition of Monomeric OMP Interaction Patches.....	116
Table 5-3 – Difference in Amino Acid Composition between the Full Surface and Patches 1 and 2 of BtuB .....	117
Table 5-4 – Difference in Amino Acid Composition between BtuB and FhuA in Patch 1 .....	118
Table 5-5 – Difference in Amino Acid Composition between BtuB and FhuA in Patch 2.....	118
Table 5-6 – Difference in Amino Acid Composition between Head and Tail LPS contacts for FhuA in Patch 2 .....	119
Table 5-7 – Z Distances of BtuB Interaction patches .....	120
Table 5-8 – Hydrophobicity analysis of the OmpC protein-detergent contact patches .....	127
Table 5-9 - Final calculated hydrophobicity values for the membrane-facing $\beta$ -barrel surface and trimer interfaces of OmpC .....	127
Table 5-10– Amino Acid Composition of OmpC Trimer Interfaces .....	128
Table 5-11 –Heat Map of Amino Acid Composition in Trimer Interfaces of OmpC.	128

Table 5-12 – Heat Map comparing OmpC Trimer Interfaces and Monomeric BtuB Interaction Patches .....	129
Table 5-13 – Hydrophobicity calculations for entire membrane-facing surface and PPI patch 1 in VDAC (2JK4).....	132
Table 5-14 - Planned Mutations in Patch 1 of BtuB to Mimic Amino Acid Composition of Patch 2 .....	136
Table 7-1 Standard Buffers and Reagents used in experiments.....	141

## Acknowledgements

Firstly, I would like to thank my supervisors, Dr Christoph Baumann and Prof Marek Brzozowski, for giving me this opportunity, and their continued advice and support throughout the project. I offer Dr Baumann extra thanks for offering the position, and his dedication to the success of the project from beginning to end. I would also like to thank my TAP members, Dr. Laurence Wilson and Dr Marjan van der Woude, for their time and feedback during my TAP meetings. The guidance of these academics has been invaluable in completing my MSc by Research. I'd like to thank all the members of the Baumann lab, for helping me stay in a healthy mental state throughout the project. I would also like to thank the entirety of the L1, staff and students alike. Throughout my project they have been incredibly helpful, lending me reagents, answering my questions and almost anything in between. They have made me incredibly welcome for the year I spent there, and for that I am deeply grateful. Special thanks go to Rhian Gore, Herman Fung, Emily Flack, Ivan Gyulev and Rosie Leaman.

I would like to thank Dr Andrew Leech for his advice and expertise, as well as the ESI-MS runs he carried out to analyse my samples. His work was vital to the advancement of my project.

Dr Martin Fascione and his lab group in the Department of Chemistry were incredibly helpful, in suggesting the bromomaleimide conjugation protocol and the fluorescent tetrapeptide. I thank Robin Brabham for his guidance on synthesising the peptide. I would also like to thank Richard Spears for carrying out LC-MS analysis of my coupled samples. Everyone in the group made working in a different lab environment much easier and calming than it had any right to be.

Final thanks go to all my friends and family, who helped ease all my stresses and worries throughout the year. Special thanks go to Katie Rooke for providing cider as motivation during my write-up period. Their emotional support and encouragement kept me going for the full duration of the course.

**Declaration**

I declare that this thesis is a presentation of original work and I am the sole author. Exceptions to this are as follows: generation of plasmids pRP1 and pREN1 by the Baumann lab, ESI-MS analysis of protein samples by Dr. Andrew Leech, LC-MS analysis of R-domain samples by Richard Spears, MALDI MS analysis of R-domain samples by Dr. David Ashford, synthesis of N-methylbromomaleimide by Dr. Martin Fascione's group. All references to others work are held within the text. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

## Chapter 1: Introduction

### 1.1 Bacterial Membranes are a Selectively Permeable Barrier

Biological membranes are barriers of selective permeability found within living cells. Together with the cell wall, membranes make up the cellular envelope of bacterial cells. Bacteria utilise this envelope to prevent the entry of harmful components, whilst facilitating the transport of desired materials into the cell. In addition, they must export waste and toxic by-products of metabolism, such as reactive oxidative species from respiration. There are two main types of bacterial cell envelope, Gram-negative and Gram-positive, originally categorised using Gram's method (Gram, 1884).

Gram-positive envelopes consist of a singular plasma membrane and a thick, peptidoglycan cell wall of many layers (30-100 nm). In contrast, Gram-negative bacteria possess an additional outer membrane (OM) with thinner peptidoglycan layers (a few nanometres wide) sandwiched between the two membranes in a region known as the periplasm (Silhavy, Kahne and Walker, 2010). Unlike typical phospholipid-containing membranes, the Gram-negative OM is asymmetrical containing lipopolysaccharide (LPS) molecules in its outer leaflet. LPS forms a permeation barrier that hinders hydrophobic molecules from passing through the OM. The majority of in-use antibiotics are hydrophobic. As a result, the permeation barrier a significant factor in antibiotic resistance. Due to this knowledge, targeting LPS and their transporters has become a strategy for discovering new antimicrobials (Gu *et al.*, 2015).

Translocation across biological membranes is a tightly regulated process, often requiring multiple protein complexes to traverse the bilayer depending on the size of the molecule. This process is even more complex in Gram-negative bacteria, given that two membranes must be translocated to access the cytoplasm. Drug delivery across the OM is an important issue in the modern-day battle against antibiotic resistance. One approach has been to weaken the cell envelope through degradation of the OM utilising inhibitors of LPS biosynthesis (Brown, 2016). Another potential strategy is to exploit the translocation machinery itself to deliver a drug intracellularly. Miller *et al.* 2015 have shown that the known properties of OccD1, an OM transporter of basic amino acids in *Pseudomonas aeruginosa*, can be exploited to enable uptake of the carbapenem class of antibiotics. This group optimised carbapenem analogues to have

an increased uptake rate through this OM channel with the aim of increasing their efficacy (Isabella *et al.*, 2015). Bypassing the OM in this manner still leaves the inner membrane (IM) as an obstacle to the antibiotic. Furthering our fundamental understanding of the OM translocation process is important to fine-tune such delivery strategies.

## 1.2 Translocation across the Outer Membrane of Gram-Negative Bacteria

The OM is a formidable barrier to both hydrophilic and hydrophobic molecules, protecting the cell from undesirable compounds. However, the cell requires an uptake mechanism to allow the entry of desired substrates and ions to maintain metabolic function. Porins, protein barrels of anti-parallel  $\beta$ -sheets arranged to form a channel, perform this task. Channels are filled with water and are size-restrictive, acting as a method of selectivity. Some channel proteins, such as FepA, possess ‘plug domains’ that can be used to control the flow of transported material through switching between ‘open’ and ‘closed’ gated states (Ferguson and Deisenhofer, 2002).

OmpF, a trimeric porin found in *Escherichia coli* (*E. coli*) OMs, is a well-studied example of a  $\beta$ -barrel translocator. The trimer has 3 barrels, each consisting of 16 anti-parallel  $\beta$  strands to create its own separate lumen. OmpF is incredibly stable allowing for extensive crystallographic studies, and evidenced when it became the first membrane protein to have its atomic-resolution structure determined (Garavito *et al.*, 1983; Cowan *et al.*, 1992; Efremov and Sazanov, 2012). OmpF allows the flow of small hydrophilic molecules (e.g. sugars and amino acids) across the membrane (Bavoil, Nikaido and von Meyenburg, 1977). This ‘diffusion of nutrients’ that OmpF allows is of great ‘physiological importance’ to *E. coli* (H Nikaido, 2003). OmpF is the entry transporter for many antibiotics, including the  $\beta$ -lactam family of antibiotics, e.g. ampicillin (Ziervogel and Roux, 2013). This has led to investigations of OmpF as a means of improving antibiotic delivery to overcome bacterial resistance to these small molecules (Nestorovich *et al.*, 2002; Mahendran *et al.*, 2009).

In contrast to the non-specific diffusion that occurs through OmpF, BtuB is a porin specific for cobalamin (Vitamin B<sub>12</sub>). BtuB is a monomeric outer membrane protein (OMP) composed of two domains: a  $\beta$ -barrel domain consisting of 22 anti-parallel  $\beta$  strands, and an N-terminal, globular plug domain to control entry (Cherezov *et al.*,

2006). Unlike OmpF, passive diffusion is insufficient to translocate cobalamin across the OM. Instead, BtuB relies on the TonB system and calcium binding to actively transport cobalamin (Chimento, Kadner and Wiener, 2003).

Following cobalamin binding, BtuB interacts with TonB (an IM protein) at the periplasmic interface via the Ton box, a component of the hatch domain. Binding of the C-terminal end of TonB to the Ton box of BtuB results in an allosteric change of the substrate binding site, an obstructive plug domain, and that allows Vitamin B<sub>12</sub> to enter the periplasm (Ferguson and Deisenhofer, 2002; Noinaj *et al.*, 2010; Sikora and Cafiso, 2016). TonB powers this energy-dependent translocation process by harnessing the proton motive force (pmf) across the IM (Cadieux and Kadner, 1999).

The established consensus is that a force-transduction pathway applies mechanical force from TonB to BtuB to remove the plug domain and allow entry (Hickman *et al.*, 2017). How this energy is transduced across the periplasm from TonB, located in the IM, to BtuB in the OM is of some debate. TonB is hypothesised to carry "conformationally stored potential energy" (Postle and Kadner, 2003). One hypothesis is that the TonB linker region is able to shorten and lengthen by residue transition. This would allow energy to transfer to the OM via a conformational change, the resulting vertical displacement of TonB extending the width of the periplasm (Hickman *et al.*, 2017). Another hypothesis is that of a propeller model. A rotary motion initiated by the TonB-associated proteins ExbB and ExbD, and powered by pmf, allows BtuB to translocate cobalamin upon association with the propeller of TonB (Postle and Kadner, 2003).

The Tol-Pal system, much like TonB, spans across the periplasm. Tol-Pal is thought to be important for OM integrity, given that Tol<sup>-</sup> mutants compromise the OM and lead to leakage of periplasmic components (Cascales *et al.*, 2007). It consists of five core proteins: TolA, TolQ, TolR, TolB, and Pal. TolA, TolQ and TolR are transmembrane proteins located in the IM. This IM complex is able to harness the pmf, with TolQ/TolR forming a proton-conducting pore (Wojdyla *et al.*, 2015). TolB and Pal form a complex associated with the inner face of the OM, with TolB acting as a bridge between both TolA and OMPs (Bonsor *et al.*, 2007). TolQ and TolR transmit pmf-derived energy to TolA resulting in conformational changes within TolA (Cascales, Llobès and Sturgis, 2001). This conformational change is thought to be required to

form a TolA-Pal complex, allowing energy transduction to the OM (Cascales and Llobès, 2004). Regardless of mechanism, both the Tol and Ton systems serve to transduce energy from the IM to the OM. In the case of Tol, this system can be parasitized by certain bacterial proteins to facilitate their translocation across the OM (Pommier *et al.*, 2005).

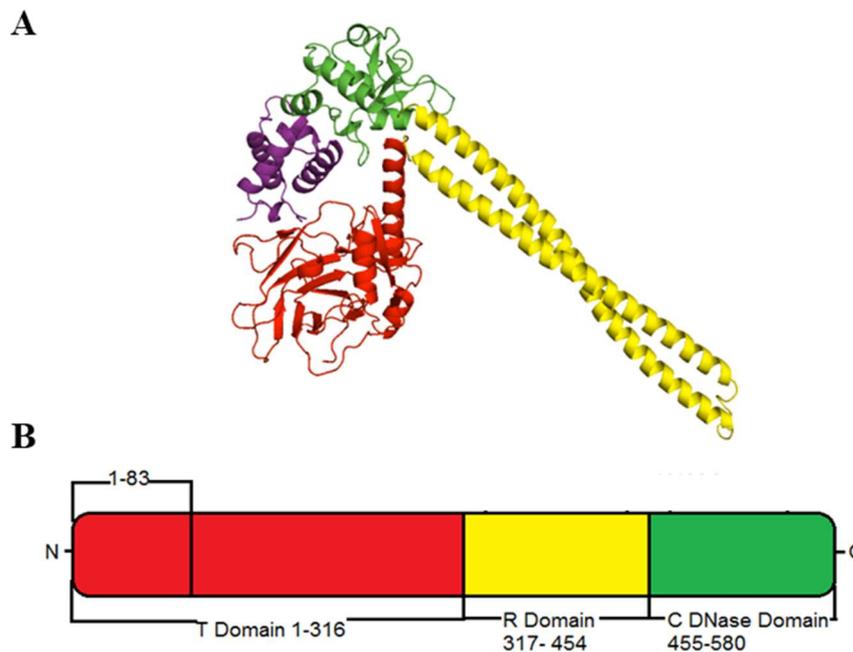
### 1.3 Bacteriocins: Agents of Bacterial Warfare

Bacteriocins are bacterial-produced, stress-induced, proteinaceous antibiotics that provide a competitive advantage to the strain that produces them. By killing rival, closely related strains, bacteriocins free up resources to help sustain the population. Their nomenclature is derived from the species that produces them, e.g. *E. coli* produce colicins. Colicins are prevalent within *E. coli* populations, according to one study “30 % of the natural *E. coli* population produces colicin” (Gordon and Riley, 1999).

Whilst their exact mechanisms vary across the groups, colicins translocate across the OM of the target bacterial cell and initiate a cytotoxic mechanism to kill the target cell. This process involves parasitizing the OM transport proteins of the target cell to circumvent the protective bacterial envelope. Colicins are produced and exported en masse to ensure their chance of effectiveness (Lazdunski *et al.*, 1998). Despite being released in large numbers, colicins are exceptionally lethal, only one successfully bound colicin molecule is necessary to kill a target cell (Papadakos, Wojdyla and Kleanthous, 2012).

There are two main classifications of colicins, Group A and Group B. Whether a colicin belongs to Group A or Group B is defined respectively by its use of either the Tol or Ton system to translocate across the bacterial cell envelope (Konisky, 1982). Colicins are further differentiated by the cytotoxic mechanism they utilise to kill their targets. Nuclease colicins translocate an enzymatic domain into the cytoplasm, where they proceed to degrade chromosomal deoxyribonucleic acid (DNA), ribosomal ribonucleic acid (RNA) or transfer RNA until the cell dies (Pommer *et al.*, 2001). Ionophoric colicins are translocated to the IM where they create a pmf-decoupling pore (Housden and Kleanthous, 2012). Groups of colicins are further subdivided by the receptors they specifically recognise in the OM.

Amongst the most studied group of colicins are the E colicins, belonging to Group A. E colicins bind to the OM receptor BtuB to initiate translocation (Zakharov *et al.*, 2012). One member of this group, Colicin E9 (ColE9), has been studied particularly extensively, with readily available structural data (Maté and Kleanthous, 2004; Klein *et al.*, 2016). ColE9 is an endonuclease colicin with a conserved domain structure (Figure 1-1) (Pommer *et al.*, 1998). ColE9 consists of 3 key domains each with an important functional role: the translocation (T) domain, the central receptor-binding (R) domain and the cytotoxic (C) domain. ColE9 parasitises FtsH, an IM AAA<sup>+</sup> ATPase/protease that translocates misfolded proteins to the cytoplasm. Partially unfolded ColE9 imitates a misfolded protein and is translocated into the cytoplasm by FtsH (Walker *et al.*, 2007). Following entry to the cytoplasm, the endonuclease within the C-domain digests DNA to kill the cell (Mosbahi *et al.*, 2002).



**Figure 1-1 – Colicin E9 3D Model and Domain Organisation**

**A** - 3D structure of Colicin E9 **B** - Linear view of domain organisation of Cole9 from N- to C-terminus. Cole9 consists of 3 domains: Translocation domain (T, red), Receptor domain (R, yellow) and Cytotoxic domain (C, green). Immunity protein (Im9, purple) is also visualised in the 3D structure, bound to the C-domain DNase. The residue boundaries are defined by residue number from the N-terminal end to the C-terminal. Residues 324 and 447 of the R-domain are mutated to cysteines to form a top lock. Residues 1-83 form the intrinsically unstructured region of the T-domain, essential for Cole9 translocation (Note: This region does not appear in the crystal structures).

#### 1.4 Colicin Translocation across the Bacterial Outer Membrane

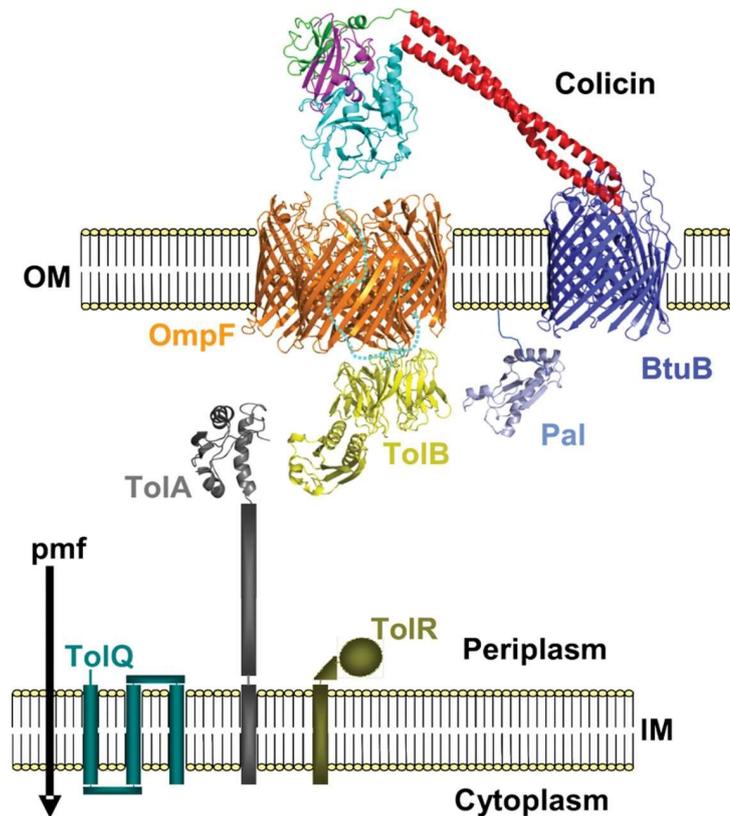
Previous studies have determined the main stages of colicin translocation across the OM of the target bacterial cell. First, the R-domain binds to the BtuB receptor, a transporter of the vitamin B<sub>12</sub>, forming a nanomolar affinity complex and initiating the translocation process (Housden *et al.*, 2005). The N-terminus of the T-domain (residues 1-83) possesses an Intrinsically Unstructured Translocation Domain (IUTD) that threads through the OmpF porin to allow Cole9 to translocate into the target cell (Figure 1-2) (Kleanthous, 2010; Housden *et al.*, 2013). Crystal structures have shown this also occurs with the IUTD of Cole3 (Yamashita *et al.*, 2008). The OmpF-Cole9-BtuB translocon is formed following lateral diffusion and collisional interaction within the OM (Spector *et al.*, 2010). OmpF is one of the most abundant OMPs on the cell surface (>10<sup>5</sup> copies of OmpF per cell) (Hiroshi Nikaido, 2003). This abundance increases the probability of an IUTD, from a BtuB-bound colicin, locating an OmpF

molecule to form the translocon complex given the confined lateral diffusion that occurs in the bacterial OM (Spector *et al.*, 2010; Rassam *et al.*, 2015).

As a Group A colicin, ColE9 utilises the Tol system to translocate across the OM (Carr *et al.*, 2000). Mutational analysis and isothermal titration calorimetry (ITC) has identified that a 13-residue binding epitope of the IUTD binds to TolB within the periplasm (Loftus *et al.*, 2006). This TolB-binding epitope (TBE) is thought to trigger the translocation of ColE9 across the membrane (Housden *et al.*, 2010). The pmf is required for the translocation of ColE9 (Papadakos, Wojdyla and Kleanthous, 2012). TolQ-R has been evidenced by mutational studies, and structural alignment to known flagellar motor proteins MotA and MotB, to transduce the pmf to TolA (Cascales, Lloubès and Sturgis, 2001). The energy transduced by the Tol system is likely used in some manner to translocate ColE9 across the OM.

Premature cytotoxicity in the colicin-producing bacterium is prevented by the production of a specific immunity protein that binds to the DNase region. The complex between the C-domain of ColE9 and its respective immunity protein (Im9) has an extremely high affinity ( $K_D = 9.3 \times 10^{-17}$  M at 25°C) (Wallis *et al.*, 1995). The ColE9-Im9 complex remains bound upon formation of the translocon between ColE9, BtuB and OmpF. Im9 release occurs during translocation and is dependent upon the pmf (Kim, Tarr and Penfold, 2014). Inactive mutants of TolA and TolB have been shown to prevent Im9 release from ColE9 (Vankemmelbeke *et al.*, 2009).

It is not known exactly how the cytotoxic domain traverses the OM, or whether protein unfolding facilitates this translocation process. One study states that the OmpF lumen is not wide enough to support the translocation of structured T- and C-domains of Colicin E3 (ColE3), and that the cytotoxic domain is unfolded prior to translocation (Zakharov *et al.*, 2006). The nuclease domain of Colicin E2 (ColE2), for example, is known to enter the cytoplasm whilst R-domain is still in contact with BtuB, the colicin molecule spanning the entire cell envelope (Duche, 2007). Since ColE2 and ColE9 use the same translocation machinery as ColE9, thus ColE9 may adopt a similar mechanism. Im9 release requires a global conformational change in all three domains of ColE9, which could be the unfolding event (Vankemmelbeke *et al.*, 2013). Evidence for the unfolding event occurring during ColE9 translocation has yet to be found.



**Figure 1-2 – Formation of the Translocon upon Colicin E9 binding to the Cell Surface**  
 Formation of the ColE9 translocon complex of ColE9-OmpF-BtuB. Immunity protein (Im9) is bound with a high affinity to the DNase domain, and dissociates upon translocation. The R-domain of ColE9 (red) binds to the BtuB receptor, the IUTD (cyan) then threads through the OmpF porin, and the TolB binding epitope (TBE) of ColE9 binds to TolB in the periplasm. This couples the translocon to the pmf harnessed by the Tol complex in the IM. This transduction of energy is thought to unfold the ColE9 at the cell surface, so it can thread through the lumen of the OmpF porin and into the periplasm. Image taken from (Bonsor, Meenan and Kleanthous, 2008).

### 1.5 Investigating the Mechanics of Colicin E9 Translocation

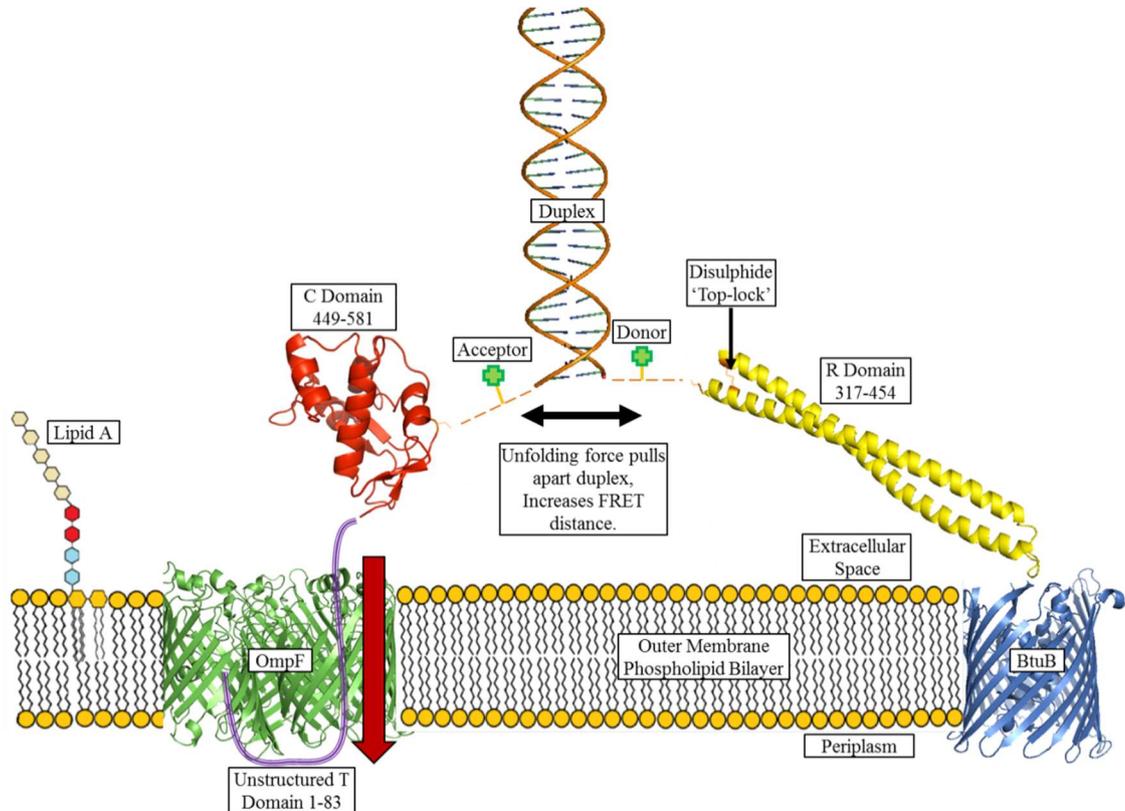
It is hypothesised that during translocation ColE9 crosses the Gram-negative envelope in an unfolded state. FRET studies have shown the distance between the T- and C-domains increases upon binding of ColE3 to BtuB, suggesting unfolding does take place (Zakharov *et al.*, 2008). If so, it must unfold at the cell surface via a mechanical unfolding force, the source of which is unknown. The pmf harnessed by the Tol complex was hypothesised to provide the energy necessary to unfold the colicin. Previous work has proven that a functional Tol complex and pmf is necessary for the dissociation of immunity protein (Im9) at the cell surface, and subsequent translocation of ColE9 across the OM (Vankemmelbeke *et al.*, 2009).

However, a mutational study has indicated that "colicins A and E2 do not use a mechanical energy source produced by TolQR for their import" (Llobès *et al.*, 2012). In an attempt to resolve this dilemma, we have designed a force probe to detect a mechanical unfolding force at the extracellular surface upon the binding of ColE9 to BtuB and formation of the translocon. To generate this probe, a novel molecular tension probe (MTP) composed of fluorescently labelled ColE9 T- and R-domains linked by a short DNA duplex has been proposed (Figure 1-3). Similar single-molecule probes have been used to detect mechanical forces within structural biology (Wang and Ha, 2013; Blakely *et al.*, 2014; Jurchenko and Salaita, 2015).

An acceptor and donor FRET fluorophore pair will be coupled to cysteine residues engineered at specific sites in the T- and R-domains and positioned on either side of the duplex (Figure 1-4). This will enable a decrease in FRET efficiency to report on unzipping of the DNA duplex. It is known that the force required to mechanically unzip a DNA strand (10-15 pN) is approximately equivalent to the force required to unfold ColE9 (< 20 pN) (Cocco, Monasson and Marko, 2002; Farrance *et al.*, 2013). If the FRET efficiency decreases during translocation of the ColE9 across the OM, then it would provide circumstantial evidence that a mechanical unfolding force can be generated at the OM surface (Bockelmann *et al.*, 2002; Cocco, Monasson and Marko, 2002; Robinson, 2013; Vankemmelbeke *et al.*, 2013). If evidence for an unfolding force exists, gene deletion bacterial strains (e.g. tolA<sup>-</sup>, tolB<sup>-</sup>) and pmf decoupling experiments can be used to determine if the Tol complex can be implicated in force generation.

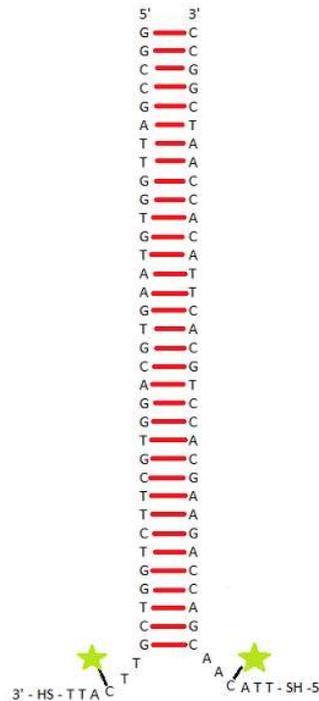
Previous work in the Baumann lab has determined that the T-domain of ColE9 is insoluble. To improve solubility, a fusion of the unstructured region of the T-domain (aa 1-83) and the C-terminal ColE9 DNase domain shall be used. Though the T-domain is not contiguous with the C-domain in ColE9, this fusion will allow the construct to be purified in a soluble state. Use of only the IUTD region of the T-domain should prevent the addition of the DNase from sterically hindering translocation. The engineered C-terminal cysteine required for chemical coupling of the single-stranded DNA shall be introduced within the C-domain, but as close as possible to residue 83 of the T-domain. To prevent the DNase domain from digesting the DNA duplex it must be inactivated. A previous study into the cytotoxic domain of

ColE9 introduced a H103A mutation to inactivate the DNase (Garinot-Schneider *et al.*, 1996). H103A shall be introduced into the DNase domain to inactivate it for use in FRET experiments as part of the MTP.



**Figure 1-3 - Molecular Tension Probe for Detecting Mechanical Unfolding During ColE9 Translocation Across Outer Membrane**

The proposed MTP to analyse the mechanics of ColE9 translocation. The probe consists of two protein-ssDNA oligonucleotide chimeras, with the ssDNA strands annealed to form a duplex. One half contains R-domain (ColE9 aa 317-454, cartoon structure in yellow), a coiled-coil domain that binds to BtuB (cartoon structure in blue) in the OM to initiate translocation. ColE9 R<sup>317-454</sup> contains a 'top-lock' disulphide bond between two engineered cysteines to prevent R-domain from translocating across the OM. The other half contains the IUTD region of T-domain (ColE9 aa 1-83, purple line) and the DNase-containing the C-domain (ColE9 aa 449-581, cartoon structure in red). The unstructured region threads through OmpF (cartoon structure in green) and, by an unknown mechanism, enables ColE9 to translocate across the OM (red arrow). Each ssDNA oligonucleotide is labelled with a fluorophore (green cross). These two fluorophores form a FRET pair, designed to detect an unfolding force significant enough to unwind and unzip the DNA duplex through a decrease in FRET efficiency. DNA unzipping would require force generation at the cell surface suggesting mechanical unfolding of ColE9 during OM translocation is possible.



**Figure 1-4 - DNA Duplex of Molecular Tension Probe**

The annealed DNA duplex at the centre of the MTP. The acceptor-donor FRET pair (green stars) consists of one fluorophore on each ssDNA oligonucleotide. Each oligonucleotide has an engineered thiol group for conjugation to respective ColE9 domains positioned at the 3' and 5' ends respectively. A force of 10-15pN is required to mechanically unzip the duplex, and should this occur, the distance between the acceptor-donor FRET pair will increase leading to a measurable decrease in FRET efficiency.

## 1.6 Aims and Objectives

The primary aim of this work is to characterise the components of the MTP, and make progress towards an assembled MTP. A fully-functional MTP would provide a tool to investigate the presence of an unfolding force at the cell surface during ColE9 translocation. This is incredibly important for the elucidation of the ColE9 translocation mechanism across the bacterial OM.

In addition, a structural bioinformatics project will investigate regions of promiscuous protein-protein interactions (PPIs) in OMPs. Lateral diffusion of OmpF and BtuB affects the likelihood of translocon formation. Any potential transient complexes of OmpF and BtuB would slow this rate of diffusion and lead to confinement. Structural databases can be mined for atomic resolution structures of OMPs, and then structural modelling tools can be used to identify and map potential interaction patches on the  $\beta$ -barrel surface of these OMPs.

To produce the novel and functional MTP, and study PPIs between OMPs, requires realisation of the following experimental objectives:

- Produce and characterise ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> and top-locked R-domain using established purification protocols (Chapter 3 and 4).
- Evaluate the effectiveness of the H103A mutation for inactivation of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (Chapter 3).
- Determine a thiol coupling protocol for the conjugation of protein and single-stranded deoxyribonucleic acid (ssDNA) components to form ssDNA-protein chimeras. R-domain will be used as a test protein for this ssDNA coupling protocol (Chapter 4).
- Determine regions of preferred contact for PPIs in OMPs, like BtuB, OmpF and the phylogenetically-related mitochondrial voltage-dependent anion channels, using structural bioinformatics tools (Chapter 5).

## Chapter 2: Materials and Methods

### 2.1 Materials

All buffers and solutions were prepared using ultra-pure deionised water (resistivity of 18.3 MΩcm) sourced from a water filtration system. Chemicals used in experimental buffers were reagent grade purity (> 98 % pure) unless otherwise specified. Bacterial strains used for transformations were *E. coli* BL21\*(DE3) and DH5α. Bacterial cells were made competent for transformation by an established protocol (Hanahan, Jessee and Bloom, 1991).

### 2.2 Methodology

#### 2.2.1 Separation of Protein and DNA by Electrophoresis

Gel electrophoresis was used to separate protein and DNA mixtures under native and denaturing conditions. Different gel compositions were used to assay both protein and nucleotide material: native agarose gels (1 % w/v) for DNA material, denaturing Sodium dodecyl sulfate (SDS)-polyacrylamide gels for protein material, and native polyacrylamide gels (12 % w/v) for analysing protein and DNA complexes. Sodium dodecyl sulphate poly-acrylamide gel electrophoresis (SDS-PAGE) analysis of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> used 12 % (w/v) denaturing polyacrylamide gels, while analysis of ColE9 R-domain used 15 % (w/v) denaturing polyacrylamide gels.

Agarose gel electrophoresis was utilised to analyse plasmid DNA and conduct assays of ColE9 DNase activity. Each run used 1 % (w/v) agarose gels cast in 1x tris-acetate-ethylenediaminetetraacetic acid (TAE) buffer (See Appendix) containing 1-2 µl of 10,000 X SYBRSafe dye (Invitrogen, California, USA). Nucleotide samples for analysis were loaded onto the gels alongside a DNA ladder (See Appendix). Electrophoresis was done in 1xTAE buffer at 50 milliAmps (constant current) until the sample bands had migrated 80-90 % into the gel (~2.5 hr). The SYBRSafe-stained DNA bands that resulted were viewed on a ultra-violet radiation (UV) transilluminator and photographed using a digital camera. Gels with free lanes after initial use were stored in 1x TAE buffer at 4 °C. Reused gels were first restained by soaking in Ethidium Bromide (2 µg/ml) in 1x TAE for 15 min, before destaining with deionised water for 20 min.

SDS-PAGE was utilised to identify fractions for pooling following protein purification, to test reduction of ColE9 R-domain thiol groups, and to evaluate thiol coupling reactions. Protein samples analysed by SDS-PAGE were first mixed with loading dye (See Appendix) and vortexed for 1 sec. SDS gels were run in 1x SDS running buffer (See Appendix). Any unused cast gels were stored in running buffer at 4 °C. Prior to loading in gels, samples were denatured by heating at 98 °C for 2 min. Once loaded alongside a respective protein ladder (See Appendix), electrophoresis was conducted at 150 V until the dye front reached the bottom of the gel (1-1.5 hr). Following this, gels were stained with Coomassie Blue stain (see Appendix), and then subsequently destained in deionised water. Gels were viewed and digitally photographed using a Gene Genius Bio Imaging System (SynGene), and GeneSnap software (version 6.00.19, SynGene).

Native poly-acrylamide gel electrophoresis (PAGE) gels were used to test ColE9 R-domain-ssDNA thiol coupling. Native gels were cast within a 1.5 mm gel cassette (NC2015, 10 well comb, Invitrogen). Tris/Tricine buffer (See Appendix) was used for Native PAGE. Samples were loaded with 4x sample buffer at a 1:3 ratio of buffer to sample. Bovine Serum Albumin (BSA) was analysed alongside ColE9 R-domain-ssDNA samples as a control; 1 in 200 dilutions of BSA were prepared from a 100 mg/ml stock for this purpose. Following sample loading, electrophoresis was done at 120 V (constant voltage) at 4 °C until the dye front reached the end of the gel (2.5+ hr). Gels were stained and destained in the same manner as for SDS-PAGE gels; however, this process generally took longer for native PAGE gels due to their greater thickness relative to SDS-PAGE gels.

### **2.2.2 Preparation of ssDNA Oligonucleotides**

The oligonucleotides used in thiol-coupling experiments are listed in Table 2-2. All sequences are given in a 5'-3' orientation and the oligonucleotides were ordered from Sigma-Aldrich.

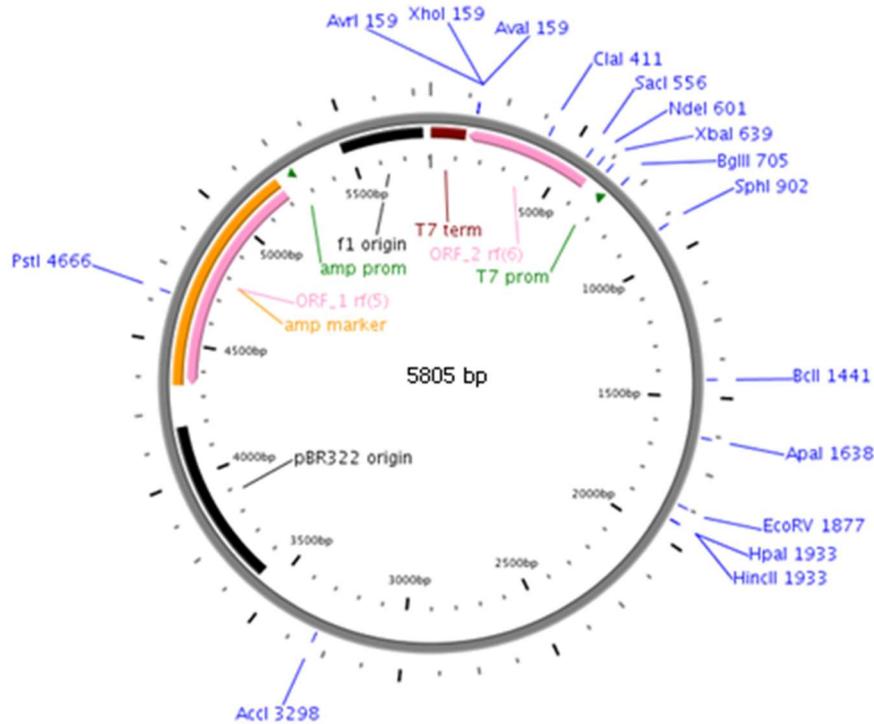
**Table 2-1 – List of ssDNA Oligonucleotides used in Thiol-Coupling Experiments**

Each oligonucleotide has its respective sequence and modifications listed. The oligonucleotides are designed to anneal to one another.

Oligonucleotide Name	Sequence 5' to 3'	Key Modifications
E9 T-domain linked oligonucleotide with 3' thiol	GGCCGATTGGTGTAAGTGCAGGTGCTTCTGGT CGTTCATT-[C <sub>3</sub> -Thiol]	3' thiol group with C <sub>3</sub> linker
E9 R-domain linked oligonucleotide with 5' thiol	[Thiol-C <sub>6</sub> ]- TTACAACGACCAGAAGCACCTGCACTTACACC AATCGGCC	5' thiol group with C <sub>6</sub> linker

### 2.2.3 Protein-Encoding Plasmids and Transformation into Competent *E. coli* Cells

The pRP1 and pREN1 plasmids encode ColE9 T<sub>1-83</sub>-TEV-DNase and ColE9 R-domain (see Appendix for coding sequence) (amino acids 317-454 with Y324C and L447C mutations which form a disulphide bond to stabilise the coiled-coil structure), respectively. The pRP1 plasmid was generated previously by Dr. Patrice Rassam. This was accomplished by excising the region encoding the E9 DNase domain (with H103A mutation) from pRJ353 via NcoI/XhoI sites, then ligating this DNA fragment into agarose gel purified pNGH25 plasmid (encoding E9 DNase domain without mutation) digested with the same restriction enzymes. The pREN1 plasmid encodes the ColE9 R-domain (amino acids 317-454 with Cys at position 2 (2C), and Y324C and L447C mutations) with a thrombin cleavable C-terminal hexa-histidine tag (His<sub>6</sub>). Both of these plasmids incorporated an ampicillin resistance gene for selection. Both plasmids also incorporated a lac controlled T7 RNA polymerase promoter upstream, designed to induce production of the protein of interest following addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the bacterial culture.



**Figure 2-1 – Plasmid map of pREN1 encoding Colicin E9 R-domain**

Plasmid map of the 5805bp long pREN1 showing: restriction sites highlighted in blue, promoter sites in green, selection markers in orange and Open Reading Frames (ORFs) in pink. ColE9 R-domain is encoded the ORF region labelled ‘ORF\_2 rf(6)’.

A published transformation protocol was utilised (Hanahan, Jessee and Bloom, 1991) and it is briefly described here. Competent *E. coli* cells (BL21\*(DE3) and DH5 $\alpha$ ) were mixed with 1  $\mu$ l of plasmid (1-10 ng) on ice (pRP1 with BL21\*(DE3) and pREN1 with DH5 $\alpha$ ). After incubation on ice for 20 min, the samples were heat-shocked at 42 °C for 45 sec, and then placed on ice for 2 min. Following this, 450  $\mu$ l of pre-warmed lysogeny broth (LB) medium (see Appendix) was added to each mixture and the samples were shaken (200 rpm) at 37 °C for 90 min. LB agar plates with 100  $\mu$ g/ml ampicillin (Table 2-1) were used to grow 100  $\mu$ l of transformed bacteria, spread by a sterile glass spreader, air-dried and then incubated agar-up at 37 °C. To test for successful transformation, plasmid was extracted from bacterial cells with a QIA Spin Mini prep kit (Qiagen, Netherlands), with all steps conducted at room temperature (20-22 °C). Plasmid DNA was kept on ice and analysed on a SybrSafe stained 1 % (w/v) agarose-TAE gel.

### 2.2.4 Over-Production of Recombinant Proteins in Bacteria

Overnight cultures of picked, transformed colonies were grown up in 15 ml LB media supplemented with 100 µg/ml ampicillin (LB-Amp). The optical density at 600 nm ( $OD_{600}$ ) of a 1-in-10 dilution of the culture was measured. An  $OD_{600}$  value  $>0.2$  confirmed successful culture growth allowing induction to proceed. Following this 5 ml of overnight culture was added to 500 ml of LB-Amp media. Cultures were incubated at 37 °C with shaking (120 rpm) and the  $OD_{600}$  was measured every 30-60 min (Figure 3-1). Once  $OD_{600}$  reached a value of  $\sim 0.5$ , 0.119 mg IPTG (Sigma Aldrich) was added to induce over-production of protein, along with an additional 50 mg ampicillin to maintain selection. The cultures were incubated at 37 °C with shaking (120 rpm) for 4 hr, with  $OD_{600}$  measured regularly (Figure 3-2), and 1 ml cell aliquots were taken every hour. Aliquots were pelleted by centrifugation (13,000 xg, 4 °C, 2 min, Thermo bench-top centrifuge). The supernatant was removed and the pellets stored at 4 °C for later analysis by SDS-PAGE.

The remaining culture was centrifuged (6,000 xg, 15 min, 4 °C, Sorvall Evolution SLC-3000 rotor). The pellet was resuspended using 30 ml of its respective supernatant, transferred to a 50 ml screw-cap Falcon plastic tube and centrifuged again (6,000 xg, 15 min, 4 °C, Sigma). The pellet was stored at -20 °C until successful induction was confirmed by SDS-PAGE. Pelleted cell aliquots were resuspended using a volume of 50 mM Tris-HCl (pH 7.5) buffer determined by their  $OD_{600}$  (20.8 µl Tris buffer per 0.1  $OD_{600}$ ). The samples for each time-point were prepared as follows for SDS-PAGE: 18 µl cells in Tris buffer, 2 µl 1 mM Dithiothreitol (DTT) in deionised water, and 10 µl 5x SDS loading dye (without DTT). Samples were analysed on SDS-PAGE gels to confirm recombinant protein over-production.

### 2.2.5 Extraction and Purification of Recombinant Proteins

Once production of the recombinant protein was confirmed, each pellet was resuspended in 35 ml of ice-cold lysis buffer. The suspended pellets had 350 µl of 100 mM phenylmethylsulfonyl fluoride (PMSF, prepared in ethanol) added and were sonicated to lyse the cells (6 x 15 sec pulses at power level 7 with 30 sec pause between pulses, equivalent to 87 Watts). A few DNase I (Sigma-Aldrich) flakes were

then added to the lysed cell suspension. After incubation on ice for 30 min with regular mixing, the resulting lysed pellet was centrifuged (48,000  $\times$ g, 4 °C, 30 min, Sorvall). Afterwards, another 350  $\mu$ l of 100 mM PMSF was added to the supernatant, it was then filtered through a 0.45  $\mu$ m syringe filter, and loaded using pump A of the fast protein liquid chromatography (FPLC) machine (BioLogic Duo-Flow, BioRad) onto an affinity purification column.

ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> was purified using a co-expressed His<sub>6</sub> tagged immunity protein (Im9) to bind the E9 DNase domain. A Ni<sup>2+</sup> affinity column (1 ml, GE Biosciences) was used for purification, charged with NiSO<sub>4</sub> prior to attachment to the FPLC machine. Following equilibration with wash buffer (see Appendix), the full 35 ml was loaded via pump A. Post-loading, the column was flushed with wash buffer (See Appendix) and the flow-through collected. Bound ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> was eluted under denaturing conditions introduced by an increasing gradient of guanidine HCl elution buffer (0 M - 6 M). Absorbance at 280 nm (OD<sub>280</sub>) was measured during elution to monitor protein content of the flow-through. Eluted fractions (1 ml) were collected, then subsequently analysed by SDS-PAGE to confirm ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> purification.

ColE9 R-domain was purified using a C-terminal His<sub>6</sub> tag which was linked to the protein by a peptide containing a thrombin cleavage site. Buffers used for R-domain purification contained reducing agent (5 mM  $\beta$ -mercaptoethanol) to prevent formation of dimers between incorporated thiol groups. A NiSO<sub>4</sub>-charged Ni<sup>2+</sup> affinity column was used again to purify by FPLC. Following equilibration with wash buffer (see Appendix), the full 35 ml was loaded as before. Post-loading, the column was washed, with flow-through collected. Bound ColE9 R-domain was eluted using a gradient of imidazole elution buffer (10 mM to 500 mM). As done previously, eluted fractions (1 ml) were collected and analysed by SDS-PAGE to confirm ColE9 R-domain purification. Once identified, protein-containing fractions were pooled to form stock solutions. Stocks were transferred into storage buffer (see Appendix) by membrane dialysis (12-14 kDa molecular weight cut-off (MWCO), Spectra/Por) with stirring at 4 °C overnight.

Protein concentrations within buffer solutions were calculated using the Beer-Lambert equation ( $A = \epsilon c l$ ; where  $A$  = absorption at given wavelength,  $\epsilon$  = extinction

coefficient at wavelength of absorption measurement,  $c$  = protein concentration, and  $l$  = cuvette pathlength) to correlate absorbance ( $A$ ) with concentration (Swinehart, 1962). Absorbance measurements at a wavelength of 280 nm ( $A_{280}$ ) were recorded for triplicate samples using NanoDrop™ ND-1000 spectrophotometer (Thermo Scientific), NanoDrop™ 1000 measurement software (version 3.7.1, Thermo Scientific), and 1.2  $\mu$ l of protein sample. Extinction coefficients ( $\epsilon$ ) for each protein were generated using the ExPASy ProtParam web tool (<http://web.expasy.org/protparam/>). The  $\epsilon$  values used for concentration calculations were as follows: ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>,  $\epsilon = 33,460 \text{ M}^{-1}\text{cm}^{-1}$ ; ColE9 R-domain post thrombin cleavage,  $\epsilon = 6990 \text{ M}^{-1}\text{cm}^{-1}$ . Where necessary, stock solutions were concentrated using spin concentrators (Vivaspin 500, Sigma Benchtop Centrifuge, 15,000  $xg$ , 10 °C). All protein stock solutions were stored at 4 °C.

### 2.2.6 Removal of His<sub>6</sub>-tag from R-Domain using Thrombin

A thrombin cleavage site was used to remove the His<sub>6</sub>-tag to ensure it had no effect on future thiol coupling experiments. 10 mg of protein (3.14 ml of 3.18 mg/ml) was digested with 20  $\mu$ l (1 Unit per  $\mu$ l) of Human Thrombin (Thrombin Cleavage Kit, GE Biosciences), made up to a 5 ml reaction with storage buffer. Thrombin cleavage was conducted at 20 °C for 16 hr. A small population of thiol-modified Col E9 R-domain forms into a dimeric form after 24 hr under oxidizing conditions. The dimer population reduces thiol coupling efficiency, necessitating removal for future experiments. Cleavage product was reduced with 1.8  $\mu$ l of 13.7 M  $\beta$ -mercaptoethanol ( $\beta$ -ME) to remove the dimer population, and subsequently dialysed into 50 mM Tris-HCl (pH 7.5) buffer (12-14 kDa MWCO, Spectra/Por). The dimer population, and residual thrombin from the cleavage reaction, was removed using size exclusion column chromatography (Superdex 75, HR26/60).

Thrombin-cleaved R-domain was loaded onto the column via injection loop after passing through a 0.22  $\mu$ m syringe filter and analysed using a flow rate of 3 ml/min (see Appendix). OD<sub>280</sub> was measured to track the presence of protein in the fractions. 15  $\mu$ l aliquots of fractions were analysed by SDS-PAGE to confirm the presence of protein. ColE9 R-domain fractions were pooled and the concentration measured as previous (Section 2.2.5). This stock was stored at 4 °C for use in later thiol coupling experiments.

### 2.2.7 Characterisation of Cole9 T<sup>1-83</sup>-TEV-H103A DNase and R-domain by ESI-MS

Electrospray ionisation mass spectrometry (ESI-MS) was used to confirm the mass of purified and chemically-modified proteins. Protein samples were removed from buffers that negatively affect ESI-MS (phosphate and Tris-containing buffers) by membrane dialysis (12-14 kDa MWCO, Spectra/Por) or buffer exchange via spin concentrators (Vivaspin 500, 12,000 MWCO). For dialysis, the sample was immersed in 1L of the desired buffer at 4°C overnight, or for 4 hr at room temperature. When using spin concentrators, the samples were pipetted into the spin column and centrifuged (15,000 xg, 10 °C, Thermo Benchtop) to reduce the volume to ~100 µL. The samples were then made up to the original volume using the desired buffer, and the process repeated at least 3 times to ensure complete buffer exchange. Samples were buffer exchanged by these methods into 25 mM ammonium acetate (pH 6.8) to allow for effective ionisation during ESI-MS. Positive-mode ESI-MS was performed by Dr. Andrew Leech in the Molecular Interactions lab of the Biology Technology Facility.

ESI-MS was conducted on a Waters LCT Premier XE system using MassLynx 4.1 software. The system was calibrated with sodium formate solution and calibration verified/corrected with horse heart myoglobin (16951.5 ± 1.5 Da). Millipore C18 ZipTips were used to purify samples following the manufacturer's protocol, except that formic acid was used in place of trifluoroacetic acid, and elution was into 1:1 acetonitrile: water containing 0.1 % (v/v) formic acid. Samples prepared in ammonium acetate or low concentration Tris buffers were diluted directly in the acetonitrile-water-formic acid solvent. Following loading by syringe, data was collected for 3 min over a scan range of 200 - 2000 m/z.

### 2.2.8 Production of pUC-ϕ600 Plasmid DNA to Test DNase Activity

The pUC-ϕ600 plasmid (3933 bp) is pUC19 derived and incorporates the Φ13 T7 RNA polymerase promoter region of T7 DNA (Cross, 2014). 50 µl of *E. coli* DH5α cells were transformed with 2 µl of pUC-ϕ600 plasmid (1 µg/ml). Following 1 hr of shaking (200 rpm), cells were spread onto LB-Amp plates and grown as described above in section 2.2.4. Single colonies were used to inoculate 10 ml volumes of LB-Amp which were grown overnight at 37°C with shaking (200 rpm). The cultures were pelleted by centrifugation (5,000 xg, 4 °C, 15 min, Sigma bench-top centrifuge).

Plasmid DNA was extracted from the cell pellets using a QIA Spin Mini prep kit (Qiagen, Netherlands), with all steps conducted at room temperature. Plasmid DNA was eluted in nuclease-free water and its concentration determined using a NanoDrop™ spectrophotometer. Final concentrations ranged from 289.9 ng/μl to 346.9 ng/μl. The product was kept on ice and analysed on a SybrSafe-stained 1 % (w/v) agarose-TAE gel.

### 2.2.9 Restriction Digest of pUC-ϕ600 by EcoRI to Linearise DNA

To gain an effective understanding of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> activity, its ability to digest both linearised and supercoiled pUC-ϕ600 plasmid DNA was assayed. To linearise pUC-ϕ600 the following reaction mixture was used: 42.5 μl of 90.4 ng/μl pUC-ϕ600 plasmid DNA (76.84 ng/μl final concentration), 5 μl Buffer H (Promega), 0.5 μl acetylated BSA (1 mg/ml; B8894, Sigma-Aldrich) and 2 μl 12 U/μl EcoRI endonuclease (Promega). The reaction mixture was gently mixed by pipetting, vortexed for 1 sec, centrifuged briefly, and then incubated at 37 °C overnight. Linearised pUC-ϕ600 plasmid product was purified using a Wizard® SV Gel and PCR Clean-Up System (Promega), then the concentration was determined by spectrophotometry, and its purity was analysed on a SybrSafe-stained 1 % (w/v) agarose-TAE gel to confirm linearisation.

### 2.2.10 Assays of DNA Degradation by ColE9 DNase Using Linear and Supercoiled pUC-ϕ600 Plasmid

DNase assays were done using wild-type ColE9 DNase (1.53 mg/ml) and ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> fusion (1.42 mg/ml) with either linearised or supercoiled pUC-ϕ600 DNA as the substrate. The following reaction mixture (final total reaction volume = ~15 μl) was used: 2 μl of linearised pUC-ϕ600 plasmid DNA (final concentration: ~16.4 ng/μl), 3 μl 50 mM MgCl (final concentration: 10 mM Mg<sup>2+</sup>), 10 μl of T<sup>1-83</sup>-TEV-H103A DNase or ColE9 DNase with varying protein amounts (0.1, 1, 2.5, 5 or 10 μg), then made up to the final volume (~15 μl) with 200 mM Tris-HCl (pH 7.5), 0.1 μl BSA (final concentration: ~1 mg/ml) and 0.2 μl 10 mM DTT (final concentration: ~0.13 mM). Stocks of 200 mM Tris-HCl and 50 mM MgCl<sub>2</sub> were filtered through 0.22 μm Millipore syringe filters before use.

Unless otherwise stated, reactions were run at 37 °C for 30 min with 3 µl of 0.5 M ethylenediaminetetraacetic (EDTA) used to quench the reaction. A negative control without Cole9 DNase added was also run under the same conditions. 10 µl of each reaction and 2 µl of DNA loading dye were run on a SybrSafe-stained 1 % (w/v) Agarose gel. For reference, 6 µl of DNA ladder (GeneRuler) was run alongside each Cole9 DNase variant. Following the gel electrophoresis, agarose-TAE gels were illuminated on a transilluminator to view the DNA bands.

The expected concentration of  $Mg^{2+}$  in the buffers used for TIRF microscopy is 2 mM. A functioning MTP will require the DNase domain to be inactive at this  $Mg^{2+}$  concentration. To determine T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> activity at 2 mM  $Mg^{2+}$ , another assay was run at this  $Mg^{2+}$  concentration using the exact same reaction mix as previously described.

#### **2.2.11 Quenching Cole9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Activity with EDTA**

The impact of EDTA on the activity of both Cole9 DNases was assayed to assess whether or not it could be used to protect oligonucleotide components during MTP formation and usage. The reaction mix was exactly the same as for previous Cole9 DNase activity assays, except the DNase per reaction was 10 µg. For both Cole9 DNase variants, three reactions were carried out: incubation at room temperature with EDTA, incubation at 37 °C with EDTA, and incubation at 37 °C without EDTA. All reactions were incubated for 25 min. Reactions were quenched with 5 µl of 0.5 M EDTA (final concentration: 125 mM). A negative control replacing Cole9 DNase with background buffer was also done for each variant. Reactions were analysed on a SYBRSafe stained 1 % (w/v) agarose-TAE gel as described above.

#### **2.2.12 Quantification of Free Thiols using Ellman's Reagent**

Analysing the change in absorption at a wavelength of 412 nm ( $A_{412}$ ) upon addition of Ellman's reagent (5, 5'-dithio-bis-(2-nitrobenzoic acid, DTNB)) allows for the quantification of free thiols within a sample (Equation 2-1) (Riener, Kada and Gruber, 2002). The Ellman's reagent-based assay is an established analytical technique for quantifying free thiol groups, such as those within cysteine residues of proteins (Riener, Kada and Gruber, 2002). DTNB (Sigma) was prepared as fresh 10 mM stocks

in the same buffer (not water) as the sample for each assay. A standard curve was generated using fresh DTT as the thiol-containing sample.

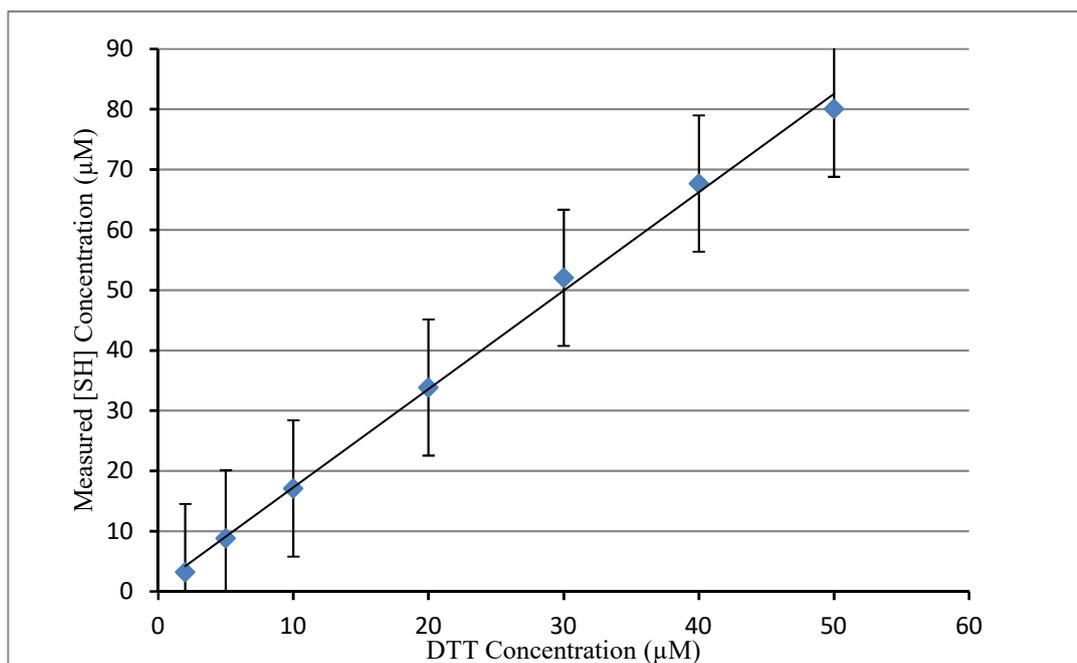
A range of DTT concentrations (2, 5, 10, 20, 30, 40 and 50  $\mu\text{M}$ ) were prepared in a final volume of 800  $\mu\text{l}$  in 1 cm-pathlength plastic cuvettes using 1 M DTT in deionised water and 50 mM Tris-HCl (pH 7.5) / 300 mM NaCl buffer. For each condition, a negative control was prepared using 820  $\mu\text{l}$  of buffer with no reagent added. Each sample to be tested had 20  $\mu\text{l}$  of fresh 10 mM Ellman's reagent added, and was mixed by inverting several times. After 5 min the  $A_{412}$  value ( $A_{412S}$ ) was measured by NanoDrop™ spectrophotometer. A reagent control ( $A_{412R}$ ) of 800  $\mu\text{l}$  Tris buffer and 20  $\mu\text{l}$  reagent, and sample controls ( $A_{412P}$ ) of 800  $\mu\text{l}$  Tris buffer and 20  $\mu\text{l}$  DTT (2-50  $\mu\text{M}$ ) were also ran. Each condition was assayed three times in total to allow for calculation of an average free thiol concentration ( $[\text{SH}]$ , Figure 2-4).

$$[\text{SH}] = (A_{412S} - A_{412R} - A_{412P}) / ((\Delta\epsilon_{412} \times l) \times (v/1 \text{ ml}))$$

$$\Delta\epsilon_{412} = 14,150 \text{ M}^{-1}\text{cm}^{-1} \quad l = 1 \text{ cm} \quad v = 0.8 \text{ ml}$$

**Equation 2-1 – Calculation of the thiol concentration of a sample using the Ellman's reagent test**

The equation used to calculate free thiol concentration within a protein sample.  $A_{412S}$  = sample absorbance at 412 nm post-Ellman's reagent addition.  $A_{412R}$  = absorbance at 412 nm of the reagent control.  $A_{412P}$  = absorbance at 412 nm of the protein control.  $\Delta\epsilon_{412}$  = extinction coefficient of Ellman's reagent at 412 nm.  $l$  = path length of the sample cuvette.  $v$  = total volume of reaction. Equation and extinction coefficient taken from Riener, Kada and Gruber (2002).



**Figure 2-2 – Standard Curve of Ellman's Test with DTT**

Measured thiol (SH) concentration of DTT sample by Ellman's test from change in absorbance at 412 nm corresponding to the actual concentration in  $\mu\text{M}$ . The curve has been fit using linear regression and error bars denote standard error of triplicate measurements.

$$y = 1.6323x + 0.9321$$

$$R^2 = 0.9974$$

### 2.2.13 Self-Coupling of Colicin E9 R-domain

For the MTP to be functional, the chemical coupling protocol using the thiol group modification must be viable. To test the coupling reaction, we attempted to couple two ColE9 R-domain molecules with a bromomaleimide bioconjugate. MTP coupling protocols were derived from an established bioconjugation technique (Smith et al, 2010). Thrombin-cleaved ColE9 R-domain ( $\sim 37.2 \mu\text{M}$ ) in 50 mM Tris-HCl (pH 7.5), 300 mM NaCl buffer was used as test material. N-Methyl-bromomaleimide (N-MBM), synthesised by Dr. Martin Fascione's lab (Department of Chemistry, University of York) was made up to a 2.6 mM stock in dimethylformamide (DMF).

The R-domain was first concentrated to  $\sim 150 \mu\text{M}$  with a spin concentrator (Vivaspin, 5000 Da MWCO, 15,000 xg, 10 °C). The sample was then dialysed (12-14 kDa MWCO, Spectra/Por) into 1 L 100 mM sodium phosphate (14.23 ml 1 M  $\text{NaH}_2\text{PO}_4$ , 85.75 ml 1 M  $\text{Na}_2\text{HPO}_4$ , pH 8.0), 150 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP)-HCl buffer overnight. Post-dialysis, the R-domain

stock was split into two separate stocks, one for N-MBM coupling, and the other for later self-coupling. As with previous samples, stock concentrations were calculated from the Beer-Lambert equation, with required absorbance measurements recorded by NanoDrop™ spectrophotometer. If R-domain stock concentration was found to be < 124 µM, it was spin-concentrated again to ~124 µM.

N-MBM (2.6 mM, made up in DMF) was added to reduced R-domain stock (~124 µM) at a 1:1 label:protein molar equivalency ratio. It was ensured that DMF solvent was never > 5 % (v/v) of the total reaction volume to prevent R-domain denaturation. The mix was vortexed for 1 sec and maintained at 0 °C for 1 hr. The reaction was dialysed (12-14 kDa MWCO, Spectra/Por) into 1 L of fresh buffer (without TCEP) overnight to remove free N-MBM. 10 µl aliquots of R-domain, pre- and post-labelling, were mixed with 3 µl of 5x loading dye (No DTT) analysed by SDS-PAGE (15 % (w/v) gel) alongside 10 µl of PAGE ladder. A molecular weight shift of +110 g/mol was indicative of successful labelling. For a more accurate assessment of modification, the sample was dialysed into Ammonium Acetate buffer and analysed by positive-mode ESI-MS by Dr Andrew Leech (Waters XE, see Appendix for details).

To investigate bromomaleimide labelling efficiency a variety of conditions were assayed: different incubation times (2 hr, 6 hr, 18 hr), increased incubation temperatures (0 °C, 20 °C, 30 °C), addition/removal of reducing agents (+/- TCEP, DTT), and varying the label:protein molar equivalency (0:1, 2:1, 20:1). After confirming the reduced state of R-domain stock by electrophoresis (10 µl protein + 3 µl 5x loading dye with no added DTT) self-coupling of R-domain was attempted. Self-coupling was initiated by addition of reduced R-domain to N-MBM-labelled R-domain at a 1:1 molar equivalence ratio. The reaction was vortexed for 1 sec and kept at 0 °C for 2 hr. Products were analysed by SDS-PAGE (15 % (w/v) gel), with the most promising results sent off for further characterisation by positive-mode ESI-MS.

#### **2.2.14 Testing Bromomaleimide Labelling of Cole9 R-domain with Ellman's Reagent**

Analysing the change in  $A_{412}$  upon addition of Ellman's Reagent (DTNB) allows for the quantification of free thiol within a sample (Equation 2-1). This was done to confirm labelling of R-domain with N-MBM at the N-Cys (2C) site without the need

for ESI-MS. Fresh 10 mM DTNB stock in sample buffer was prepared for each test. A standard curve was previously generated using DTT as the thiol-containing sample to be tested (Chapter 2.1.11).

The standard protocol used for the Ellman's tests was as follows. A range of sample concentrations was made up to a final volume of 800  $\mu$ l each in 1 cm-pathlength plastic cuvettes. For each condition a respective control was made up to 820  $\mu$ l with no reagent added ( $A_{412P}$ ). Test samples had 20  $\mu$ l of fresh 10 mM DMTB added, then they were mixed by inversion, and after 5 min had their  $A_{412}$  value measured by spectrophotometry ( $A_{412S}$ ). A reagent control of 800  $\mu$ l Tris buffer and 20  $\mu$ l reagent was also run ( $A_{412R}$ ) for each experiment. Where possible, each condition was repeated twice and used to provide an average for [SH] calculation.

With a standard curve established to confirm Ellman's reagent efficacy, the assay was used to test proteins with thiol modifications. To conserve purified protein, smaller reaction volumes were used and  $A_{412}$  values recorded by NanoDrop™ spectrophotometer. SasG, a thiol-containing protein, was used to check the viability of the Ellman's test. SasG and reduced SasG stocks (20 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM EDTA, +/- 2.5 mM DTT) were reacted with N-NBM (100 mM in DMF) at a 20:1 molar equivalency ratio of label:protein. Two different incubation conditions, 20 °C and 30 °C, were tested for both reduced and unreduced SasG labelling reactions. After a 1 hr incubation, the reactions were dialysed overnight at 4 °C (12-14 kDa MWCO, Spectra/Por) into 1 L of fresh buffer to remove free label. Ellman's test reaction mixes for SasG were prepared as follows: 5  $\mu$ l labelled SasG, 5  $\mu$ l Tris buffer and 2  $\mu$ l 10 mM DTNB in Tris buffer. DTNB was added last and the sample was mixed by pipetting. After 5 min, 1.2  $\mu$ l of reaction had its  $A_{412}$  measured by NanoDrop™ spectrophotometry to determine [SH].

The [SH] of unlabelled, reduced R-domain and N-MBM-labelled R-Domain was then determined to confirm labelling. A 1 ml aliquot of uncleaved R-domain (~154  $\mu$ M) was treated with 13.3  $\mu$ l 75 mM DTT to reduce any dimer present and activate the thiol group. After incubation for 1 hr at room temperature, the sample was dialysed at 4 °C (12-14 kDa MWCO, Spectra/Por) into 1 L 100 mM sodium phosphate (pH 7.0) / 300 mM NaCl overnight to remove DTT. This prevented thiol groups in the reducing agent from producing false positive values during measurement of  $\Delta A_{412}$ . After

determining the concentration of R-domain ( $\sim 141 \mu\text{M}$ ), 450  $\mu\text{l}$  of sample was labelled with 12.69  $\mu\text{l}$  N-MBM (100 mM in DMF) at a 20:1 label:protein molar ratio. After vortexing for 1 sec, the reaction was incubated at 30 °C for 6 hr, and then dialysed overnight at 4 °C (12 -14 kDa MWCO, Spectra/Por) into 1 L of fresh buffer to remove free label.

As with previous samples, a fresh 10 mM Ellman's stock was made up in R-domain sample buffer. The reaction mix for the test was prepared as follows: 10  $\mu\text{l}$  R-domain and 2  $\mu\text{l}$  10 mM Ellman's reagent. As done previously, relevant protein and reagent controls were set up alongside the samples. Reactions were mixed by pipetting and the  $\Delta A_{412}$  measured after 5 min by NanoDrop™ spectrophotometry. This enabled the calculation of the free thiol population within R-domain stocks and N-MBM labelling efficiency. The same protocol was used to test R-domain labelled with N-MBM for extended incubation periods (6 hr and 18 hr) at 30 °C.

#### **2.2.15 Use of MALDI-MS to test for Bromomaleimide labelling of Cole9 R-domain**

Matrix-assisted laser desorption/ionization (MALDI MS) uses a softer form of ionisation compared to ESI-MS. MALDI-MS allows for an easier characterisation of the mass change upon labelling after fragmentation of the labelled protein by a suitable protease. Fragments of unlabelled/labelled R-domain can be analysed by MALDI MS to confirm N-MBM labelling of the 2C site.

Uncleaved R-domain ( $\sim 184 \mu\text{M}$ ) in sodium phosphate buffer (100 mM Na Phosphate pH 7.0, 300 mM NaCl) was labelled with N-MBM (100 mM in DMF) at a 20:1 label:protein molar ratio for 6 hr at 30 °C. Unlabelled R-domain was analysed alongside the labelling reaction as a control. Samples were separated by SDS-PAGE (15 % (w/v) gel) alongside 10  $\mu\text{l}$  PAGE protein ladder to isolate the protein for MALDI MS. The sample mix was prepared as follows: 4  $\mu\text{l}$  protein, 1  $\mu\text{l}$  1 mM DTT and 2  $\mu\text{l}$  5x Loading Dye. Bands were excised from the gel using an ethanol-flamed scalpel and stored in labelled 1.5 ml microcentrifuge tubes. Tubes were spun briefly in a bench-top centrifuge to deposit samples to the bottom of tubes with residual buffer removed using a pipette.

Samples were submitted to Dr. David Ashford for analysis (Technology Facility, Department of Biology, University of York). Trypsin was used to digest R-domain and generate peptide fragments within the correct size range for MS analysis. Trypsin cleaves at the C-terminal side of Lys or Arg residues unless the next residue is Pro. Following digestion, samples were analysed by liquid chromatography–mass spectrometry (LC-MS). Samples were loaded onto a nanoAcquity Itra performance liquid chromatography system (Waters) equipped with a nanoAcquity Symmetry C<sub>18</sub> 5 µm trap (180 µm x 20 mm Waters) and a nanoAcquity high strength silica T3 1.8 µm C<sub>18</sub> capillary column (75 µm x 250 mm, Waters). The trap wash solvent was 0.1 % (v/v) aqueous formic acid and the trapping flow rate was 10 µL/min. The trap was washed for 5 min before switching flow to the capillary column.

Liquid chromatography separation used a gradient elution of two solvents (solvent A: aqueous 0.1 % (v/v) formic acid; solvent B: acetonitrile containing 0.1 % (v/v) formic acid). The capillary column flow rate was 350 nL/min and the column temperature was 60°C. The gradient profile was linear 2-35 % B over 20 min. All runs then proceeded to wash with 95 % solvent B for 2.5 min. The column was returned to initial conditions and re-equilibrated for 25 min before subsequent injections. The sequences of the identified peptide fragments were analysed using BLAST (<http://www.ncbi.nlm.nih.gov/blast>) (McGinnis and Madden, 2004; Johnson *et al.*, 2008) and MASCOT (Perkins *et al.*, 1999).

#### **2.2.16 Labelling ColE9 R-domain with N-MBM and Cysteine**

Attempts were made to conjugate cysteine to N-MBM-labelled ColE9 R-domain as a confirmation of labelling. The coupling protocol mimicked the planned ssDNA oligonucleotide coupling for assembling the MTP. R-domain was labelled with N-MBM as in previous protocols. Uncleaved, N-MBM-labelled R-domain (~176 µM) in sodium phosphate buffer (100 mM sodium phosphate pH 7.0, 300 mM NaCl) was reacted with cysteine (100 mM in sodium phosphate buffer) at a 5:1 Cys:protein molar ratio for 6 hr at 30 °C. The reaction was then dialysed into 1 L of fresh sodium phosphate buffer at 4 °C (12-14 kDa MWCO, Spectra/Por) overnight to remove free cysteine.

Following removal from dialysis, the reaction product R-domain-N-MBM-Cys had its concentration determined as done previously. R-domain-N-MBM-Cys was then analysed by SDS-PAGE (15 % (w/v) gel) alongside unlabelled R-domain to identify any molecular weight shift indicative of successful coupling. The sample mix was prepared as follows: 4  $\mu$ l R-domain, 2  $\mu$ l 10 mM DTT and 2  $\mu$ l 5x loading dye. The entire mix was loaded into gel wells alongside 10  $\mu$ l PAGE ladder. R-domain-N-MBM-Cys was delivered to Richard Spears (Fascione Lab, Department of Chemistry, University of York) for analysis by positive-mode LC-MS to detect the precise mass change indicative of successful conjugation (+121.16 g/mol).

### **2.2.17 Coupling Fluorescent Tetrapeptide to ColE9 R-domain**

Attempts were made to couple a synthesised fluorescent tetrapeptide to ColE9 R-domain using thiol chemistry. To initiate coupling, a 20 mM stock of peptide in sodium phosphate buffer (100 mM Na Phosphate, pH 7.5, 150 mM NaCl) was added to unlabelled/labelled R-domain in equivalent buffer at a 20:1 peptide:protein molar ratio. Coupling reactions were incubated at 30 °C for 18 hr to mimic labelling conditions. Coupling products were analysed by SDS-PAGE (15 % (w/v) gel) using the following sample mix: 10  $\mu$ l protein, 5  $\mu$ l 4x loading dye with either +/- 3  $\mu$ l DTT or +/- 1.67  $\mu$ l 137 mM beta-mercaptoethanol ( $\beta$ -ME). Two controls were run alongside reaction products, one containing only free peptide, and another containing unreacted R-domain. Images of gels were photographed by digital camera with UV illumination to detect fluorescent bands. Following this, gels were stained with Coomassie Blue dye with another photograph of the protein bands taken post-destaining.

Subsequent peptide coupling experiments had longer incubation times with reducing agents (1 hr and 4 hr with  $\beta$ -ME) prior to gel loading. Certain experiments removed free peptide from peptide-coupled R-domain by buffer exchange. Sample aliquots (~500  $\mu$ l) were added to spin concentrator columns (Vivaspin 500, MWCO 5,000 Da) and reduced to a volume of ~100  $\mu$ l by centrifugation (15,000 xg, 4 °C, Thermo bench-top centrifuge). Columns were refilled up to 500  $\mu$ l with fresh buffer and centrifuged to reduce volume to 100  $\mu$ l. This process was repeated 3-4 times to remove free and weakly-associated fluorescent peptide. Resultant samples were again analysed by SDS-PAGE (15 % (w/v) gel) to detect R-domain-peptide conjugate.

Cysteine was coupled to R-domain in an attempt to block the 2Cys residue. Cysteine-HCl (Sigma Aldrich, 100 mM in water) was added to uncleaved R-domain at a 20:1 cysteine:protein molar ratio. Both unlabelled and N-MBM-labelled R-domain stocks were reacted with cysteine in these experiments. Unlabelled/labelled R-domain-Cys samples were incubated with peptide as previously described. Resulting samples were analysed by SDS-PAGE (15 % (w/v) gel) as done previously. R-domain-peptide reaction samples were analysed on the same gel to evaluate the effect of blocking with cysteine.

Selected samples were chosen for further analysis by LC-MS; these runs were conducted by Richard Spears (Fascione Lab, Department of Chemistry, University of York). Pre-packed disposable (PD) MiniTrap G-25 (GE Biosciences) desalting columns were used to transfer samples into ammonium acetate buffer (100 mM, pH 6.8). First, columns were equilibrated by adding ~8 ml of ammonium acetate buffer to the column bed, and the flow-through was discarded. Up to 500 µl of sample was added to the packed bed, made up to 500 µl with additional buffer, and the flow-through was again discarded. For sample elution, 1 ml of ammonium acetate buffer was used, and the flow-through collected. If necessary, the sample was spin concentrated (Vivaspin 500, MWCO 5,000 Da, 15,000 xg, 4 °C, Thermo bench-top centrifuge) to ensure the protein concentration was suitable for ESI-MS ( $\geq 100$  µM). A small aliquot of each sample was diluted 100-fold into 90 % water: 10 % (v/v) acetic acid and analysed by LC-MS.

### **2.2.18 Coupling ColE9 T-domain Linked 3' Thiol Oligonucleotide to ColE9 R-domain**

ssDNA oligonucleotides with respective 5' and 3' thiol group modifications were ordered (Table 2, Sigma-Aldrich) to act as test molecules for MTP assembly. Test oligonucleotides were ordered without the fluorescent labels they would possess in the final MTP. Lyophilised ColE9 T-domain linked 3' thiol oligonucleotide was first briefly centrifuged, then suspended in 200 µl nuclease-free water to generate a 300 µM stock solution. 100 µl of this stock was taken, and to it was added: 170 µl of sodium phosphate buffer (100 mM sodium phosphate pH 7.0, 150 mM NaCl), and 30 µl DTT (1 M in sodium phosphate buffer). This reduced sample was vortexed for 1 sec and incubated at room temperature for 1 hr. The sample was then buffer exchanged using a

spin concentrator (Vivaspin 500, 5,000 Da MWCO) to remove DTT from the solution. All 300  $\mu$ l was added to the column, and spun down (15,000 xg, 10 °C, Thermo bench-top centrifuge) to ~100  $\mu$ l. Volume of stock was increased to 300  $\mu$ l with fresh buffer without DTT, and the process repeated three times to ensure complete DTT removal.

Both N-MBM-labelled R-domain and unlabelled R-domain stocks (~154  $\mu$ M) were reacted with T-domain linked 3' thiol oligonucleotide at a 4:1 protein:oligonucleotide molar ratio. Oligonucleotide stock concentrations were calculated from absorbance measurements recorded by NanoDrop™ spectrophotometry. The reaction was vortexed for 1 sec, and then incubated at 30 °C for 24 hr. Aliquots were taken at 3 hr and 24 hr and analysed by SDS-PAGE (15 % (w/v) gel) alongside controls of unreacted R-domain. Sample mixes for gel analysis were as follows: 10  $\mu$ l of protein/protein-DNA, and 5  $\mu$ l 4x protein ladder (no DTT). The entire reaction mix for each lane was loaded onto the gel alongside 10  $\mu$ l protein ladder.

To aid identification of potential R-domain-ssDNA complexes, aliquots were digested with DNase I, then analysed by SDS-PAGE alongside existing samples. For these reactions, a pre-existing stock of DNase I (Sigma-Aldrich, 10 mg/ml in water) was diluted to 0.5  $\mu$ g/ $\mu$ l (0.2 U per  $\mu$ l). The reaction mix was as follows: 5  $\mu$ l protein-oligonucleotide (~150  $\mu$ M), 1  $\mu$ l 0.5  $\mu$ g/ $\mu$ l DNase I and 0.6  $\mu$ l 10 mM MgCl<sub>2</sub>. All reactions were incubated at 37 °C for 1 hr. Each mix was loaded into SDS-PAGE gels with 2.5  $\mu$ l of 4x loading dye. A follow-up analysis added 1  $\mu$ l 0.5 M EDTA to each sample to chelate Mg<sup>2+</sup> ions.

### **2.2.19 Annealing of ColE9 R-domain linked 5' Thiol Oligonucleotide to ColE9 T-domain linked 3' Thiol Oligonucleotide**

ColE9 R-domain linked 5' thiol oligonucleotide was annealed to unlabelled/labelled R-domain using the same coupling protocol as for T-domain linked 3' thiol oligonucleotide. A 200  $\mu$ M stock of lyophilised 5' thiol oligonucleotide was made up using 144  $\mu$ l water, after centrifuging the sample tube briefly. A 100  $\mu$ l aliquot was taken and 170  $\mu$ l 100 mM Sodium Phosphate (pH 7.0), 100 mM NaCl buffer and 30  $\mu$ l 1 M DTT in phosphate buffer (Final concentration: 100 mM) was added to it. Reduced oligonucleotide was vortexed for 1 sec and incubated at room temperature for 1 hr. The sample was buffer exchanged in a spin concentrator (Vivaspin 500, 5,000 MWCO) to

remove DTT. All 300  $\mu$ l was added to the column, and centrifuged to reduce the volume to  $\sim$ 100  $\mu$ l (15,000 xg, 10 °C, Thermo bench-top centrifuge). The volume of the stock was increased to 300  $\mu$ l again with fresh buffer without DTT, and the process repeated three times to ensure complete DTT removal.

Products of R-domain linked 5' thiol oligonucleotide coupling treated with DNase were analysed by SDS-PAGE (15 % (w/v) gel) as described previously. Sample concentrations were determined by NanoDrop<sup>TM</sup> spectrophotometry. Aliquots of R-domain linked 5' thiol oligonucleotide and R-domain linked 3' thiol oligonucleotide were reacted at a 1:1 molar ratio. Unlabelled R-domain linked 5' thiol oligonucleotide was reacted with unlabelled R-domain linked 3' thiol oligonucleotide, and labelled R-domain linked 5' thiol oligonucleotide with labelled R-domain linked 3' thiol oligonucleotide, respectively. After vortexing for 1 sec to mix, the reaction was incubated at 30°C for 18 hr. This reaction was analysed by native PAGE (12 % (w/v) gel). BSA was used as a marker (100 mg/ml stock was diluted 1 in 200). The samples were prepared as follows: 30  $\mu$ l protein-oligonucleotide reaction mix and 10  $\mu$ l 4x native PAGE sample buffer (Invitrogen). Native gel electrophoresis was done for 5 hr at 4 °C (1 M Tris/1 M Tricine pH 8.2 running buffer). The gel was stained with Coomassie blue dye stain overnight, and then destained in water until bands became clear ( $\sim$ 24 hr). Samples were also analysed by LC-MS as described previously. A modification to this protocol was the use of 30 kDa MWCO spin concentrators (Vivaspin 500) to remove unannealed DNA molecules.

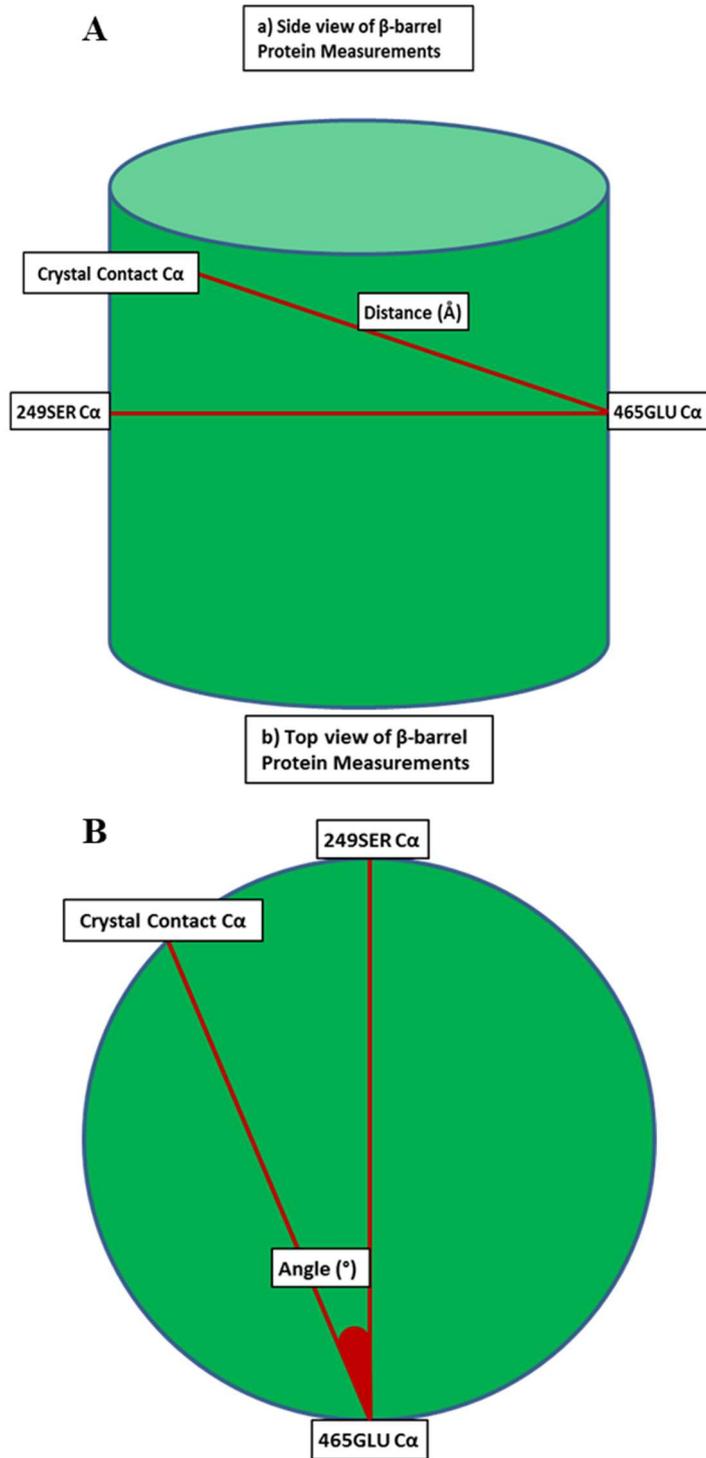
## **2.3 Methods for Structural Bioinformatics Investigation of Protein-Protein, Protein-Lipopolysaccharide and Protein-Detergent Interactions in Crystal Structures of Bacterial Outer Membrane Proteins**

### **2.3.1 Mining protein data bank (PDB) data for monomeric OMP crystal contacts in WinCOOT**

FhuA (1QFF) was used as a reference structure for searching three-dimensional (3D) alignment in PDBeFold (<http://www.ebi.ac.uk/msd-srv/ssm/>). The initial search cut-off was  $\geq$  50 % SSE (fraction of secondary structure of query chain identified in the target protein), and resulted in 69 matches. Individual monomeric structures were selected with a relatively low root-mean-square deviation (RMSD) value (4 Å) for further

analysis of their interactions. Structures from PDB files were analysed in WinCOOT using symmetry analysis in a 25 Å radius from the macromolecular centre of the reference protein (Emsley *et al.*, 2010). This radius of 25 Å was deemed to be sufficient to include all relevant crystal contacts. Structures with biologically relevant interactions in the membrane plane were noted for further study. Residues directly or indirectly involved with detergent, LPS or protein interactions were catalogued.

For detailed 3-D analysis, BtuB (3RGM) was used as a reference because FhuA (1QFF) had no interactions in the plane of the membrane bilayer. Each monomeric OMP structure was superimposed onto BtuB and the distance (Å) and angle of the crystal contact residues relative to established reference positions (Glu465 and Ser249) was measured. The reference residues are positioned at about 50 % of the height of the macromolecule with a short, linear distance from one end of the  $\beta$ -barrel to the other (Figure 2-3). Distances were measured from the C $\alpha$  of Glu465 of BtuB to the C $\alpha$  of the residue in question from the superimposed protein. The angle to the C $\alpha$  of this residue was measured relative to a reference line intersecting Ser249 C $\alpha$  and Glu465 C $\alpha$ . The distances from the Glu465 and Ser249 reference residues to the nearest respective residue of the superimposed protein were also measured. X, Y and Z coordinates of the crystal contact residues were recorded to generate 3D plots in the future. In addition, the distance of each intermolecular contact residue to the closest atom of the symmetry partner was noted, i.e. the atom of the generated symmetry image of the protein that was the shortest distance from the contact residue.



**Figure 2-3 – Reference residues for distance and angle measurements**

Images showing the reference residues Glu465 and Ser249 mapped onto a cylinder representing a  $\beta$ -barrel membrane protein. The distance between the residues was used for measuring the angles from Glu465. **A-** Shows the side-on view in the plane of the membrane bilayer, **B-** shows a top-down view of the extracellular protein surface.

Using the measurements obtained, the distance vs angle to the reference point (Glu465) was plotted for the crystal contacts. It was also noted whether contacts were from protein-protein, protein-detergent or protein- LPS interactions. One monomeric OMP analysed, FhuA (4CU4), was crystallized with LPS bound. These protein-LPS interactions were mapped in 3D onto BtuB (3RGM) using PyMOL (Schrödinger, 2013).

### **2.3.2 Mining the Crystal Contacts in Other Outer Membrane Protein Structures**

Comparison and 3D alignment in PDBeFold was done using a cut-off of  $\geq 60$  %SSE (<http://www.ebi.ac.uk/msd-srv/ssm/>). Crystal structures of individual suitable trimers and monomers were searched for using OmpA (1QJP), OmpC (1OSM) and OmpF (1OPF). An RMSD value of 4 Å was the threshold for the selected proteins. However, a limited number of individual trimer crystal structures were available, and only 3/38 matches were individual trimers which could be analysed by these methods (including the reference OmpC).

Crystal contacts were searched for in the plane of the membrane bilayer by symmetry analysis in a 40 Å radius from the macromolecular centre of the trimeric OMP. The increased symmetry radius was used to account for the larger radius of the trimer compared to the monomer. The standard 25 Å radius was used for monomeric OmpA. These contacts were subsequently plotted in PyMOL to visualise them in 3D. Residues involved in the trimerisation interfaces of OmpC were noted for further analysis. Distances and angles were measured between the C $\alpha$  of contact residues and the Tyr231 and Ala84 reference residues of OmpC (1OSM) monomer.

### **2.3.3 Mining the Crystal Contacts of Mitochondrial Voltage-Dependent Anion Channels**

Searching and comparison was done in PDBeFold with a cut-off of  $\geq 70$  % SSE as described previously. The RMSD cut-off was 4 Å and only individual monomers were used for analysis, resulting in 6/54 suitable PDB files. Symmetry analysis was conducted in WinCOOT with a 25 Å radius from the macromolecular centre. Interactions in the plane of the membrane bilayer were noted following the same procedure as described previously. Using human voltage-dependent anion channel (hVDAC) as a reference (2JK4, chain A), matching proteins were superimposed onto

hVDAC, and the distance (Å) and angle of the C $\alpha$  of crystal contact residues relative to a reference (Gly195 and Gly85) were measured. hVDAC reference residues were identified by structural alignment to previous reference residues (3RGM, Glu465 and Ser249). Distances were measured from the C $\alpha$  of Gly195 in hVDAC to the C $\alpha$  of the residue in the superimposed protein. An angle was measured between the following residues: Gly85 C $\alpha$ , Gly195 C $\alpha$  and the specific residue C $\alpha$ . The distances of the Gly195 and Gly85 reference residues to the nearest respective residue of the superimposed protein were also measured.

### 2.3.4 Hydrophobicity Analysis

To understand what differentiates these patches from the rest of the  $\beta$ -barrel surface, the amino acids within the patches and their relative hydrophobicities were also investigated. It has been established that  $\beta$ -barrel proteins contain an 'aromatic girdle' of Trp/Tyr residues at the extremes of the membrane bilayer interface that may serve as swim-stabilisers (Elofsson and von Heijne, 2007). These girdle residues were mapped onto BtuB (3RGM) to determine whether they participated in the predicted promiscuous PPI patches. The hydrophobicity of these patches may differ relative to the overall membrane bilayer facing  $\beta$ -barrel surface. The aromatic girdles have also been used to define the boundaries of the hydrophobic membrane bilayer interface to calculate the overall hydrophobicity of this  $\beta$ -barrel surface.

Using established hydrophobicity scales (Kyte and Doolittle, 1982; White and Wimley, 1998) and amino acid side chain volumes (Häckel, Hinz and Hedwig, 1999), the patches were compared to the rest of the membrane bilayer facing  $\beta$ -barrel surface. To do this, a parameter similar to Matthew's coefficient (Daltons Å<sup>-3</sup>) was generated, instead substituting the hydrophobicity for each amino acid (Tables 2-2, 2-3 and Figure 2-4) (Matthews, 1968). A total hydrophobicity coefficient for the entire patch was calculated from the amino acids in the patch, and then divided by the total number of amino acids to obtain the average hydrophobicity per residue. This was also completed for the entire membrane-facing  $\beta$ -barrel surface as a comparison.

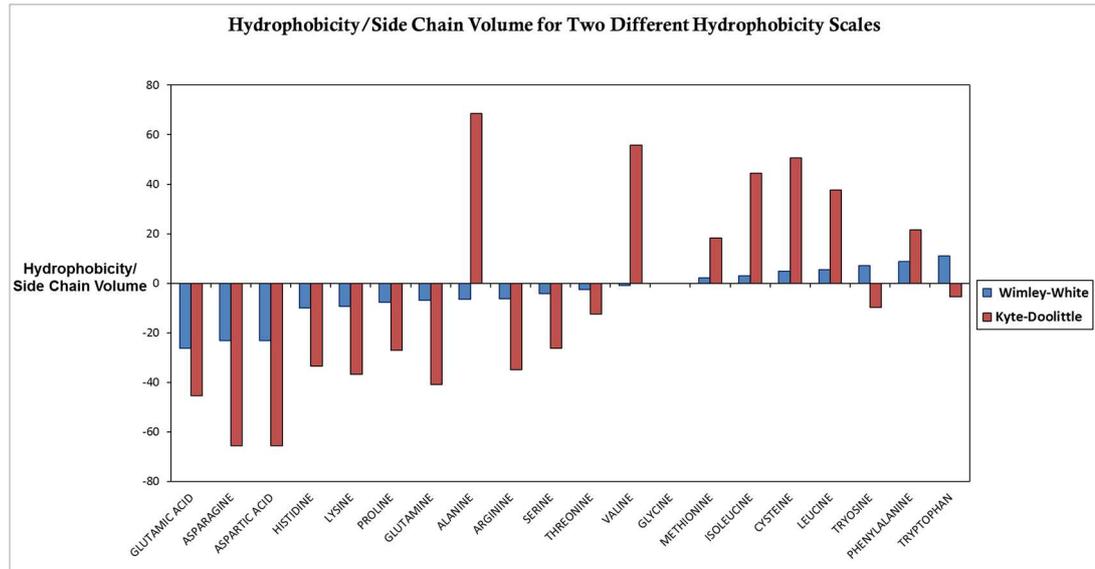
**Table 2-2 – Terms and definitions for the  $\beta$ -barrel surface hydrophobicity calculations**  
Parameters were used to generate the Hydrophobicity Contribution Coefficient for each amino acid at the  $\beta$ -barrel surface.

<b>Term</b>	<b>Definition</b>
Volume of Amino Acid Side Chain ( $SC_p$ ) ( $\text{\AA}^3 \text{residue}^{-1}$ )	Experimentally determined volumes of membrane lipid facing amino acid side chains. Found by subtracting the volume of the glyceryl residue from the overall residue volume, e.g. $SC_p(\text{Ala}) = V_p(\text{Ala}) - V_p(\text{Gly})$ (Harpaz, Gerstein and Chothia, 1994)
Hydrophobicity of Amino Acid ( $HP_{AA}$ ) ( $\text{kcal mol}^{-1}$ )	Used two different published Hydrophobicity Scales for calculations.  <b>WW</b> = Hydrophobicity scale using whole-residue free energy transfer $\Delta G$ from water to POPC interface (White and Wimley, 1998).  <b>KD</b> = Hydrophobicity scale using whole-residue free energy transfer $\Delta G$ from water to vapour (Kyte and Doolittle, 1982).
Hydrophobicity Contribution Coefficient per amino acid ( $HPCC_{AA}$ ) ( $\text{kcal mol}^{-1} \text{\AA}^{-3} \text{residue}$ )	$HPCC_{AA} = HP_{AA}/SC_p$  $\Sigma HPCC_{WW/KD} = (HP_{GLY} + HP_{LEU} + \dots)$
Hydrophobicity of $\beta$ -barrel surface (or patch surface) ( $HPBS_{WW/KD}$ )	The overall hydrophobicity of the surface is calculated by multiplying the $HPCC_{AA}$ for each amino acid by the number of that specific amino acid present at the $\beta$ -barrel surface, adding these values for all surface exposed amino acids, then dividing it by the total number of outward-facing residues at the surface.  $HP_{BBS} = \Sigma((HPCC_{ILE} * \Sigma No. ILE) + (HPCC_{GLU} * \Sigma No. GLU) + \dots) / \Sigma No. AA$

**Table 2-3 – Calculated hydrophobicity contributions for each amino acid**

Side chain volumes and hydrophobicity for each amino acid were obtained from published sources. See Table 5-1 for detailed definitions of terms used. Values are to four significant figures.

Amino Acid	SC <sub>P</sub> (Å <sup>3</sup> residue <sup>-1</sup> )	HP <sub>AAWW</sub> (kcal mol <sup>-1</sup> )	HP <sub>AAKD</sub> (kcal mol <sup>-1</sup> )	HPCC <sub>AAWW</sub> (kcal mol <sup>-1</sup> Å <sup>-3</sup> residue)	HPCC <sub>AAKD</sub> (kcal mol <sup>-1</sup> Å <sup>-3</sup> residue)
Glycine	0	-0.01	-0.4	0	0
Arginine	129	-0.81	-4.5	-6.279	-34.88
Histidine	95.5	-0.96	-3.2	-10.05	-33.51
Lysine	106.2	-0.99	-3.9	-9.322	-36.72
Aspartic acid	53.3	-1.23	-3.5	-23.08	-65.67
Glutamic acid	77	-2.02	-3.5	-26.23	-45.45
Serine	30.4	-0.13	-0.8	-4.276	-26.32
Threonine	56.2	-0.14	-0.7	-2.491	-12.46
Asparagine	53.3	-1.23	-3.5	-23.08	-65.67
Glutamine	85.6	-0.58	-3.5	-6.776	-40.89
Cysteine	49.4	0.24	2.5	4.858	50.61
Proline	59.3	-0.45	-1.6	-7.589	-26.98
Alanine	26.3	-0.17	1.8	-6.464	68.44
Isoleucine	101.1	0.31	4.5	3.066	44.51
Leucine	100.8	0.56	3.8	5.556	37.70
Methionine	103.9	0.23	1.9	2.214	18.29
Phenylalanine	129.7	1.13	2.8	8.712	21.59
Tryptophan	167.9	1.85	-0.9	11.02	-5.360
Tryosine	133.3	0.94	-1.3	7.052	-9.752
Valine	75.3	-0.07	4.2	-0.9296	55.78



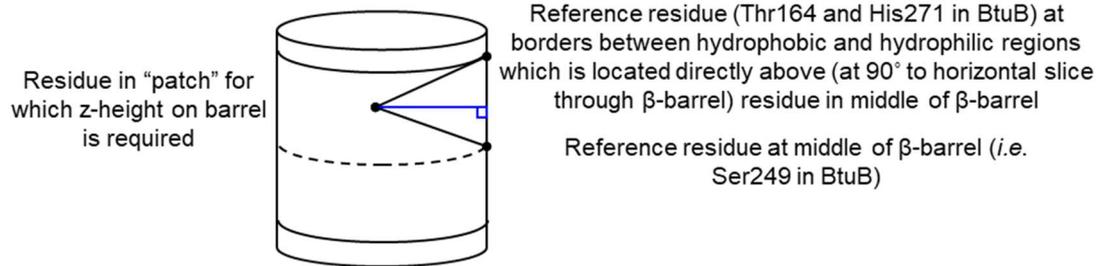
**Figure 2-4 – Hydrophobicity/side chain volume ( $\text{kcal mol}^{-1} \text{Å}^{-3}$ ) for each amino acid**  
Graph showing the calculated hydrophobicity divided by side chain volume for each amino acid. This Hydrophobicity Contribution Coefficient (HPCC) was calculated using both the Wimley-White (blue) and Kyte-Doolittle (red) hydrophobicity scales. These values were used to generate a HPCC value for any interaction patch identified on the  $\beta$ -barrel surface.

### 2.3.5 Amino Acid Composition and Location Analysis

The differences in amino acid composition were compared between Patch 1 and Patch 2 of monomeric OMPs mapped on to BtuB and the entire  $\beta$ -barrel surface, and the respective patches of BtuB and FhuA. This was done to determine if the amino acid composition within identified PPI patches affects their ability to form interactions. Heat mapping was used to highlight the most distinct differences in percent amino acid composition.

To determine if certain amino acids cluster at certain points in height along the  $\beta$ -barrel, a series of reference residues (Thr164, Ser249 and His271) were identified, and the distance and angle to each interface residue measured relative to these reference residues (Figure 2-5). The reference residues form a vertical cross-section ending at the boundaries of the hydrophobic region of the  $\beta$ -barrel OMP, determined by an electrostatic surface potential map of BtuB (3RGM) generated in PyMOL. Utilising trigonometry and the measurement tools of PyMOL and WinCOOT in 3D, a series of z distances from the horizontal cross-section marked at Ser249 were calculated. Z values

were positive if closer to the extracellular face of the OMP, and negative if closer to the periplasmic face.



**Figure 2-5 – Diagram showing the protocol used to determine Z distances of interface residues**

Black circles represent the positions of the reference residues Thr164, Ser249 and His271. The dotted line shows the horizontal cross-section of the  $\beta$ -barrel at Ser249 at the mid-point of the total  $\beta$ -barrel height.

## Chapter 3: Production and Analysis of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Construct

### 3.1 Introduction

#### 3.1.1 Background to ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Construct and Incorporation in MTP Design

The translocation (T) domain of ColE9 forms one half of the MTP designed to quantify the mechanical unfolding forces that may be generated at the bacterial cell surface during ColE9 cell entry. T-domain binds to OmpF and facilitates ColE9 translocation across the OM; thus is essential to a viable MTP. Previous attempts to purify the unstructured region of the translocation domain (residues 1-83 of ColE9, T<sup>1-83</sup>-domain) failed due to problems with protein solubility. To purify the T<sup>1-83</sup>-domain, the ColE9 DNase-containing C domain was fused to its C-terminus to improve its solubility. This construct is encoded by the pRP1 plasmid (constructed by Dr. Patrice Rassam) where the T<sup>1-83</sup>-domain and DNase domain (residues 455 -580 of ColE9) are separated by a TEV cleavage site (ENLYFQG).

ColE9 DNase is a non-specific endonuclease with a distinctive H-N-H motif, and as such is highly dependent on the presence of divalent metal cations (e.g. Mg<sup>2+</sup>) to function. The DNase domain is capable of endonuclease activity against a range of substrates from RNA, to ssDNA and dsDNA (Pommer *et al.*, 2001). The phosphate groups of the substrate, being negatively charged, bind favourably to a positive cleft on the DNase surface where the H-N-H motif is located. Activation of a hydrolytic water molecule by His103, acting as a general base, hydrolyses the phosphodiester linkage of its nucleic acid substrate (Pommer *et al.*, 2001).

A divalent metal cation (e.g. Mg<sup>2+</sup>) is usually required to stabilise the penta-coordinate transition state of this reaction and serve as the general acid, protonating the 3' oxygen leaving group. His127 and His102 coordinate the divalent metal cation so it can act as an effective cofactor (Maté and Kleanthous, 2004). Random mutagenesis studies support the importance of Arg96, His127 and Glu100 to the ColE9 DNase mechanism (Garinot-Schneider *et al.*, 1996).

Whilst allowing the translocation domain to be soluble, and thus making the MTP viable, introducing the DNase region to the MTP does come with a drawback. Since the MTP contains ssDNA oligonucleotide components, the DNase region must be inactivated to prevent the probe from degrading itself. Two methods of inactivating the DNase were proposed: purification with ColE9 immunity protein (Im9), or the introduction of inactivation mutations to the DNase domain. Given its extremely high binding affinity for DNase ( $K_d \approx 10^{-14}$  M), Im9 would prove difficult to remove (Farrance *et al.*, 2013). Mutation of His103 to an alanine residue has been characterised as an inactive mutant of ColE9 DNase (Walker *et al.*, 2002). Introduction of H103A to T<sup>1-83</sup>-TEV-DNase was planned to inactivate the DNase domain without resorting to inhibition by Im9. Previous work incorporated this H103A mutation into the pNGH25 plasmid by site-directed mutagenesis (prepared by Dr. Patrice Rassam), resulting in the final plasmid, pRP1, encoding the protein ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>.

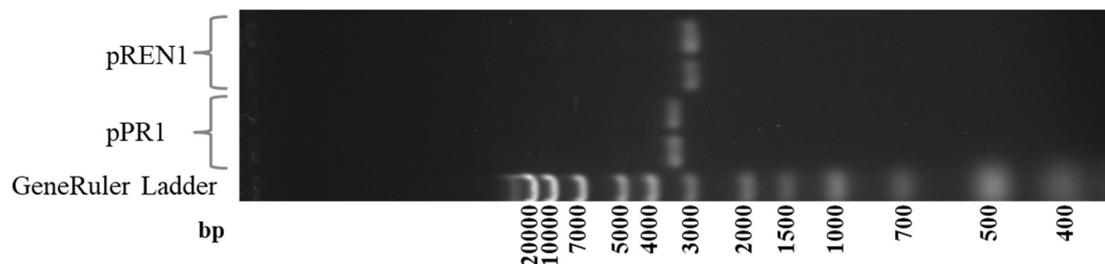
Before ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct can be introduced as a component of the MTP, it must be determined whether or not the H103A mutation has successfully inactivated the DNase activity. To assay this, an easily purified pUC- $\phi$ 600 plasmid DNA was used as a substrate for the DNase. Once purified, pUC- $\phi$ 600 was used to assay plasmid degradation across a range of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> concentrations. Wild-type ColE9 DNase was assayed alongside ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> to judge the comparative effectiveness of the inactivating H103A mutation.

The MTP was designed for use in fluorescence microscopy experiments, and the established buffer used for this work contains 2 mM Mg<sup>2+</sup> (i.e. M9 supplemented media used to ensure bacterial cell viability during microscopy). Mg<sup>2+</sup> ions are cofactors in ColE9 DNase degradation of ssDNA. Therefore, it is important that ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> activity be assayed under these conditions to ensure inactivation is maintained. Should the ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> display evidence of plasmid degradation, EDTA would be investigated for use as a quenching agent. EDTA could potentially quench remaining DNase activity within ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> to allow its use as part of the MTP. These experiments will probe whether or not it is possible to protect the oligo component of the MTP using H103A.

### 3.2 Results

#### 3.2.1 Over-Production of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Construct

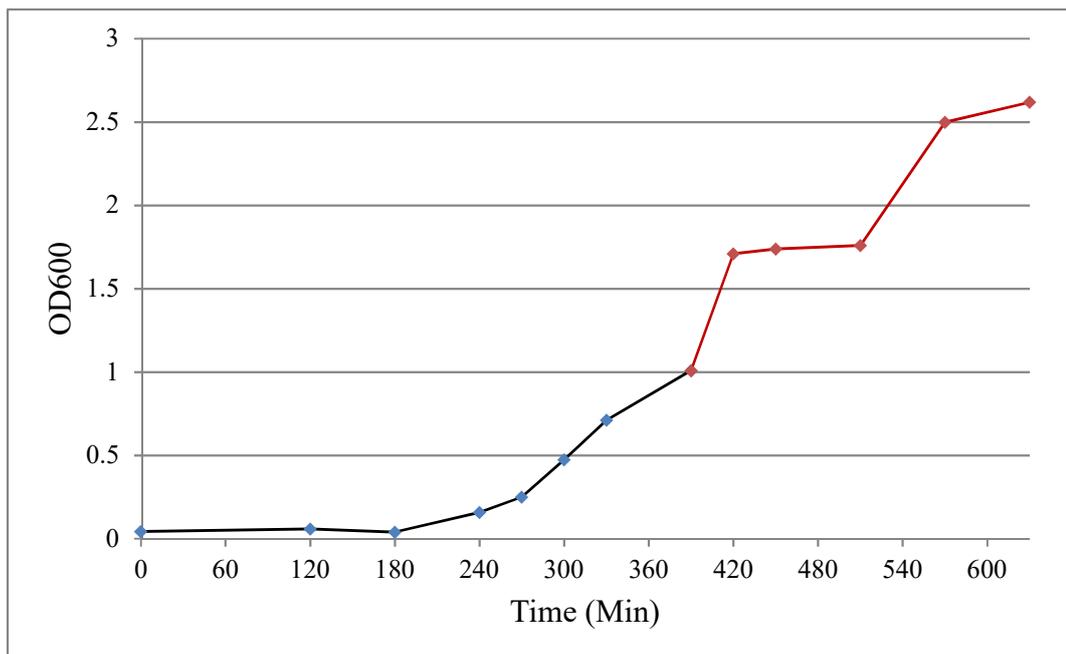
Agarose-TAE gel electrophoresis of extracted plasmids showed a single distinct DNA band, indicating the successful production of intact supercoiled plasmid (Figure 3-1). Bands in pREN1 lanes were observed at ~2,900 bp (expected size = ~5,800 bp). The full sequence for pRP1 is unavailable. Plasmid identity could not be confirmed from expected size. Plasmid stocks proceeded to transformation, with identity to be confirmed from identification of over-produced encoded protein product.



**Figure 3-1 – 1 % (w/v) Agarose-TAE gel of pREN1/pPR1 Plasmid Purification**

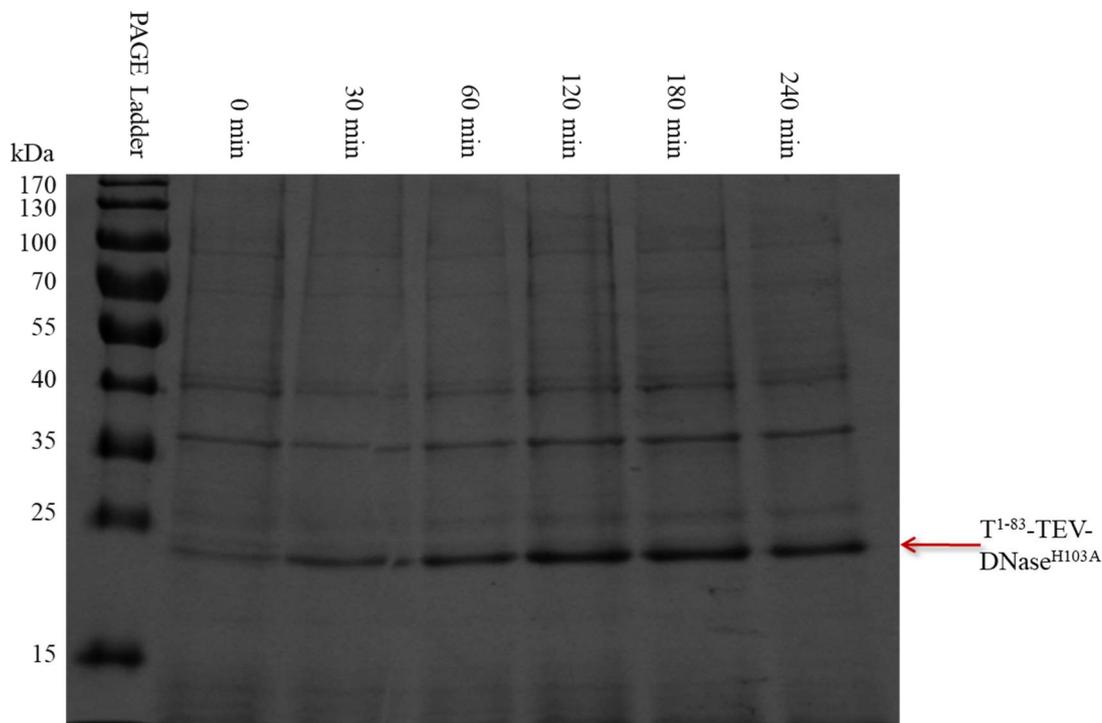
Colonies were purified from transformed BL21\*(DE3) and DH5 $\alpha$  bacterial cells for pPR1 and pREN1, respectively, each bacterial culture was split in half for purification, resulting in the four samples analysed on the gel. Thermo Scientific GeneRuler DNA Ladder Mix was used for size reference.

Competent *E. coli* cells (BL21\*(DE3)) transformed with pPR1 were grown successfully, generating a visible log phase in the growth curve (Figure 3-2). This growth rate was maintained following the addition of IPTG to induce production of protein. The cultures continued to be shaken and grown at 37 °C for 4 hr, with OD<sub>600</sub> measured regularly (Figure 3-2), and 1 ml aliquots taken every hour for analysis by SDS-PAGE (Figure 3-3). A band that increased in intensity over time at ~23,000 Da was observed, thought to be ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (expected MW = 23267.2 Da). With evidence suggesting successful over-production of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>, samples proceeded immediately to purification.



**Figure 3-2 – Growth Curve of BL21\*(DE3) bacterial cells transformed with pPR1 plasmid during induction with IPTG**

Optical density (OD<sub>600</sub>) as a function of time (min) for BL21\*(DE3) cells transformed with the plasmid pPR1 and engaged in logarithmic growth. Blue data points are readings taken pre-induction with IPTG, and red data points are taken post-induction with IPTG.



**Figure 3-3 – 12 % (w/v) SDS-PAGE gel analysis of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> over-production in BL21\*(DE3) bacterial cells**

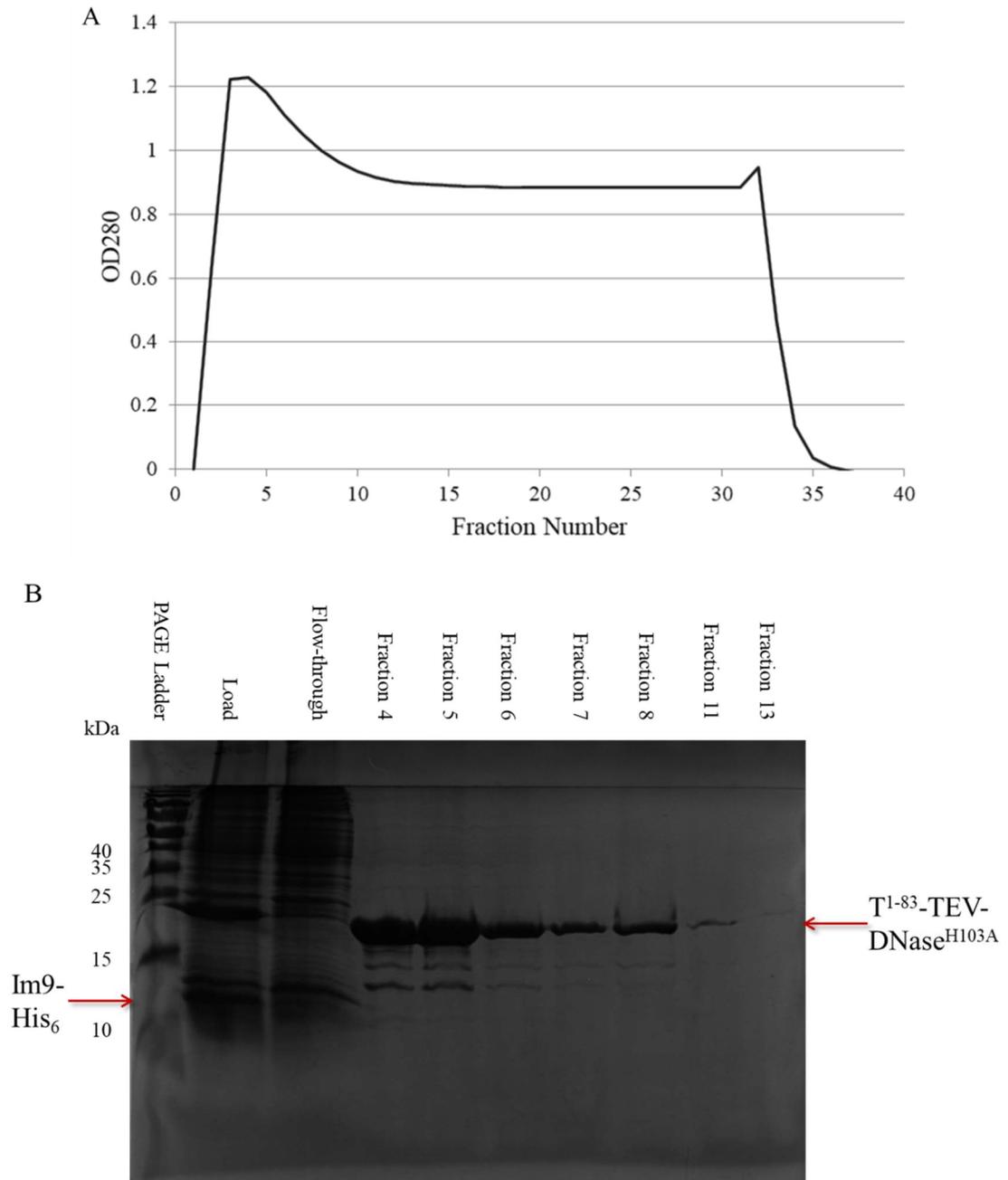
12 % (w/v) SDS PAGE gel analysis of cell aliquots from over-production of Cole9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> in BL21\*(DE3) *E. coli*. Induction was carried out over 4 hr with aliquots taken at regular time points. The band that increases in intensity over time at ~23,000 Da is attributed to the Cole9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct.

### 3.2.2 Purification and Characterisation of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>

Fractions taken from Ni<sup>2+</sup>-affinity purification indicated the presence of pure ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> product. OD<sub>280</sub> readings from the FPLC increased to form an absorbance peak during the elution of fractions 4 through 10, consistent with the presence of protein (Figure 3-4A). Analysis by SDS-PAGE revealed intense protein bands at ~23,000 Da, suggesting the purified product was ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (expected MW = 23267.2 Da) (Figure 3-4B). The majority of bands present in the load and flow-through were not present in the eluted fractions, further evidence in support of successful purification using co-produced Im9-His<sub>6</sub>.

Mobility shift analysis of the purification gel was conducted to estimate the molecular weight of observed bands. The relative mobility (Rf) of the protein molecular weight ladder was plotted against log molecular weight (Figure 3-5). The resulting linear trendline was used to estimate the molecular weight of observed bands. The molecular weight of the band thought to be ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (expected MW = 23267.2 Da) was estimated at 23320 Da (4sf). The molecular weight of the band thought to be Im9-His<sub>6</sub> (expected MW = 10647.7 Da) was estimated at 11520 Da (4sf). Although not exact, these estimates suggested that the observed bands are ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> and Im9-His<sub>6</sub>. Further validation was required to confirm product identity.

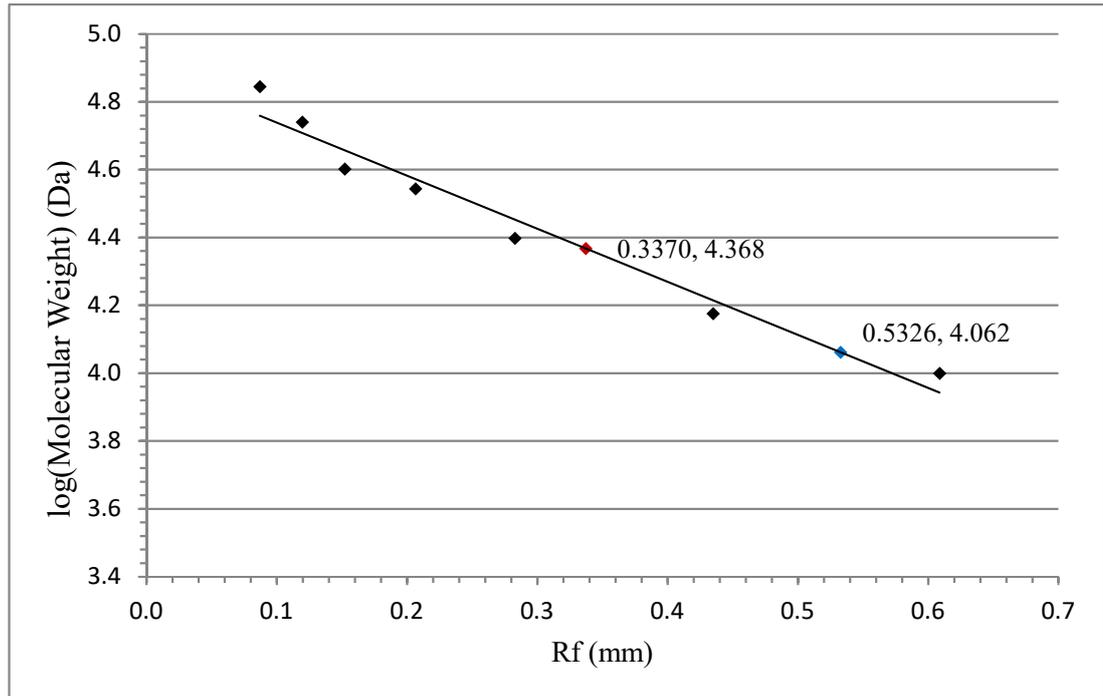
The addition of the N-terminal DNase domain resulted in a soluble product that was subsequently analysed by SDS PAGE and ESI-MS. ESI-MS detected a major peak that suggested a product mass of 23136.2 Da, which is 131 Da less than the expected mass, but consistent with the loss of the N-terminal methionine (Figure 3-6). This loss of the N-terminal methionine is a fairly common occurrence within *E. coli* as methionine aminopeptidase often removes N-terminal residue to improve the stability of certain proteins (Liao *et al.*, 2004). To conclude, ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> was successfully produced and purified, then stored at a concentration of ~61.4 μM. The ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct exists in a soluble form, and was stored effectively in potassium phosphate buffer (See Appendix). This makes it a viable component for our MTP provided the DNase is completely inactivated.



**Figure 3-4 – Analysis of Fractions eluted from HisTrap Column**

**A** - Optical density at 280 nm (solid line) during elution of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>, construct bound by Im9-His<sub>6</sub> tag to the column, was eluted using 6 M guanidine-HCl (See Appendix). The first peak likely represents the protein beginning to elute from the column.

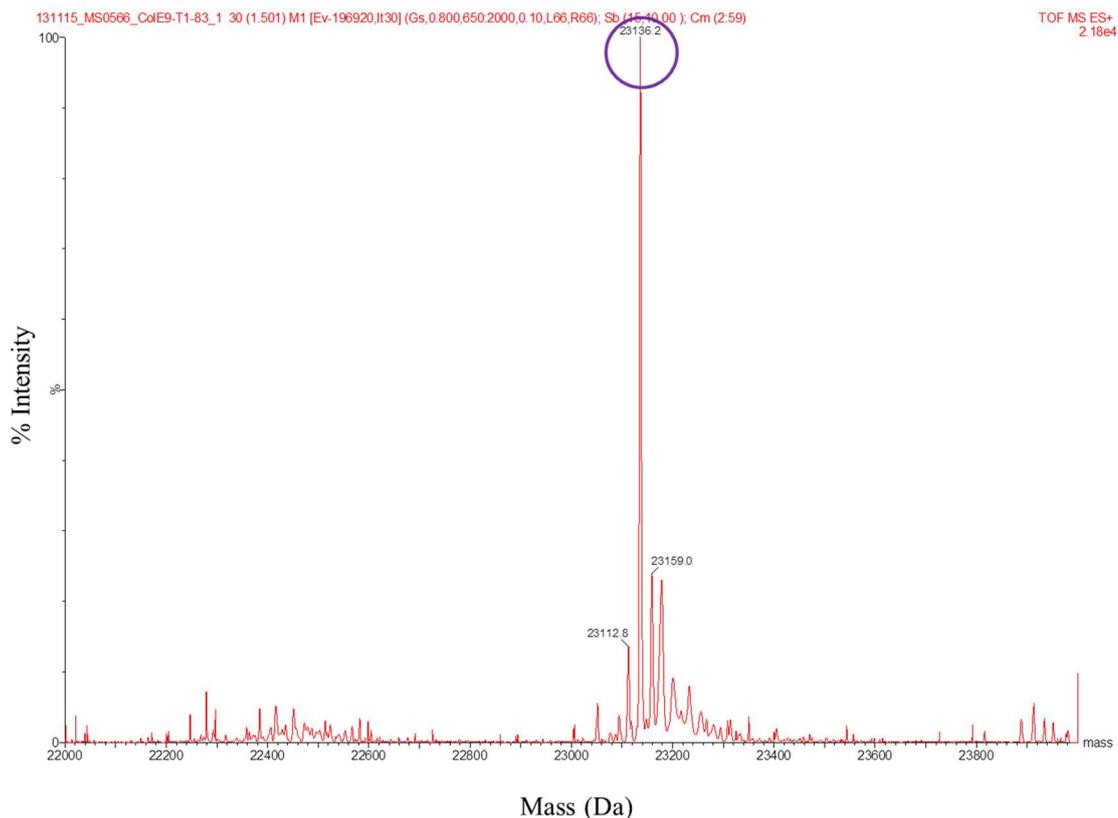
**B** – 12 % (w/v) SDS-PAGE gel analysis of fractions 4, 5, 6, 7, 8, 11 and 13. Fractions (1 ml) were eluted from HisTrap column using 6 M guanidine-HCl. The ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> bands at ~23,000 Da in the fractions coincide with the maximum in the OD<sub>280</sub> signal in panel A.



**Figure 3-5 – Mobility Shift Analysis of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>**

Plot of relative mobility (Rf) against log of molecular weight (Da) of protein bands present during SDS-PAGE analysis of HisTrap elution fractions (Figure 3-4B). Rf is the distance migrated in millimetres from the top of the gel (at well) divided by the length of gel in millimetres (92 mm). Rf was determined for the bands in the protein molecular weight ladder (black), the observed band thought to be T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (red), and the observed band thought to be Im9-His<sub>6</sub> (blue). The molecular weight of the unknown species was estimated using the linear equation of see plot (see below). The measured Rf value for the observed bands was substituted into the x position to determine y (log (MW)).

$$y = -1.548x + 4.895$$



**Figure 3-6 – ESI-MS mass spectrum of purified ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>**

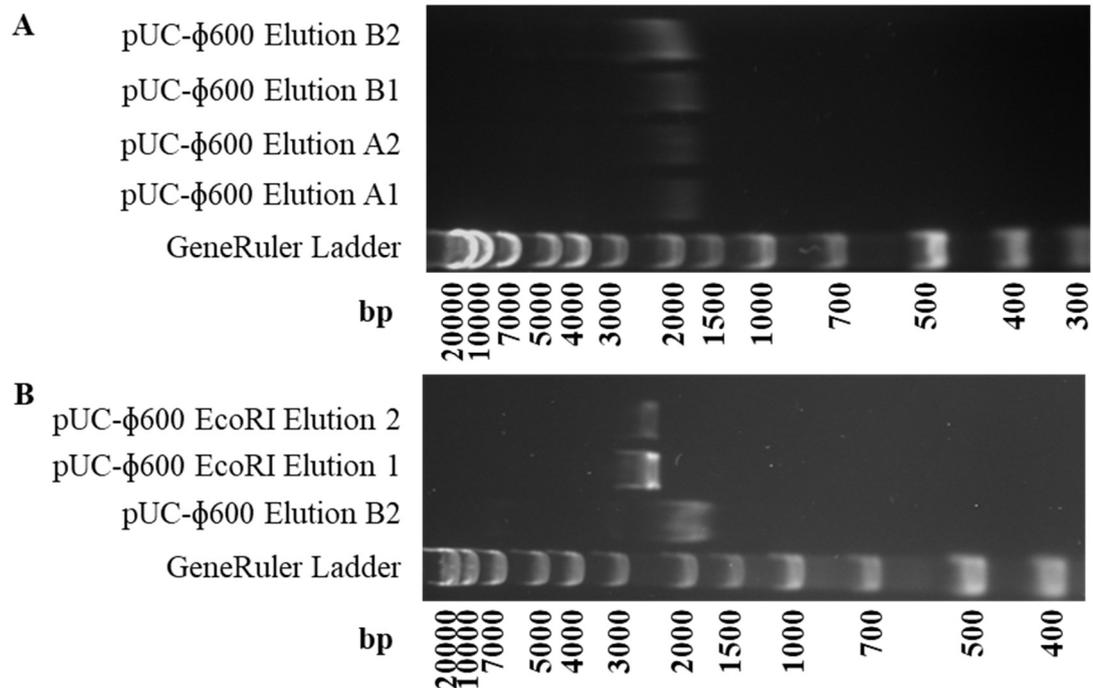
ESI-MS mass spectrum of ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> done in positive mode. The analysis was done on 2  $\mu$ l of sample diluted 75-fold in a MS compatible buffer. Major peak (purple circle) is at 23136.6 Da, which is consistent with the expected peak of 23267.2 Da minus 131 Da for loss of the N-terminal methionine. Peaks are calibrated using a myoglobin external standard.

### 3.2.3 Production and Purification of pUC- $\phi$ 600 Plasmid DNA

The plasmid pUC- $\phi$ 600 was successfully transformed into DH5 $\alpha$  *E. coli* competent cells and extracted using a commercial DNA mini-prep kit. Each sample was eluted twice as per the manufacturer's guide generating four pUC- $\phi$ 600 products: A1, A2, B1 and B2. Samples were analysed on a SybrSafe-stained 1 % (w/v) agarose-TAE gel to confirm pUC- $\phi$ 600 purification (Figure 3-7). The presence of a clear DNA band in each product lane at the same molecular weight confirmed successful purification of supercoiled pUC- $\phi$ 600 plasmid DNA (Figure 3-7).

A mix of B1 and B2 pUC- $\phi$ 600 plasmid DNA stocks were successfully linearised by restriction digest to produce linearised pUC- $\phi$ 600 at a concentration of 53.1 ng/ $\mu$ l (2.665  $\mu$ g DNA was purified). This was confirmed by 1 % (w/v) agarose-TAE gel

electrophoresis with DNA bands in EcoRI-digested samples displaying a mobility shift consistent with linearisation (Figure 3-7). The pUC- $\phi$ 600 plasmid DNA can now be used as a substrate for assaying ColE9 DNase and determining whether the H103A mutation inactivates the DNase.



**Figure 3-7 – Linearisation of pUC- $\phi$ 600 by EcoRI and 1 % (w/v) Agarose-TAE Gel Analysis**

**A** - Purified pUC- $\phi$ 600 plasmid DNA extracted from DH5 $\alpha$  *E. coli* cells visualised on a 1 % (w/v) agarose gel. Plasmid was purified using QIASpin mini prep kit and eluted in nuclease-free water. Each sample, A and B, went through two separate elution steps from separate spin-columns thus generating two purified samples.

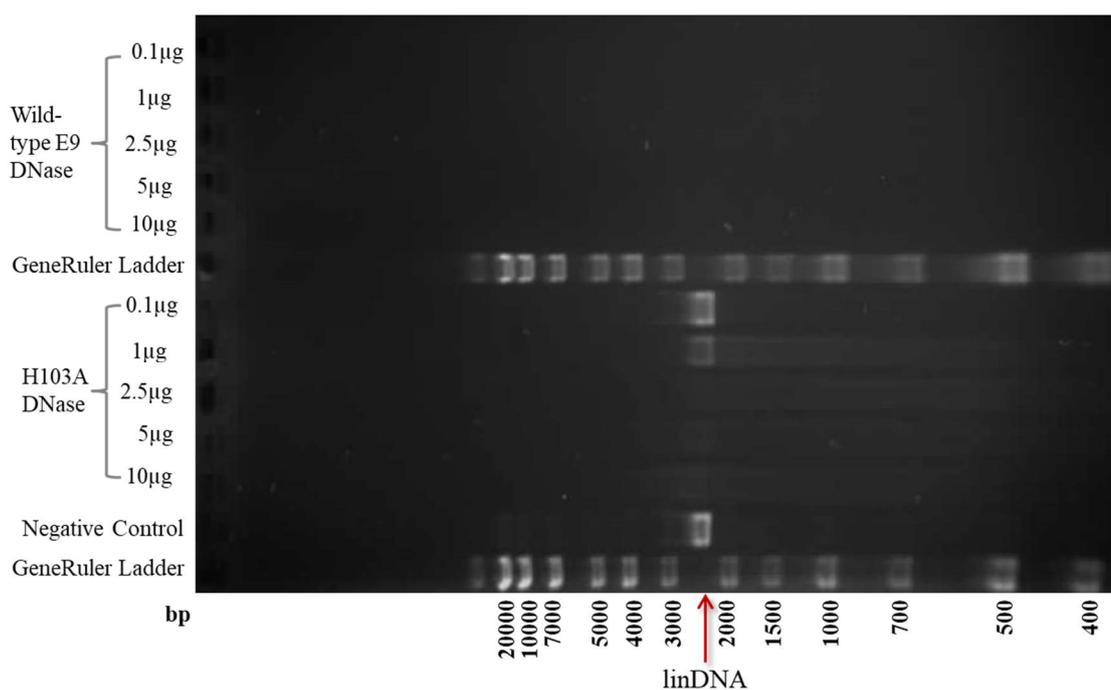
**B** - Restriction digest of pUC- $\phi$ 600 (sample B2) with EcoRI to linearise the supercoiled plasmid DNA. The resultant reaction mixture was purified and eluted in two steps from the spin-column resulting in two separate samples as labelled. The visible shift in molecular weight of the DNA bands exposed to EcoRI indicates that the reaction was successful.

### 3.2.4 Testing Endonuclease Activity of ColE9 T1-83-TEV-DNase<sup>H103A</sup> compared to wild-type ColE9 DNase

Both ColE9 DNase and ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (labelled as “H103A DNase” on agarose gel images) were tested at a range of concentrations (0.1  $\mu$ g - 10  $\mu$ g) against ~100 ng of linearised pUC- $\phi$ 600 plasmid DNA. DNase activity was visualised through the reduction in intensity of DNA bands on agarose gels due to degradation. DNA was completely digested by the wild-type ColE9 DNase regardless of enzyme

concentration, which is indicated by the absence of DNA bands (Figure 3-8, lanes labelled “wild-type Cole9 DNase”).

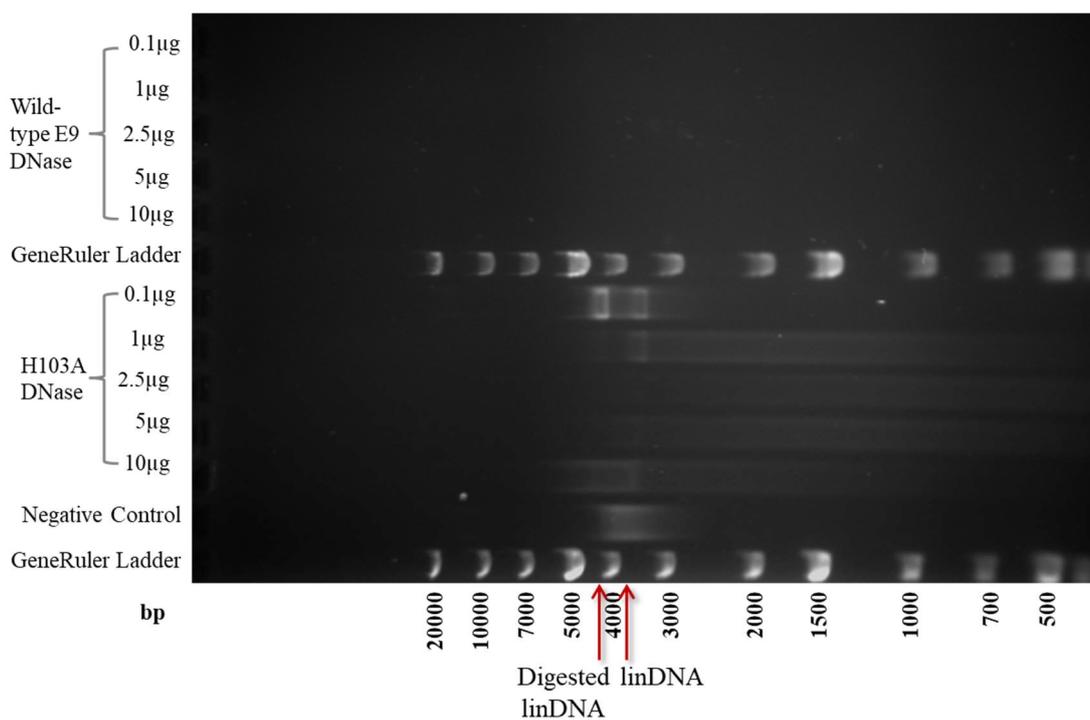
DNA bands were still visible in T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> lanes. The DNA bands became progressively less intense as the amount of DNase added to the assay increased, suggesting that some DNase activity remained. Even in the assay containing the most T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> (10 µg), a faint DNA band was visible. It should be noted that none of the bands present in T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> lanes were as intense as the negative control. The negative control, whilst not containing DNase, was incubated in the exact same buffer as the other reactions for the same length of time. This control possessed a clear band, confirming that sufficient pUC-ϕ600 plasmid DNA was loaded on the 1 % (w/v) agarose-TAE gel. This control band also confirmed that residual endonuclease was successfully removed by the DNA clean-up kit, and did not affect the assay.



**Figure 3-8 – 1 % (w/v) Agarose-TAE Gel of DNA Digestion Assay analysing Cole9 DNase Variants**

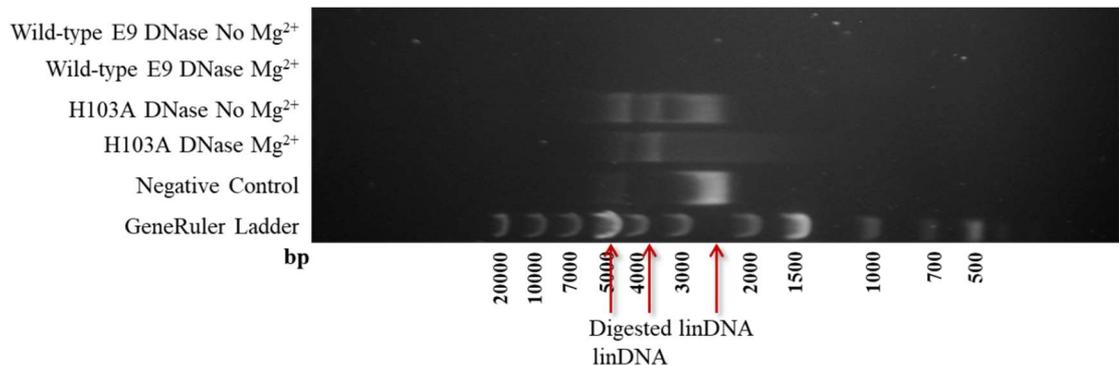
DNA digestion assays of two Cole9 DNase variants using linearised pUC-ϕ600 plasmid DNA as a substrate and analysed on a 1 % (w/v) agarose-TAE gel. The faint bands present in the negative control can be attributed to an overflow of the GeneRuler DNA ladder. Concentration of Mg<sup>2+</sup> used for the reactions was 10 mM. All DNA was digested in the wild-type Cole9 DNase assays, while the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> showed degradation that increased as the amount of DNase present in the sample increased from 0.1 µg (top) to 10 µg (bottom). Linearised DNA is highlighted as linDNA.

The previous assays contained an Mg<sup>2+</sup> concentration of 10 mM as Mg<sup>2+</sup> is a cofactor of ColE9 DNase. The M9 media used for live cell fluorescence microscopy imaging contains 2 mM Mg<sup>2+</sup>. A digestion assay was done with 2 mM Mg<sup>2+</sup> to investigate if the H103A mutation would be sufficient to inactivate T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> at a 5-fold lower Mg<sup>2+</sup> concentration. As observed previously, wild-type ColE9 DNase assays showed a complete absence of linearised pUC- $\phi$ 600 plasmid DNA in all samples (Figure 3-9). In contrast, undigested DNA bands were present in the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> assays. The intensity of linearised pUC- $\phi$ 600 plasmid DNA bands generally decreased as the amount of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> present increased. It is difficult to properly quantify the amount of undigested DNA due to the diffuse nature of the bands. The negative control is blurred, potentially due to degradation of stock DNA. However, in contrast to the equivalent assays with 10 mM Mg<sup>2+</sup>, the assays with 1-10  $\mu$ g T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> added retained more undigested DNA.



**Figure 3-9 – 1 % (w/v) Agarose-TAE Gel Analysis of Digestion Assay with 2 mM Mg<sup>2+</sup>**  
 Digestion assay of two DNase variants using linearised pUC- $\phi$ 600 DNA as a substrate analysed on a 1 % (w/v) Agarose-TAE gel. Concentration of Mg<sup>2+</sup> used for the reactions was 2 mM. All DNA was digested in the wild-type ColE9 DNase assays. T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> degraded the DNA less and the amount of degradation increased with the amount of DNase present in the assay from 0.1  $\mu$ g (top) to 10  $\mu$ g (bottom). Degradation by ColE9 H103A DNase mutant was slightly less effective when compared to the DNase assay at 10 mM Mg<sup>2+</sup> (Figure 3-8). Incomplete linearisation of stock used may have caused smearing of the control band. Linearised DNA is highlighted as linDNA.

It appeared that the lower Mg<sup>2+</sup> concentration further inhibited T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>. A separate experiment assayed DNase activity of both variants in the presence and absence of Mg<sup>2+</sup> (10 mM). For this assay, a fixed amount of DNase (2.5 µg) and linearised pUC-ϕ600 was utilised. Wild-type Cole9 DNase remained active with and without added Mg<sup>2+</sup> as all plasmid DNA was digested (Figure 3-10). T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> assays showed reduced amounts of undegraded linear pUC-ϕ600 DNA, in contrast to the negative control, under both conditions. DNase displayed, as expected, greater activity in the presence of Mg<sup>2+</sup> ions, with bands less intense in lanes with Mg<sup>2+</sup> present. Overall, while the H103A mutation has partially inactivated T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>, it still maintains significant endonuclease activity in the presence and absence of Mg<sup>2+</sup>.



**Figure 3-10 – 1 % (w/v) Agarose-TAE Gel Analysis of Digestion Assay with and without added Mg<sup>2+</sup>**

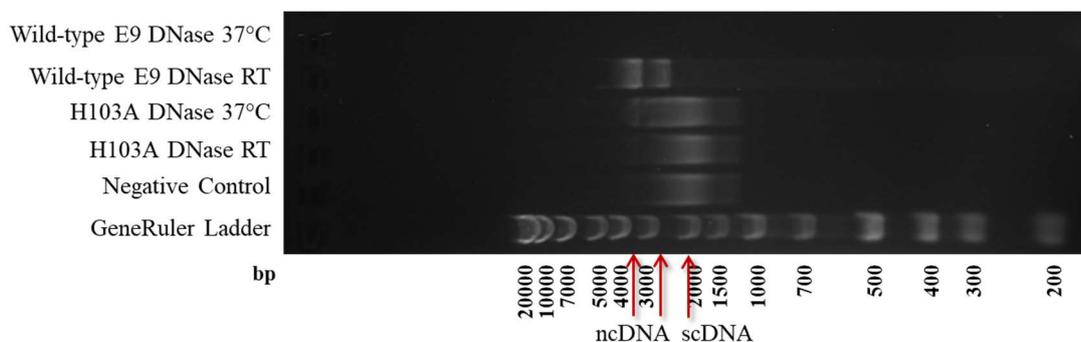
Digestion assay of DNase variants using linearised pUC-ϕ600 DNA substrate analysed on a 1 % (w/v) agarose-TAE gel. Concentration of Mg<sup>2+</sup> was 10 mM. DNA was completely digested in the wild-type Cole9 DNase assays, while assays with T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> showed some DNA degradation when compared to the negative control. Removal of Mg<sup>2+</sup> did not prevent DNA degradation by T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>. Incomplete linearisation of stock may have caused smearing of the control band. Linearised DNA is highlighted as linDNA.

### 3.2.5 Single-Strand DNA Nicking Assay of Cole9 DNase Variants utilising supercoiled pUC-ϕ600 plasmid DNA

DNase activity was assayed using supercoiled pUC-ϕ600 plasmid DNA to determine if single-stranded DNA nicking occurred, i.e. the ability of the DNase to cleave one strand of supercoiled plasmid DNA. It has the advantage of being more sensitive to lower enzyme concentrations than standard digestion assays (Pommer *et al.*, 2001). Wild-type Cole9 DNase and T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> were assessed for their ability to

nick and degrade supercoiled pUC- $\phi$ 600 plasmid DNA. The background concentration of Mg<sup>2+</sup> was 10 mM. Samples were analysed following incubation at either RT or 37 °C for 25 min. Wild-type Cole9 DNase after incubation at 37 °C had completely degraded all supercoiled pUC- $\phi$ 600 plasmid DNA present. This was evidenced by a complete lack of undigested supercoiled plasmid DNA (Figure 3-11).

Following incubation at RT, the supercoiled pUC- $\phi$ 600 plasmid DNA was merely nicked by the wild-type Cole9 DNase, which was evidenced by the presence of an additional band consistent with nicked closed circular plasmid DNA (Figure 3-11). T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> showed very little difference relative to the negative control. The pUC- $\phi$ 600 plasmid DNA in this test appeared as smeared bands, though this may be due to degradation of stock DNA. The lane containing wild-type Cole9 DNase incubated at RT showed better defined DNA bands suggesting supercoiling was involved (Figure 3-11). The agarose-TAE gel analysis of supercoiled pUC- $\phi$ 600 plasmid DNA also showed smeared bands (Figure 3-7). It appeared that the H103A mutation had a significant effect on the ability of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> to degrade and nick supercoiled plasmid DNA. However, T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> incubated at 37°C still showed some DNase activity (Figure 3-11).



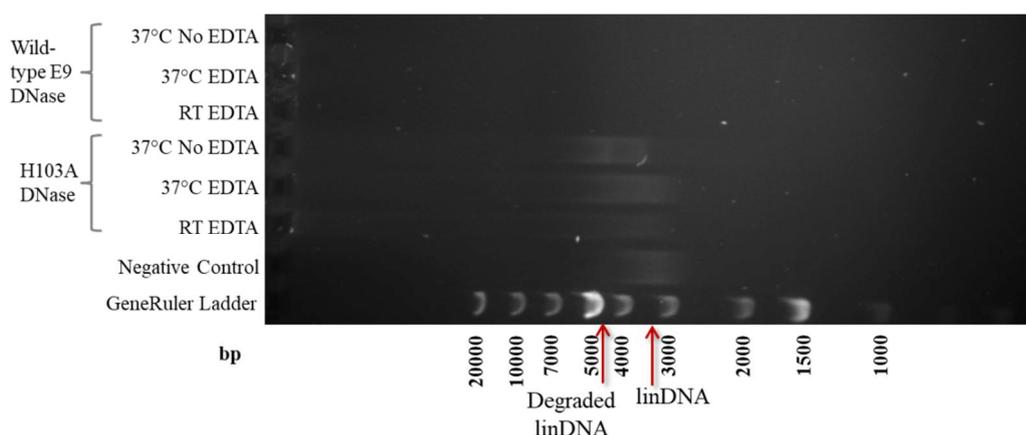
**Figure 3-11 – 1 % (w/v) Agarose-TAE Gel Analysis of Single-Strand DNA Nicking Assay for DNase variants**

Nicking assay for the two DNase variants using supercoiled pUC- $\phi$ 600 plasmid DNA as the substrate analysed on a 1 % (w/v) agarose-TAE gel. DNA was completely digested by wild-type Cole9 DNase after incubation at 37 °C. Wild-type Cole9 DNase showed evidence of DNA nicking even with incubation at RT. T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> showed little evidence of degradation compared to the negative control with the exception of the sample incubated at 37°C. Supercoiled DNA is highlighted as scDNA; nicked circular DNA is highlighted as ncDNA.

### 3.2.6 Inhibiting Wild-Type Cole9 DNase and T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Endonuclease Activity with EDTA

It was speculated that the remaining endonuclease activity of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> could be inhibited by addition of EDTA. Sufficient amounts of EDTA (~125 mM), if proven to inhibit DNase activity completely, could then be added to future storage buffers for T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> coupled to single-stranded DNA. However, this approach would not work for future fluorescence microscopy experiments where the addition of millimolar concentrations of Mg<sup>2+</sup> is required for bacterial cell growth. Samples containing wild-type Cole9 DNase with added EDTA lacked DNA bands (Figure 3-12). The concentration of Mg<sup>2+</sup> used was 10 mM.

Even with incubation at RT, wild-type Cole9 DNase displayed significant endonuclease activity against the pUC- $\phi$ 600 plasmid DNA substrate. There were undigested pUC- $\phi$ 600 DNA bands present in T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> samples, though these appeared less intense than the negative control (Figure 3-12). For unknown reasons, the linear pUC- $\phi$ 600 DNA bands, including those in the negative control, appeared very diffuse and smeared. It was concluded that EDTA had no discernible affect on the endonuclease activity of wild-type Cole9 DNase and T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>.



**Figure 3-12 – 1 % (w/v) Agarose-TAE Gel Analysis of Endonuclease Activity for DNase variants with and without added EDTA**

Endonuclease activity assays for two DNase variants with and without added EDTA using linearised pUC- $\phi$ 600 plasmid DNA as a substrate and analysed on a 1 % (w/v) agarose-TAE gel. All wild-type samples showed complete degradation of linear pUC- $\phi$ 600 plasmid DNA. T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> samples displayed less DNA degradation. Undigested linear pUC- $\phi$ 600 plasmid DNA appeared smeared in all samples.

### 3.3 Discussion

#### 3.3.1 Purification and Characterisation of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Construct

The T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct can be easily purified using a previously established protocol. A 500 ml overnight culture of DH5 $\alpha$  bacterial cells in LB-Amp media produced ~10 mg of purified protein. If the DNase activity of this construct can be inactivated, while maintaining its ability to bind OmpF, it could be integrated into the MTP. The ability to produce T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> on demand is vital for future experiments and assembly of the MTP. It also allows for potential further sequence modifications via mutagenesis of the pRP1 plasmid DNA. It was also easily analysed by ESI-MS, producing an easily interpreted spectrum with clearly identifiable mass peaks. This allows any future modifications to be confirmed and characterised using ESI-MS.

#### 3.3.2 Assaying the DNase Inactivating H103A Mutation in the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> Construct

The DNase assays conducted all provided evidence that the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct retained residual endonuclease activity regardless of the assay conditions or DNA substrate. Furthermore, T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> remained active with the Mg<sup>2+</sup> concentration of M9 media used for live-cell fluorescence microscopy (2 mM Mg<sup>2+</sup>). This meant that T<sup>1-83</sup>-TEV-H103A DNase would still be active during FRET experiments utilising the MTP, which consists of ssDNA oligonucleotide components. According to previous work, ColE9 DNase is able to digest ssDNA with an even greater efficiency than the double-stranded DNA assayed here (Pommer *et al.*, 2001).

The single-strand DNA nicking assay results suggested some inactivation of the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> compared to double-stranded DNA digestion assays. However, the supercoiled pUC- $\phi$ 600 plasmid DNA substrate used is not exactly comparable to the oligonucleotide components of the MTP. Based on the results from these assays further steps must be taken to inactivate the DNase domain before T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> can be used as a component of the MTP.

EDTA is a metal-chelating chemical compound; it will sequester divalent metal ions present, e.g. Mg<sup>2+</sup>. An excess concentration of 125 mM was theorised to completely

inhibit the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> by removing its metal cation cofactor. Attempts to use EDTA to inhibit the endonuclease activity of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> proved unsuccessful (Figure 3-12). The linear pUC- $\phi$ 600 DNA bands that were present appeared diffuse and smeared, even in the negative control. Given that pUC- $\phi$ 600 plasmid DNA in other assays did not display this effect, it was hypothesised that the excess EDTA added may be the culprit. T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> still possessed residual endonuclease activity against linear pUC- $\phi$ 600 plasmid DNA in the absence of Mg<sup>2+</sup> (Figure 3-10). The DNase activity observed in absence of added Mg<sup>2+</sup> was consistent with the inability of added EDTA to inhibit T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>. As a result, EDTA cannot be used in future attempts to inactivate the DNase; a mutation-based strategy appears to be the only effective means of completely inactivating the ColE9 DNase.

### 3.3.3 Future Work

The H103A mutation has not completely inactivated the ColE9 DNase domain, thus a new strategy must be introduced to enable the use of the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct in the MTP. The reasons for this are unclear. The H103A DNase mutant does retain the coordination of the Mg<sup>2+</sup> ion, which may result in retention of weak endonuclease activity (Walker *et al.*, 2002). One solution would be to introduce another mutation, alongside H103A, to completely remove any remaining endonuclease activity from T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>. His127 (residue 575 for full-length colicin E9) is an essential residue for the endonuclease activity of ColE9 DNase. A His127Ala mutation has been shown in a bacterial cell killing assays to result in a colicin-insensitive phenotype (Garinot-Schneider *et al.*, 1996). This study used competitive binding assays to confirm that this mutation does not interfere with colicin E9 binding to BtuB making it suitable for future MTP studies.

Future strategies should incorporate the H127A mutation into T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct. Site-directed mutagenesis could be used to introduce the necessary mutation into the pRP1 plasmid. Primers have already been designed for site-directed mutagenesis by PCR using the pRP1 plasmid DNA as a template (Table 3-1). After over-production and purification, the T<sup>1-83</sup>-TEV-DNase<sup>H103A/H127A</sup> construct can be assessed for DNase activity using the endonuclease assays presented here.

In addition, a method for producing large quantities of a ssDNA substrate for endonuclease assays should be investigated. This would provide a more accurate picture of whether the DNase would compromise the MTP, as the oligonucleotide components themselves are single-stranded. ColE9 DNase struggles to bind and digest shorter length double-stranded DNA; the oligo components annealed to form 40 bp duplexes, thus they may escape degradation (Pommer *et al.*, 2001). Should the new mutation result in the complete inactivation of the DNase, the T<sup>1-83</sup>-TEV-DNase<sup>H103A/H127A</sup> construct can then be used as a component in the MTP. Thiol coupling experiments would follow to ensure it is able to couple to the oligonucleotide component.

**Table 3-1 - PCR Primers for site-directed mutagenesis of the pRP1 plasmid to create the T<sup>1-83</sup>-TEV-DNase<sup>H103A/H127A</sup> construct**

Primers designed to introduce the H127A mutation into the pRP1 plasmid via site-directed mutagenesis. Oligonucleotide DNA sequence, percent GC content, length in bp, and melting temperature are all provided. The codon for the alanine residue is highlighted in red.

Primer	DNA Sequence	Length (bp)	GC Content (%)	Melting Temperature (°C)
Forward	5'- CGAGTGACTACACCTAAGCGA <b>GCG</b> ATCGATATTCACCGAGGTAAG-3'	45	51.11	87.46
Reverse	5'- CTTACCTCGGTGAATATCGAT <b>CGCT</b> CGCTTAGGTGTAGTCACTCG -3'	45	47.62	84.95

Results reported here describe reproducible protocols for the over-production and purification of the T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct. Protocols have also been established to assay DNase activity for this construct. In the future these assays can be used to assess whether additional mutations in the ColE9 DNase domain remove the observed residual endonuclease activity. This will facilitate progress towards assembly of a fully functional MTP for FRET studies of outer membrane translocation by colicin E9.

## Chapter 4: Developing a Thiol Coupling Protocol for the Molecular Tension Probe using Colicin E9 R-Domain

### 4.1 Introduction

ColE9 R-domain<sup>317-454</sup> (Y324C and L447C, N-terminal Cys at position 318) will form the second half of the molecular tension probe (MTP) (Figure 4-1). As part of the MTP, the R-domain will bind to BtuB and localise it at the cell surface. Cysteine residues were introduced at positions 324 and 447 of the ColE9 R-domain resulting in the formation of a disulfide bond. This ‘top-lock’ disulfide was incorporated to prevent unfolding of the coiled-coil R-domain structure at the cell surface and the subsequent translocation across the OM. Should mechanical unfolding forces be generated at the cell surface, the anchored R-domain will generate tension in the MTP at the OM. This tension will mechanically unzip the DNA duplex within the MTP and lead to separation of the donor-acceptor Förster Resonance Energy Transfer (FRET) fluorophore pair.

Another cysteine mutation was engineered at position 318 of ColE9 during previous work by Dr. Patrice Rassam. This cysteine was introduced to enable conjugation to the ssDNA oligonucleotide component of the MTP via a disulfide link. Generation of protein-DNA chimeras between thiol groups has been accomplished by other groups (Cecconi *et al.*, 2008). Maleimide groups are highly selective for thiol groups compared to other nucleophilic side chains, e.g. lysine amine groups (Smith *et al.*, 2010). Using this property, the R-domain can be selectively labelled at a specific cysteine residue. Recent work has incorporated bromomaleimide reactive groups, allowing for reversible modification through treatment with an excess of reducing agent if required (Tedaldi *et al.*, 2009). Using these established protocols for bioconjugation, it was attempted to couple the ColE9 R-Domain to ssDNA oligonucleotide at Cys318.

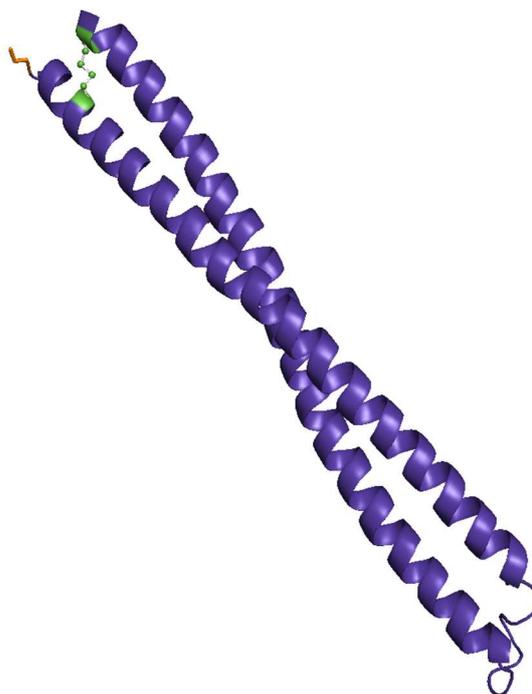
N-methyl-bromomaleimide (N-MBM) was chosen as the maleimide reactive group for the bioconjugation. The ability to reverse any covalent modification formed by treatment with extreme reducing conditions, e.g. 100 molar equivalent of  $\beta$ -mercaptoethanol, is extremely useful (Smith *et al.*, 2010). This allowed for recycling

of R-domain material when necessary. Following successful labelling of R-domain with N-MBM, conjugation of the ssDNA oligonucleotide handle to the maleimide label would then be attempted (Figure 4-2). A positive result would confirm the viability of the thiol conjugation protocol for assembling the MTP.

Initial experiments would attempt coupling of N-MBM-labelled R-domain to another free R-domain molecule. This prevented unnecessary wastage of expensive oligonucleotide materials whilst the thiol conjugation protocol was refined. To confirm N-MBM labelling, and subsequent conjugation with R-domain, reaction products were analysed by positive-mode ESI-MS and SDS-PAGE (Kent, 1995). A limitation of SDS-PAGE analysis was that the increase in mass from successful N-MBM-labelling (+110 g/mol) would be difficult to resolve as a mobility shift. ESI-MS, being a more precise technique, was incorporated to make up for this shortfall. If the R-domain-N-MBM-R-domain conjugation product was identified, conjugation of N-MBM-labelled R-domain to its respective ssDNA oligonucleotide component would be attempted (Col E9 R-domain Coupling 5' Thiol).

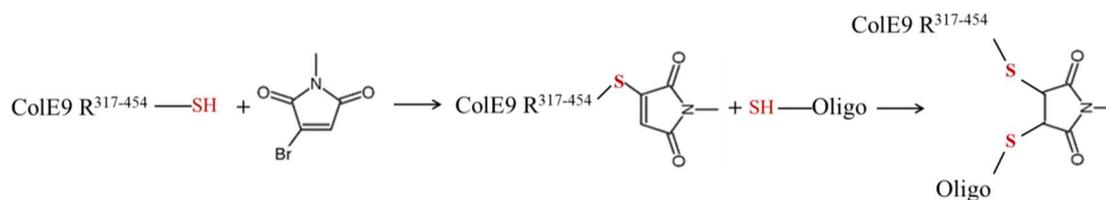
Not all proteins are capable of generating a clear, discernible ESI-MS spectrum. Should positive-mode ESI-MS fail to detect the R-domain, alternative methods of analysis were required to confirm successful labelling of Cys318. Exposure of Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB) to free thiol groups, like Cys318 in unlabelled R-domain, results in the formation of a mixed disulphide and a free TNB<sup>2-</sup> ion. The release of TNB<sup>2-</sup> ions leads to a measurable change in the molar absorptivity at 412 nm ( $\Delta\epsilon_{412}$ ). The concentration of TNB<sup>2-</sup> ions is directly equivalent to the concentration of thiol groups in the sample. Ellman's reagent-based assays were used to calculate free [SH] in R-domain samples pre- and post-N-MBM labelling. A calculated [SH] of  $\sim 0 \mu\text{M}$  is indicative of a successful labelling protocol.

Following labelling with N-MBM, conjugation of R-domain to its respective ssDNA oligonucleotide handle would be attempted. Successful formation of the R-domain-ssDNA chimera would be confirmed by ESI-MS and/or native gel electrophoresis. Experiments were designed to refine and confirm a viable coupling protocol for assembling the MTP. Optimal conjugation reaction conditions elucidated in these experiments could then be repeated for coupling of the ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct.



**Figure 4-1 – Structure of R-Domain**

R-Domain visualised in 3D using PyMOL. R-domain forms a coiled-coil tertiary structure consisting of two  $\alpha$  helices coiled around one another. The cysteine residue at position 318, the labelling site, is coloured in orange and shown in stick form with its thiol side chain. The two incorporated cysteine residues that form the disulphide bridge (Cys324 and Cys447) are coloured in green, with the disulfide bridge ‘top-lock’ shown as spheres.



**Figure 4-2 – Diagram of R Domain-Oligo Conjugation utilising N-MBM**

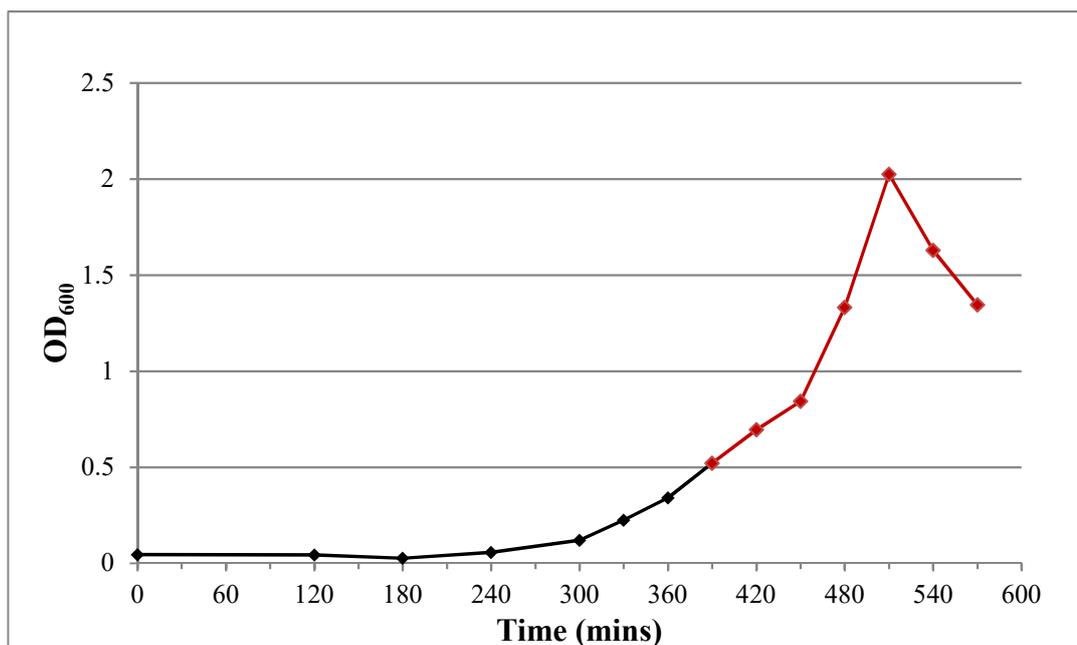
Representative diagram of the planned protocol for the conjugation of R-domain with an ssDNA oligonucleotide handle (oligo) to form a protein-ssDNA chimera. N-MBM is used as a motif to first label the free thiol group in the R-domain (Cys318), then to conjugate to the thiol group of its respective oligonucleotide. The thiol groups utilised in the reaction protocol are highlighted in red.

## 4.2 Results

### 4.2.1 Transformation of *E. coli* cells with plasmid pREN1

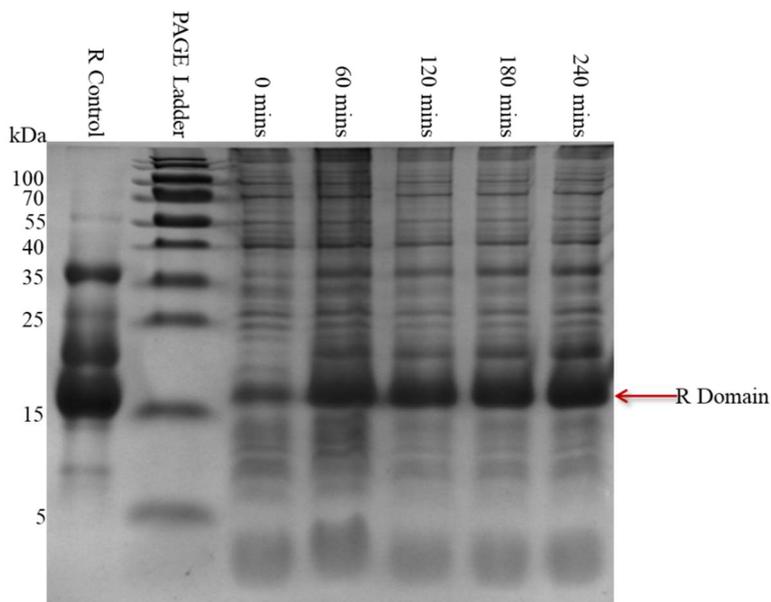
Plasmid pREN1 was purified from transformed competent *E. coli* cells (DH5 $\alpha$ ) grown with ampicillin selection. Purified pREN1 plasmid was analysed on a 1 % (w/v) agarose gel (Figure 2-1) as defined bands indicating that the plasmid was intact. Following reference to DNA ladder (GeneRuler 1 kb Plus, Thermo) the observed band was estimated to be  $\sim 3000$  bp. This did not match the expected length of the plasmid, 5805 bp (Figure 2-1).

Addition of IPTG to transformed *E. coli* BL21(DE3) cells induced over-production of the R-domain. Cultures generated a visible log growth phase (Figure 4-3) and growth was continued following addition of IPTG to induce R-domain over-production. Analysis of cell culture aliquots sampled during cell growth by SDS-PAGE (Figure 4-4) showed a protein band at a molecular weight of  $\sim 17,000$  Da that increased in intensity over time. This was indicative of successful over-production of the R-domain product (expected MW =  $\sim 17,381$  Da).



**Figure 4-3 – Growth curve for *E. coli* BL21(DE3) bacterial cells transformed with pREN1 plasmid during induction with IPTG**

Optical density (OD<sub>600</sub>) was measured as a function of time (min) for *E. coli* BL21(DE3) bacterial cells transformed with the plasmid pREN1 and engaged in logarithmic growth. Black data points are readings taken pre-induction with IPTG addition, and red data points are taken post-induction with IPTG.



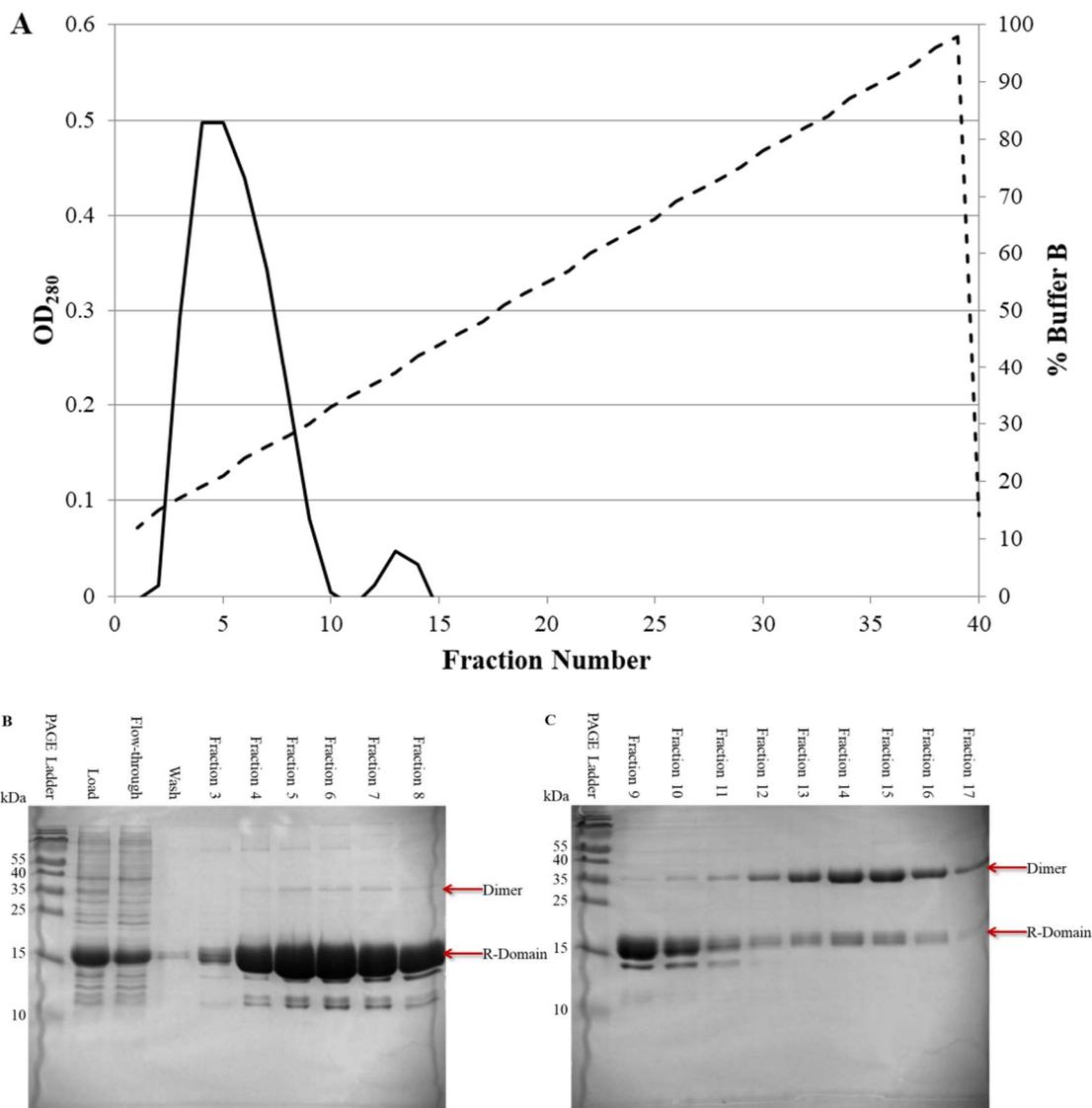
**Figure 4-4 – 15 % (w/v) SDS-PAGE analysis of R-domain over-production in *E. coli* BL21(DE3) bacterial cells**

15 % (w/v) SDS-PAGE gel analysis of cell from over-production of R-domain in *E. coli* BL21(DE3) cells. Induction was carried out over 4 hr with aliquots taken at regular time points. The band that increased in intensity over time at ~17,000 Da is attributed to R-domain.

#### 4.2.2 Purification and Characterisation of ColE9 R-domain

Fractions eluted during Ni<sup>2+</sup>-affinity chromatography indicated the presence of pure R-domain following analysis by SDS-PAGE and spectrophotometry. Fractions 3 to 17 were deemed to potentially contain purified protein. OD<sub>280</sub> readings form multiple absorbance peaks during elution of fractions. The absorbance peak of greatest intensity was measured during elution of fractions 3 to 10 (Figure 4-5A). These fractions were analysed by SDS-PAGE, revealing an intense band at ~ 17,000 Da, mirroring the post-induction SDS-PAGE gel (Figures 4-5B and 4-5C). However, an additional, less intense peak was also present during elution of fractions 12 to 17 (Figure 4-5A). SDS-PAGE analysis revealed protein bands within these fractions with a molecular weight of ~ 34,000 Da (Figure 4-5C). This is double the expected MW of R-domain, suggesting the presence of a dimeric R-domain population.

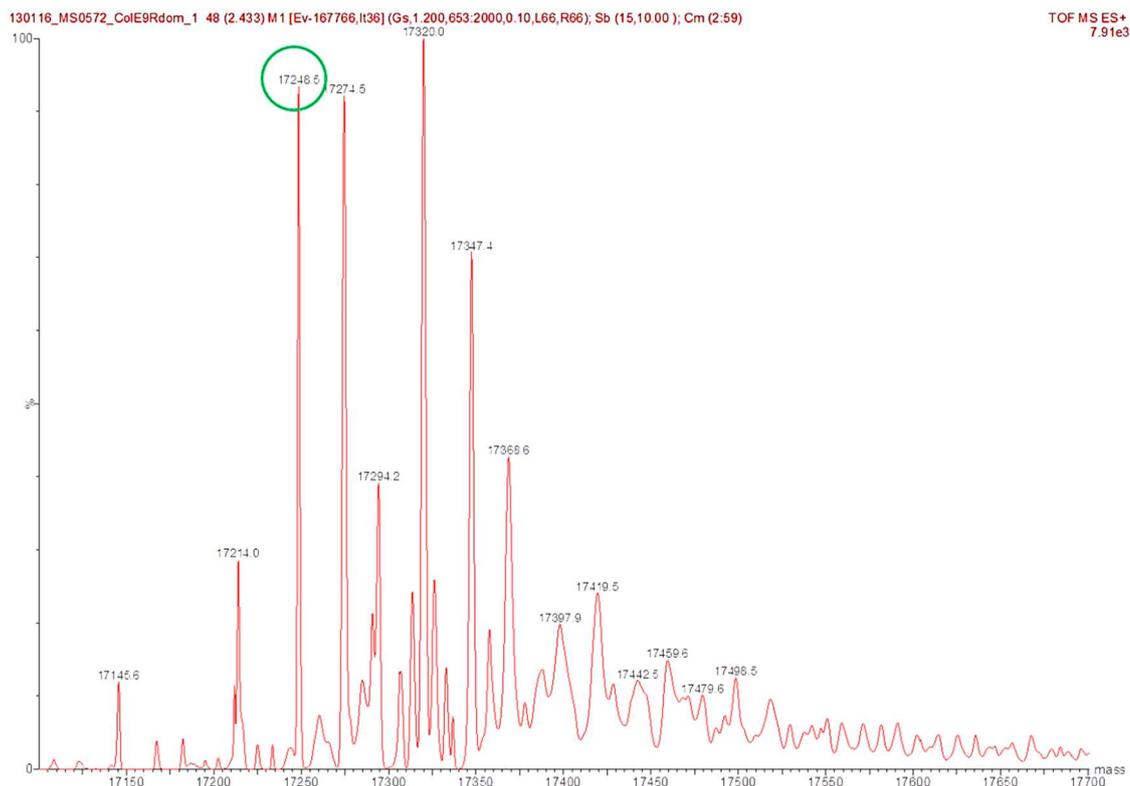
Fractions 3 to 11 contained a significant monomeric population of R-domain and were combined (~9 ml). Nanodrop<sup>TM</sup> spectrophotometry was used to determine the concentration of protein (~ 154 µM). Multiple forms of the ionised protein were detected by positive-mode ESI-MS in the pooled fractions. This complicated identification of the ColE9 R-domain by ESI-MS analysis. ESI-MS detected major peaks at 17249.2 Da, 17275.2 Da, 17294 Da, 17320.7 Da, 17348.1 Da and 17369.3 Da (Figure 4-6). The peak at 17249.2 Da most likely represents ColE9 R-domain, consistent with the expected mass of R-domain minus the N-terminal methionine (-131 Da).



#### Figure 4-5 –Analysis of Eluted FPLC Fractions by OD<sub>280</sub> and SDS-PAGE

**A-** Optical density at 280 nm (OD<sub>280</sub>, solid line) during elution of R-domain bound to the HisTrap Ni<sup>2+</sup> affinity column by His<sub>6</sub> tag. R-domain was eluted using an imidazole gradient by increasing amount of Buffer B (See Appendix) from 2 % to 100 % (dashed line). Protein was detected within fractions 3 to 17. The two peaks likely represent protein eluting from the column.

**B/C-** 15 % (w/v) SDS PAGE analysis of fractions 3 to 17 (B- 3 to 8, C- 9 to 17) eluted from the HisTrap Ni<sup>2+</sup> affinity column using an imidazole gradient (10 mM to 500 mM). The R-domain band at ~17,000 Da in fractions 3 to 11 coincide with the maximum OD<sub>280</sub> signal in panel A. R-domain dimer bands at ~35 kDa coincide with the second peak in OD<sub>280</sub> signal in panel A.



**Figure 4-6 – ESI-MS mass spectrum of purified ColE9 R-domain**

ESI-MS mass spectrum of R-domain done in positive mode. The analysis was done on 1  $\mu$ l of sample diluted 500-fold in a MS compatible buffer. The peak circled in green is corrected to 17249.2 Da, which is consistent with the expected peak of 17381.3 Da minus 131 Da for loss of the N-terminal methionine. Peaks are calibrated using a myoglobin external standard.

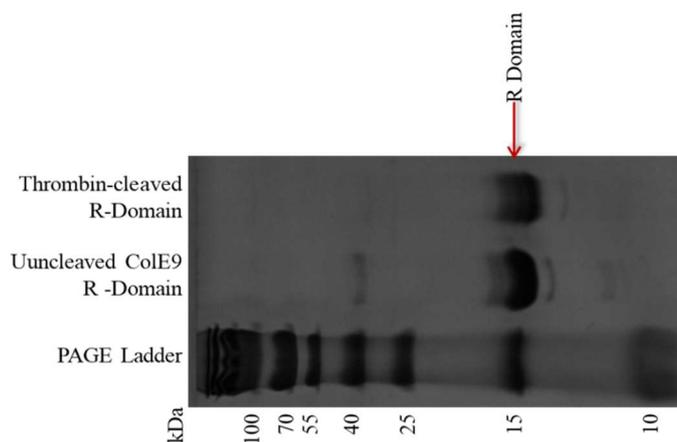
#### 4.2.3 Removing the Purification Tag and Dimer population of R-domain

R-domain was subject to cleavage by thrombin to remove its His<sub>6</sub> tag. Reaction product was analysed by SDS-PAGE alongside unreacted R-domain. It displayed a mobility shift between the two bands indicative of successful cleavage (Figure 4-7). The expected reduction in molecular weight for cleaved R-domain was ~1,227 Da. The cleavage product was analysed by positive-mode ESI-MS to confirm the molecular weight. ESI-MS detected a greater number of ionised protein forms compared to the uncleaved sample (Figure 4-8). The detected mass peak closest to the expected mass of ~16,040.8 Da was the peak of greatest intensity within the spectrum, measured at 16,048 Da. The peak was most likely due to protonation of residues with charged side chains, e.g. arginine and lysine. The peak at ~16,048 Da is consistent with

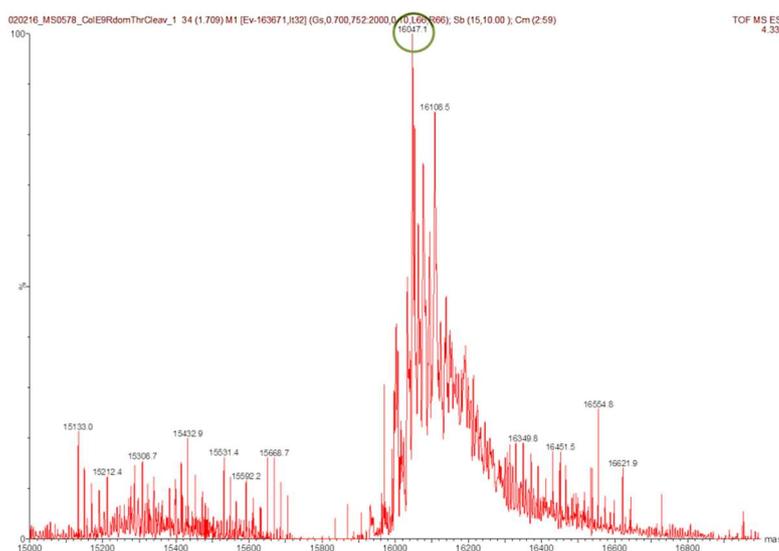
the increase in mass should all acidic and basic side chains of R-domain be protonated (+8.04 Da).

Cleaved R-domain was loaded onto a size exclusion column to remove both residual thrombin and the inherent dimer population. Following analysis by SDS-PAGE,  $\beta$ -Mercaptoethanol ( $\beta$ -ME) was judged to be the most effective reducing agent to reduce the inherent dimer population of R-domain (see Appendix). Cleaved R-Domain was reacted with  $\beta$ -ME prior to loading to remove as much of the dimer population as possible. OD<sub>280</sub> measurements of eluents indicated the presence of three protein species within the samples (Figure 4-9A). The first, and least intense, peak likely represented thrombin. As the protein of the greatest MW, it was expected that thrombin would elute off the size exclusion column first. The next peak (Fractions 20 to 28) likely represented dimer, expected to elute off second. Finally, the most intense peak (Fractions 34 to 48) represented the elution of monomeric, cleaved R-domain.

Analysis of fractions by SDS-PAGE confirmed the identity of R-domain populations (Figure 4-9B/C). Fractions 20 to 28 contained a protein species with a MW of  $\sim 32,000$  Da, confirming the presence of dimer. Fractions 34 to 48 contained a protein species with a MW of  $\sim 16,000$  Da, indicative of thrombin-cleaved R-domain monomer. Following SDS-PAGE analysis, fractions 37 to 45 were combined to form a stock. From the initial uncleaved R-domain stock of  $\sim 9$  ml,  $\sim 36$  ml of cleaved R domain was produced, the concentration of which was determined by Nanodrop<sup>TM</sup> spectrophotometry to be  $\sim 37$   $\mu$ M.

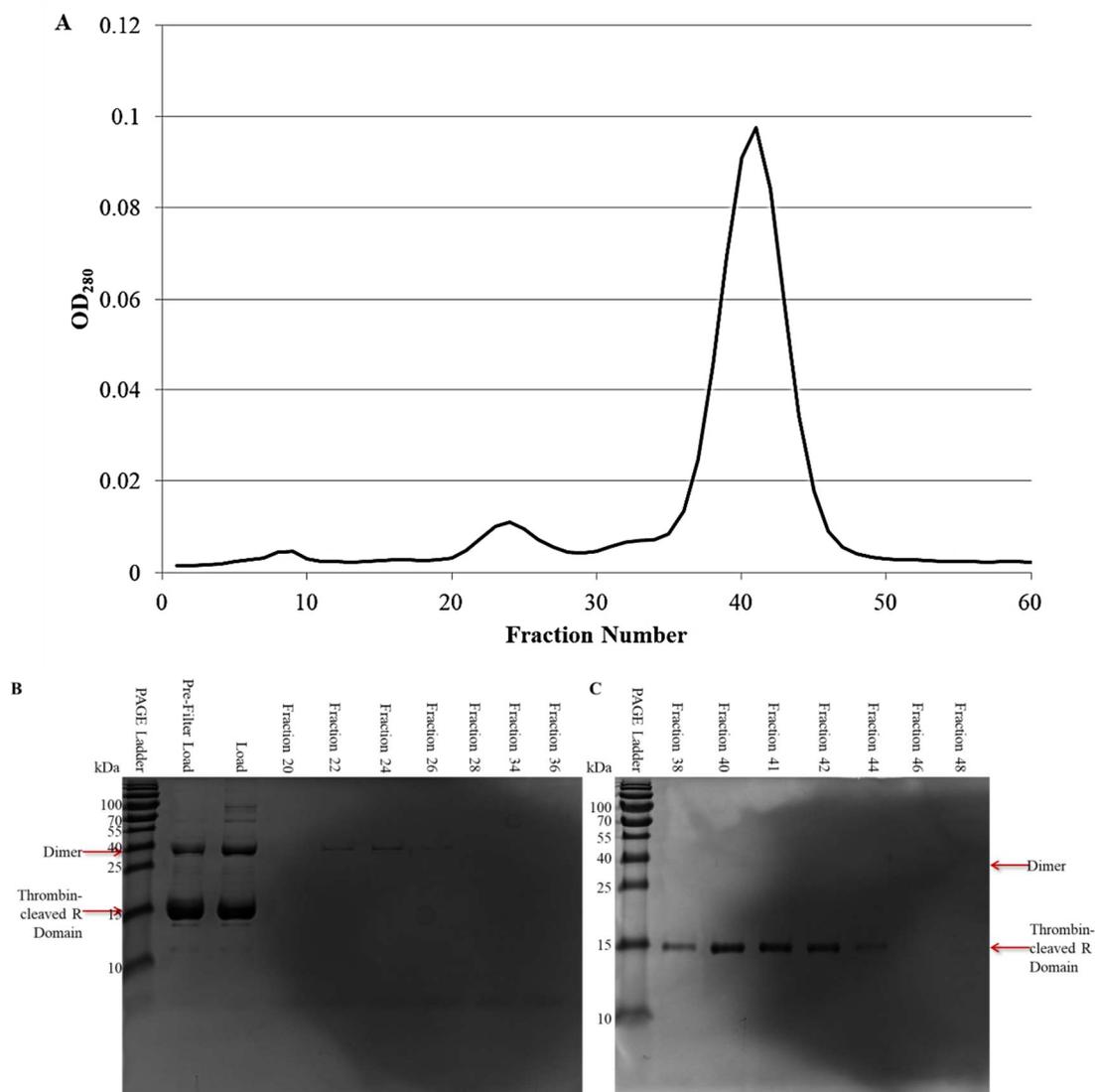


**Figure 4-7 – 15% (w/v) SDS gel of R Domain pre- and post-Thrombin Cleavage**  
 15% (w/v) SDS PAGE analysis of R-domain samples: one reacted with thrombin, the other unreacted to serve as a control. There is a noticeable shift of the thrombin-treated sample to a lower molecular weight, evidence of removal of the His<sub>6</sub> tag. This is consistent with the expected reduction in molecular weight of ~1227.27 Da following cleavage at the thrombin site.



**Figure 4-8 – ESI-MS mass spectrum of Thrombin-cleaved CoIE9 R-Domain**

ESI-MS mass spectrum of Thrombin-cleaved R-Domain run in positive mode. The analysis was done on 1  $\mu$ l of samples diluted 150-fold in a MS compatible buffer. The peak circled in green is at 16,048 Da, which is consistent with the expected peak of ~16,040.8 Da, plus 8.04 Da for the protonation of acidic and basic side chains. Peaks are calibrated using a myoglobin external standard.



**Figure 4-9 – Analysis of Fractions eluted from Size Exclusion column**

**A** - Optical density at 280 nm (OD<sub>280</sub>, solid line) during elution of thrombin and R-Domain from a Superdex75 size exclusion column. The first peak likely represents thrombin eluting from the column first, the second peak likely represents dimer eluting next, and the maximum peak likely represents cleaved R-Domain.

**B and C**- 15 % (w/v) SDS PAGE gel analysis of fractions: 20, 22, 24, 26, 28, 34, 36, 38, 40, 41, 42, 44, 46 and 48. Fractions were eluted from a Superdex75 size exclusion column. (B) Analysis of the second notable peak detected by OD<sub>280</sub>, bands at ~ 32,000 Da are consistent with dimer. (C) Analysis of the maximum OD<sub>280</sub> peak, bands at ~16,000 Da consistent with the expected molecular weight of cleaved R-Domain.

#### 4.2.4 Initial Attempts to Self-Couple R-domain using Bromomaleimide

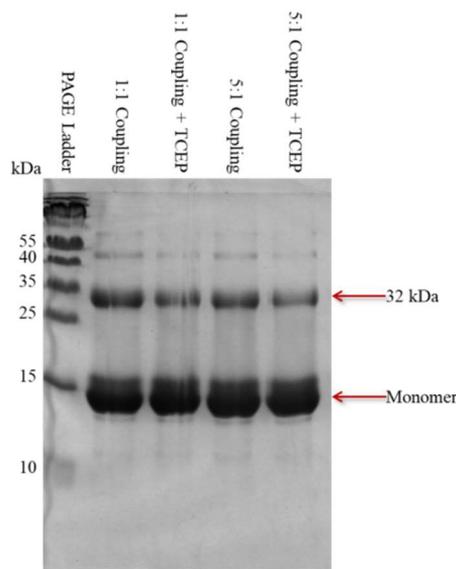
##### 4.2.4.1 Analysing Self-Coupling of R-Domain

Visualising N-MBM labelling of R-domain proved difficult by SDS-PAGE as expected. Determining the exact mass of R-domain from ESI-MS spectra in earlier

experiments was difficult due to the number of peaks produced (Figures 4-6 and 4-8). Positive-mode ESI-MS runs of N-MBM-labelled R-domain produced spectra that were even more difficult to read. No peaks were identified at ~16,150 Da, which would have provided proof of successful N-MBM labelling. For this reason, self-coupling of the R-domain was used to establish successful N-MBM labelling.

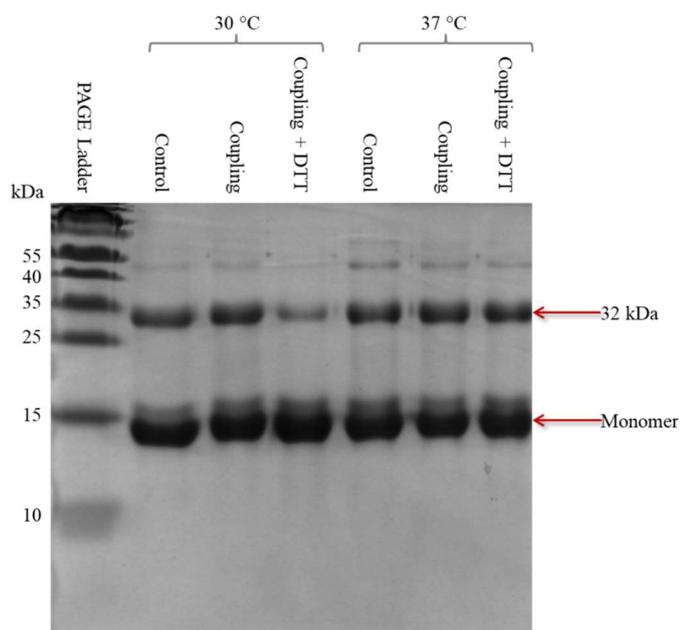
Self-coupling was first attempted at room temperature over an incubation time of 22 hrs. Reduced and N-MBM-labelled R-domain was reacted at 1:1 and 5:1 molar ratios; each condition had a separate run with 50 mM TCEP added. This initial coupling attempt was inconclusive as bands at ~32,000 Da appear to resemble coupled product (Figure 4-10). However, addition of TCEP reduces the intensity of this band suggesting that it could be disulfide-linked dimer. The similarity of molecular weight between dimer and the expected self-coupled product made confirming conjugation difficult.

To assay if the increase in intensity of the ~32,000 Da band was due to successful labelling, reactions were conducted at incubation temperatures of 30 °C and 37 °C with DTT instead of TCEP as the reducing agent (Figure 4-11). SDS-PAGE analysis revealed a band at ~32,000 Da which had the same intensity within the control and coupling lanes of both incubation conditions. The exception was the 30 °C coupling reaction with DTT, where the 32,000 Da band was less intense than both the control and its equivalent reaction without added DTT, and the negative control. Interestingly, this reduction in intensity of the 32,000 Da band was not present in the coupling reaction carried out at 37 °C with DTT. All SDS-PAGE gels of coupling reactions showed the monomeric band at ~16,000 Da was the most intense band. The struggle to determine potential conjugated product from dimer by SDS-PAGE analysis confirms the requirement for an alternative method of confirming labelling.



**Figure 4-10 – 15 % (w/v) SDS-PAGE Analysis of Labelled R-domain Self-Coupling at Room Temperature**

Self-coupling of N-MBM-labelled R-domain and reduced R-domain at room temperature analysed by 15 % (w/v) SDS-PAGE. The monomeric band is at ~16,000 Da, and the potential coupled product/dimeric band is at ~32 kDa. Additional faint higher molecular weight bands are present at ~16,000 Da intervals suggesting potential polymerisation.

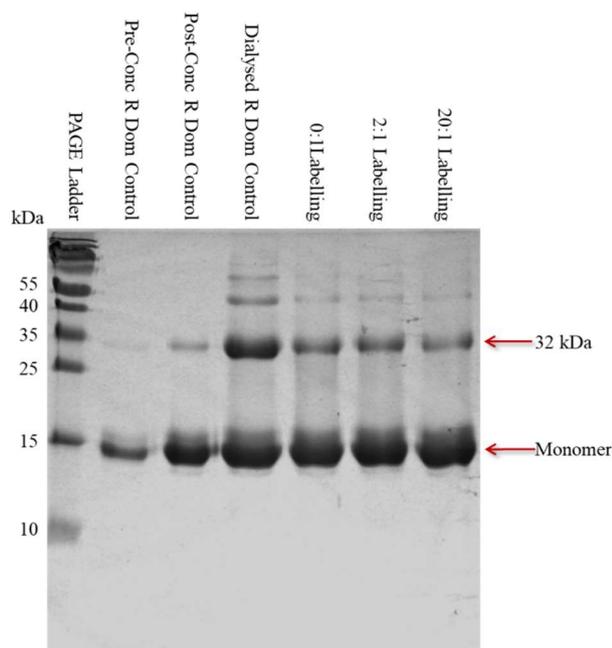


**Figure 4-11 – SDS-PAGE Analysis of Labelled R-domain Self-coupling at 30 °C and 37 °C**

Self-coupling of N-MBM-labelled R-domain and reduced R-domain at 30 °C and 37 °C analysed by 15 % (w/v) SDS-PAGE. The monomeric band is at ~16,000 Da, and the potential coupled product/dimeric band is at ~32 kDa. Additional faint higher molecular weight bands are present at ~16,000 Da intervals suggesting potential polymerisation, which are most prominent in the 37 °C reactions.

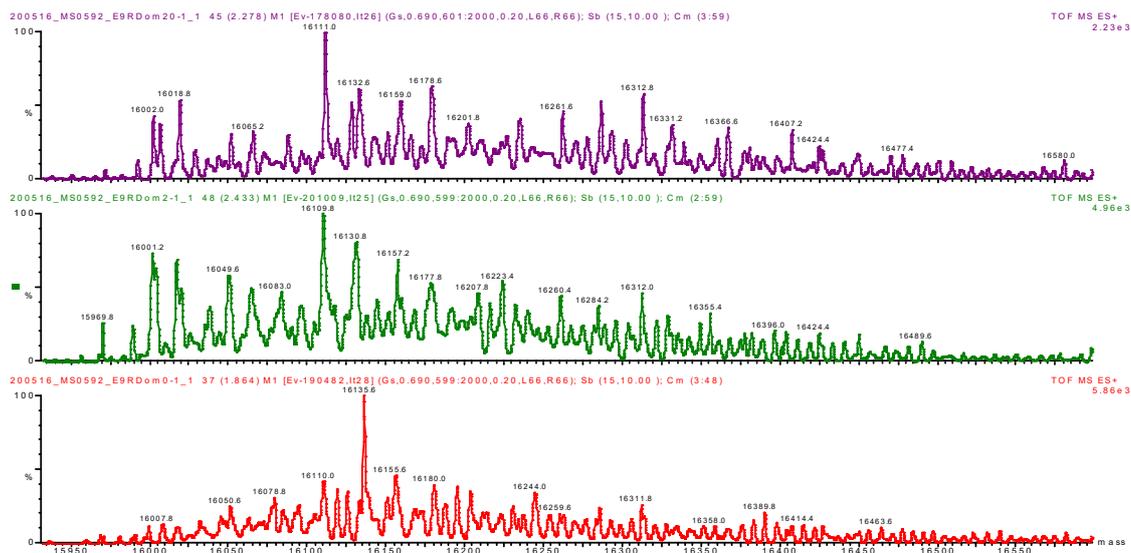
#### 4.2.4.2 Labelling R-Domain with N-MBM using different label:protein molar ratios

To review the best conditions for N-MBM labelling of R-domain, reactions were run at label:protein molar ratios of 0:1, 2:1 and 20:1. Temperature was kept at room temperature for all reactions. SDS-PAGE analysis of each condition revealed little difference between them, with the monomeric band at ~16 kDa not showing any mobility shift indicative of a change in molecular weight (Figure 4-12). Following dialysis into ammonium acetate buffer, each sample was analysed individually by positive-mode ESI-MS. All of the resulting spectra contained multiple peaks making them difficult to interpret (Figure 4-13). None of the spectra gave the expected peak of ~ 16,150 Da representative of N-MBM-labelled, cleaved R-domain. The most intense peak of the 0:1 control unexpectedly was of a greater mass than those of the 2:1 and 20:1 labelling reactions. The 2:1 and 20:1 molar equivalency samples possessed a very similar ESI-MS spectrum.



**Figure 4-12 – 15% (w/v) SDS-PAGE analysis of R-Domain Labelling with Bromomaleimide**

Attempts to label R-Domain with N-methyl-bromomaleimide at label:protein molar ratios of 0:1, 2:1 and 20:1 analysed by 15 % (w/v) SDS-PAGE. Bands at ~16 kDa represent monomer, bands at ~32 kDa represent dimer, and additional bands at increasing intervals of ~16 kDa represent further polymerisation. The first two controls consist of R-Domain before and after concentration by VIVASPIN 500 spin concentrating columns.



**Figure 4-13 –ESI-MS spectra of 0:1, 2:1 and 20:1 N-MBM labelling reactions**

This image compares the ESI-MS spectra of these reactions across the range 15950 - 16550 Da for the three samples (0:1 red, 2:1 green, 20:1 purple). ESI-MS was run in positive mode. 10  $\mu$ l of 20:1 and 2:1 sample were processed by zip-tip. 5  $\mu$ l of 0:1 sample was run by zip-tip. The most intense peaks for 20:1, 2:1 and 0:1 was at 16112.3 Da, 16111.1 Da and 16136.9 Da respectively. All peaks were corrected to a myoglobin external standard.

## 4.2.5 Testing N-MBM Labelling via Ellman's Test

### 4.2.5.1 Generation of a Standard Curve for Ellman's Reagent

To begin, a standard curve was generated using DTT as the test thiol (Figure 2-2). A single DTT molecule contains two thiol groups, therefore the measured [SH] concentration by Ellman's test equates to twice the [DTT]. As a result, the expected gradient of the standard curve [SH] against [DTT] was  $\sim 2$ . In reality, the actual gradient of the standard curve produced was  $\sim 1.67$ . This standard curve was used as a reference for future analysis of thiol-containing proteins. Throughout this testing, it was concluded that pH 7.0 was the condition most conducive for an accurate Ellman's test. As a result, labelled R-domain was transferred by membrane dialysis (Spectra/Por, 12,000-15,000 Da Molecular Weight Cut-Off (MWCO)) into pH 7.0 Sodium Phosphate buffer.

### 4.2.5.2 Ellman's Test of N-MBM-labelled Uncleaved R-Domain

Since the entire cleaved R-domain stock was in use to test successful labelling, uncleaved R-domain was used for Ellman's tests to save time and material. Uncleaved

R-domain, labelled at 20 °C and 30 °C with N-MBM, underwent an Ellman's test alongside unlabelled R-domain and another thiol-containing protein, Bovine Serum Albumin (BSA), for comparison. BSA contains one free thiol site for flagging by Ellman's reagent. The  $\Delta\epsilon_{412}$  for the labelled R-domain samples was unexpectedly greater than that of the unlabelled sample (Table 4-1). Surprisingly, the calculated [SH] was also greater than that of BSA. This was deduced to be due to the presence of TCEP, reducing the disulphide bridge within DTNB, and leading to the increase in  $A_{412}$ . Residual TCEP from the labelling protocol was removed from test samples by dialysis before later Ellman's tests.

**Table 4-1 – Ellman's Test Results of R-Domain-N-MBM labelling samples**

Table of results from Ellman's test of unlabelled/labelled R-Domain samples following reaction with/without N-MBM at 20 °C and 30 °C incubation temperatures. Bovine Serum Albumin (BSA, 150  $\mu\text{M}$ ) was used as a control. Initial protein concentrations were measured by Nanodrop<sup>TM</sup> spectrophotometer.  $\Delta\epsilon_{412}$  for the Ellman's test was measured in the same manner and used to calculate [SH] (See Equation 2-1). All numbers are given to 3 significant figures. TCEP present in R-Domain samples appears to have affected the results.

<b>Samples</b>	<b>Reduced R Domain</b>	<b>Labelled R Domain 20 °C Incubation</b>	<b>Labelled R Domain 30 °C Incubation</b>	<b>Bovine Serum Albumin (BSA)</b>
Nanodrop [Protein]	5.35	5.48	4.88	9.15
Expected [SH] in $\mu\text{M}$	5.35	0.00	0.00	9.15
Measured [SH] in $\mu\text{M}$	22.4	27.2	38.4	9.96

#### 4.2.5.3 Labelling of SasG with N-MBM and subsequent Ellman's Tests

To determine ideal conditions for labelling SasG (Cys-G51-link-G57-Cys), a protein containing two free thiol groups, was test labelled with N-MBM. Labelling reactions were conducted at 20 °C and 30 °C and, alongside equivalent unlabelled controls, were subject to Ellman's testing to determine [SH]. Small-scale reactions were conducted, with the resulting  $\Delta\epsilon_{412}$  measured by Nanodrop<sup>TM</sup> spectrophotometer. Ellman's tests determined [SH] within samples of SasG reduced at 20 °C to be greater than the equivalent sample at 30 °C (Table 4-2). N-MBM labelled SasG samples reacted at 20 °C and 30 °C possessed ~52 % and ~3 % free thiol as a proportion of their population.

Measured [SH] values of free thiol in SasG samples suggested that labelling had occurred at a relatively low efficiency, 30 °C providing the best incubation temperature for labelling. As a result, future labelling reactions using N-MBM were carried out at 30 °C and at pH 7.0.

**Table 4-2 – Ellman’s Test Results of SasG-N-MBM labelling samples**

Table of results from Ellman’s test of unlabelled/labelled SasG following reaction with/without N-MBM at 20 °C and 30 °C incubation temperatures. Initial protein concentrations were measured by Nanodrop™ spectrophotometer.  $\Delta\epsilon_{412}$  for the Ellman’s test was measured by nanodrop spectrophotometer and used to calculate [SH] (See Equation 2-1). All numbers are given to 3 significant figures.

<b>Samples</b>	<b>Reduced SasG 20 °C</b>	<b>Reduced SasG 30 °C</b>	<b>Labelled SasG 20 °C Incubation</b>	<b>Labelled SasG 30 °C Incubation</b>
Nanodrop [Protein]	14.5	37.4	135	75.8
Expected [SH] in $\mu\text{M}$	14.5	37.4	0.00	0.00
Measured [SH] in $\mu\text{M}$	14.5	24.4	71.1	2.05
% Free Thiol in Population	<b>100</b>	<b>65.2</b>	<b>52.7</b>	<b>2.73</b>

#### 4.2.5.4 Returning to N-MBM Labelling of Uncleaved R-domain

Based upon the results from the Ellman’s test of SasG, a fresh aliquot of uncleaved R-domain was labelled with N-MBM for 6 hrs at 30 °C in pH 7.0 buffer. Small aliquots of the reaction underwent an Ellman’s test, the resulting  $\Delta\epsilon_{412}$  measured by Nanodrop™ spectrophotometer. Unlabelled R-domain was measured to contain ~25.5 % free thiol, whilst labelled R-domain contained no detectable free thiol (Table 4-3). This suggests that under the new labelling conditions, R-domain has been successfully conjugated with N-MBM. However, whilst the reaction is fairly efficient in this example, it is limited by the lack of free thiol in the R-domain stock. Following these results, I then sought to confirm N-MBM labelling at the 318Cys residue using alternative ESI-MS methods.

**Table 4-3 – Ellman’s Test Results of R-Domain-N-MBM labelling samples**

Table of results from Ellman’s test of labelled/unlabelled R-Domain samples following reaction with/without N-MBM at 30 °C and pH 7.0. TCEP has been dialysed out of the buffer for this test. Bovine Serum Albumin (BSA, 150  $\mu$ M), containing one free thiol site, was used as a control. Initial protein concentrations were measured by Nanodrop™ spectrophotometer.  $\Delta\epsilon_{412}$  for the Ellman’s tests were also measured by a Nanodrop™ spectrophotometer and used to calculate [SH] (See Equation 2-1). All numbers are given to 3 significant figures.

<b>Samples</b>	<b>Reduced R-Domain</b>	<b>Labelled R-Domain</b>
Nanodrop [Protein]	118	118
Expected [SH] in $\mu$ M	118	0.00
Measured [SH] in $\mu$ M	30.0	0.00
% Free Thiol in Population	<b>25.5</b>	<b>0.00</b>

#### 4.2.6 Use of MALDI and LC-MS to confirm N-MBM labelling of R-domain

Following digestion into smaller fragments by Trypsin, both unlabelled and labelled samples of uncleaved R-domain were run by MALDI LC-MS. The detected fragments were NCBI nucleotide BLAST searched (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) against the R-domain sequence (Table 4-4). I sought to identify a fragment containing N-MBM successfully conjugated to 318Cys. Detected fragments of both samples covered 41 % of the total R-domain sequence. None of these sequence fragments contained the N-terminal 318Cys labelling site, giving no confirmation of successful N-MBM conjugation. A Mascot search of the detected fragments gave 60 % sequence coverage for unlabelled R-domain, and 57 % sequence coverage for labelled R-domain (Table 4-5). Despite the increased sequence coverage, the Mascot search again could not detect the N-terminal 318Cys labelling site. As such, I was unable to confirm successful N-MBM-labelling of R-domain from MALDI LC-MS analysis.

**Table 4-4 - NCBI BLAST Search Results of R-Domain samples from MALDI LC-MS**  
 NCBI Blast search results of detected fragments following MALDI LC-MS analysis of R-Domain samples. Two runs were undertaken, one of unlabelled R-Domain, the other of N-MBM –labelled R-Domain. The score, the assessed biological relevance of each finding, is stated for each sample. The % of the sequence the fragments cover, as well as where in the sequence they cover (highlighted in red) is also given. Neither run covered the labelling site, 2Cys. Sequences are listed from left to right, N-terminal to C-terminal respectively.

Samples	Score	Sequence Coverage in %	Visualised Sequence Coverage
Unlabelled Uncleaved R-Domain	544	41	CVEAAERN CERARA <b>ELNQANEDVARNQERQ</b> AK <b>AVQVYNSRKSELDAANKTLADAI</b> AEIKQF NR <b>FAHDPMAGGHRMWQMAGLKAQRAQT</b> DVNNKQAAFDA <b>AAKEKSDADAALSAAQER</b> RKQKENKEKDAKDKCDKESKRNLVPR <b>GSLE</b> <b>HHHHHH</b>
N-MBM Labelled Uncleaved R-Domain	389	41	CVEAAERN CERARA <b>ELNQANEDVARNQERQ</b> AK <b>AVQVYNSRKSELDAANKTLADAI</b> AEIKQF NR <b>FAHDPMAGGHRMWQMAGLKAQRAQT</b> DVNNKQAAFDA <b>AAKEKSDADAALSAAQER</b> RKQKENKEKDAKDKCDKESKRNLVPR <b>GSLE</b> <b>HHHHHH</b>

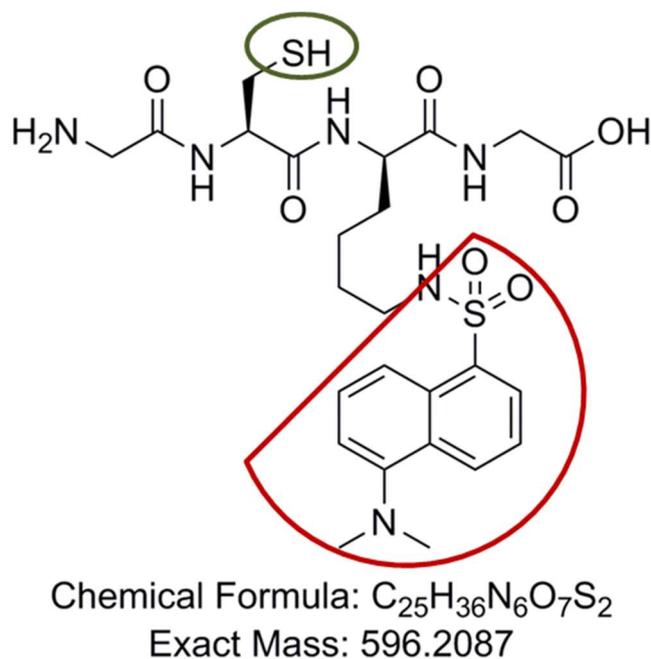
**Table 4-5 – Mascot Search Results of MALDI LC-MS run of R-Domain**

Mascot search results of detected fragments following MALDI LC-MS analysis of R-Domain samples. Two runs were undertaken, one of unlabelled R-Domain, the other of N-MBM –labelled R-Domain. The score, the assessed biological relevance of each finding, is stated for each sample. The % of the sequence the fragments cover, as well as where in the sequence they cover (highlighted in red) is also listed. Neither run covered the labelling site, 2Cys. Sequences are listed from left to right, N-terminal to C-terminal respectively.

Samples	Score	Sequence Coverage in %	Visualised Sequence Coverage
Unlabelled Uncleaved R-Domain	708	60	CVEAAERN CERARA <b>ELNQANEDVARNQERQ</b> AK <b>AVQVYNSRKSELDAANKTLADAI</b> AEIKQF FNR <b>FAHDPMAGGHRMWQMAGLKAQRAQT</b> DVNNKQAAFDA <b>AAKEKSDADAALSAAQER</b> RKQKENKEKDAKDKCDKESKRNLVPR <b>GSLE</b> <b>HHHHHH</b>
N-MBM Labelled Uncleaved R-Domain	667	53	CVEAAERN CERARA <b>ELNQANEDVARNQERQ</b> AK <b>AVQVYNSRKSELDAANKTLADAI</b> AEIKQF NR <b>FAHDPMAGGHRMWQMAGLKAQRAQT</b> DVNNKQAAFDA <b>AAKEKSDADAALSAAQER</b> RKQKENKEKDAKDKCDKESKRNLVPR <b>GSLEH</b> <b>HHHHH</b>

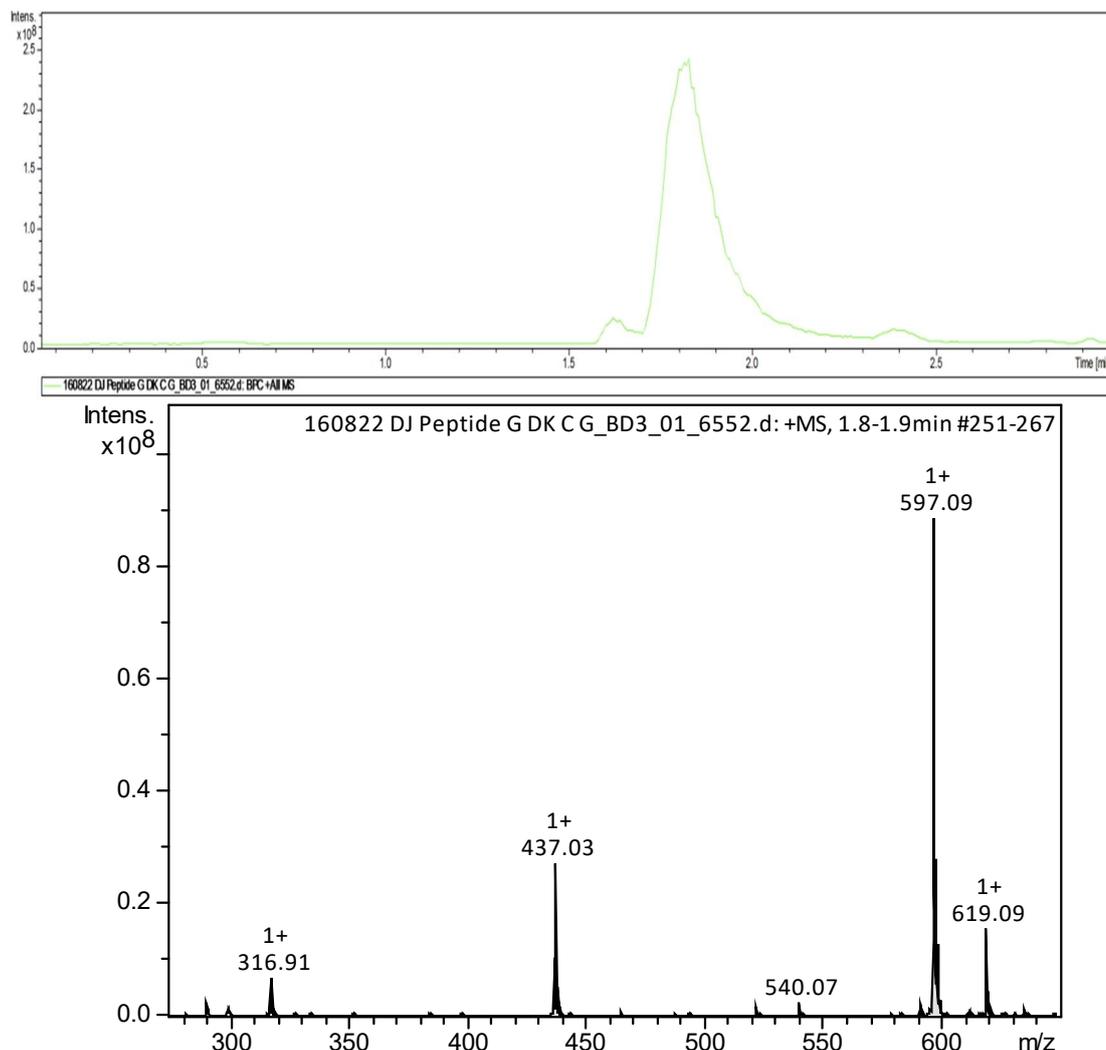
#### 4.2.7 Visualising Thiol Coupling of R-domain using Fluorescent Tetrapeptide

Previous attempts to visualise successful labelling and/or coupling by SDS-PAGE analysis were unsuccessful. The increase in mass of labelled R-domain was too small to observe between (+ ~110 g/mol), and the dimer band at ~ 32 kDa obscured any potential self-coupling product. To circumvent this, a fluorescent, thiol-containing tetrapeptide (G-C-DK-G) (Figure 4-14) was synthesised, using a standard solid-state polypeptide synthesis protocol (Kumarn, Chimnoi and Ruchirawat, 2013). Purity and successful synthesis of the peptide was confirmed by LC-MS, a peak at 597.09 Da matching the expected mass of 596.21 Da (Figure 4-15). The peptide was synthesised with a Dansyl-Lysine group to emit fluorescence. Using a UV transilluminator, peptide fluorescence was then visualised as bands on an SDS gel. Referencing these bands to the relative position of R-domain on a white image of the same gel would confirm successful thiol conjugation. This would validate the conjugation protocol at the 318Cys labelling site.



**Figure 4-14 – Chemical structure of synthesised fluorescent tetrapeptide**

The visualised: chemical structure, formula and molecular weight of G-C-DK-G fluorescent tetrapeptide. The fluorescent Dansyl group was used to visualise successful coupling reactions (red). The thiol group of the cysteine (green) was used to conjugate labelled/unlabelled R-domain.



**Figure 4-15 – LC-MS analysis of synthesised fluorescent tetrapeptide**

Liquid chromatography and mass spectrometry analysis of synthesised fluorescent tetrapeptide. Liquid chromatography measurements of intensity against time show one main peak, the peptide, and a small splinter peak. The smaller peak is likely some contaminants, but the very low intensity suggests it is insignificant. ESI-MS shows a clear mass peak at 597.09 g/mol, the expected molecular weight of the tetrapeptide. This confirms that the solid-state synthesis has been successful.

The first attempt to conjugate the custom-made peptide to R-domain proved successful. Both unlabelled and labelled samples, displayed a fluorescent band on an SDS gel at a molecular weight of ~17 kDa (Figure 4-16). A detectable fluorescence emission from a band at ~32 kDa was also present. The expected molecular weight shift (+ ~596 g/mol) appeared in reaction lanes of SDS gels viewed under white light (Figure 4-14). Addition of reducing agent (~1.67 mM DTT) led to a minimal reduction in intensity of the fluorescent band. Negative controls of unlabelled/labelled R-domain were completely absent of fluorescent bands in their lanes. Another negative control,

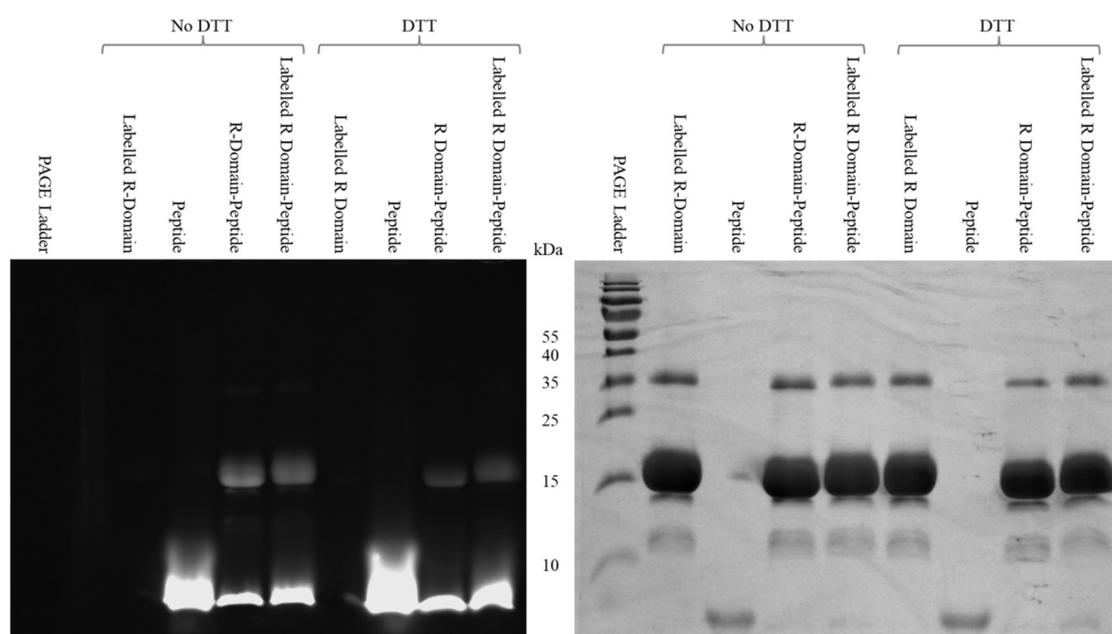
containing only peptide, showed fluorescence at the bottom of the gel, a completely different position to reaction lanes. The white image showed lower molecular weight bands indicative of potential degradation of the R-domain protein.

Treatment of peptide-coupled R-domain with a powerful reducing agent ( $\beta$ -Mercaptoethanol, 100:1 molar equivalent) led to a visible reduction in intensity of the fluorescent band (Figure 4-17). Increasing the incubation time from 1 hrs to 4 hrs did not lead to any further reduction in fluorescent band intensity (see Appendix). Further experiments reacted R-domain with Cysteine prior to peptide addition, to act as a blocking group at the 318Cys/N-MBM site. Resulting fluorescent bands were visualised by SDS-PAGE analysis (Figure 4-18). Lanes where Cysteine was present suffered a slight reduction in fluorescence emission compared to equivalent bands in standard peptide-R-domain reaction lanes.

To ensure visualised fluorescence was due to conjugation, not association, free peptide was removed by buffer exchange. Fluorescence visible at the bottom of previous transilluminated SDS gels was absent in these buffer exchanged samples (Figure 4-19). An exception was the lane representing unlabelled, Cys-treated reaction, which contained a significant fluorescent band. Post-buffer exchange, fluorescent bands at  $\sim 17$  kDa are still visible within peptide-R-domain reaction lanes. As previous, lanes without Cysteine present contained bands of a lower fluorescent intensity. Bands within lanes of unlabelled R-domain, as opposed to N-MBM-labelled R-domain, displayed greater fluorescence emission (Figure 4-19). Flow-through lanes showed a faint presence of free thiol, with the exception of the unlabelled, Cys-treated lane. The white image displayed a complete absence of free peptide bands in comparison to previous gels, confirming peptide removal (Figure 4-19). Comparing SDS gels pre- and post-buffer exchange confirmed complete maintenance of fluorescence emission throughout the peptide removal process.

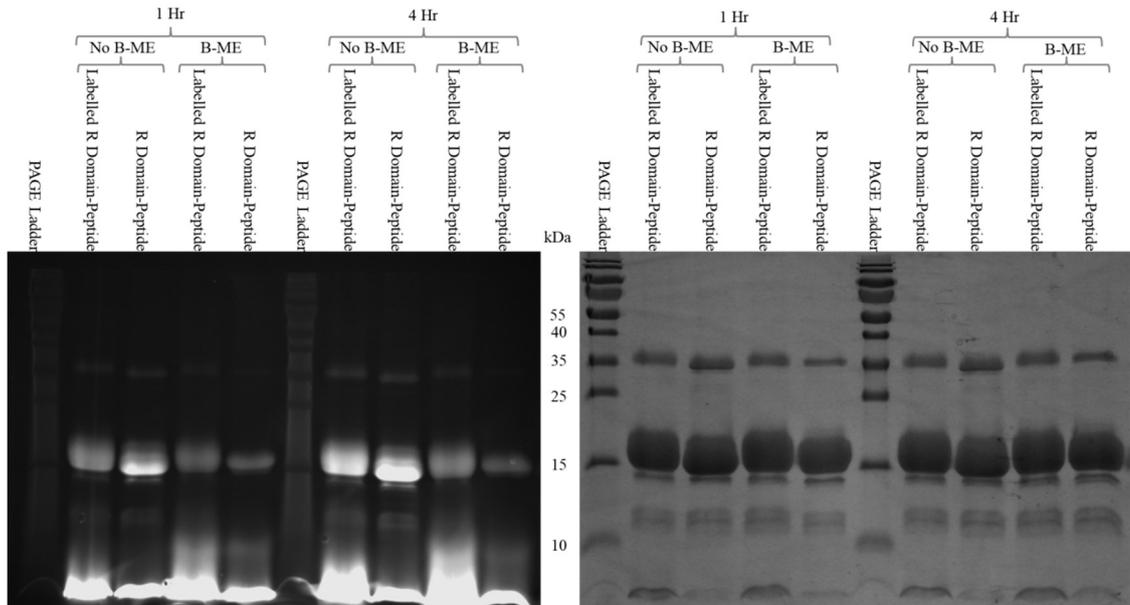
The impact on the intensity of fluorescent bands, following addition of reducing agent (100 m.e.  $\beta$ -ME, 30 °C, 1 hr incubation), was further tested free peptide removed samples.  $\beta$ -ME addition led to reduced intensity of fluorescent bands as previous (Figure 4-20). Free peptide removal allowed visualisation of increased free peptide in  $\beta$ -ME lanes. The reduction in fluorescence emission was most significant in samples with unlabelled R-domain. Bands at  $\sim 32$  kDa in N-MBM-labelled R-domain lanes

displayed lower fluorescence emission. These bands were completely absent in images of  $\beta$ -ME lanes. The potential degradation bands highlighted in previous SDS gels were absent in SDS gels post-buffer exchange. White images also hinted at a slight molecular shift of bands containing labelled R-domain compared to unlabelled R-domain (Figure 4-20).



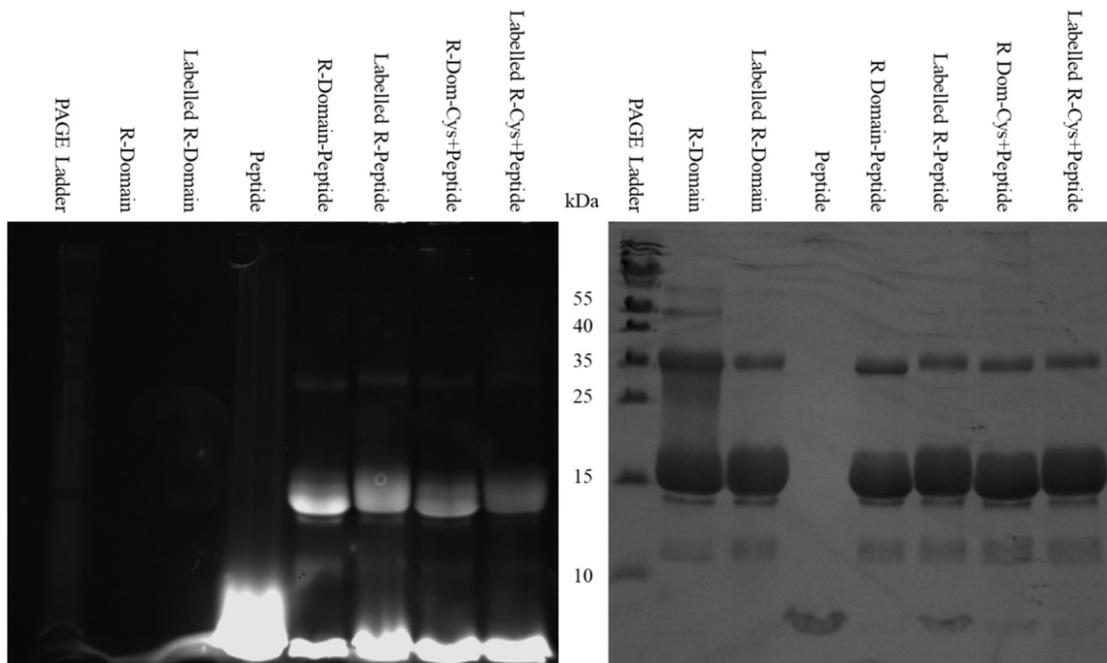
**Figure 4-16 – 15 % (w/v) SDS Gel of Peptide and R-Domain Coupling Reaction**

15% (w/v) SDS gel visualising coupling of fluorescent peptide to unlabelled and N-MBM-labelled R-Domain. The effect of DTT (~1.67 mM) on the reaction and peptide is also analysed. The image on the left is a UV transilluminated photograph revealing any fluorescent bands present. The fluorescent protein band is visible in coupling reaction lanes with and without DTT. The image on the right is a white image of the same gel.



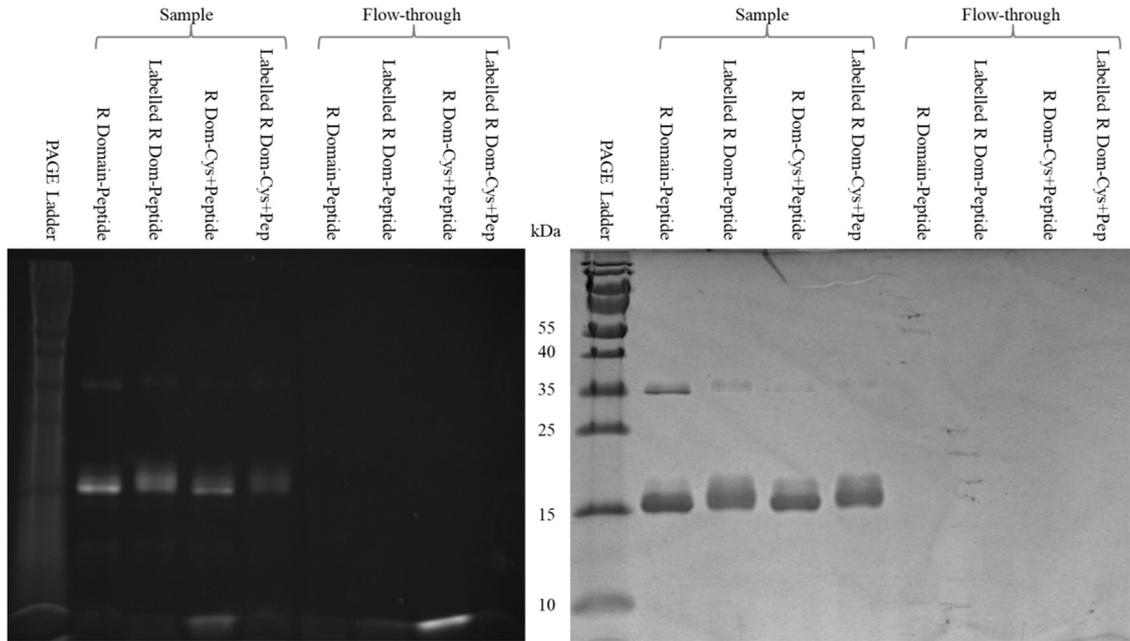
**Figure 4-17 – 15% (w/v) SDS Gel of Coupling Reactions treated with  $\beta$ -ME**

15% (w/v) SDS gel visualising coupling of fluorescent peptide to unlabelled and N-MBM-labelled R-Domain. The effect of  $\beta$ -ME (~20.8 mM) on the coupling reaction is analysed here at 1 hr and 4 hr incubation times. The image on the left is a transluminated image revealing any fluorescent bands present. The image on the right is a white image of the same gel.



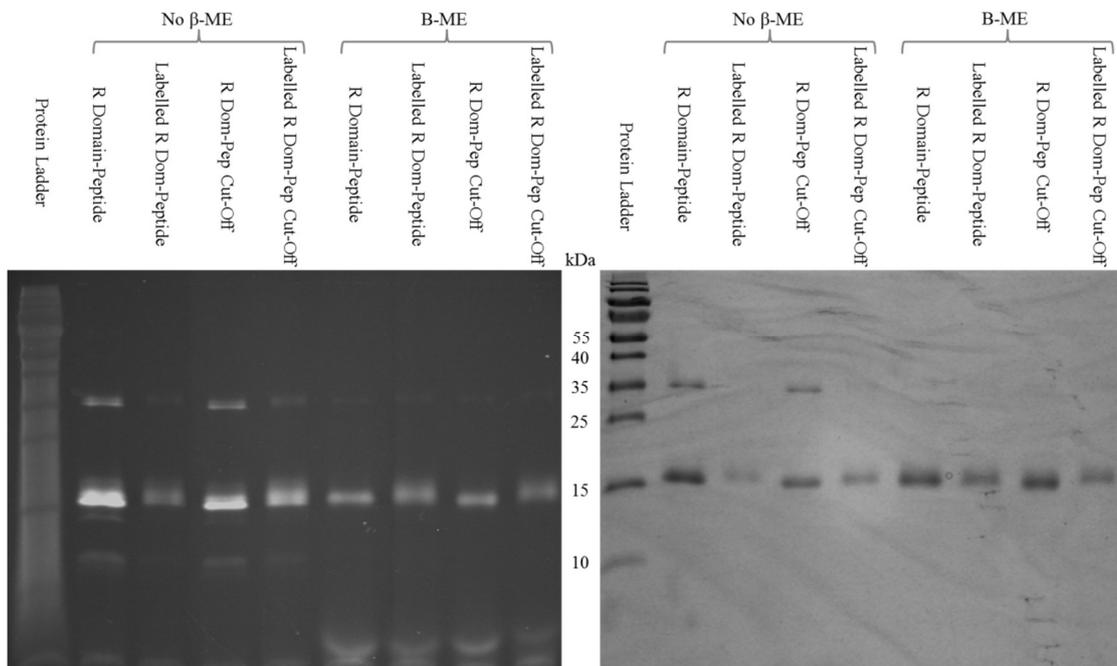
**Figure 4-18 – 15% (w/v) SDS Gel of Coupling Reactions with Cysteine Blocker**

15% (w/v) SDS gel visualising coupling of fluorescent peptide to unlabelled and N-MBM-labelled R-Domain. Selected R-Domain samples were reacted with Cysteine to act as a blocking group for the 2Cys/N-MBM coupling site. The image on the left is a transluminated image revealing any fluorescent bands present. The image on the right is a white image of the same gel.



**Figure 4-19 – 15% (w/v) SDS gel of Peptide Coupled Samples Post Buffer Exchange**

15% (w/v) SDS gel visualising R-Domain-peptide coupled samples post buffer exchange. Samples were buffer exchanged using VIVASPIN 500 spin concentrators. Buffer exchanged and flow-through samples for each condition are labelled. The image on the left is a transluminated image revealing any fluorescent bands present. The image on the right is a white image of the same gel.



**Figure 4-20 15 % (w/v) SDS gel of R-domain-Peptide Post-Cut-off with  $\beta$ -ME**

15 % (w/v) SDS gel of unlabelled/labelled R-domain-Peptide coupled samples. Selected samples have been reduced with  $\beta$ -Mercaptoethanol (~100 molar equivalent) as indicated. The image on the left is a transluminated image revealing any fluorescent bands present. Notably the transluminated image was taken at a 4 sec exposure. The image on the right is a white image of the same gel.

Liquid Chromatography Mass Spectrometry (LC-MS) of peptide-coupled samples, conducted by Richard Spears (Fascione Lab, Department of Chemistry, University of York), were met with mixed success. Unlabelled R-domain reacted with peptide produced a different spectrum, with a peak at 17808 Da indicating a covalent modification (Figure 4-21). All LC-MS spectra of N-MBM-labelled R-domain samples produced too much noise to identify any presence of protein material. Attempts to run samples of N-MBM-labelled R-domain with Cysteine blocker were also unsuccessful. The labelling process clearly negatively affected R-domain's already tentative ability to fly by Mass Spectrometry.

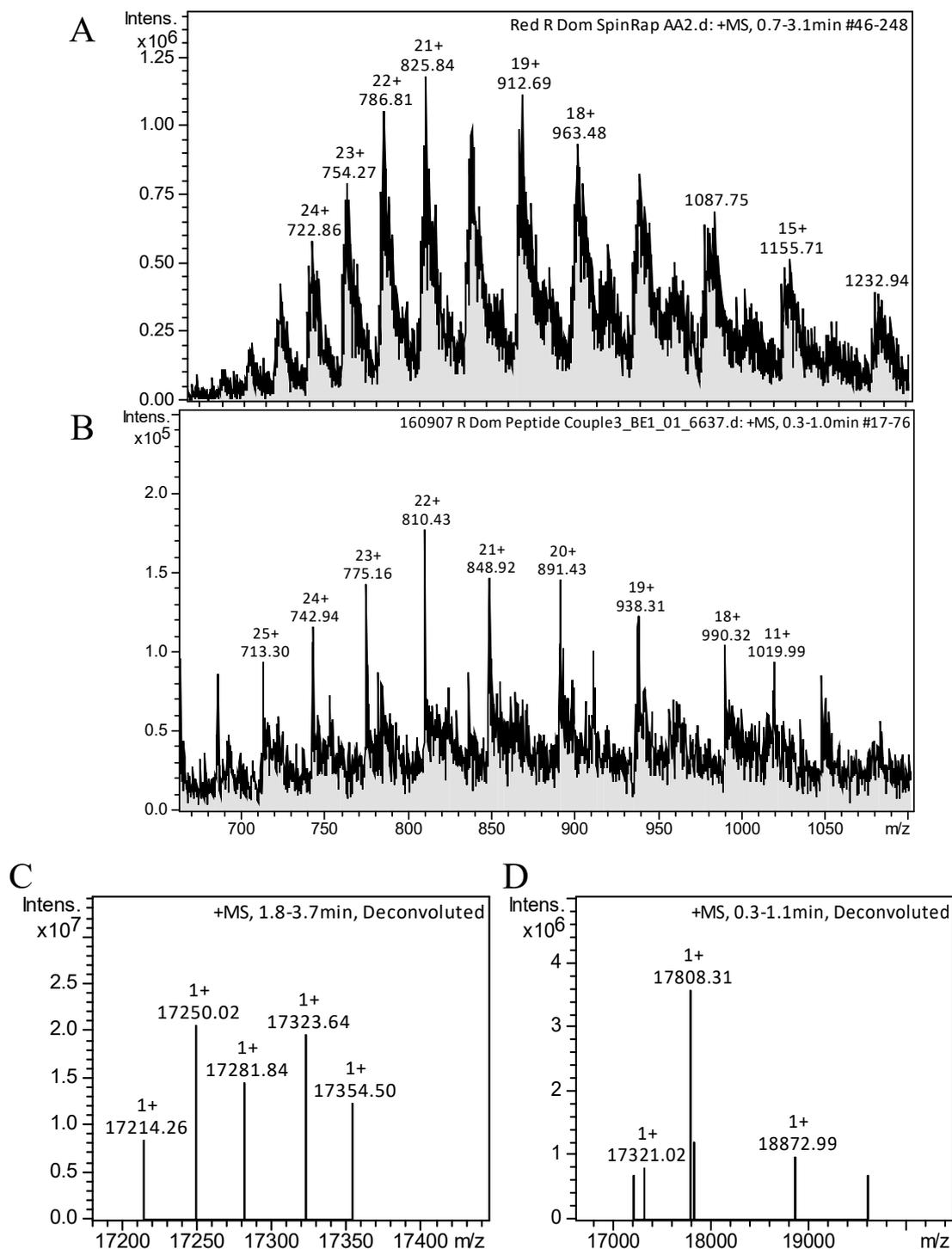
#### **4.2.8 Thiol Coupling R-domain to Oligo Components of the MTP**

Following peptide analysis of N-MBM labelling and thiol coupling, I proceeded to attempt coupling of R-domain to its respective oligonucleotide component. Aliquots of unlabelled and labelled R-domain stocks were reacted with E9 T Coupling 3' Thiol Oligo at a 4:1 protein to oligo ratio. Two incubation time lengths were tested, 3 hrs and 24 hrs. SDS-PAGE analysis displayed protein bands at ~17 kDa and ~32 kDa, similar to other experiments (Figure 4-22). Unlabelled R-domain incubated with oligo for 24 hrs produced an additional faint band within its lane at ~28 kDa. The band at ~32 kDa, commonly associated with dimer, was noticeably thicker in samples that underwent 24 hr incubation (Figure 4-22).

Initial attempts to identify oligo-coupled bands, using DNase I as a digestion tool, were unsuccessful due to unexpected protease activity. Addition of EDTA to counteract this did not fully inhibit protease activity. This resulted in visible degradation bands on SDS gels (Figure 4-4). A control exposing unlabelled/labelled R-domain to DNase I generated no observable difference between them and oligo reaction samples (Figure 4-23). The same band observed at ~28 kDa in the previous gel was visible in DNase I treated samples, suggesting it to be completely proteinaceous. No discernible difference could be seen between N-MBM-labelled R-domain lanes that had been reacted with oligo and those that had not.

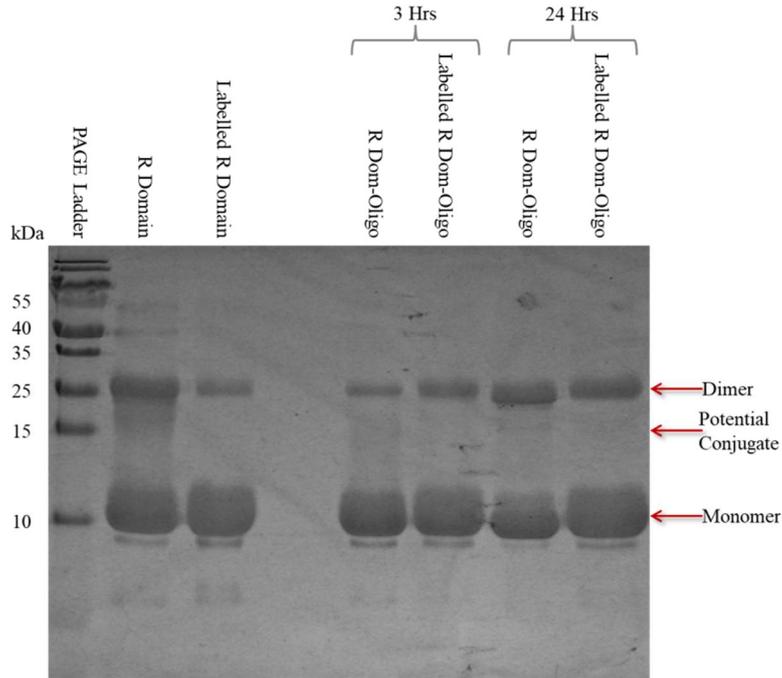
The second attempt to couple R-domain to oligo, using E9 R Coupling 5' Thiol Oligo instead, met with similar results (Figure 2-24). There were no obvious signs of a shift

indicative of oligonucleotide coupling. Lanes with DNase I present showed little difference, other than removal of the ~32 kDa band, and slight degradation of the monomeric band. Attempts to identify a present protein-oligo conjugate by LC-MS were unsuccessful. It was impossible to determine peaks belonging to protein or nucleotide material from spectra produced by R-domain-oligo reaction samples. Visualising R-domain-oligo samples on 12 % (w/v) native gels gave no clear results; bands struggled to translate through the gel. Time constraints prevented any further coupling reactions and analysis of existing samples. To conclude, I could not obtain conclusive evidence of successful R-domain-oligo conjugation from these methods.



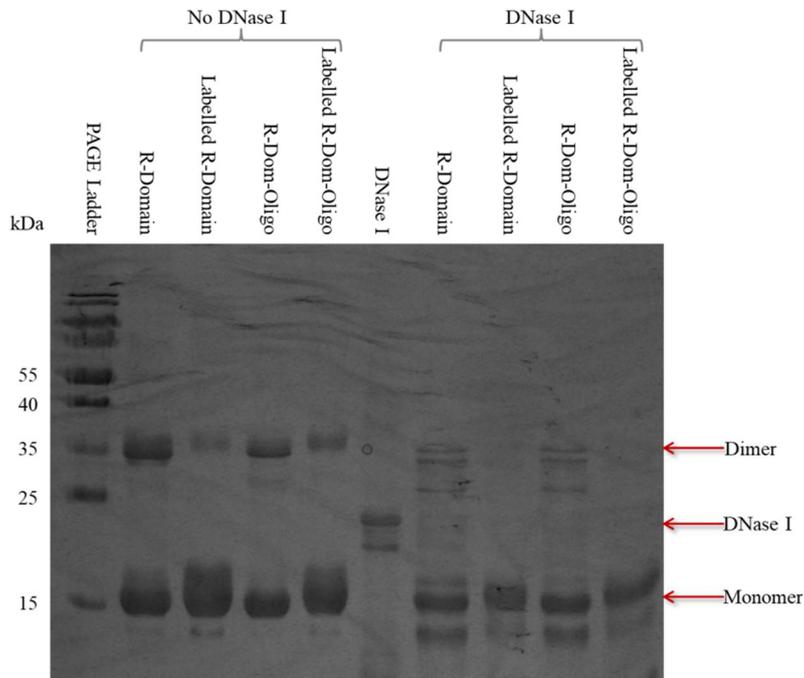
#### Figure 4-21- ESI-MS of Uncleaved R-Domain Peptide Coupling Samples

Positive ESI-MS spectra showing Intensity vs Charge of Uncleaved R-Domain samples, including those coupled to fluorescent tetrapeptide G-C-DK-G. **A**- spectrum is of R-Domain that has not been coupled to peptide, with an expected mass of ~17250 Da. **B**-spectrum visualises the results of R-Domain labelled with peptide. **C**- Detailed view of Intensity vs m/z for key peaks of Reduced R-domain sample. **D**- Detailed view of Intensity vs m/z for key peaks of R-domain-peptide reaction sample. Successful conjugation of peptide to R-Domain would result in an expected mass increase of ~600 Da.



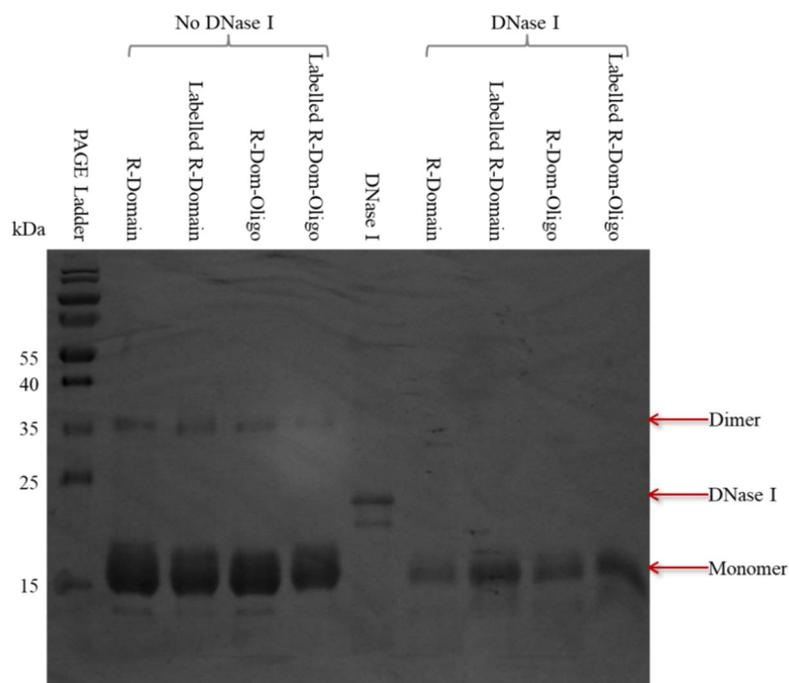
**Figure 4-22– 15% (w/v) SDS gel of R-Domain 3’Oligo ssDNA Coupling**

15% (w/v) SDS gel of unlabelled/labelled R-Domain samples reacted with E9 T Coupling 3’ Thiol ssDNA Oligo. Coupling reactions were run for two different incubation times: 3 hrs and 24 hrs. Bands at ~17 kDa, ~32 kDa, and ~28 kDa represent monomeric, dimeric and potential product respectively. The potential product band is barely visible in the unlabelled R-Domain-Oligo 24 hr lane.



**Figure 4-23 – 15 % (w/v) SDS gel of R-Domain-Oligo mix Reacted with DNase I**

15% (w/v) SDS gel of unlabelled/labelled R-Domain reacted with E9 T Coupling 3’ Thiol ssDNA Oligo. Certain labelled samples were reacted with DNase I at 37 °C for 1 hr and then quenched with EDTA. Protease activity is evident in the DNase I treated samples in spite of EDTA quenching.



**Figure 4-24- 15 % (w/v) SDS gel of R-Domain and 5'Oligo ssDNA Coupling**  
 15% (w/v) gel visualising the coupling reaction of R-Domain and E9 R Coupling 5' Thiol. Certain labelled samples were reacted with DNase I at 37 °C for 1 hr, and then quenched with EDTA. Some protease activity is evident in lanes where DNase I was present.

## 4.3 Discussion

### 4.3.1 Initial Attempts of N-MBM-Labeling and Self-Coupling of R-domain

R-domain has been shown here to be: readily producible, in a soluble state, at an acceptable yield, and on demand. Successful characterisation of R-domain confirms its suitability as a building block for the MTP. This established R-domain purification protocol proved useful in providing material towards thiol conjugation experiments. With the protein confirmed as viable, I moved to develop a protocol to effectively synthesise the MTP. This was to be achieved through thiol coupling of R-domain to its respective ssDNA oligonucleotide.

Whilst purification of R-domain was a smooth process, confirming N-MBM-labelling of the protein was less so. Positive ESI-MS runs of N-MBM-reacted R-domain samples were unreliable, often producing ESI-MS spectra with multiple peaks, or no discernible peaks at all. The spectra displayed multiple forms of ionised R-domain. These could be adducts of solution ions such as: phosphate (+ 98 Da), acetate (+60 Da), or calcium (+ 45 Da). One such example was the peak at 17348.1 Da from ESI-

MS spectra of R-domain (Figure 4-6), a mass difference of + 98.9 Da from the accepted R-domain peak of 17249.2 Da. This peak appears to have resulted from the association of a residual phosphate ion. This likely originated from storage buffer, and failed to be removed during dialysis. Despite this example, many of these additional peaks on the ESI-MS spectra cannot be explained by ion association. The additional peaks may have arisen from different amino acid compositions in the R-domain sample, due to degradation or modification. Several SDS gels did show evidence of degradation, supporting this theory (Figures 4-16, 4-17, 4-18). However, speculative deletions of R-domain residues from the N-terminus and/or C-terminus in the web tool ExPASy ProtParam (<http://web.expasy.org/protparam/>) failed to generate mass values that matched spectra peaks.

Another concern was that further modifications to the R-domain, such as N-MBM labelling, or thrombin cleavage, further reduced its ability to fly by positive ESI-MS. This resulted in either: the generation of even more peaks on the spectra, or a complete absence of peaks altogether. Confirmation of N-MBM-labelling was impossible, in part due to the inability to identify expected mass peak of N-MBM-R-domain (16150 Da) by positive ESI-MS (Figure 4-8). Adjusting the label-to-protein molar equivalency ratio did not result in the expected increase in molecular weight of ~110 g/mol as hoped (Figure 4-13). The unreliable nature of performing ESI-MS with R-domain became a significant obstacle to the work. It required bypassing with alternative methodology to confirm R-domain modifications.

As predicted, the relatively small molecular weight shift (+110 g/mol) made confirmation by SDS-PAGE analysis next to impossible. R-domain protein appears on an SDS gel as a fairly broad band in its monomeric form. However, a much thinner band at ~32 kDa is also present, showing the tendency of a small population to form a dimer. This dimer likely occurs due to natural formation of a disulphide bond between the thiol groups (318Cys) of two R-domain molecules. Unless removed by reducing agent, the dimer band was always present, with oxidation occurring over a time span less than 24 hrs. The dimer band itself became a massive hindrance to self-coupling experiments. Due to the similar expected molecular weight, any potential self-coupled product was completely obscured by the dimer band. To remove the dimer band, reducing agent was introduced (Figure 4-17). Following these results, R-domain stocks

were reduced by TCEP.HCl, and subsequently stored at -20 °C. This was introduced to reduce the impact of reoxidation on the availability of free thiol for labelling.

One positive result was a visible depletion in the monomeric band, accompanied by an increase in intensity of the ~32 kDa band, of self-coupled R-domain reacted at 30 °C (Figure 4-10). The addition of TCEP to the same reaction conditions largely reversed this, leading to confusion about whether or not the change is due to self-coupling, or increased dimer formation due to the temperature. Despite this concern, future reactions were carried out at 30 °C compared to the initial temperature of 0 °C stated in the literature. Reactions at 37 °C were deemed too unstable for the protein to maintain for long incubation periods. A standard incubation time of 18 hrs was also incorporated into the N-MBM labelling protocol to raise the chances of successful conjugation. Despite these adjustments to the protocol, dimer interference prevented confirmation of N-MBM labelling by SDS-PAGE analysis. Throughout SDS-PAGE analysis of self-coupling, monomeric R-domain remained the most intense band on the gel. This suggested that the thiol conjugation protocol had not produced the desired product efficiently. Further confirmation was required to ascertain if the maleimide conjugation protocol was a success.

#### **4.3.2 Ellman's Tests and MALDI LC-MS to confirm N-MBM Labelling**

I then pursued alternative methods to confirm the viability of the coupling protocol. Ellman's tests were carried out to confirm labelling at the 318Cys thiol site of R-domain. Initial tests failed to detect any signs of N-MBM labelling (Table 4-1) via reduction in [SH]. Upon adjustment of the background buffer to pH 7.0, a value closer to the optimal pH for Ellman's reagent, results became clearer to understand. The lowered pH value also had the potential to improve the efficiency of thiol coupling reactions (Bednar, 1990). Measurement of  $\Delta\epsilon_{412}$  by Nanodrop<sup>TM</sup> spectrophotometer appeared to improve accuracy further.

Following these adjustments, Ellman's tests suggested that N-MBM had labelled to free thiol to an extent in both SasG and R-domain (Tables 4-2 and 4-3). Tests of unlabelled R-domain samples revealed that the free thiol labelling site (318Cys) was not present in the majority of the R-domain population (Table 4-3). Lack of 318Cys labelling sites within the R-domain population would negatively impact labelling

efficiency. This may have been due to degradation at the N-terminus. Some evidence of degradation has been noted during SDS-PAGE analysis of labelling experiments, several months after initial purification of the R-domain stock (Figures 4-16 to 4-19). This could also explain the multiple ionisation peaks within MS spectra (ESI and LC) of R-domain, but not why the labelling protocol increases the issue further.

Alternatives to the standard ESI-MS protocol were pursued in an attempt to confirm N-MBM labelling. MALDI LC-MS analysis of R-domain Trypsin-digested fragments also failed to identify successful labelling. When searched by BLAST and MASCOT, it became clear the detected fragments did not cover the entirety of R-domain. None of these sequence fragments contained the N-terminus, where 318Cys resides, rendering MALDI unsuccessful at confirming N-MBM labelling of R-domain (Tables 4-4 and 4-5). LC-MS runs by Richard Spears succeeded in visualising unlabelled R-domain, but were completely unable to detect any protein material within runs of labelled samples. N-MBM-labelled R-domain's ability to fly during mass spectrometry remained unreliable despite the range of techniques used. These results highlighted a necessity for a technique that would confirm thiol conjugation to R domain without direct need for ESI-MS confirmation.

### **4.3.3 Fluorescent Peptide Coupling of R-domain**

This technique was established through solid-state synthesis of the fluorescent tetrapeptide, G-C-DK-G. The Dansyl-Lysine residue allowed conjugation of thiol groups between peptide and R-domain to be visualised as fluorescent bands on an SDS gel. Optimal reaction conditions elucidated from previous experiments (pH 7.0 and 30 °C) were incorporated into the peptide coupling protocol. Peptide was synthesised without issue, and characterised successfully by LC-MS (Figure 4-15). Initial attempts garnered positive results; fluorescent peptide was shown to effectively couple both unlabelled and labelled R-domain (Figure 4-16). Further analysis proved that this thiol conjugation was resistant to low concentrations of reducing agent (Figure 4-17). An attempt to use Cysteine as a blocker to isolate peptide coupling to one site was unsuccessful (Figure 4-18). This does not necessarily mean the peptide isn't bound at 318Cys, perhaps by being a smaller molecule; Cysteine struggles to find an energetically favourable alignment to couple where the peptide does not.

Buffer-exchanged samples of R-domain-peptide conjugation retained their fluorescent bands upon re-examination, evidence that the two components have conjugated, not just merely associated (Figure 4-19). The intensity of these bands only decreased due to the presence of a reducing agent in an extreme excess, a further indication that a disulphide bond forms the linkage (Figure 4-20). Whilst N-MBM-labelled samples failed, unlabelled R-domain-peptide coupling samples ran on LC-MS produced peaks with a mass increase indicative of peptide conjugation (Figure 4-21). Such a result can only come via formation of a covalent bond between the two; a weakly associated peptide complex would never fly by LC-MS. From these results, it was clear that conjugation at the 318Cys thiol site was possible at the conditions explored.

#### **4.3.4 Coupling of ssDNA Oligonucleotide Handles to R-domain**

With a degree of confidence established in the ability of R-domain to conjugate to thiol-containing molecules, coupling of R-domain to ssDNA oligonucleotide handles was attempted. SDS-PAGE analysis turned up no confirmation of successful conjugation, DNase I digestion was unable to identify mobility shifts in bands as a result of digestion of nucleotide material. The protease contamination within the DNase I stock severely impacted the appearance of R-domain on the gel (Figure 4-23). This made it incredibly difficult to identify potential protein-oligo product bands.

LC-MS runs of these samples were unable to detect any peaks related to protein-oligo material at all, even for unlabelled R-domain samples. The few attempts at visualising the reaction by native gel electrophoresis were unsuccessful due to failure of bands to translocate sufficiently through the gel. Had time been sufficient, a lower % (w/v) Native gel could have provided a clearer picture on the success of the reaction. In addition, a fresh stock of purified DNase I could have been acquired to assess for potential R-domain-oligo product by SDS-PAGE analysis. As it is, further work needs to be done to clarify the thiol coupling protocol for synthesis of the MTP.

#### **4.3.5 Future Work**

In summary, R-domain has shown promise as a component of the MTP in its current state, but requires further testing to prove the viability of the thiol coupling protocol. There are concerns about the stability of R-domain, especially at the N-terminus, the location of the labelling site. These concerns are alleviated in part by the positive

results of fluorescent peptide coupling. Absence of evidence from sample analysis by positive ESI-MS has made confirmation of N-MBM labelling, and the thiol conjugation protocol, difficult. As a result, the validity of the N-MBM coupling protocol is still in question. The results produced from Ellman's tests lacked consistency, suggesting it is an unreliable method of confirming N-MBM labelling. Adjustments made to the Ellman's test protocol during the course of the project, such as the change in background buffer pH, may have influenced this. To counteract this, further Ellman's tests of labelled samples should be run using the most recent methodology. Until these tests provide consistent results, they should only be used in support of peptide and ESI-MS analysis.

Future investigations into R-domain-oligo coupling should assess both unlabelled and labelled R-domain samples to determine if N-MBM has a positive effect on conjugating thiol-containing molecules. Further assessment by Native gel electrophoresis is required to confirm any R-domain-oligo product present in current samples. From there, the optimum molar equivalency can be established for the protein-to-oligo coupling protocol to ensure maximum coupling efficiency. This approach can be further supported by Ellman's tests and fluorescent peptide experiments. Results here have determined several reaction conditions for thiol coupling, providing a platform to move towards an established MTP synthesis protocol.

## **Chapter 5: Investigating Protein-Protein, Protein-Lipopolysaccharide and Protein-Detergent Interactions in Crystal Structures of Bacterial Outer Membrane Proteins**

### **5.1 Research Aims**

Molecular dynamics (MD) simulations of BtuB-OmpF and BtuB-BtuB interactions in model membranes (PE:PG 3:1; protein volume occupancy  $\approx$  20-30 %) have investigated the presence of promiscuous protein-protein interactions (PPIs) (Kleanthous, Rassam and Baumann, 2015; Ortiz-Suarez *et al.*, 2016). These studies have hypothesised the presence of regions of preferred sites for PPIs on the membrane-facing  $\beta$ -barrel surface of such proteins. PPIs are thought to influence the spatio-temporal organisation of OM proteins, and drive several vital OM processes (Kleanthous, Rassam and Baumann, 2015). This chapter describes work done to identify such regions within OMPs, like BtuB or OmpF, and phylogenetically related mitochondrial Voltage-Dependent Anion Channels (VDACs).

If membrane-facing interfaces exist in  $\beta$ -barrel proteins where PPIs are more likely to occur, it would suggest that these interfaces or patches might be a conserved structural feature evolved in an early common ancestor. Readily available crystal structures within the protein database (PDB) on the European Bioinformatics Institute website (EBI, <http://www.ebi.ac.uk/>) were marked for analysis. Symmetry analysis was used to identify crystal contacts in the plane of the membrane bilayer for these structures. PPI patches would be assumed to be coincident with sites for these crystal contacts. To identify any characteristic features of these PPI patches, their amino acid composition and hydrophobicity was analysed relative to the entire membrane-facing  $\beta$ -barrel surface.

### **5.2 Introduction**

Cellular membranes are highly dynamic environments in a constant state of flux. The components of the membrane, namely phospholipids and membrane proteins, are constantly moving by lateral diffusion and can form part of larger lipid rafts and membrane domains. These lipid rafts, a form of higher order organisation within the OM (LaRocca *et al.*, 2013; Bramkamp and Lopez, 2015), are integral to certain

cellular processes (López and Kolter, 2010). For example, evidence suggests that lipid rafts crucial for the organisation of signal transduction processes (Simons and Toomre, 2000). Membrane proteins are highly prevalent in this environment, making up ca. 50% of the membrane by mass and 25% of the surface area (Dupuy and Engelman, 2008). The crowded environment in the membrane leads to a tendency for membrane proteins to form transient interactions with one another. Whether or not membrane proteins themselves have evolved structurally to form these interactions is a topic of interest.

As crystallization techniques have improved, greater numbers of membrane protein crystal structures have become available in established databases (Chabanon, Mangin and Charcosset, 2016). The European Bioinformatics Institute (EMBL, <https://www.ebi.ac.uk/pdbe/node/1>) has a publicly available database of structures, including their pdb files. As well as providing information on these proteins, this database provides an opportunity to structurally align proteins in PDBeFOLD. Using this tool, and other programs such as WinCOOT (Emsley *et al.*, 2010) and PyMOL (version 1.8, Schrödinger), provided the opportunity to study the structural features of membrane proteins in 3D. Symmetry analysis can provide an insight into the intermolecular interactions that occur in the plane of the membrane.

Current and past research into the interactions between membrane components has focussed on MD simulations. MD simulations can be used to study membrane dynamics *in silico* by simulating the lateral diffusion of lipids and membrane proteins to gain insight into their behaviour. These simulations typically utilise an established computational model, such as MARTINI, to simulate the force fields involved (Periole *et al.*, 2007). Coarse-grained MD simulations have indicated that membrane proteins slow the lateral diffusion of the lipids close to them (Koldso and Sansom, 2015). In such simulations, LPS has been shown to influence OMP-OMP interactions (Patel, Qi and Im, 2017). Some proteins exhibit tight interactions with lipid molecules that facilitate PPIs between them, such as the dimerization of NanC (Dunton *et al.*, 2014). Lipids have been shown to be essential to the function of several membrane proteins; for example, signalling by the human epidermal growth factor receptor (EGFR) (Bocharov *et al.*, 2016). Specific interaction sites between membrane lipids and some OMPs have been identified, e.g. OmpF (Arunmanee *et al.*, 2016).

Alpha-helical transmembrane proteins, like glycophorin A and bacteriorhodopsin, have had the protein-protein interactions that occur in the membrane environment extensively characterised (Doura and Fleming, 2004; Periolo *et al.*, 2007). These interactions are direct dimer interactions, rather than non-specific interactions mediated by membrane lipids.  $\beta$ -barrel membrane proteins tend to have less obvious sites of interactions, with the obvious exception being trimeric OMPs.  $\beta$ -barrel proteins found in mitochondrial membranes are descended from ancestral bacterial OMPs (Höhr *et al.*, 2015). A more subtle mechanism may exist to facilitate clustering of  $\beta$ -barrel membrane proteins, and may be present in  $\beta$ -barrel proteins of completely different membrane environments.

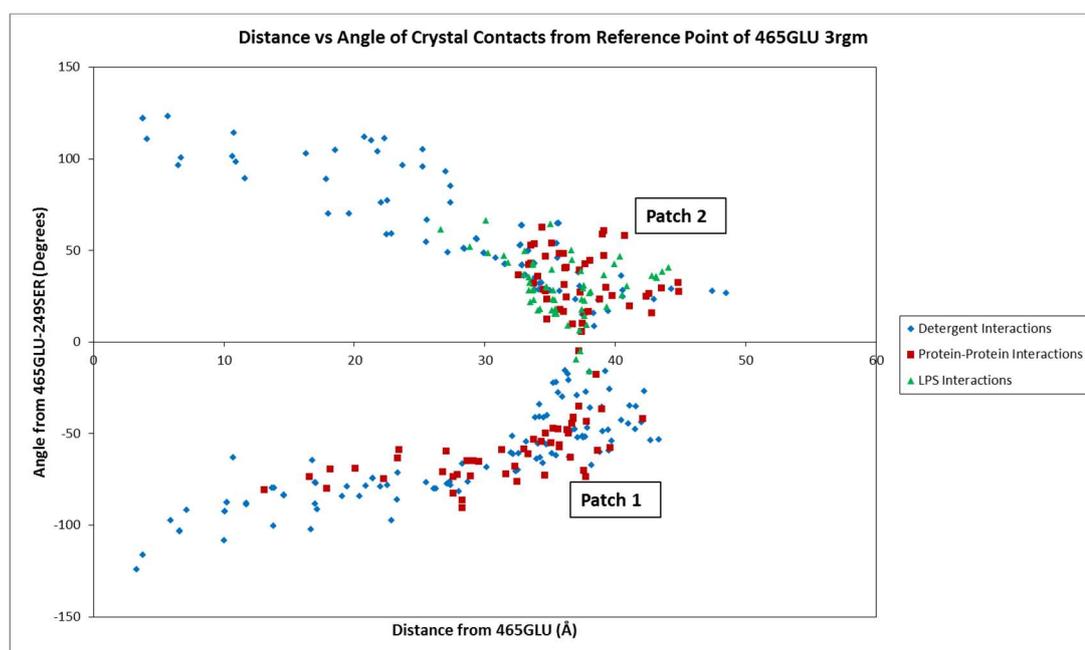
The established theory of hydrophobic mismatching suggests that a mismatch between the height of a  $\beta$ -barrel transmembrane domain and the surrounding bilayer facilitates protein-protein interactions in the membrane. Multiple coarse-grained MD simulations and qualitative analyses have promoted the idea of this hydrophobic mismatch being a genuine property of membrane proteins (Harroun *et al.*, 1999; Schmidt, Guigas and Weiss, 2008; Parton, Klingelhoefer and Sansom, 2011). Being able to analyse membrane protein interactions in 3D will provide insight into whether or not this hydrophobic mismatch hypothesis is credible in a variety of different membrane proteins. As well as observing the hydrophobic mismatch in terms of the barrel height, there is a potential to study the amino acid composition of the  $\beta$ -barrel surface and their hydrophobicity. By comparing the hydrophobicity of a PPI patch to the  $\beta$ -barrel surface as a whole, we can establish whether or not transient interactions might be promoted by the chemical properties of the protein surface.

If conserved PPI patches are identified within OMPs, this would expand our understanding of how an OMP interacts with other components in the OM. Are PPIs a conserved trait within OMPs? If so, what properties enable them to form such interactions? Phylogenetically related mitochondrial Voltage-Dependent Anion Channels (VDACs) have been shown to form oligomeric states under physiological conditions (Geula *et al.*, 2012). If PPI patches are identified within VDACs, their relative position on the  $\beta$ -barrel surface compared to monomeric OMPs can be analysed to determine if they might be evolutionarily conserved.

### 5.3 Results

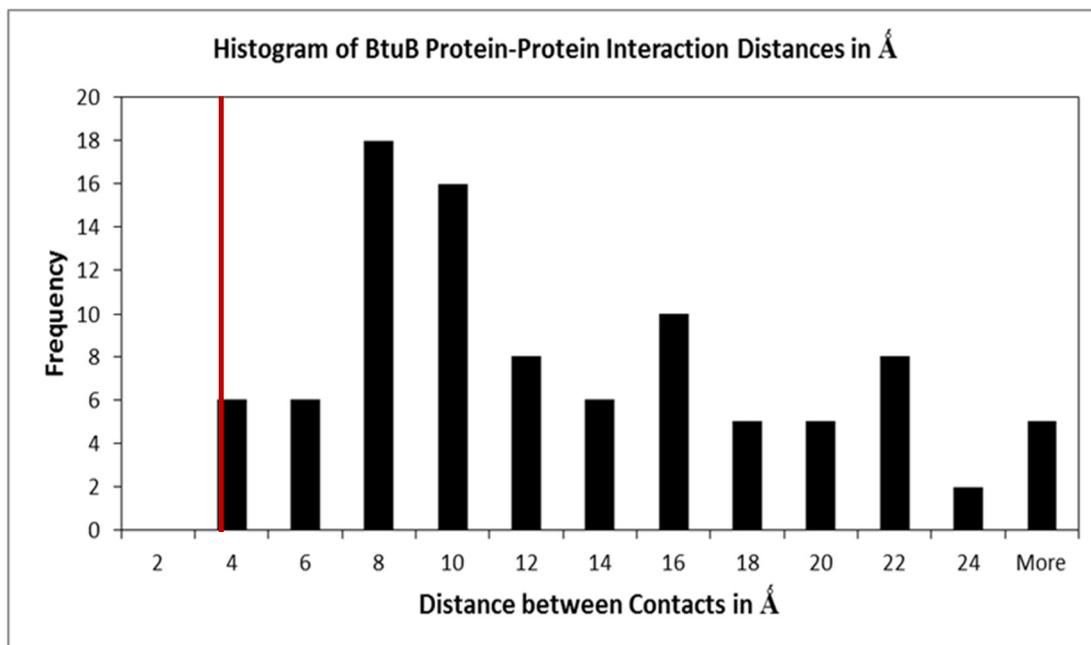
### 5.3.1 Interactions in Monomeric OMPs

The distance versus angle data showed two groupings of intermolecular interactions on the BtuB (3RGM)  $\beta$ -barrel surface. Detergent interactions were present across the entire  $\beta$ -barrel surface, suggesting that they are non-specific (Figure 5-1). A histogram of interaction distances determined interactions to be most prevalent at a distance of 8-10 Å (Figure 5-2). Given that the van der Waals radius of an  $\text{CH}_3$  group is  $\sim 2$  Å, the potential van der Waals interactions were expected to occur at a distance of about 4 Å (Bondi, 1964; Alvarez, 2013). The majority of the observed interactions occurred over a greater distance, implying disordered detergent molecules participate in the identified intermolecular interactions.



**Figure 5-1 –Distance vs Angle of Mapped Crystal Contacts Relative to BtuB**

Distance (Å) from the  $\text{C}\alpha$  of the crystal contact residue to the  $\text{C}\alpha$  of the reference residue (Glu465) plotted against the angle (degrees) of the crystal contact  $\text{C}\alpha$  relative to the line between Glu465 and Ser249. Protein-detergent (blue diamond), protein-protein (red square) and protein-LPS interactions (green triangle) of monomeric OMPs mapped onto BtuB are shown. The two discovered patches of protein-protein interaction are labelled.



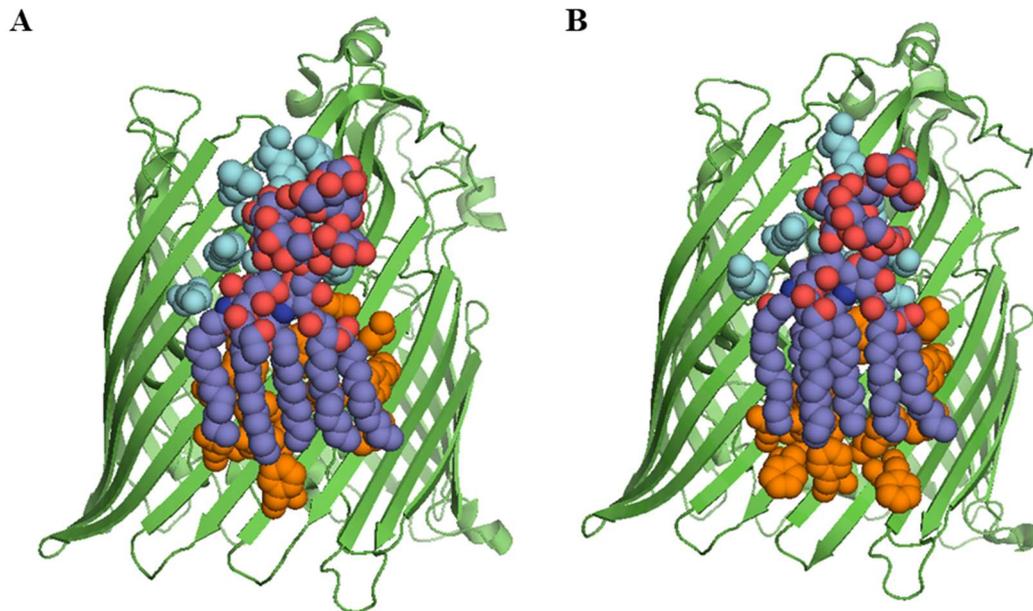
**Figure 5-2 – Histogram of distances between protein-protein interaction residues and the closest symmetry residue**

Distances (Å) plotted against frequency. Monomeric OMPs with a 3D structural similarity to FhuA (1QFF) were analysed by symmetry analysis in WinCOOT. Symmetry interaction distances between residues in the plane of the membrane were recorded. Red vertical line represents 4 Å, the predicted distance for van der Waals (vdW) interactions. Distances are widely spread but cluster in the range 8-10 Å, which is greater than the vdW interaction distance.

FhuA (4CU4) was crystallized with LPS bound in a similar orientation to that occurring in the OM. When analysed, it was shown that the LPS was bound in the same region of the molecule as the original reference structure (FhuA, 1QFF; Figure 5-3). These LPS interactions grouped in the same area as one of the protein-protein groupings (Figure 5-1) implying that this may indeed be a patch of preferred intermolecular interaction *in vivo*. Notably, there was no difference in the backbone structure of pdb files of FhuA crystallized with and without an LPS molecule bound.

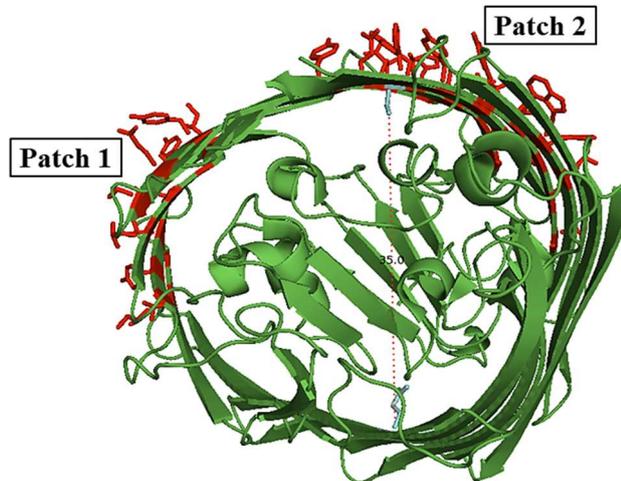
Interestingly, LPS bound to 1QFF and 4CU4 was bound in the same relative position on the  $\beta$ -barrel (Figure 5-6). Many residues involved in these LPS interactions are shared between the two structures. Protein-protein crystal contacts were mapped onto BtuB (3RGM) by structural alignment and revealed two distinct patches for PPIs (Figure 5-7). This allowed groupings of interactions, previously identified in distance vs angle plots (Figure 5-1), to be visualised in 3D. Patch 1 was noticeably shorter compared to the rest of the barrel, a potential structural property for the promotion of PPIs. After mapping aromatic girdle residues onto BtuB (3RGM) (Figure 5-5), it

became clear that they are not specifically located in the PPI patches, rather they are fairly evenly distributed across the barrel surface.



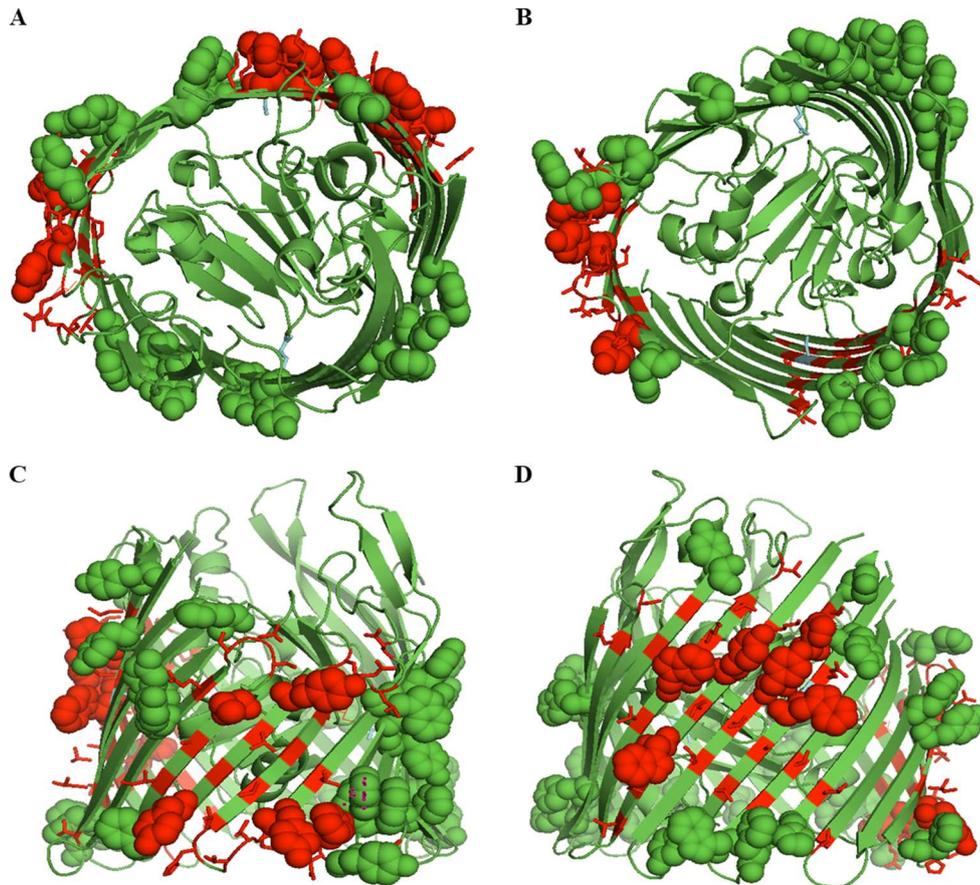
**Figure 5-3 – LPS binding site of FhuA structures 1QFF and 4CU4**

3D PyMOL face-on image of A- FhuA (1QFF) and B- FhuA (4CU4) LPS (Purple-Red-Blue Spheres) binding site. Residues bound to the LPS head groups are shown as cyan-coloured spheres and residues bound to LPS tail groups are shown as orange spheres.



**Figure 5-4 – Intermolecular interaction patches of Monomeric OMPs**

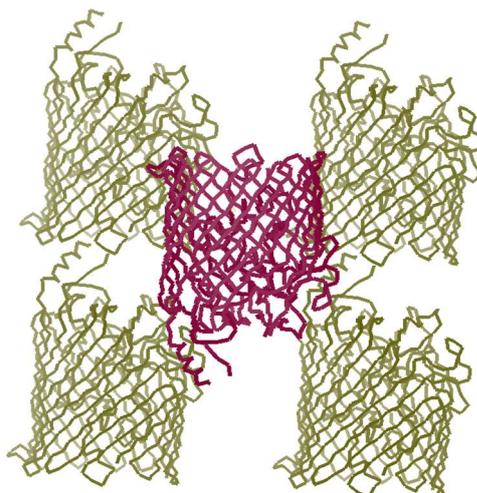
3D PyMOL image, taken from an extracellular view, of BtuB (3rgm) with residues of monomeric OMP protein-protein contacts mapped onto the barrel in red, side chains are shown in stick form. These crystal contact residues are grouped into two distinct patches, patch 1 and patch 2, labelled as above. These patches visualise in 3D crystal contact groupings noted in distance vs angle measurements (Figure 5-1).



**Figure 5-5 – Aromatic girdle residues of BtuB in relation to PPI patches**

3D PyMOL images of BtuB (3RGM) with identified PPI patches mapped on the  $\beta$ -barrel and shown in red. Outward-facing Tyrosine and Tryptophan residues, likely part of the aromatic girdle, are shown as spheres. A- Extracellular view of BtuB with girdle residues and PPIs shown, B- Periplasmic view, C- face-on view of patch 1, D- face-on view of patch 2.

*Neisseria meningitidis* has been shown to survive despite the removal of Lipid A from the outer membrane (Steeghs *et al.*, 1998). FrpB (4AIQ), an iron transporter in *Neisseria meningitidis*, was analysed as part of the monomeric OMP dataset. The symmetry data for this transporter possessed two interfaces for protein contacts; in the same location as Patches 1 and 2, but the  $\beta$ -barrels interacted in an inverted orientation relative to one another (Figure 5-6). Two of the three BtuB-related structures that displayed protein-protein interactions had inverted orientations relative to their symmetry partners. This is significant and could affect the validity of the patches discovered for this organism. As more pdb files of BtuB-related molecules are released, perhaps these findings can be validated as a common feature amongst the majority of  $\beta$ -barrel membrane proteins.



**Figure 5-6 – Symmetry analysis for FrpB iron transporter**

3D WinCOOT image of FrpB (4AIQ) structure from *Neisseria meningitides* with generated symmetry partners. Red is the original molecule and the dark green structures are the generated symmetry partners.

Hydrophobicity analysis of the PPI patches of BtuB revealed Patch 1 is more hydrophobic than the average membrane-facing  $\beta$ -barrel surface, and Patch 2 is less hydrophobic (Table 5-1). The degree of variation of the patches from the average surface hydrophobicity was very similar, about  $\pm 0.7$  for the WW scale.

**Table 5-1 – Hydrophobicity analysis of the PPI patches of BtuB**

Final calculated hydrophobicity values (see Methods section 2.3.4) for the full  $\beta$ -barrel surface and PPI patches of BtuB are shown. For resulting hydrophobicity values (in bold), the more positive the result, the greater the hydrophobicity of the analysed  $\beta$ -barrel surface. Terms are defined in Table 2.2.

Surface Analysed	Full Surface	Patch 1	Patch 2
$\Sigma\text{No.}_{AA}$	101	14	23
$(\Sigma\text{No.}_{TYR,TRP})$	26	4	8
$\Sigma\text{HPCC}_{WW}$	176.5	34.31	23.82
$\Sigma\text{HPCC}_{KD}$	1351	268.9	145.7
$\Sigma\text{HP}_{WW}/\Sigma\text{No.}_{AA}$	386.5	392.9	363.9
$\Sigma\text{HP}_{KD}/\Sigma\text{No.}_{AA}$	837.6	1260	262.3
<b>HPBB<sub>WW</sub></b>	<b>1.748</b>	<b>2.451</b>	<b>1.036</b>
<b>HPBB<sub>KD</sub></b>	<b>13.38</b>	<b>19.21</b>	<b>6.335</b>

**Table 5-2 – Amino Acid Composition of Monomeric OMP Interaction Patches**

Amino acid composition of the full barrel surface and PPI patches of BtuB (3RGM). The proportion of each amino acid is shown as a percent value of the surface and highlighted with cell bars.

Amino Acid	Full Face 3rgm	Amino Acid Percentage	Patch 1	Amino Acid Percentage	Patch 2	Amino Acid Percentage
GLYCINE	11	10.89	2	14.29	1	4.35
ARGININE	1	0.99	0	0.00	0	0.00
HISTIDINE	2	1.98	0	0.00	2	8.70
LYSINE	1	0.99	0	0.00	0	0.00
ASPARTIC ACID	1	0.99	0	0.00	0	0.00
GLUTAMIC ACID	0	0.00	0	0.00	0	0.00
SERINE	4	3.96	0	0.00	0	0.00
THREONINE	4	3.96	0	0.00	2	8.70
ASPARAGINE	1	0.99	0	0.00	1	4.35
GLUTAMINE	4	3.96	1	7.14	1	4.35
CYSTEINE	1	0.99	0	0.00	0	0.00
PROLINE	0	0.00	0	0.00	0	0.00
ALANINE	8	7.92	1	7.14	1	4.35
ISOLEUCINE	7	6.93	2	14.29	0	0.00
LEUCINE	11	10.89	2	14.29	3	13.04
METHIONINE	0	0.00	0	0.00	0	0.00
PHENYLALANINE	7	6.93	0	0.00	0	0.00
TRYPTOPHAN	11	10.89	1	7.14	4	17.39
TRYOSINE	15	14.85	3	21.43	4	17.39
VALINE	12	11.88	2	14.29	4	17.39
<b>SUBTOTALS</b>	101		14		23	

**Table 5-3 – Difference in Amino Acid Composition between the Full Surface and Patches 1 and 2 of BtuB**

The heat map highlighting the percent difference between the amino acid composition of the full barrel surface of BtuB (3RGM) and interaction patches 1 and 2. Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	Patch 1	Difference from Full Barrel	Patch 2	Difference from Full Barrel	Difference of Patch 1 from Patch 2
GLYCINE	14.29	3.39	4.35	-6.54	9.94
ARGININE	0.00	-0.99	0.00	-0.99	0.00
HISTIDINE	0.00	-1.98	8.70	6.72	-8.70
LYSINE	0.00	-0.99	0.00	-0.99	0.00
ASPARTIC ACID	0.00	-0.99	0.00	-0.99	0.00
GLUTAMIC ACID	0.00	0.00	0.00	0.00	0.00
SERINE	0.00	-3.96	0.00	-3.96	0.00
THREONINE	0.00	-3.96	8.70	4.74	-8.70
ASPARAGINE	0.00	-0.99	4.35	3.36	-4.35
GLUTAMINE	7.14	3.18	4.35	0.39	2.80
CYSTEINE	0.00	-0.99	0.00	-0.99	0.00
PROLINE	0.00	0.00	0.00	0.00	0.00
ALANINE	7.14	-0.78	4.35	-3.57	2.80
ISOLEUCINE	14.29	7.36	0.00	-6.93	14.29
LEUCINE	14.29	3.39	13.04	2.15	1.24
METHIONINE	0.00	0.00	0.00	0.00	0.00
PHENYLALANINE	0.00	-6.93	0.00	-6.93	0.00
TRYPTOPHAN	7.14	-3.75	17.39	6.50	-10.25
TRYOSINE	21.43	6.58	17.39	2.54	4.04
VALINE	14.29	2.40	17.39	5.51	-3.11

The amino acid composition of these patches (Table 5-2) was noted alongside the percentage composition relative to the total number of amino acids (Table 5-3). This analysis revealed an absence of phenylalanine residues in PPI patches relative to the full  $\beta$ -barrel surface. Notably, Patch 1 possessed more isoleucine residues than the full  $\beta$ -barrel surface, whereas Patch 2 was noticeably devoid of isoleucine. Patch 1 contains noticeably fewer tryptophan residues (1 residue) compared to Patch 2 (4 residues).

**Table 5-4 – Difference in Amino Acid Composition between BtuB and FhuA in Patch 1**  
Heat map highlighting the percent difference between the amino acid composition of Patch 1 in BtuB (3RGM) and FhuA (1QFF). Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	3rgm BtuB Patch 1	Amino Acid Composition	1qff FhuA Patch 1 Equivalen	Amino Acid Composition	Difference
GLYCINE	2	14.29	0	0.00	14.29
ARGININE	0	0.00	0	0.00	0.00
HISTIDINE	0	0.00	0	0.00	0.00
LYSINE	0	0.00	0	0.00	0.00
ASPARTIC ACID	0	0.00	0	0.00	0.00
GLUTAMIC ACID	0	0.00	0	0.00	0.00
SERINE	0	0.00	1	7.14	-7.14
THREONINE	0	0.00	1	7.14	-7.14
ASPARAGINE	0	0.00	0	0.00	0.00
GLUTAMINE	1	7.14	0	0.00	7.14
CYSTEINE	0	0.00	0	0.00	0.00
PROLINE	0	0.00	0	0.00	0.00
ALANINE	1	7.14	3	21.43	-14.29
ISOLEUCINE	2	14.29	0	0.00	14.29
LEUCINE	2	14.29	1	7.14	7.14
METHIONINE	0	0.00	0	0.00	0.00
PHENYLALANINE	0	0.00	3	21.43	-21.43
TRYPTOPHAN	1	7.14	0	0.00	7.14
TRYOSINE	3	21.43	1	7.14	14.29
VALINE	2	14.29	4	28.57	-14.29
<b>SUBTOTALS</b>	<b>14</b>		<b>14</b>		

**Table 5-5 – Difference in Amino Acid Composition between BtuB and FhuA in Patch 2**  
Heat map shows the percent difference between the amino acid composition of Patch 2 in BtuB (3RGM) and FhuA (1QFF). Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	3rgm BtuB Patch 2	Amino Acid Composition	1qff FhuA Patch 2	Amino Acid Composition	Difference
GLYCINE	1	4.35	0	0.00	4.35
ARGININE	0	0.00	4	14.81	-14.81
HISTIDINE	2	8.70	0	0.00	8.70
LYSINE	0	0.00	4	14.81	-14.81
ASPARTIC ACID	0	0.00	1	3.70	-3.70
GLUTAMIC ACID	0	0.00	1	3.70	-3.70
SERINE	0	0.00	0	0.00	0.00
THREONINE	2	8.70	0	0.00	8.70
ASPARAGINE	1	4.35	1	3.70	0.64
GLUTAMINE	1	4.35	2	7.41	-3.06
CYSTEINE	0	0.00	0	0.00	0.00
PROLINE	0	0.00	1	3.70	-3.70
ALANINE	1	4.35	0	0.00	4.35
ISOLEUCINE	0	0.00	0	0.00	0.00
LEUCINE	3	13.04	4	14.81	-1.77
METHIONINE	0	0.00	0	0.00	0.00
PHENYLALANINE	0	0.00	6	22.22	-22.22
TRYPTOPHAN	4	17.39	0	0.00	17.39
TRYOSINE	4	17.39	1	3.70	13.69
VALINE	4	17.39	2	7.41	9.98
<b>SUBTOTALS</b>	<b>23</b>		<b>27</b>		

**Table 5-6 – Difference in Amino Acid Composition between Head and Tail LPS contacts for FhuA in Patch 2**

Heat map shows the percent difference between the amino acid composition of FhuA residues in contact with LPS head groups, and FhuA residues in contact with lipid A acyl chains. Residue data was taken from FhuA (1QFF) in Patch 2. Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

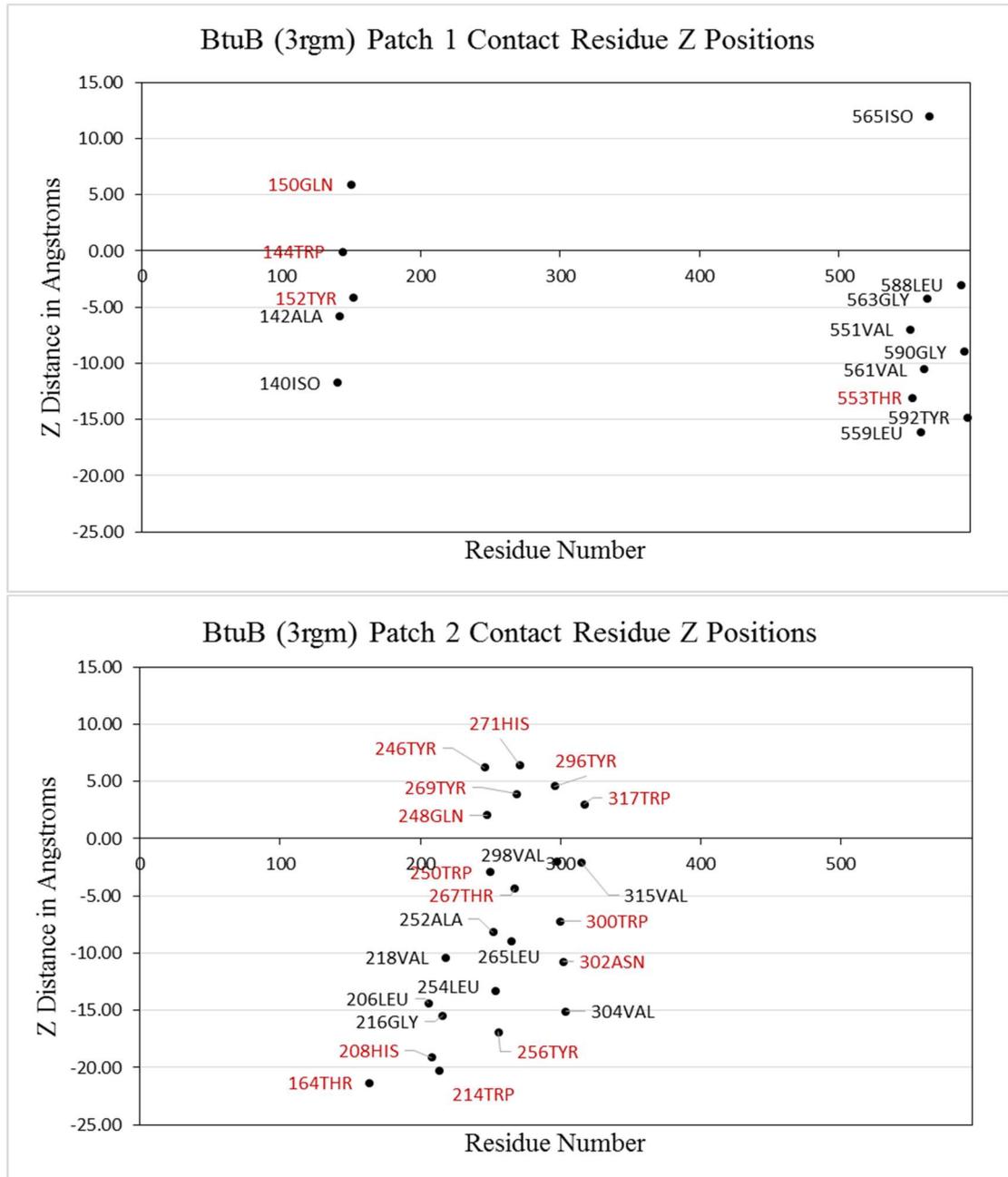
Amino Acid	1qff FhuA Patch 2 TAIL LPS	Amino Acid Composition	1qff FhuA Patch 2 HEAD LPS	Amino Acid Composition	Difference
GLYCINE	0	0.00	0	0.00	0.00
ARGININE	0	0.00	4	30.77	-30.77
HISTIDINE	0	0.00	0	0.00	0.00
LYSINE	1	8.33	3	23.08	-14.74
ASPARTIC ACID	0	0.00	1	7.69	-7.69
GLUTAMIC ACID	1	8.33	0	0.00	8.33
SERINE	0	0.00	0	0.00	0.00
THREONINE	0	0.00	0	0.00	0.00
ASPARAGINE	0	0.00	1	7.69	-7.69
GLUTAMINE	1	8.33	1	7.69	0.64
CYSTEINE	0	0.00	0	0.00	0.00
PROLINE	1	8.33	0	0.00	8.33
ALANINE	0	0.00	0	0.00	0.00
ISOLEUCINE	0	0.00	0	0.00	0.00
LEUCINE	1	8.33	3	23.08	-14.74
METHIONINE	0	0.00	0	0.00	0.00
PHENYLALANINE	6	50.00	0	0.00	50.00
TRYPTOPHAN	0	0.00	0	0.00	0.00
TRYOSINE	1	8.33	0	0.00	8.33
VALINE	2	16.67	0	0.00	16.67
<b>SUBTOTALS</b>	<b>14</b>		<b>13</b>		

Both patches of BtuB contain less phenylalanine residues than their respective patches in FhuA (Table 5-4 and 5-5). This is especially notable in patch 2, where the LPS molecule is bound. Comparing the head and tail LPS contact residues of FhuA reveals a greater prevalence of phenylalanine residues in tail contacts and arginine in the Head contacts (Table 5-6). The hydrophobicity profiles match expectations for the respective contacts to head and tail regions of the LPS molecule.

**Table 5-7 – Z Distances of BtuB Interaction patches**

The calculated z distance (Å) of the patch residues in BtuB (3RGM) from the horizontal plane of reference residue Ser249 in BtuB. The distances are heat mapped with blue showing a positive z distance (towards extracellular face) and red showing a negative z distance (towards periplasmic face).

<b>Interface 1</b>	<b>Z Distance from 249SER</b>	<b>Interface 2</b>	<b>Z Distance from 249SER</b>
140ISO	-11.72	164THR	-21.32
142ALA	-5.77	206LEU	-14.38
144TRP	-0.06	208HIS	-19.10
150GLN	5.92	214TRP	-20.27
152TYR	-4.16	216GLY	-15.45
551VAL	-6.97	218VAL	-10.41
553THR	-13.08	246TYR	6.24
559LEU	-16.17	248GLN	2.06
561VAL	-10.52	250TRP	-2.93
563GLY	-4.26	252ALA	-8.17
565ISO	12.00	254LEU	-13.30
588LEU	-3.00	256TYR	-16.88
590GLY	-8.99	265LEU	-8.95
592TYR	-14.81	267THR	-4.36
		269TYR	3.89
		271HIS	6.47
		296TYR	4.58
		298VAL	-1.95
		300TRP	-7.24
		302ASN	-10.76
		304VAL	-15.08
		315VAL	-2.08
		317TRP	3.01



**Figure 5-7 – Scatter plots showing the Z Distance of Patch Residues and their Residue Number**

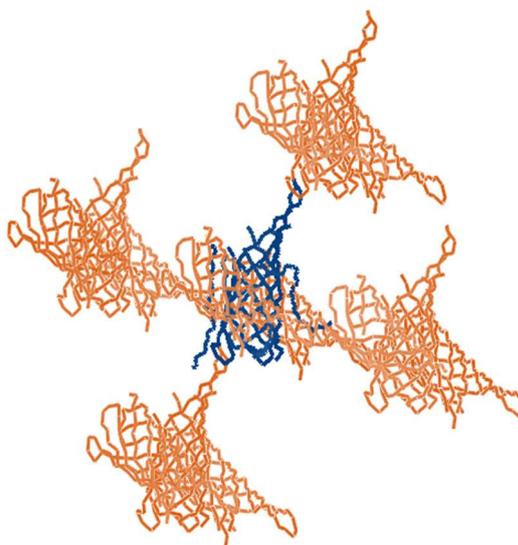
Residues are plotted as black circles with their respective number and amino acid code labelled. Hydrophobic residues are labelled in black text; polar residues are labelled in red text. Positive Z distance refers to residues above Ser249 and closer to the extracellular face; negative Z values are closer to the periplasmic face.

As one would expect for the aromatic girdle residues, tyrosine and tryptophan are located close to the upper and lower boundaries of BtuB (Table 5-7). Histidine and threonine residues gather at the boundaries of the hydrophobic and hydrophilic regions of BtuB (Figure 5-7), while less bulky alanine residues are located close to the centre

of the hydrophobic region. Polar amino acids tend to gather at the borders of the hydrophobic region of the  $\beta$ -barrel.

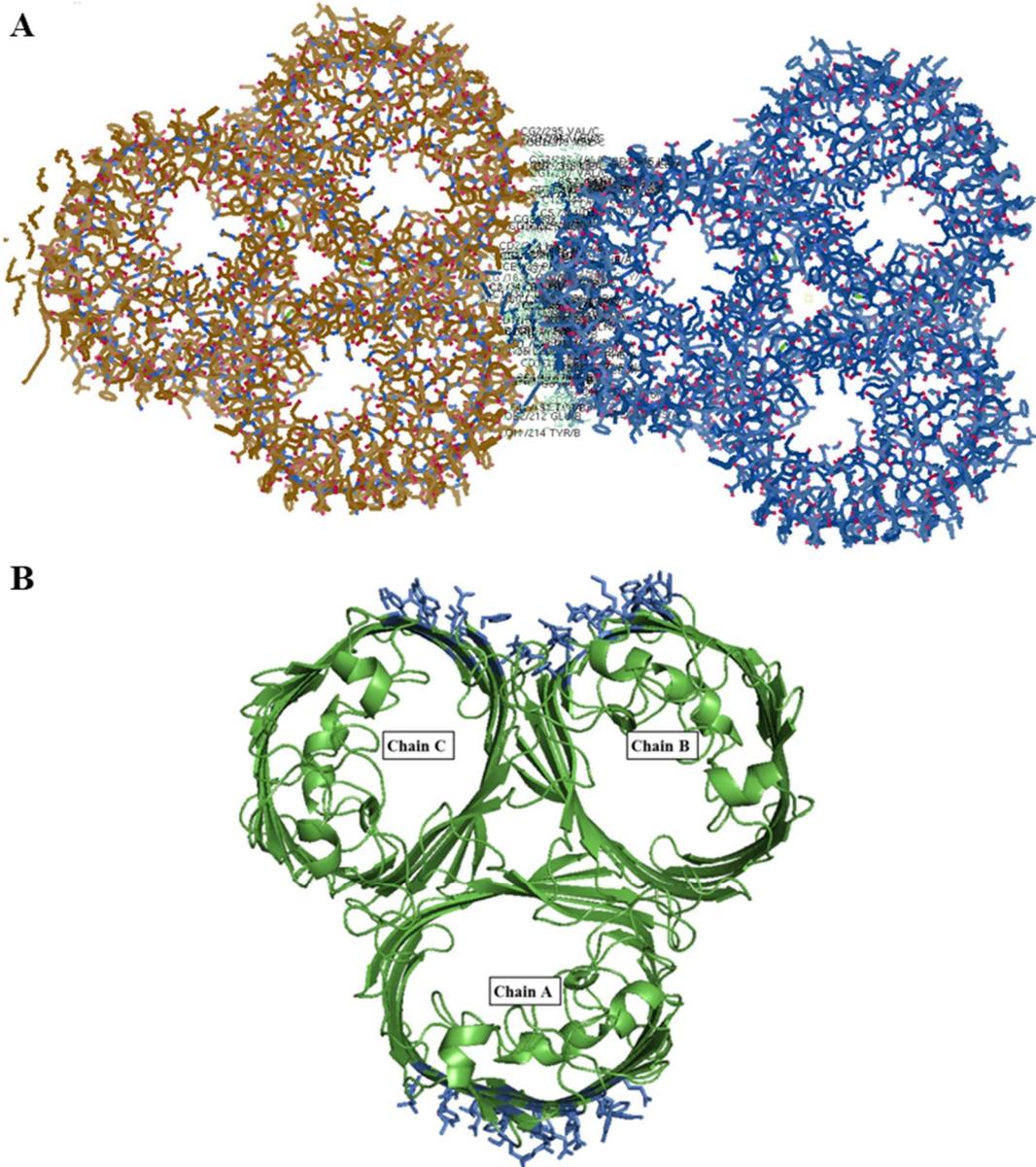
### 5.3.2 Intermolecular Interactions in OmpA and OmpC

Monomeric OmpA displayed only oblique, detergent-mediated interactions in the plane of the membrane bilayer (Figure 5-8). The reference trimeric OmpC (1OSM) showed planar interactions characterised as detergent-based interactions (Figure 5-9A). These were subsequently visualised in 3D using PyMOL (Figure 5-9B). Unfortunately, the lack of individual trimer structures within pdb database prevented the detection of PPI patches for trimeric porin-porin interactions. Residues involved in the trimerisation interfaces of OmpC were noted. The trimer interfaces are located in similar positions on the  $\beta$ -barrel relative to patch 1 and patch 2 of BtuB (Figure 5-10). Notably, the patches are closer together, as shown by data points that link the two groupings, and by mapping trimer interfaces onto the  $\beta$ -barrel itself (Figures 5-11 and 5-12).



**Figure 5-8 – Symmetry simulation of OmpA (1QJP)**

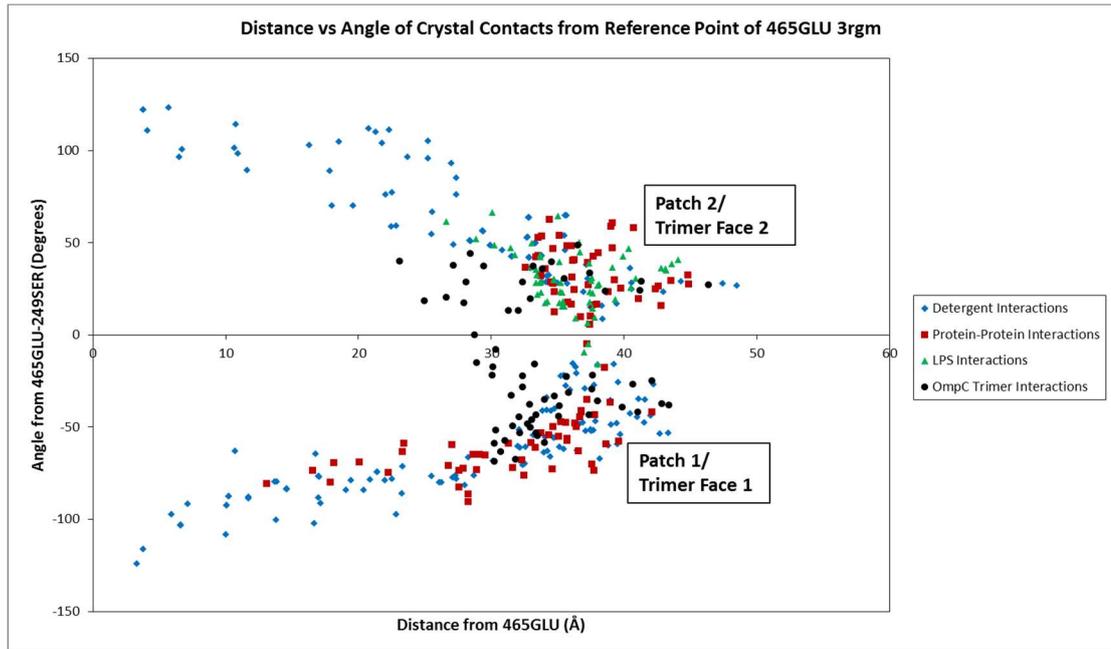
WinCOOT 3D image of OmpA (blue) and its simulated symmetry partners (orange). The detergent bound to OmpA forms all symmetry interactions. The symmetry partners are also not in the plane of the membrane bilayer and have oblique orientations.



**Figure 5-9 – Trimer-Trimer Interactions in OmpC (1OSM)**

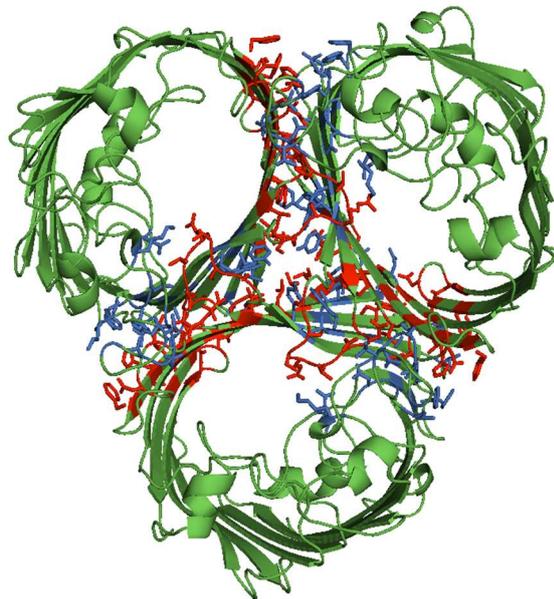
**A** – 3D PyMOL image of two OmpC molecules (brown and blue) with raw distances versus angles of interactions between them. All interactions are formed between the bound decane detergent molecules.

**B** – 3D PyMOL image of OmpC molecule with symmetry interaction residues highlighted (blue) and side chains in stick form. Chains A, B and C of OmpC are



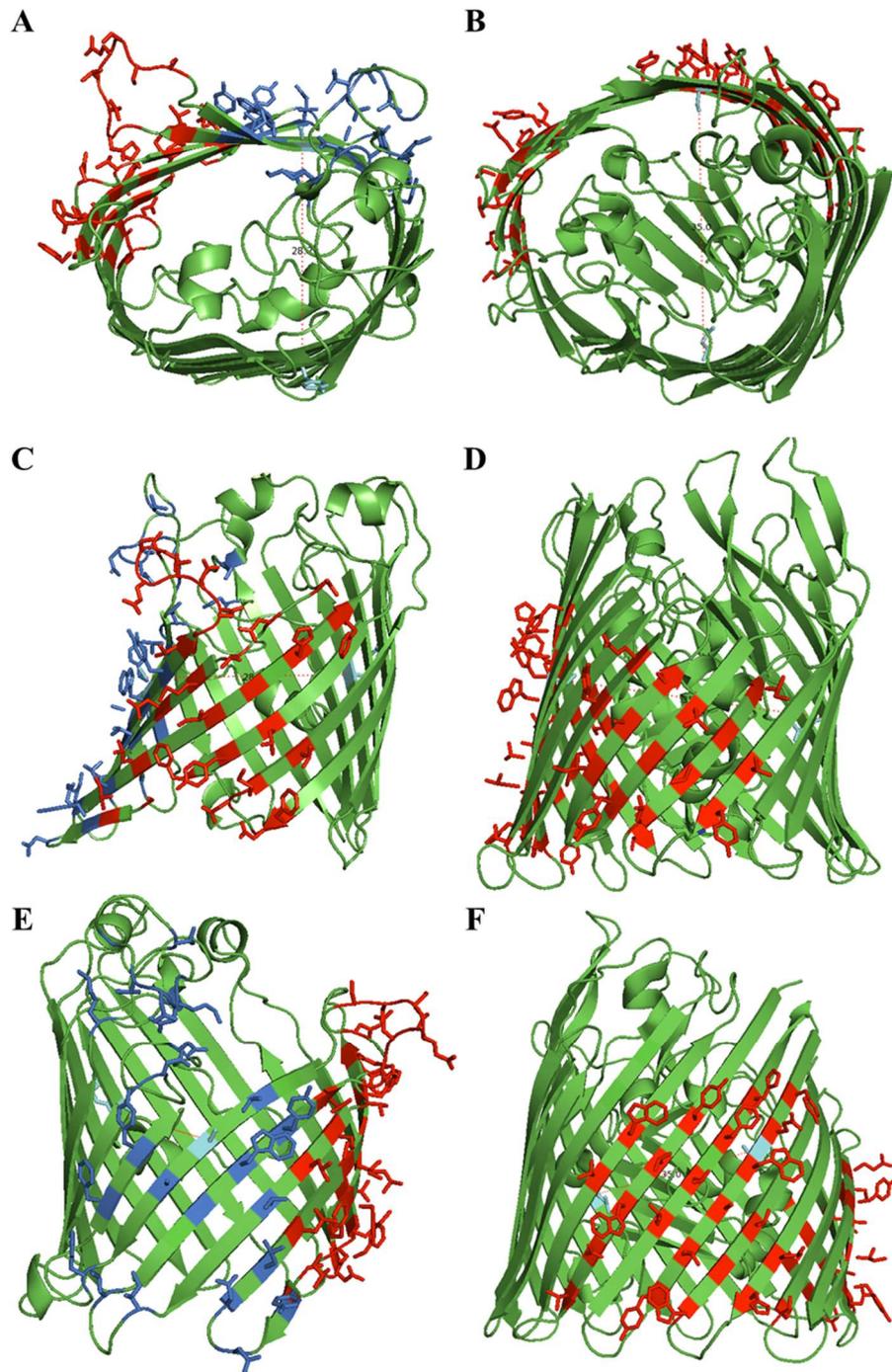
**Figure 5-10 – Plot of OmpC trimer contacts as distance vs angle relative to contacts in BtuB**

Distance vs angle measurements for OmpC trimer contacts and BtuB protein-protein interactions relative to reference residues Glu465 and Tyr231-Ser249. BtuB (3RGM) contacts are measured from Glu465 and OmpC (1OSM) contacts from Tyr231, the closest residue by  $C\alpha$  distance when superimposed.



**Figure 5-11 – Interactions at OmpC Trimer Interface**

3D PyMOL cartoon image of OmpC (1OSM) with trimer interface residues labelled. Trimer interfaces 1 (red) and 2 (blue) have their side chains visualised in stick form to better highlight the residues involved.



**Figure 5-12 – Interaction Patches in OmpC and BtuB**

Comparative image of OmpC and BtuB interaction patches. OmpC (1OSM) trimer interfaces 1 (red) and 2 (blue) are highlighted. BtuB (3RGM) PPI patches are highlighted in red. A - Extracellular top-down view of OmpC monomer, B- Extracellular top-down view of BtuB, C - Face-on view of OmpC trimer face 1, D - Face-on view of BtuB PPI patch 1, E - Face-on view of OmpC trimer face 2, F - Face-on view of BtuB PPI patch 2. Final image shows OmpC with trimer interfaces 1 and 2 mapped onto the protein in red and blue, respectively. All interface residues have their side chains shown in stick form.

Analysing the hydrophobicity of the trimer-trimer contacts present in OmpC (Table 5-8) revealed that the contact patches, and by extension most of the exposed face in the trimer, possess a lower hydrophobicity score than the monomer  $\beta$ -barrel surface. Interestingly, patches on chains B and C that interact with chain A from another trimer, possess a lower and higher hydrophobicity score than the chain A patch, respectively. This pairing of regions of higher and lower hydrophobicity appears to be present in many  $\beta$ -barrel interactions, like the monomeric OMP PPI patches analysed previously.

Comparative 3D visualisation of trimer interfaces and PPI patches mapped on OmpC and BtuB displayed certain structural characteristics (Figure 5-12). The membrane-facing  $\beta$ -barrel surfaces of trimer interface 1 and PPI patch 1 are much shorter compared to the rest of the  $\beta$ -barrel. This is consistent with the hydrophobic mismatch theory established by MD simulations (Schmidt, Guigas and Weiss, 2008). There was a visible gap across the  $\beta$ -barrel between the two PPI patches of BtuB, whereas the trimer interfaces have no such gap. It could be argued that it is one large interface. Trimer patch 2, much like patch 1, has a relatively short membrane-facing  $\beta$ -barrel surface. In contrast, the similarly located patch 2 in BtuB has a much higher  $\beta$ -barrel surface. This difference may be a key structural factor in defining whether or not a membrane protein is able to form trimeric structures. The hydrophobicity of these trimer interfaces was analysed (Table 5-9) using the same methodology as the PPI patches in BtuB. Trimer interfaces 1 and 2 were of lower and higher hydrophobicity, respectively, than the overall  $\beta$ -barrel membrane-facing surface.

**Table 5-8 – Hydrophobicity analysis of the OmpC protein-detergent contact patches**

In terms of the resulting hydrophobicity values (in bold), the more positive the result, the greater the hydrophobicity of the analysed  $\beta$ -barrel surface. Terms are defined in Table 2.1. Patches are visualised in 3D in Figure 5-9.

Surface Analysed	Full Surface	Chain A Patch	Chain B Patch	Chain C Patch
$\Sigma N_{O_{AA}}$	87	30	20	14
$(\Sigma N_{O_{TYR,TRP}})$	16	6	3	2
$\Sigma HP_{CC_{WW}}$	36.15	-39.53	-35.95	-18.16
$\Sigma HP_{CC_{KD}}$	1072	142.4	-137.5	130.2
$\Sigma HP_{WW}/\Sigma N_{O_{AA}}$	238.5	28.67	117.5	162.1
$\Sigma HP_{KD}/\Sigma N_{O_{AA}}$	993.1	156.7	-330.0	957.1
$HP_{BB_{WW}}$	<b>0.4155</b>	<b>-1.318</b>	<b>-1.797</b>	<b>-1.297</b>
$HP_{BB_{KD}}$	<b>12.33</b>	<b>4.746</b>	<b>-6.87</b>	<b>9.298</b>

**Table 5-9 - Final calculated hydrophobicity values for the membrane-facing  $\beta$ -barrel surface and trimer interfaces of OmpC**

In terms of the resulting hydrophobicity values (in bold), the more positive the result, the greater the hydrophobicity of the analysed  $\beta$ -barrel surface. Terms are defined in table 2.1.

Surface Analysed	Full Surface	Patch 1	Patch 2
$\Sigma N_{O_{AA}}$	87	18	15
$(\Sigma N_{O_{TYR,TRP}})$	16	1	4
$\Sigma HP_{CC_{WW}}$	36.15	-1.702	24.35
$\Sigma HP_{CC_{KD}}$	1072	15.97	243
$\Sigma HP_{WW}/\Sigma N_{O_{AA}}$	238.5	183.9	438
$\Sigma HP_{KD}/\Sigma N_{O_{AA}}$	993.1	1306	676.6
$HP_{BB_{WW}}$	<b>0.4155</b>	<b>-0.09458</b>	<b>1.623</b>
$HP_{BB_{KD}}$	<b>12.33</b>	<b>8.874</b>	<b>16.2</b>

**Table 5-10– Amino Acid Composition of OmpC Trimer Interfaces**

Amino acid composition for the full  $\beta$ -barrel surface and the trimer interfaces of OmpC (1osm). The relative proportion of each amino acid is shown as a percentage of the surface and highlighted with cell bars.

Amino Acid	Full Face Residues	Full Monomer Barrel	Trimer Face 1	Trimer Interface 1	Trimer Face 2	Trimer Interface 2
GLYCINE	6	6.90	1	5.56	1	6.67
ARGININE	0	0.00	0	0.00	0	0.00
HISTIDINE	1	1.15	1	5.56	0	0.00
LYSINE	2	2.30	0	0.00	0	0.00
ASPARTIC ACID	3	3.45	2	11.11	0	0.00
GLUTAMIC ACID	1	1.15	0	0.00	0	0.00
SERINE	0	0.00	0	0.00	0	0.00
THREONINE	6	6.90	2	11.11	1	6.67
ASPARAGINE	2	2.30	0	0.00	1	6.67
GLUTAMINE	3	3.45	0	0.00	0	0.00
CYSTEINE	0	0.00	0	0.00	0	0.00
PROLINE	1	1.15	0	0.00	0	0.00
ALANINE	6	6.90	0	0.00	2	13.33
ISOLEUCINE	7	8.05	2	11.11	1	6.67
LEUCINE	12	13.79	3	16.67	2	13.33
METHIONINE	2	2.30	2	11.11	0	0.00
PHENYLALANINE	10	11.49	3	16.67	2	13.33
TRYPTOPHAN	1	1.15	0	0.00	1	6.67
TRYOSINE	15	17.24	1	5.56	3	20.00
VALINE	9	10.34	1	5.56	1	6.67
<b>SUBTOTAL</b>	<b>87</b>		<b>18</b>		<b>15</b>	

**Table 5-11 –Heat Map of Amino Acid Composition in Trimer Interfaces of OmpC**

Heat map shows the percent difference between the amino acid composition of the entire membrane-facing  $\beta$ -barrel surface of OmpC (1OSM) and trimer interfaces 1 and 2. Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	Trimer Interface 1	Difference from Full Barrel	Trimer Interface 2	Difference from Full Barrel
GLYCINE	5.56	-1.34	6.67	-0.23
ARGININE	0.00	0.00	0.00	0.00
HISTIDINE	5.56	4.41	0.00	-1.15
LYSINE	0.00	-2.30	0.00	-2.30
ASPARTIC ACID	11.11	7.66	0.00	-3.45
GLUTAMIC ACID	0.00	-1.15	0.00	-1.15
SERINE	0.00	0.00	0.00	0.00
THREONINE	11.11	4.21	6.67	-0.23
ASPARAGINE	0.00	-2.30	6.67	4.37
GLUTAMINE	0.00	-3.45	0.00	-3.45
CYSTEINE	0.00	0.00	0.00	0.00
PROLINE	0.00	-1.15	0.00	-1.15
ALANINE	0.00	-6.90	13.33	6.44
ISOLEUCINE	11.11	3.07	6.67	-1.38
LEUCINE	16.67	2.87	13.33	-0.46
METHIONINE	11.11	8.81	0.00	-2.30
PHENYLALANINE	16.67	5.17	13.33	1.84
TRYPTOPHAN	0.00	-1.15	6.67	5.52
TRYOSINE	5.56	-11.69	20.00	2.76
VALINE	5.56	-4.79	6.67	-3.68

**Table 5-12 – Heat Map comparing OmpC Trimer Interfaces and Monomeric BtuB Interaction Patches**

Heat map shows the percent difference between the amino acid composition of the trimer interfaces of OmpC (1OSM) and the interaction patches of BtuB (3RGM). Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	Trimer Interface 1	Difference from BtuB Patch 1	Trimer Interface 2	Difference from BtuB Patch 2
GLYCINE	5.56	-8.73	6.67	2.32
ARGININE	0.00	0.00	0.00	0.00
HISTIDINE	5.56	5.56	0.00	-8.70
LYSINE	0.00	0.00	0.00	0.00
ASPARTIC ACID	11.11	11.11	0.00	0.00
GLUTAMIC ACID	0.00	0.00	0.00	0.00
SERINE	0.00	0.00	0.00	0.00
THREONINE	11.11	11.11	6.67	-2.03
ASPARAGINE	0.00	0.00	6.67	2.32
GLUTAMINE	0.00	-7.14	0.00	-4.35
CYSTEINE	0.00	0.00	0.00	0.00
PROLINE	0.00	0.00	0.00	0.00
ALANINE	0.00	-7.14	13.33	8.99
ISOLEUCINE	11.11	-3.17	6.67	6.67
LEUCINE	16.67	2.38	13.33	0.29
METHIONINE	11.11	11.11	0.00	0.00
PHENYLALANINE	16.67	16.67	13.33	13.33
TRYPTOPHAN	0.00	-7.14	6.67	-10.72
TRYOSINE	5.56	-15.87	20.00	2.61
VALINE	5.56	-8.73	6.67	-10.72

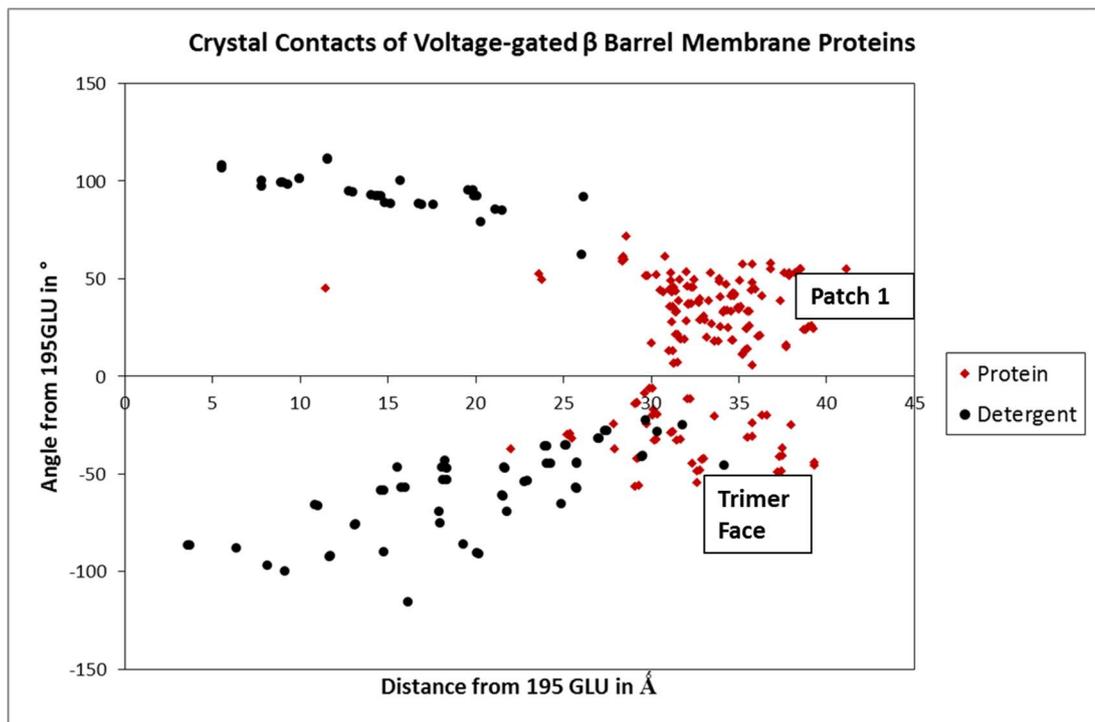
The trimer interfaces have been compared to their corresponding interaction patches in BtuB in terms of relative position on the  $\beta$ -barrel (Table 5-12). Notably, trimer interface 1 has significantly fewer tyrosine residues than the full monomer barrel. It also has a greater proportion of methionine and aspartic acid residues compared to the full membrane-facing  $\beta$ -barrel surface.

There are some significant differences in the amino acid composition of the trimer interfaces relative to the monomeric OMP interaction patches. Both trimer interfaces have significantly more phenylalanine residues than the interaction patches of BtuB. In comparison to the full membrane-facing surface, interface 1 has fewer tyrosine residues than patch 1 in BtuB. Both trimer interfaces have fewer tryptophan and valine residues than the corresponding patches in BtuB. Finally, trimer interface 1 has more methionine, aspartic acid and threonine residues than patch 1 in BtuB.

### 5.3.3 Intermolecular Interactions in Voltage-Dependent Anion Channels

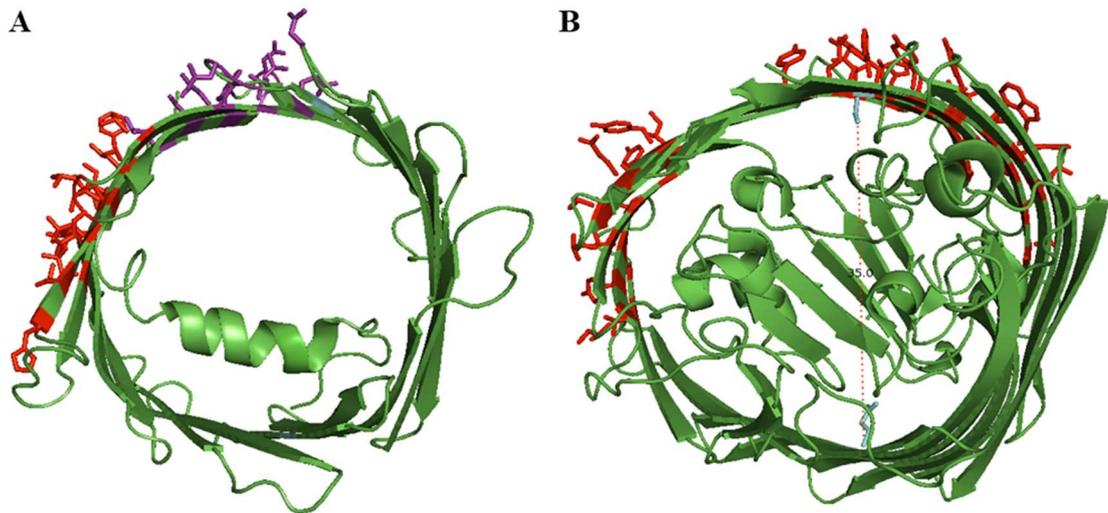
Mitochondrial VDAC proteins show a similar grouping to monomeric OMP proteins in PPI patch 1 and to a lesser extent PPI patch 2 (Figure 5-13). Protein-protein

interactions mapped onto the human VDAC (2JK4)  $\beta$ -barrel relative to the PPI patches of BtuB suggest that this patch is in the same position as patch 1 of BtuB (Figure 5-14). The data suggest that a PPI patch is a common feature of most  $\beta$ -barrel membrane proteins. Noticeably, the PPI patch is located in an area where the  $\beta$ -barrel is shorter and free from aromatic girdle residues, which is consistent with the hydrophobic mismatch theory (Figure 5-15). Patches 1 and 2 in BtuB contain 7 aromatic girdle residues (tyrosine and tryptophan), whereas patch 1 in VDAC contains no girdle residues.



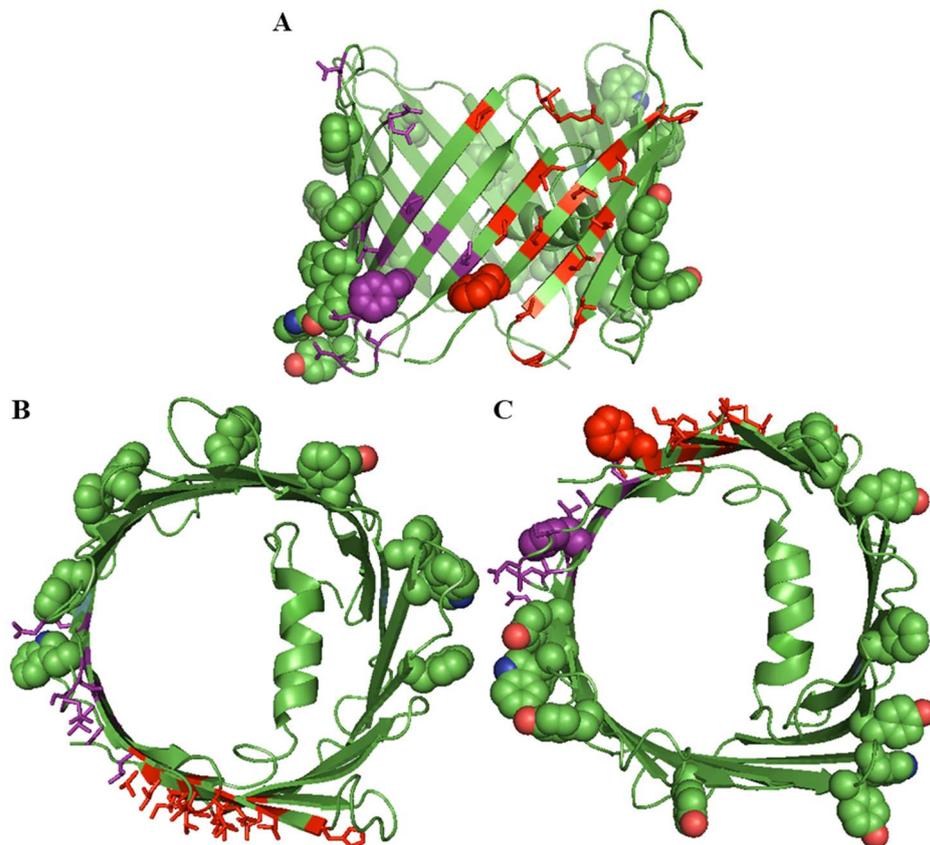
**Figure 5-13 – Distance versus angle for crystal contacts in VDACs**

Distance between  $C\alpha$  of crystal contact residues versus angle of  $C\alpha$  of Glu195 relative to reference residues Glu195 and Gly85. Protein-protein contacts are shown as red diamonds and protein-detergent contacts are shown as black circles.



**Figure 5-14 – Relative PPI patches of BtuB (3RGM) and VDAC (2JK4)**

PPI patches are highlighted in red and the additional trimerisation patch is highlighted in purple. Both  $\beta$ -barrel proteins (BtuB, bottom and VDAC, top) are orientated relative to the same reference residues (Glu465-Ser249 and Glu185-Gly85, highlighted in cyan) and viewed from the extracellular face.



**Figure 5-15 – Location of aromatic girdle residues and intermolecular interaction patches in VDAC**

PyMOL 3D cartoon image of VDAC (2JK4). Extracellular (Top), periplasmic (Middle), patch face-on (Bottom) views are shown as spheres and coloured orange. PPI patch is shown in red; trimerisation patch is shown in purple.

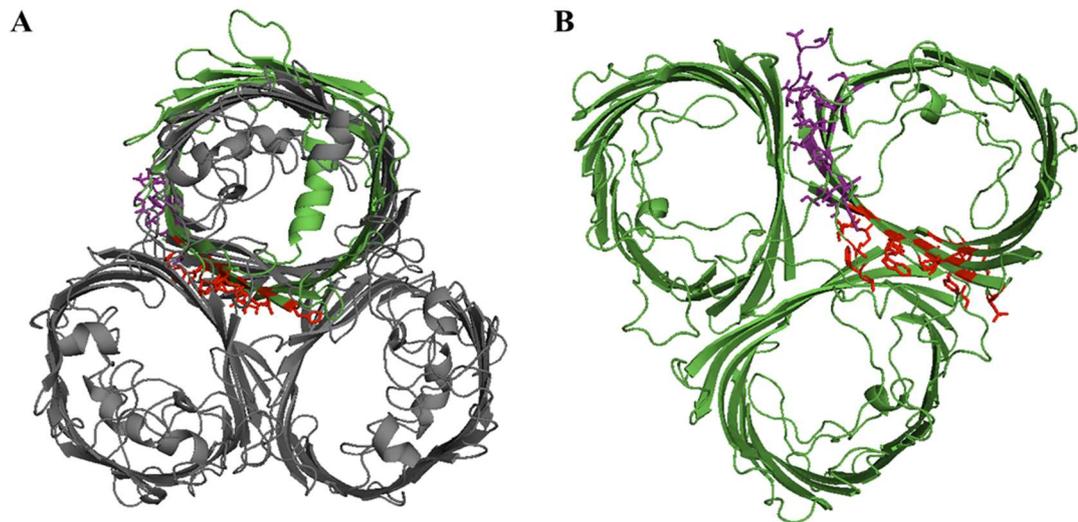
In terms of hydrophobicity, the two different scales lead to different conclusions when comparing the VDAC patch and the average membrane-facing  $\beta$ -barrel surface (Table 5-13). The White and Wimley (1998) hydrophobicity scale indicated that patch 1 is less hydrophobic than the rest of the membrane-facing  $\beta$ -barrel surface, while the Kyte and Doolittle (1982) scale indicated the opposite. VDAC PPI patch 1, just like BtuB PPI patch 1, is shorter in height compared to the rest of the membrane-facing surface (Figure 5-15). This appeared to be a conserved feature amongst patches in this region of  $\beta$ -barrel membrane proteins.

**Table 5-13 – Hydrophobicity calculations for entire membrane-facing surface and PPI patch 1 in VDAC (2JK4)**

Final values using both hydrophobicity scales are highlighted in bold. Terms are defined in Table 5.1.

Surface Analysed	Full Surface	Patch 1
$\Sigma\text{No.}_{\text{AA}}$	83	12
$(\Sigma\text{No.}_{\text{TYR,TRP}})$	11	0
$\Sigma\text{HPCC}_{\text{WW}}$	202.4	26.57
$\Sigma\text{HPCC}_{\text{KD}}$	1685	384.8
$\Sigma\text{HP}_{\text{WW}}/\Sigma\text{No.}_{\text{AA}}$	0.4017	0.2983
$\Sigma\text{HP}_{\text{KD}}/\Sigma\text{No.}_{\text{AA}}$	1.788	3.258
<b>HPBB<sub>WW</sub></b>	<b>2.439</b>	<b>2.214</b>
<b>HPBB<sub>KD</sub></b>	<b>20.3</b>	<b>32.07</b>

VDAC was superimposed onto the OmpC trimer to compare corresponding interaction patches. PPI Patch 1 of the VDAC was in the same position on the  $\beta$ -barrel surface as one of the trimerisation interfaces of OmpC (Figure 5-16). This suggests that this patch may be conserved between BtuB, OmpC and VDAC. The trimerisation interface present in certain VDACs (1PRN) is not involved.



**Figure 5-16 – Interaction patches of VDAC superimposed onto OmpC**

Comparison between interaction patches of trimeric OMPs and mitochondrial VDAC proteins. **A**- Human VDAC (2JK4, green) superimposed onto OmpC trimer (1OSM, grey) viewed from extracellular face. Patch 1 shown in red and trimer interface in purple. **B** – VDAC trimer (1PRN) with the same colourin as 2JK4 for the trimer patches viewed from extracellular face. Interaction patches have amino acid side chains displayed in stick form.

## 5.4 Discussion

### 5.4.1 Overview of Symmetry Analysis

Through symmetry analysis, it was possible to identify preferred patches of intermolecular interaction across all three of the membrane protein families studied. In addition, one of the patches (Patch 1) is conserved between monomeric OMPs and mitochondrial VDACs. It could be that this interaction patch is essential to the viability of membrane proteins and has been evolutionarily conserved since it first arose in a distant common ancestor. Strong evidence for evolutionary conservation between these two proteins already exists, because VDAC was shown to insert and assemble within the bacterial outer membrane (Walther *et al.*, 2010).

Whilst OmpA and OmpC based searches did not produce direct intermolecular interaction patches, this was severely hindered by the availability of pdb files with symmetry partners in the plane of the membrane bilayer. Patch 1 from VDAC lies in the same relative position as the OmpC trimer interface (Figure 5-16). Given that this patch was also present in monomeric OMPs, it suggests that this PPI patch evolved into a trimerisation interface.

### 5.4.2 Hydrophobicity of PPI Patches

When investigating what differentiates patches from the rest of the  $\beta$ -barrel surface, hydrophobicity appears to be a key factor. Proteins with multiple PPI patches or trimer interfaces have higher and a lower hydrophobicity, respectively, than the average  $\beta$ -barrel surface (Tables 5-1 and 5-9). Hydrophobicity of these patches would drive intermolecular interactions, and manipulate the bilayer between them through transient, energetically favourable conformations of LPS. PPI patches and trimer interfaces mapped onto BtuB (3RGM) and OmpC (1OSM), respectively, each possess a patch/interface of high hydrophobicity and patch/interface of low hydrophobicity. The higher and lower hydrophobicity patches are in opposite relative locations on the  $\beta$ -barrel surface. The main differences between the PPI patches of BtuB and trimer interfaces of OmpC are the greater vertical height of the  $\beta$ -barrel, and the greater separation distance between patches on the  $\beta$ -barrel surface.

The patches identified in this investigation generally possess a positive hydrophobic mismatch, with the exception of patch 2 in BtuB. Both trimer interfaces in OmpC display a positive hydrophobic mismatch compared to BtuB, where PPI patches 1 and 2 are positive and negative, respectively (Table 5-9). This may provide a structural rationale as to why BtuB remains a monomer in the membrane, whilst OmpC exists as a trimer. There is potential to explore this hypothesis further, engineering a BtuB monomer through amino acid mutations to possess a greater hydrophobic mismatch, and then determine if it forms a trimer *in vivo*.

In terms of the usefulness of the hydrophobicity scales themselves, the White-Wimley scale appears more relevant in this study. This scale was determined by measuring the partitioning of host-guest peptides at membrane bilayer interfaces, thus it is derived from an environment far closer to the *in vivo* membrane bilayer (White and Wimley, 1998). In contrast, the Kyte-Doolittle scale was calculated for transfer of the amino acid side chain from water to gas phase. Relative hydrophobicity calculated using the WW scale shows a lower hydrophobicity for patch 1 in VDACS, whilst relative hydrophobicity calculated using the K-D scale indicated it was significantly more hydrophobic (Table 5-13). Apart from this disagreement, the two scales mostly provided similar conclusions on other patches in this investigation.

Investigating the amino acid composition of the interaction patches for these  $\beta$ -barrel OMPs has revealed some interesting differences. When comparing the composition of the patches in BtuB with the rest of the membrane-facing  $\beta$ -barrel surface, the lack of phenylalanine residues is the most notable difference (Table 5-5). Interestingly, the LPS binding patch in FhuA is located in the same position as patch 2 of BtuB. These patches are rich in phenylalanine which is in contact with the acyl chain region of the LPS molecule (Table 5-3). In addition, the trimer interfaces of OmpC are located in a similar position on the  $\beta$ -barrel surface to patches 1 and 2, which have a propensity to contain phenylalanine (Table 5-11).

Patches also differ in amino acid composition within the same protein. For example, Patch 1 of BtuB has more isoleucine residues than Patch 2, as well as less tryptophan (Table 5-3). It is unclear how significant these differences are given that the patch consists of a limited number of amino acids (14 and 23 residues, respectively). When the Z distances of the contact residues for BtuB were calculated, the expected groupings of polar residues at the boundary of the hydrophobic region of the  $\beta$ -barrel surface were observed (Figure 5-8).

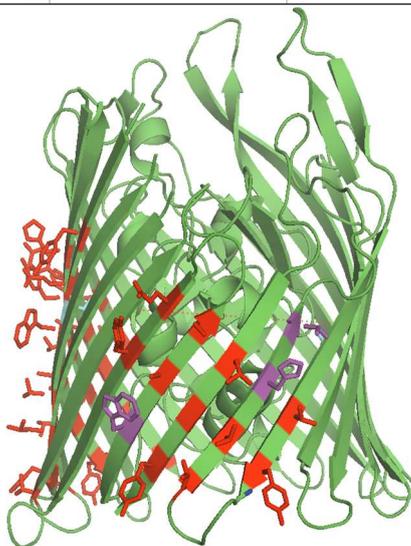
### 5.4.3 Future Work

For future experiments, it would be of interest to identify the most likely residues for amino acid substitution experiments to determine essential PPI patch residues. Other authors have also proven that it is possible to produce eukaryotic membrane proteins in a prokaryotic host, but they form intracellular inclusion bodies (Freigassner, Pichler and Glieder, 2009). To determine if the PPI patch is important for  $\beta$ -barrel assembly, the established protocol for incorporating VDAC into the prokaryotic OM would be used to incorporate VDAC with a modified patch 1. Key residues determined by hydrophobicity and amino acid composition studies would be substituted by site-directed mutagenesis, and assayed for impact on VDAC  $\beta$ -barrel assembly.

**Table 5-14 - Planned Mutations in Patch 1 of BtuB to Mimic Amino Acid Composition of Patch 2**

Heat map showing the change in percent amino acid composition in patch 1, and the difference to patch 2, after planned mutations to BtuB. Green indicates that the patch has a greater proportion of that particular amino acid relative to the full  $\beta$ -barrel surface and red indicates that the patch has less of that particular amino acid. Greater intensity in colour indicates a greater difference.

Amino Acid	Patch 1 Post Mutations	Amino Acid Percentage	Difference from Patch 2	New Difference from Patch 2
GLYCINE	1	7.14	9.94	2.80
ARGININE	0	0.00	0.00	0.00
HISTIDINE	1	7.14	-8.70	-1.55
LYSINE	0	0.00	0.00	0.00
ASPARTIC ACID	0	0.00	0.00	0.00
GLUTAMIC ACID	0	0.00	0.00	0.00
SERINE	0	0.00	0.00	0.00
THREONINE	1	7.14	-8.70	-1.55
ASPARAGINE	0	0.00	-4.35	-4.35
GLUTAMINE	1	7.14	2.80	2.80
CYSTEINE	0	0.00	0.00	0.00
PROLINE	0	0.00	0.00	0.00
ALANINE	1	7.14	2.80	2.80
ISOLEUCINE	0	0.00	14.29	0.00
LEUCINE	2	14.29	1.24	1.24
METHIONINE	0	0.00	0.00	0.00
PHENYLALANINE	0	0.00	0.00	0.00
TRYPTOPHAN	2	14.29	-10.25	-3.11
TRYOSINE	3	21.43	4.04	4.04
VALINE	2	14.29	-3.11	-3.11
<b>SUBTOTAL</b>	14			

**Figure 5-17– BtuB mutations to mimic Patch 2**

3D PyMOL cartoon image of BtuB (3RGM) with interaction patch residues highlighted in red and their side chains displayed as sticks. The planned mutations (I140W, G563H and I565T) to adjust the properties of patch 1 to match patch 2 are visualised in magenta.

Mutations have been designed to modify residues in patch 1 of BtuB to be closer to the amino acid composition of patch 2. These mutations would cause the patch to favour interactions with a lipid molecule over PPIs with other  $\beta$ -barrel membrane proteins.

The planned mutations (I140W, G563H and I563T) target the observed differences in isoleucine, tryptophan, histidine, threonine and glycine composition between the patches (Table 5-14). Two of the mutations are placed in close proximity on the same  $\beta$ -strand (Figure 5-17) to simplify mutagenesis. Proving the impact of these mutations *in vivo* would establish that the composition of amino acids in these membrane-facing  $\beta$ -barrel surface patches is important in determining whether a monomeric OMP engages in protein-lipid interactions or PPIs.

## Chapter 6: Conclusions and Future Work

### 6.1 Progress Towards a Functional Molecular Tension Probe

ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> and R-domain constructs have been successfully over-produced in *E. coli* cells. These proteins were subsequently purified using protocols optimised during this research project. The molecular weight of both purified recombinant proteins were confirmed by ESI-MS analysis. The recombinant proteins remain soluble in storage buffer and are suitable for future MTP assembly.

His103 has been identified as the ‘putative general base’ in the ColE9 DNase mechanism (Pommer *et al.*, 2001). However, the H103A mutation failed to completely inactivate the ColE9 DNase. The published literature has indicated that this mutation would be sufficient to inactivate the DNase (Walker *et al.*, 2002), but DNA degradation and nicking assays of the ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct have detected residual DNase activity. These results suggest that another residue besides His103 may serve as the general base in the mechanism, or that another residue is able to replace this function in the mutated DNase. Effective protocols for assaying ColE9 DNase activity have been established. Results from these activity assays can be used as a benchmark for a future ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct with additional DNase inactivating mutations.

Experiments assessing the N-methylbromomaleimide conjugation strategy have not provided conclusive evidence of conjugation via the maleimide reactive group. The dimerised R-domain present in the samples has made confirmation of coupling via N-MBM conjugation extremely challenging. Ellman’s reagent-based assays provided some evidence of successful N-MBM labelling, though not all of these assays were conclusive in this regard. ESI-MS analysis of N-MBM-R-domain samples was unsuccessful as the samples failed to produce interpretable spectra.

To address the aforementioned challenges, a novel conjugation method using a fluorescent tetrapeptide has been developed for testing thiol coupling in modified proteins without the need for confirmation by ESI-MS. This protocol could see wider use for analysing thiol conjugation in proteins for which ESI-MS analysis cannot be done. The R-domain was successfully conjugated to fluorescent tetrapeptide as

confirmed by the presence of fluorescent protein bands on SDS-PAGE gels. However, whether or not the peptide is coupling via the N-MBM reactive group, or directly at Cys318, is still unknown. Revised conditions for the thiol conjugation protocol have been tested and increases in the following parameters were implemented compared to the literature protocol (Tedaldi *et al.*, 2009): reaction temperature (from 0 °C to 37 °C), molar ratio of label:protein (from 1:1 to 20:1) and reaction incubation time (from 1 hr to 18 hr). This troubleshooting of the thiol conjugation protocol should facilitate future progress towards assembly of a functional MTP.

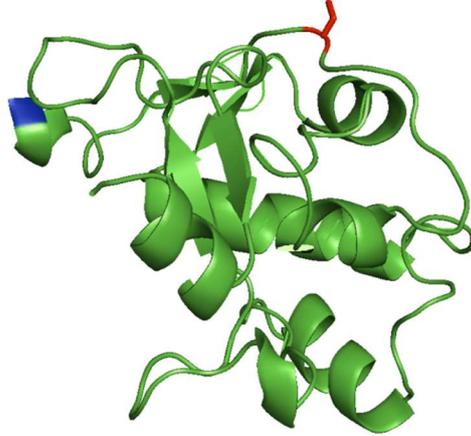
## 6.2 Identification of Putative Intermolecular Interaction Patches in Outer Membrane Proteins

A structural bioinformatics project has discovered evidence of conserved PPI patches in OMPs which are present in phylogenetically-related prokaryotic and eukaryotic membrane protein families. Analysis of these patches has revealed a link between monomeric OMP interaction patches and the trimerisation interfaces of multimeric OMPs, e.g. OmpC. Discovery of these patches validates past MD simulations suggesting a propensity for transient interactions between OMPs (Kleanthous, Rassam and Baumann, 2015; Rassam *et al.*, 2015). However, analysis of intermolecular distances between closely packed OMPs in protein crystals suggests that these interactions are mediated by the acyl chains of membrane lipids, rather than direct van der Waals interactions between OMPs.

## 6.3 Future Work

The His127Ala mutation, identified in the literature as sufficient to inactivate ColE9 DNase (Pommer *et al.*, 2001), could be incorporated into the ColE9 T<sup>1-83</sup>-TEV-H<sup>103A</sup>DNase construct by site-directed mutagenesis of the pRP1 plasmid. ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H<sup>103A</sup></sup> also requires a cysteine modification for thiol-mediated coupling to its respective ssDNA. The site where the Cys residue is engineered must be solvent accessible, distal to potential fluorescence quenching amino acids, and proximal to the N-terminus of the DNase domain to enable efficient labelling. A site for this mutation has been identified by 3D structural modelling: Lys14 in the ColE9 DNase domain (Lys462 in complete colicin E9 sequence). After confirming complete DNase

inactivation in the T<sup>1-83</sup>-TEV-DNase<sup>K14C, H103A, H127A</sup> construct, coupling to its respective ssDNA handle could be attempted.



**Figure 6-1 – Location of Proposed Cysteine Mutation in Colicin E9 DNaseH103A**

The DNase region of T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup> construct visualised in 3D using PyMOL. The DNase domain is coloured in green, the proposed K14C mutation is coloured in red, and the thiol group visualised in stick form. The N-terminal methionine is coloured in blue. The C-terminal end of T<sup>1-83</sup> in the construct joins at the N-terminal methionine. The mutation is designed to be close to the N-terminal methionine, both in sequence and 3D position.

ESI-MS and SDS-PAGE analysis has raised concerns about the stability of the R-domain. Degradation of the R-domain could result in a loss of free thiol sites (318Cys) within the protein sample, for which evidence was provided by Ellman's reagent-based assays. A loss of reduced Cys residues for thiol-mediated coupling would negatively impact attempts at coupling of ssDNA to the R-domain to form one half of the MTP. Prior to future labelling attempts, circular dichroism studies could be used to determine if the R-domain still retains its coiled-coil tertiary structure (Provencher and Glockner, 1981).

## Appendix

### A – Buffers and Reagents

**Table 7-1 Standard Buffers and Reagents used in experiments**

Concentrations are set to 1X; text indicates higher concentrations used where required.

Buffer or Reagent	Components
Agarose Gel Loading Dye 4x	10 mM Tris-HCl, 50 mM EDTA, 0.025 % (w/v) bromophenol blue, 60 % (v/v) glycerol. Adjusted to pH 7.6.
Biotinylated Thrombin	Approx. 1 U/ $\mu$ l in 50 mM sodium citrate (pH 6.5), 200 mM NaCl, 0.1 % (w/v) PEG-8000, 50 % (v/v) glycerol.
Coomassie Blue Stain (SDS/Native Gels)	0.25 g Coomassie Brilliant Blue R250 in 90 ml 1:1 MeOH:water and 10 ml glacial acetic acid. Solution filtered through Whatman No. 1 filter (GE Healthcare, Little Chalfont, UK).
LB Agar (for plates)	Same as LB. 10 % (w/v) Agarose added.
Lysogeny broth (LB, Miller)	10 g 1 % (w/v) tryptone, 5 g yeast extract, 10 g NaCl, made to 1 L using MilliQ H <sub>2</sub> O and adjusted to pH 7.5.
Native Gel Loading Buffer	50 mM BisTris, 6N HCl, 50 mM NaCl, 10% (v/v) Glycerol, 0.001 % (w/v) Ponceau S. Adjusted to pH 7.2.
Native Gel Running Buffer	10x: 1 M Tris/1 M Tricine, made to 1 L using water and adjusted to pH 8.2.
Promega Buffer H	90 mM Tris-HCl (pH 7.5), 10 mM MgCl <sub>2</sub> , 50 mM NaCl.
Quiagen Elution Buffer (EB)	50 mM Tris-HCl (pH 8.1-8.2), 1.4 M NaCl.
R-domain Dialysis/Storage Buffer	50 mM Tris-HCl, 300 mM NaCl, made to 1 L using water and adjusted to pH 7.5.
R-domain Elution Buffer (Buffer B)	20 mM NaH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> , 300 mM NaCl, 500 mM Imidazole, made to 1 L using water and adjusted to pH 7.7. 5mM $\beta$ -Mercaptoethanol is added prior to filtering.

R-domain Lysis Buffer	20 mM NaH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> , 300 mM NaCl, 10 mM Imidazole, made to 1 L using water and adjusted to pH 7.7.
R-domain Wash Buffer (Buffer A)	20 mM NaH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> , 300 mM NaCl, 10 mM Imidazole, made to 1 L using water and adjusted to pH 7.7. 5 mM β-Mercaptoethanol is added prior to filtering
SDS-PAGE Loading Dye	50 mM Tris-HCl (pH 6.8), 100 mM DTT, 2% (w/v) SDS, 0.1 % (w/v) bromophenol blue and 10 % (w/v) glycerol.
SDS-PAGE Running Buffer	25 mM Tris, 264 mM glycine and 0.1 % (w/v) SDS.
T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Buffer	20 mM KH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> , 0.5 M NaCl, made to 1 L using water and adjusted to pH 7.0.
T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Dialysis/Storage Buffer	20 mM KH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> , 0.5 M NaCl, made to 1 L using water and adjusted to pH 7.0.
T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Elution Buffer	20 mM KH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> , 0.5 M NaCl, 6 M Guanidine HCl, made to 1 L using water and adjusted to pH 7.0. <b>NB.</b> When adding Guanidine HCl the reaction is highly endothermic so add gradually onto a hot plate. Start out with a very low volume of water.
T <sup>1-83</sup> -TEV-DNase <sup>H103A</sup> Wash Buffer	20 mM KH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> , 0.5 M NaCl, 5 mM Imidazole, made to 1 L using water and adjusted to pH 7.0.
TAE Agarose Electrophoresis Buffer	20 mM acetic acid, 40 mM Tris, 1 mM EDTA. Adjusted to pH 8.0.
Thermo Scientific GeneRuler DNA Ladder Mix	6X DNA Loading Dye: 10 mM Tris-HCl (pH 7.6), 0.03 % (v/v) bromophenol blue, 0.03% (v/v) xylene cyanol FF, 60 % (v/v) glycerol and 60 mM EDTA.
Thermo Scientific PageRuler Prestained Protein Ladder (10-180kDa)	Dye-stained proteins in 62.5 mM Tris-H <sub>3</sub> PO <sub>4</sub> (pH 7.5 at 25 °C), 1 mM EDTA, 2 % (w/v) SDS, 10 mM DTT, 1 mM NaN <sub>3</sub> and 33 % (v/v) glycerol.

Thrombin Cleavage Buffer	20 mM Tris-HCl (pH 8.4), 150 mM NaCl, 2.5 mM CaCl <sub>2</sub> .
Thrombin Cleavage Size Exclusion Buffer	20 mM KH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> , 0.5 M NaCl, made to 1 L using water and adjusted to pH 7.0.

## B - Protein Construct Coding Sequences

### I. ColE9 T<sup>1-83</sup>-TEV-DNase<sup>H103A</sup>: Im9 His<sub>6</sub>

ATGAGCGGTGGAGATGGACGCGGCCATAACACGGGCGCGCATAGCACAAG  
 TGGTAACATTAATGGTGGCCCGACCGGGATTGGTGTAAAGTGGTGGTGCTTC  
 TGATGGTTCAGGATGGAGTTCGGAAAATAACCCGTGGGGTGGTGGTTCG  
 GTAGCGGCATTCACTGGGGAGGTGGCTCCGGTCGTGGTAATGGCGGGGGT  
 AATGGCAATTCGGTGGTGGCTCGGGAACAGGCGGTAATTTGTCAGCAGA  
 AACCTGTATTTCCAGGGCGCCATGGAGAGTAAACGGAATAAGCCAGGGA  
 AGGCGACAGGTAAAGGTAAACCAGTTGGTGATAAATGGCTGGATGATGCA  
 GGTAAGATTCAGGAGCGCCAATTCAGATCGCATTGCTGATAAGTTGCGT  
 GATAAAGAATTTAAAAGCTTCGACGATTTTCGGAAGGCTGTATGGGAAGA  
 GGTGTGCGAAAGATCCTGAGCTTAGTAAAAATTTAAACCCAAGCAATAAGT  
 CTAGTGTTTCAAAGGTTATTCTCCGTTTACTCCAAGAATCAACAGGTCG  
 GAGGGAGAAAAGTCTATGAACTTCATGCTGACAAGCCAATTAGTCAAGGT  
 GGTGAGGTTTATGACATGGATAATATCCGAGTGACTACACCTAAGCGACAT  
 ATCGATATTCACCGAGGTAAGTAAAATGGAAGTGAAGCATAGCATTAGTG  
 ATTATACAGAAGCTGAATTTTACAGCTTGTAACAACAATTTGTAATGCGG  
 ACACTTCCAGTGAAGAAGAAGTGGTTAAATTGGTTACACACTTTGAGGAAA  
 TGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
 ATGATGACTCACCTTCAGGTATTGTAAACACAGTAAAACAATGGCGAGCC  
 GCTAACGGTAAGTCAGGATTTAAACAGGGCCTCGAGCACCACCACCACCA  
 CCACTGA

### I. R-domain (Y324C, L447C), N-Cys (318C), C-Thrombin Cl.S-His<sub>6</sub>

ATGTGCGTTGAAGCGGCTGAGCGAAATTGTGAACGCGCGCGTGCAGAGCT  
 CAATCAGGCAAATGAAGATGTTGCCAGAAATCAGGAGCGACAGGCTAAAG

CTG TTCAGG TTTATA ATTCGCGT AAAAGCGAACTTGATGCAGCGAATAAAA  
CTCTTGCTGATGCAATAGCTGAAATAAAACAATTTAATCGATTTGCCCATG  
ACCCAATGGCTGGCGGTCACAGAATGTGGCAAATGGCCGGGCTTAAAGCT  
CAGCGGGCGCAGACGGATGTAAATAATAAGCAGGCTGCATTTGATGCTGC  
TGCAAAGAGAAGTCAGATGCTGATGCTGCATTAAGTGCCGCGCAGGAGC  
GCCGCAAACAGAAGGAAAATAAAGAAAAGGACGCTAAGGATAAATGCGA  
TAAGGAGAGTAAACGGAATCTGGTGCCGCGCGGCAGCCTCGAGCACCACC  
ACCACCACCACTGA

**List of Abbreviations**

A <sub>280</sub>	-	Absorbance measurements at a wavelength of 280 nm
BSA	-	Bovine serum albumin
C	-	Colicin E9 cytotoxic domain
ColE2	-	Colicin E2
ColE3	-	Colicin E3
ColE9	-	Colicin E9
DMF	-	Dimethylformamide
DNA	-	Deoxyribonucleic acid
DTNB	-	5, 5'-dithio-bis-2-nitrobenzoic acid
DTT	-	Dithiothreitol
<i>E. coli</i>	-	<i>Escherichia coli</i>
EDTA	-	Ethylenediaminetetraacetic
ESI-MS	-	Electrospray ionisation mass spectrometry
FPLC	-	Fast protein liquid chromatography
FRET	-	Förster resonance energy transfer
hVDAC	-	Human voltage-dependent anion channel
IM	-	Inner membrane
Im9	-	Colicin E9 Immunity Protein
IPTG	-	Isopropyl β-D-1-thiogalactopyranoside
ITC	-	Isothermal titration calorimetry
IUTD	-	Intrinsically unstructured translocation domain
LB	-	Lysogeny broth
LC-MS	-	Liquid chromatography–mass spectrometry
LPS	-	Lipopolysaccharide
MALDI MS	-	Matrix-assisted laser desorption/ionization mass spectrometry
MD	-	Molecular dynamics
MTP	-	Molecular tension probe
MWCO	-	Molecular weight cut-off
N-MBM	-	N-Methyl-bromomaleimide
OD	-	Optical density
OD <sub>280</sub>	-	Optical density at 280 nm
OD <sub>600</sub>	-	Optical density at 600nm

OM	-	Outer membrane
OMP	-	Outer membrane protein
PAGE	-	Poly-acrylamide gel electrophoresis
PD	-	Pre-packed disposable
PDB	-	Protein data bank
Pmf	-	Proton motive force
PMSF	-	Phenylmethylsulfonyl fluoride
PPI	-	Promiscuous protein-protein interactions
R	-	Colicin E9 receptor-binding domain
RMSD	-	Root-mean-square deviation
RNA	-	Ribonucleic acid
SDS	-	Sodium dodecyl sulfate
SDS-PAGE	-	Sodium Dodecyl Sulphate Poly-Acrylamide Gel Electrophoresis
ssDNA	-	Single-stranded deoxyribonucleic acid
T	-	Colicin E9 translocation domain
TAE	-	Tris-acetate-EDTA
TBE	-	TolB-binding epitope
TCEP	-	Tris(2-carboxyethyl)phosphine
UPLC	-	Ultra performance liquid chromatography
UV	-	Ultra-violet radiation
$\beta$ -ME	-	Betamercaptoethanol

## References

- Alvarez, S. (2013) 'A cartography of the van der Waals territories', *Dalton Transactions*, 42(24), pp. 8617–8636. doi: 10.1039/c3dt50599e.
- Arunmanee, W., Pathania, M., Solovyova, A. S., Le Brun, A. P., Ridley, H., Baslé, A., van den Berg, B. and Lakey, J. H. (2016) 'Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 113(34), pp. E5034–43. doi: 10.1073/pnas.1602382113.
- Bavoil, P., Nikaido, H. and von Meyenburg, K. (1977) 'Pleiotropic transport mutants of Escherichia coli lack porin, a major outer membrane protein', *Molecular & General Genetics*, 158(1), pp. 23–33. doi: 10.1007/BF00455116.
- Bednar, R. (1990) 'Reactivity and pH dependence of thiol conjugation to N-ethylmaleimide: detection of a conformational change in chalcone isomerase', *Biochemistry*, 29(15), pp. 3684–3690. doi: 10.1021/bi00467a014.
- Blakely, B., Dumelin, C., Trappmann, B., McGregor, L., Choi, C., Anthony, P., Duesterberg, V., Baker, B., Block, S., Liu, D. and Chen, C. (2014) 'A DNA-based molecular probe for optically reporting cellular traction forces', *Nature Methods*, 11(12), pp. 1229–32. doi: 10.1038/nmeth.3145.
- Bocharov, E., Lesovoy, D., Pavlov, K., Pustovalova, Y., Bocharova, O. and Arseniev, A. (2016) 'Alternative packing of EGFR transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane', *Biochimica et Biophysica Acta - Biomembranes*. Elsevier B.V., 1858(6), pp. 1254–1261. doi: 10.1016/j.bbamem.2016.02.023.
- Bockelmann, U., Thomen, P., Essevaz-Roulet, B., Viasnoff, V. and Heslot, F. (2002) 'Unzipping DNA with optical tweezers: high sequence sensitivity and force flips', *Biophysical journal*. Elsevier, 82(3), pp. 1537–1553. doi: 10.1016/S0006-3495(02)75506-9.
- Bondi, A. (1964) 'van der Waals Volumes and Radii', *The Journal of Physical Chemistry*, 68(3), pp. 441–451. doi: 10.1021/j100785a001.

- Bonsor, D., Grishkovskaya, I., Dodson, E. and Kleanthous, C. (2007) 'Molecular mimicry enables competitive recruitment by a natively disordered protein', *Journal of the American Chemical Society*, 129(15), pp. 4800–4807. doi: 10.1021/ja070153n.
- Bonsor, D., Meenan, N. and Kleanthous, C. (2008) 'Colicins exploit native disorder to gain cell entry: a hitchhiker's guide to translocation', *Biochemical Society Transactions*, 36(1), pp. 1409–1413. doi: 10.1042/BST0361409.
- Bramkamp, M. and Lopez, D. (2015) 'Exploring the existence of lipid rafts in bacteria', *Microbiology and Molecular Biology Reviews*, 79(1), pp. 81–100. doi: 10.1128/MMBR.00036-14.
- Brown, D. (2016) 'Drug discovery strategies to outer membrane targets in Gram-negative pathogens', *Bioorganic and Medicinal Chemistry*, 24(24), pp. 6320–6331. doi: 10.1016/j.bmc.2016.05.004.
- Cadieux, N. and Kadner, R. (1999) 'Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter', *Proceedings of the National Academy of Sciences of the United States of America*, 96(19), pp. 10673–10678. doi: 10.1073/pnas.96.19.10673.
- Carr, S., Penfold, C., Bamford, V., James, R. and Hemmings, A. (2000) 'The structure of TolB, an essential component of the tol-dependent translocation system, and its protein-protein interaction with the translocation domain of colicin E9', *Structure*, 8(1), pp. 57–66. doi: 10.1016/S0969-2126(00)00079-4.
- Cascales, E., Buchanan, S., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S. and Cavard, D. (2007) 'Colicin Biology', *Microbiology and Molecular Biology Reviews*, 71(1), pp. 158–229. doi: 10.1128/MMBR.00036-06.
- Cascales, E. and Lloubès, R. (2004) 'Deletion analyses of the peptidoglycan-associated lipoprotein Pal reveals three independent binding sequences including a TolA box', *Molecular Microbiology*, 51(3), pp. 873–885. doi: 10.1046/j.1365-2958.2003.03881.x.
- Cascales, E., Lloubès, R. and Sturgis, J. (2001) 'The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB', *Molecular Microbiology*, 42(3), pp. 795–807. doi: 10.1046/j.1365-2958.2001.02673.x.

Cecconi, C., Shank, E., Dahlquist, F., Marqusee, S. and Bustamante, C. (2008) 'Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers', *European Biophysics Journal*, 37(6), pp. 729–738. doi: 10.1007/s00249-007-0247-y.

Chabanon, E., Mangin, D. and Charcosset, C. (2016) 'Membranes and crystallization processes: State of the art and prospects', *Journal of Membrane Science*, 509(1), pp. 57–67. doi: 10.1016/j.memsci.2016.02.051.

Cherezov, V., Yamashita, E., Liu, W., Zhahnina, M., Cramer, W. and Caffrey, M. (2006) 'In Meso Structure of the Cobalamin Transporter, BtuB, at 1.95 Å Resolution', *Journal of Molecular Biology*, 364(4), pp. 716–734. doi: 10.1016/j.jmb.2006.09.022.

Chimento, D., Kadner, R. and Wiener, M. (2003) 'The Escherichia coli outer membrane cobalamin transporter BtuB: Structural analysis of calcium and substrate binding, and identification of orthologous transporters by sequence/structure conservation', *Journal of Molecular Biology*, 332(5), pp. 999–1014. doi: 10.1016/j.jmb.2003.07.005.

Cocco, S., Monasson, R. and Marko, J. (2002) 'Force and kinetic barriers to initiation of DNA unzipping of the DNA double helix', *Proceedings of the National Academy of Sciences*, 65(4), p. 23. doi: 10.1103/PhysRevE.65.041907.

Cowan, S., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R., Jansonius, J. and Rosenbusch, J. (1992) 'Crystal structures explain functional properties of two E. coli porins', *Nature*, 358(6389), pp. 727–733. doi: 10.1038/358727a0.

Doura, A. and Fleming, K. (2004) 'Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer', *Journal of Molecular Biology*, 343(5), pp. 1487–1497. doi: 10.1016/j.jmb.2004.09.011.

Duche, D. (2007) 'Colicin E2 is still in contact with its receptor and import machinery when its nuclease domain enters the cytoplasm', *Journal of Bacteriology*, 189(11), pp. 4217–4222. doi: 10.1128/JB.00092-07.

Dunton, T., Goose, J., Gavaghan, D., Sansom, M. and Osborne, J. (2014) 'The Free Energy Landscape of Dimerization of a Membrane Protein, NanC', *Public Library of*

- Science Computational Biology*, 10(1). doi: 10.1371/journal.pcbi.1003417.
- Dupuy, A. and Engelman, D. (2008) 'Protein area occupancy at the center of the red blood cell membrane', *Proceedings of the National Academy of Sciences*, 105(8), pp. 2848–52. doi: 10.1073/pnas.0712379105.
- Efremov, R. and Sazanov, L. (2012) 'Structure of Escherichia coli OmpF porin from lipidic mesophase', *Journal of Structural Biology*. Elsevier Inc., 178(3), pp. 311–318. doi: 10.1016/j.jsb.2012.03.005.
- Elofsson, A. and von Heijne, G. (2007) 'Membrane protein structure: prediction versus reality', *Annual review of biochemistry*, 76(1), pp. 125–140. doi: 10.1146/annurev.biochem.76.052705.163539.
- Emsley, P., Lohkamp, B., Scott, W. and Cowtan, K. (2010) 'Features and development of Coot', *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 66(4), pp. 486–501. doi: 10.1107/S0907444910007493.
- Farrance, O., Hann, E., Kaminska, R., Housden, N., Derrington, S., Kleanthous, C., Radford, S. and Brockwell, D. (2013) 'A Force-Activated Trip Switch Triggers Rapid Dissociation of a Colicin from Its Immunity Protein', *Public Library of Science Biology*, 11(2), p. e1001489. doi: 10.1371/journal.pbio.1001489.
- Ferguson, A. and Deisenhofer, J. (2002) 'TonB-dependent receptors - Structural perspectives', *Biochimica et Biophysica Acta - Biomembranes*, 1565(2), pp. 318–332. doi: 10.1016/S0005-2736(02)00578-3.
- Freigassner, M., Pichler, H. and Glieder, A. (2009) 'Tuning microbial hosts for membrane protein production', *Microbial Cell Factories*, 8(1), p. 69. doi: 10.1186/1475-2859-8-69.
- Garavito, M., Jenkins, J., Jansonius, J., Karlsson, R. and Rosenbusch, J. (1983) 'X-ray diffraction analysis of matrix porin, an integral membrane protein from Escherichia coli outer membranes', *Journal of Molecular Biology*, 164(2), pp. 313–327. doi: 10.1016/0022-2836(83)90079-7.
- Garinot-Schneider, C., Pommer, A., Moore, G., Kleanthous, C. and James, R. (1996) 'Identification of Putative Active-site Residues in the DNase Domain of Colicin E9 by

- Random Mutagenesis', *Journal of Molecular Biology*, 260(5), pp. 731–742. doi: 10.1006/jmbi.1996.0433.
- Geula, S., Naveeds, H., Liangs, J. and Shoshan-Barmatz, V. (2012) 'Structure-based analysis of VDAC1 protein: Defining oligomer contact sites', *Journal of Biological Chemistry*, 287(3), pp. 2179–2190. doi: 10.1074/jbc.M111.268920.
- Gordon, D. and Riley, M. (1999) 'A theoretical and empirical investigation of the invasion dynamics of colicinogeny', *Microbiology*, 145(3), pp. 655–661. doi: 10.1099/13500872-145-3-655.
- Gram, H. (1884) 'Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten', *Fortschritte der Medizin*, 2, pp. 185–189.
- Gu, Y., Stansfeld, P., Zeng, Y., Dong, H., Wang, W. and Dong, C. (2015) 'Lipopolysaccharide is inserted into the outer membrane through an intramembrane hole, a lumen gate, and the lateral opening of lptd', *Structure*. The Authors, 23(3), pp. 496–504. doi: 10.1016/j.str.2015.01.001.
- Häckel, M., Hinz, H.-J. and Hedwig, G. (1999) 'Partial molar volumes of proteins: amino acid side-chain contributions derived from the partial molar volumes of some tripeptides over the temperature range 10-90 degrees C.', *Biophysical Chemistry*, 82(1), pp. 35–50. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17030339&retmode=ref&cmd=prlinks%5Cnpapers3://publication/uuid/0489FCEB-2D35-4B91-9646-61E0D8C6C023>.
- Hanahan, D., Jessee, J. and Bloom, F. (1991) 'Plasmid Transformation of Escherichia coli and Other Bacteria', *Methods in Enzymology*, 204(1970), pp. 63–113. doi: 10.1016/0076-6879(91)04006-A.
- Harpaz, Y., Gerstein, M. and Chothia, C. (1994) 'Volume changes on protein folding', *Structure*, 2(7), pp. 641–649. doi: 10.1016/S0969-2126(00)00065-4.
- Harroun, T., Heller, W., Weiss, T., Yang, L. and Huang, H. (1999) 'Theoretical Analysis of Hydrophobic Matching and Membrane-Mediated Interactions in Lipid Bilayers Containing Gramicidin', *Biophysical journal*, 76(6), pp. 3176–3185. doi:

10.1016/S0006-3495(99)77469-2.

Hickman, S., Cooper, R., Bellucci, L., Paci, E. and Brockwell, D. (2017) 'Gating of TonB-dependent transporters by substrate-specific forced remodelling', *Nature Communications*, 8(14804), p. 14804. doi: 10.1038/ncomms14804.

Höhr, A. I. C., Straub, S. P., Warscheid, B., Becker, T. and Wiedemann, N. (2015) 'Assembly of ??-barrel proteins in the mitochondrial outer membrane', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., 1853(1), pp. 74–88. doi: 10.1016/j.bbamcr.2014.10.006.

Housden, N., Hopper, J., Lukoyanova, N., Rodriguez-Larrea, D., Wojdyla, J., Klein, A., Kaminska, R., Bayley, H., Saibil, H., Robinson, C. and Kleanthous, C. (2013) 'Intrinsically Disordered Protein Threads Through the Bacterial Outer-Membrane Porin OmpF', *Science*, 340(6140), pp. 1570–1574. doi: 10.1126/science.1237864.

Housden, N. and Kleanthous, C. (2012) 'Colicin translocation across the Escherichia coli outer membrane', *Biochemical Society Transactions*, 40(6), pp. 1475–9. doi: 10.1042/BST20120255.

Housden, N., Loftus, S., Moore, G., James, R. and Kleanthous, C. (2005) 'Cell entry mechanism of enzymatic bacterial colicins: porin recruitment and the thermodynamics of receptor binding', *Proceedings of the National Academy of Sciences*, 102(39), pp. 13849–13854. doi: 10.1073/pnas.0503567102.

Housden, N., Wojdyla, J., Korczynska, J., Grishkovskaya, I., Kirkpatrick, N., Brzozowski, M. and Kleanthous, C. (2010) 'Directed epitope delivery across the Escherichia coli outer membrane through the porin OmpF', *Proceedings of the National Academy of Sciences*, 107(50), pp. 21412–21417. doi: 10.1073/pnas.1010780107.

Isabella, V., Campbell, A., Manchester, J., Sylvester, M., Nayar, A., Ferguson, K., Tommasi, R. and Miller, A. (2015) 'Toward the rational design of carbapenem uptake in pseudomonas aeruginosa', *Chemistry and Biology*. Elsevier Ltd, 22(4), pp. 535–547. doi: 10.1016/j.chembiol.2015.03.018.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T.

(2008) 'NCBI BLAST: a better web interface', *Nucleic acids research*, 36(Web Server issue), pp. W5–W9. doi: 10.1093/nar/gkn201.

Jurchenko, C. and Salaita, K. (2015) 'Lighting Up the Force: Investigating Mechanisms of Mechanotransduction Using Fluorescent Tension Probes', *Molecular and Cellular Biology*, 35(15), pp. 2570–82. doi: 10.1128/MCB.00195-15.

Kent, C. (1995) 'Eukaryotic phospholipid biosynthesis', *Annual Review of Biochemistry*, 64(1), pp. 315–343. doi: 10.1146/annurev.bi.64.070195.001531.

Kim, Y., Tarr, A. and Penfold, C. (2014) 'Colicin import into E. coli cells: A model system for insights into the import mechanisms of bacteriocins', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., 1843(8), pp. 1717–1731. doi: 10.1016/j.bbamcr.2014.04.010.

Kleanthous, C. (2010) 'Swimming against the tide: progress and challenges in our understanding of colicin translocation.', *Nature reviews. Microbiology*. Nature Publishing Group, 8(12), pp. 843–848. doi: 10.1038/nrmicro2454.

Kleanthous, C., Rassam, P. and Baumann, C. (2015) 'Protein-protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins', *Current Opinion in Structural Biology*, 35, pp. 109–115. doi: 10.1016/j.sbi.2015.10.007.

Klein, A., Wojdyla, J., Joshi, A., Josts, I., Mccaughey, L., Housden, N., Kaminska, R., Byron, O., Walker, D. and Kleanthous, C. (2016) 'Structural and biophysical analysis of nuclease protein antibiotics', *Biochemical Journal*, 473(1), pp. 2799–2812. doi: 10.1042/BCJ20160544.

Koldso, H. and Sansom, M. (2015) 'Organization and Dynamics of Receptor Proteins in a Plasma Membrane', *Journal of the American Chemical Society*, 137(46), pp. 14694–14704. doi: 10.1021/jacs.5b08048.

Konisky, J. (1982) 'Colicins and Other Bacteriocins with Established Modes of Action', *Annual review of microbiology*, 36(1), pp. 125–144. doi: 10.1146/annurev.mi.36.100182.001013.

Kumarn, S., Chimnoi, N. and Ruchirawat, S. (2013) 'Synthesis of integerrimide A by an on-resin tandem Fmoc-deprotection-macrocyclisation approach', *Organic &*

*biomolecular chemistry*, 11(44), pp. 7760–7. doi: 10.1039/c3ob41447g.

Kyte, J. and Doolittle, R. (1982) ‘A simple method for displaying the hydropathic character of a protein’, *Journal of Molecular Biology*, 157(1), pp. 105–132. doi: 10.1016/0022-2836(82)90515-0.

LaRocca, T., Pathak, P., Chiantia, S., Toledo, A., Silvius, J., Benach, J. and London, E. (2013) ‘Proving Lipid Rafts Exist: Membrane Domains in the Prokaryote *Borrelia burgdorferi* Have the Same Properties as Eukaryotic Lipid Rafts’, *Public Library of Science Pathogens*, 9(5). doi: 10.1371/journal.ppat.1003353.

Lazdunski, C., Bouveret, E., Rigal, A., Journet, L., Llobès, R., Bénédicti, H. and Lloubé, R. (1998) ‘Colicin Import into *Escherichia coli* Cells’, *Journal of Bacteriology*, 180(19), pp. 4993–5002. doi: <https://www.ncbi.nlm.nih.gov/pubmed/9748429>.

Llobès, R., Goemaere, E., Zhang, X., Cascales, E. and Duché, D. (2012) ‘Energetics of colicin import revealed by genetic cross-complementation between the Tol and Ton systems’, *Biochemical Society Transactions*, 40(6), pp. 1480–1485. doi: 10.1042/BST20120181.

Loftus, S., Walker, D., Maté, M. J., Bonsor, D., James, R., Moore, G. and Kleanthous, C. (2006) ‘Competitive recruitment of the periplasmic translocation portal TolB by a natively disordered domain of colicin E9’, *Proceedings of the National Academy of Sciences of the United States of America*, 103(33), pp. 12353–8. doi: 10.1073/pnas.0603433103.

López, D. and Kolter, R. (2010) ‘Functional microdomains in bacterial membranes’, *Genes and Development*, 24(17), pp. 1893–1902. doi: 10.1101/gad.1945010.

Mahendran, K., Mach, T., Bessonov, A., Tomberli, B., Nickel, J. and Berry, M. (2009) ‘The Biophysics Of Antibiotics Translocation Through OmpF Revealed By Computer Simulations’, *Biophysical Journal*, 96(3), p. 41a. doi: 10.1016/j.bpj.2008.12.105.

Maté, M. and Kleanthous, C. (2004) ‘Structure-based analysis of the metal-dependent mechanism of H-N-H endonucleases’, *Journal of Biological Chemistry*, 279(33), pp. 34763–34769. doi: 10.1074/jbc.M403719200.

- Matthews, B. (1968) 'Solvent content of protein crystals', *Journal of Molecular Biology*, 33(2), pp. 491–497. doi: 10.1016/0022-2836(68)90205-2.
- McGinnis, S. and Madden, T. L. (2004) 'BLAST: At the core of a powerful and diverse set of sequence analysis tools', *Nucleic Acids Research*, 32(WEB SERVER ISS.), pp. 20–25. doi: 10.1093/nar/gkh435.
- Mosbahi, K., Lemaître, C., Keeble, A., Mobasheri, H., Morel, B., James, R., Moore, G., Lea, E. and Kleanthous, C. (2002) 'The cytotoxic domain of colicin E9 is a channel-forming endonuclease', *Nature structural biology*, 9(6), pp. 476–484. doi: 10.1038/nsb797.
- Nestorovich, E., Danelon, C., Winterhalter, M. and Bezrukov, S. (2002) 'Designed to penetrate: time-resolved interaction of single antibiotic molecules with bacterial pores', *Proceedings of the National Academy of Sciences*, 99(15), pp. 9789–9794. doi: 10.1073/pnas.152206799.
- Nikaido, H. (2003) 'Molecular basis of bacterial outer membrane permeability revisited', *Microbiol Mol.Biol.Rev.*, 67(4), pp. 593–656. Available at: [papers3://publication/uuid/0D579192-8F26-4DE3-A218-E6A61D785035](https://pubmed.ncbi.nlm.nih.gov/152206799/).
- Nikaido, H. (2003) 'Molecular basis of bacterial outer membrane permeability revisited', *Microbiology and Molecular Biology Reviews*, 67(4), pp. 593–656. doi: 10.1128/MMBR.67.4.593.
- Noinaj, N., Guillier, M., Barnard, T. and Buchanan, S. (2010) 'TonB-Dependent Transporters: Regulation, Structure, and Function', *Annual Review of Microbiology*, 64(1), pp. 43–60. doi: 10.1146/annurev.micro.112408.134247.
- Ortiz-Suarez, M., Samsudin, F., Piggot, T., Bond, P. and Khalid, S. (2016) 'Full-Length OmpA: Structure, Function, and Membrane Interactions Predicted by Molecular Dynamics Simulations', *Biophysical Journal*. Biophysical Society, 111(8), pp. 1692–1702. doi: 10.1016/j.bpj.2016.09.009.
- Papadakos, G., Wojdyla, J. and Kleanthous, C. (2012) *Nuclease colicins and their immunity proteins*, *Quarterly Reviews of Biophysics*. doi: 10.1017/S0033583511000114.

- Parton, D., Klingelhoefer, J. and Sansom, M. (2011) 'Aggregation of model membrane proteins, modulated by hydrophobic mismatch, membrane curvature, and protein class', *Biophysical Journal*. Biophysical Society, 101(3), pp. 691–699. doi: 10.1016/j.bpj.2011.06.048.
- Patel, D., Qi, Y. and Im, W. (2017) 'Modeling and simulation of bacterial outer membranes and interactions with membrane proteins', *Current Opinion in Structural Biology*. Elsevier Ltd, 43, pp. 131–140. doi: 10.1016/j.sbi.2017.01.003.
- Periole, X., Huber, T., Marrink, S. and Sakmar, T. (2007) 'G protein-coupled receptors self-assemble in dynamics simulations of model bilayers', *Journal of the American Chemical Society*, 129(33), pp. 10126–10132. doi: 10.1021/ja0706246.
- Perkins, D., Pappin, D., Creasy, D. and Cottrell, J. (1999) 'Probability-based protein identification by searching sequence databases using mass spectrometry data', *Electrophoresis*, 20(18), pp. 3551–67. doi: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.
- Pommer, A., Cal, S., Keeble, A., Walker, D., Evans, S., Kühlmann, U., Cooper, A., Connolly, B., Hemmings, A., Moore, G. R., James, R. and Kleanthous, C. (2001) 'Mechanism and cleavage specificity of the H-N-H endonuclease colicin E9', *Journal of Molecular Biology*, 314(4), pp. 735–749. doi: 10.1006/jmbi.2001.5189.
- Pommer, A., Wallis, R., Moore, G., James, R. and Kleanthous, C. (1998) 'Enzymological characterization of the nuclease domain from the bacterial toxin colicin E9 from *Escherichia coli*', *Biochemical journal*, 334(1), pp. 387–392. Available at: <http://www.biochemj.org/bj/334/0387/bj3340387.htm>.
- Pommier, S., Gavioli, M., Cascales, E. and Llobes, R. (2005) 'Tol-Dependent Macromolecule Import through the *Escherichia coli* Cell Envelope Requires the Presence of an Exposed TolA Binding Motif', *Journal of Bacteriology*, 187(21), pp. 7526–7534. doi: 10.1128/JB.187.21.7526.
- Postle, K. and Kadner, R. (2003) 'Touch and go: Tying TonB to transport', *Molecular Microbiology*, 49(4), pp. 869–882. doi: 10.1046/j.1365-2958.2003.03629.x.
- Provencher, S. and Glockner, J. (1981) 'Estimation of globular protein secondary

- structure from circular dichroism', *Biochemistry*, 20(1), pp. 33–37. doi: 10.1021/bi00504a006.
- Rassam, P., Copeland, N., Birkholz, O., Tóth, C., Chavent, M., Duncan, A., Cross, S., Housden, N., Kaminska, R., Seger, U., Quinn, D., Garrod, T., Sansom, M., Piehler, J., Baumann, C. and Kleanthous, C. (2015) 'Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria', *Nature*, 523(1), pp. 333–336. doi: 10.1038/nature14461.
- Riener, C., Kada, G. and Gruber, H. (2002) 'Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4'-dithiodipyridine', *Analytical and Bioanalytical Chemistry*, 373(4–5), pp. 266–276. doi: 10.1007/s00216-002-1347-2.
- Robinson, R. (2013) 'A Small Force Remodels a Large Protein to Unleash Its Deadly Potential', *Public Library of Science Biology*, 11(2), pp. 1–2. doi: 10.1371/journal.pbio.1001486.
- Schmidt, U., Guigas, G. and Weiss, M. (2008) 'Cluster formation of transmembrane proteins due to hydrophobic mismatching', *Physical Review Letters*, 101(12), pp. 1–4. doi: 10.1103/PhysRevLett.101.128104.
- Schrödinger, L. (2013) 'The PyMOL Molecular Graphics System, Version 1.6.x', *DeLano Scientific LLC*.
- Sikora, A. and Cafiso, D. (2016) 'TonB Binding Produces Allosteric Changes in the Outer Loops and Substrate Binding Site of the TBDT BTUB', *Biophysical journal*, 110(3), p. 514a. doi: 10.1016/j.bpj.2015.11.2748.
- Silhavy, T., Kahne, D. and Walker, S. (2010) 'The bacterial cell envelope', *Cold Spring Harbor perspectives in biology*, 2(5), pp. 1–16. doi: 10.1101/cshperspect.a000414.
- Simons, K. and Toomre, D. (2000) 'Lipid rafts and signal transduction', *Nature reviews. Molecular cell biology*, 1(1), pp. 31–39. doi: 10.1038/35036052.
- Smith, M., Schumacher, F., Ryan, C., Tedaldi, L., Papaioannou, D., Waksman, G., Caddick, S. and Baker, J. (2010) 'Protein modification, bioconjugation, and disulfide bridging using bromomaleimides', *Journal of the American Chemical Society*, 132(6),

pp. 1960–1965. doi: 10.1021/ja908610s.

Spector, J., Zakharov, S., Lill, Y., Sharma, O., Cramer, W. and Ritchie, K. (2010) ‘Mobility of BtuB and OmpF in the Escherichia coli outer membrane: Implications for dynamic formation of a translocon complex’, *Biophysical Journal*. Biophysical Society, 99(12), pp. 3880–3886. doi: 10.1016/j.bpj.2010.10.029.

Steeghs, L., den Hartog, R., den Boer, A., Zomer, B., Roholl, P. and van der Ley, P. (1998) ‘Meningitis bacterium is viable without endotoxin’, *Nature*, 392(6675), pp. 449–450. doi: 10.1038/33046.

Swinehart, D. (1962) ‘The Beer-Lambert Law’, *Journal of Chemical Education*, 39(7), p. 333. doi: 10.1021/ed039p333.

Tedaldi, L., Smith, M., Nathani, R. and Baker, J. (2009) ‘Bromomaleimides: new reagents for the selective and reversible modification of cysteine’, *Chemical Communications*, (43), pp. 6583–6855. doi: 10.1039/b915136b.

Vankemmelbeke, M., Housden, N., James, R., Kleanthous, C. and Penfold, C. (2013) ‘Immunity protein release from a cell-bound nuclease colicin complex requires global conformational rearrangement’, *MicrobiologyOpen*, 2(5), pp. 853–861. doi: 10.1002/mbo3.122.

Vankemmelbeke, M., Zhang, Y., Moore, G., Kleanthous, C., Penfold, C. and James, R. (2009) ‘Energy-dependent immunity protein release during tol-dependent nuclease colicin translocation’, *Journal of Biological Chemistry*, 284(28), pp. 18932–18941. doi: 10.1074/jbc.M806149200.

Walker, D., Georgiou, T., Pommer, A., Walker, D., Moore, G., Kleanthous, C. and James, R. (2002) ‘Mutagenic scan of the H-N-H motif of colicin E9: implications for the mechanistic enzymology of colicins, homing enzymes and apoptotic endonucleases’, *Nucleic acids research*, 30(14), pp. 3225–3234. doi: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC135741/>.

Walker, D., Mosbahi, K., Vankemmelbeke, M., James, R. and Kleanthous, C. (2007) ‘The role of electrostatics in colicin nuclease domain translocation into bacterial cells’, *Journal of Biological Chemistry*, 282(43), pp. 31389–31397. doi:

10.1074/jbc.M705883200.

Wallis, R., Moore, G. R., James, R. and Kleanthous, C. (1995) 'Diffusion-Controlled Association and Femtomolar Binding for the Cognate Complex?', pp. 13743–13750.

Walther, D., Bos, M., Rapaport, D. and Tommassen, J. (2010) 'The mitochondrial porin, VDAC, has retained the ability to be assembled in the bacterial outer membrane', *Molecular Biology and Evolution*, 27(4), pp. 887–895. doi: 10.1093/molbev/msp294.

Wang, X. and Ha, T. (2013) 'Defining single molecular forces required to activate integrin and notch signaling', *Science*, 340(6135), pp. 991–994. doi: 10.1126/science.1231041.

White, S. and Wimley, W. (1998) 'Hydrophobic interactions of peptides with membrane interfaces', *Biochimica et Biophysica Acta - Reviews on Biomembranes*, 1376(3), pp. 339–352. doi: 10.1016/S0304-4157(98)00021-5.

Wojdyla, J., Cutts, E., Kaminska, R., Papadakos, G., Hopper, J., Stansfeld, P., Staunton, D., Robinson, C. and Kleanthous, C. (2015) 'Structure and function of the Escherichia coli Tol-Pal stator protein TolR', *Journal of Biological Chemistry*, 290(44), pp. 26675–26687. doi: 10.1074/jbc.M115.671586.

Yamashita, E., Zhalnina, M., Zakharov, S., Sharma, O. and Cramer, W. (2008) 'Crystal structures of the OmpF porin: function in a colicin translocon', *The European Molecular Biology Organization Journal*, 27(15), pp. 2171–80. doi: 10.1038/emboj.2008.137.

Zakharov, S., Sharma, O., Zhalnina, M. and Cramer, W. (2008) 'Primary events in the colicin translocon: FRET analysis of colicin unfolding initiated by binding to BtuB and OmpF', *Biochemistry*, 47(48), pp. 12802–12809. doi: 10.1021/bi800865h.

Zakharov, S., Sharma, O., Zhalnina, M., Yamashita, E. and Cramer, W. (2012) 'Pathways of colicin import: utilization of BtuB, OmpF porin and the TolC drug-export protein', *Biochemical Society transactions*, 40(6), pp. 1463–1468. doi: 10.1042/BST20120211.

Zakharov, S., Zhalnina, M., Sharma, O. and Cramer, W. (2006) 'The colicin E3 outer

membrane translocon: Immunity protein release allows interaction of the cytotoxic domain with OmpF porin', *Biochemistry*, 45(34), pp. 10199–10207. doi: 10.1021/bi060694+.

Ziervogel, B. and Roux, B. (2013) 'The binding of antibiotics in OmpF porin', *Structure*, 21(1), pp. 76–87. doi: 10.1016/j.str.2012.10.014.