

**Dissecting protein-protein interactions facilitating  
secretion of the virulence factor EspC by  
enteropathogenic *Escherichia coli*.**

**Keith Bishop BSc (Hons).**

**Thesis submitted to The University of Nottingham for the degree of  
Master of Philosophy, June 2009.**

**Supervisors: Dr R. Delahay and Dr K. Hardie.**

# CONTENTS

Contents	i
Index of tables	v
Index of figures	vi
Acknowledgements	ix
Abstract	xi
<b>1      Introduction</b>	<b>1</b>
<b>1.1     Enteropathogenic <i>Escherichia coli</i></b>	<b>1</b>
<b>1.1.1    Attaching and Effacing (A/E) lesions</b>	<b>1</b>
<b>1.1.2    Genes encoding secreted proteins</b>	<b>2</b>
<b>1.1.3    EPEC proteins secreted by the T3SS</b>	<b>2</b>
<b>1.1.3.1    The translocator proteins</b>	<b>2</b>
<b>1.1.3.2    The translocated LEE effector proteins</b>	<b>3</b>
<b>1.1.3.3    The translocated non-LEE effector proteins</b>	<b>3</b>
<b>1.1.4    Function of other virulence factors</b>	<b>4</b>
<b>1.1.5    EspC</b>	<b>4</b>
<b>1.2     Overview of Gram-negative bacterial protein secretion</b>	<b>6</b>
<b>1.2.1    Introduction</b>	<b>6</b>
<b>1.2.2    Inner membrane (IM) mediated translocation</b>	<b>7</b>
<b>1.2.2.1    SecB-dependent secretion</b>	<b>7</b>
<b>1.2.2.2    Signal Recognition Pathway</b>	<b>10</b>
<b>1.2.2.3    Twin arginine translocation (Tat) pathway</b>	<b>10</b>
<b>1.2.2.4    YidC</b>	<b>11</b>
<b>1.2.3    The type I secretion system (T1SS) – ABC transporters</b>	<b>12</b>
<b>1.2.4    The type II secretion system (T2SS) – The secreton dependent pathway (Main Terminal Branch)</b>	<b>14</b>
<b>1.2.4.1    Inner membrane (IM) associated T2S components</b>	<b>14</b>
<b>1.2.4.2    Outer membrane (OM) associated T2S components</b>	<b>16</b>
<b>1.2.5    The type III secretion system (T3SS)</b>	<b>17</b>
<b>1.2.5.1    Introduction</b>	<b>17</b>
<b>1.2.5.2    The non-flagellar T3SS – the contact dependent pathway</b>	<b>17</b>

<b>1.2.6</b>	The type IV secretion system (T4SS) – the conjugational secretion system	20
<b>1.2.7</b>	The type V secretion system (T5SS)	23
<b>1.2.7.1</b>	Introduction	23
<b>1.2.7.2</b>	The N-terminal signal peptide	23
<b>1.2.7.3</b>	The autotransporter passenger-domain	25
<b>1.2.7.4</b>	The autotransporter chaperone-domain	26
<b>1.2.7.5</b>	The autotransporter β-domain	26
<b>1.2.7.6</b>	Mechanisms of autotransporter secretion across the OM	26
<b>1.2.7.6.1</b>	N-terminus first – the threading model	27
<b>1.2.7.6.2</b>	C-Terminus first – the hairpin model	27
<b>1.2.7.6.3</b>	The multimeric model	28
<b>1.2.7.6.4</b>	The BamA model	28
<b>1.2.8</b>	The type VI secretion system (T6SS)	31
<b>1.3</b>	YbgC	32
<b>1.4</b>	Aims of this study	33
<b>2</b>	<b>Materials and Methods</b>	<b>34</b>
<b>2.1</b>	<i>E. coli</i> strains and growth media	34
<b>2.2</b>	Yeast strains and growth media	34
<b>2.3</b>	Buffers and reagents	34
<b>2.4</b>	Oligonucleotides and Bioinformatics Resources	34
<b>2.5</b>	Structural modelling of EspC domains	40
<b>2.6</b>	DNA preparation, manipulation and analysis	42
<b>2.6.1</b>	Preparation of <i>E. coli</i> genomic DNA	42
<b>2.6.2</b>	Cleavage of DNA with restriction enzymes	42
<b>2.6.3</b>	Agarose gel electrophoresis	42
<b>2.6.4</b>	Amplification of DNA using Polymerase Chain Reaction (PCR)	43
<b>2.6.5</b>	Colony PCR	43
<b>2.6.6</b>	Purification of DNA	44
<b>2.6.7</b>	Calf intestinal alkaline phosphatase	44
<b>2.6.8</b>	Ligation of DNA fragments	44
<b>2.6.9</b>	Electroporation and Chemical transformation	45

<b>2.6.9.1</b>	Electroporation	45
<b>2.6.9.2</b>	Rubidium chloride transformation	45
<b>2.6.10</b>	Screening of transformants	45
<b>2.6.10.1</b>	Colony Screening	45
<b>2.6.11</b>	DNA Sequencing	46
<b>2.6.12</b>	Mutagenesis	46
<b>2.6.12.1</b>	$\lambda$ Red Recombination – making an <i>espC</i> deletion mutant	46
<b>2.6.12.2</b>	Site-directed mutagenesis	47
<b>2.6.13</b>	The yeast two-hybrid (Y2H) system	47
<b>2.6.13.1</b>	Generation of yeast two-hybrid Gal4 fusion constructs	47
<b>2.6.13.2</b>	Lithium acetate transformation of <i>S. cerevisiae</i>	48
<b>2.6.13.3</b>	$\beta$ -galactosidase assay	48
<b>2.6.13.4</b>	Statistical analysis of $\beta$ -galactosidase activity	48
<b>2.6.14</b>	Construction of pMALC2X and pET30a over-expression vectors	48
<b>2.7</b>	Protein analysis and purification	49
<b>2.7.1</b>	Preparation of <i>E.coli</i> secreted protein samples	49
<b>2.7.2</b>	Preparation of <i>E.coli</i> whole cell lysates.	50
<b>2.7.3</b>	SDS-PAGE	50
<b>2.7.4</b>	Western blot Analysis	50
<b>2.7.5</b>	Electro-elution	51
<b>2.7.6</b>	Mass spectrometry identification of proteins	51
<b>2.7.7</b>	Maltose binding protein (MBP) and His-tag purifications	51
<b>2.7.8</b>	MBP column co-purification	52
<b>2.8</b>	Polyclonal antibody generation	52
<b>3</b>	<i>In silico/bioinformatic analyses of EspC and YbgC</i>	<b>54</b>
<b>3.1</b>	Introduction.	54
<b>3.2</b>	Identification of the predicted structural/functional domains of the AT EspC	54
<b>3.3</b>	Primer design for PCR amplification of <i>espC</i> -encoded domains.	54
<b>3.4</b>	Analysis of homologous serine protease ATs.	62
<b>3.5</b>	EspC inter-domain structural modelling.	75
<b>3.6</b>	<i>In silico</i> analysis of YbgC.	78
<b>3.7</b>	Discussion.	84

<b>4</b>	<b>Generation of an EspC mutant and polyclonal antibody tools</b>	<b>86</b>												
<b>4.1</b>	Introduction.	86												
<b>4.2</b>	Generation of an <i>espC</i> EPEC mutant – Lambda red mutagenesis.	86												
<b>4.3</b>	Purification of wild-type EspC and generation of specific anti-EspC antibodies.	86												
<b>4.4</b>	Characterisation of EspC mutant with polyclonal anti-EspC sera	88												
<b>4.5</b>	Generation of specific polyclonal anti-sera to EspB and FliC <table> <tr> <td><b>4.5.1</b></td> <td>Construction of an EspB over-expression vector</td> <td>90</td> </tr> <tr> <td><b>4.5.2</b></td> <td>Construction of a FliC over-expression vector</td> <td>90</td> </tr> <tr> <td><b>4.5.3</b></td> <td>Purification of EspB and FliC</td> <td>92</td> </tr> <tr> <td><b>4.5.4</b></td> <td>Generation and validation of EspB and FliC polyclonal anti-sera</td> <td>92</td> </tr> </table>	<b>4.5.1</b>	Construction of an EspB over-expression vector	90	<b>4.5.2</b>	Construction of a FliC over-expression vector	90	<b>4.5.3</b>	Purification of EspB and FliC	92	<b>4.5.4</b>	Generation and validation of EspB and FliC polyclonal anti-sera	92	90
<b>4.5.1</b>	Construction of an EspB over-expression vector	90												
<b>4.5.2</b>	Construction of a FliC over-expression vector	90												
<b>4.5.3</b>	Purification of EspB and FliC	92												
<b>4.5.4</b>	Generation and validation of EspB and FliC polyclonal anti-sera	92												
<b>4.6</b>	Examination of EPEC protein secretion in the <i>espC</i> mutant	92												
<b>4.7</b>	Discussion	95												
<b>5.</b>	<b>Construction and use of yeast two-hybrid vectors</b>	<b>98</b>												
<b>5.1</b>	Introduction	98												
<b>5.2</b>	Generation of <i>espC</i> passenger-domain and <i>ybgC</i> Y2H vector constructs	98												
<b>5.3</b>	Orientation of <i>espC</i> domain fragments and <i>ybgC</i> within pGBT9 and pGAD424.	101												
<b>5.4</b>	Y2H protein-protein interactions utilising <i>espC</i> and <i>ybgC</i> Y2H constructs <table> <tr> <td><b>5.4.1</b></td> <td>Plasmid combinations used in Y2H plate assays</td> <td>103</td> </tr> <tr> <td><b>5.4.2</b></td> <td>Analysis of quantitative β-galactosidase assays</td> <td>105</td> </tr> </table>	<b>5.4.1</b>	Plasmid combinations used in Y2H plate assays	103	<b>5.4.2</b>	Analysis of quantitative β-galactosidase assays	105	101						
<b>5.4.1</b>	Plasmid combinations used in Y2H plate assays	103												
<b>5.4.2</b>	Analysis of quantitative β-galactosidase assays	105												
<b>5.5</b>	Further definition of the interacting inter- and passenger-domain regions <table> <tr> <td><b>5.5.1</b></td> <td>Generation of Y2H constructs encoding truncated passenger-domain</td> <td>107</td> </tr> <tr> <td><b>5.5.2</b></td> <td>Selection and amplification of passenger sub-domains</td> <td>107</td> </tr> </table>	<b>5.5.1</b>	Generation of Y2H constructs encoding truncated passenger-domain	107	<b>5.5.2</b>	Selection and amplification of passenger sub-domains	107	105						
<b>5.5.1</b>	Generation of Y2H constructs encoding truncated passenger-domain	107												
<b>5.5.2</b>	Selection and amplification of passenger sub-domains	107												
<b>5.6</b>	Generation of Y2H constructs encoding truncated inter-domain	110												
<b>5.7</b>	Validation of Y2H interactions <table> <tr> <td><b>5.7.1</b></td> <td>Construction of over-expression vectors encoding EspC passenger- and inter-domains</td> <td>112</td> </tr> <tr> <td><b>5.7.2</b></td> <td>Expression of recombinant EspC domains</td> <td>112</td> </tr> <tr> <td><b>5.7.3</b></td> <td>Co-purification studies</td> <td>116</td> </tr> </table>	<b>5.7.1</b>	Construction of over-expression vectors encoding EspC passenger- and inter-domains	112	<b>5.7.2</b>	Expression of recombinant EspC domains	112	<b>5.7.3</b>	Co-purification studies	116	110			
<b>5.7.1</b>	Construction of over-expression vectors encoding EspC passenger- and inter-domains	112												
<b>5.7.2</b>	Expression of recombinant EspC domains	112												
<b>5.7.3</b>	Co-purification studies	116												
<b>5.8</b>	Discussion	119												

<b>6.</b>	<b>Targeted mutagenesis of the EspC inter-domain and secretion analysis</b>	<b>122</b>
<b>6.1</b>	Introduction	122
<b>6.2</b>	Identification of mutagenesis targets on EspC	122
<b>6.3</b>	Rationale for detection of phenotypic defects following mutagenesis of conserved residues.	125
	<b>6.3.1</b> Site-directed mutagenesis and analysis of mutated <i>espC</i> encoded by pLAC1 and pLAC2 transformants.	125
	<b>6.3.2</b> Secretion analysis in WT and <i>espC</i> mutant EPEC strains transformed with mutated pLAC2 plasmids.	127
<b>6.4</b>	Discussion.	134
<b>7.</b>	<b>General Discussion</b>	<b>136</b>
<b>8.</b>	<b>References</b>	<b>142</b>

## Index of tables

<b>1.1</b>	Representative selection of T5SS proteins and their observed (or proposed) function(s)	24
<b>2.1</b>	Bacterial and yeast strains used in this study.	35
<b>2.2</b>	Oligonucleotides used for Sequencing	36
<b>2.3</b>	Oligonucleotides used for PCR	37
<b>2.4</b>	Bioinformatic resources utilised during this study	38
<b>2.5</b>	Plasmids used/constructed in this study	39
<b>2.6</b>	Antibody concentration, targets and sources	53
<b>3.1</b>	Identification of homologous proteins utilising a PSI-BLAST search of EspC whole sequence	55
<b>3.2</b>	Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pi/Mw, BioEDIT and GoIV analyses of EspC signal peptide	68
<b>3.3</b>	Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pi/Mw, BioEDIT and GoIV analyses of EspC passenger-domain	69

<b>3.4</b>	Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pI/Mw, BioEDIT and GoIV analyses of EspC inter-domain	70
<b>3.5</b>	Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pI/Mw, BioEDIT and GoIV analyses of EspC β-domain	71
<b>3.6</b>	Identification of YbgC homologous proteins utilising a PSI-BLAST. pI values Obtained using EDITSEQ of DNASTAR.	79
<b>5.1</b>	PCR reaction mixtures and conditions for amplification of DNA encoding the functional EspC domains and YbgC	99
<b>5.2</b>	Summary of Y2H interactions	104

## Index of figures

<b>1.1</b>	Schematic of transport across the inner membrane (IM).	8
<b>1.2</b>	A model of the T1SS.	13
<b>1.3</b>	Schematic of the T2SS (Main Terminal Branch).	15
<b>1.4</b>	Schematic of the T3SS apparatus core components.	19
<b>1.5</b>	Schematic of the T4SS - the conjugational secretion system.	22
<b>1.6</b>	Schematic of the multimeric model of AT translocation.	28
<b>1.7</b>	Diagrammatic representation of BamA model of AT translocation.	30
<b>1.8</b>	Schematic Type VI Secretion System (T6SS)	31
<b>3.1</b>	ClustalW alignment of the autotransporter proteins	56
<b>3.2</b>	Schematic representation of <i>espC</i> .	61
<b>3.3</b>	Hydrophilicity profile analysis of EspC and homologous protein signal peptides.	64
<b>3.4</b>	Hydrophilicity profile analysis of EspC and homologous protein passenger-domains.	65
<b>3.5</b>	Hydrophilicity profile analysis of EspC and homologous protein inter-domains.	66
<b>3.6</b>	Hydrophilicity profile analysis of EspC and homologous protein β-domains.	67
<b>3.7</b>	Phylogenetic tree generated from entire AT sequences.	72
<b>3.8</b>	Phylogenetic tree generated from AT signal peptide and passenger-domain sequences	73

<b>3.9</b>	Phylogenetic tree generated from AT inter- and β-domain sequences.	74
<b>3.10</b>	Alignment of the EspC inter-domain sequence subgroup against Hbp2 (truncated Hbp)	76
<b>3.11</b>	Crystal structure of haemoglobin protease AT	77
<b>3.12</b>	ClustalW alignment of YbgC amino acid sequence compared against homologous sequences identified by a PSI-BLAST search.	80
<b>3.13</b>	Hydrophilicity profile analysis of YbgC and homologous proteins.	81
<b>3.14</b>	Phylogenetic analyses of homologous YbgC proteins	82
<b>3.15</b>	Crystal structure of Hydrolase (Accession: 1S5U).	83
<b>4.1</b>	Generation and characterisation of an <i>espC</i> mutant	87
<b>4.2</b>	Verification of EspC purification and characterisation of an <i>espC</i> deletion mutant	89
<b>4.3</b>	Cloning of <i>espB</i>	91
<b>4.4</b>	Generation of purified EspB and FliC and validation of the rabbit polyclonal serum generated against them	93
<b>4.5</b>	Western immunoblot characterisation of EPEC E2348/69Δ <i>espC</i> .	94
<b>5.1</b>	Optimisation of PCR amplification of DNA encoding EspC β-, inter- and passenger-domains, and <i>ybgC</i> .	100
<b>5.2</b>	Verification of pGAD424 and pGBT9 constructs by PCR and <i>EcoRI</i> restriction digests.	102
<b>5.3</b>	Quantitative assay of β-galactosidase activity.	106
<b>5.4</b>	Schematic representation of passenger sub-domains and truncated inter-domain fragments	108
<b>5.5</b>	Optimisation of PCR amplification of DNA encoding EspC passenger sub-domains.	109
<b>5.6</b>	<i>EcoRI</i> restriction digests of truncated passenger- and inter-domain fragments cloned in Y2H vectors pGAD424 and pGBT9, and over-expression vector pMALC2X.	111
<b>5.7</b>	Verification of over-expression constructs	113
<b>5.8</b>	Verification of pET30a constructs encoding either EspC passenger- or inter-domain.	114
<b>5.9</b>	Verification of pMALC2X constructs encoding either EspC passenger- or inter-domain.	115

<b>5.10</b>	Assessment of passenger- and inter domain interaction by co-purification	117
<b>5.11</b>	Assessment of EspC passenger- and inter-domain interaction by column co-purification	118
<b>6.1</b>	Multiple sequence alignment of the EspC subfamily of ATs with Hbp2 identifying conserved mutagenesis targets in the EspC inter-domain	123
<b>6.2</b>	Location of the conserved GFS ( $\text{Gly}^{976}$ $\text{Phe}^{977}$ $\text{Ser}^{978}$ ) motif on the EspC inter-domain model structure	124
<b>6.3</b>	Bioedit alignment of DNA and corresponding amino acid (AA) sequence showing successful site-directed mutagenesis of arabinose independent construct pLAC1 (pBAD18:: <i>espC</i> ) using the Stratagene Quikchange II XL Site-directed mutagenesis kit	126
<b>6.4</b>	Analysis of EspC secretion in WT EPEC and its <i>espC</i> deficient mutant hosting mutated and non-mutated <i>espC</i> expression construct pLAC2	129
<b>6.5</b>	Analysis of growth in WT EPEC and its <i>espC</i> deficient mutant hosting mutated and non-mutated <i>espC</i> expression construct pLAC2 following arabinose induction	130
<b>6.6</b>	Western immunoblot analysis of TCA precipitated secreted proteins from the supernatant cultures of the EPEC <i>espC</i> deficient mutant complemented with mutated pLAC2 constructs over 3 h.	131
<b>6.7</b>	Western immunoblot analysis of EPEC <i>espC</i> deficient mutant complemented with mutated pLAC2 constructs	132
<b>6.8</b>	Western immunoblot analysis of whole cell lysates and TCA secreted proteins of the EPEC <i>espC</i> deficient mutant hosting pPheMut2.	133

## Acknowledgements

Firstly I would like to thank my supervisors Dr Rob Delahay and Dr Kim Hardie, whose patience, guidance and support contributed immeasurably to the completion of both research and this thesis. I would also like to thank Institute Directors, Prof Paul Williams and Prof Yash Mahida.

To those in the Institute of Infection, Immunity and Inflammation whom I worked with and alongside, and who helped make my time at the University of Nottingham good fun even when the going was tough, you know who you are – my thanks and appreciation.

Honourable mentions for those who worked with me during in my last days in QMC III, including Darren, Neil, Jo, Geraldine, Kevin, Richard, Karen, Marylin, Katie, Rebecca, Emma, Charlotte, James, Graham, Ricardo, Rupert, and Prof John Atherton. In addition, many thanks to Sue Dodson (who like me was there from the beginning), Dr Alan Cockayne, Dr Klaus Winzer, Darren Furniss, Paul Roach and Steve Ashmore (IT Support).

The staff of the B.M.S.U, Neil, Pip, Ian and Sue, and not forgetting Kevin Bailey from B.S.A.U without you guys sections of this and other work in III would not have been possible and somehow I never got the chance to say a big thank you before I left.

I can not leave out my current work colleagues from ADAS - Ben, Kevin, Helen, Jon, Shila, Claire, and Maged they deserve a nod for the good humour and encouragement to keep going in the moments I really had had enough of writing the thesis.

Dorothée and Benjamin, whose support, patience, humour, love and sacrifice I really could not have done without. The world is indeed a much better place to live because of you both. You have kept me going, picked me up when I was down, gave endless encouragement, countless cups of tea and reminded me constantly that writing a thesis is not in the grand scheme of things all that important. I love you both very much. I must also mention the ‘bump’ aka Benjamin’s little brother who I am looking forward to meeting in September 2009 (unless you turn up before then – gulp!).

Mum, Dad, Linda, for all your love and support, my love and thanks.

**Dedicated to my friends**

**Faye Williams (1976 – 2001)**

**Patrick Landers (1973 – 2002)**

**and**

**Samuel Oladapo Akintoye (1968 – 2008)**

**I miss them still.**

## **Abstract**

This study investigates the domain structure and secretion of the autotransporter (AT) EspC, secreted by enteropathogenic *Escherichia coli* (EPEC). The boundaries of the archetypical signal peptide and  $\beta$ -domain surrounding the passenger domain of EspC were established. Then, structural modelling was used to define a region of EspC between the passenger- and  $\beta$ -domains, termed the inter-domain, which may comprise a passenger- chaperone domain. Yeast two-hybrid and co-purification approaches were subsequently used to assess protein-protein interactions between individual EspC domains and with a putative secretion accessory factor, YbgC. Direct interaction between the EspC passenger- and inter-domain was observed, consistent with a proposed chaperone function for the inter-domain. Structural modelling identified conserved surface motifs within the inter-region as targets for mutagenesis to determine their influence upon EspC secretion. Complementation of an EPEC *espC* mutant strain with inter-domain EspC mutants showed profound affects on secretion. Furthermore, these mutants had a dominant-negative affect on wildtype EspC secretion, affecting bacterial cell viability.

These results provide the first characterisation of the putative EspC inter-domain and implicate it in facilitating efficient AT secretion. Observations from this work provide a foundation and rational direction for future research and a broad complement of research tools which will greatly facilitate future studies.

## 1. Introduction

### 1.1 Enteropathogenic *Escherichia coli*

*Escherichia coli* (*E. coli*) is a typical member of the Enterobacteriaceae, being a facultative anaerobic, non-sporing, Gram-negative bacillus. *E. coli* is usually motile with peritrichous flagella, and is often fimbriate. It was originally described and named by Theodore Escherich as *Bacterium coli commune* (Begue et al, 1998). It was later renamed *E. coli*.

The overwhelming majority of known *E. coli* strains are non-pathogenic and are the predominant facultative anaerobe of the human intestine throughout life (Begue et al, 1998). Enterovirulent *E. coli* have been divided into seven different classes of diarrhoeagenic *E. coli* based primarily on their differing modes of pathogenicity, namely EPEC (enteropathogenic), EIEC (enteroinvasive), EHEC (enterohaemorrhagic), ETEC (enterotoxigenic), EaggEC (enteroaggregative), DHEC (diarrhoea-associated haemolytic) and CDT producing (cytotoxic distending toxin) (Clark, 2001). The evolution of *E. coli* from commensal to gastrointestinal pathogen has occurred through horizontal transfer of DNA (Baümler, 1997), this transition has likely occurred several times accounting for the differences in modes of virulence (Pupo et al, 1997). EHEC for example is a major cause of serious food borne illnesses in the developed world, and is reported to have arisen very recently by the addition of toxin encoding genes (Whittam et al, 1993) from EPEC. EPEC is a major cause of gastroenteritis in children in developing countries (Nataro and Kaper, 1998). Diseases caused by enterovirulent *E. coli* are numerous and include meningitis, urinary tract infections, diarrhoea, wound infections, septicaemia and endocarditis. *E. coli* is probably the commonest global cause of bacterial diarrhoeal disease affecting both the developing and developed world.

#### 1.1.1 Attaching and Effacing (A/E) Lesions

The hallmark of EPEC pathogenicity is the attaching and effacing (A/E) lesion (Frankel et al, 1998), in which the adherence of EPEC to the epithelium of the small intestine results in a characteristic effacement of the brush border microvilli and the condensation of actin and other cytoskeletal proteins in a pedestal on which the bacteria sit (Hueck, 1998; Knutson et al, 1989; Nataro and Kaper, 1998; Kaper, 1998). While attached to the

luminal surface of intestinal epithelial cells, EPEC inject effector proteins directly into the host cell cytosol through a filamentous type III secretion system (filiT3SS) (Jarvis *et al*, 1995; Sekiya *et al*, 2001), which has been likened to a form of molecular syringe (Vallance *et al*, 2004).

### 1.1.2 Genes encoding secreted proteins

A/E lesion formation relies upon the gene products encoded on a pathogenicity island (PAI) termed the locus of enterocyte effacement (LEE) (Hueck, 1998; Nataro and Kaper, 1998) which encodes the filiT3SS and its secreted products. The LEE Region has a core of 41 open reading frames (Elliot *et al*, 1998). Genes homologous to *Yersinia* T3S (*ysc*) genes have been named '*E. coli* secretion' (*esc*). Those without *Yersinia* homologues but which have a defined role in secretion of *E. coli* proteins are called 'secretion of *E. coli* proteins' (*sep*). Genes encoding secreted proteins and their chaperones are termed '*E. coli* secretion protein' (*esp*) and 'Chaperone for *E. coli* secretion' (*ces*). The remaining genes encode putative T3S proteins currently without homologues in other systems, or that have retained original nomenclature, for example *eae* (Intimin) (Jerse *et al*, 1990) and *ler* (LEE encoded regulator, Mellies *et al*, 1999).

The LEE-encoded T3S components are transcribed from three polycistronic operons designated LEE1, LEE2 and LEE3 (Mellies *et al*, 1999). Many of the secreted Esp proteins are part of a fourth polycistronic operon designated LEE4 (Mellies *et al*, 1999). Genes involved in intimate attachment to the host cell, *tir*, *cesT*, and *eae*, were found to be transcribed from a fifth polycistronic operon termed LEE5.

### 1.1.3 EPEC proteins secreted by the T3SS

#### 1.1.3.1 The translocator proteins.

The translocator proteins are a subset of EPEC T3SS proteins which form an extracellular filamentous structural assembly on the surface of the bacterium which contacts the host cell surface and mediates delivery of effector proteins through the host plasma membrane (see overview T3SSs). EPEC translocator proteins include EspA (Knutton *et al*, 1998; Hueck, 1998; Kenny, 2002; Elliot *et al*, 2001, Vallance *et al*, 2004), EspB (Foubister *et al*, 1994; Knutton *et al*, 1998; Taylor *et al*, 1998; Kresse *et al*, 1999;

Warawa *et al*, 1999; Tacket *et al*, 2000) and EspD (Knutton *et al*, 1998; Wachter *et al*, 1999; Daniel *et al*, 2001).

### 1.1.3.2 The translocated LEE effector proteins.

Following assembly of the T3SS EspA filament and pore structure in the host cell membrane, EPEC secretes a set of effector proteins into the host cytosol via the translocator assembly which profoundly interfere with host cell signalling processes. These translocated effector proteins include Map (Vallance *et al*, 2004; Kenny, 2002; Jepson *et al*, 2003), Tir (Jerse and Kaper, 1991; Kenny *et al*, 1997; Knutton *et al*, 1997; Vallance and Finlay, 2000), EspF (McNamara *et al*, 2001; Kenny, 2002; Crane *et al*, 2001), EspG (Elliot *et al*, 2001; Tomson *et al*, 2005), EspH (Tu *et al*, 2003), SepD (Deng *et al*, 2004), SepL (Deng *et al*, 2004; Pallen *et al*, 2005), SepQ (O'Connell *et al*, 2004) and SepZ (Kanack *et al*, 2005).

### 1.1.3.3 The translocated non-LEE effector proteins.

It has also been established that in addition to the translocator and translocated LEE effector proteins, that there are other transported effector proteins, which are not encoded by LEE. These are known as the translocated non-LEE effector proteins, and include Espl/NleA (Mundy *et al*, 2004), NleB, NleC, NleD, NleE, NleF and NleG identified through the systematic mutagenesis of *C. rodentium* LEE genes (Deng *et al*, (2004), Cif (cycle inhibiting factor) (Marches *et al*, 2003), TccP (Tir cytoskeleton coupling protein) (Garmendia *et al*, 2004) and EspJ (Dahan *et al*, 2005). Some of these effector proteins have been experimentally demonstrated to play roles in bacterial colonization (Espl), arresting the cell cycle of Hela cells (Cif), actin assembly (TccP), and influencing the dynamics of pathogen clearance from a host (EspJ).

More recently, Tobe *et al.* (2006) provided evidence for the presence of a greatly extended repertoire of effectors encoded in the EHEC Sakai genome. 39 of the candidate T3S proteins were confirmed experimentally to be effectors through either proteomics or translocation assays or both. The majority of these additional functional effector genes were found to be encoded by nine exchangeable effector loci that lie within lambdoid prophages.

#### **1.1.4 Function of other secreted virulence factors**

In addition to the proteins encoded by the LEE, EPEC possess two other known virulence factors. The plasmid-encoded bundle forming pilus (BFP) has an important role in adhesion and the formation of a 3D micro-colony through inter-bacterial interactions (Knutton *et al*, 1999; Nataro and Kaper, 1998). Virulence studies in human volunteers have shown that EPEC containing a *bfpF* mutation (*bfpF* encodes a nucleotide binding protein, which is required for the dispersal phase of EPEC autoaggregation: Tobe and Sasakawa, 2001) exhibit reduced virulence, suggesting an important role for the dispersal of bacteria in the progression of the infection (Bieber *et al*, 1998). EHEC do not possess BFP (Nataro and Kaper, 1998). The second virulence factor is EspC, which belongs to the serine protease autotransporters of Enterobacteriaceae (SPATE) subfamily of autotransporter (AT) proteins (Stein *et al*, 1996; Mellies *et al*, 2001; Navarro-García *et al*, 2004).

#### **1.1.5 EspC**

*espC* is located within a unique region of DNA not found in commensal or laboratory strains of *E. coli*, designated the *espC* PAI which is distinct from the LEE PAI of EPEC (strain E2348/69) and the EAF plasmid (encoding the BFP and per regulator) (Mellies *et al*, 2001). The *espC* PAI contains at least two loci associated with virulence, the *espC* gene (Stein *et al*, 1996) and *orf3*, which shows predicted amino acid similarity with VirA of *Shigella flexneri* (Uchiya *et al*, 1995; Mellies *et al*, 2001) and rORF2 of EPEC (Friedberg *et al*, 1999; Mellies *et al*, 2001). The 40.5% G+C content of the unique 15195-bp region, is substantially lower than that of *E. coli* K-12 (50.8%) (Blattner *et al*, 1997), suggesting that this region was acquired by horizontal transfer (Aoyama *et al*, 1994). The *espC* PAI inserted at a chromosomal site adjacent to a tRNA-like gene, *ssrA*, which is also the site of insertion of the PAI of *V. cholerae* (Karaolis *et al*, 1998). Several open reading frames (ORFs) with predicted protein similarity to a variety of mobile genetic elements are contained in this region. The *espC* PAI was found only in a subset of EPEC, suggesting that EspC plays a role as an accessory virulence factor in some but not all EPEC strains.

EspC has enterotoxin activity *in vitro* (Elliot *et al*, 2000). Vidal and Navarro-García, (2006) have demonstrated that purified EspC can enter a cell by pinocytosis. However

when they infected HEp-2 cells (a human epithelial cell line) with EPEC the rate of EspC uptake was massively increased, augmenting cytoskeletal damage. Efficient internalisation/delivery into intestinal cells of EspC is thought to occur by two possible mechanisms (I) it is directly related to the T3SS (though how exactly is not clear), and (II) [Vidal and Navarro-García](#), have hypothesised that, like *Salmonella* and *Shigella*, EPEC may increase pinocytosis at the contact site. The basis of this is also unclear, but in *Salmonella* SopB induces profuse macropinocytosis events to accompany bacterial uptake, and in *Shigella flexneri*, cytoskeletal rearrangements occur as a result of translocated effector proteins resulting in bacterial internalisation. In both cases, the T3SS is involved. Recent work by [Vidal and Navarro-García \(2008\)](#) demonstrates that EspC is linked to the T3SS through an uncharacterized interaction with EspA (part of the molecular syringe).

Although EspC may use the T3SS to enter eukaryotic cells, it does not use it to exit the bacterial cell. Instead, like the majority of proteins secreted from Gram negative bacteria, it is an AT utilizing the type V secretion system (T5SS) (described in section 1.2.7). EspC shows amino acid similarity to members of the immunoglobulin A (IgA) protease family of ATs which include the IgA protease of *Neisseria gonorrhoeae* ([Pohlner, 1995](#)) the first AT to be identified. Several members of the AT family of proteins, including Tsh, SepA, ShMu, EspP, Pet, Pic, and EspC, have a conserved serine protease motif. None of these proteins, however, cleave IgA. It was demonstrated that Pet of EAEC is an enterotoxin that induces loss of actin microfilaments ([Eslava et al, 1998; Navarro-Garcia et al, 1998](#)). The precise function of EspC of EPEC, however, is unknown ([Drago-Serrano et al, 2006](#)). A mutation in espC did not affect the ability of EPEC to disrupt cytoskeletal rearrangement, phosphorylate a 90-kDa host protein (Tir protein), or adhere to or invade three different tissue culture cell lines ([Stein et al, 1996](#)). Recently [Drago-Serrano et al, \(2006\)](#) demonstrated, that EspC interacts and degrades haemoglobin, and hypothesized that this may contribute to the utilization of haem and haemoglobin iron for bacterial growth.

Secretory IgA antibodies to EspC that elicit protective immunity against EPEC disease have been isolated from breast milk of Mexican women ([Manjarrez-Hernandez et al, 2000](#)). This demonstrated that EspC is secreted *in vivo* and is immunogenic, and as such may prove to be an important component of any vaccine developed to protect against EPEC.

In addition to the uncertainty surrounding the role of EspC in EPEC pathogenesis the precise mechanism it employs to exit the cell is unclear; many aspects of AT biogenesis, such as interaction with chaperones or accessory factors and the source of energy for passenger domain translocation are not yet well understood (Dautin and Berstein, 2007). However ATs have been demonstrated to have distinct structural features in common (Dautin and Berstein, 2007). Before describing the current models of AT secretion, a brief overview of the known bacterial secretion pathways will be presented.

## 1.2 Overview of Gram-negative bacterial protein secretion.

### 1.2.1 Introduction.

Gram-negative bacteria have evolved multiple pathways for protein secretion (Thanassi *et al*, 2000). Proteins targeted for secretion must cross the periplasm and outer membrane (OM) in addition to the inner membrane (IM). Six major secretion pathways (Filloux *et al*, 2008; Bingle *et al*, 2008) are currently known to facilitate translocation across the OM of Gram-negative bacteria, these include (I) T1SS (ABC transporters), (II) T2SS (Secreton-dependent pathway including the chaperone usher (CU) pathway), (III) T3SS (Flagellar/Virulence-associated), (IV) T4SS (Conjugational secretion system), (V) T5SS (including ATs, Two-partner secretion or TPS and Trimeric ATs) (Jacob-Dubuisson *et al*, 2004; Newman *et al*, 2004; Henderson *et al*, 2004; Thanassi *et al*, 2005), and (VI) T6SS (Filloux *et al*, 2008; Bingle *et al*, 2008). As part of this overview, IM export pathways, Sec, Signal Recognition particle (SRP), twin arginine translocase (TAT), and YidC which assist in the integration, folding, and assembly of IM and OM secreted proteins ((Luirink *et al*, 2005; Lee *et al*, 2006; Dreissen and Nouwen, 2008) are also described.

The T2SS, T5SS, and CU pathway use the Sec translocase and a signal peptide at their N-terminus for the targeting and transport of their effector proteins across the IM. Following signal peptide cleavage and folding, the mature protein is then released from the periplasmic space across the OM (Filloux, 2004). The T5SS and CU pathways are comparatively simple systems allowing secretion across the OM without the need for input of energy from the IM (Thanassi *et al*, 2005). The T1SS, T3SS, and T4SSs translocate proteins across both membranes in a single energy-coupled step (Saier,

2006). T6SSs can also directly translocate a potential effector protein into eukaryotic cells (Das *et al*, 2000; Pukatzki *et al*, 2006).

The CU pathway forms an additional terminal branch of the general secretory pathway (Thanassi *et al*, 2002) and is discussed only briefly here. The CU pathway is dedicated to the construction and secretion of a superfamily of virulence-associated surface structures (Thanassi *et al*, 1998). This pathway takes its name from the components of its secretion machinery, which consist of a periplasmic chaperone that works in concert with an integral OM protein (usher). Following export across the IM by the Sec system, the pilus subunits are bound to the chaperone proteins, which prevents the self-assembly of pili/fimbriae (rod-like fibers) in the periplasm (Saier, 2006). Interactions between the chaperone and usher proteins release the pilus subunits, which are subsequently exported through the usher protein across the OM as a prelude to pilus assembly on the outer surface of the OM. The CU pathway may also assemble amorphous, capsular structures (Thanassi *et al*, 2005).

### **1.2.2 Inner membrane (IM) mediated translocation**

Three distinct IM protein export pathways have been identified (I) the Sec-dependent pathway (Pugsley, 1993; Dreissen and Nouven, 2008), comprising (A) SecB-dependent pathway, and (B) the signal recognition particle (SRP) mediated pathway (Luirink *et al*, 1992, Wolin, 1994), (II) the Tat system (Voulhoux *et al*, 2001; Lee *et al*, 2006), and (III) YidC (Dalbey and Kuhn, 2004; van der Laan *et al*, 2005; Dreissen and Nouven, 2008) an IM transport system that can act both in concert with Sec-dependent systems (Pohlschröder *et al*, 2005) or independently (Chen *et al*, 2002) (see Fig 1.1). All essential Sec and SRP- related genes, and at least one YidC (or homologue), are present in every bacterium whose genome has become available for investigation (Economou, 1999; Pohlschröder *et al*, 2005).

#### **1.2.2.1 SecB-dependent translocation**

The Sec pathway is the only known universally conserved protein translocation pathway (Pohlschröder *et al*, 2005). Initially the substrate or protein is targeted to translocase sites in the IM determined by the N-terminal extension known as the signal peptide. Generally 20-30 amino acids long, the signal peptides do not share strict conservation of

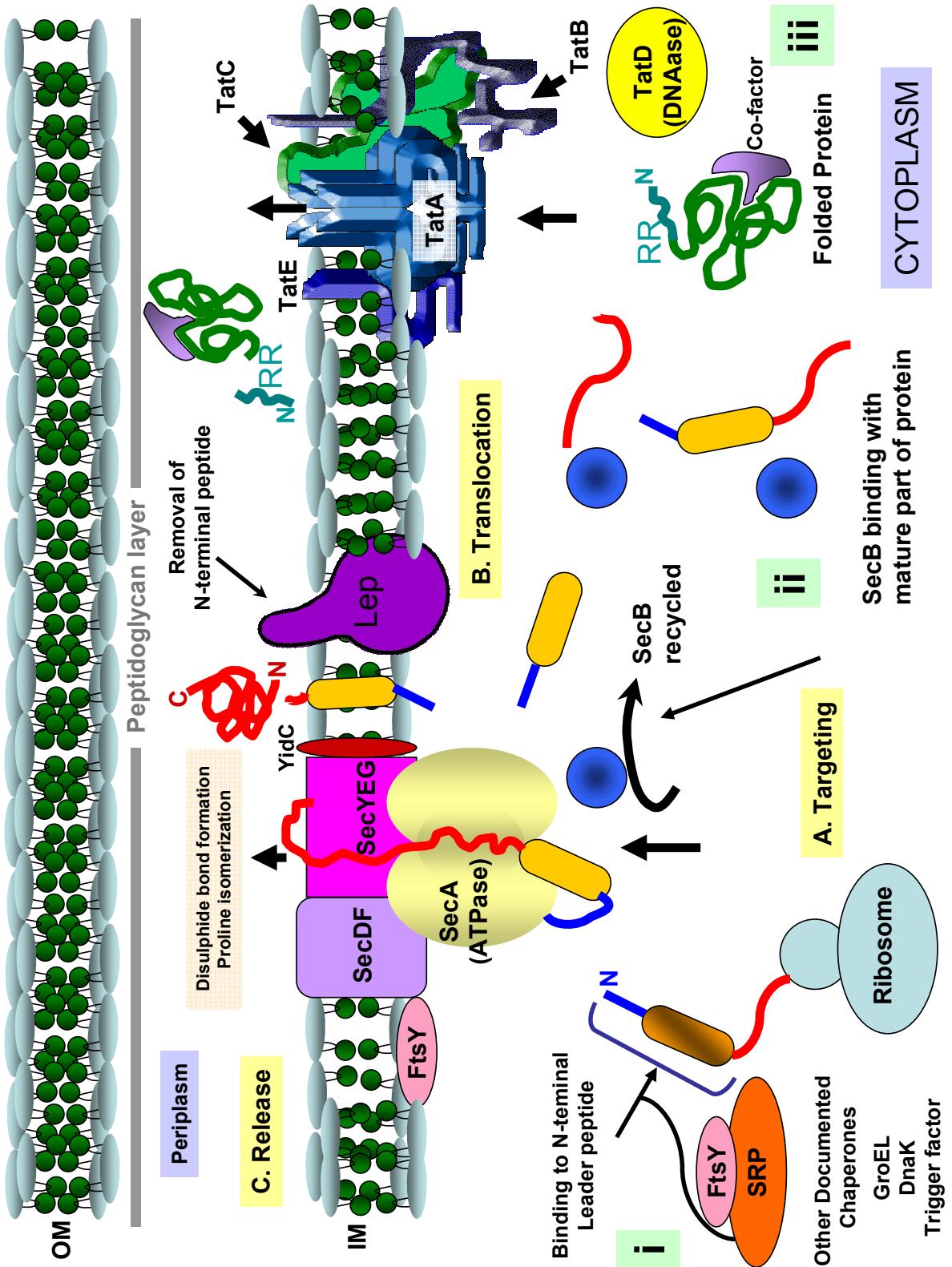


Figure 1.1: Schematic of transport across the inner membrane (IM). Three different mechanisms are shown for transport across the IM (i) Signal recognition particle (SRP) pathway, (ii) SecB dependent pathway, (iii) Twin arginine translocase (Tat) pathway. The three stages of translocation are also shown (A) Targeting by N-terminal peptide either directly or by chaperone through binding to IM proteins, (B) Translocation across the IM and (C) release into the periplasm and N-terminal leader peptide removal by serine peptidase (Lep). The OM (OM) and peptidoglycan layer are both indicated on the schematic. Relative protein sizes shown are not shown to scale or accurate. YidC has been included in this schematic as a part of the Sec translocon.

primary sequence. However they do share similar chemical properties, including (I) a positively charged amino terminus, (II) a central hydrophobic core and (III) polar residues flanking the other side of the hydrophobic region (Pugsley, 1993).

For some proteins, targeting also requires a chaperone called SecB, a protein also utilised in T1S (Delepelaire and Wandersman, 1998; Wolff *et al*, 2003). SecB binds the pre-protein, preventing premature folding, and facilitates attachment at the extreme carboxyl terminus of SecA (Fekkes *et al*, 1997). When SecB binds, the pre-protein is transferred onto SecA which is bound to SecYEG at the IM (Fekkes *et al*, 1998); it is then released and recycled (Fekkes *et al*, 1997). In addition to SecB, other chaperones have been implicated in Sec dependent translocation including GroEL, DnaK and trigger factor (Economou, 1999).

Translocation is catalysed by a pre-protein translocase, which is comprised of two distinct complexes (I) SecYEG and (II) SecDF-YajC-YidC (Duong and Wickner, 1997, Driessen *et al*, 1998). The SecYEG translocase forms a membrane-spanning domain containing heterotrimers of SecY, SecE and SecG (Meyer *et al*, 1999; Bost and Belin, 1997). SecDF-YajC-YidC though non-essential for IM translocation in *E. coli*, increases the rate of export 10 fold (Pogliano and Beckwith, 1994). The SecDFYajC complex has been suggested to (I) promote IM cycling of SecA (II) stabilize SecG and (III) facilitate the release of Sec substrates from the pore.

SecA, an ATPase, is unique to bacteria and essential, operating as the motor of translocation across the IM (Pohlschroder *et al*, 1997). SecA is large, elongated and forms a dimer (Duong *et al*, 1997; Economou, 1998). Cycling of ATP causes conformational changes in SecA resulting in the release of the polypeptide causing its movement through the translocase (Economou and Wickner, 1994; Economou *et al*, 1995). Acidic phospholipids have also been determined to be an important component of the active pre-protein translocase; they are required for (I) activation of SecA, (II) correct signal peptide interaction and (III) SecYEG stability (van Klompenburg and deKruijff, 1998). Two sources of energy are used in Sec translocation (I) chemical (ATP) and (II) electrochemical (proton motive force or PMF). ATP is essential, PMF enhances translocation rate. (Economou, 1999).

A large serine protease IM protein called Lep, with a periplasmically exposed domain ([Paetzel et al, 1998](#)), facilitates protein release by removal of the signal peptide after which folding can occur. A variety of proteins can act as folding factors in the periplasmic space, including SurA ([Lazar and Kolter, 1996](#)), Skp, ([Bulieris et al, 2003](#)), and DegP ([Skorko-Glonek et al, 2007](#)).

### **1.2.2.2 Signal recognition particle (SRP) pathway**

The signal recognition particle (SRP) pathway involves cytosolic factors strongly resembling components involved in targeting proteins to the eukaryotic endoplasmic reticulum membrane ([Luirink and Dobberstein, 1994](#); [Wolin, 1994](#)). This pathway is particularly important for the targeting of polytopic (multiple-membrane-spanning) membrane proteins. The SRP consists of a ribonucleotide/protein complex, which is composed of the Ffh protein (homologous to a 54 kDa eukaryotic SRP subunit, hence Ffh - fifty four homologue) ([Luirink et al, 1992](#)) and a 4.5S-RNA species ([Economou, 1998](#)).

The SRP acts like a chaperone facilitating insertion into the translocase co-translationally. SRP, unlike the SecB, recognises the signal peptide and in combination with FtsY (a GTPase) targets mostly hydrophobic polytopic proteins. FtsY associates with the IM and cytoplasm ([Herskovits et al, 2000](#)), and interacts directly with IM phospholipids to release the SRP in an energy dependant step involving GTP. The pre-protein is then targeted for export through the SecYEG translocase. The SRP pathway is a SecB-independent step ([Valent et al, 1998](#)) and has been shown to be required in the translocation of non-IM proteins e.g. the AT haemoglobin protease (Hbp) ([Sijbrandi et al, 2003](#)).

### **1.2.2.3 Twin arginine translocase (Tat) pathway**

The Tat pathway is involved in a wide range of cellular functions including metabolism, OM biosynthesis and bacterial pathogenesis etc (see [Gohlke et al, 2005](#); [Lee et al, 2006](#)). It is distinctive in that it not only transports folded proteins ([Berks, 1996](#); [Wu et al, 2000](#)), but discriminates against misfolded proteins ([de Lisa et al, 2003](#)). Originally described in eukaryotes as a mechanism of protein transport across the chloroplast

thylakoid membrane (Settles *et al*, 1997), TAT homologues have been identified in many bacteria, but are not ubiquitous, nor considered essential to survival (Wu *et al*, 2000).

A signal peptide at the N-terminus is required, which is structurally similar to Sec-dependent leader sequences. In general the Tat signal peptide is (I) longer (II) hosts a positively charged Sec-avoidance signal (Cristóbal *et al*, 1999) and (III) includes a twin arginine motif (Berks, 1996; Wexler *et al*, 1998), subsequently demonstrated not to be essential (Stanley *et al*, 2000; Hinsley *et al*, 2001). The majority of known Tat-dependent proteins contain co-factors including (I) nucleotides (II) simple metal ions, and (III) complex metal co-factors. Most proteins cannot be transported via the Tat pathway without co-factor (Palmer and Berks *et al*, 2003), though examples have been described (Jongbloed *et al*, 2000; Oschner *et al*, 2002).

Two major Tat complexes have been isolated from the IM. The first complex contains TatB and TatC with a small proportion of TatA (Bolhuis *et al*, 2001; de Leeuw *et al*, 2002), and has been shown to act as the receptor element of the translocation pathway (Cline and Mori, 2001; de Leeuw *et al*, 2002). TatC forms a stable core within the TatBC complex upon which the TatB component assembles (Orriss *et al*, 2007). The second major complex is comprised of TatA with a small proportion of TatB (Sargent *et al*, 2001; de Leeuw *et al*, 2002). Recent studies suggest TatA transiently associates with TatBC-substrate during active protein translocation, where TatA assembles in ring-like structures to create a protein conducting channel, capable of varying its size depending on the substrate protein (Gohlke *et al*, 2005). Translocation via Tat is entirely dependant upon the PMF (Yahr and Wickner, 2001), while polypeptide binding to the Tat is independent of PMF (Ma and Cline, 2000).

#### 1.2.2.4 YidC

YidC an evolutionary conserved IM protein (Dalbey and Kuhn, 2004; van der Laan *et al*, 2005), that operates in conjunction with and without the SRP- and Sec- system (Houben *et al*, 2002; van der Laan *et al*, 2004). It has been suggested to be involved directly in membrane insertion, folding and assembly of proteins (Dalbey and Kuhn, 2004; Pohlschröder *et al*, 2005), when the ribosome is in proximity to the Sec translocase. Recent work by Ravaud *et al*, (2008) confirmed the C-terminus of YidC is conserved in YidC homologues and required for the YidC insertase function. This role may assist

ribosome docking to allow co-translational insertion (de Gier and Luirink, 2003). However no evidence has been uncovered to suggest YidC is involved in targeting or translocation across the IM (Sijbrandi *et al*, 2003). Proteins strictly dependent on YidC or its homologues (Alb3, Oxa1) include F1F0 ATPase (van der Laan *et al*, 2004). YidC dependent proteins are involved in energy transduction and translocation, and have profound direct and indirect effects when depleted (Yi *et al*, 2003). YidC may be thought of as an integral membrane chaperone which interacts transiently with membrane proteins during their biogenesis and stimulates their correct assembly (de Gier and Luirink, 2003).

### 1.2.3 The type I secretion system (T1SS) - ABC transporters

The T1S pathway is common to many organisms (including *Agrobacterium tumefaciens*, *E. coli*, *Pseudomonas aeruginosa*, *Synechococcus spp*) and allows proteins of diverse size and function to be translocated (Delepelaire, 2004). In Gram-negative bacteria T1S is carried out by a translocator consisting of three proteins that span the inner and outer cell membranes (Delepelaire, 2004). The first T1S pathway to be characterised was that utilized by  $\alpha$ -haemolysin of *E. coli* (Felmlee *et al*, 1985). Secretion across both membranes occurs in a single step (Fig 1.2).

The proteins that form the transporter have been well characterised and a detailed model based on the X-ray structure has been described (Thanabalu *et al*, 1998; Koronakis *et al*, 2000). Of these three proteins one is related to the *E. coli* OM protein, TolC (Wandersman and Delepelaire 1990). TolC is a trimeric protein, anchored in the OM by  $\beta$ -strands with a long periplasmic part made almost uniquely of  $\alpha$ -helices (Delepelaire, 2004). The other two components are IM proteins. The first is an ATP-binding cassette (ABC) protein (Fath and Kolter, 1993), which is likely a dimer (Ghang, 2003), and provides the energy for protein secretion (Hueck, 1998). The second protein is the membrane fusion (or adaptor) protein (MFP) (Dinh *et al*, 1994), and is likely a trimer (Thanabalu *et al*, 1998). A targeting peptide at the C-terminus of the substrate (protein requiring translocation) triggers the stepwise recruitment of the secretion apparatus following an initial interaction with the ABC protein. The substrate is not cleaved during translocation and may or may not interact with the MFP in the process.

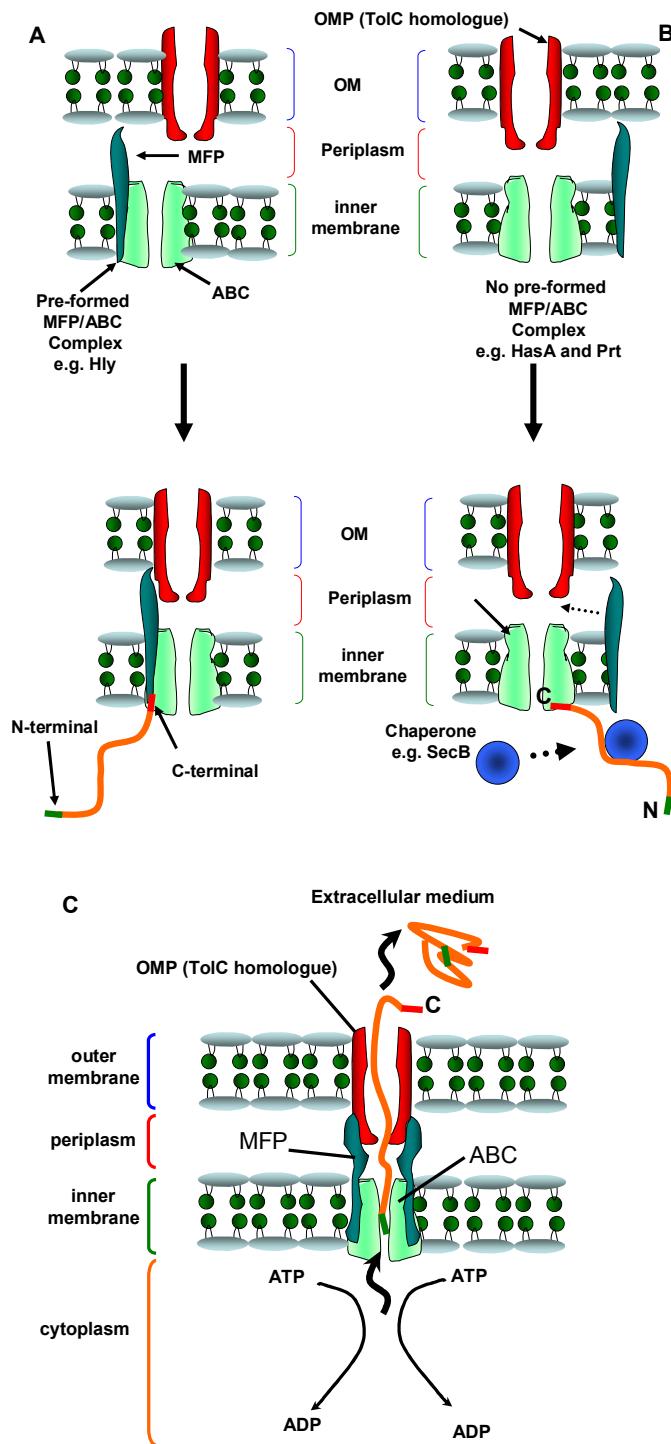


Figure 1.2: A model of the T1S pathway. Several distinct T1S assemblies have been defined. In the Hly system (A) a pre-formed complex between the MFP and ABC protein is apparent though it is conjectured that this may be an experimental artefact. In contrast in the Has and Prt assembly no preformed complex (B) between the MFP and ABC protein has been observed. The schematic demonstrates the assembly of the secretion complex driven by the C-terminal secretion signal which remains uncleaved after transport (Koronakis *et al*, 1989), and the possible involvement of a chaperone to prevent folding; though the only known example is HasA which binds the SecB chaperone. ATP hydrolysis is hypothesised to drive translocation and is indicated on the schematic (C). The stoichiometry of the different components is speculative. (Delepelaire, 2004) (Schematic adapted from Delepelaire, 2004).

#### **1.2.4 The type II secretion system (T2SS) – The secreton dependent pathway (Main Terminal Branch)**

The T2SS is widely distributed among pathogens of the Proteobacteria (Sandkvist, 2001), representatives of which are present in all subdivisions of the phylum. Often referred to as the main terminal branch (MTB) of the Sec-dependent General Secretory Pathway (GSP), the components of the secretion system were originally described in *Klebsiella oxytoca* (d'Enfert *et al*, 1987).

The T2SS is a two-step secretion process whereby the protein substrate is first exported across the IM into the periplasmic space via one of the translocases (described in section 1.2.2). Subsequent secretion of the substrate across the OM is facilitated by a dedicated secretion apparatus comprising 12-16 proteins called the secreton, which spans both the IM and OM (Sandkvist, 2001; Henderson *et al*, 2004) (see Fig 1.3). The loci encoding MTB components are relatively conserved, with most genes (*gspA* to *gspO*) transcribed from a single operon (Francetic and Pugsley, 1996; Thanassi and Hultgren, 2000).

##### **1.2.4.1 Inner membrane (IM) associated T2S Components.**

Gsp proteins associated with the IM (G, H, I, J and K) are homologous to the pilus subunit pilin and are called 'pseudopilins' (see Fig 1.3). Pseudopilins have been shown in several organisms to undergo processing by GspO (a specialised peptidase), assembling into a pilus structure which spans the IM from cytoplasm to periplasm (Bally *et al*, 1992; Pugsley *et al*, 1992; Howard *et al*, 1993).

GspE, a member of the T2S/T4S superfamily of NTPases (Planet *et al*, 2001), is suggested to utilise pseudopilins as a substrate in a manner similar to the cycle of insertion and de-insertion of SecA in a role as a traffic ATPase for transport across the IM. GspE lacks hydrophobic domains to anchor it in the IM with which it is associated, implying it interacts with additional Gsp components to form a functional translocator, likely GspL and GspF (Py *et al*, 2001). GspF is thought to aid GspE association with the IM or be integral to pore formation. GspM is essential for GspL stability, and *vice versa*, implying they also interact with one another (Possot *et al*, 2000; Sandkvist *et al*, 1999).

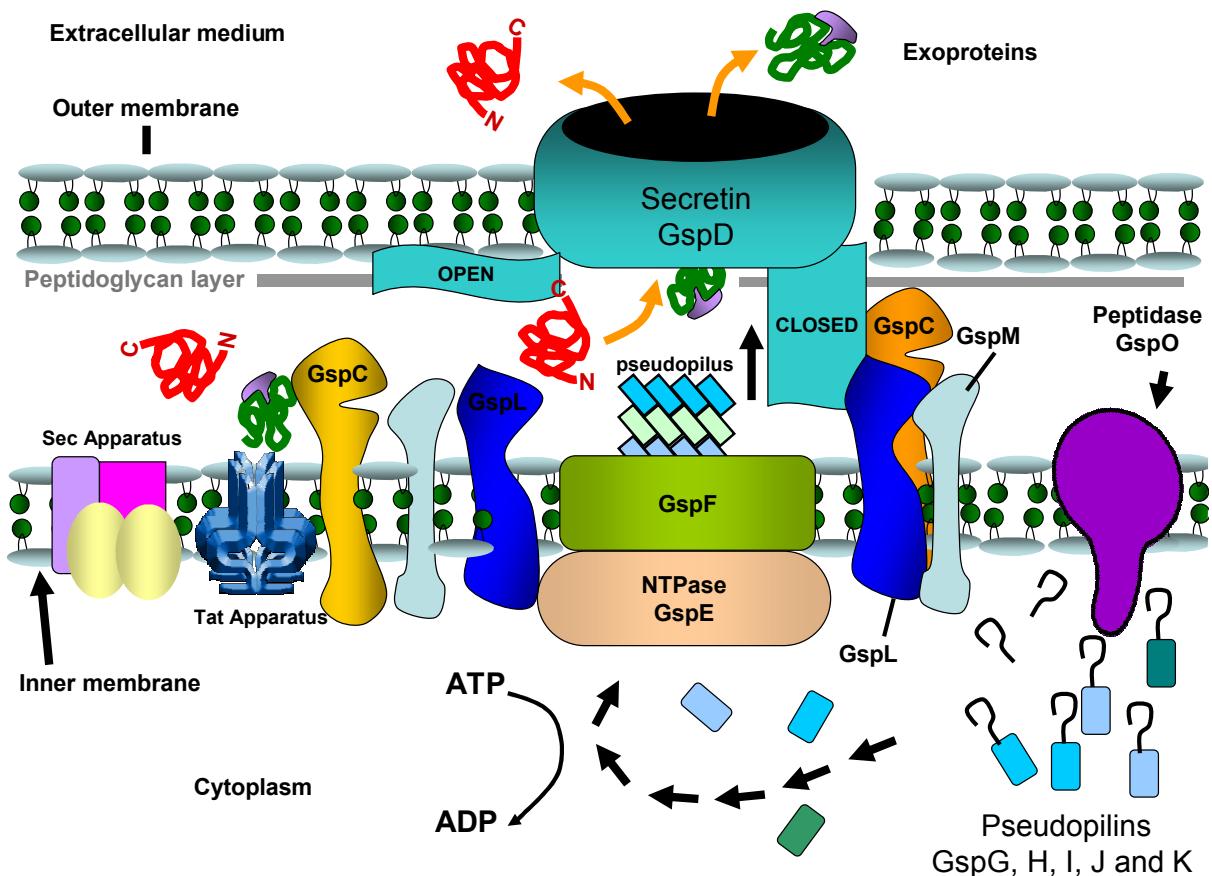


Figure 1.3: Schematic of the T2SS (Main Terminal Branch). Proteins are recognised by the secretin and transported across the OM by the secretin, GspD (shown as ring like structure, forming a large channel). GspM and GspL are proposed to interact with one another, and are directed to a specific site in the cell envelope. GspL also interacts with GspE, anchoring it to the cytoplasmic membrane. GspF (shown in green in the cytoplasmic membrane) has also been demonstrated to interact with GspL and GspE, it may therefore help associate the GspE with the membrane or be integral to pore formation allowing pseudopilin translocation. GspE is an NTPase (shown in pale red colour) and its opening and closure has been linked to ATP hydrolysis cycles and promotes the assembly of pseudopilins via the central cavity of GspE, which then form a pseudopilus structure that pushes the exoproteins through the secretin channel. When GspC interacts with the periplasmic N-terminal domain of GspD shown in bright orange on the right the 'gate' of secretin is closed. In the absence of this interaction, the central cavity of GspD is accessible ('gate' shown as open, GspC shown on left in gold to denote a different conformation) to exoproteins such as those translocated via the Sec (indicated as red protein above) and Tat machinery (indicated as green and purple protein above). The pseudopilins (GspG, H, I, J, K) undergo removal of the leader peptides (shown as hook) by GspO (prepropeptidase). GspG is the major pseudopilin component. Pseudopilin GspK is proposed to arrest elongation of the pseudopilus. GspA, GspB, GspN and GspS have not been represented. (schematic adapted from [Filloux, 2004](#)).

GspC is structurally like TonB (Bleves *et al*, 1996), and is thought to have a role in energy transduction and channel gating of the secretin in a similar siderophore-mediated manner (Possot *et al*, 1999). GspA (hosts an ATP binding site) and GspB (with TonB homology), appear to be crucial but have not been found in most known T2SSs. Homologues are found in the *E. coli* gsp gene cluster however, and are thought to form a complex in the IM to facilitate protein secretion by converting energy from ATP hydrolysis, although the exact role has not yet been determined (Francetic and Pugsley, 1996; Schoenhofen *et al*, 1998; Filloux, 2004).

#### 1.2.4.2 OM associated T2S Components

Only two known T2SS proteins are located in the OM, GspD and GspS, which comprise the Secreton.

GspD proteins are generally called secretins and form multimers consisting of 12-15 subunits (Bitter, 2003; Chen *et al*, 1996). The GspD C-terminus is proposed to be required for insertion of protein into the OM (Bitter *et al*, 1998), while the N-terminus extends into the periplasm to facilitate protein interactions. GspD has been shown to form oligomeric ring shaped structures (pores) (Bitter *et al*, 1998), large enough to accommodate folded proteins. In addition, GspD is thought to be gated, in a manner analogous to TonB to prevent cell death (Filloux, 2004).

The N-terminus of secretin is suggested to act as the channel gate through a mechanism involving conformational change of the receptor and proton motive force (Nouen *et al*, 2000). This may be achieved by (I) folding back into the cavity formed by the C-terminus, or (II) interaction with other proteins (GspC, GspN or GspB proteins are considered to be suitable for this function).

GspS is an OM lipoprotein found in some but not all T2SSs. One GspS (PulS) is essential for OM insertion of its cognate secretin (PulD), helps protect PulD against proteolytic degradation (Hardie *et al*, 1996), and is possibly a part of a secretin complex (Nouen *et al*, 2000).

## **1.2.5. The type III secretion system (T3SS)**

### **1.2.5.1. Introduction**

T3SSs fall into two major groups: those that (I) are associated with flagellar biosynthesis, or (II) which mediate interactions between bacteria and eukaryotic cells (virulence-associated T3SSs) (Pallen *et al*, 2005). These T3SSs are designated here as flagellar or non-flagellar T3SSs respectively (flagT3SS or nflagT3SS). A subset of the nflagT3SSs linked with A/E pathogens (e.g. EPEC) can also be referred to as filamentous T3SSs (filiT3SS) as these uniquely possess long filaments extending beyond the needle complex.

flagT3SSs contribute to pathogenesis by facilitating (I) movement to optimal environments (II) interaction with host cells, (III) formation of microcolonies (Girón *et al*, 2002), (IV) invasion, (V) biofilm formation, and (VI) mediating translocation of proteins into eukaryotic cells (Konkel *et al*, 2004). flagT3SSs and nflagT3SSs are structurally and evolutionarily related (Hueck, 1998). In *E. coli* BipA GTPase and IHF (integration host factor) co-ordinate regulation of flag- and nflagT3SSs, as these systems require significant energy resources, by enhancing LEE gene expression and repressing flagellar biosynthesis (Grant *et al*, 2003). Similar systems have been revealed in *Salmonella* (Goodier *et al*, 2001; Kelly *et al*, 2004). Whilst nflagT3SSs appear restricted to two bacterial phyla (proteobacteria and the Chlamydia/ Verrucomicrobium), flagT3SSs have been found in six (Proteobacteria, Spirochaetes, Aquificae, Thermotogae, Firmicutes, and Planctomycetes) (Pallen *et al*, 2005). Additionally, bioinformatics analyses have revealed a second cryptic, though non-functional T3SS in EHEC, which has been designated ETT2 (*E. coli* Type-III secretion system 2) (Pallen *et al*, 2005). Though non-functional as a T3SS due to inactivating deletions/mutations, regulators of LEE were identified in ETT2 (Ren *et al*, 2004) and are considered to apply to other systems Pallen *et al* (2005).

### **1.2.5.2. The non-flagellar T3SS – the contact dependent pathway**

nflagT3SSs are complex, supramolecular structures which bridge, and form a continuous channel across the IM, periplasmic space, OM, extracellular space and host cellular membrane (Gerlach and Hensel, 2007). The nflagT3SS is essentially a

molecular pump, utilising hydrolysis of ATP to export proteins from the bacterial cytoplasm across the IM, periplasm and OM, ensuring no periplasmic intermediates are required. This pathway is triggered when a pathogen comes into contact with the host cell, and is known as the ‘contact-dependent secretion system’, ‘injectisome’ or ‘molecular needle’ because of its shape, structure and ability to translocate proteins directly into host cell cytoplasm (Cornelis, 2006). At least 20 different proteins comprise the nflagT3SS machinery, nine of which are present in all known systems. The key component of the membrane-spanning basal body is formed by two oligomeric ring complexes, one found in the IM (constructed from multiple copies of three or more proteins), and the second in the OM consisting of secretins (Fig 1.4). An ATPase is bound at the cytoplasmic interface and is presumed to form a double hexameric ring, which enables and energises translocation (Müller *et al*, 2006). The OM secretins are thought to stabilize and anchor the hollow needle complex, which forms the protein-conducting channel. The needle itself is made of several hundred copies of proteins of the YscF family of proteins (Cordes *et al*, 2003) and allows translocation of at least partially unfolded substrates. Needle length appears critical for proper T3S (Mota *et al*, 2005) and the mechanism controlling it remains under debate (Journet *et al*, 2005; Cornelis, 2006; Marvolits *et al*, 2006; Mota *et al*, 2005; West *et al*, 2005). Uniquely, A/E pathogens like EPEC possess a long filament structure composed of EspA which extends beyond the needle (Knutton *et al*, 1998). The EspA filament length can extend up to ten times further than the average nflagT3SS needle, and is hypothesised to be a critical adaptation enabling penetration of the thick glycocalyx layer of the intestinal epithelium by the T3S apparatus of A/E organisms (Ghosh, 2004; Yip and Strynadka, 2006). A structure known as the ‘needle extension’ is found in some nflagT3SSs, this is comprised of EspB and EspD in EPEC which associate with both EspA and the host membrane (Wolff *et al*, 1998; Wachter *et al*, 1999; Ide *et al*, 2001).

A subset of nflagT3SS effector proteins characteristically require small acidic cytosolic chaperones for delivery to the IM secretion machinery, and have similar physical properties including (I) low molecular mass, (II) acidic pI, and (III) an amphipathic C-terminal  $\alpha$ -helix. Since SycE of *Yersinia* was originally identified (Wattiau and Cornelis, 1993), a number of EPEC chaperones have been described including (I) CesAB (Creasey *et al*, 2003a; ), (II) CesT (Abe *et al*, 1999; Elliot *et al*, 1999; Creasey *et al*, 2003b), (III) CesD, (Wainwright and Kaper, 1998) (IV) CesD2, (Neves *et al*, 2003), and (V) CesF, (Elliot *et al*, 2002).

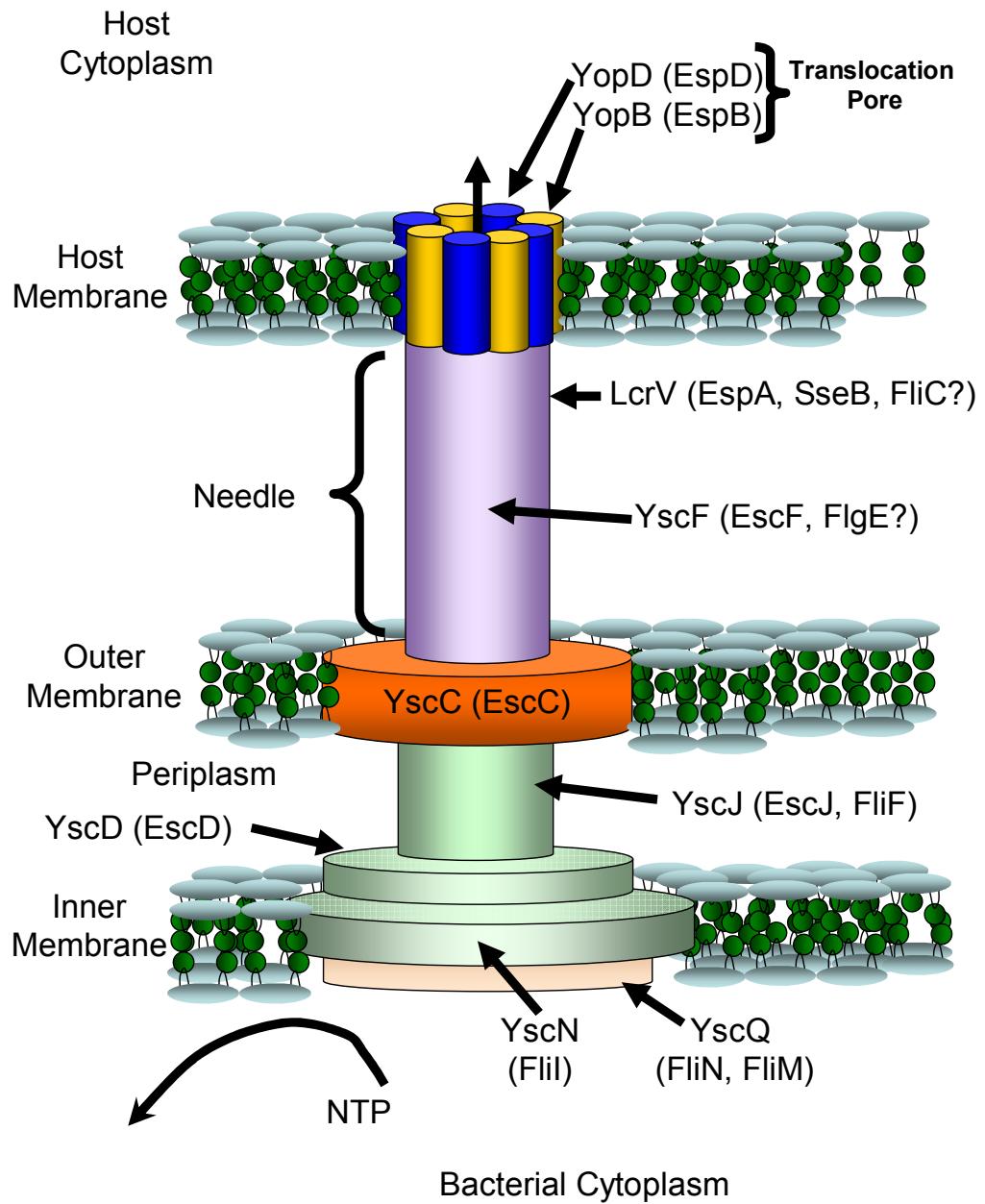


Figure 1.4: Schematic of the T3SS apparatus core components. The contact-dependent secretion apparatus of the non-flagellar/virulence-associated pathway is highly similar to the bacterial flagellar apparatus. Both possess IM (light green) and OM ring structures (orange), a membrane associated ATPase (pink) and a helical extracellular structure (blue). The nflagT3SS exports structural components (needle, needle extension, and translocation pore sub-units (e.g. EspB/EspD), to the host cell during assembly. On assembly, bacterial proteins are translocated directly into the cytoplasm of the host cell, a development that requires the needle extension (purple) and translocation pore formed by two proteins on the target membrane (blue/yellow). Components of the T3S apparatus are referred to above by *Yersinia* family names with *E. coli* and flagellar equivalents indicated in parentheses.

Chaperones are thought to stabilize, prevent premature folding and mask domains of proteins required for membrane targeting in the cell (Letzelter *et al*, 2006). After chaperone release from an effector-chaperone complex, energy is released resulting in, ATP hydrolysis and unfolded protein being loaded into the T3S apparatus (Akeda and Galan, 2005). Interestingly, recent work by Paul *et al*, (2008) has uncovered that proton motive force (PMF) alone can drive T3S, demonstrating that the flagellar secretion apparatus functions as a proton-driven translocator with no requirement for ATP hydrolysis.

How effector proteins are recognised as substrates for T3S is unclear, as no classical signal peptides have been identified. However, there is support for a protein based signal sequence, (Ghosh, 2004); one such putative protein based signal has been demonstrated for YopE, and is degenerate (Lloyd *et al*, 2001). Analysis of predicted signal peptides of effectors from *E. coli*, *Salmonella*, *Shigella*, and others have revealed enrichment of I, S, N, and T amino acid residues (Lloyd *et al*, 2002).

#### **1.2.6 The type IV secretion system (T4SS) – The conjugal secretion system.**

Type IV secretion systems (T4SSs) are often referred to as the conjugal secretion system, due to the high sequence similarity with bacterial conjugation machinery from which they are thought to have evolved (Gerlach and Hensel, 2007). T4SSs are of particular relevance and concern because they are the dominant apparatus for the transmission amongst bacteria of both antibiotic resistance genes (DNA) and virulence factors (proteins) (Ochman *et al*, 2000).

Examples of bacteria that use T4SSs include *Agrobacterium tumefaciens*, *Helicobacter pylori*, and *Neisseria gonorrhoeae*. Of the two subclasses of T4SSs, T4A is the most studied e.g. *A. tumefaciens* (Burns, 1999; Christie *et al*, 2005). The best characterized T4B system is the virulence associated *dot* (defect in organelle trafficking)/*icm* (intracellular multiplication) machinery of *Legionella pneumophila* uncovered during experiments to screen for mutants unable to survive within macrophages (Vogel and Isberg, 1999).

DNA translocated via this route is single stranded and is contained within a nucleoprotein particle with a relaxase (Christie, 2004). The relaxase binds the oriT (origin of transfer) of a mobile DNA element generating a site, and strand specific nick, which initiates conjugal processing of substrate DNA. Additionally, DNA transfer is unidirectional (5' to 3'), suggesting that relaxase pilots the DNA through the secretion channel (Christie *et al*, 2005).

The classic *A. tumefaciens* T4SS comprises 12 proteins named VirB1 through 11 and VirD4 (Fronzes *et al*, 2009). T4SSs in other bacteria can display additional components or have homologs for only a subset of those proteins (Christie *et al*, 2005). T4SSs are anticipated to span the periplasm and both bacterial membranes, structurally resembling a transenvelope secretion channel, with some T4SSs associated with an extracellular pilus, through which substrates can pass (Christie *et al*, 2005; Fronzes *et al*, 2009) (Fig 1.5). The transenvelope complex components are members of the mating-pair formation (Mpf) protein family. Protein-protein interaction studies have identified a core T4SS complex consisting of four proteins, VirB7, 8, 9, and 10 (Christie *et al*, 2005; Das *et al*, 2000). Together with VirB4, this T4SS core complex is the minimal functional entity (Lui and Binns, 2003; Harris *et al*, 2001). Protein transfer can be divided into several distinct steps which includes substrate processing, docking, and translocation. Initially a core secretion system complex comprising the most highly conserved subunits of transenvelope components is formed (Fig. 1.5). This is followed by early recruitment of VirB1, a peptidoglycan hydrolase, which facilitates T-pilus assembly by (I) disruption of the peptidoglycan wall at specific sites of Vir-T4SS assembly, initiating insertion of the T4SS, and (II) T pilus assembly via interaction with T-pilus subunits (Ward *et al*, 2002; Christie *et al*, 2005; Zupan *et al*, 2007).

In the second stage, the T-pilus is completed (Zupan *et al*, 2007), and the Mpf structure formed. The large VirB11 IM ATPase is then recruited (Cao and Saier, 2001) and as a consequence of ATP-dependent conformational changes is proposed to promote the formation of the transenvelope structure. VirD4 (the ‘coupling protein’) recruits substrates utilising a positively charged export signal (Das and Xie, 1998; Llosa *et al*, 2003; Jakubowski *et al*, 2004). ATP energy and proton motive force (pmf) have both been demonstrated to be important in translocation (Christie *et al*, 2005). All known secretion substrates have translocation signals with consensus Arginine motifs and similar hydrophilic profile at their C-terminus (Vergunst *et al*, 2000; Vergunst *et al*, 2005).

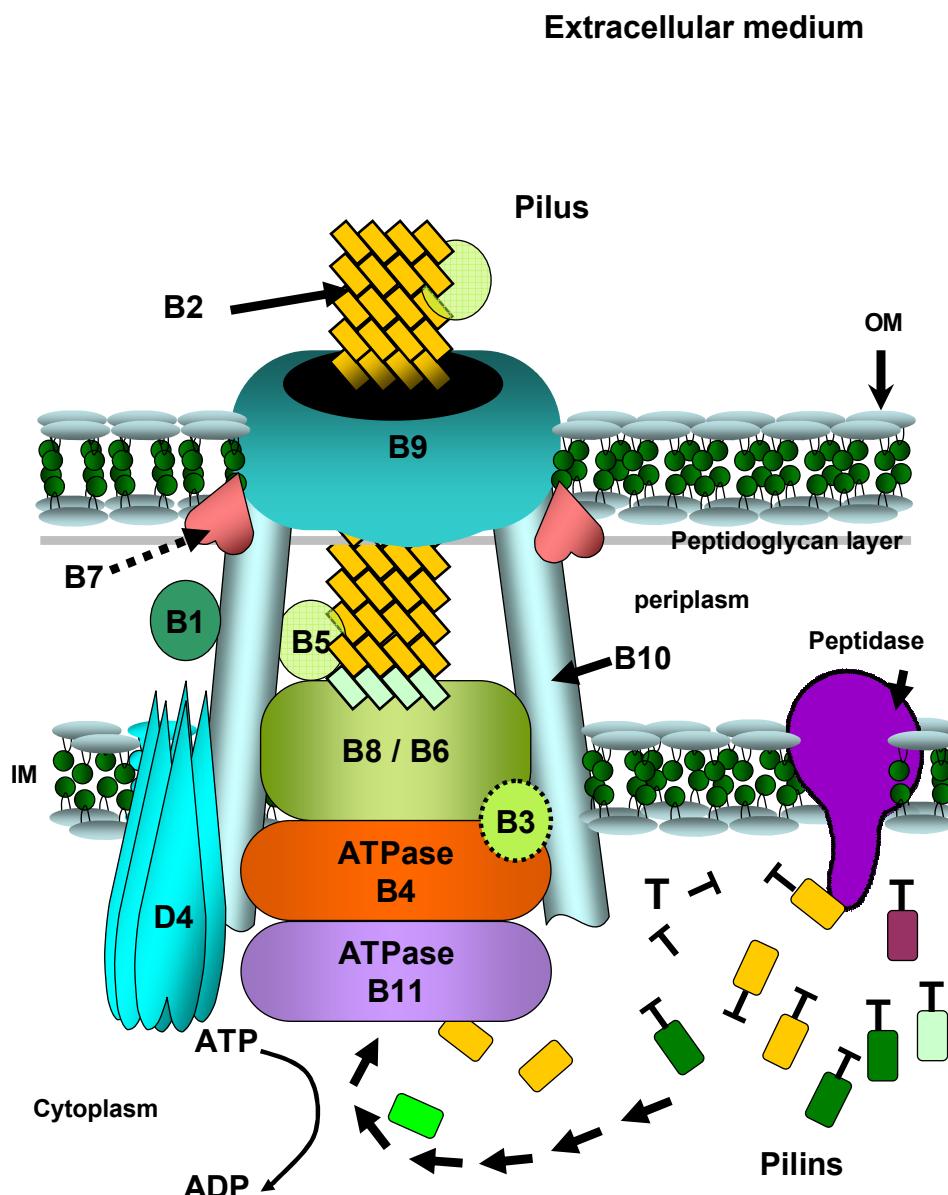


Figure 1.5: Schematic T4SS – the conjugational secretion system. The above schematic is based on the VirB/D4 T4SS (e.g. *A. tumefaciens*). The system has components homologous to the T2SS. The T4 supramolecular structure spans the cytoplasm, IM, periplasm and OM into the extracellular medium. Initially a core complex comprising the most highly conserved subunits VirB4, B8, B9 and B10 is formed (which can include lipoprotein VirB7), followed by early recruitment of VirB1 to facilitate core assembly. In the second stage, VirB2, VirB3, and VirB5 are recruited to assist pilus assembly. The Mpf structure now formed recruits the VirB11 ATPase which undergoes ATP – dependent conformational changes which affects membrane binding and establishment of Mpf contacts; this is believed to stimulate the formation of the transenvelope structure, composed of a VirB2 polymer (pilus) and OM pore complex. VirD4 recruits substrates, via a positively charged export signal, aiding docking of targeted substrates to the translocation channel (VirB10, B8, and B6). (Figure adapted from Fronzes et al, 2009; Christie et al, 2005 and Filloux, 2004)

## **1.2.7. The type V secretion system (T5SS)**

### **1.2.7.1. Introduction.**

The Type V secretion system (T5SS) is almost ubiquitous among Proteobacteria, and is the largest known secretion pathway with in excess of 800 ATs identified ([Pallen et al, 2003; Kajava and Steven, 2006](#)). All the characterized ATs produced by bacterial pathogens contribute to virulence by participating in a wide diversity of processes including adhesion, aggregation, invasion, biofilm formation, serum resistance and cytotoxicity ([Henderson et al, 2004](#)) (See Table 1).

There are currently three known groups of T5SSs; (I) Autotransporters (T5A or type V<sub>A</sub>), (II) Trimeric autotransporters (T5C or type V<sub>C</sub>), which though initially grouped with IgA proteases (T5A) were subsequently found to require trimerization to promote their secretion ([Surana et al 2004; Yeo et al 2004; Cotter et al 2005; Meng et al 2006](#)), and finally (III) the two partner secretion system or TPS (T5B or type V<sub>B</sub>) that is closely related to ATs ([Jacob-Dubisson et al, 2004](#)).

ATs, in contrast to the TPS system are synthesized as single polypeptides with two major domains, (I) secreted (passenger domain), and (II) mediating the secreted domains translocation. In the TPS system the two domains are independently manufactured. T5SS like T2SS utilises the Sec-dependent apparatus for IM translocation of proteins. However, while T2 proteins are folded and activated in the periplasm prior to secretion, T5 secreted proteins are activated and folded after OM secretion ([Grass and St. Geme<sup>3rd</sup> 2000; Jacob-Dubisson et al, 2001](#)), though limited folding appears to be tolerated ([Viega et al, 2004](#)).

### **1.2.7.2 The N-terminal signal peptide**

All ATs have a sec-dependent signal peptide at their N-terminus adjacent to the passenger domain. [Szabady et al, \(2005\)](#) noted the unusual length of signal peptides, which contain a unique highly conserved sequence motif, which is found in all signal peptides of the serine protease ATs of *E. coli* and *Shigella*.

Table 1.1: Representative selection of T5SS proteins and their observed (or proposed) function(s). The T5SS is comprised of three groups including, (I) conventional ATs (T5<sub>A</sub> type), (II) Two-partner secretion (TPS) (T5<sub>B</sub> type) which are closely related to ATs and (III) Trimeric ATs (T5<sub>C</sub> type). Conventional ATs in particular have been demonstrated to have a diverse array of functions.

<b>Organism</b>	<b>Protein</b>	<b>Function</b>	<b>Reference</b>
<b>T5<sub>A</sub> - Autotransporters (Conventional ATs)</b>			
<i>E. coli</i>	EspC	Enterotoxin, Haem binding protein/ Haemoglobin protease	Mellies <i>et al</i> , 2001; Drago-Serrano <i>et al</i> , 2006
	EatA	Enterotoxin	Patel <i>et al</i> , 2004
	Pet	Proteolytic toxin	Eslava <i>et al</i> , 1998
	Ag43	Facilitates biofilm formation	Kjaergaard <i>et al</i> , 2000a Kjaergaard <i>et al</i> , 2000b
	Hbp	Haemoglobin protease	Otto <i>et al.</i> , 1998,
	Pic	Mucinase	Henderson <i>et al.</i> , 1999
	AIDA-1	Adhesin	Benz and Schmidt, 1989
<i>H. pylori</i>	VacA	Toxin	Reyrat <i>et al</i> , 1999
<i>Neisseria</i> spp.	IgA1 protease	IgA1 Protease	Pohlner <i>et al</i> , 1987
<i>Shigella flexneri</i>	IcsA	Facilitates intracellular motility	Goldberg <i>et al</i> , 1993
<i>Bordetella</i> spp.	BrkA	Serum resistance	Fernandez, and Weiss, 1994
<i>Haemophilus influenzae</i>	Hap	Adhesin/ protease	St. Geme <sup>3rd</sup> <i>et al</i> , 1994
<b>T5<sub>B</sub> - Two-partner secretion or TPS</b>			
<i>Bordetella pertussis</i>	FHA/FhaC	Adhesin	Guédin <i>et al</i> , 2000
<i>Serratia marcescens</i>	ShlA/ShlB	Haemolysin/Cytolysin	Hertle <i>et al</i> , 1999 Hertle, 2002
<i>Edwardsiella tarda</i>	EthA/EthB	Cytolysin	Hirono <i>et al</i> , 1997
<i>Proteus mirabilis</i>	HpmA/HpmB	Haemolysin/Cytolysin	Uphoff and Welch, 1990 Jacob-Dubuisson <i>et al</i> , 1997
<i>H. influenzae</i>	HMW1/HMW1B HMW2/HMW2B	Adhesin	Barenkamp and Leininger, 1992
	HxuA1/HxuB	Haem-binding protein	Cope <i>et al</i> , 1994; Cope <i>et al</i> , 1998
<b>T5<sub>C</sub> - Trimeric Autotransporters</b>			
<i>H. influenzae</i>	Hia	Adhesin	Surana <i>et al</i> , 2004
<i>Yersinia enterocolitica</i>	YadA	Adhesin, Serum Resistance	Roggenkamp <i>et al</i> , 2003
<i>E.coli</i>	EibA	Immunoglobulin binding protein	Sandt and Hill, 2000 Sandt and Hill, 2001
<i>Neisseria meningitidis</i>	NadA	Adhesin	Comanducci <i>et al</i> , 2002

The extended signal peptide of an AT is at least 42 to 60 amino acids in length, with a similar extension found in the closely related TPS pathway ([Jacob-Dubuisson et al, 2001](#)). The signal peptides can be roughly divided into two domains, a C-terminal segment which resembles a classical Sec signal peptide, and a unique N-terminal extension of approximately 25 amino acids ([Dautin and Bernstein, 2007](#)). The C and N-terminal domains can be divided into two further sub-domains. The N-terminus of the signal peptide extension hosts (I) a highly conserved N-terminal motif (designated n1) and small region of hydrophobic/aromatic residues (designated h1), and (II) a degenerate C-terminus comprising an unusual number of charged amino acids (designated n2) ([Brunder et al, 1997](#)), a hydrophobic core (designated h2), and consensus cleavage site (C). The N-terminal extension inhibits SRP mediated IM translocation, ensuring post-translational targeting ([Peterson et al, 2006; Desvaux et al, 2007](#)). [Peterson et al. \(2006\)](#) suggest this delay slows dissociation from the Sec machinery to control periplasmic folding.

#### 1.2.7.3 The autotransporter passenger-domain

The passenger domain (mature protein), which is translocated to the extracellular surface, can vary considerably in size ranging from approximately 600 to 3000 amino acid residues in length ([Barnard et al, 2007](#)). Between ATs there is considerable sequence variability, and even those with significant homology can have very different functions ([Henderson et al, 2004](#)), indicating that function is specific to the host-bacterium interaction.

Once translocated across the IM via the Sec machinery, the signal peptide is cleaved and the AT is ferried across the periplasm by periplasmic chaperones such as Skp and Sur ([Sklar et al, 2007](#)). The C- terminus of this periplasmic AT bears the  $\beta$ -domain which is integrated into the OM. The passenger domain is then secreted across the OM and remains either cell-associated or is cleaved and released into the extracellular milieu ([Henderson et al, 1998](#)). The putative mechanisms of translocation across the OM are discussed in more detail later. The passenger domains of ATs (T5A) also undergo a range of post-synthesis processing steps, which include cleavage from the bacterial cell surface, lipidation, glycosylation, oligomerisation and localisation to distinct regions of the OM ([Dautin and Bernstein, 2007](#)).

#### **1.2.7.4 The autotransporter chaperone-domain.**

Within the passenger domain, adjacent to the  $\beta$ -barrel domain is a stretch of amino acids, predicted by bioinformatics analyses to encode an  $\alpha$ -helix. Using a series of PCR generated in-frame deletion mutants in this region, Oliver *et al*, (2003a) revealed, a chaperone-like activity thought to prevent premature folding of the protein until it has crossed the OM (Hendrixson *et al*, 1997; Oliver *et al*, 2003a), and enhance translocation efficiency possibly by (I) stabilizing the  $\beta$ -Barrel (Mogenson *et al*, 2005) or (II) by encouraging correct folding (Ohnishihi *et al*, 1994; Oliver *et al*, 2003b; Verlarde and Nataro, 2004). The autochaperone is proposed to be involved in forming a hairpin-like structure which leads secretion of the passenger domain through the channel formed by the  $\beta$ -core (Henderson *et al*, 1998). Recent work on EspP supports this view and shows the catalytic site for cleavage of the passenger domain to lie inside the  $\beta$ -barrel (Barnard *et al*, 2007).

#### **1.2.7.5 The autotransporter $\beta$ -domain**

The  $\beta$ -domains of the ATs are conserved in both sequence and function, and comprise approximately 250-300 amino acid residues ( $T5_A$ ) and 70 amino acid residues ( $T5_C$  type) respectively (Barnard *et al*, 2007). The  $\beta$ -domains form  $\beta$ -barrel structures in the OM, facilitating the translocation of the passenger domain. Three crystal structures for  $\beta$ -domains, all with a highly similar architecture, have been determined, (I) NalP (Oomen *et al*, 2004), (II) EspP (Barnard *et al*, 2007) and (III) Hia (a trimeric AT) (Meng *et al*, 2006). Dautin and Bernstein (2007) hypothesise that despite sequence divergence of the translocated passenger domain, the conserved sequence and architecture of the  $\beta$ -barrel and presence of one or more  $\alpha$ -helices inside the  $\beta$ -barrel imply that ATs possess similar domain translocation mechanisms. A number of models (Oomen *et al*, 2004; Meng *et al*, 2006; Jong *et al*, 2007; Skillman *et al*, 2005) have been proposed to explain passenger domain translocation and are explored in more detail in the following section.

#### **1.2.7.6 Mechanisms of autotransporter secretion across the OM**

The mechanism of autotransport first described for the IgA protease of *N. gonorrhoeae* (Pohlner *et al*, 1987) postulated the precursor protein being exported beyond the IM in a Sec-dependent manner, coupled to the cleavage of a signal peptide. The Sec-

dependent SRP has been shown to be involved in the export of at least one AT, namely haemoglobin protease (Hbp) ([Sijbrandi et al, 2003](#)). How the  $\beta$ -barrel inserts into the OM and transports the passenger-domain through is unclear. However, recent work on OM complexes such as BamA (Omp85/YaeT) ([Voulhoux et al, 2003; Wu et al, 2005](#) [Knowles et al, 2009](#)) have provided clues as to how this might occur. Four putative secretion mechanisms for crossing the OM have been proposed including (I) the threading model, (II) the hairpin model, (III) the multimeric model and (IV) the BamA model.

#### **1.2.7.6.1 N-terminus first – the threading model**

Translocation starts with the N-terminus. The  $\beta$  barrel accommodates one extended strand at a time for which the pore observed in the crystal structure is wide enough. Doubt has been cast on this particular model because (I) truncated passenger domains lacking N-terminal residues, and (II) passenger domain chimeras lacking N-terminal targeting signals can still be effectively translocated across the OM ([Dautin and Bernstein, 2007](#)).

#### **1.2.7.6.2 C-terminus first – the hairpin model**

A temporary hairpin fold located at the translocator/passenger domain interface forms within the pore of the translocator domain possibly during the assembly of the  $\beta$ -barrel ([Jose et al, 1995](#)). This hairpin may compromise the autochaperone found in many ATs ([Oliver et al, 2003](#)), but could also be essential for the translocation of some ATs such as BrkA ([Oliver et al, 2003](#)) and AIDA-I ([Maurer et al, 1999](#)).

In both the hairpin and threading model, the passenger folds extracellularly and must remain unfolded at the periplasmic side of the membrane, perhaps assisted by chaperones ([Henderson et al 1998, Purdy et al 2002](#)). Recent work by [Letley et al, \(2006\)](#), on proximal cysteine residues within the passenger domain of the VacA AT showed that a small hairpin loop could potentially be formed as in IgA1 protease of *N. gonorrhoeae*. See [Rutherford and Mourez, \(2006\)](#) or [Dautin and Bernstein, \(2007\)](#) for a schematic of the proposed model.

### 1.2.7.6.3 The multimeric model

Viega *et al.* (2002) proposed translocation occurs via a central pore formed by a multimeric complex of  $\beta$ -domains (Fig 1.6). The hypothesis is based on observations utilising the NalP AT which include (I) translocation (albeit at low efficiency) of a folded disulphide-bond containing artificial passenger domain (Viega *et al.*, 1999), and (II) ring-like structures revealed by electron microscopy of the *N. meningitidis* IgA protease translocator domain after its over-production and purification from *E. coli* OM preparations (Viega *et al.*, 2002). It is unclear however whether oligomers form prior to OM translocation, and how the oligomer forms a protein-translocating channel. In addition, bioinformatics analyses indicate that IgA hosts a conventional  $\beta$ -domain, implying that a conventional  $\beta$ -barrel structure is formed (Yen *et al.*, 2002; Viega *et al.*, 2002). It is proposed that only a subset of ATs, including the IgA protease, may utilise a distinct multimer based secretion process possibly following  $\beta$ -barrel insertion into the OM (Oliver *et al.*, 2003; Dautin and Bernstein, 2007; Viega *et al.*, 2008). A number of studies however have now demonstrated that ATs are monomeric including Tsh (Stathopoulos *et al.*, 1999; Hritonenko *et al.*, 2006), EspP (Skillman *et al.*, 2005), NalP (Khalid and Sansom, 2006), AIDA-I (Müller *et al.*, 2005), and Hbp (Oomen *et al.*, 2004).

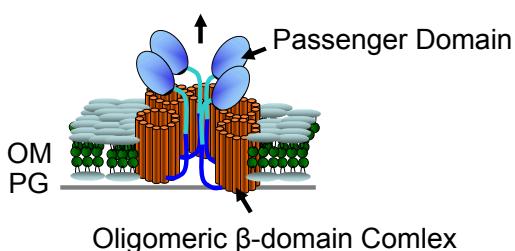


Figure 1.6: Schematic of the multimeric model of AT translocation. The multimeric model suggests that the passenger domain folds, at least partially, in the periplasm and is then translocated across the peptidoglycan layer (PG) and OM via a large channel. The large channel results from oligomerisation of the  $\beta$ -domain (adapted from Dautin and Bernstein, 2007).

### 1.2.7.6.4 The BamA Model.

A network of proteins responsible for folding and inserting OMPs (including ATs) into the OM has recently been uncovered, the core complex of which is now known as the  $\beta$ -barrel assembly machinery (BAM) (Knowles *et al.*, 2009). Nearly all known OMPs require the BAM complex for folding (Werner and Misra, 2005; Wu *et al.*, 2005; Doerrler and Raetz, 2005; Malinverni *et al.*, 2006), though there may be exceptions including the secretin PulD, which inserts into the OM in the absence of the BAM complex (Collin *et al.*, 2007; Guilvout *et al.*, 2008).

BamA (formerly Omp85) was originally characterised in *Neisseria meningitidis*, shown to be required for the assembly of integral OM proteins and additionally facilitate AT secretion (Voulhoux *et al*, 2003; Werner and Misra, 2005; Jain and Goldberg, 2007). BamA belongs to a highly conserved family of proteins present in all domains of life, except Archaea. In *E. coli* BamA (formerly known as YaeT, an orthologue of Omp85), forms a multi-protein complex with four lipoproteins including BamB (YgfL), BamC (NlpB), BamD (YfiO), and BamE (SmpA) (Knowles *et al*, 2009; Stegmeier and Andersen, 2006). BamA and BamD are essential for cell viability (Doerrler and Raetz, 2005; Wu *et al*, 2005; Sklar *et al*, 2007), and phenotypic effects in deletion mutants (of *bamC* and *bamB* in particular) suggest the others also contribute to the maintenance of cell integrity. BamA is present in all Gram-negative bacteria and has two major components (I) a set of five POTRA (polypeptide transport-associated) domains oriented towards the periplasm and (II) a carboxy-terminal β-barrel inserted into the OM (Sanchez-Pulido *et al*, 2003). BamA and other BamA-like proteins have in addition been shown to multimerize *in vitro* which is thought to imply the native BAM complex itself is multimeric. Additional evidence of BamA involvement in AT secretion includes its homology to TPS transporter proteins (Yen *et al*, 2002), studies suggesting it is able to form channels of sufficient size (Robert *et al*, 2006), and its ability to catalyse secretion of extracellular loops which can exceed 30 amino acids in length (Dautin and Bernstein, 2007). It is suggested that once the leader sequence has been processed by the signal peptidase, an unfolded AT associates with periplasmic chaperones (e.g. DegP) which are thought to comprise two pathways (I) the SurA pathway and (II) the Skp–DegP pathway, facilitating transport of ATs across the periplasm to the OM (Sklar *et al*, 2007; Onufryk *et al*, 2005; Rizzitello *et al*, 2001). The passive BAM complex is then utilized to integrate the β-barrel domain and the mature AT secreted (see Fig 1.7).

How BamA functions in OMP/β-barrel assembly remains unclear, however it is beginning to emerge that its POTRA domains might have a role in binding unfolded OMPs (Sanchez-Pulido *et al*, 2003). Bos *et al*, (2007) have suggested that the BAM complex recognizes a specific recognition motif encoded in the C-terminal β-strand of OMPs, as the targeting motif appears to differ between *E. coli* and *N. meningitidis* implying that OMP sorting is species specific. The precise sequence of events and route to the OM remains unknown, though several mechanisms by which substrate OMPs might feasibly be inserted into the OM have recently been discussed by Knowles *et al*. (2009), which are outlined in (Fig 1.7).

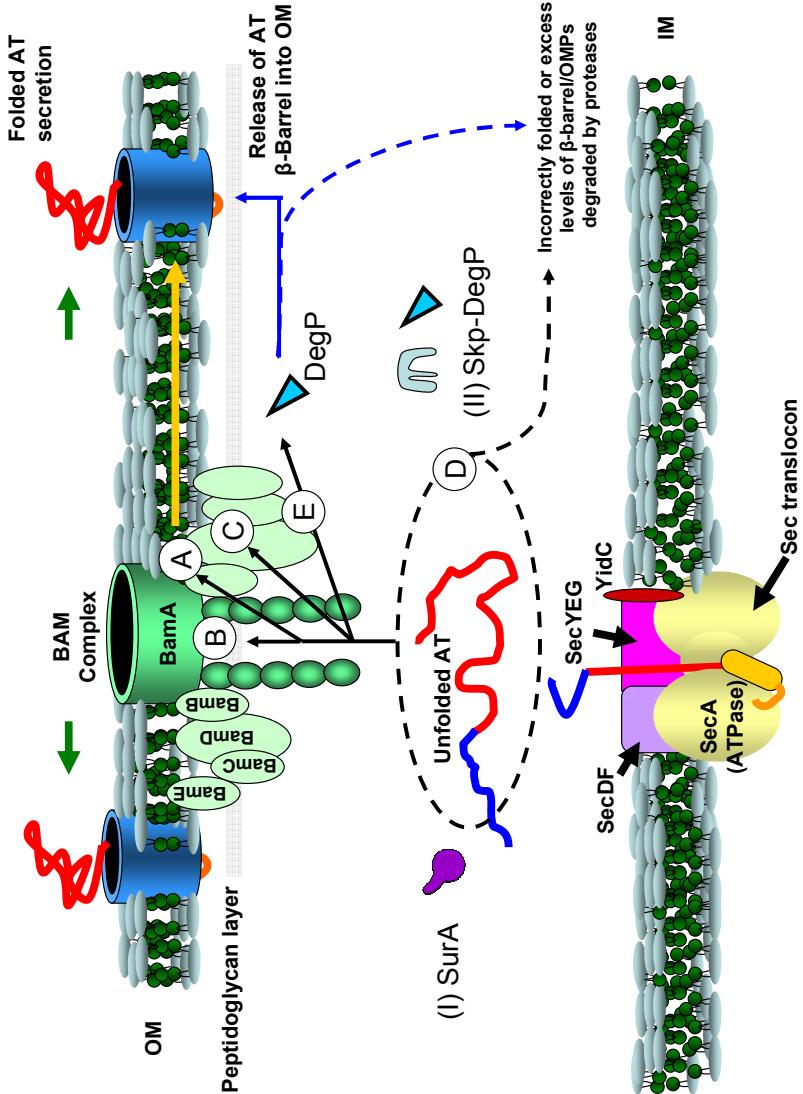


Fig 1.7: Diagrammatic representation of BamA model of AT translocation (adapted from Knowles *et al*, 2009). In *E.coli* the BAM complex is known to be comprised of at least four lipoproteins BamB, BamC, BamD (Wu *et al*, 2005), and BamE (Sklar *et al*, 2007), in addition to BamA. ATs are initially exported across the IM by the Sec-translocon, where unfolded AT is then thought to be recruited by two possible chaperone pathways, (I) the SurA and (II) the Skp–DegP pathway for transport across the periplasm to the OM. Excess levels of unfolded ATs in the periplasm are degraded by proteases (e.g. DegP). Folding and insertion of unfolded ATs is thought to occur via the BAM complex (Sklar *et al*, 2007), and though not currently clear how, a number of mechanisms have been proposed including (A) the pore-folding model, (B) the complex pore-folding model, (C) the barrel-folding model, (D) the chaperone-folding model, and (E) the accessory folding model. Firstly, in (A) the β-barrel pore of BamA is used for insertion of the unfolded AT β-barrel into the membrane, and the POTRA domains and/or accessory components act to thread the AT β-barrel into the pore. In the second model (B) the central core is formed by a multimeric BAM complex which acts as the point of insertion into the OM, where release of the AT β-barrel could then occur through dissociation of the multimeric BAM complexes. In (C) the β-barrel of BamA provides the template for barrel folding in the vicinity of the BAM complex. In (D) periplasmic chaperones (e.g. DegP), act to fold and protect the protein from degradation during its passage through the periplasm, suggesting the BAM complex only functions to insert the β-barrel of the AT into the OM. In the final model (E) the BAM complex is hypothesized to fold the embryonic AT, with no involvement in OM insertion. Once folded the AT β-barrel could then be released to DegP either facilitating removal of incorrectly folded ATs, or permitting insertion into the OM by an unknown mechanism that may/may not involve the BAM complex.

### 1.2.8 The type VI secretion system (T6SS).

Genomic screening of the Type IVB component IcmF revealed a group of PAIs previously termed IcmF-associated homologous protein (IAHP) cluster now designated the virulence associated secretion (VAS) genes e.g. in *V. cholerae*. The novel secretion system was further defined as a type VI secretion (T6SS), where the secretion of substrates into the extracellular space is accomplished without N-terminal signal peptides (Pukatzki *et al*, 2007). A total of 18 genes have been identified to date in *V. cholerae* T6SS (Shrivastava *et al*, 2008). T6SSs are widely distributed among Gram-negative bacteria (Bingle *et al*, 2008), including *Edwardsiella tarda* *S. enterica* *V. cholerae*, *P. aeruginosa* and enteroaggregative *E. coli* (Rao *et al*, 2004; Parsons and Heffron, 2005; Pukatzki *et al*, 2007; Mougous *et al*, 2006; Dudley *et al*, 2006) (Fig 1.8).

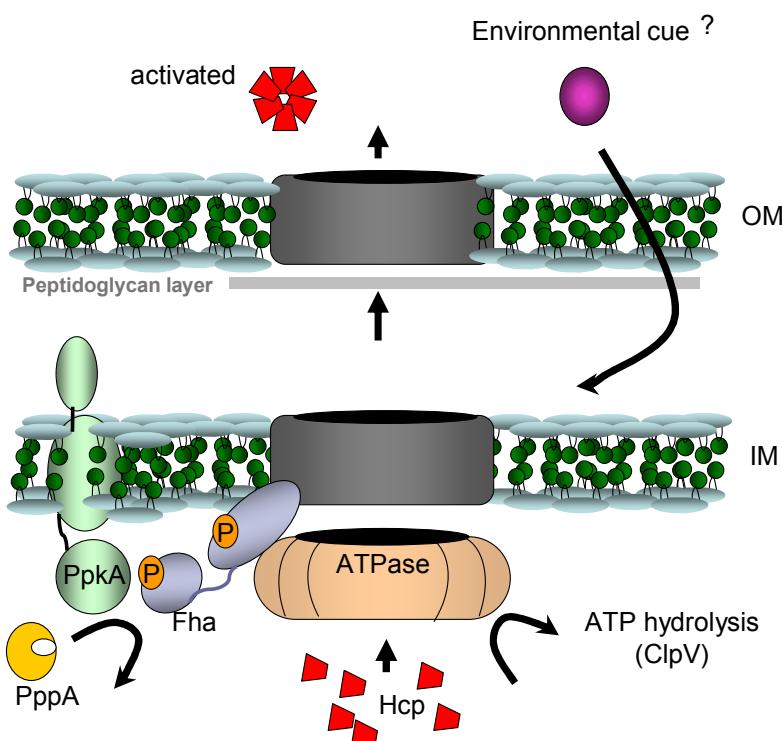


Figure 1.8: Schematic Type VI Secretion system (T6SS). Many of the basic components of the secretion system have not yet been characterised (shown in grey), though the VgrG proteins of *Vibrio* and their orthologs have been hypothesised to form heterotrimeric complexes capable of acting as a pore forming channel (Pukatzki, *et al*, 2007). PpkA and PppA regulate Hcp secretion via phosphorylation of Fha. The system is triggered by an external environmental cue (purple), which triggers dimerization of the PpkA, this causes it to autophosphorylate resulting in binding and phosphorylation of Fha (shown by orange spheres with P). A candidate for the environmental sensor is PpkA itself as it contains a large periplasmic domain. The phosphorylated Fha results in an activated state allowing Hcp secretion to be powered by the ATPase (ClpV in *P. aeruginosa* Weibezahl *et al*, 2004)). The Hcp form ring structures (red) that are thought to stack forming tubes through which a transport conduit may be formed capable of delivering effector proteins into a host cell. (Figure adapted from Mougous *et al*, 2007)

Four distinct proteins are secreted extracellularly, including (I) Hcp, (II) VgrG-1, (III) VgrG-2, and (IV) VgrG-3, secretion of Hcp being the hallmark characteristic of a functional T6SS. Hcp, VgrG-1 and VgrG-2 apparently contribute to the function of T6S apparatus despite being transported substrates of the system. Bioinformatics analyses show VrgG proteins have structural features analogous to the phage tail of bacteriophage T4, and are a fusion of the (I) phage tail protein, and (II)  $\beta$ -helix repeat (which forms the cell puncturing needle of the phage tail) ([Pukatzki et al, 2007](#)). Such a  $\beta$ -helical needle could function as an extracellular translocon, puncturing host cell membranes to deliver effector proteins into the target cytosol. In addition both VgrG-1 and VgrG-3 contain a C-terminal actin-cross-linking domain (ACD) (suggesting direct secretion of proteins into a host cell, analogous to the T3SS) and peptidoglycan (PG) binding domain. Alternate extended domains have been identified including (I) zinc metalloprotease domain (*P. aeruginosa*), (II) mannose binding domain (*V. intermedia*), and (III) Tromomyosin-like domain (*Yersinia pestis*). Together this suggests a multi-functional role for VgrG proteins (Fig 1.8).

### 1.3. YbgC

YbgC was identified in a parallel study by Dr Louise Arnold as a putative accessory factor for AT secretion, since transposon mutagenesis of the *ybgC* gene had a deleterious effect on EspC secretion. The cytoplasmic protein YbgC ([Vianney A et al, 1996](#)) is part of the Tol-Pal system of the *E. coli* cell envelope. The Tol-Pal system has been implicated in a number of possible functions including (I) anchoring the OM to the peptidoglycan layer (II) catalysis of porin biogenesis (III) regulation of porin activity and (IV) energising macromolecule transport through the OM ([Wallberger et al, 2002](#)). The Tol-Pal proteins are encoded by genes organized into two operons. The first encodes YbgC, TolQ, TolR, TolA, TolB, Pal and YbgF and the second operon encodes TolB, Pal and YbgF ([Walburger et al, 2002](#)). Amino acid identity between paired YbgC proteins in *E. coli*, *Haemophilus influenzae*, *Vibrio cholerae*, *P. aeruginosa* and *Rhodobacter capsulatus* ranges from 41 to 62% indicating a high level of conservation between different organisms. *H. influenzae* YbgC is an acyl-coenzyme A thioesterase ([Zhuang et al, 2002](#)), it catalyses the hydrolysis of short chain aliphatic acyl-CoA thioesters. YbgC is representative of a large group of structurally related ‘unknown proteins’ in bacteria whose encoding genes nestle amongst gene clusters functioning in the synthesis and/or export of extracellular substances ([Zhuang et al, 2002](#)). YbgC has previously been

shown to interact with acyl carrier protein by yeast two-hybrid (Y2H) analysis ([Walburger et al, 2002](#); [Butland et al, 2005](#)) and to exist as a dimer under non-reducing conditions ([Djamel Gully, unpublished data](#)). [Thoden et al, \(2003\)](#), have shown from the crystallographic structure of the 4-hydroxybenzoyl-CoA thioesterase from *Arthobacter* sp. that the enzyme is tetrameric and is a dimer of dimers, with all the subunits adopting a so called ‘hot dog fold’ composed of anti-parallel β-sheets flanked on one side by an α-helix. Though the quaternary structure of that observed in *Pseudomonas* and *Arthobacter* are quite different, they have been shown to possess equivalent catalytic efficiencies, substrate specificity, and metabolic functions.

#### **1.4 Aims of the Study**

This study aimed to explore aspects of AT secretion, primarily definition of a hitherto undescribed EspC inter-domain, initially by utilising the yeast two-hybrid system to investigate protein-protein interactions between individual domains of the AT EspC from EPEC. Interactions between EspC and putative accessory factors which might aid AT secretion, such as YbgC, were also considered. Toward these goals, this thesis describes;

- a) Detailed bioinformatics analysis and structural modelling of EspC to clearly define component EspC sub-domains (Signal, Passenger, Inter and Beta-barrel) and their borders.
- b) Molecular approaches for the generation of diverse constructs encoding EspC sub-domains for use in the yeast two-hybrid system and the expression of recombinant protein.
- c) Identification of pairwise interactions between EspC sub-domains which suggest the presence of a functional EspC Inter-domain.
- d) The construction of an EPEC *espC* mutant strain and antibodies specific for EPEC secreted proteins EspC, EspB and FliC.
- e) Analysis of wildtype and mutant EPEC protein secretion profiles to determine the contribution of the putative EspC Inter-domain to EspC export.

## **2. Materials and Methods**

### **2.1. *E. coli* strains and growth media.**

The bacterial strains used in this study are listed in Table 2.1. Bacterial strains EPEC E2348/69, BL21(DE3), Novablue, *E. coli* XL-1 Blue, and *E.coli* DH5 $\alpha$  were cultured in Luria broth or on LB agar plates. Liquid cultures were incubated at 37°C with shaking at 225-250 rpm; agar plates were incubated in a 37°C static incubator. Where appropriate, either 50  $\mu$ g carbenicillin mL $^{-1}$  or 25  $\mu$ g kanamycin mL $^{-1}$  was added to the media. For long term storage strains were frozen at -80°C in LB containing 20% (v/v) glycerol. When appropriate either 1 mM IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) or 0.002% (w/v) arabinose was added to induce cultures for expression of recombinant protein.

### **2.2 Yeast strains and growth media**

Yeast strain *Saccharomyces cerevisiae* strain PJ69-4A was cultured in complete medium, YPD (20 g peptone L $^{-1}$  (Oxoid), 10 g yeast extract L $^{-1}$  (Oxoid), pH 5.8), or in yeast minimal medium, YMM (6.7 g yeast nitrogen base without amino acids L $^{-1}$  (Difco), pH 5.8). Media were supplemented with 2% (w/v) glucose and the appropriate amino acids and bases at the following concentrations: Ade 20  $\mu$ g mL $^{-1}$ , Ura 20  $\mu$ g mL $^{-1}$ , Met 20  $\mu$ g mL $^{-1}$ , His 20  $\mu$ g mL $^{-1}$ , Trp 20  $\mu$ g mL $^{-1}$  and Leu 30  $\mu$ g mL $^{-1}$ . Liquid yeast cultures were incubated at 30 °C with shaking at 200 rpm; agar plates were incubated in a 30 °C static incubator.

### **2.3 Buffers and Reagents**

All buffers and reagents used were prepared in accordance with manufacturers' instructions, or [Sambrook et al, \(1989\)](#). All chemicals are supplied by Sigma unless otherwise stated. Stock solutions were prepared according to [Sambrook et al, \(1989\)](#).

### **2.4 Oligonucleotides and Bioinformatics Resources**

Oligonucleotide primers for sequencing or for use in the polymerase chain reaction (PCR) reaction were synthesised by Sigma-Genosys. Primers for site-directed mutagenesis experiments were obtained from MWG-Biotech. All primers were diluted to

a stock concentration of 10 pmol  $\mu\text{L}^{-1}$  for use in PCR reactions. A list of oligonucleotides used in this study for sequencing, their name and sequence are listed in Table 2.2. A list of oligonucleotides used in this study for PCR, their name sequence, restriction site can be seen in Table 2.3. A list of bioinformatics resources used in this study can be seen in Table 2.4. In addition, a list of the plasmids used (original vectors and constructs) are also provided in Table 2.5.

Table 2.1: Bacterial and Yeast Strains used in this study.

<b>Strains</b>	<b>Description/genotype</b>	<b>Source</b>
EPEC E2348/69	wild-type EPEC O127:H6 isolated from an outbreak in Taunton, UK	Levine <i>et al</i> , 1978
<i>E.coli</i> BL21(DE3)	F <sup>-</sup> ompT hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>+</sup> )gal DCM(DE3),	Novagen
<i>E.coli</i> Novablue	<i>endA1 hsdR17(r<sub>K12</sub><sup>-</sup>m<sub>K12</sub><sup>+</sup>) supE44 thi-1 recA1 gyrA96 relA1 lac</i> [F' <i>proA<sup>+</sup>B<sup>+</sup> lac<sup>R</sup>ZΔM15::Tn10 (Tet<sup>R</sup>)</i> ]	Novagen
<i>E. coli</i> XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lac<sup>R</sup>ZΔM15 Tn10 (Tet<sup>R</sup>)</i> ]c	Stratagene
<i>E.coli</i> DH5 $\alpha$	( <i>SupE44, ΔlacU169, deoR, (φ80lacZΔM15), hsdR17 (r<sub>k</sub><sup>-</sup>,m<sub>k</sub><sup>+</sup>) recA1, endA1, gyrA96, thi-1, relA1, phoA, F<sup>-</sup>, λ</i>	Hanahan, 1983.
<i>Saccharomyces cerevisiae</i> strain PJ69-4A	( <i>MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2 : : GAL1-HIS3 GAL2-ADE2 met2 : : GAL7-lacZ;</i>	James <i>et al</i> , 1996

Table 2.2: Oligonucleotides used for Sequencing.

Primer Name	Primer_Sequence	Primer Source*	Stock Concentration pmol $\mu\text{L}^{-1}$
<b>Internal passenger-domain primers</b>			
Int_Pas_F1	gctggtgccgaccaatg	S-G	10
Int_Pas_F2	tagtcatgtatgccttg	S-G	10
Int_Pas_F3	tatgcttggactttatt	S-G	10
Int_Pas_F4	atcatgttgttcagggg	S-G	10
Int_Pas_F5	tcaagtatgtaaaataac	S-G	10
Int_Pas_F6	tagtacaggacaggjtata	S-G	10
Int_Pas_F7	aaaagtccagtaaaacaa	S-G	10
Int_Pas_F8	tgggataacaaaagacca	S-G	10
Int_Pas_F9	aagtttgttgtgaaaca	S-G	10
Int_Pas_F10	gttccgggttctatct	S-G	10
Int_Pas_F11	gctcttgaacgttatg	S-G	10
Int_Pas_F12	atagcatccagtggtgc	S-G	10
Int_Pas_F13	ctgggacgacttatac	S-G	10
Int_Pas_R1	ctatatgtattaccattt	S-G	10
Int_Pas_R2	ttcagaatagacccaaca	S-G	10
Int_Pas_R3	ccacaccataacgttca	S-G	10
Int_Pas_R4	tcttttttacgttatac	S-G	10
Int_Pas_R5	aaaatctactgtaaatctt	S-G	10
Int_Pas_R6	gtttcattataaaaaagggt	S-G	10
Int_Pas_R7	ctaacgtaaaactgtcac	S-G	10
Int_Pas_R8	tcccagtttaactgaa	S-G	10
Int_Pas_R9	taactggacttttgagt	S-G	10
Int_Pas_R10	tatgtattaccatttt	S-G	10
Int_Pas_R11	agaatagacccaacatct	S-G	10
Int_Pas_R12	ctcaccattgtatccac	S-G	10
Int_Pas_R13	ggaattggAACATTG	S-G	10
<b>Y2H vector primers</b>			
pGAD-F	cgcgtttggaaatcactacag	S-G	10
pGAD-R	gcacgtatgcacagtgaagt	S-G	10
pGBT-F	tgccgtcacagatagatgg	S-G	10
pGBT-R	cctgagaaaggcaacctgacc	S-G	10

\*Primer Source: Sigma Genosys denoted as S-G

Table 2.3: Oligonucleotides used for PCR

\*Primer Source: Sigma Genosys denoted as S-G, MWG-Biotech denoted as MWG

Table 2.4: Bioinformatic resources used during this study to obtain DNA and protein sequences, align sequences, primer design, primer orders, sequencing, protein mapping, and modelling, codon usage, and to obtain free analytical software.

Resource	Use	Web address
National Centre for Biotechnology (NCBI)	Source of DNA and protein sequences	<a href="http://www.ncbi.gov">www.ncbi.gov (NCBI Databases)</a>
	BLAST SEARCH	<a href="http://www.ncbi.nlm.nih.gov/BLAST/">www.ncbi.nlm.nih.gov/BLAST/</a>
	Conserved Domain Architecture Retrieval Tool (CDART)	<a href="http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi">www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi</a>
ClustalW	Sequence alignments	<a href="http://www.ebi.ac.uk/clustalw">www.ebi.ac.uk/clustalw</a>
SignalP	Predicts the presence and location of signal peptide cleavage	<a href="http://www.cbs.dtu.dk/services/SignalP/">www.cbs.dtu.dk/services/SignalP/</a>
Compute pI/Mw	Computation of theoretical pI (isoelectric point) and Mw (molecular weight)	<a href="http://www.expasy.ch/tools/pi_tool.html">www.expasy.ch/tools/pi_tool.html</a>
Primer3	Primer design and restriction analysis	<a href="http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi">frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi</a>
Premierbiosoft	Primer design and restriction analysis	<a href="http://www.premierbiosoft.com/netprimer/netpraunch/netpraunch.html">www.premierbiosoft.com/netprimer/netpraunch/netpraunch.html</a>
Cut2	Primer design and restriction analysis	<a href="http://www.firstmarket.com/cutter/cut2.html">www.firstmarket.com/cutter/cut2.html</a>
Sigma Genosys	Primer orders	<a href="http://www.sigmaaldrich.com/Brands/Sigma_Genosys.html">www.sigmaaldrich.com/Brands/Sigma_Genosys.html</a>
MWG Biotech	Primer orders	<a href="http://www.mwg-biotech.com/">www.mwg-biotech.com/</a>
Swiss-Model	Protein modelling	<a href="http://swissmodel.expasy.org//SWISS-MODEL.html">swissmodel.expasy.org//SWISS-MODEL.html</a>
CPHmodels	Protein modelling	<a href="http://www.cbs.dtu.dk/services/CPHmodels/">www.cbs.dtu.dk/services/CPHmodels/</a>
GOR IV	Secondary structure predictions	<a href="http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html">npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html</a>
Proteomics Server	Bioinformatic software	<a href="http://www.expasy.org">www.expasy.org</a>
PeptIdent (now Aldente)	Experimentally measured peptide masses are compared with theoretical peptides calculated for all proteins in a protein sequence database	<a href="http://ca.expasy.org/tools/peptident.html">ca.expasy.org/tools/peptident.html</a>
Graphical Codon Usage Analyser (GCUA)	Compares codon usage in gene to genomic DNA	<a href="http://gcua.schoedl.de/">gcua.schoedl.de/</a>
Bioedit	Sequence analysis Software	<a href="http://www.mbio.ncsu.edu/BioEdit/bioedit.html">www.mbio.ncsu.edu/BioEdit/bioedit.html</a>
Chromas	Sequence analysis software	<a href="http://www.technelysium.com.au">www.technelysium.com.au</a>
Rasmol	Protein Modelling Software	<a href="http://www.umass.edu/microbio/rasmol/">www.umass.edu/microbio/rasmol/</a>
Clustal	Phylogenetic Tree Software	<a href="http://ftp.ebi.ac.uk/pub/software/dos/clustalw/">ftp://ftp.ebi.ac.uk/pub/software/dos/clustalw/</a>
Institute Links	Links to Bioinformatic resources	<a href="http://www.nottingham.ac.uk/links/tools.htm">www.nottingham.ac.uk/links/tools.htm</a>
Geneservice	Sequencing	<a href="http://www.geneservice.mrc.ac.uk">www.geneservice.mrc.ac.uk</a>

Table 2.5: Plasmids used/constructed in this study.

Plasmid	Description	Source
pGAD424	<i>oriColE1 ori2μ LEU1 P<sub>ADH</sub></i> :: <i>GAL4'</i> activator domain :: MCS Amp <sup>R</sup> , Accession No: U07647	Bartel <i>et al</i> , 1993
pGBT9	<i>oriColE1 ori2μ TRP1 P<sub>ADH</sub></i> :: <i>GAL4'</i> binding domain :: MCS Amp <sup>R</sup> , Accession No: U07646	Supplied by Dr R. Delahay. Bartel <i>et al</i> , 1993
pGAD::β	pGAD424 vector containing Beta Barrel Domain <i>espC</i> insert	This Study
pGAD::inter	pGAD424 vector containing Inter Domain <i>espC</i> insert	This Study
pGAD::pas	pGAD424 vector containing Passenger Domain <i>espC</i> insert	This Study
pGAD:: <i>ybgC</i>	pGAD424 vector containing <i>ybgC</i>	This Study
pGBT::β	pGBT9 vector containing Beta Barrel Domain <i>espC</i> insert	This Study
pGBT::inter	pGBT9 vector containing Inter Domain <i>espC</i> insert	This Study
pGBT::pas	pGBT9 vector containing Passenger Domain <i>espC</i> insert	This Study
pGBT:: <i>ybgC</i>	pGBT9 vector containing <i>ybgC</i>	This Study
pGBT:: <i>tir</i>	pGBT9 containing <i>tir</i>	Hartland <i>et al</i> , 1999
pGBT:: <i>cesT</i>	pGBT9 containing <i>cesT</i>	Delahay <i>et al</i> , 2002
pGAD:: <i>cesT</i>	pGAD424 containing <i>cesT</i>	Delahay <i>et al</i> , 2002
pGADF	<i>Bam</i> HI frameshift derivative of pGAD424	R. Delahay (unpublished)
pGADH	<i>Bam</i> HI frameshift derivative of pGAD424	R. Delahay (unpublished)
pGBT9::PS3	pGBT9 vector containing insert encoding truncated EspC passenger sub domain, 801 to 1200bp.	This Study
pGBT::IF2	pGBT9 vector encoding truncated EspC inter domain,	This Study
pGBT::IF3	pGBT9 vector encoding truncated EspC inter domain,	This Study
pGAD::IF2	pGAD424 vector encoding truncated EspC inter domain,	This Study
pGAD::IF3	pGAD424 vector encoding truncated EspC inter domain,	This Study
pET30a	Kan <sup>R</sup> , T7lac (promoter), His•Tag, S•Tag, Thrombin/Enterokinase protease sites, Multiple cloning site.	Novagen (Merck Biosciences)
pET30a::pas	pET30a containing EspC passenger domain	This Study
pET30a:: <i>espB</i>	pET30a containing EspB (Accession: AAC38396.1) secreted protein. <i>Bam</i> HI/ <i>Ncol</i> Restriction sites	This Study
pET30a:: <i>fliC</i>	pET30a containing <i>FliC</i> (Accession AAC74990.1) a flagellin protein. <i>Eco</i> RI/ <i>Hind</i> III	Ferrández <i>et al</i> , 2005
pMALC2X	Amp <sup>R</sup> , lac promoter, Maltose binding protein fusion vector	NEB
pMAL::pas	pMalC2X encoding EspC passenger Domain	This Study
pMAL::inter	pMalc2X encoding EspC inter domain	This Study
pSTBlue1	Dual opposed SP6/T7 promoters, Amp <sup>R</sup> , Kan <sup>R</sup> , Dual <i>Eco</i> RI sites flank insert. Used for sub-cloning PCR products.	Merck Biosciences
pSTBlue1::pas	pSTBlue-1 plasmid encoding EspC passenger domain	This Study
pSTBlue1:: <i>ybgC</i>	pSTBlue-1 plasmid encoding <i>ybgC</i> (Accession: NP_308798)protein	This Study
pSTBlue1::PS2	pSTBlue-1 vector encoding EspC Passenger sub Domain, 401 to 800bp	This Study
pSTBlue1::PS3	pSTBlue-1 vector encoding EspC Passenger sub Domain, 801 to 1200bp	This Study
pSTBlue1::PS4	pSTBlue-1 vector encoding EspC Passenger sub Domain, 1201 to 1600bp	This Study
pCR2.1	AmpR, KanR, pUC origin, f1 origin, M13 primer sites, MCS, LacZα. Used for sub-cloning PCR products	Invitrogen
pCR2.1:: <i>espB</i>	pCR2.1 encoding EspB (Accession: AAC38396.1) protein.	This Study
pKD4	Contains Kanamycin cassette, Accession No: AY048743	Ferrández <i>et al</i> , 2005
pKD46	Red Recombinase expression plasmid, <i>Bet</i> and <i>Exo</i> genes, Accession No: AY048746	Datsenko <i>et al</i> , 2000
pBAD18	Amp <sup>R</sup> , Arabinose promoter, Derived from plasmids, pRS48, pRS4 and pCR48, pBR322 ori, <i>amp</i> '.	Guzman <i>et al</i> , 1995
pLAC1	pBAD18 with <i>espC</i> inserted at <i>Xba</i> I/ <i>Sac</i> I within multi-cloning site. Not under arabinose induction	Dr L. Arnold & Dr K. Hardie
pLAC2	pBad18 with <i>espC</i> inserted at <i>Xba</i> I/ <i>Sac</i> I within multi-cloning site. Arabinose inducible.	Dr L. Arnold & Dr K. Hardie
pPheMut1A	pLAC1 subjected to site directed mutagenesis, Phenylalanine (Phe <sup>977</sup> ) replaced with alanine (Phe <sup>977Ala</sup> ).	This Study
pPheMut1	pLAC2 subjected to site directed mutagenesis, Phenylalanine (Phe <sup>977</sup> ) replaced with alanine (Phe <sup>977Ala</sup> ).	This Study
pPheMut2	pLAC2 subjected to site directed mutagenesis, Phenylalanine (Phe <sup>977</sup> ) replaced with Glycine (Phe <sup>977Gly</sup> ).	This Study
p3AlaMut	pLAC2 subjected to site directed mutagenesis, GFS motif replaced with alanine substitutions (Gly <sup>976Ala</sup> Phe <sup>977Ala</sup> Ser <sup>978Ala</sup> )	This Study

## 2.5. Structural modelling of EspC domains.

Protein and nucleotide sequences for EspC were required for primer design, identification of domains and structural features, and comparison with other ATs. The EspC pathogenicity island sequence found in enteropathogenic *E. coli* strain E2348/69 (Accession: AF297061), (Mellies *et al*, 2001), the *espC* gene sequence (bases 5309 to 9226) and EspC protein sequence (Accession: AAG37043) were obtained from the NCBI database.

PSI-BLAST searches (Altschul *et al*, 1997; Attwood and Parry-Smith, 1999) were performed with the EspC AT protein sequence ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) to obtain homologous protein sequences with which to identify similar ATs sequences and establish any conserved domains. The Conserved Domain Architecture Retrieval Tool (CDART, see Table 2.4) (Geer *et al*, 2002), was used to identify a selection of serine protease ATs with high homology to EspC. E-values (Expect values) of each of the whole ATs were recorded.

Following identification of a number of homologous ATs, the sequence information of each protein was obtained and saved (see Table 3.1). Prior to performing a multiple sequence alignment, the EspC protein sequence was subjected to secondary structure analysis using GorIV (see Table 2.4) to identify and avoid disruption of potentially important secondary structural features such as  $\alpha$ -helices or  $\beta$ -turns when designing the primers. Protein sequences were also submitted to SignalP (see Table 2.4) to establish any signal peptide sequence, in conjunction with referral to the literature to determine any empirically determined signal peptides. EspC was then aligned against the selected homologous ATs identified from the PSI-BLAST search and CDART domain analysis. A ClustalW alignment ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)) was performed to identify conserved and variable boundaries of the individual ATs identified; the GorIV structural analysis was included in the sequence alignment (see Figure 3.1). Consequently, using this analysis as a guide, three EspC domains and their boundaries were defined (excluding the signal peptide) and subsequently used as the basis for generation of plasmid fusion constructs for use in yeast-two hybrid (Y2H) interaction studies. Domains comprised, (I) the passenger-domain, (II) the autochaperone (inter-domain) and, (III) the  $\beta$ -barrel domain (see Fig 3.2A).

Utilising the ClustalW multiple sequence analysis, phylogenetic trees were generated revealing the most closely related homologues of EspC. In addition BioEDIT was used to generate multiple hydrophobicity profiles of each of the identified domains. Each of the identified domains of EspC (including the N-terminal signal peptide) was also subjected to (I) a PSI-BLAST search, the E-values obtained for each of the individual domains giving a measure of confidence in the match between the selected ATs and that of EspC, and a (II) ClustalW alignment to create a series of phylogenetic trees, (visualised in Bioedit), to discern possible evolutionary relationships.

Having obtained key sequence information, identified homologous proteins, and the likely domain boundaries, sequence locations close to domain boundaries were analysed further. Bioinformatics software; Primer3, Premierbiosoft and Cut2 (Table 2.4) were used to design the most appropriate PCR primers for amplification of each domain. As *espC* did not contain an *EcoR1* restriction site, the *EcoR1* restriction site was incorporated into all primers to enable non-directional cloning into the yeast two-hybrid (Y2H) vectors. *EcoR1* restriction sites were positioned within the primers so that the EspC domains could be cloned in-frame in the yeast-two hybrid vectors, making viable fusions with the Gal4 activation or binding domain fragments present in vectors pGAD424 or pGBT9 respectively.

An EspC phylogenetic sub-group, identified in earlier analyses was also subjected to a sequential series of alignments in order to establish conserved and invariant residues from identified homologous sequences. Equivalent alignments were performed with and without sequence of the haemoglobin protease (designated Hbp2) for which a crystallographic structure of the passenger domain (which includes the putative chaperone or inter domain) has been determined (1WXR, haemoglobin protease, Otto *et al*, 2005). The *espC* domain was truncated until sufficient identity allowed the mapping of *espC* domain, using Swissmodel (see Table 2.4) over Hbp co-ordinates i.e. the Hbp passenger domain was used as a template for EspC structure. Once *espC* had been modelled, a PDB file was created and subsequently visualised in RasMol (See Table 2.4). The model was used as the basis for the identification of conserved and possibly functionally/structurally important residues in the inter-region which were subsequently targeted for site-directed mutagenesis. The structural model also informed the design of constructs for use in Y2H studies which comprised progressively smaller *espC* inter-domain fragments.

## **2.6. DNA Preparation, manipulation and analysis**

### **2.6.1 Preparation of *E. coli* genomic DNA**

Purified genomic DNA was prepared using the Genelute bacterial genomic DNA Kit (Cat No: NA2100, Sigma) and Qiagen Genomic DNA Genomic-tip (20/G), according to manufacturer's protocols. Purified DNA was stored at 4°C until required; DNA was diluted 1:10 in distilled water for PCR reactions. Crude genomic DNA preparations used for template in some PCR reactions were prepared as follows. 500 µL broth from an overnight EPEC culture was centrifuged at maximum speed in an Eppendorf refrigerated microfuge (model 5415R). The supernatant was removed, and the pellet washed in 1 mL sterile deionised water. The bacterial suspension was centrifuged as before. The supernatant was again removed and the pellet resuspended in 500 µL sterile deionised water. The suspension was then heated on a heating block to 100°C and maintained at this temperature for 10 min. The suspension was again centrifuged as before; a 250 µL fraction of the supernatant was removed taking care not to disturb the pellet. The 250 µL fraction was then used as crude template genomic DNA.

### **2.6.2 Cleavage of DNA with restriction enzymes**

Purified DNA was digested with restriction enzymes obtained from either Promega or New England Biolabs. Digests were carried out according to manufacturer's instructions. Cleaved vector DNA for use in non-directional cloning approaches was treated with Calf-intestinal alkaline phosphatase (Promega) immediately following digest to prevent re-annealing.

### **2.6.3 Agarose gel electrophoresis**

Plasmids, DNA restriction fragments and PCR products were analysed by agarose gel electrophoresis according to [Sambrook et al. \(1989\)](#). All samples and markers (1 kb ladder, Promega) were loaded onto the gel using 5x DNA loading buffer (10x TAE, 20% (w/v) Glycerol, 0.05% Bromophenol Blue). 1% agarose gels were prepared in TAE buffer containing 10 µg mL<sup>-1</sup> ethidium bromide (Sigma), and electrophoresis performed using 1x TAE buffer in horizontal tanks (Anachem) at 60-90 V. DNA fragments were then

visualised using either a UV transluminator (Anachem) or DNA visualisation equipment (Uvitec Gel Doc System DOC-CF08.XD).

#### **2.6.4 Amplification of DNA using the Polymerase Chain Reaction (PCR)**

The polymerase chain reaction ([Saiki et al, 1988](#)) was used for the amplification of specific regions of DNA. Primers for amplification of *espC* domains and *ybgC* (Sigma Genosys), and for mutagenesis (MGW Biotech) were designed containing restriction enzymes as required (Table 2.3). To optimise the PCR amplifications, a variety of conditions were altered including (I) annealing temperature, (II) addition of dimethylsulphoxide (DMSO) to minimise DNA secondary structure ([Chakrabarti and Schutt, 2001](#)), (III) addition of bovine serum albumin (BSA), (IV) alteration of magnesium concentration, and (V) glycerol. Either crude or pure genomic DNA isolated from EPEC E2348/69 was used as template, together with appropriate primers to amplify specific *espC* or *ybgC* fragments (Table 2.3).

100 µL reaction mixtures contained 5 µL of each primer at concentration 10 pmol, 16 µL dNTPs (dATP, dTTP, dGTP, dCTP)(Promega) at 2.5 mM concentration, 2 µL template DNA, and 5 µL DMSO (dimethylsulphoxide, Sigma), 1 µL hi-fidelity proof reading enzyme and the correct dilution of 10X enzyme buffer was used. DNA polymerase was added to the reaction after a 5 min hot start at 95°C. A typical amplification involved 30 cycles of template denaturation at 95°C for 30 s, followed by annealing of primers to template at 55°C for 30 s, then extension at 68°C (generally 1 min kb<sup>-1</sup> DNA). The run was finished with an extension at 68°C for 10 min to ensure all strands were completed before being cooled to 4°C until collected. Two proof-reading polymerases were utilised, either Deep Vent (New England Biolabs) or KOD Hotstart (Merck Biosciences), the latter requiring an extension temperature of 68°C. KOD Hotstart was used only in amplification and cloning of the passenger domain. In addition, GoTaq polymerase (Promega) was used in PCR mutagenesis reactions. Mutagenesis conditions were determined according to the QuikChange XL II Site Directed Mutagenesis Kit protocol (Stratagene).

#### **2.6.5. Colony PCR**

A bacterial colony was picked with an sterile toothpick and swirled in 100 µL of distilled deionised water in an autoclaved microfuge tube and briefly vortexed, and the area was marked from where the colony was picked (bottom of the plate). The colony suspension

was heated in a boiling water bath (100°C) for 5 min, then microfuged at 13000 rpm for 5 min. 80 µL of the supernatant was then transferred in to a new microfuge tube. 1 to 2 µL of the supernatant was taken for use as template DNA in a 20 µL PCR standard PCR reaction.

#### **2.6.6. Purification of DNA**

PCR amplified DNA fragments were cleaned using a Qiagen PCR purification kit and digested with the relevant enzymes. Digested fragments were run on an agarose gel as described above and the correct band was excised and DNA extracted using a Qiagen Gel Extraction Kit according to manufacturers' instructions. DNA was eluted in 50 µL sterile distilled water. Plasmids were obtained from overnight cultures using a QIAprep Spin Miniprep Kit (Qiagen).

#### **2.6.7. Calf intestinal alkaline phosphatase**

Dephosphorylation of plasmid DNA was carried out using calf intestinal alkaline phosphatase (CIAP) (Promega, UK). The enzyme was used at 1:100 dilution in the supplied buffer and added in two 0.5 µL aliquots to the digested plasmid mixture, incubating for 30 min at 37°C after each addition. The dephosphorylation reaction was stopped by heating to 65°C for 15 min.

#### **2.6.8. Ligation of DNA fragments**

T4 DNA ligase was used to catalyse the ligation of DNA ends by forming a phosphodiester bond between the 3' –OH group and one of the 5' –phosphate groups of another end ([Weiss et al, 1968](#)). Ligations were set up in a reaction volume of 20 µL or 30 µL (for large ratios). Typically, the molar ratio of fragment: vector used was 3:1, 6:1 and 10:1, using 1 U T4 DNA ligase (1 Unit µL<sup>-1</sup>, Promega) and 1x ligase buffer (Promega). Ligations were incubated at ~16°C overnight. Prior to electroporation of *E. coli* strains, the DNA to be introduced into the cells was dialysed to remove excess salt. Dialysis was achieved by pipetting the DNA onto a 0.025 µm Ultra filtration disc (Millipore), which was floating on distilled water, for 20 min at room temperature.

## **2.6.9. Electroporation and Chemical Transformation**

### **2.6.9.1. Electroporation**

Electrocompetent *E. coli* cells were prepared according to the protocol described by [Sambook et al, \(1989\)](#). Briefly, cells were grown to an OD<sub>600</sub> of between 0.4 - 0.5, cells were then pelleted at 4°C, 10,000 rpm for 10 min (Beckman Avanti™ 30) and washed five times in ice cold 10% glycerol. Finally the cells were resuspended in 1 in 500 of their original volume; 50 µL aliquots were stored at -80°C until required.

Plasmids were introduced into electro-competent *E. coli* cells by electroporation as described originally by [Dower et al. \(1988\)](#). An electroporator (Biorad Gene Pulser) was used with 2mm pre-chilled cuvettes. 60 µL electroporation mix (*E. coli* and plasmid DNA construct) was added to individual cuvettes which were then pulsed at 1.9 kV, 200 Ω, 25 µF. 940 µL warm sterile LB medium was added immediately after the pulse to recover the transformed cells. Following the addition of fresh medium cells were incubated at 37°C (static) for 1 h before being plated onto antibiotic selective agar plates to isolate cells containing recombinant DNA.

### **2.6.9.2 Rubidium chloride transformation**

Plasmids were introduced into chemically competent cells made initially according to the method described by [Inoue et al \(1990\)](#). However, a method using Rubidium chloride ([Kushner, 1978](#)) was employed subsequently to make all competent cells in *E. coli*, and this method was found to be more effective. All competent cells were stored as aliquots of 100 to 200 µL for no more than 3 months at -80°C.

## **2.6.10. Screening of transformants.**

### **2.6.10.1 Colony Screening**

Individual colonies were streaked onto fresh LA plates and grown in LB broth overnight containing the appropriate antibiotic for selection. Plasmids were recovered as described in section 2.5.6. Plasmids were then digested with appropriate restriction enzymes and the resulting fragments visualised by UV irradiation following separation through a 1%

TAE/agarose gel. Plasmid constructs were also confirmed by PCR either directly or by colony PCR to determine the orientation of the insert where the same restriction sites had been used in the forward and reverse primers. Constructs were confirmed by sequencing.

Plasmids pSTBlue1 and pCR2.1 contain a polylinker inserted into the reading frame of *lacZ*, enabling blue/white colony selection. Where the insert is correctly ligated into the polylinker, the *lacZ* gene will be disrupted and the colony will be white. Where no insert is incorporated into the plasmid, the *lacZ* gene is left intact and it is able to produce β-galactosidase and the colony will turn blue. Blue/white selection was performed on LB agar plates containing 50 µg mL<sup>-1</sup> Carb<sup>R</sup>, and 12.5 µg mL<sup>-1</sup> Tet<sup>R</sup>, supplemented with Isopropyl β-D-Thiogalactopyranoside (IPTG) at 80 µg mL<sup>-1</sup> and 5-Bromo-4-chloro-3indolyl β-D-galactopyranoside (X-Gal) at 70 µg mL<sup>-1</sup>. White colonies picked were subjected to restriction digests, to check for insert. Ligations were transformed into Novablue competent cells (Merck Biosciences). .

### **2.6.11 DNA Sequencing**

The majority of sequencing reactions were performed using primers designed from plasmid sequence immediately flanking the multiple cloning sites into which insert had been cloned. For large insert sequences such as the passenger domain, internal sequences in the domain were also used (Table 2.2). Sequencing reactions were used to verify that no base changes had been introduced during PCR amplification. The sequencing reactions were performed by Geneservice (Cambridge, UK). Sequence files were formatted for use with the sequence editing program Chromas enabling analysis of sequence data in either Chromas or the BioEdit suite of programs.

### **2.6.12. Mutagenesis**

#### **2.6.12.1. λ Red Recombination – making an *espC* deletion Mutant**

The *espC* gene was replaced with the kanamycin/*espC* flanking region PCR product according to the protocol described in [Murphy et al, 2003](#). Primers designated EspCMutF and EspCMutR (Table 2.3), incorporating the flanking EspC N-terminus and C-terminus regions, were used to generate an in-frame deletion of EspC.

In brief, initially a kanamycin resistance cassette (from the plasmid pKD4) incorporating *espC* flanking regions was amplified; this was purified, digested with *DpnI* (New England Biolabs), and separated by agarose gel electrophoresis. The *DpnI* digested product was excised, purified and electroporated into EPEC E2348/69 hosting plasmid pKD46 ([Datsenko and Wanner, 2000](#)). Transformants were recovered on 25 µg mL<sup>-1</sup> kanamycin plates at 37°C for 2 h, and gene deletion verified by PCR. All PCR reaction conditions were modified as necessary to achieve amplification.

#### **2.6.12.2. Site-directed mutagenesis**

Experiments were carried out according to the principles and protocol described in the QuikChange® II XL Site-Directed Mutagenesis Kit Instruction manual (Stratagene). The primers (I) PheMut\_For and PheMut\_Rev and (II) 3AlaMut\_For and 3AlaMut\_Rev were designed around the target region and the minimum number of residues altered. A codon usage web site ([gcua.schoedl.de](http://gcua.schoedl.de)) was utilised to compare the target gene to wild type genomic sequence to ensure successful amino acid substitution.. Mutagenesis was confirmed by full sequencing of the *espC* gene. Once verified, the plasmids were transformed into wild type EPEC and the *espC* EPEC mutant for further study of AT export.

#### **2.6.13. The yeast two-hybrid (Y2H) system**

The two-hybrid system is a method of detecting protein-protein interactions in budding yeast cells (*S. cerevisiae*) with a high degree of sensitivity and specificity. It was originally described in late 1980's-early 90's ([Fields and Song, 1989](#) and [Chien et al, 1991](#)) and developed on the back of earlier work on the modular nature of eukaryotic transcription factors by [Ma and Ptashne \(1988\)](#).

##### **2.6.13.1 Generation of yeast-two hybrid Gal4 fusion constructs**

Oligonucleotide primers were designed to amplify *espC* domain fragments from EPEC E2348/69 genomic DNA which were subsequently purified and cloned to the *EcoRI* site of yeast two-hybrid vectors pGBT9 (harbouring the yeast *TRP1* marker gene) and pGAD424 (harbouring the yeast *LEU2* marker gene). A PCR insert encoding the putative accessory factor YbgC was also cloned to both vectors. A third plasmid pSTBlue1 (Merck Biosciences) was employed as an intermediate cloning vehicle for the

*espC* fragment encoding the passenger domain and also the *ybgC* gene, both of which proved difficult to clone directly to the yeast vectors. pSTBlue1 inserts were excised with *Eco*RI for subsequent cloning into the yeast two-hybrid vectors, creating fusions to the binding and activation domains of the yeast transcriptional activator *GAL4* respectively.

#### **2.6.13.2 Lithium acetate transformation of *S. cerevisiae***

pGBT9 ‘bait’ construct and pGAD424 ‘prey’ constructs were introduced simultaneously into *S. cerevisiae* strain PJ69-4A ([James et al, 1996](#)) using the high-efficiency lithium acetate transformation procedure described by [Geitz & Schiestl, \(1995\)](#). Co-transformants were replica-plated according to the protocol of [Lederberg and Lederberg, \(1952\)](#). The replica plater and velveteen squares were supplied by Sigma.

#### **2.6.13.3 β-galactosidase assay**

Activation of the *lacZ* reporter gene in yeast strains harbouring pairs of interacting constructs were further assessed by quantification of β-galactosidase activity in cell extracts according to the liquid assay method of [Miller, \(1972\)](#) using *o*-nitrophenyl β -D-galactopyranoside (ONPG) as substrate.

#### **2.6.13.4 Statistical analysis of β-galactosidase activity**

All data are reported as mean ± standard deviation (SD). The results were compared statistically by using Student’s test for unpaired data. Significance was indicated by a two-tailed *P*-value of less than 0.05.

### **2.6.14. Construction of pMALC2X and pET30a over-expression vectors.**

Standard methods for the manipulation of DNA were adhered to ([Sambrook et al, 1989](#)). Briefly, over-expression vectors encoding EspC domains were constructed as follows; pGAD424 or pGBT9 harbouring cloned PCR domains (passenger-, inter- or passenger sub-domain) were excised by *Eco*RI restriction digest, the DNA inserts were then agarose gel purified and then ligated, using a 6:1 insert:vector ratio overnight into linearised (*Eco*RI digested), dephosphorylated (alkaline phosphatase (CIAP) treated) and gel purified over-expression vectors (I) pMALC2X or (II) pET30a.

pET30a harbouring either (I) *espB* or (II) *fliC*, were constructed using standard methods (Sambrook *et al*, 1989). EspB sequence (Accession AAC38396.1) in the LEE region of EPEC strain E2348/69 (Accession AF297061, Elliot *et al*, 1998) was obtained from the NCBI database ([www.ncbi.gov](http://www.ncbi.gov)), and the *espB* gene sequence bases 32115..33080 saved. FliC sequence (Accession AAC74990.1) in *E. coli* K12, strain MG1655 (Accession AE000285, Blattner *et al*, 1997) was also obtained from the NCBI database and used to design primers. The *fliC* gene sequence bases 1711..3207 (complement) and surrounding bases were also saved. The primers EspB\_For, EspB\_Rev and FliC\_For, FliC\_Rev (seeTable 2.2) were used to amplify *espB* and *fliC* respectively. and included the restriction sites *Bam*H/*Nco*I and *Eco*R1/*Hind*III respectively.

## 2.7. Protein analysis and purification.

### 2.7.1 Preparation of *E. coli* secreted protein samples.

Bacteria from a 10 mL overnight LB culture were harvested by centrifugation for 10 min at 4000 g, RT (using a Jouan Centrifuge B4i CE0122). The pellet was resuspended in 10 mL of fresh, pre-warmed (37°C) LB broth containing 0.2% glucose. Cells were grown to an OD<sub>600</sub> between 0.3-0.4. 1.75 mL was used to inoculate a fresh 35 mL culture of the same medium containing the appropriate levels of antibiotic and 0.02% (w/v) arabinose. Typically samples of culture were taken at 0, 30, 60, 120 and 180 min, and a more defined time range of 0, 20, 40 and 60 min. 1.5 mL aliquot taken and the OD<sub>600</sub> recorded. Bacteria were pelleted in a microfuge for 2 min and 1 mL of supernatant was taken. Secreted proteins were concentrated by adding 0.2 mL trichloroacetic acid and stored at 4°C overnight. The precipitated secreted proteins were spun at maximum speed in a refrigerated microfuge (Eppendorf Refrigerated Microfuge 5415R), the supernatant was decanted and remaining TCA solution removed carefully by pipette, taking care not to disturb the pellet. The pellet was washed by re-suspending in 500 µL of ice cold acetone (stored at -20°C) and re-pelleted as previously in the microfuge. The pellet was resuspended in 50 mM Tris-HCl pH 8.0 to ensure neutralisation of any remaining TCA. The volume resuspended in was based on the OD<sub>600</sub> previously recorded e.g. pellet resuspended in 100 µL if the OD<sub>600</sub> = 1.0. Where pellets did not easily resuspend, samples were briefly sonicated at amplitude 5 µm for several seconds. Samples were stored at -20°C until required.

### **2.7.2. Preparation of *E. coli* whole cell lysates**

The OD<sub>600</sub> of broth cultures taken and typically 5 mL removed from the culture and used to prepare secreted proteins,. Samples were centrifuged at 4000 rpm for 15 min (Jouan Centrifuge B4i CE0122). Supernatants were removed and cell pellets frozen at -20°C until required. Cells were resuspended in a volume of phosphate buffer saline (PBS) determined by the OD<sub>600</sub> and volume of culture taken to give a standardised OD<sub>600</sub> of 1. Cells were sonicated, using an MSE Soniprep 150, for a maximum of 3 min in 30 s bursts on ice. The sonicated bacterial cells were then centrifuged at 4000 rpm for 5 min to remove larger cell debris. 5x SDS-PAGE sample buffer was added to sonicated samples to make a final volume of 100 µL. The samples were boiled for 10 min.

### **2.7.3. SDS- PAGE**

SDS-PAGE was carried out as previously described by Laemmli, (1970). 10-20 µL of sample was loaded on to an 11.5% (unless otherwise stated) polyacrylamide gel and proteins separated using a Biorad Minigel system under constant voltage (per gel) of 125 V through the stacking gel, and 150 V through the resolving gel. Whole cell lysates were equalised for cell number by dilution before application (section 2.6.2). Proteins were either visualised by Coomassie stain (0.1% (w/v) Coomassie brilliant Blue, 45% (v/v) methanol, 9% (v/v) glacial acetic acid, 0.2 µm filtered) for 60 min, followed by washing with 10% (v/v) methanol – 7% (v/v) glacial acetic acid, or transferred on to nitrocellulose membrane (see below). 5 µL Pre-stained markers (Precision Plus®, Biorad) were used to determine molecular weight, separation on SDS-PAGE gels and demonstrate transfer in western blots.

### **2.7.4 Western blot analysis**

Protein separation was achieved by passage through an SDS-PAGE gel (as described above), proteins were then transferred to 0.2 µm, nitrocellulose blotting membrane (Protan® Schliecher and Schuell, Scientific Laboratory Supplies), pre-soaked in transfer buffer using constant current of 80 mA overnight. Western blot membranes were blocked in PBST (PBS with 0.5% Tween 20) containing 1% (w/v) BSA (Bovine Serum Albumin (Sigma) for 1 h shaking at room temperature (RT). Primary antibodies (Table 2.6) were diluted in 1% BSA/PBST and incubated either for 3 h shaking at RT or

alternatively overnight. The membrane was subjected to 5 x 3 min washes in PBST before being incubated for 2 h in secondary antibody, Protein A alkaline phosphatase conjugate (Calbiochem), diluted in 1% BSA/PBST. Prior to development, the membrane was washed again, 5x 3 min in PBST followed by 2x 1 min in dH<sub>2</sub>O (sterile de-ionised water). The membrane was finally developed using 5-bromo-4-chloro-3-indolyl-phosphate (BCIP)/nitroblue tetrazolium salt (NBT) according to manufacturers' instructions. Antibody concentrations utilised are listed in Table 2.6.

Blotting was also conducted in an alternative manner. Separation by SDS-PAGE was performed as described in section 2.7.3. Protein was then transferred onto Hybond-P Poly-vinylidene-difluoride (PVDF) membrane (Amersham Life Sciences) in 3-[cyclohexylamino]-1-propanesulphonic acid (CAPS) buffer, pH 11 carried out at 80 mA overnight. Prior to blotting PVDF membrane was activated according to the manufacturers' protocol, by immersing the PVDF membrane in methanol.

#### **2.7.5. Electro-elution**

The Biorad electro-elution apparatus was set up according to the manufacturer's instructions. Electro-elution was used solely to purify EspC protein for the purpose of generating a Rabbit anti-EspC polyclonal antibody.

#### **2.7.6. Mass spectrometry identification of proteins**

Proteins were resolved on 11.5% SDS-PAGE gels run at 150 V (constant voltage), visualised with 0.1% (w/v) Coomassie blue (in 45.4% methanol, 4.6% acetic acid, 0.45 µm filtered) for 1 h, and de-stained with sterile distilled water until bands visible. Target proteins were excised from the gel and subjected to an enzymatic trypsin digest, before being applied to a MALDI-TOF mass spectrometer, carried out by Kevin Bailey at the Biopolymer Synthesis and Analytical Unit (BSAU, Queen's Medical Centre, Nottingham).

#### **2.7.7. Maltose binding protein (MBP) and His-tag purifications**

Induced bacterial cultures of either (I) *E. coli* XL1Blue harboring plasmid pMALC2X or pMALC2X constructs (passenger- or inter-domain MBP fusions) or (II) *E. coli* BL21DE3 harboring pET30a or pET30a constructs (passenger- or inter-domain his-tag fusions), were lysed by sonication using an MSE Soniprep 150, for a maximum of 3 min in 30 s

bursts on ice. After centrifugation, the soluble fraction (lysate) was loaded onto the appropriate columns and the fusion proteins purified according to manufacturers instructions (New England Biolabs and Merck Biosciences respectively).

Briefly, clarified lysate (0.22 µm filtered) was passed through 0.5 mL washed MBP resin (New England Biolabs) contained in 10 mL disposable columns (Sigma). The resin was washed with 30 mL MBP column buffer. At this stage the MBP fusion was either eluted with MBP buffer / 10 mM maltose or the second clarified lysate from the *E. coli* BL21(DE3) containing one of the pET30a plasmid constructs was added to the column to assess the protein interaction as below. His-tagged proteins were recovered by eluting with binding buffer containing 250 mM imidazole. Protein concentration was estimated using Bradford reagent (Sigma) and bovine serum albumin as a standard. Protein purity was assessed by SDS-PAGE.

#### **2.7.8. MBP column co-purification**

Immobilized MBP-tagged fusion protein (comprising inter- or passenger-domains) was also used in protein-protein interaction experiments. Following binding of the MBP-fusion protein to the maltose resin, and the subsequent wash steps (described in manufacturers instructions and briefly above), lysates of either His-tagged passenger- or inter-domain (5-20 mL broths), also prepared in MBP column buffer, were washed through the column. Following a further wash step, interacting proteins were eluted in 5 mL elution buffer and collected in 500 µL fractions in separate tubes. Samples of the pre-purified MBP fusion, column flow through, His-tagged protein applied to the column and column flow through were retained for analysis. Co-elution of interacting proteins was determined by SDS-PAGE and Western immunoblot using. (I) rabbit anti-MBP (New England Biolabs) (II) rabbit anti-His (Santa Cruz Biotechnology) or (III) rabbit anti-EspC antiserum (Table 2.6) as appropriate.

#### **2.8. Polyclonal antibody generation**

Rabbit polyclonal antibodies were raised according to a protocol based on [Harlow and Lane, \(1988\)](#). Animal handling, immunisations, and test bleeds were conducted by the Bio-Medical Services Unit (BMSU) at the Queen's Medical Centre, Nottingham. NZ

White rabbit serum was incubated overnight at 4°C, centrifuged at 3000 g for 15 min to remove residual red blood cells, then serum removed and stored at –20°C until required. Specificity of antibodies was determined by immunoblot.

Table 2.6: Antibody concentrations, targets and sources

Primary antibody	Antigen target	Antibody source	Protein-A AP conjugate concentration	Notes on primary antibody
Anti- EspC (1: 1000)	EspC	Rabbit	1: 1000	Generated with the help of the BMSU and Sue Dodson using electro-eluted EspC
Anti- EspB (1: 1000)	EspB	Rabbit	1: 1000	Generated with the aid of the BMSU using purified His-EspB (this study).
				Utilised in published work of <a href="#">Ferrández et al (2005)</a> .
Anti- FliC (1: 3000)	FliC	Rabbit	1: 1000	Generated with the help of the BMSU using purified His-FliC (this study).
				Utilised in published work of <a href="#">Ferrández et al (2005)</a> .
Anti - MBP (1:1000)	Maltose binding protein	Rabbit	1: 1000	New England BioLabs Cat No: E8030S
Anti – His (1:1000)	His-Tag	Rabbit	1: 1000	Santa Cruz Biotechnology Cat No: sc-803

### **3. *In silico*/bioinformatics analyses of EspC and YbgC.**

#### **3.1 Introduction**

As a first step towards identifying potential co-factor involvement in EspC secretion, we intended to investigate direct protein-protein interactions between EspC and YbgC, and also EspC domain-domain interactions that might facilitate autotransport. To this end, an initial in depth bioinformatics analysis of the structural features of EspC and YbgC was undertaken in order to first define relevant domains and their boundaries.

#### **3.2 Identification of the predicted structural/functional domains of the AT EspC**

EspC and homologous protein sequences (Table 3.1) were subjected to *in silico* analyses as described in section 2.4.2. A general domain structure of the EspC AT was successfully established (Fig 3.1). As expected, SignalP (Table 2.4) revealed a Sec-dependent signal peptide at the N-terminus of EspC analogous to the longer sequences found on other serine protease ATs (see reviews by [Desvaux et al, \(2004\)](#) and [Henderson et al, \(2004\)](#)). The C-terminal portion of the protein showed similarity to other ATs predicting formation of the signature  $\beta$ -barrel required for OM translocation. A PSI-BLAST search with the remaining central part of EspC (passenger domain) revealed a section with homology to IgA1 protease. [Oliver et al, \(2003a\)](#) have shown that the region between the  $\beta$ -barrel domain and IgA1 protease like domain (termed inter-domain hereafter) may be a chaperone region which prevents premature folding of the passenger prior to secretion. A number of other structural features have been highlighted in the alignment (Fig 3.1) including (I) the cleavage site between the N-terminal domain and passenger-domain, (II) a serine protease catalytic site (GDSGS) ([Dutta et al, 2002](#)), and (III) the experimentally determined cleavage site of the inter-/ $\beta$ -domains ([Barnard et al, 2007](#)).

#### **3.3 Primer design for PCR amplification of *espC*-encoded domains**

Oligonucleotide primers were designed to amplify the three of the defined domains of EspC including (I) the passenger-domain, (II) the putative autochaperone domain (inter-domain) and, (III) the  $\beta$ -domain (see Fig 3.1, Fig 3.2A). The signal peptide was excluded as it is cleaved once targeted to the IM.

Table 3.1: Identification of homologous proteins utilising a PSI-BLAST search of EspC whole sequence. Information was obtained from the NCBI database and relevant literature and pI values were obtained from analysis in EDITSEQ of DNASTAR

AT	Molecular Weight kDa	Length (amino acids)	Identity (%)	pI	Function	Accession Number	E-value	Organism	Reference
<b>EspC</b>	140.8	1305	100	5.53	Enterotoxin Haem binding protein Haemoglobin protease	Q9EZE7	0.0	<i>E. coli</i>	Mellies <i>et al</i> , 2001 Drago-Serrano <i>et al</i> , 2006
<b>Pet</b>	139.8	1295	51.8	6.76	Proteolytic toxin	O68900	0.0	<i>E. coli</i>	Eslava <i>et al</i> , 1998
<b>Sat</b>	140.0	1295	50.3	6.02	Vacuolating Cytotoxin	Q8FDW4	0.0	<i>E. coli</i>	Guyer <i>et al</i> , 2000
<b>EpeA</b>	141.7	1300	51.8	6.88	Protease Mucinase	YP_308812	0.0	<i>E. coli</i>	Leyton <i>et al</i> , 2003.
<b>SigA</b>	139.8	1285	53.0	8.19	Unknown	NP_708742	0.0	<i>Shigella flexneri</i>	Al-Hasani <i>et al</i> , 2000
<b>SepA</b>	146.0	1364	52.2	6.35	Protease Inflammation Invagination	NP_858203	0.0	<i>Shigella flexneri</i>	Benjelloun-Touimi <i>et al</i> , 1995
<b>AidA</b>	147.2	1366	47.9	6.28	Adhesin	ZP_00714135	0.0	<i>E. coli</i>	Rasko <i>et al</i> , 2005 (NCBI)
<b>Eata</b>	147.8	1364	46.5	6.54	Enterotoxin	Q84GK0	0.0	<i>E. coli</i>	Patel <i>et al</i> , 2004
<b>Tsh</b>	148.2	1377	33.8	5.49	Haemagglutinin Haemoglobin binding	Q47692	0.0	<i>E. coli</i>	Provence and Curtis <sup>3rd</sup> , 1994
<b>Hbp</b>	148.3	1295	33.8	5.55	Haemoglobin protease	O88093	0.0	<i>E. coli</i>	Otto <i>et al</i> , 1998
<b>Vat</b>	148.3	1377	33.5	6.18	Vacuolating Cytotoxin	AAO21903	1e-156	<i>E. coli</i>	Parreira and Gyles, 2003
<b>Boa</b>	149.3	1384	34.8	6.13	Unknown	AAW66606	1e-179	<i>Salmonella bongorii</i>	Henderson, 2005 (NCBI)
<b>Pic</b>	146.6	1371	39.6	6.23	Protease ?	NP_752289	0.0	<i>E. coli</i>	Welch <i>et al</i> , 2002

The figure displays two sequence logos for the Signal Sequence and Passenger Domain. The top logo, titled "Signal Sequence", shows conservation across positions 10 to 80. The bottom logo, titled "Passenger Domain", shows conservation across positions 90 to 160. Both logos use a color scheme where green indicates conservation, yellow indicates variability, and red indicates high variability. A vertical arrow points from the top logo to the bottom logo, indicating the transition from signal sequence to passenger domain.



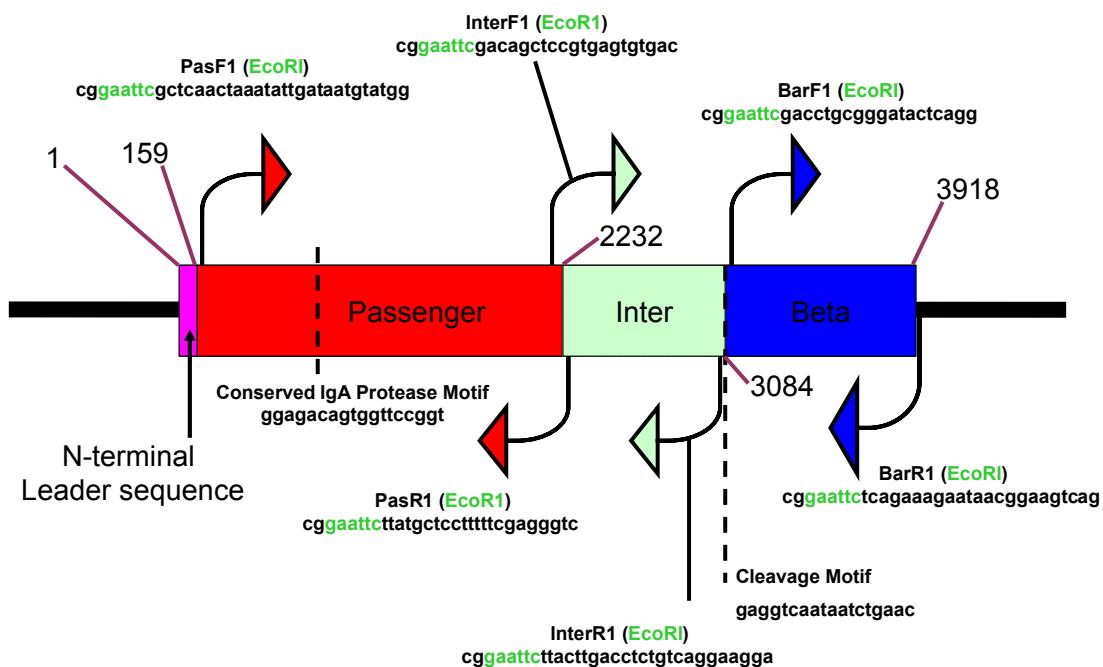


### **AT β-domain/helper domain**

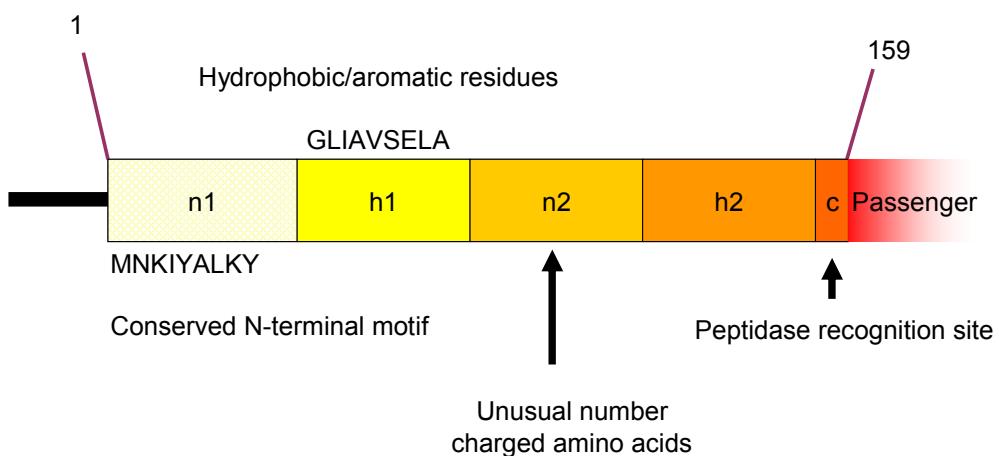
## Inter domain/β-domain cleavage site

Gor	IV	.....	eeeee	1305
EspC			FRYSF	1305
Pet			FRYSF	1295
Sat			FRYSF	1295
EpeA			FRYSF	1300
SigA			FRYSF	1285
SepA			FRYVF	1364
AidA			I RYYF	1366
EatA			FRYYF	1364
Tsh			I RYSF	1377
Hbp			I RYSF	1377
Vat			I RYSF	1377
Boa			FRYMF	1384
Pic			FRYMF	1371
Clustal	:	** *		410

Figure 3.1: ClustalW alignment of autotransporter proteins. EspC is shown aligned against homologous serine ATs identified by PSI-BLAST searches. Secondary structure analysis of EspC using GoriV is also shown (Alpha Helix (h) 19.54%, Extended Strand (e) 25.44%, and Random Coil (c), 55.02%). Four domains have been identified in the protein sequence; I) Extended signal sequence (pink) II) passenger domain (red), with the first amino acid of the mature protein highlighted in yellow III) a putative chaperone region or Inter domain (green) and IV) AT  $\beta$ -domain (blue). The IgA protease consensus catalytic site (GDSGSG) and experimentally determined cleavage site of the AT (EVN..NLN) ([Barnard et al, 2007](#)) have been highlighted in green. The ClustalW consensus indicates the level of homology across the serine protease ATs, with increasing levels of homology indicated between inter- and  $\beta$ -domains.



A: Schematic EspC Domain structure



B: Schematic of N-terminal leader sequence

Figure 3.2: Schematic representation of *espC*. *espC* has been divided into four domains. (A) In pink is the N-terminal leader sequence (motif structure shown in more detail in B); in red the secreted passenger which shows homology with Immunoglobulin A1 (IgA1) protease of *Neisseria meningitidis*; in blue is the  $\beta$ -barrel domain, where identity is shown with presumed integral AT  $\beta$ -domains; and finally in green the putative autochaperone (based on work by Oliver et al, 2003) we have called this the 'inter' domain. In (A) relative primer locations and DNA sequences (5' to 3') are indicated (*espC* pathogenicity island Accession: AF297061, obtained from NCBI). Primer names and restriction sites have been incorporated into the primer and are shown in green. Other features highlighted in the schematic by a dashed line, denote the IgA protease consensus catalytic site within the passenger domain (non-functional in serine protease ATs), and secondly the inter-domain/ $\beta$ -domain cleavage site (Oliver et al, 2003; Barnard et al, 2007). (B) A schematic of the N-terminal leader sequence highlights the main features of the signal peptide, including the conserved N-terminal motif, hydrophobic/aromatic residues and the high number of charged amino acids in domain n2.

### **3.4. Analysis of homologous serine protease ATs**

As a measure of global relatedness, a hydrophobicity plot using The Kyte and Doolittle Scale was initially generated for each of the established AT domains using BIOEDIT (Fig 3.3, 3.4, 3.5, 3.6). EspC and all homologous serine protease ATs domains selected had similar hydrophobicity profiles despite their disparate sequence similarities, showing broad conservation of underlying physical property and structure. The signal peptide was found to have a highly conserved hydrophobicity profile closest to the N-terminus, and more divergent profile adjacent to the passenger domain (corresponding to the n2 and h2 domains of the extended serine protease signal peptides). The hydrophobicity profiles of passenger-domains were similar but not identical whilst both the inter-, and  $\beta$ -domain hydrophobicity profiles, were almost indistinguishable.

Having established domain boundaries for EspC, the individual domains were subjected to a second round of PSI-BLAST searches against the NCBI non-redundant database to establish the level of similarity between individual EspC domains and other selected serine protease ATs. High similarity between inter- and  $\beta$ - domains comensurate with their role in export was evident (E-values shown in Tables 3.2, 3.3, 3.4 and 3.5). The sequence variability within the passenger domains reflects the evolution of functional diversity, though even highly similar ATs can have distinct substrate specificities despite being nearly 100% identical ([Henderson et al, 2004](#)).

Structural analysis of the signal peptides indicate they are (I) 5.1 to 6.6 kDa in size, (II) basic in nature and (III) have a sequence identity to EspC signal peptide ranging from 32.7 to 52.7%. GoriV analysis of the peptides shows wide variance in the composition of various structural features which includes high proportion of residues forming  $\alpha$ -helices in EspC contrasting with an absence of  $\alpha$ -helices in both EpeA and SigA. The level of amino acids forming extended stand (or  $\beta$ -strand) also varies considerably, under 10% in EspC, Hbp, Tsh and SepA, but greater than 38% in both SigA and EpeA. Significant random coiling is present suggesting that the signal region for many ATs may be largely unstructured.

EspC and its homologues all had passenger-domains of equivalent size (71.7 to 81.3 kDa) with identity amongst the selected sequences ranging from 22.7 to 52.2%. GoriV analysis revealed that there was a similar proportion of  $\alpha$ -helices among the

homologous proteins with the exception of Vat, that had approximately half that of the others. All had relatively similar ratios of extended or  $\beta$ -strand and random coiling (see Table 3.3). The inter-domains were of similar size (~29.8 to 31.9 kDa), with either weakly or mildly acidic pI values. The level of identity was also generally higher than that found amongst both signal peptide and passenger domains ranging from 30.7% to 85.3% (see Table 3.4). All AT  $\beta$ -domains showed >60% sequence identity to EspC, and >90% in the case of SepA. The  $\beta$ -domains were of uniform length, 277 amino acids (~30 kDa), with a slightly wider diversity of pI values in contrast to the inter-domain. It is thought that this might reflect the different environments into which the passenger-domain is secreted (see Table 3.5).

Phylogenetic trees allow the degree of evolutionary relatedness between protein or nucleotide sequences to be examined. Therefore, to gain some indication of the sequence (and by inference, structural) similarity of different ATs and their component domains, full length AT sequence and estimated sub-domains comprising signal peptide, passenger-, inter- and  $\beta$ -domains regions were aligned and phylogenetic relationships subsequently determined using CLUSTALW (Fig 3.7, 3.8 and 3.9).

Phylogenetic analyses of the individual domains revealed some noteworthy differences and similarities from that of the full sequence analysis. In the full sequence analysis, EspC was shown not to cluster closely with any other AT suggesting evolutionary divergence of EspC domains. However, AT domains are functionally discrete, and with the exception of the passenger-domain are considered to perform similar functions in export. Consistent with this, phylogenetic analysis of individual domain sequences revealed more defined family groups of ATs. Both the inter- and  $\beta$ -domain sequences of EspC (Fig 3.9A and Fig 3.9B respectively) clustered with those of SepA, AidA and EatA. Interestingly however, both the signal peptide and passenger-domains (Fig 3.8A and Fig 3.8B respectively) clustered with a different subset comprising Pet, Sat and Pic (signal) or SigA (passenger). This contrasts with the more stable family sub-groupings of the other ATs, whereby the same family groups are still evident regardless of which domain is used in the analysis. This suggests that EspC may have a mosaic-like domain composition, having acquired different domains from different AT subsets.

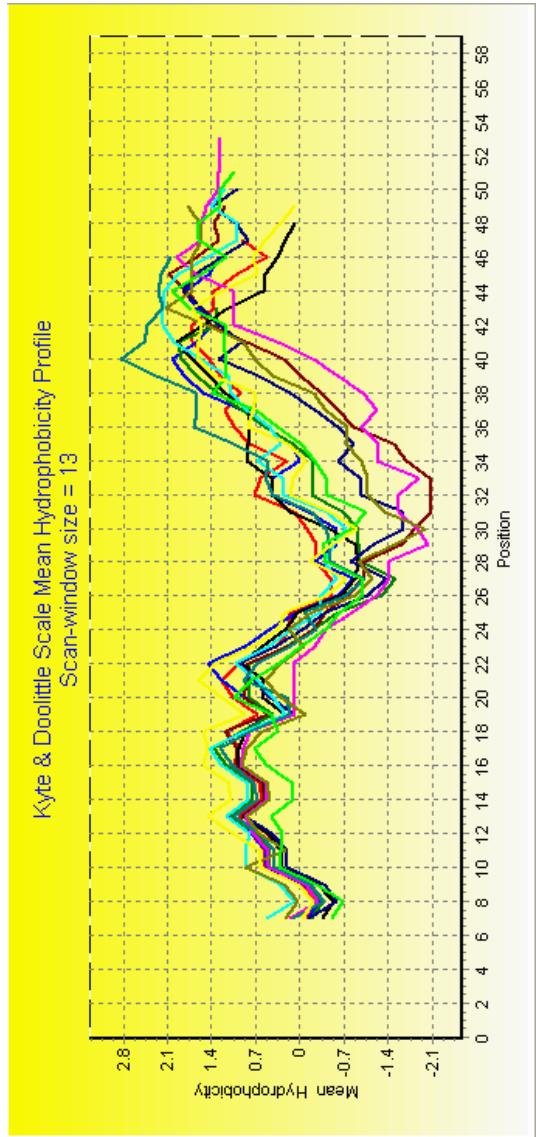


Figure 3.3 Hydrophilicity profile analysis of EspC and homologous protein signal peptides using the Kyte and Doolittle Scale Mean Hydrophobicity Profile (BioEdit analysis). In red, EspC (EPEC), in blue Pet (*E.coli*), in green Sat (*E.coli*), in brown EpeA (*E.coli*), in black SigA (*Shigella flexerni*), in dark blue SepA (*S. flexerni*), in pink AidA (*E.coli*), in light blue EatA (*E.coli*) in purple Tsh (*E.coli*), in turquoise Hbp (*E.coli*), in olive Vat (*E.coli*), in light green Boa (*Salmonella bongorii*) and yellow Pic (*E.coli*). The analysis shows very similar hydrophobicity profiles for the N-termini (equivalent to the n1 and h1 domains) of the signal peptides. However there is greater variation in the hydrophobicity profiles amongst the C-termini of the signal peptides (corresponding to the n2 and h2 domains).

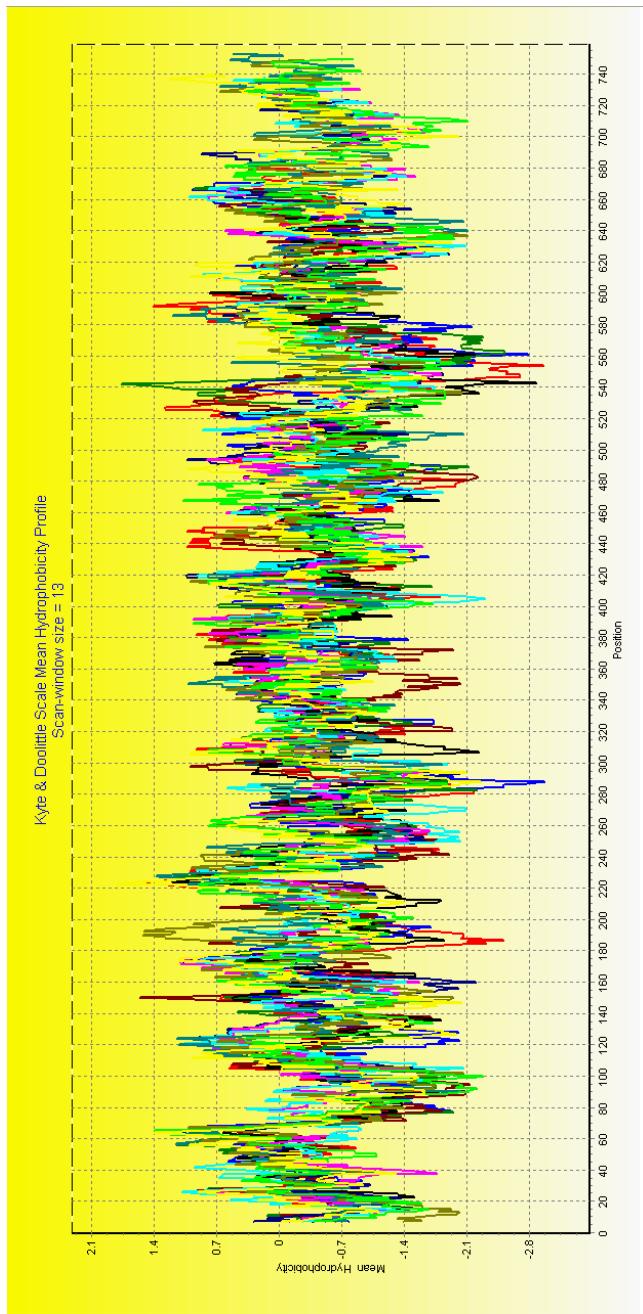


Figure 3.4 Hydrophilicity profile analysis of EspC and homologous protein passenger-domains using the Kyte and Doolittle Scale Mean Hydrophobicity Profile (BioEdit analysis). In red, EspC (EPEC), in blue Pet (*E.coli*), in green Sat (*E.coli*), in brown EpeA (*E.coli*), in black SigA (*Shigella flexerni*), in dark blue SepA (*S. flexerni*), in pink AidA (*E.coli*), in light blue EatA (*E.coli*) in purple Tsh (*E.coli*), in turquoise Hbp (*E.coli*), in olive Vat (*E.coli*), in light green Boa (*Salmonella bongorii*) and in yellow Pic (*E.coli*). The analysis shows similar hydrophobicity profiles traces.

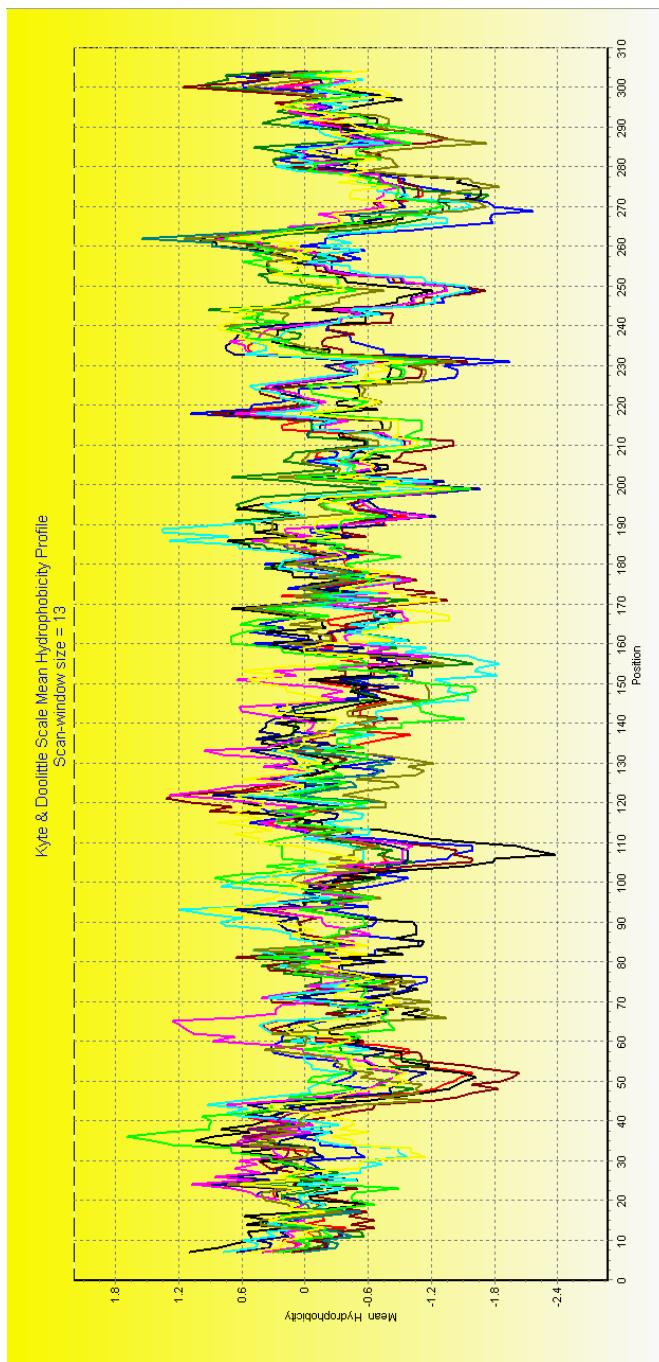


Figure 3.5 Hydrophilicity profile analysis of EspC and homologous protein inter-domains using the Kyte and Doolittle Scale Mean Hydrophobicity Profile (BioEdit analysis). In red, EspC (EPEC), in blue Pet (*E. coli*), in green Sat (*E. coli*), in brown EpeA (*E. coli*), in black SigA (*Shigella flexerni*), in dark blue SepA (*S. flexerni*), in pink AidA (*E. coli*), in light blue EatA (*E. coli*) in purple Tsh (*E. coli*), in turquoise Hbp (*E. coli*), in olive Vat (*E. coli*), in light green Boa (*Salmonella bongorii*) and in yellow Pic (*E. coli*). The analysis shows similar hydrophobicity profile traces.

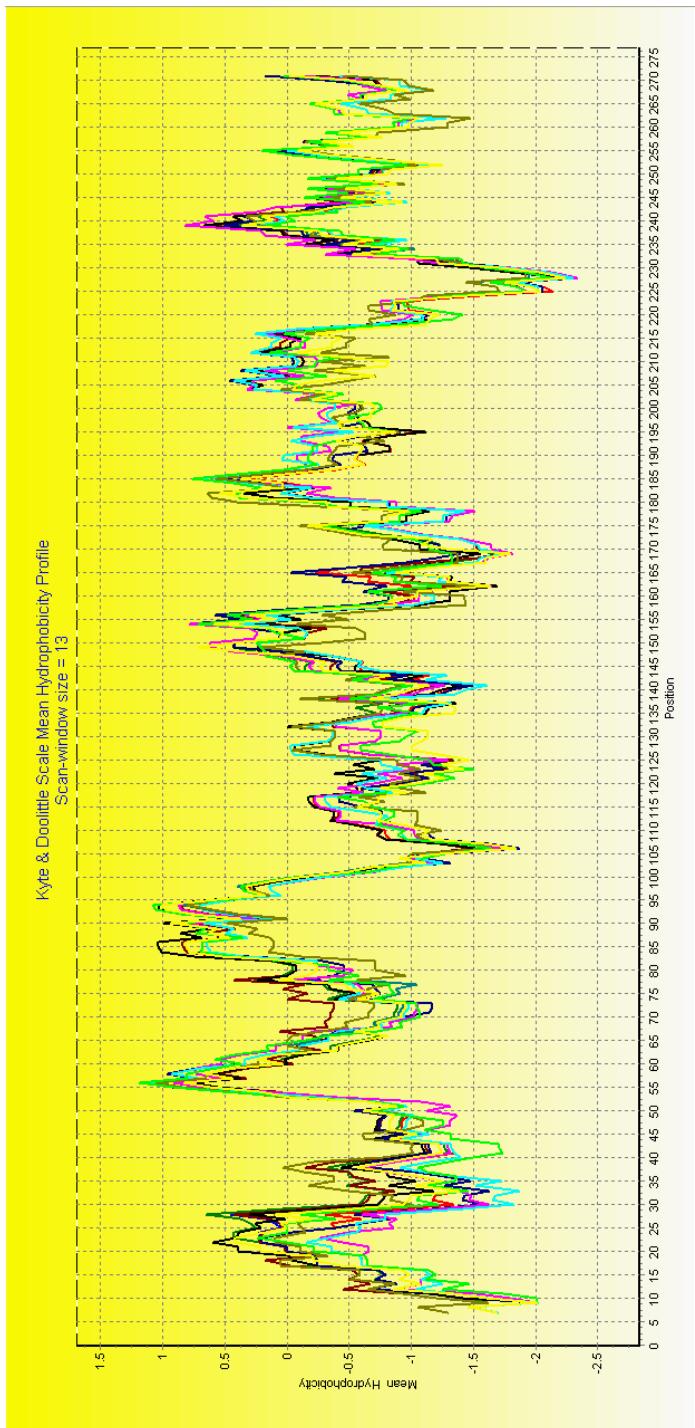


Figure 3.6 Hydrophilicity profile analysis of EspC and homologous protein  $\beta$ -domains using the Kyte and Doolittle Scale Mean Hydrophobicity Profile (BioEdit analysis). In red, EspC (EPEC), in blue Pet (*E.coli*), in green Sat (*E.coli*), in brown EpeA (*E.coli*), in black SigA (*Shigella flexerni*), in dark blue SepA (*S. flexerni*), in pink AidA (*E.coli*), in light blue EatA (*E.coli*) in purple Tsh (*E.coli*), in turquoise Hbp (*E.coli*), in olive Vat (*E.coli*), in light green Boa (*Salmonella bongorii*) and in yellow Pic (*E.coli*). The analysis shows highly similar hydrophobicity profile traces, with less variation than that seen in the traces for both the passenger- and inter- domain.

Table 3.2: Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pl/Mw, BioEDIT and GorIV analyses of EspC signal peptide.

AT	Signal Peptide								PSI-BLAST	
	Compute pl/Mw				BioEDIT	GorIV Analyses (% Composition)				
	MW (kDa)	Length	pl	ID (%)	α-Helix	Extended Strand	Random Coil	E-Value		
<b>EspC</b>	5.6	53	9.90	100	59.49	9.43	32.08	1e-21		
<b>Pet</b>	5.6	52	9.85	49.0	42.31	17.31	40.38	2e-04		
<b>Sat</b>	5.1	49	10.38	50.9	44.90	17.31	40.38	7e-05		
<b>EpeA</b>	6.0	55	10.30	37.5	0.00	38.18	61.82	8e-04		
<b>SigA</b>	5.8	54	9.90	37.5	0.00	42.59	57.41	4e-04		
<b>SepA</b>	6.4	56	10.67	43.1	42.86	8.93	48.21	2e-04		
<b>AidA</b>	6.6	59	10.63	38.9	16.95	16.95	66.10	0.002		
<b>EatA</b>	5.9	56	10.46	33.9	33.93	23.21	42.86	0.72		
<b>Tsh</b>	5.8	52	10.92	35.8	32.69	9.62	57.69	0.97		
<b>Hbp</b>	5.8	52	10.92	35.8	32.69	9.62	57.69	0.97		
<b>Vat</b>	5.9	55	10.46	34.4	36.36	30.91	32.73	0.010		
<b>Boa</b>	6.2	57	10.29	32.7	36.84	17.54	45.61	0.061		
<b>Pic</b>	6.1	55	9.76	52.7	32.73	27.27	40.00	2e-06		

Table 3.3: Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pI/Mw, BioEDIT and GorIV analyses of EspC passenger-domain.

AT	Passenger Domain								PSI-BLAST	
	Compute pI/Mw				BioEDIT	GorIV Analyses (% Composition)				
	MW (kDa)	Length	pI	ID (%)	α-Helix	Extended Strand	Random Coil	E-Value		
<b>EspC</b>	74.8	691	5.52	100	12.01	30.97	57.02	0		
<b>Pet</b>	73.2	677	6.47	46.3	15.95	30.28	53.77	2e-172		
<b>Sat</b>	73.3	674	5.62	45.7	12.91	30.27	56.82	6e-165		
<b>EpeA</b>	73.6	675	6.33	46.8	18.52	26.67	54.81	9e-175		
<b>SigA</b>	71.7	655	8.69	52.2	18.63	25.80	55.57	0		
<b>SepA</b>	79.1	744	5.68	24.4	14.25	30.38	55.38	3e-29		
<b>AidA</b>	79.4	744	5.29	23.8	11.16	32.53	56.32	7e-29		
<b>EatA</b>	80.2	744	5.46	23.9	13.58	30.51	55.91	1e-31		
<b>Tsh</b>	80.7	759	4.88	22.9	13.44	28.19	58.37	2e-34		
<b>Hbp</b>	80.7	759	4.93	22.9	13.44	28.19	58.37	2e-34		
<b>Vat</b>	80.6	755	5.46	22.7	7.55	32.58	59.87	2e-30		
<b>Boa</b>	81.3	756	5.31	23.3	15.21	24.74	60.05	3e-35		
<b>Pic</b>	79.5	746	5.63	24.1	11.93	28.15	59.92	3e-46		

Table 3.4: Physical/structural properties of homologous proteins obtained using a PSI-BLAST search, Compute pi/Mw, BioEDIT and GorIV analyses of EspC inter-domain.

AT	Inter Domain								PSI-BLAST E-Value	
	Compute pi/Mw				BioEDIT	GorIV Analyses (% Composition)				
	MW (kDa)	Length	pi	ID (%)	α-Helix	Extended Strand	Random Coil			
EspC	29.8	284	4.89	100	17.25	27.82	54.93		1e-159	
Pet	30.8	289	6.02	38.2	31.83	19.72	48.44		7e-45	
Sat	31.4	295	5.47	34.8	23.73	22.71	53.56		1e-42	
EpeA	31.8	293	5.64	36.3	30.03	20.82	49.15		1e-43	
SigA	31.9	299	5.66	33.1	24.75	24.08	51.17		1e-39	
SepA	30.1	287	5.43	85.3	19.16	24.74	56.10		1e-134	
AidA	30.5	286	5.96	73.1	18.18	26.67	55.24		2e-111	
EatA	30.8	287	6.50	71.0	30.66	20.91	48.43		2e-109	
Tsh	31.1	289	5.06	31.0	19.03	25.26	55.71		1e-32	
Hbp	31.1	289	5.06	31.0	16.61	26.30	57.09		1e-32	
Vat	31.3	290	4.95	31.0	15.17	27.93	56.90		6e-32	
Boa	31.2	294	5.38	30.7	21.77	23.47	54.76		1e-35	
Pic	30.4	293	5.57	38.3	25.94	24.91	49.15		6e-52	

Table 3.5: Physical/structural properties of homologous AT proteins obtained using a PSI-BLAST search, Compute pI/Mw, BioEDIT and GorIV analyses of EspC β-domain.

AT	β-Domain								PSI-BLAST E-Value	
	Compute pI/Mw				BioEDIT		GorIV Analyses (% Composition)			
	MW (kDa)	Length	pI	ID (%)	α-Helix	Extended Strand	Random Coil			
EspC	30.6	277	5.18	100	25.63	17.69	56.68		6e-163	
Pet	30.3	277	5.52	80.8	20.94	23.83	55.23		5e-136	
Sat	30.3	277	5.52	80.1	19.13	25.63	55.23		3e-135	
EpeA	30.4	277	5.92	81.2	20.94	25.99	53.07		2e-137	
SigA	30.4	277	6.00	80.5	16.25	28.16	55.60		2e-135	
SepA	30.5	277	5.49	94.2	24.91	20.22	54.87		9e-154	
AidA	30.8	277	5.72	85.9	14.44	29.60	55.96		2e-140	
EatA	30.9	277	7.26	82.6	15.88	28.16	55.96		1e-137	
Tsh	30.6	277	6.26	61.7	27.08	23.47	49.46		7e-100	
Hbp	30.6	277	6.26	61.7	27.08	23.47	49.46		7e-100	
Vat	30.5	277	6.48	61.7	27.80	23.10	49.10		1e-99	
Boa	30.6	277	6.41	69.6	19.86	25.99	54.15		5e-116	
Pic	30.6	277	5.92	79.0	22.02	21.66	56.32		2e-131	

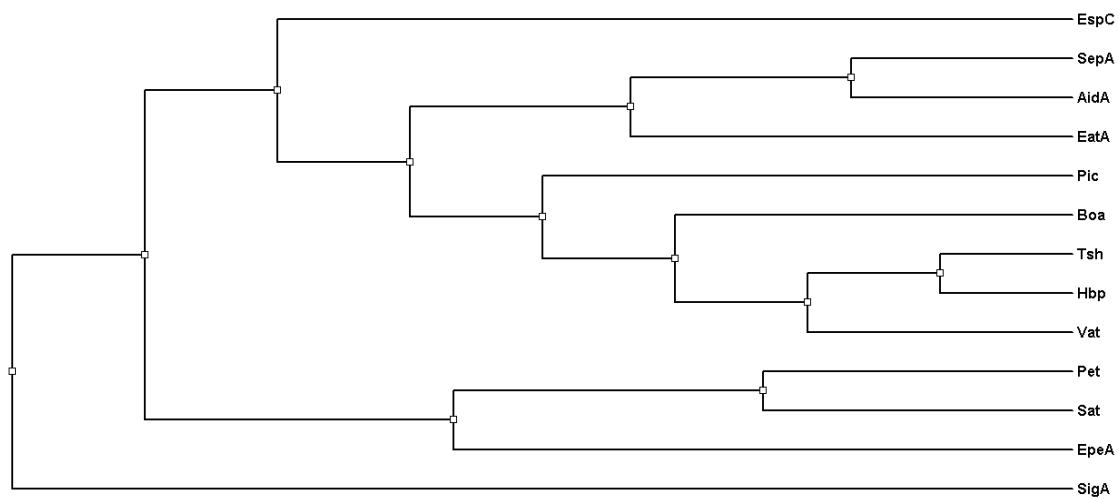
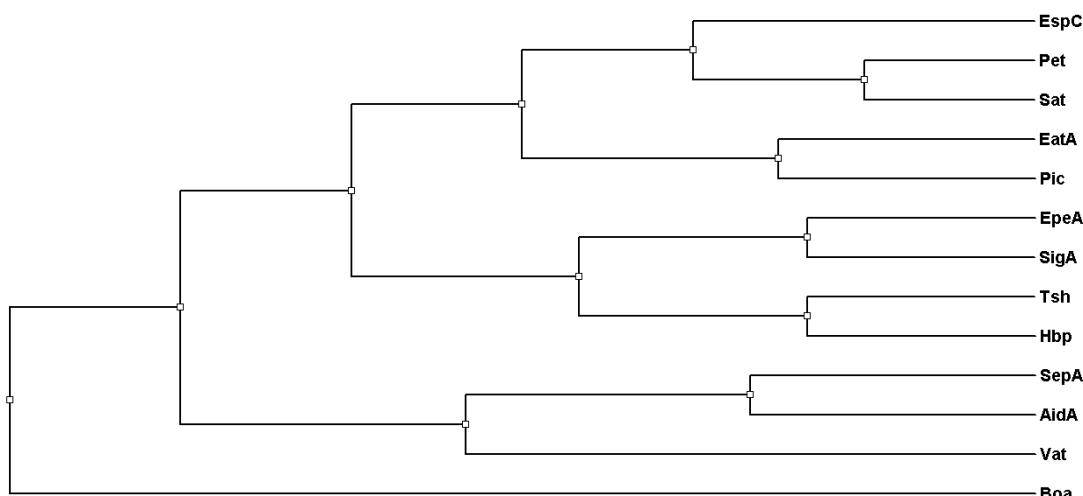
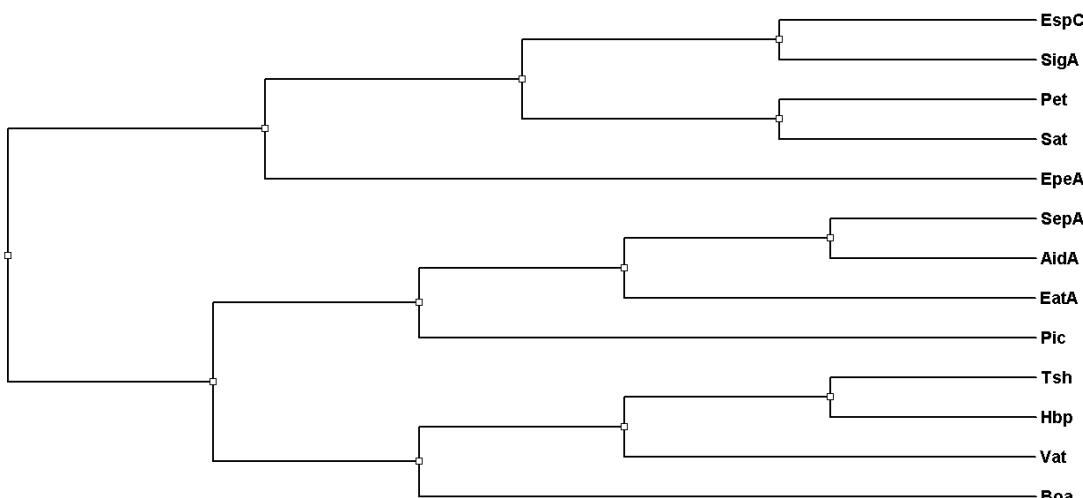


Figure 3.7 Phylogenetic tree generated from entire AT sequences. AT protein sequences including EspC were obtained from the NCBI database, the accession numbers and references for each of them is listed in Table 3.1. The phylogenetic tree was constructed from the pileup of sequences in ‘ClustalW 1.83’. Over its entire length, EspC is shown not to cluster closely with any other AT.

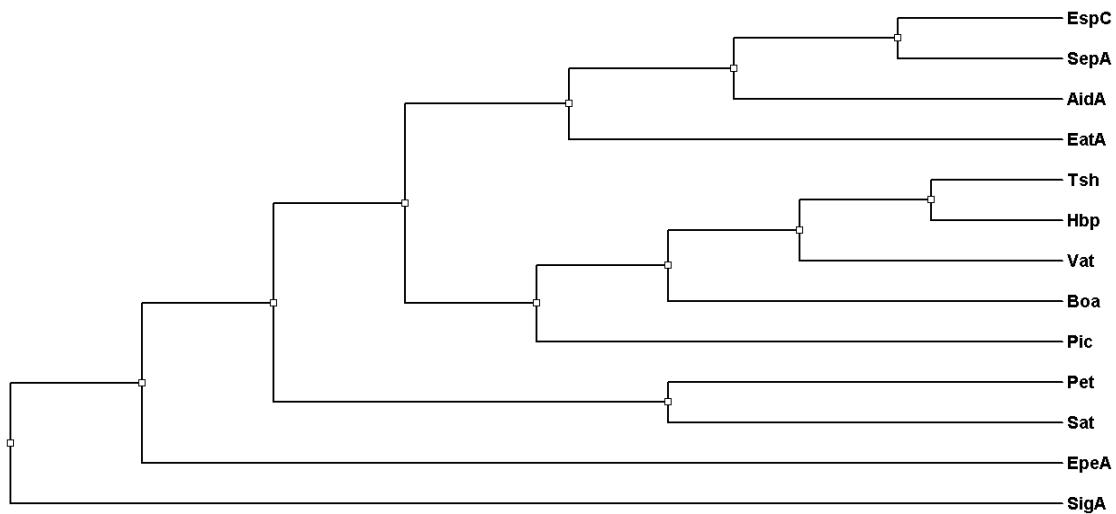


**A: Phylogenetic tree generated from AT signal peptide sequence**

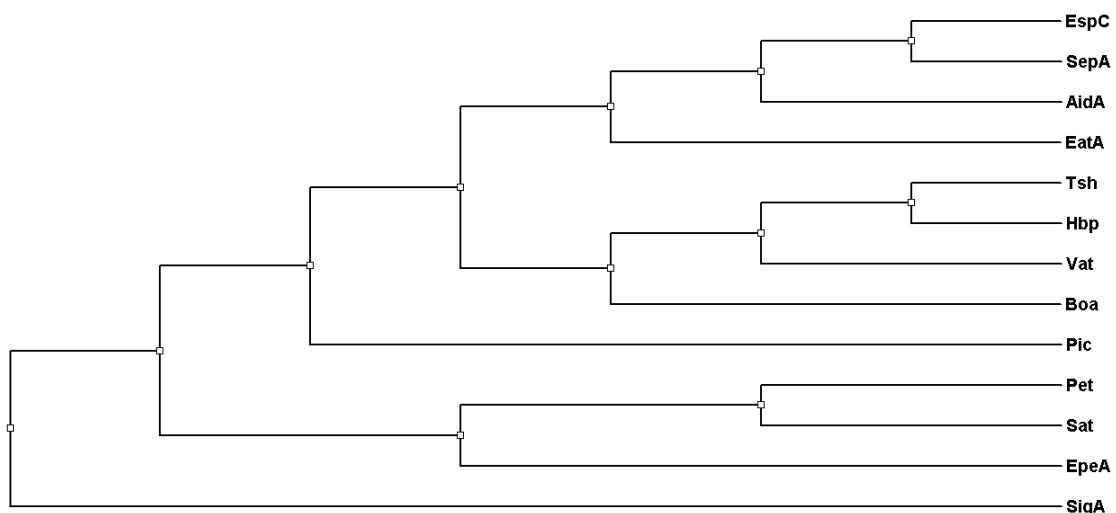


**B: Phylogenetic tree generated from AT passenger-domain sequence**

Figure 3.8 Phylogenetic trees generated from AT signal peptide and passenger-domain sequences. AT protein sequences including EspC were obtained from the NCBI database, the accession numbers and references for each of them is listed in Table 3.1. The signal peptide (A) and passenger domain (B) phylogenetic trees were constructed from the pileup of sequences in 'ClustalW 1.83'. The EspC signal region is most similar to that of the equivalent region in Pet and Sat, and additionally clusters with EatA and Pic. EspC passenger domain is most similar to that of the equivalent region in SigA, but as with the signal region, also clusters with Pet and Sat, and additionally EpeA.



**A: Phylogenetic tree generated from AT inter-domain sequence**



**B: Phylogenetic tree generated from AT β-domain sequence**

Figure 3.9 Phylogenetic tree generated from AT inter- and β-domain sequences. AT protein sequences including EspC were obtained from the NCBI database, the accession numbers and references for each of them is listed in Table 3.1. The inter-domain (A) and β-domain (B) phylogenetic trees were constructed from the pileup of sequences in ‘ClustalW 1.83’. EspC inter-domain is most similar to that of the equivalent region in SepA. Surprisingly, given the apparent sequence relatedness of passenger-domains, EspC inter-region sequence does not cluster closely with SigA, Pet or Sat inter-regions, but instead clusters with SepA, AidA and EatA. Consistent with the results for the Inter-domain, EspC β-domain is shown to be most closely related to SepA, AidA and EatA.

### **3.5 EspC inter-domain structural modelling.**

Phylogenetic analysis of the full AT sequence uncovered a smaller sub-group (identified in the phylogenetic guide tree to have the most similar inter- and β-regions (Fig 3.9). This sub-group comprised EspC, EpeA, SigA, and SepA, which were then used to identify conserved and invariant residues by a series of sequential alignments (not shown) against haemoglobin protease (which we have denoted as Hbp2) for which a crystallographic structure of the passenger-domain (which includes the inter-domain) has been determined ([Otto et al, 2005](#)). Subsequent structural modelling of the EspC inter-domain against the crystal structure of Hbp2 revealed structural features of EspC within the inter-domain amenable for potential exploitation by site-directed mutagenesis. Additionally, the ability to model the EspC inter-domain on the equivalent Hpb2 structure provided some degree of confidence that the EspC protein also comprised a putative chaperone domain.

Hbp2 has less identity with EspC than the other sequences so aligning without it gave a better indication of conserved/invariant residues. The alignment shown includes Hbp2 (Fig 3.10). Conserved and invariant residues are again indicated on the sequence alignment by the ClustalW consensus. For structural modelling, sequential alignments of small inter-domain regions using only EspC and Hbp2 were carried out using ClustalW, then each alignment subsequently submitted to SwissModel. Sequences to be aligned were progressively truncated towards the inter/β-domain boundary until similarity between sequences was sufficient for SwissModel to generate a homology model. The start and end of the ‘Alignment Interface’ used to paste in the final ClustalW alignment is indicated by horizontal arrows (Fig 3.10). The final sequence alignment was entered into SwissModel using Hbp2, chain A (PDB code:1WXR) as the model template and a PDB file comprising EspC sequence overlaying the Hbp2 structure generated. The homology model was subsequently viewed and annotated using the Rasmol structural visualisation program which allowed predicted EspC inter-domain structural features to be observed (Fig 3.11).

### EspC Inter construct start highlighted in green

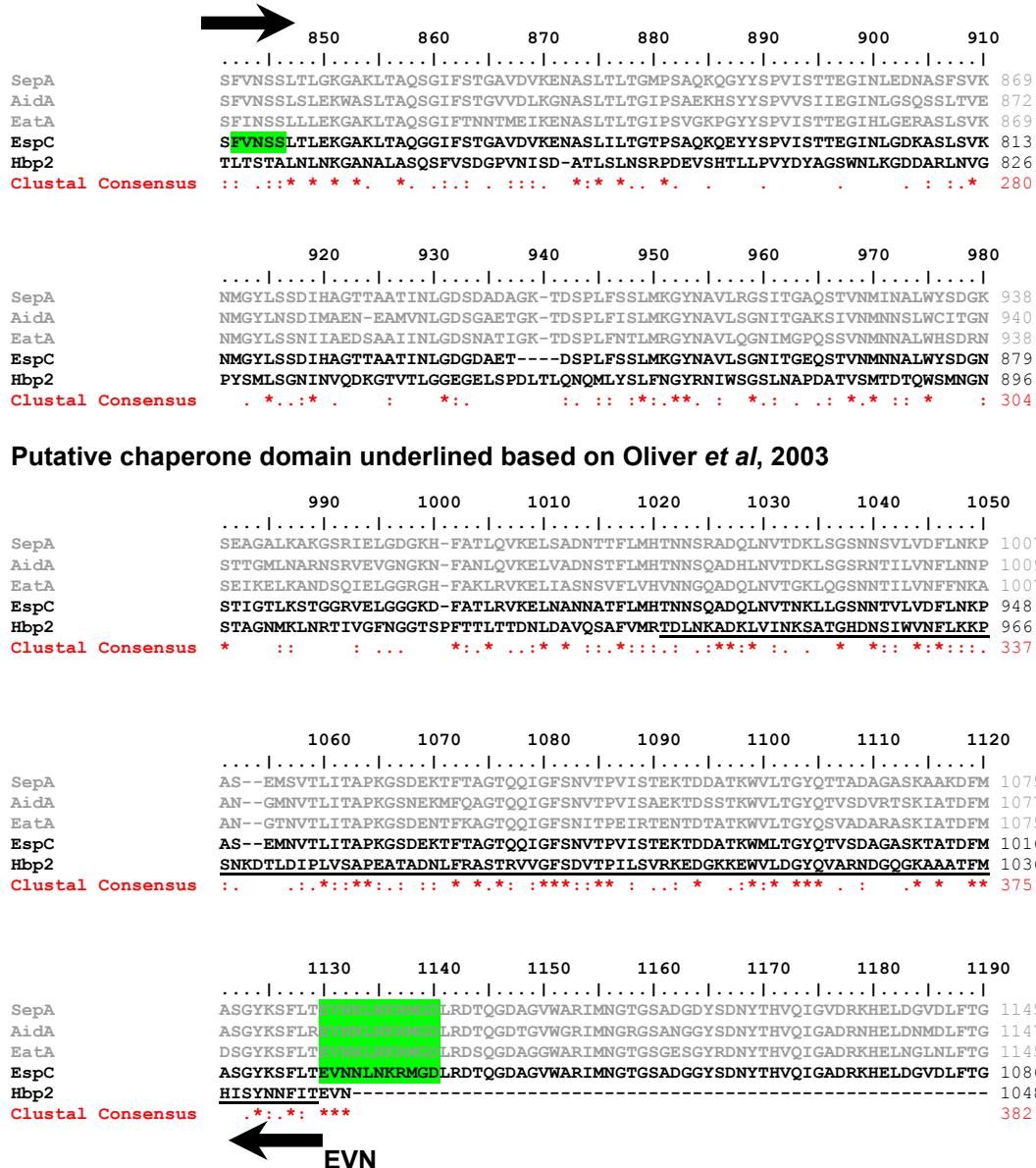


Figure 3.10: Alignment of the EspC inter-domain sequence subgroup against Hbp2 (truncated Hbp). Conserved and invariant residues are indicated by the ClustalW consensus by a star (\*). The start/end of the EspC/Hbp2 truncation which allowed mapping of EspC onto Hbp is indicated in the above schematic by horizontal arrows with the start/end sequence. Highlighted in green is the start/end of the original inter-domain sequence selected for subsequent cloning into a variety of vectors.

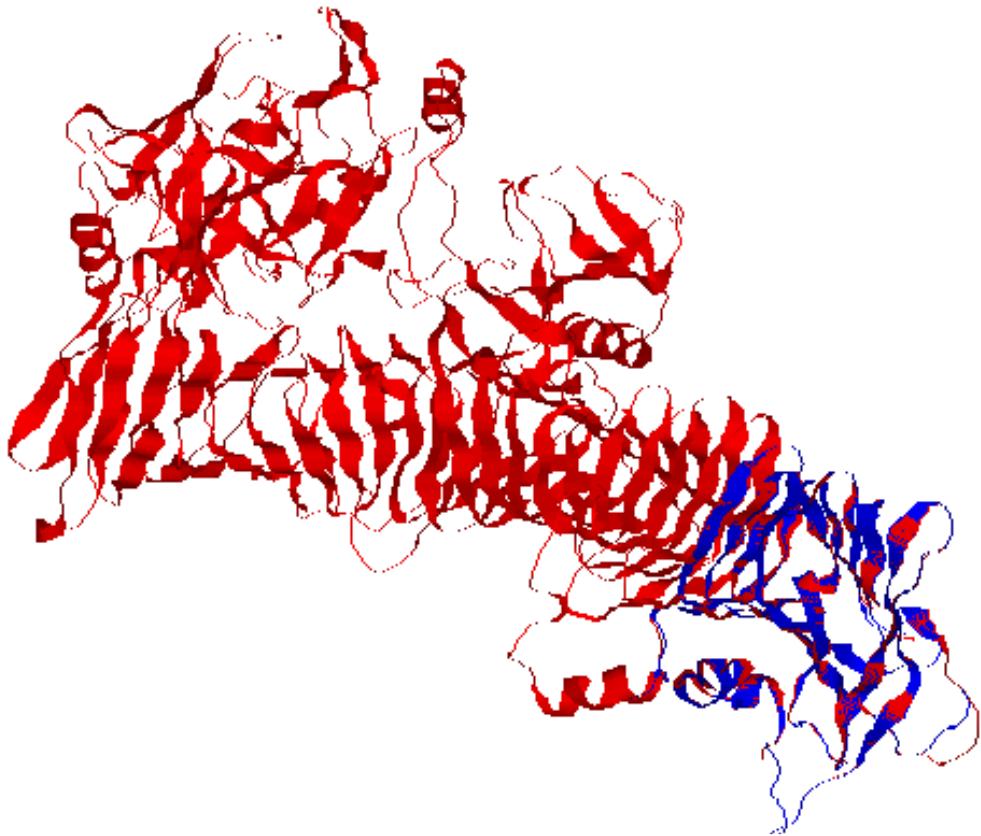


Figure 3.11: Crystal structure of haemoglobin protease AT (Hbp, Accession No: 1WXR), obtained from SWISS-PROT, visualized using RASMOl. The terminal region of the Hbp2 structure, comprising the putative inter-domain was used for homology modelling with EspC sequence. The modelled EspC inter-domain is shown overlaying the Hbp2 structure in blue.

### **3.6 *In silico* analysis of YbgC.**

YbgC is a cytoplasmic protein and as such lacks a signal peptide. As with *espC* sequences, the sequence of EHEC (Sakai strain) *ybgC* was also obtained from the NCBI (Accession: NP\_308798), and was used to identify homologous proteins (Table 3.6). In addition following a ClustalW alignment, GorIV analysis and referral to the literature, a general structure of the YbgC protein was also successfully established (Fig 3.12) and primers designed to amplify the entire gene.

No obvious sub-domains within YbgC were identified. A hydrophobicity analysis of YbgC and its homologues using the Kyte and Doolittle Scale (Fig 3.13) revealed highly similar profiles shared by the recently released crystal structure of YbgC (Kim *et al*, 2004), a hydrolase with homology to 4-HBA-CoA (Accession: 1S5U). A phylogenetic tree was again generated to identify closely related homologues (Fig 3.14).

YbgC is predicted to be a thioesterase (Zhuang *et al*, 2002) that aids phospholipid incorporation into the OM (Gully and Bouveret, 2006). Structural analysis of a homologue in *Haemophilus influenzae* has identified the location of the active site of the thioesterase based on the X-Ray crystal structure of 4-HBA-CoA (*Pseudomonas* sp.) (Zhuang *et al*, 2002). It contains a  $\beta$ -strand, and  $\alpha$ -helix in close proximity to the N-terminus. An aspartic acid residue at position 17 (Asp17) located between these secondary structures functions in base/nucleophilic catalysis (Zhuang *et al*, 2002). Asp18 is thought to perform the same role in the YbgC *E. coli* homologue. Consistent with the small size of the protein, no specific functional domains were apparent.

To illustrate the possible multimeric nature of YbgC, the sequence obtained from NCBI was submitted to the CPHmodels server ([www.cbs.dtu.dk/services/CPHmodels/](http://www.cbs.dtu.dk/services/CPHmodels/)) for modelling. The server returned the closest match as being YbgC (Accession: 1S5U), chain E (E-value 1e-72). In the model generated the active site has been emphasised using Rasmol software (Fig 3.15). The YbgC molecule, appears to form a homotetrameric macromolecule, that is a dimer of two tetrameric molecules.

Table 3.6 Identification of YbgC homologous proteins utilising a PSI-BLAST. pI values obtained using EDITSEQ of DNASTAR.

Name	Length (Amino acids)	MW	pI	Function	Accession Number	Organism	E-value	Reference
YbgC	134	15.6	7.40	Hypothetical protein ECs0771	NP_308798 (BAB34194)	<i>E. coli</i> EHEC Sakai	5e-67	<a href="#">Makino et al, 1999</a>
YbgC	134	15.6	7.40	Putative esterase YbgC	YP_668672	<i>E. coli</i> 536 06:K15:H31	5e-67	<a href="#">Brzuszkiewicz et al, 2006</a>
APEC O1	134	15.6	7.40	putative thioesterase	YP_851840	<i>E. coli</i> APEC 01:K1:H7	5e-67	<a href="#">Johnson et al, 2007</a>
Protein YbgC	134	15.6	7.40	Acyl-CoA thioester hydrolase	NP_752745	<i>E. coli</i> CFT073	5e-67	<a href="#">Welch et al, 2002</a>
YbgC	134	15.6	7.40	Predicted acyl-CoA Thioesterase	NP_415264	<i>E. coli</i> K12 MG1655	5e-67	<a href="#">Blattner et al, 1997</a>
YbgC	134	15.6	7.40	Acyl-CoA thioester hydrolase (YbgC) Z0904	NP_286464	<i>E. coli</i> EHEC 0157:H7	7e-66	<a href="#">Perna et al, 2001</a>
COG0824	134	15.6	7.40	Predicted thioesterase	ZP_00721390	<i>E. coli</i> F11	8e-66	<a href="#">NCBI, 2005</a>
Chain A	138	15.9	7.46	Hypothetical Protein (Hydrolase) Ec709	1S5U_A	<i>E. coli</i>	5e-67	<a href="#">Kim et al, 2004</a>
YbgC	134	15.7	8.15	Putative esterase	NP_459729	<i>Salmonella typhimurium</i> LT2	1e-62	<a href="#">McCellend et al, 2001</a>
YbgC	139	15.9	7.27	Tol-pal associated Acyl-CoA thioesterase	ZP_01985153	<i>Vibrio harveyi</i>	7e-31	<a href="#">Heidelberg, 2007 (NCBI)</a>
YbgC	134	15.3	7.45	Unknown, cytoplasmic protein	CAC82705	<i>Erwinia chrysanthemi</i>	4e-47	<a href="#">Dubuisson et al, 2005</a>
YbgC	136	15.3	5.90	Short chain Acyl-CoA Thioesterase YbgC.	YP_204340	<i>Vibrio fischeri</i>	8e-33	<a href="#">Ruby et al, 2005</a>

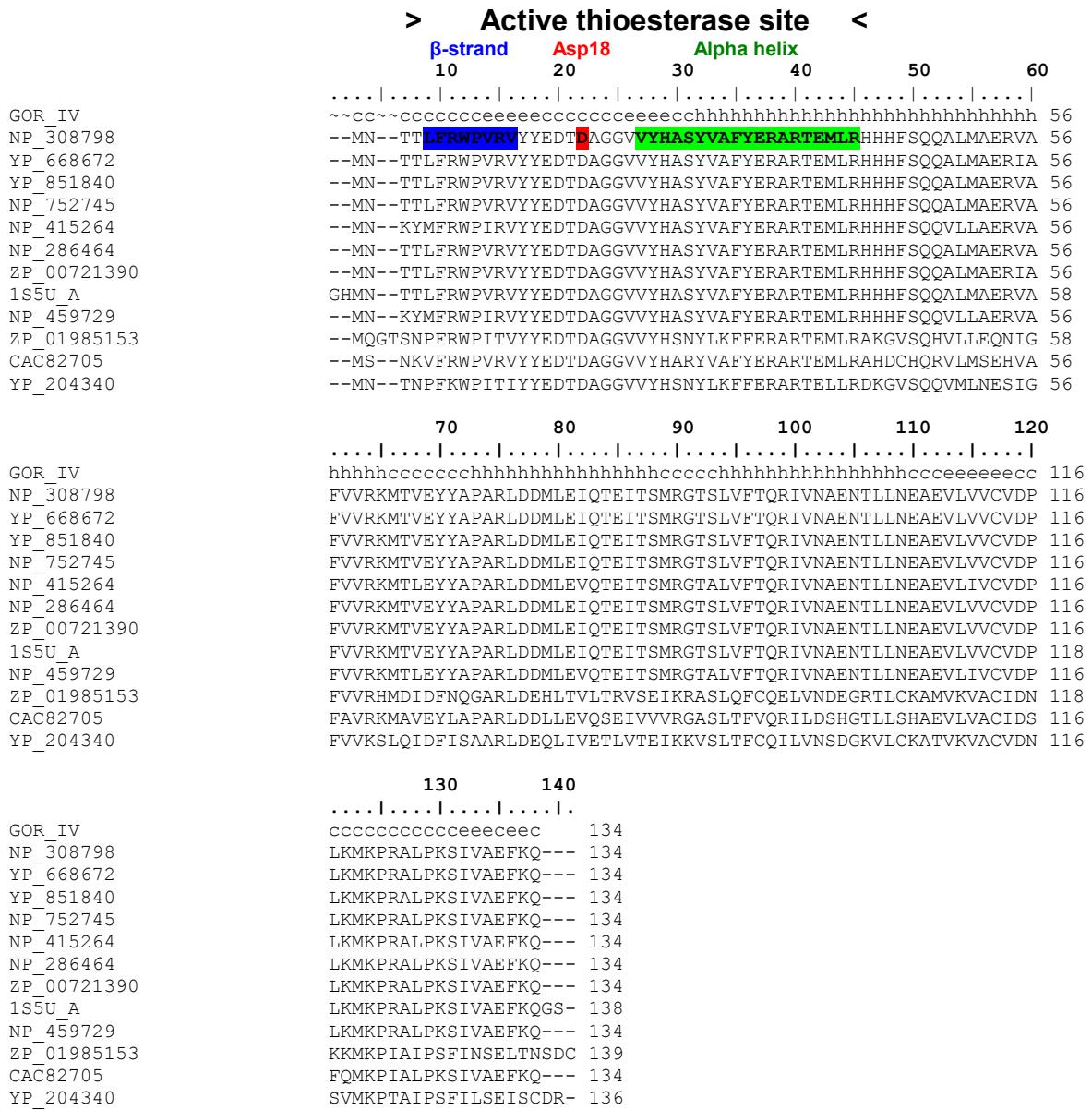


Figure 3.12: ClustalW alignment of YbgC amino acid sequence compared against homologous sequences identified by a PSI-BLAST search. Secondary structure analysis of YbgC (Accession NP\_308798) by GORIV is included (Alpha Helix (h) 49.25%, Extended Strand (e) 14.93%, and Random Coil (c), 35.82%). No individual domains have been identified in the protein sequence. However, from the literature the probable active site of the putative thioesterase has been highlighted, the  $\beta$ -strand (highlighted in blue), the Asp18 residue (red) and the N-terminus of the first alpha helix (green).

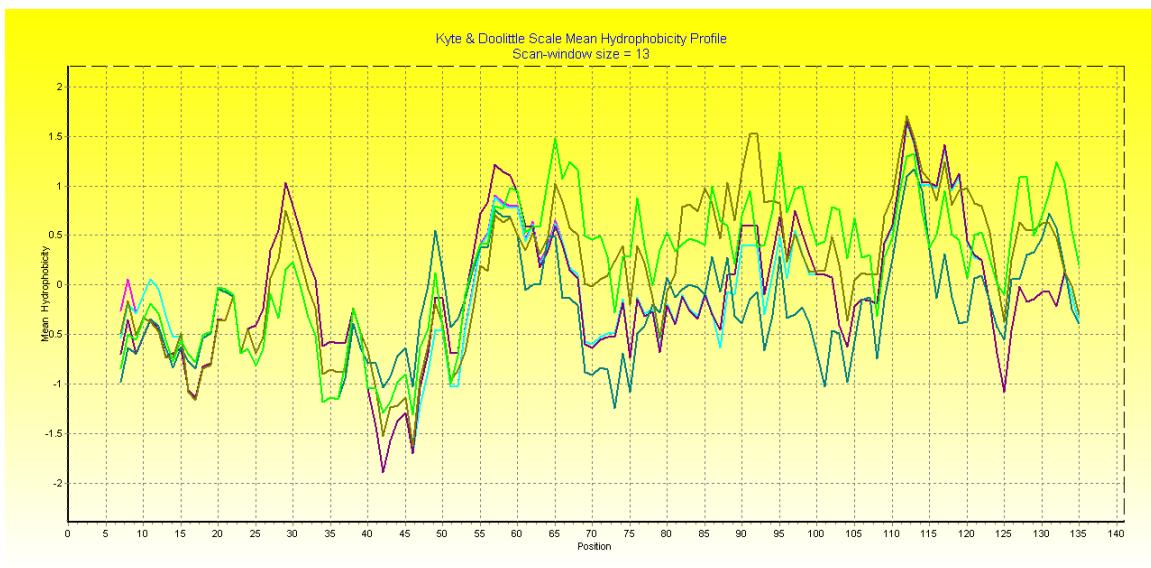


Figure 3.13 Hydrophilicity Profile Analysis of YbgC and homologous proteins using Kyte and Doolittle Scale Mean Hydrophobicity Profile (BioEdit analysis). In red, NP\_308798 (EHEC, Sakai), in blue YP\_668672 (*E.coli*, 536), in green YP\_851840 (APEC), in brown NP\_752745 (*E.coli*, CFT073), in black NP\_415264 (*E.coli*, K12, MG1655), in dark blue NP\_286464 (EHEC), in pink ZP\_00721390 (*E.coli*, F11), in light blue 1S5U chain A (*E.coli*, Ec709) in purple NP\_459729 (*Salmonella typhimurium*), in turquoise ZP\_01985153 (*Vibrio harveyi*), in olive CAC82705(*Erwinia chrysanthemi*), in light green YP\_204340 (*Vibrio fischeri*). The analysis shows a similar hydrophilicity profile amongst YbgC and its homologous proteins.

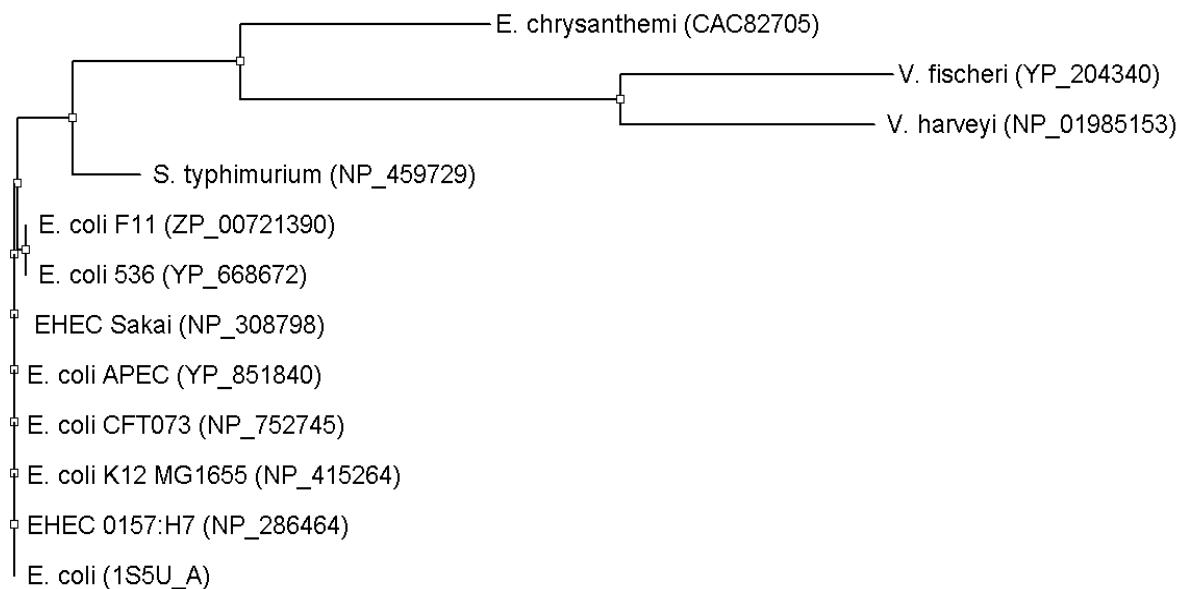


Figure 3.14 Phylogenetic analyses of homologous YbgC proteins. YbgC protein sequences were obtained from the NCBI database, the accession numbers and references for each of them is listed in Table 3.1. The phylogenetic tree was constructed from the pileup of sequences in ‘ClustalW 1.83’. In the above tree, EHEC Sakai YbgC (NP\_308798), used to design primers has high homology ( $\geq 97\%$ ) to all *E. coli* YbgC proteins, including YbgC (Accession: 1S5U) a hydrolase (Kim et al, 2007) whose crystal structure has been shown to be similar to 4HBT, a thioesterase known to form a tetrameric structure from dimers of the individual protein. EHEC Sakai YbgC (NP\_308798) also has high level of identity (>91%) with *S. typhimurium*, though with the remaining non *E. coli* strains a lower level of homology is apparent.

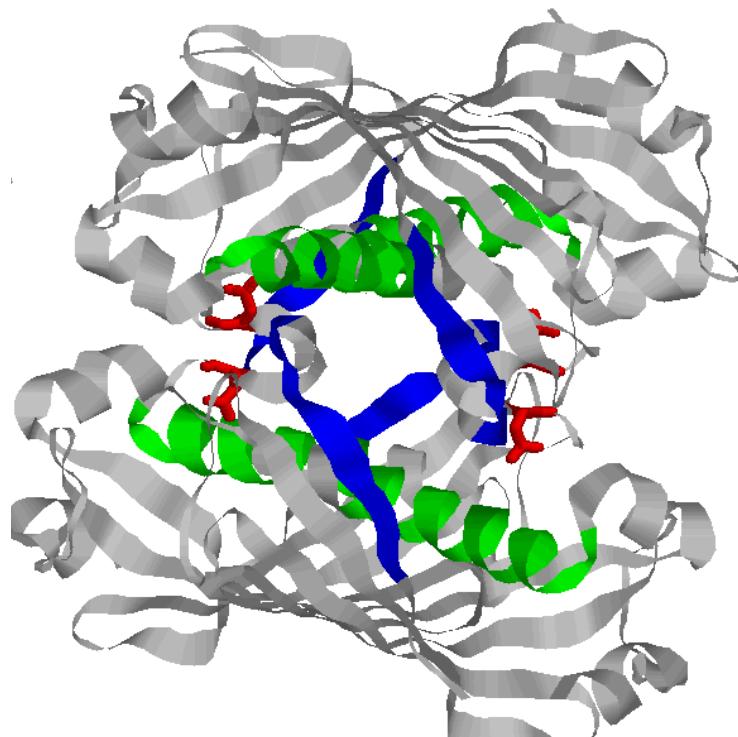


Figure 3.15: Crystal structure of Hydrolase (Accession: 1S5U). Hydrolase found in *E.coli* that is homologous to 4HBT a thioesterase with homology to YbgC. YbgC (Accession: NP\_308798) was obtained from NCBI and mapped most closely utilising CHP Models 2.0 to 1S5U\_E (E-value: 1e-72). The individual YbgC model molecules can be shown to form a dimer of two tetrameric complexes. The active site elements including Asp19 (Red), the  $\beta$ -strand (Blue) and  $\alpha$ -helix (Green) have been highlighted, only one tetrameric molecule is shown in this figure.

### 3.7 Discussion.

Establishing the domain boundaries within EspC with confidence, facilitated the design of oligonucleotide primers for future PCR amplification of three separate EspC domains including I) the passenger-domain II) the inter-domain and III) the  $\beta$ -barrel domain.

Phylogenetic analyses revealed that the EspC inter- and  $\beta$ -domains may derive from the same AT subset, which is very intriguing. The inter-domain is considered to assist folding of the passenger in at least some ATs. Given the functional dependence that this implies, it might be expected that the inter- and passenger-domains closely co-evolved. However, the observation that the EspC inter- and  $\beta$ -domains appear more closely related suggests that in EspC, the inter-domain may be important for function of the  $\beta$ -domain, perhaps assisting membrane localization or export. This domain partition is also reminiscent of the T5B (TPS) system comprising two separate proteins, one (signal and passenger) being the functional exported protein and the other ( $\beta$ ) being the translocation accessory protein. [Davis et al, \(2001\)](#), have previously identified such DNA transfer between species for the gene *lav*, which encodes a novel virulence-associated AT of *Haemophilus influenzae* which their analyses found to have transferred by horizontal acquisition to *Neisseria*. This work built on the evidence of [Kroll et al, \(1998\)](#) who first reported natural transfer of *H. influenzae* sequences to *N. meningitidis*.

Evidence that YbgC forms multimers is supported by unpublished experimental work (Djamel Gully, personal communication) showing that under non-reducing conditions YbgC migrated in SDS-PAGE gels as a dimer. Experimental evidence also exists that YbgC is a dimer based on the crystallographic structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthobacter* sp. ([Thoden et al, 2003](#)), of which YbgC is a homologue. [Thoden et al's](#) study demonstrated that the thioesterase enzyme from *Arthobacter* sp. is tetrameric and is composed of a dimer of dimers, with all subunits adopting a 'hot dog fold'. Similar structures have been observed in *Pseudomonas* thioesterase ([Benning et al, 1998](#)) though its quaternary structure is different they possess equivalent catalytic efficiencies, substrate specificity, and metabolic functions. The recently published crystal structure of hydrolase (Accession: 1S5U) with significant similarity to YbgC (Accession: NP\_308798) has been included in the *in silico* analysis of YbgC to demonstrate that *E. coli* YbgC is likely multimeric in nature, consisting of up to 8 chains forming a homotetrameric structure. When using the modeling server CHPmodels, YbgC protein

sequence was found to closely match that of Chain E. In addition, two different studies, by Rain *et al*, (2001) and Terradot *et al*, (2004) established through a Y2H protocol and *in vitro* pull down assay respectively, that *Helicobacter pylori* protein HP0496, encoding a YbgC homologue (HpYbgC) interacts with the cytotoxin associated gene protein, CagA. Recent work by Angelini *et al*, (2008) to characterize HpYbgC, established that this protein is an acyl-Co-A thioesterase, a member of the family of hot-dog proteins, has a tetrameric arrangement, and is highly homologous to *E. coli* YbgC. No specific phenotype was attributable to YbgC however. Gully and Bouveret (2006) have suggested that YbgC activity is linked to the Tol-Pal system without being crucial to its function. They hypothesize that YbgC may be required for the synthesis, termination and release of ACP of specific lipids, that are transferred to phospholipid synthetic enzymes.

## **4. Generation of an EspC mutant and polyclonal antibody tools**

### **4.1 Introduction**

A number of tools were required for this study including (I) an EspC deficient mutant in which native EspC, EspC domains or mutated derivatives could be expressed to determine the effects on EPEC EspC secretion without interference from the wild type EspC, and (II) anti-EspC antibodies to identify EspC in complexes or perform affinity purification of complexes containing it. To assist characterisation of the EspC mutant, further polyclonal antibodies were generated including anti-EspB, and anti-FliC antibodies. The versatility of these other antibodies is illustrated by their wider use for related studies which have already been published ([Ferrández et al, 2005](#)).

### **4.2 Generation of an espC EPEC mutant – Lambda red mutagenesis**

Oligonucleotides EspCMutF and EspCMutR (Table 2.3) were used to successfully generate a 1.5 kb PCR fragment encoding a kanamycin cassette containing *espC* flanking regions in three of four magnesium concentrations tried (Fig 4.1A). Following electroporation of the gel purified 1.5 kb PCR fragment (Kan-*espC* flank) into electro-competent EPEC E2348/69 cells hosting the Lambda Red plasmid pKD46, transformants were recovered on LB kanamycin plates. Primer pairs (I) pKD4\_For and pKD4\_Rev and (II) pKD4\_For and Barr1 were used to confirm the deletion of EspC by colony PCR (4.1B). In addition, the selected colony was confirmed to be EPEC rather than a contaminant by the PCR amplification of both *espB* and *escN* – two EPEC specific T3SS genes (Fig 4.1C) using the primer pairs (I) EspB\_For and EspB\_Rev, and (II) EscN\_For and EscN\_Rev respectively ([Creasy et al, 2003](#)) (see Table 2.3 for primer sequences). An in-frame deletion of EspC was successfully generated

### **4.3 Purification of wild-type EspC and generation of specific anti-EspC antibodies.**

EPEC secreted proteins were precipitated from overnight broth supernatants of strain E2348/69 as described in Section 2.6.1, separated on 11.5% SDS-PAGE gels, and visualised with Coomassie Blue stain (section 2.7.3). In collaboration with Sue Dodson the EspC protein (~110 kDa) was excised and electro-eluted (Section 2.7.5).

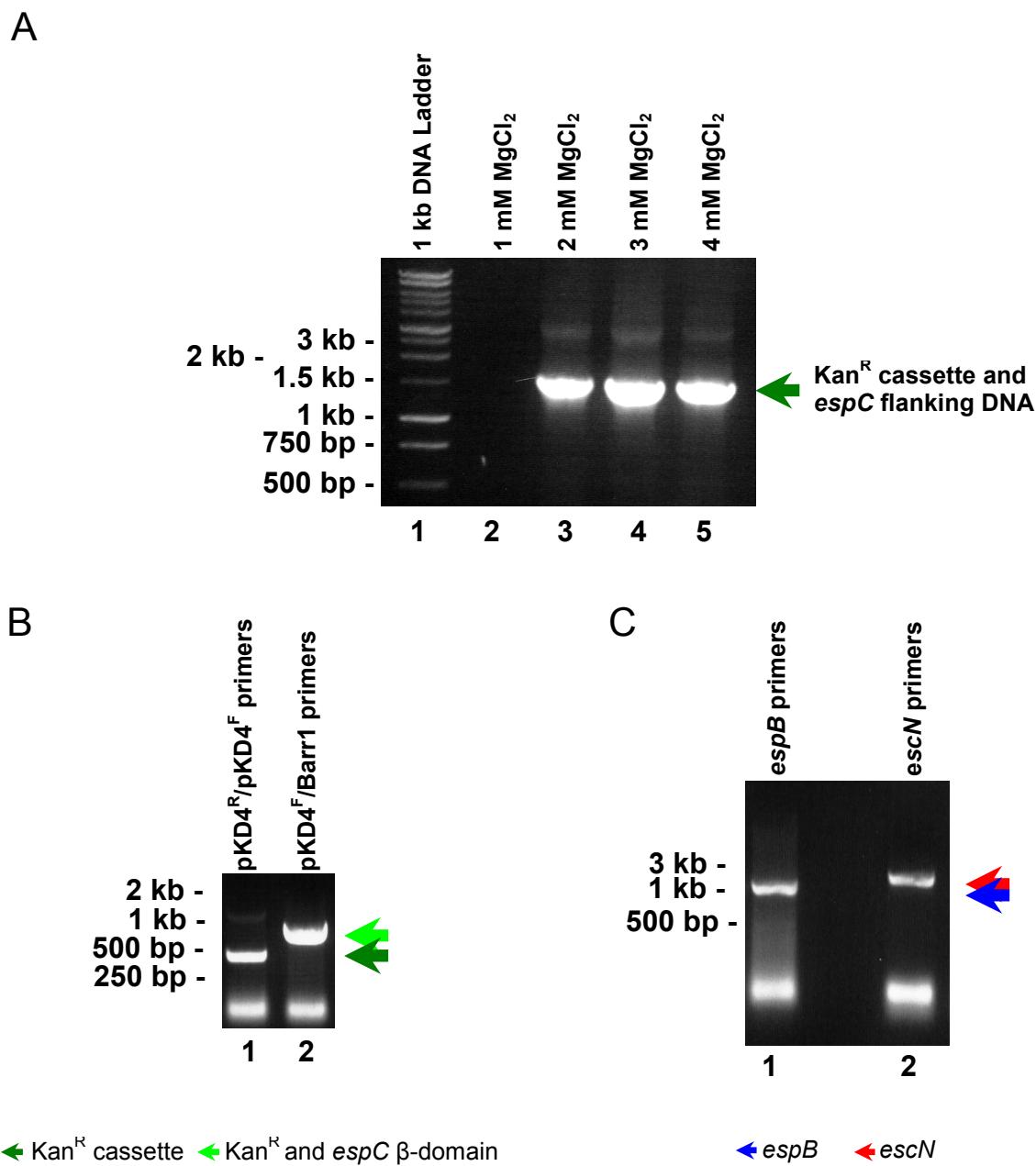


Figure 4.1: Generation and characterisation of an *espC* mutant. DNA encoding a kanamycin resistance cassette from plasmid pKD4 with EPEC *espC* flanking regions was successfully amplified (A). PCR conditions utilising 2 mM (Lane A3), 3 mM (Lane A4), and 4 mM  $MgCl_2$  (Lane A5) concentrations resulted in successful generation of a 1.5 kb PCR product of the expected size by comparison to DNA markers (1 kb, Lane A1). EPEC containing pKD46 was then transformed by electroporation with the 1.5 kb PCR product. Transformants recovered were characterised by colony PCR to demonstrate the majority of *espC* had been successfully deleted (B) and the strain confirmed as being EPEC. (C). The gene deletion was confirmed by the presence of the kanamycin cassette (B1) and its proximity to remaining *espC* β-domain DNA (B2) using PCR amplification. The strain was confirmed as that of enteropathogenic *E.coli* (EPEC) by PCR amplification of *espB* (C1) and *escN* (C2). PCR products for (B) and (C) were analysed on a 1% agarose gel by comparison to 1kb markers but are not shown above.

To verify purification by electro-elution was successful, electro-eluted protein was applied to another SDS-PAGE gel and transferred onto PVDF membrane in CAPS buffer at 80 mA overnight for visualisation with Naphthalene black (Fig 4.2A). The identity of electro-eluted protein material was confirmed by Kevin Bailey at B.S.A.U. (QMC, Nottingham) using a MALDI-TOF mass spectrometer. Following trypsin digestion and cysteine treatment with Iodoactemide (forms carbiodomethyl-cysteine or Cys-CAM); peptide mixtures were applied to the mass spectrometer (EspC histogram shown Fig 4.2B). The Peptident (now Aldente) peptide mass fingerprinting tool ([ca.expasy.org/tools/peptident.html](http://ca.expasy.org/tools/peptident.html)) was used to analyse peptide masses obtained (Wilkins and Williams, 1997), and probable matches found. 18 peptide matches were found to EspC (Accession P77070) and Enterotoxin EspC (Accession Q9EZE7) corroborating successful purification. Electro-eluted EspC was subsequently administered by subcutaneous immunisations to NZ White rabbits by the Bio-Medical Services Unit (B.M.S.U., Q.M.C) to generate polyclonal serum. Antibody specificity was confirmed against TCA precipitated EPEC secreted proteins of both wild type and its EspC mutant (this study) (Fig 4.2C and D).

#### **4.4 Characterisation of EspC mutant with polyclonal anti-EspC Sera.**

EPEC E2348/69 wild type and its isogenic *espC* mutant strain, were grown to OD<sub>600nm</sub> 0.4 - 0.5 (Time 0 min). Supernatant samples were then taken at 30 min intervals and TCA precipitated. These were then applied to an 11.5% SDS-PAGE gel and Western blotted with rabbit anti-EspC polyclonal antisera (Fig 4.2E). As the Figure shows, the polyclonal antiserum specifically recognises EspC secreted from wild-type EPEC, but not from the EPEC *espC* mutant. *espC*, incorporating *SacI* and *XbaI* restriction sites, was previously amplified from strain E2348/69 and cloned into the tightly controlled arabinose-inducible expression vector pBAD18 (Guzman *et al*, 1995) by Louise Arnold. Two different pBAD18 constructs producing EspC were obtained and designated pLAC1 and pLAC2 (Arnold, 2003). EspC was found to be constitutively expressed by pLAC1, in contrast to pLAC2 where tight expression (inducible) was achieved. pLAC2 was transformed into the EPEC *espC* mutant strain and selected on LB Carb 50 µg mL<sup>-1</sup>. Subsequent broth cultures were induced for 3 h with 0.02% arabinose. Western blotting with Rabbit anti-EspC polyclonal sera of supernatant proteins revealed complementation of EspC secretion (Fig 4.2E).

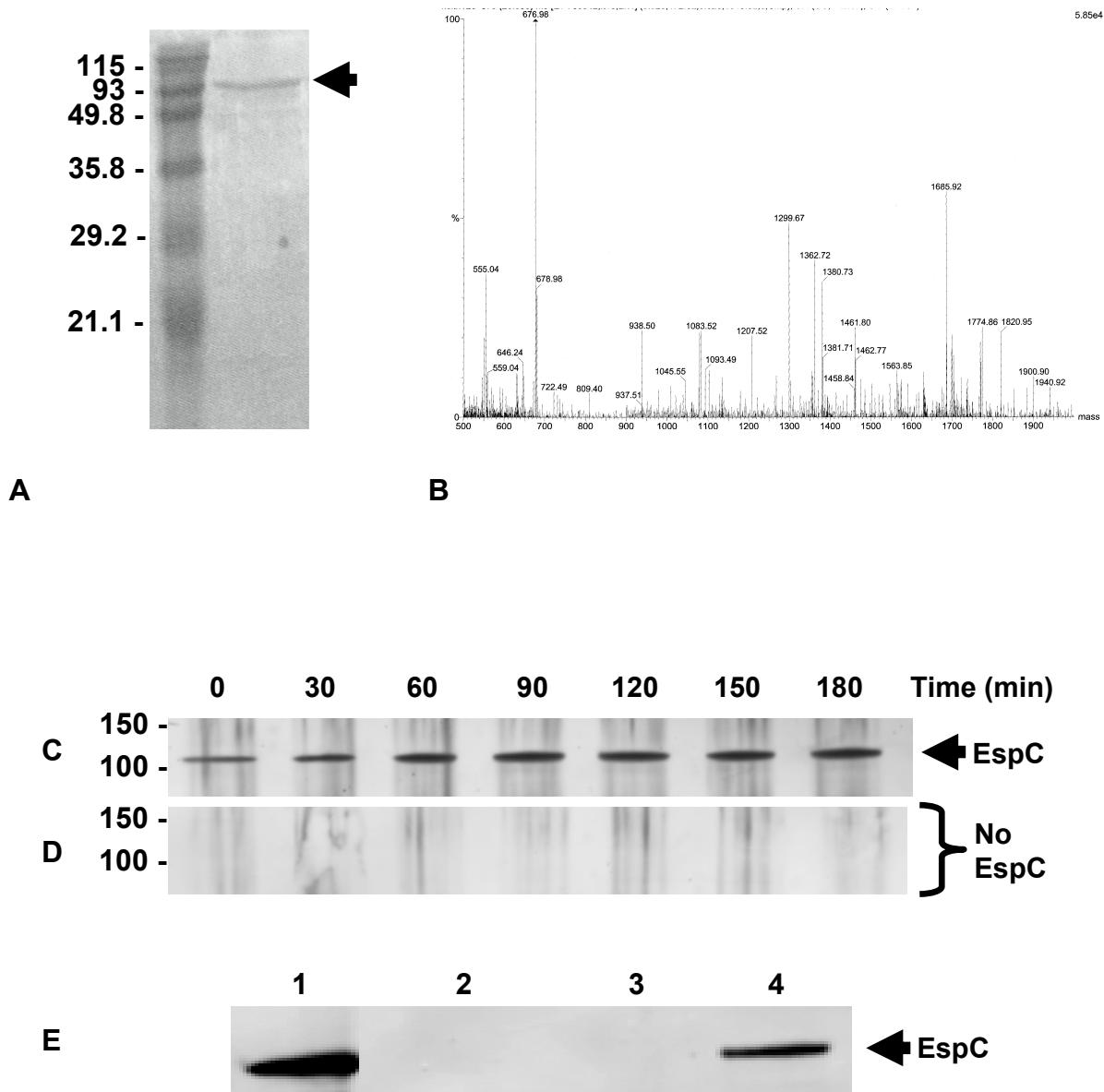


Figure 4.2: Verification of EspC purification and characterisation of an *espC* deletion mutant. Electro-eluted EspC was applied to an SDS-PAGE gel, and a single lane of EspC sample, in comparison with markers was Western blotted onto PVDF membrane and subjected to naphthalene black staining to confirm purification and concentration (A). Four lanes of electro-eluted EspC protein also separated by SDS-PAGE were excised and submitted for mass spectrometry using a MALDI-TOF mass spectrometer. Part of the mass spectrometer histogram is shown in (B). The Peptident peptide mass fingerprinting tool was used to analyse the peptide masses retrieved in the mass spectrometer histogram, the highest number of probable matches confirmed EspC as the purified protein. Polyclonal antiserum generated from electro-eluted EspC was confirmed to specifically recognises EspC in WT EPEC supernatants taken at 30 min intervals (C). It also confirmed EspC is absent from the supernatants of the EPEC *espC* mutant (D), corroborating the PCR amplifications demonstrating the *espC* gene deletion. The EspC mutant with and without plasmid pBAD18 (E2 and E3 respectively) was also confirmed by complementation with pLAC2 (E4) which over expresses EspC inducibly. Wild type strain E2348/69 secreted proteins were also analysed for comparison (E1) and shown above to specifically recognise EspC.

## **4.5. Generation of specific polyclonal anti-sera to EspB and FliC**

In order to assess any possible effects of EspC deletion on other known EPEC secretion systems implicated in the pathogenesis of EPEC, specific antisera was also generated to two key secreted components from (I) the virulence associated T3SS (EspB) and (II) the flagellar T3SS (FliC).

### **4.5.1. Construction of an EspB over-expression vector**

*espB* was successfully amplified by PCR (Fig 4.3A) using primers EspB\_For and EspB\_Rev in combination with an Elongase PCR optimisation kit according to manufacturers protocol. The *espB* DNA fragment was ligated into pCR2.1 and constructs recovered in host strain *E.coli* Novablue (Merck) by selection on LB kanamycin plates. pCR2.1 containing *espB* was purified using a Qiagen kit, and the *espB* fragment excised using an *Ncol/BamHI* restriction digest (Fig 4.3B) and then purified with a Qiagen gel extraction kit. Following ligation into pET30a, and transformation into Novablue with selection on LB kanamycin plates, the pET30a::*espB* construct was confirmed by *Ncol/BamHI* restriction digest (Fig 4.3D). Purified pET30a::*espB* was transformed into *E.coli* strain BL21(DE3) for expression trials.

### **4.5.2 Construction of a FliC over-expression vector**

*fliC* was cloned into the pET30a over-expression plasmid as described in section 2.6.14. *fliC* was amplified successfully by PCR, utilising primers FliC\_For and FliC\_Rev (Table 2.3), in all magnesium concentrations utilised (not shown). Unlike *espB* it proved possible to cut the amplified *fliC* PCR fragment with restriction enzymes *HindIII* and *EcoRI*, and directly ligate it into the pET30a over-expression vector (cut with *HindIII* and *EcoRI*). *E.coli* Novablue transformants were recovered on LB kanamycin plates. Colony PCR was utilised to identify plasmid constructs containing insert. Three transformants screened were found to contain pET30a::*fliC* (not shown). pET30a::*fliC* was purified using a Qiagen kit and transformed into chemically competent cells of *E.coli* strain BL21(DE3) to confirm expression of the cloned *fliC*.

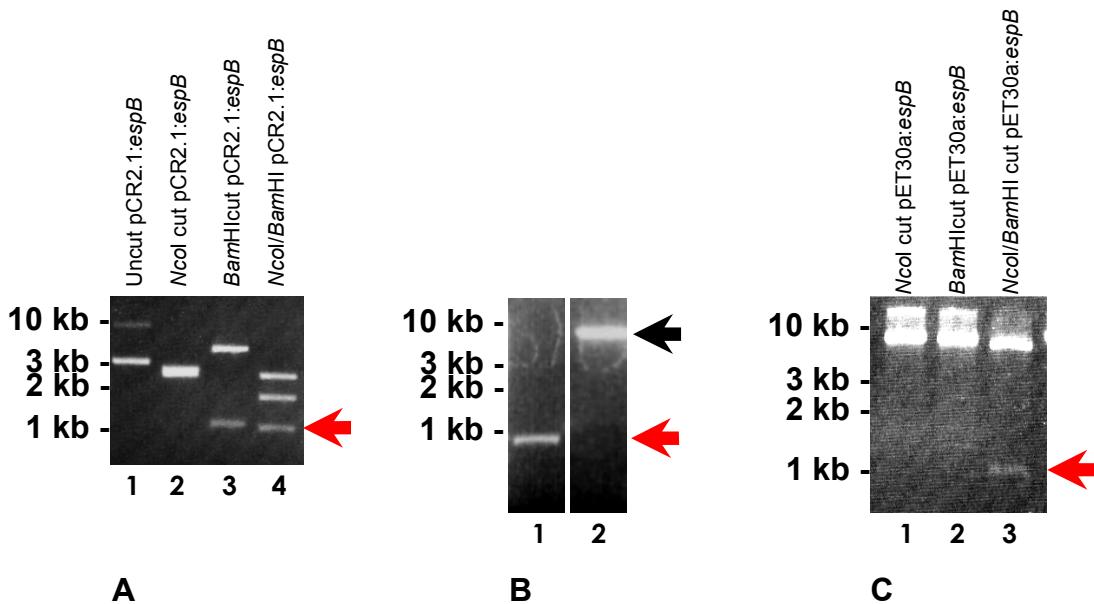


Figure 4.3 Cloning of *espB*. A single PCR product of approximately 1kb encoding *espB* was amplified successfully (not shown). The amplified fragment was integrated into PCR cloning vector pCR2.1, and plasmid construct recovered from *E.coli* Novablue transformants. Purified pCR2.1::*espB* was then subjected to restriction analysis (A). *espB* was cut from pCR2.1::*espB* utilising *NcoI* /*BamHI* (A4), the fragment was excised and gel purified (B1) along with *NcoI/BamHI* digested pET30a (B2). *espB* and pET30a were ligated together; the resulting construct was again recovered in *E.coli* Novablue transformants. Purified plasmid pET30a::*espB* was validated by restriction analysis (D) and confirmed by double restriction digest *NcoI/BamHI* (C3). pET30a::*espB* was then transformed into BL21(DE3) to confirm expression. The red arrow indicates *espB* and the black arrow indicates pET30a.

#### **4.5.3 Purification of EspB and FliC**

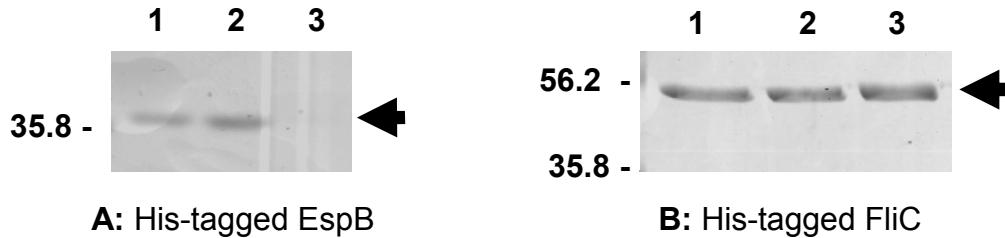
His-tagged EspB and FliC proteins were induced and over-expressed in *E.coli* strain BL21(DE3) and purified on a Novagen nickel binding column (Novagen, Merck Biosciences) (Fig 4.4). His-tag purified EspB and FliC were resolved by 11.5% SDS-PAGE and submitted for mass spectrometry as before to confirm the identity of the electro-eluted protein. His-tagged EspB was authenticated by 11 peptide matches to EspB (*eaeB*, Accession Q05129) and His-tagged FliC protein corroborated by 9 matches to FliC (Accession PO4949) (histograms not shown).

#### **4.5.4 Generation and validation of EspB and FliC polyclonal antisera.**

EspB and FliC polyclonal sera were checked for specificity and the presence of non-specific antibodies to EPEC proteins as follows; the effect of salt concentration on secretion by EPEC described by [Kenny et al. \(1997\)](#) was used to induce T3 protein secretion. Antibody specificity was tested by comparing immunoreactivity to proteins secreted by EPEC WT grown in L-Broth at either high ( $10\text{ g L}^{-1}$ ) or low salt concentration ( $0\text{ g L}^{-1}$ ), and also against purified His-tagged proteins. Secreted protein samples were concentrated by TCA precipitation (see Section 2.7.1), separated by SDS-PAGE, then immunoblotted with rabbit anti-EspB (Fig 4.5A) or rabbit anti-FliC (Fig 4.5B). Rabbit antibodies were detected with a protein-A alkaline phosphatase conjugate and visualised with NBT/BCIP chromogenic substrate. Both anti-EspB and FliC recognised purified protein confirming their specificity. Production of EspB protein was found to decrease at the lower salt concentration (corroborated by [Kenny et al, 1997](#)) whereas FliC was found to increase at lower salt concentration. FliC polyclonal antiserum exhibited some cross-reactivity with EspB, possibly because the NZ White rabbits used may have been previously infected with a rabbit specific enteropathogenic *E.coli* strain.

#### **4.6. Examination of EPEC protein secretion in the espC mutant**

The EspC mutation was further validated since no EspC was produced in an extended incubation compared to WT EPEC (Fig 4.5A). Whilst a similar level of EspB was observed in the WT and EspC mutant strain (Fig 4.5B), an elevated level of FliC protein was noted in the EspC mutant strain, thought likely to result from more efficient processing at the cell surface, more efficient secretion or possibly different growth rates.



A: His-tagged EspB

B: His-tagged FliC

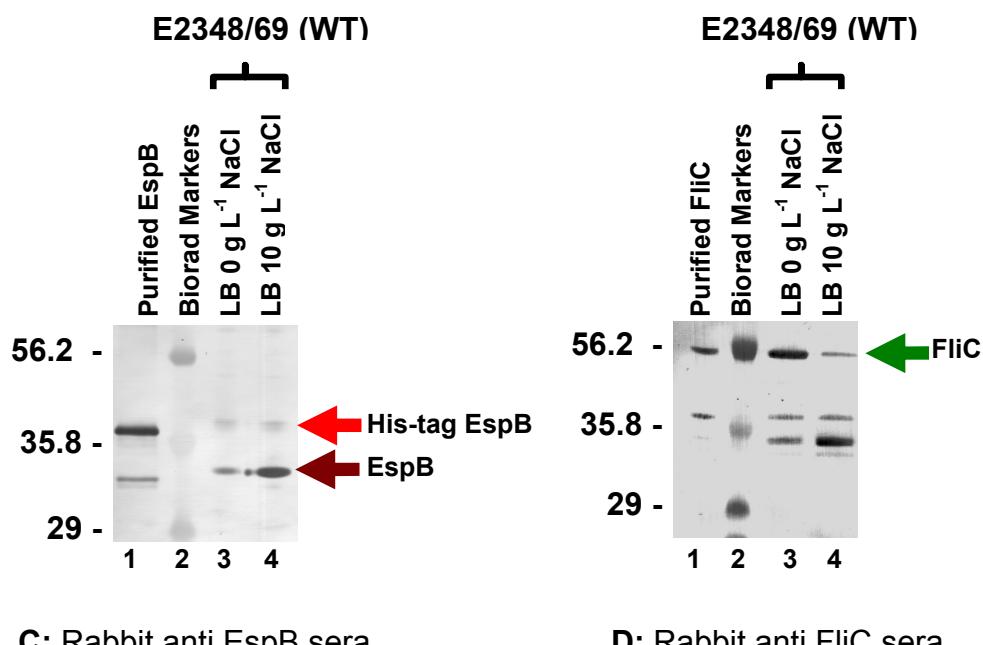


Figure 4.4: Generation of purified EspB and FliC and validation of the rabbit polyclonal serum generated against them. Both His-tagged EspB (A) and FliC (B) were affinity-purified using Nickel columns (Novagen). Individual elution fractions containing purified protein are shown. Both purified proteins were analysed by mass spectrometry. His-tagged EspB and FliC were run in parallel with TCA precipitated secreted proteins from the supernatant of EPEC E2348/69 grown in LB containing low and high salt concentrations, then separated by SDS-PAGE and immunoblotted using the specific rabbit anti-sera. Rabbit anti-EspB, recognised both His-tagged purified protein (C1) and non-his-tagged protein found in TCA preparations. In LB containing low salt concentration (C3) production of EspB was diminished compared to LB media with higher salt (C4). In contrast while rabbit anti-FliC also recognised purified protein(D1) in LB media with low salt concentration (D3) the production of FliC was found to be higher compared to LB media with higher salt (D4). Native FliC appears to be of similar size as the His-tagged counterpart, possibly due to methylation or glycosylation of FliC. Some non-specific cross-reactivity was also observed with a protein of the size expected for EspB. Pre-stained broad range markers (Biorad) were run to allow approximation/comparison of size.

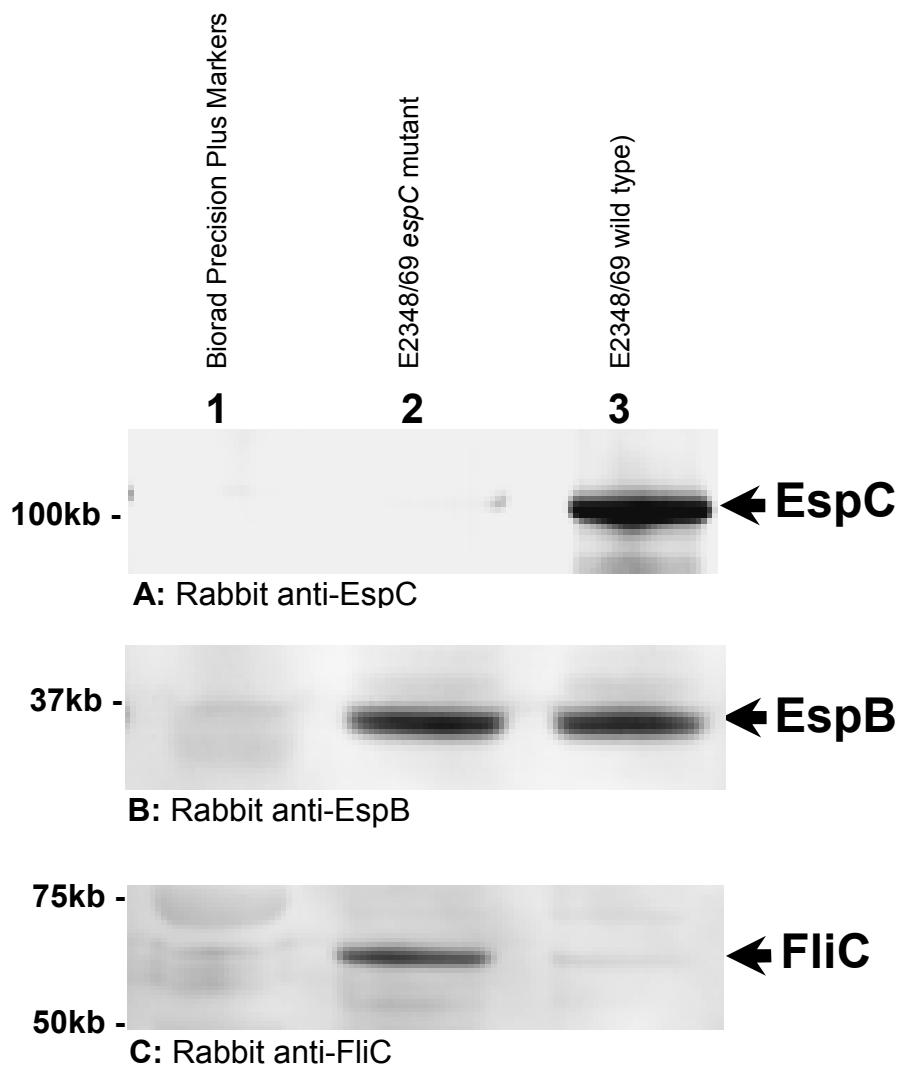


Figure 4.5: Western immunoblot characterisation of EPEC E2348/69 $\Delta$ espC. TCA precipitated supernatants from overnight cultures of EPEC E2348/69 (wild type) and its *espC* mutant were prepared and applied to an 11.5% SDS-PAGE gel then western blotted. Rabbit anti-EspC polyclonal antiserum confirmed EspC was not expressed over a longer time course in the EspC mutant in contrast to the wild type EPEC (A). Rabbit anti-EspB antiserum, additionally corroborated PCR analysis that both strains are EPEC and that EspB protein is secreted in equivalent quantities (B). Rabbit anti-FliC polyclonal antiserum, revealed that an elevated level of FliC protein is observed in the *espC* mutant compared to the wild type.

#### 4.7. Discussion

An EspC mutant was successfully generated using  $\lambda$  red mutagenesis and validated by (I) PCR, (II) complementation with pLAC2, and (III) a Western immunoblot using a specific anti-EspC polyclonal sera generated from electro-eluted EspC. An anti-EspB antiserum revealed similar levels of EspB were produced by both WT and the EspC mutant. Intriguingly however, and unanticipated, an elevated level of FliC in the *espC* mutant was uncovered by the use of anti-FliC antiserum. It is unclear if the difference is an artefact of *espC* mutagenesis, or indicative of a hitherto undescribed regulatory effect of EspC on FliC expression or secretion. Further investigation is warranted to resolve the issue.

FliC has been demonstrated to be up-regulated in conditions where salt concentration is lower, presumably as a function of an increase in motility, to move the bacteria to more favourable conditions for colonisation. The connection between motility and pathogenicity is apparent in many bacteria, and functional flagellae are often required for virulence (see Harshey and Toguchi, 1996), though in others motility must be suppressed (Akerley *et al*, 1995). Flagella (comprised of FliC) may also act as adhesins which aid binding of EPEC to microvilli.

Can EspC and flagella be linked in the literature? BipA is a member of the ribosome binding GTPase superfamily, and is known to be widely distributed in bacteria and in plants. Grant *et al*, (2003), demonstrated that it regulates multiple cell surface- and virulence-associated components in EPEC (strain E2348/69). The regulated components intriguingly encompass bacterial flagella, the *espC* pathogenicity island as well as the T3SS. It positively regulates *espC* and, independently of previously characterised regulators such as Per, positively regulates Integration host factor (IHF), H-NS and LEE gene clusters via control of Ler (LEE-encoded regulator). In contrast Grant *et al*, (2003) found that flagella- mediated motility was negatively regulated in a Ler-independent manner. BipA is believed to play a crucial role in the switch between the planktonic (motile) and virulent phase (non-motile, where T3SS proteins are expressed including EspC). The deletion of *espC* may have inadvertently disrupted this BipA regulatory pathway leading to increased expression of FliC or alternatively may be a part of the regulatory pathway itself. It had been thought that EspC may be involved in intimin degradation, but this was disproved in the work by Grant *et al*, (2003) Perhaps

EspC is instead a protease which acts to degrade flagella and/or bfp which are no longer required once intimate attachment has taken place. EspC and flagella can therefore be potentially linked in at least one possible regulatory pathway. Initial future work could address this possibility by using anti-BipA antibodies to investigate the expression of the regulator in both wild type and *espC* deletion strains.

Recent work by Soscia *et al*, (2007) showed that T3SS and flagella assembly are demonstrably linked in *Pseudomonas aeruginosa*. It is suggested that co-ordinate but inverse cross-regulation of the flagella and T3SSs occurs; the T3SS was found to be up-regulated in strains lacking flagella, and conversely flagella-gene expression and motility was increased when ExsA, a master regulator of T3SS was over-produced.

T3SSs have been described as the target for additional global regulators and two-component systems. In *E. coli* (and a number of other pathogenic species), SirA orthologs have also been shown to affect motility and virulence (Goodier and Ahmer, 2001). GacA in *P. aeruginosa* has a negative effect on T3SS whereas SirA has the opposite effect in *S. enterica* serovar Typhimurium. Soscia *et al*, (2007) also confirm a link to non-conventional sensors RetS and LadS. The RetS and LadS pathways act via a small regulatory RNA rsmZ, a target gene of GacA. RmsA negatively controls the production of homoserine lactone (HSL) and has a positive effect on T3S (Mulcahy *et al*, 2006; O'Grady *et al*, 2006). The data of Soscia *et al*, (2007) also confirms previous work indicating that GacA activates synthesis of N-butanoylhomoserine lactone (C<sub>4</sub>-HSL), a quorum sensing molecule, and that a transcriptional regulator RhIR binds C<sub>4</sub>-HSL to repress the T3SS regulon. Soscia *et al* (2007) further assert that in *P. aeruginosa* there is direct cross-regulation of T3SS and flagella and an indirect control, mediated by GacA that act in concert to co-ordinate expression of T3SS and motility genes. EspC may form an important link in this type of regulatory pathway. The *espC* deletion mutant generated in this study may have inadvertently restored FliC expression which otherwise would be inhibited.

Together with the bundle forming pili, mutation of *fliC* has also been shown to reduce adherence (Girón *et al*, 2002). Pratt and Kolter (1998) have demonstrated through microscopic analyses of motility mutants that motility is important in initial interaction with, and movement along a surface. Motility would likely be a disadvantage in a situation where translocation of proteins via the T3SS was favoured. The T3SS is also

an anti-phagocytic system (Frithz-Lindsten *et al*, 1997; Hauser *et al*, 1998). Could EspC inhibit phagocytosis of EPEC? In pathogenic strains of *S. typhi*, Eichelberg and Galán (2000) found that FliA (a flagellar Sigma factor) when deleted had significant impact on invasion. Anti-phagocytosis has been suggested to be a pathogenic mechanism for EPEC (Celli *et al*, 2001). In recently published work, EspF and EspJ of EPEC and EHEC have been revealed to have roles in anti-phagocytic mechanisms though the molecular basis for it remains unclear (Marchès *et al*, 2008). In *P. aeruginosa*, defects in T3SS proteins PepB and PepD, homologues of YopB and YopD (*Yersinia* sp.) and EspB and EspD (EPEC), result in loss of cytotoxicity and an increase in the level of internalisation (Hauser *et al*, 1998). Presumably EspC would function differently as EspB/EspD are associated with the molecular syringe, explaining the reduction of cytotoxicity when these proteins are defective. When EspC is absent, the conditions may no longer be suitable for EPEC colonisation – the bacteria may be internalised and likely destroyed by intestinal cells. The requirement to find more suitable environmental conditions may also explain the up-regulation of FliC. In the absence of EspC, FliC is apparently expressed at increased levels. Possibly FliC is degraded by EspC and in its absence is detectable at increased concentrations.

Further λ-red mutagenesis could be instigated to either delete *fliC* from the EspC mutant generated in this study, or EspC from the FliC mutant already generated previously (Ferrandiz *et al*, 2005). Complementation studies are also required to determine whether FliC can be down-regulated again once EspC expression is restored; over-expression plasmids encoding EspC are already available for this. The converse should also be investigated in the double mutant background – does complementation with a plasmid encoding *fliC* reduce EspC expression or impact on T3SS? Potential interactions could also be investigated between the two proteins, as the *fliC* gene has already been cloned and His-tag is known to be detectable in Western blots. This is a relatively straight forward experiment to follow up. In the absence of a direct FliC-EspC interaction, there is a possibility that a common regulatory gene or pathway (discussed earlier) may have been affected instead. However, time constraints did not allow these possibilities to be further examined.

## **5: Construction and use of yeast two-hybrid vectors.**

### **5.1. Introduction**

To identify interactions occurring between the EspC domains or between EspC and YbgC, the Y2H system was utilised. To enable this analysis, DNA encoding each EspC domain and also YbgC was cloned into the Y2H system vectors pGBT9 and pGAD424. This created fusions of the test protein to either the binding (pGBT9) or activation (pGAD424) domain of the yeast transcriptional activator GAL4. Both Y2H vectors bear an ampicillin resistance gene (*Amp<sup>R</sup>*) for selection in bacterial hosts. Furthermore, pGBT9 and pGAD424 respectively carry *TRP1* or *LEU2* markers to permit highly specific nutritional selection upon transformation into the multiple auxotrophic *S. cerevisiae* strain PJ69-4A. Additionally, because the Y2H system is occasionally subject to the identification of false or physiologically irrelevant protein interactions, demonstration of an interaction using other biochemical approaches is required for validation. To that end, the over-expression vectors pET30a and pMALC2X were used for the preparation of recombinant His-tagged or MBP fusion protein respectively. Purified recombinant protein was then used in co-purification experiments for the validation of protein-protein interactions.

### **5.2 Generation of *espC* passenger-domain and *ybgC* Y2H vector constructs.**

The amplification of each target DNA fragment comprising (I) EspC β-domain, (II) EspC inter-domain, (III) EspC passenger-domain, and (IV) YbgC was successfully achieved using the primer pairs (A) Barf1 and Barr1, (B) Interf1 and Interr1, (C) PasFor1 and PasRev1, and (D) YbgCF1 and YbgCR1a respectively. There was minimal contamination from secondary non-specific PCR products. The final set of PCR reaction mixtures and conditions utilised are shown in Table 5.1A and 5.1B respectively. The amplified PCR products comprising fragments of 846 bp (β-domain), 812 bp (inter-domain), 2038 bp (passenger-domain), and 410 bp (*ybgC*) (Fig 5.1A-D) were then subjected to restriction digestion with *EcoRI*. Y2H vectors pGBT9 and pGAD424 were also digested with *EcoRI* and in addition dephosphorylated by treatment with alkaline phosphatase. Vector DNA was prepared in bulk and purified by gel extraction. Direct cloning of the *espC* passenger-domain and *ybgC* into pGBT9 and pGAD424 initially proved to be unsuccessful. A multi-purpose blunt cloning vector, pSTBlue1 (Merck Biosciences) was therefore used for initial cloning of these PCR products and blue/white

**A: Reaction Mixtures**

PCR Component	β-Barrel domain PCR	Inter-domain PCR	Passenger-domain PCR	YbgC PCR
Primers (see Table 2.3) Primer Volume (10pmol/μL)	BarF1 & BarR1 5 μL each primer	Interf1 & Interr1 5 μL each primer	PasFor1 & PasRev1 5 μL each primer	YbgCF1 & YbgCR1a 5 μL each primer
10x Buffer 10X Buffer Volume	Thermpol (NEB) 10 μL	Thermopol (NEB) 10 μL	KOD (Merck) 10 μL	Thermopol (NEB) 10 μL
dNTPs (2.5 mM) (Promega)	16 μL	16 μL	16 μL	16 μL
MgSO <sub>4</sub> (final Concentration)	3, 10, 20 mM	3, 10, 20 mM	1.25, 2.5, 3.75 mM	3, 10 mM
Genomic DNA	2 μL	2 μL	2 μL	2 μL
Polymerase Polymerase Volume	Deep Vent (NEB) 1 μL	Deep Vent (NEB) 1 μL	KOD Hotstart (Merck) 1 μL	Deep Vent (NEB) 1 μL
DMSO (Sigma)	5 μL	5 μL	5 μL	0 and 5 μL
BSA 2 mg ml <sup>-1</sup> (Promega)	-	-	-	0 and 2 μL
dH <sub>2</sub> O	Add to 100 μL	Add to 100 μL	Add to 100 μL	Add to 100 μL

**B: Reaction Conditions**

	Temperature	Time	No of Cycles
<b>Initial Denaturation</b>	95°C	5 min	1 cycle
<b>Amplification (β &amp; inter-domain)</b>			
Denaturation	95°C	30 s	
Annealing	55°C	30 s	30 Cycles
Extension	68°C	1 min	
<b>Amplification (passenger-domain)</b>			
Denaturation	95°C	45 s	
Annealing	55°C	45 s	30 Cycles
Extension	68°C	2 min 15 s	
<b>Amplification (ybgC)</b>			
Denaturation	95°C	45 s	
Annealing	55°C	30 s	30 Cycles
Extension	68°C	1 min	
<b>Final Extension</b>	68°C	15 min	1 cycle

Table 5.1. PCR reaction mixtures and conditions for amplification of DNA encoding the functional EspC domains and YbgC. The main variations to the reaction included altering the DNA polymerase, 10X polymerase buffer or MgSO<sub>4</sub>. BSA and DMSO were also utilised to help circumvent secondary structure formation in optimisation steps. All PCR reaction mixtures (A) had a total volume of 100 μL. Initial denaturation and final extension conditions for the three functional EspC domains and YbgC remained constant, only the amplification stage was varied for the three EspC domains and YbgC (B).

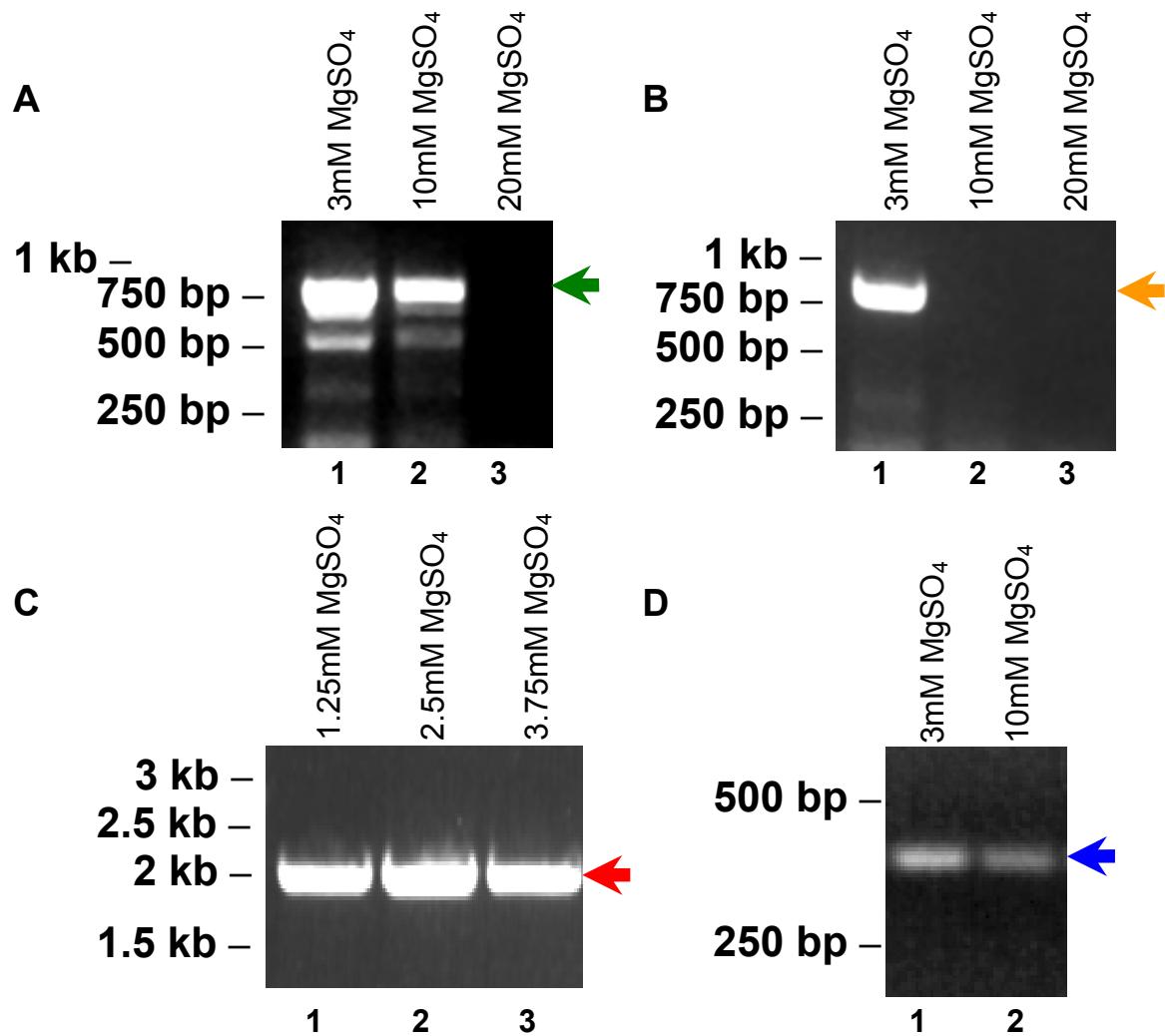


Fig 5.1: Optimisation of PCR amplification of DNA encoding EspC β-, inter- and passenger-domains, and *ybgC*. DNA encoding either the EspC β- (A), inter- (B), and passenger-domains (C) or *ybgC* (D) from EPEC E2348/69 was successfully amplified by PCR under differing concentrations of MgSO<sub>4</sub>. Products are shown following separation through a 1% agarose gel in comparison with 1 kb ladder (Promega). Arrows indicate products of the expected size following optimisation of buffer and cycling conditions. PCR products were excised, and then purified using a Qiagen gel extraction kit. Purity was verified by agarose gel electrophoresis. Following restriction digestion, DNA fragments were included in ligation reactions with pGBT9 and pGAD424.

colony selection used to select transformants hosting plasmid with an insert. The insert DNA was excised from the pSTBlue1 derivatives using *Eco*RI, and gel purified. Subsequent ligations resulted in successful cloning of both fragments into the *Eco*RI/alkaline phosphatase treated Y2H vectors.

### **5.3 Orientation of *espC* domain fragments and *ybgC* within pGBT9 and pGAD424.**

As all DNA fragments inserted into pGBT9 and pGAD424 bore *Eco*RI at both ends, it was necessary to determine their orientation in the vectors. This was achieved by PCR amplification. In pGBT9 constructs, pGBT-F primer was used in all amplifications in combination with (I) BarR1 primer (Fig 5.2A) to amplify β-domain, (II) Interr1 primer to amplify inter-domain (Fig 5.2B), (III) PasR1 to amplify the passenger-domain (Fig 5.2C), and (IV) YbgC1a to amplify *ybgC* (Fig 5.2D). In pGAD424 constructs, pGAD-F primer was used with the same gene-specific primers as noted above. Additionally, in confirming the inter-domain an alternative set of primers, in (I) pGBT9; pGBT-R and Interf1, and in (II) pGAD424; pGAD-R and Interf1 were also employed (Fig 5.2B). *espC* β- and inter-domains were cloned successfully into pGBT9 and pGAD424 in the correct orientation (Fig 5.2A and 5.2B). In addition, all pGBT9 and pGAD424 constructs were further successfully validated by *Eco*R1 restriction digests (Fig 5.2E and 5.2F).

For pGBT9 and pGAD424 constructs, the primer pairs (I) pGBT-F and pGBT-R and (II) pGAD-F and pGAD-R were used respectively for sequencing. In addition, for the longer passenger-domain a series of 26 internal primers were also utilised (see Table 2.2). Subsequent sequencing of inserts showed no errors had been introduced. However, difficulties were encountered in cloning the passenger-domain and *ybgC* gene, and were respectively overcome by (I) by changing to the proof reading enzyme KOD Hi-fidelity HotStart Polymerase, which permitted the use of a much shorter extension time (2min 15s) for the passenger-domain compared to that of Deep Vent polymerase used previously and, (II) synthesis of a new reverse primer (*ybgCR1a*) allowing generation of error-free *ybgC* constructs (early constructs containing *ybgC* were found to contain frameshift mutations).

### **5.4. Y2H protein-protein interactions using *espC* and *ybgC* Y2H constructs.**

The *ybgC* and *espC* Y2H constructs generated were utilised to investigate potential interactions between the encoded proteins. A series of test interactions were devised so

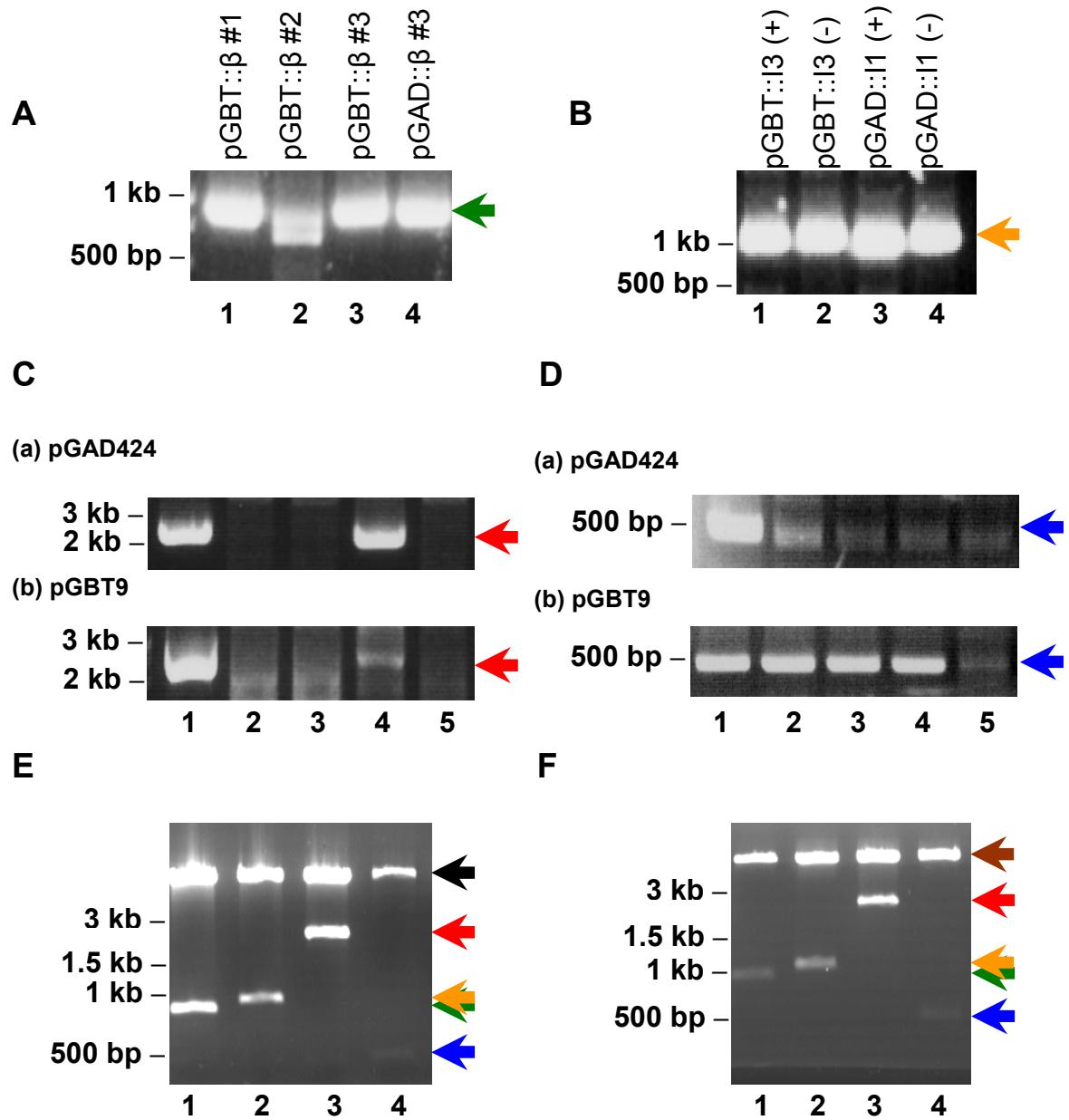


Figure 5.2: Verification of the pGAD424 and pGBT9 constructs by PCR, and EcoRI restriction digests. Purified Y2H plasmid containing β- (A), inter- (B), passenger domain (C) or *ybgC* insert (D) was used in PCR amplifications utilising vector forward and insert reverse primer combinations to test for the positive orientation. Where an alternative combination of primers was utilised (vector reverse/insert forward) the original primer set has been denoted as (+) and the alternative as (-). In addition EcoRI restriction of the pGAD424 (E) and pGBT9 (F) constructs was also carried out. Products are shown following separation through a 1% agarose gel in comparison with 1 kb ladder (Promega). Arrows indicate products of the expected size and are indicated by coloured arrows; green (β-domain), orange (inter-domain), red (passenger-domain) *ybgC* (blue), pGAD424 (black) and pGBT9 (brown). The constructs were sequenced to verify that no errors had been introduced during PCR amplification.

that a construct encoding (I) the *GAL4* binding domain or (II) the *GAL4* activation domain fused to either (I) *ybgC* or (II) an *EspC* segment were paired with each other (Table 5.2). Where interactions occurred between fusion proteins, the *GAL4* transcriptional unit was re-established facilitating growth on selective media, and identification and elimination of self-activating or non-interacting constructs. Interactions were initially selected using media plate assays. Negative control interactions comprised pGAD424 or pGBT9 constructs encoding either *YbgC* or *EspC* domains in combination with the appropriate Y2H vector without insert. Positive control constructs (intermediate and strong) were provided by Dr R Delahay ([Delahay et al, 2002](#)).

#### **5.4.1 Plasmid combinations used in Y2H plate assays**

pGAD424/pGBT9 fusion constructs were initially co-transformed into *S. cerevisiae* and grown for 3-5 days on MUHA plates (YMM plates supplemented with Methionine, Uracil, Histidine and Adenine) to select for yeast cells transformed with both vectors. Colonies grew in all cases as expected. Colonies were replica plated from MUHA plates onto MUH (intermediate stringency YMM plates), MUA (low stringency selection) or MU (high stringency selection) plates. Functional restoration of Gal4 transcriptional activity by interaction between test fusions was indicated by growth on selective media after 3-5 days incubation at 30°C (see Table 5.2). Selective plate assays were conducted on three separate occasions and the plates rated according to number of colonies. A subjective rating system based upon the number of colonies observed and their characteristics was employed in the initial screen; X: no growth occurred. 1\* : 0-10 colonies, 2\* : 11-20 colonies, 3\* : 21-30 colonies, 4\* : 31 to 50 colonies and 5\* :  $\geq 50$  colonies. Colony size, and/or pigmentation were used to adjust colony counts; where colonies were found to be pinkish/red (normally white/cream) and/or grew very poorly on selection media (colonies appeared shrivelled and in poor condition), the rating was reduced. Both reduced size/poor growth and discolouration was considered to imply that only partial or transient restoration of the *GAL4* transcriptional activator had occurred. Test interactions were compared to control interactions to resolve whether further quantitative screening was warranted. In the initial screening process, self-activating and non-interacting constructs were eliminated from further study, including: (I) pGBT9:: $\beta$  (Table 5.2) which facilitated high levels of growth on selective media when used in combination with (A) empty pGAD424 vector and (B) other *espC* constructs, and (II) *YbgC*, where no demonstrable interaction with either  $\beta$ -, inter- or passenger-domain was

Test Plate Number	Plasmid Combination Tested in <i>Saccharomyces cerevisiae</i> PJ69-4A	Test / Control	Initial Screening				Secondary Screening β-gal assay β-Gal Units (mean ± SD, n=9)	
			Growth on MUH Plates					
			1	2	3	AV		
<b>No Insert Controls</b>								
12	pGBT9: No Insert/ pGAD424: No Insert	Neg	X	X	X	X	3.89 ± 0.26	
13	pGBT9: No Insert /pGAD424::β	Neg	X	1*	1*	1*	ND	
14	pGBT9: No Insert /pGAD424::inter	Neg	X	X	X	X	3.61 ± 0.22	
15	pGBT9: No Insert /pGAD424::passenger	Neg	1*	1*	1*	1*	4.15 ± 0.22	
16	pGBT9: No Insert /pGAD424::ybgC	Neg	X	X	X	X	6.51 ± 1.17	
21	pGBT9::β/ pGAD424:No Insert	Neg	5*	5*	5*	5*	ND	
26	pGBT9::inter / pGAD424:No Insert	Neg	3*	3*	2*	3*	17.54 ± 4.41	
31	pGBT9::passenger / pGAD424:No Insert	Neg	1*	1*	2*	1*	4.76 ± 0.56	
36	pGBT9::ybgC / pGAD424:No Insert	Neg	X	X	X	X	1.038 ± 0.16	
<b>Positive Controls – Supplied by Dr R Delahay</b>								
37	pGBT9: <i>tir</i> / pGAD424:cesT	Pos	5*	5*	5*	5*	121.40 ± 4.69	
38	pGBT9: cesT/ pGAD424:cesT	Pos	4*	3*	3*	3*	8.16 ± 2.94	
<b>Test Interactions</b>								
17	pGBT9::β/ pGAD424::β	Test	5*	5*	5*	5*	ND	
18	pGBT9::β/ pGAD424::inter	Test	5*	5*	5*	5*	ND	
19	pGBT9::β/ pGAD424::passenger	Test	5*	5*	5*	5*	ND	
20	pGBT9::β/ pGAD424::ybgC	Test	5*	5*	5*	5*	ND	
22	pGBT9::inter / pGAD424::β	Test	4*	3*	3*	3*	ND	
23	pGBT9::inter / pGAD424::Inter	Test	4*	4*	5*	4*	ND	
24	<b>pGBT9::inter / pGAD424::passenger</b>	<b>Test</b>	<b>5*</b>	<b>5*</b>	<b>5*</b>	<b>5*</b>	<b>21.77 ± 3.39</b>	
25	pGBT9::inter / pGAD424::ybgC	Test	4*	4*	2*	3*	ND	
27	pGBT9::passenger / pGAD424::β	Test	2*	1*	2*	2*	ND	
28	<b>pGBT9::passenger / pGAD424::inter</b>	<b>Test</b>	<b>5*</b>	<b>3*</b>	<b>4*</b>	<b>4*</b>	<b>16.03 ± 2.86</b>	
29	<b>pGBT9::passenger / pGAD424::passenger</b>	<b>Test</b>	<b>4*</b>	<b>2*</b>	<b>2*</b>	<b>3*</b>	<b>8.33 ± 3.81</b>	
30	pGBT9::passenger / pGAD424:ybgC	Test	1*	1*	1*	1*	ND	
32	pGBT9::ybgC / pGAD424::β	Test	X	X	1*	X	ND	
33	pGBT9::ybgC / pGAD424::inter	Test	X	X	X	X	ND	
34	pGBT9::ybgC / pGAD424::passenger	Test	X	X	X	X	ND	
35	<b>pGBT9::ybgC / pGAD424::ybgC</b>	<b>Test</b>	<b>4*</b>	<b>4*</b>	<b>3*</b>	<b>4*</b>	<b>50.93 ± 4.78</b>	

Abbreviations: NC – Negative Control, Test – Test Interaction, Pos – Positive control, ND – Not done

Table 5.2: Summary of Y2H interactions. Initial screening was conducted on three occasions and plates rated according to number of colonies. X: no growth occurred. 1\* : 0-10 colonies, 2\* : 11-20 colonies, 3\* : 21-30 colonies, 4\* : 31 to 50 colonies and 5\* : ≥50 colonies. Subjective criteria including size and pigmentation was used to reduce the score of a plate (reduced sized or altered pigmentation were indicative of stress, and consequently a non/transiently functional *GAL4* transcriptional unit). Initial screening aided selection of test interactions for secondary screening (β-galactosidase activity). All controls and test interactions subjected to secondary screening are indicated in bold above. (in blue – negative controls, in green – positive controls, and in red the putative test interactions).

revealed (Table 5.2). Where growth on selective media was good compared to negative controls, constructs were selected for quantitative analysis of  $\beta$ -galactosidase activity.

#### 5.4.2. Analysis of quantitative $\beta$ -galactosidase assays

Four putative test interactions were further assessed by the  $\beta$ -galactosidase liquid assay including (I) pGBT::inter / pGAD424::passenger (24), (II) pGBT::passenger / pGAD424::inter (28), (III) pGBT::passenger / pGAD424::passenger (29), and (IV) pGBT::*ybgC* / pGAD424::*ybgC* (35) (all highlighted in emboldened red text in Table 5.2).  $\beta$ -galactosidase activities have been averaged for three transformants, and each assayed in triplicate ( $\pm$  standard deviations) (Fig 5.3 and Table 5.2).

The Student's test for unpaired data was used to carry out statistical analysis of the  $\beta$ -galactosidase activity revealed that all four test interactions were significant compared to each of their appropriate negative controls ( $p<0.05$ ). The relatively high level of background  $\beta$ -galactosidase activity for the pGBT9::inter construct in particular, suggested it self-activates to a low level, however the statistical analysis has uncovered this is not sufficient to mask the interaction involving this construct. Furthermore, (II) and (IV) were highly significant when compared to each appropriate negative control ( $p<0.01$ ). The strength of the self interaction between YbgC proteins suggests that a dimer or multimer may be formed.

#### 5.5. Further definition of the interacting inter- and passenger-domain regions.

In this study, the EspC inter-domain boundary was approximated on the basis of a sequence alignment with the Hbp AT for which the inter-domain has been more precisely defined. Not all ATs bear an inter-domain, and sequence variation between those that do blurs the precise domain boundary. Consequently, we sought to more precisely define the putative inter-domain region of EspC. We also aimed to further define the region of the passenger-domain responsible for interactions with the inter-domain. To this end, several truncations of both inter and passenger-domains were considered for Y2H interaction screens.

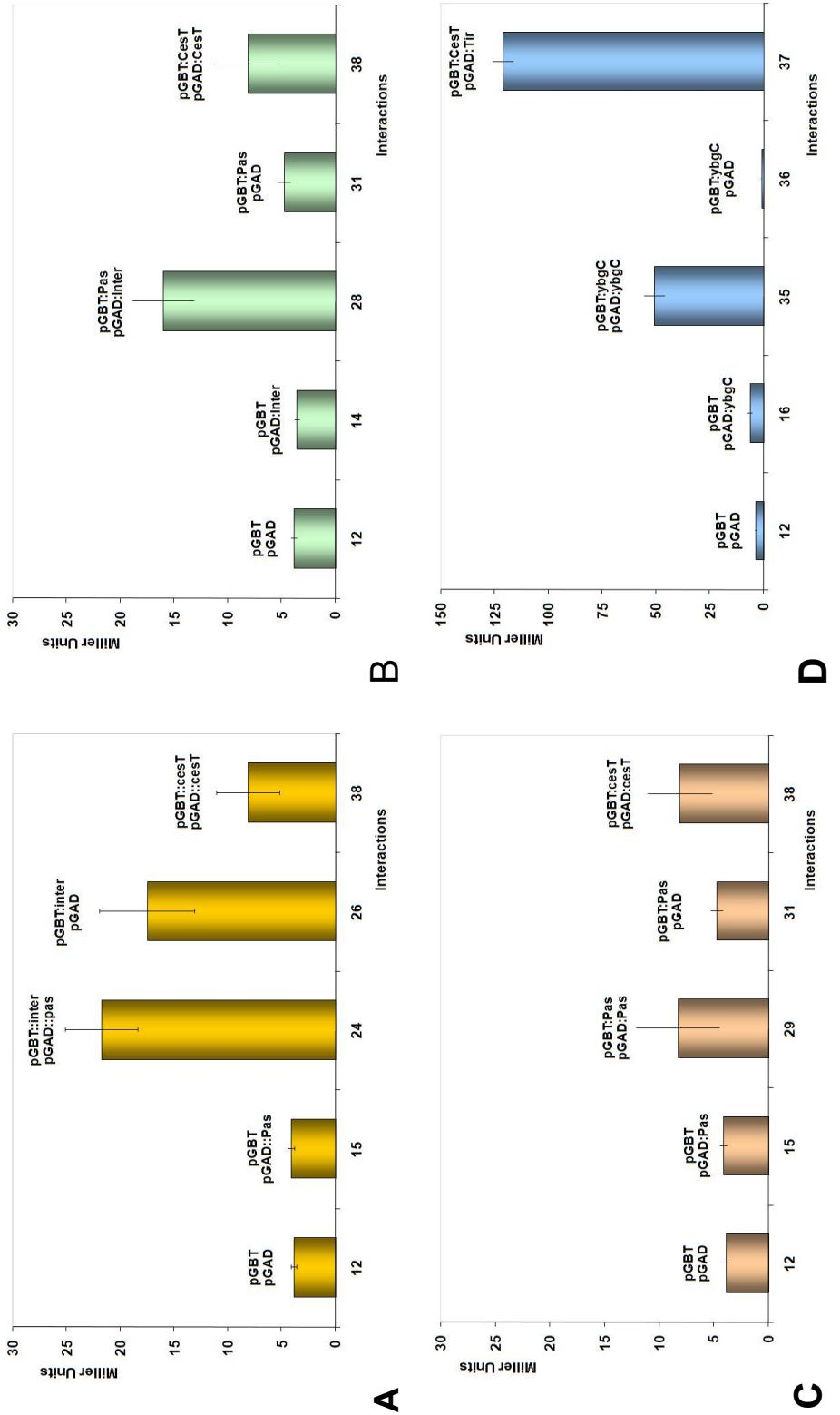


Figure 5.3: Quantitative assay of  $\beta$ -galactosidase activity. Four putative interactions (A) pGBT::inter / pGAD424:::passenger (24), (B) pGBT::passenger / pGAD424:::inter (28), (C) pGBT::passenger / pGBT::ybgC / pGAD424:::ybgC (35) were assayed and compared against the appropriate negative and positive controls in host *S. cerevisiae* strain P J69 -4A. Negative controls include (I) both Y2H vectors harbouring no insert (12), or (II) a single insert fused to either the binding or activation domain of GAL4 as hosted in the test interaction as follows (A) 15, and 31, (B) 14 and 31, (C) 15 and 31, and (D) 16 and 26.  $\beta$ -Galactosidase specific activities are calculated as Miller Units. The values are averages for three transformants, each assay triplicate ( $\pm$  standard deviations). All test interactions were significant compared to each appropriate negative control ( $p < 0.05$ ).

### **5.5.1. Generation of Y2H constructs encoding truncated passenger-domain.**

A schematic of the primer locations to amplify the five passenger sub-domains are indicated in Fig 5.4A, and have been designated PS1, PS2, PS3, PS4, and PS5. EPEC E2348/69 genomic DNA was used as template in PCR to generate passenger sub-domain fragments (~450 bp) using the following primer combinations for PS1, PS2, PS3, PS4, and PS5 respectively (I) PasF1 and PS1\_Rev1, (II) PS2\_For1 and PS2\_Rev1, (III) PS3\_For1 and PS3\_Rev1, (IV) PS4\_For1 and PS4\_Rev1, and (V) PS5\_For1 and PasR1.

All primers incorporated *Eco*RI restriction sites, and PCR amplification was carried out as described previously for the passenger-domain with the following modifications; MgSO<sub>4</sub> stock was diluted to an extended range of final concentrations of 1.25 mM, 2.5 mM and 3.75 mM, 5.0 mM, and 6.25 mM. 2 µL BSA (2 mg mL<sup>-1</sup> concentration) was included in all PCR reaction mixtures, annealing temperature was reduced to 52°C, 45 s, and extension time reduced to 45 s (final extension time remained the same).

### **5.5.2 Selection and amplification of passenger sub-domains**

The DNA encoding passenger sub-domains (PS1, PS2, PS3, PS4, and PS5) were successfully amplified in all cases (Fig 5.5). Direct (ligated following restriction digest) and indirect cloning (use of a PCR cloning vector prior to restriction digest) strategies were employed for all passenger sub-domains. No direct cloning strategy was successful.

Despite multiple attempts, only three passenger sub-domains, PS2, PS3, and PS4 were successfully cloned into the PCR cloning vector pSTBlue1 (not shown). Ultimately only sub-domain PS3 was successfully cloned into one Y2H vector, pGBT9 (Fig 5.6A). It was also successfully cloned into the pMALC2X over-expression vector (Fig 5.6B). A number of efforts to vary ligation conditions did not result in cloning of PS2 or PS4 into Y2H vectors.

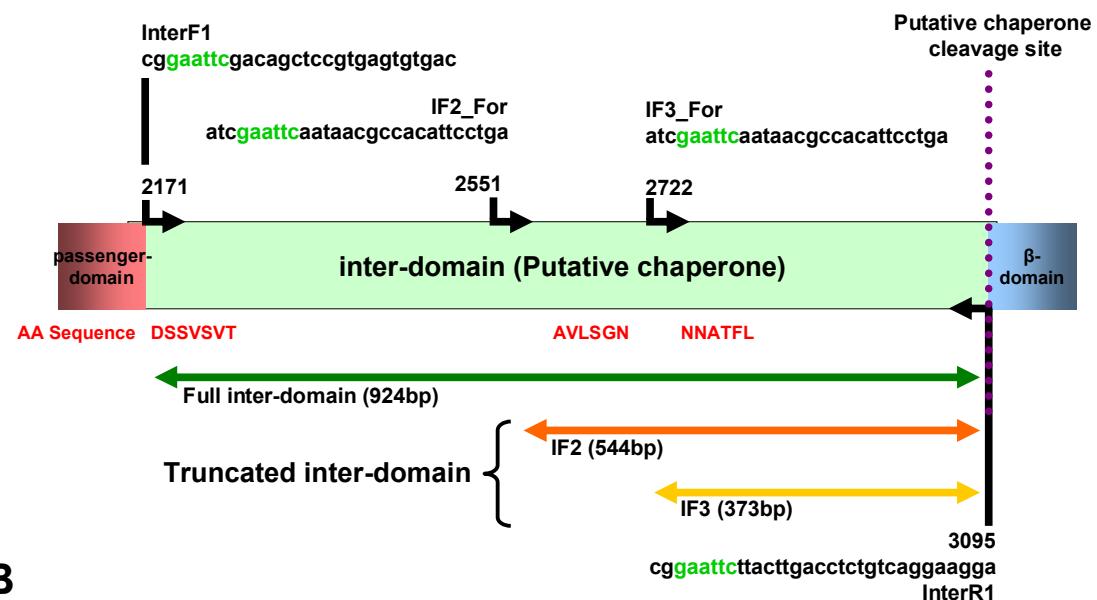
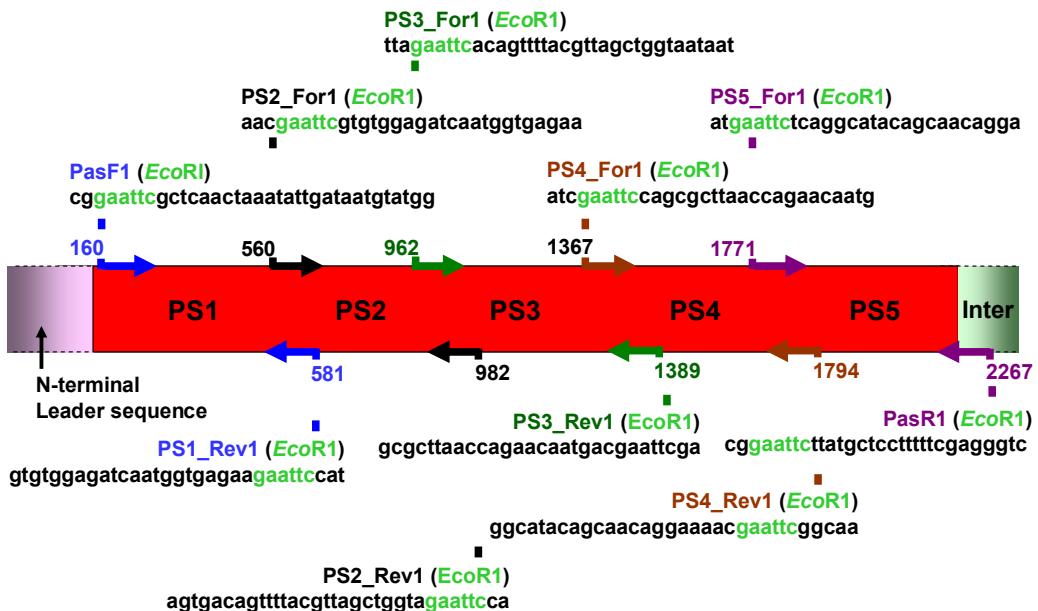


Figure 5.4: Schematic representation of passenger sub-domains and truncated inter-domain fragments. The name and sequence of primers used to amplify *espC* encoded passenger sub-domains (A) and truncated inter-domains (B) are indicated in bold type, their relative positions and direction are specified by arrows and the primer name highlighted in bold type. For each of the passenger sub-domains, PS1 (blue), PS2 (black), PS3 (dark green), PS4 (brown), and PS5 (purple) the primer pairs are highlighted in the same colour. For the inter-domain (B) the length of the DNA fragments encoded by the sets of the primers are also indicated including the full length inter domain (dark green), IF2 truncated inter-domain (orange) and IF3 truncated inter domain (gold). Also indicated are the approximate amino acid (AA) sequence equivalent to primer binding sites and the location of the putative chaperone cleavage site. *EcoRI* restriction sites incorporated into the primer shown in green within the primer sequence. All primer sequences are shown in 5' to 3' direction.

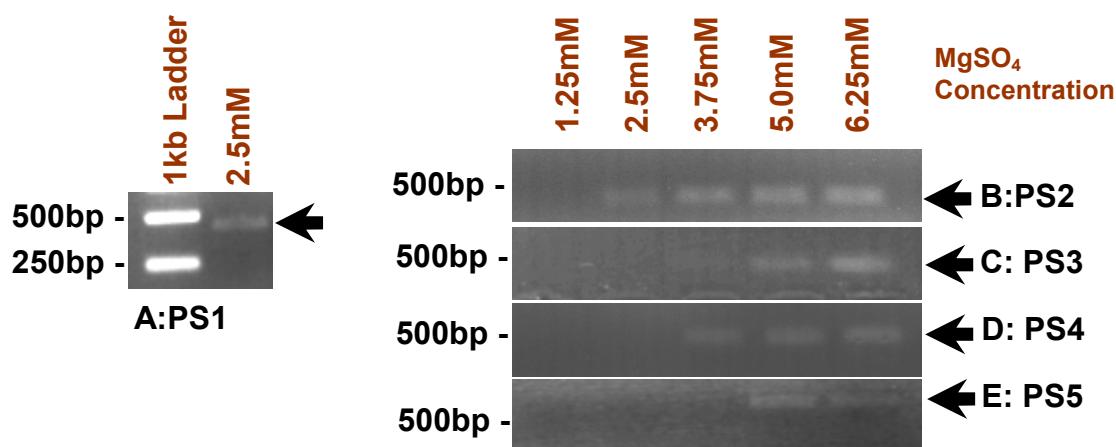


Figure 5.5: Optimisation of PCR amplification of DNA encoding EspC passenger sub-domains. Crude genomic DNA from EPEC E2348/69 was used as template with KOD Hotstart *Taq* Polymerase (Merck Biosciences) in the supplied buffer containing DMSO and differing concentrations of  $MgSO_4$  for amplification of EspC Passenger sub-domains PS2, 3, 4, and 5. PCR Products are shown following separation through a 1% agarose gel in comparison with 1 kb ladder (Promega). Product of the expected size for the Passenger sub-domains was excised from the agarose gel and purified using a Qiagen gel extraction kit. Purity was verified using agarose gel electrophoresis and following restriction digestion this DNA fragment was included in ligation reactions with pGBT9, pGAD424, pMALC2X and pET30a.

## **5.6. Generation of Y2H constructs encoding truncated inter-domain.**

Using sequence alignments (Fig 3.1 and Fig 3.10) in conjunction with the modelled structure, optimum defined truncations of the inter region were suggested. Two further primers, IF2\_For and IF3\_For (Table 2.3) were subsequently designed to amplify both truncated regions in combination with the Interr1a reverse primer (Figure 5.4B). PCR amplification, restriction digest and ligations were carried out as described previously for the inter-domain.

Two sets of constructs were made in pGAD424 and pGBT9 (Fig 5.6C). The resulting constructs were sequenced and compared to the inter-domain cloned in both pGAD424 and pGBT9. pGAD:IF3 and pGBT:IF2 were found to be without deletion or base substitutions (not shown). However, in pGAD:IF2 there was a base substitution (G for an A nucleotide), resulting in alteration of the protein sequence from A to T. In pGBT:IF3 a base deletion occurred, altering the in-frame sequence, extending the inter-domain TEVK amino acid motif to TEVSKNSRGSVDLQPS. Insufficient time to re-clone the IF2/IF3 fragments remained.

## **5.7. Validation of Y2H interactions**

Two Y2H interactions of note were revealed previously in this study (I) a strong YbgC-YbgC interaction and (II) an EspC passenger- and inter-domain interaction. Importantly no YbgC interaction with any EspC domain was detected, the strong self interaction implying it forms dimers or multimers, likely precluding detectable interaction with EspC domains. YbgC was therefore excluded from further study.

In order to validate the passenger/inter-domain Y2H interaction, the relevant *espC*-encoded domains were cloned into both pET30a and pMALC2X expression vectors for the preparation of recombinant His-tagged or MBP fusion protein respectively. Purified EspC sub-domain proteins could then be utilised in column co-purification experiments for assay of protein-protein interactions.

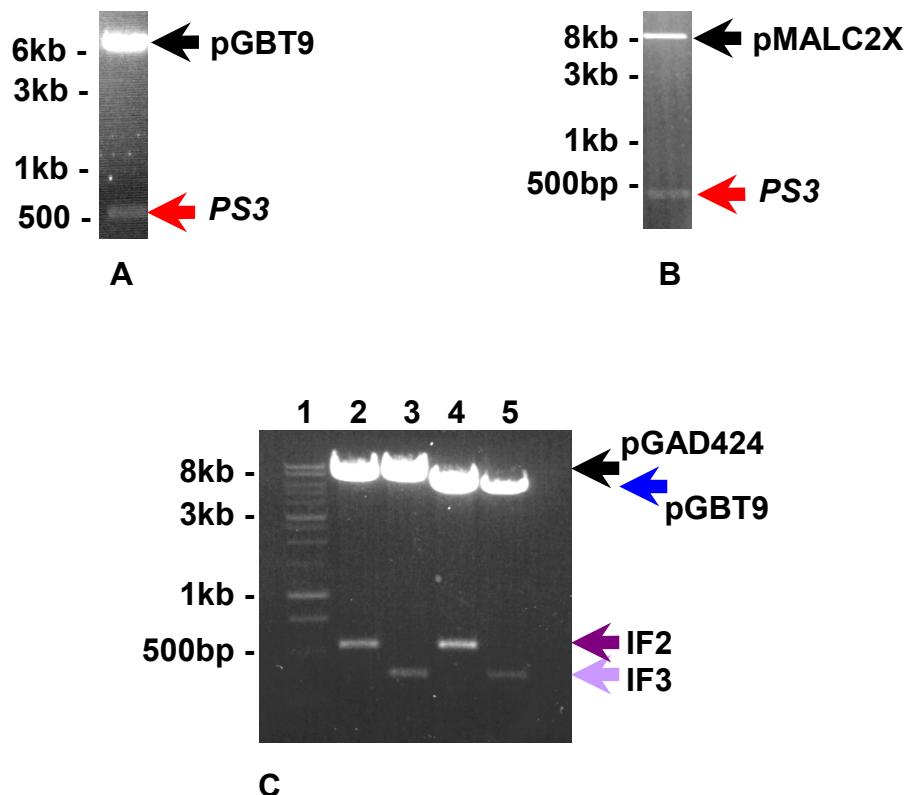


Figure 5.6: *EcoRI* digests of truncated passenger and inter-domain fragments cloned in Y2H vectors pGAD424 and pGBT9, and over-expression vector pMALC2X. EspC passenger sub-domain fragment PS3 was successfully excised from Y2H vector pGBT9 (A) and pMALC2X over-expression vector (B). Truncated inter-domain products (C) of two different lengths IF2 (approximately 520 bp) and IF3 (approximately 390 bp) were excised from both Y2H vectors pGAD424 and pGBT9 respectively. Restriction digest products were separated through a 1% agarose gel in comparison with a 1 kb ladder. Arrows indicate products of the expected size passenger sub-domain (red), IF2 (purple), IF3 (violet), pGAD424 (black) and pGBT9 (blue).

### **5.7.1 Construction of over-expression vectors encoding EspC passenger- and inter-domains**

Inserts encoding the passenger- and inter-domains were excised by *Eco*RI restriction digest from either pGBT9 or pGAD424 constructs and separately ligated into pET30a and pMALC2X. Prior to ligation, pET30a and pMALC2X were also *Eco*RI digested, and dephosphorylated. Both vector and insert fragment were gel purified and ligated in a variety of vector:insert ratios. Constructs recovered from transformants were subjected to *Eco*RI restriction digest to check for the presence of insert. Where insert was confirmed, expression of the EspC domains from (I) pMALC2X::inter (Fig 5.7A), (II) pMALC2X::passenger (Fig 5.7B), and (III) pET30a::passenger (Fig 5.7C) was validated in the appropriate strain backgrounds (BL21(DE3) for pET30a and *E. coli* XL1Blue for pMALC2X). Transformants containing pET30a::inter were also recovered, and validated by colony PCR using pET28AF and Interr1 (Fig 5.7).

### **5.7.2. Expression of recombinant EspC domains.**

Four of each pET30a constructs expressing either (I) recombinant EspC passenger- or (II) inter-domain were recovered. Passenger- and inter-domain were both clearly detectable with rabbit anti-EspC polyclonal sera (Fig 5.8A and 5.8B respectively), but only very weakly with rabbit anti-His-tag antibodies (Fig 5.8C and 5.8D respectively) perhaps as a consequence of steric masking of the His-tag, poor antibody reactivity, or cleavage of the His-tag. Attempts to purify His-tagged passenger- and inter-domains on a nickel binding column proved unsuccessful supporting the suggestion that the His-tag is either lost, or obscured/buried in the expressed folded protein. Of note is the extent of degradation which occurred in some samples (see Fig 5.9 and Fig 5.10 for example). Protease inhibitor PMSF used to inhibit degradation of the recombinant protein proved ineffective alone.

pMALC2X constructs were shown to host and produce the passenger-domain in two of six constructs by reactivity with both rabbit anti-EspC (Fig 5.9A) and rabbit anti-MBP antisera (Fig 5.9B). All pMALC2X constructs selected for testing were shown to produce inter-domain fusion protein. One construct was used for trial purification on an amylose resin column. Rabbit anti-EspC (Fig 5.9C) and rabbit anti-MBP antibodies (Fig 5.9D) were used in comparison with control plasmid pMALC2X (containing no DNA insert) to

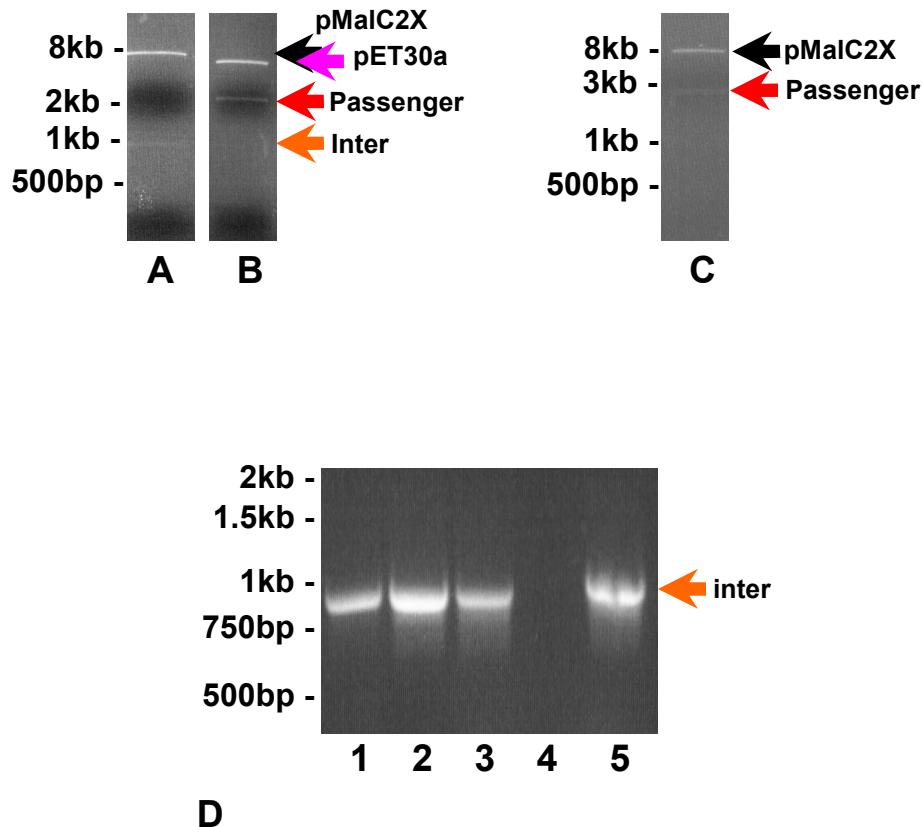


Figure 5.7: Verification of over-expression constructs. Utilising *Eco*RI restriction digestion EspC domains were excised from Y2H constructs, and then ligated separately into over-expression plasmids pET30a and pMALC2X, and constructs recovered in Novablue or XL1 Blue transformants respectively. Constructs were purified, subjected to *Eco*RI restriction digest and analysed on a 1% agarose gels. Three over-expression constructs pMALC2X::inter (A), pET30a::passenger (B), and pMALC2X::passenger (C) were found to contain EspC fragments encoding either passenger- or inter-domain. In addition, pET30a::inter constructs were verified by colony PCR (D). Novablue pET30a::Inter transformant clones 1 to 5 were analysed on a 1% agarose gel following colony PCR using pET28AF and Interr1 primers (confirming orientation). pET30a was found to encode EspC inter-domain in all clones tested except 4. Arrows indicate products of the expected size, passenger-domain (red), inter-domain (orange), pMALC2X (black), and pET30a (pink).

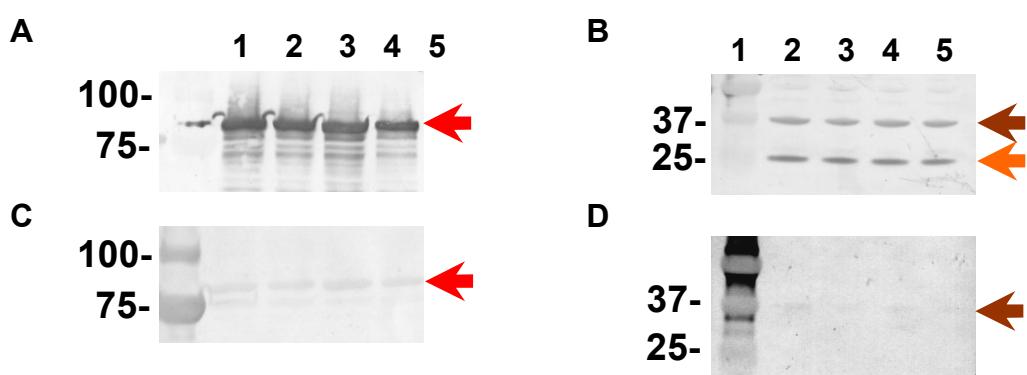


Figure 5.8: Verification of pET30a constructs encoding either EspC passenger- or inter-domain. Following induction of *E. coli* BL21(DE3) transformants hosting putative pET30a:passenger constructs, four transformants were found to over produce passenger-domain as observed following detection with rabbit anti-EspC (A) or rabbit anti-His-tag (B) antibodies. In addition, four induced *E. coli* BL21(DE3) transformants hosting putative pET30a:inter constructs were found to over produce the EspC inter-domain as observed following detection with rabbit anti-EspC (C) or Rabbit anti-His-tag (D) antibodies. However reactivity with the anti-His-tag antibody was extremely poor for both passenger- and inter-domain, but just sufficient to detect his-tagged proteins. The lower inter-domain band was not detected by the anti-His antibody and was assumed to be an N-terminal degradation product. Biorad Precision Plus Markers were used to estimate size. Arrows indicate products of the expected size, his-tagged passenger-domain (red), his-tagged inter-domain (brown), inter-domain (orange).

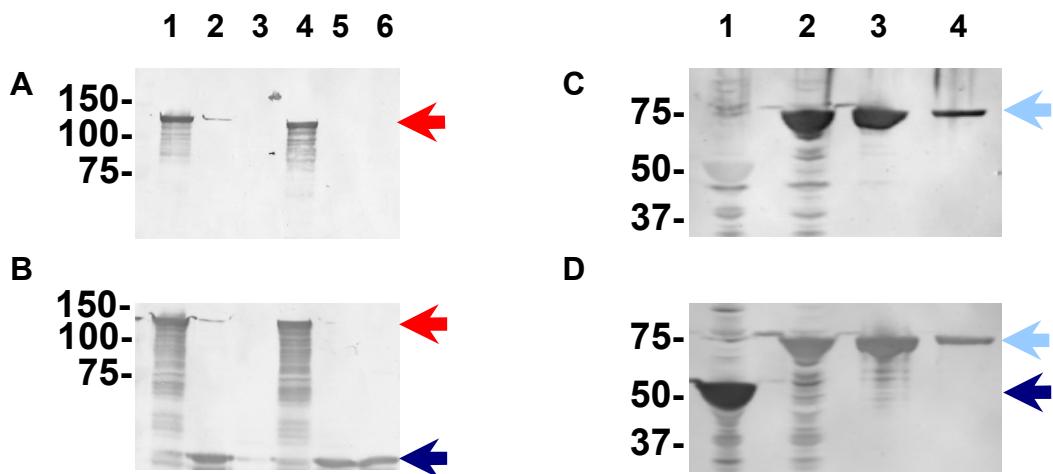


Figure 5.9: Verification of pMALC2X constructs encoding either EspC passenger- or inter-domain. Following induction of *E. coli* XL1Blue transformants hosting putative pMALC2X:passenger constructs, two of six selected transformants were found to over produce the EspC passenger-domain as observed following detection with rabbit anti-EspC (A1 and A4) and rabbit anti-MBP (B1 and B4) antibodies. Following induction of *E. coli* XL1Blue transformants hosting putative pMALC2X:inter constructs, four of six selected transformants were revealed to over-produce the EspC inter-domain. One was selected to conduct a trial purification of recombinant fusion protein MBP-inter as observed following detection with rabbit anti-EspC (C) and rabbit anti-MBP (D) antibodies. A control lysate XL1 Blue pMALC2X (C1 and D1) was used for comparison with unpurified lysate of XL1 Blue pMALC2X::inter (C2 and D2) and two eluted fractions (C3, C4 and D3,D4). MBP-inter domain was successfully purified. Arrows indicate products of the expected size, MBP-tagged passenger-domain (red), MBP-tagged inter-domain (pale blue), and MBP fusion protein (dark blue).

demonstrate that the inter-domain-MBP fusion can be successfully purified using a standard protocol.

### 5.7.3. Co-purification studies

Co-purification experiments were conducted by immobilising EspC-MBP fusions on amylose resin, then applying lysates of pET30a-espC/BL21(DE3) induced cultures in an attempt to capture interacting proteins. Due to an inability to purify the His-tagged domains, it was not possible to immobilise His-tag proteins for use as the capture protein.

Two main experiments were performed (I) immobilisation of the MBP-EspC passenger-domain fusion with subsequent attempts to capture His-EspC inter-domain contained in clarified lysates, and (II) immobilisation of MBP-EspC inter-domain fusion protein with similar attempts to capture His-EspC passenger-domain contained in clarified lysates. Control experiments, whereby un-fused MBP was used as the capture protein were performed in parallel.

Fractions were separated by SDS-PAGE and eluted proteins detected using anti-EspC and anti-MBP antibodies. Using MBP-EspC inter-domain as capture protein, no co-purification of His-tagged passenger-domain was evident (Fig 5.10), the control experiment was similarly negative as expected (Fig 5.10). However, when MBP-EspC passenger-domain fusion was utilised as the capture protein, two proteins corresponding to cleaved and uncleaved His-tagged EspC inter-domain co-eluted (Fig 5.11A). Corresponding proteins were observed in pET30a-inter/BL21(DE3) (Fig 5.8), suggesting successful co-purification.

Neither protein was evident when control lysate (induced cultures of pET30a/BL21(DE3) was passed through columns containing bound MBP-EspC passenger-domain fusion (Fig 5.11B), demonstrating specific capture of the His-inter-domain by the MBP-passenger-domain fusion. The interaction though weak, confirms the passenger-inter-domain interaction and validates the Y2H observations.

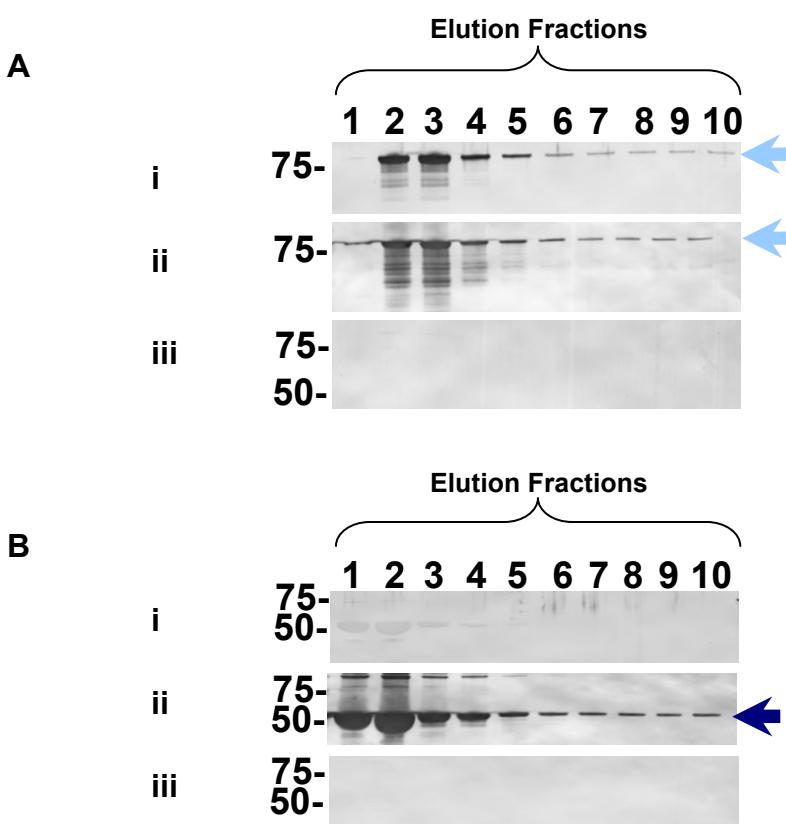


Figure 5.10: Assessment of passenger- and inter-domain interaction by co-purification. Amylose resin containing immobilised either (A) MBP-inter-domain fusion or (B) MBP (control), was used to affinity purify His-tagged EspC passenger-domain from BL21DE3 pET30a:passenger lysate passed through the same amylose column. Excess lysate was washed through and 10 mM maltose in column buffer used to elute bound proteins. Following SDS-PAGE and Western transfer, blots were probed with rabbit anti-EspC (i), rabbit anti-MBP (ii) and Rabbit anti-His-tag (iii). Lanes 1 to 10 indicate the elution fractions. The Western blots (Ai and ii) demonstrate that MBP-tagged inter-domain has been eluted and purified successfully, showing evidence of degradation of the protein; however no co-purified His-tagged EspC passenger-domain was detectable (Ai and iii). In the control experiment (B), lanes 1 to 10 indicate the elution fractions. MBP protein was eluted and purified successfully (Bii), however no co-purified passenger-domain was detectable using either rabbit anti-EspC or rabbit anti-His-tag antisera (Bi and iii). Arrows indicate products of the expected size, MBP-tagged inter-domain (pale blue), and MBP fusion protein (dark blue).

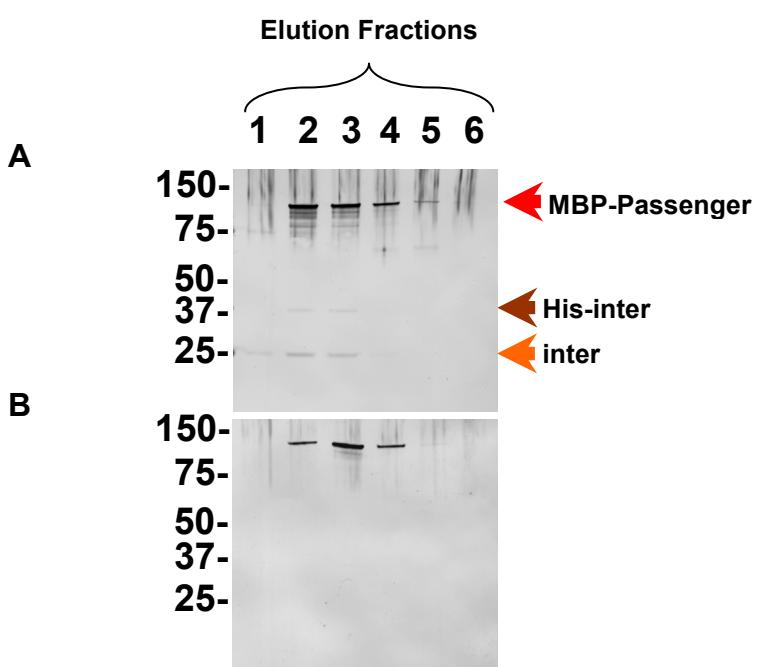


Figure 5.11: Assessment of EspC passenger- and inter-domain interaction by column co-purification. Amylose resin containing bound MBP-passenger-domain fusion was used to affinity purify His-tagged inter-domain from BL21DE3 pET30a:*inter* lysate (A). This was contrasted with control lysate BL21DE3 pET30a (B) passed through amylose resin also containing bound MBP-passenger-domain. Both Western blots were probed with rabbit anti-EspC antibodies. Bands corresponding to the His-inter-domain and a possible degradation product can be seen to co-purify with the MBP-passenger-domain fusion in elution fractions 2 and 3 (A). Arrows indicate products of the expected size, MBP-tagged passenger-domain (red), his-tagged inter-domain (brown), and inter-domain (orange).

## 5.8 Discussion

The Y2H interaction screens provide evidence for homo-dimerization of the passenger-domain, interaction of the passenger-domain with the chaperone-like inter-domain, and finally a self interaction between YbgC protein molecules, though no evidence of a YbgC-EspC domain interaction. Additional independent techniques undertaken to validate the Y2H observations (excluding YbgC) involved the use of affinity co-purification. Although weak, and only demonstrated when MBP-passenger-domain fusion was used as the capture protein, a passenger-inter-domain interaction was demonstrated with recombinant protein.

[Oliver et al, \(2003a\)](#), demonstrated that the region between the passenger- and  $\beta$ -barrel domain (spanned by the inter-domain) of the *B. pertussis* AT BrkA acts as a chaperone in the export of the mature passenger-domain. Utilizing a deletion-based strategy, a series of in-frame *brkA* deletion mutants were constructed by PCR mutagenesis, and several mutants were found not to secrete BrkA to the bacterial surface. It was concluded that (I) BrkA did not possess any auto-proteolytic activity, (II) the  $\beta$ -domain in itself was insufficient to aid translocation to the cell surface, and (III) the  $\beta$ -core and a preceding linker region form the minimum translocation unit necessary to export BrkA. Our observations support the work of [Oliver et al](#), as they indicate there is an interaction between the inter- and passenger-domain of EspC. In a subsequent study [Oliver et al \(2003b\)](#) identified that the conserved domain is located at the C-terminus of the BrkA passenger-region, which was further defined and determined to confer stability to the BrkA passenger-domain. In our ClustalW alignment, this region maps towards the C-terminus of the region we have defined as the inter-domain (see Fig 3.1, and also Fig 3.10). Deletion of the conserved region rendered the protein susceptible to degradation by OMP proteases OmpP, OmpT and by Trypsin. [Oliver et al, \(2003b\)](#) analyzed the BrkA fusions generated by the use of far-UV circular dichroism (far-UV CD) spectroscopy and proteolytic analysis. BrkA fusions containing deletion mutations (of residues E<sup>601</sup> to A<sup>692</sup>) had a far-UV CD profile, in contrast to that which did not contain any mutation, revealing that non-mutated protein was likely rich in  $\beta$ -structure, implying protein was (I) folded, and (II) had a higher tolerance to proteolytic enzymes. Other important residues and protein segments were also reported which appear to encode information that is necessary to initiate or trigger folding of the BrkA passenger-domain. Complementation mutants designed to be surface exposed rescued the instability of

mutants lacking the putative chaperone deleted region. The intramolecular chaperone region identified in the PrtS AT of *Serratia marcescens* (Ohnishi *et al*, 1994) is conserved in a number of ATs with serine protease activity. Oliver *et al*, (2003b) suggest that the region may vary depending on the particular AT. Complementation experiments showed that oligomeric AT interactions were also possible, as multiple BrkA ATs appeared to interact, though how many were involved in the interaction remained unknown.

Passenger-passenger interactions have previously been demonstrated in an *E. coli* sub-group known as self-associating autotransporters or SAATs, loosely defined on the basis of (I) their amino acid homology (translocator domain) and (II) presence of repetitive amino acid motifs (passenger domain) (Henderson and Nataro, 2001; Klemm *et al*, 2006) they include (I) TibA, an adhesin/invasin associated with ETEC, (II) AIDA, an adhesin from diarrhea-causing *E. coli* and (III) Ag43 autoaggregation factor found in the majority of *E. coli* strains. All three promote bacterial aggregation and enhance biofilm formation (Klemm *et al*, 2006). However EspC belongs to another sub-group of ATs known as SPATEs (Henderson *et al*, 2004); in addition to which the self association is weak and as yet unconfirmed by the use of alternate biochemical means (co-purification for example). The self-association implied by the Y2H interaction however could be indicative of versatility in the function of EspC which could explain the activity demonstrated as both an enterotoxin (Mellies *et al*, 2001) and Haem-binding protein/protease (Drago-Serrano *et al*, 2006).

No interaction between YbgC with any of the EspC domains was observed in Y2H assays, though strong evidence of YbgC dimerization was obtained. Despite utilization in bacterial two-hybrid experiments, no YbgC self-interaction was been previously demonstrated (Gully *et al*, 2006). YbgC was instead shown by Gully *et al*, (2006) to interact with a co-factor central to fatty acid and phospholipid synthesis, acyl carrier protein (ACP); suggesting ACP may be the physiological substrate of YbgC. Gully *et al*, implied YbgC may be one of several proteins involved in phospholipid synthesis regulation, acting on acyl-ACP under certain conditions. Experimental evidence based on the crystallographic structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthobacter* sp. (Thoden *et al*, 2003), of which YbgC is a homologue, suggests that YbgC is tetrameric and is composed of a dimer of dimers, with all subunits adopting a 'hot dog fold'. Similar structures have been observed in the thioesterase proteins from

*Pseudomonas* spp (Benning *et al*, 1998). Recently published crystallography work on the structure of a hydrolase protein (Accession: 15SU) with high homology to YbgC (Accession: NP\_308798) also clearly showed the likely multimeric nature of this molecule. The lack of observed interactions with EspC domains in the Y2H system could be explained by YbgC dimerization or multimerisation. If YbgC dimerization is a prerequisite for interaction with EspC then the appropriate complex is unlikely to be competent for Gal4 activation since restoration of Gal4 activity relies upon a pair-wise interaction. Alternatively, YbgC homo-dimers may form preferentially and preclude interaction of monomeric YbgC with EspC. One possible additional explanation for the lack of an EspC-YbgC interaction could be that YbgC interacts indirectly with EspC possibly combining with an as yet uncharacterized factor which then interacts with EspC.

Membrane proteins, because of their hydrophobic nature may localise within the nuclear membrane or aggregate, the  $\beta$ -domain was therefore anticipated to be non-functional in the Y2H system due to insolubility or incorrect localisation, this appears to be the case for the  $\beta$ -domain. An alternative Y2H approach to identify potential  $\beta$ -domain interactions is likely required, such as the split-ubiquitin system (Johnsson and Vashavsky, 1994) which has been modified specifically for use with membrane proteins (Stagljar *et al*, 1998).

## **6. Targeted mutagenesis of the EspC inter-domain and secretion analysis**

### **6.1 Introduction**

A weak/intermediate interaction between the passenger- and inter-domains of EspC was revealed by Y2H analyses, with the co-purification of recombinant proteins providing additional support. To further define the nature of this interaction, the modelling of the truncated EspC inter-domain based on the crystal structure of the Hbp AT was further exploited to identify potentially important residues for mutagenesis. SepA, AidA, EatA, and EspC, previously identified in phylogenetic studies as a closely related subgroup of the AT family, were subjected to a sequential series of alignments to find a sequence region of sufficient homology with the extreme C-terminal passenger region of haemoglobin protease (designated here as Hbp2). This facilitated subsequent modelling of the EspC inter-domain, and permitted identification of strictly conserved residues, either within the EspC subgroup of ATs, between the subgroup and Hbp2 or between Hbp2 and EspC. The positions of conserved residues or sequence motifs, which by implication are suggested to be important for structure or function, were subsequently mapped onto the modelled EspC structure. The model was then scrutinised to identify sequence features comprising the conserved sequence, such as surface exposed loops, which were considered to be of potential structural/functional significance.

### **6.2 Identification of mutagenesis targets on EspC**

The PDB file created by SwissModel, comprising modelled EspC coordinates based on the haemoglobin protease AT crystal structure (Accession: 1WXR), was visualised in Rasmol allowing conserved motifs in the mapped EspC inter-domain to be highlighted (Fig 3.10). Following scrutiny of the structure with reference to the multiple sequence alignments, several conserved residues and sequence motifs were identified. All putative target motifs are highlighted in the sequence alignment with Hbp2 (Fig 6.1). Of these, a conserved GFS (Gly, Phe, Ser) sequence motif, comprising  $G^{976}$ ,  $F^{977}$  and  $S^{978}$  was shown to map to an exposed loop in the modelled structure and was subsequently selected as a target for site directed mutagenesis (Fig 6.2). In addition, a similarly conserved GYQ (Gly, Tyr, Gln) sequence motif comprising  $G^{999}$ ,  $Y^{1000}$ , and  $Q^{1001}$  also mapped to an exposed loop presenting another attractive mutagenesis target (Fig 6.1). Other conserved potential targets include a TPW (Thr, Pro, Trp) motif comprising  $T^{981}$ ,

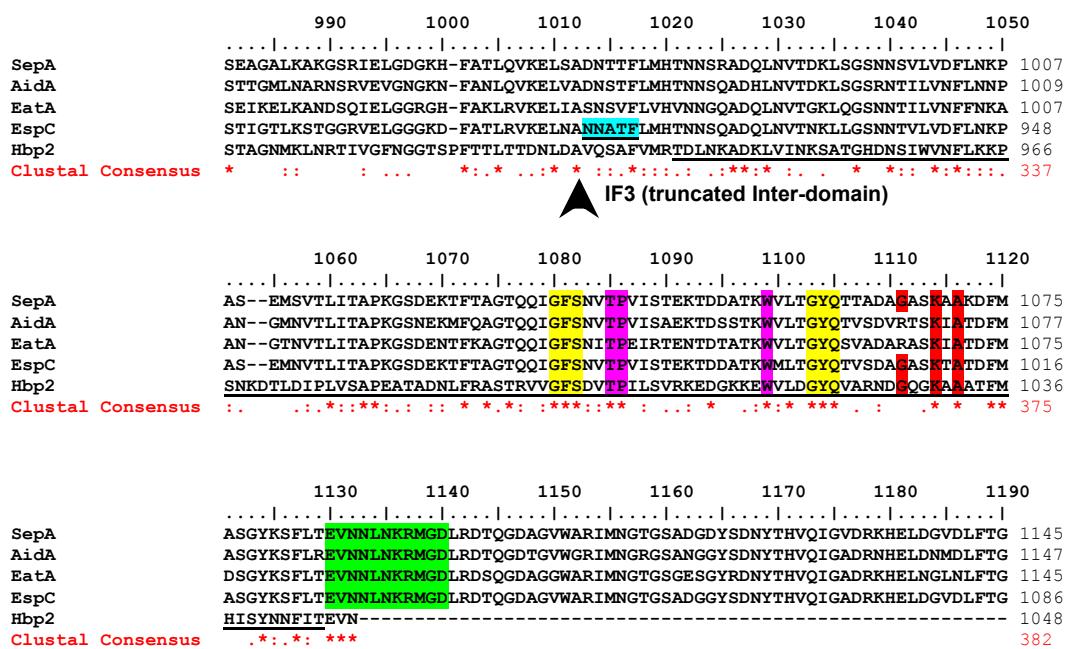


Figure 6.1: Multiple sequence alignment of the EspC subfamily of ATs with Hbp2 identifying conserved mutagenesis targets in the EspC inter-domain. Regions highlighted indicate potential targets for mutagenesis. Underlined and highlighted in cyan indicates the start of the sequence region selected to construct the inter-domain truncated product IF3. In yellow, are residues that have been found by EspC modelling to map onto the 3D structure of Hbp2 in exposed loop regions. G<sup>976</sup>, F<sup>977</sup>, S<sup>978</sup> and G<sup>999</sup>, Y<sup>1000</sup>, Q<sup>1001</sup> motifs are both found to be located in the proposed putative chaperone region (underlined in Hbp2) (Oliver *et al*, 2003). Other possible targets located in vicinity of loop regions (T<sup>981</sup>, P<sup>982</sup>, and W<sup>995</sup>) are highlighted in pink. Highlighted in red are non-adjacent residues (K<sup>1010</sup> and A<sup>1012</sup>) lying on an exposed loop. Finally, the sequence highlighted in green marks the end of the inter-domain fragment used in the protein-protein interaction studies.

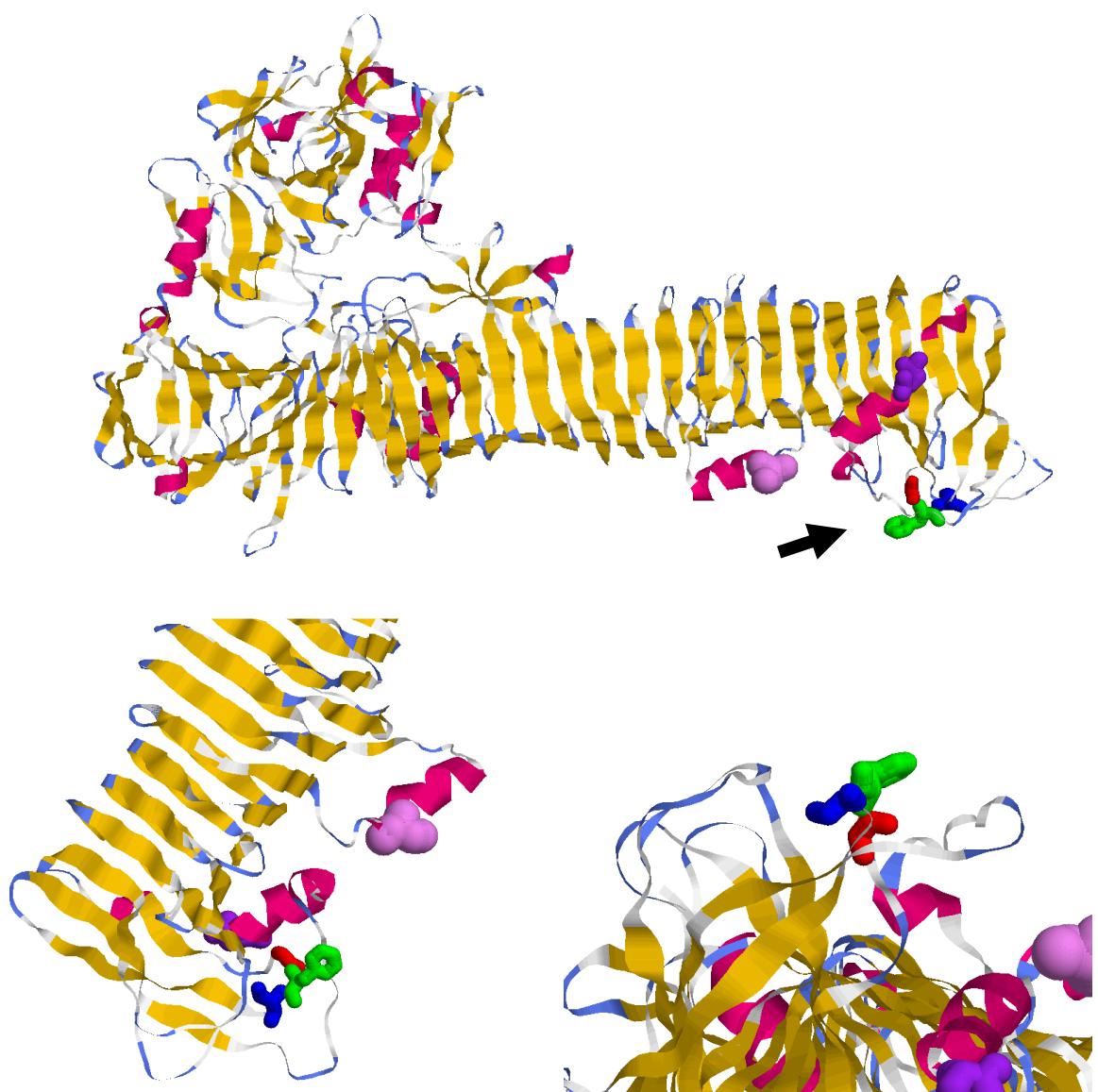


Figure 6.2: Location of the conserved GFS ( $G^{976}$ ,  $F^{977}$ ,  $S^{978}$ ) motif on the EspC inter-domain model structure. The structure onto which EspC has been mapped has been highlighted for  $\alpha$ -helices (pink),  $\beta$ -sheets (yellow), turns (pale blue) and other residues (white). The N- and C-terminus of the mapped EspC are highlighted  $S^{847}$  (violet) and  $D^{1028}$  (purple) respectively. The GFS motif appears to contribute to a surface loop. Glycine is shown in red, phenylalanine in blue, and serine in green. The motif is indicated on the full structure by an arrow.

P<sup>982</sup> and W<sup>995</sup> (Fig 6.1) in the vicinity of exposed loop regions, and finally individual conserved residues, K<sup>1010</sup> and A<sup>1012</sup> (Lys and Ala respectively) which lie non-adjacently on an exposed loop (not shown).

### **6.3. Rationale for detection of phenotypic defects following mutagenesis of conserved residues**

The conserved GFS motif lies within the region of the inter-domain proposed by Oliver *et al*, (2003), to be a putative chaperone. As the hydrophobic Phe (F) residue in this motif is unusually surface exposed, it was hypothesised that the exposed GFS loop may have functional or structural importance. Consequently, mutation of residues in this motif might be expected to affect function of the putative EspC inter-domain. If this region of EspC indeed functions to chaperone passenger-domain folding to facilitate export, then mutation of structurally/functionally important residues might be expected to show secretion defects in a mutant background. Such secretion defects can be assessed by comparison of secretion profiles obtained from WT and *espC* mutant derivatives created by complementation of the *espC* mutant strain with modified pLAC-*espC* plasmids.

#### **6.3.1. Site-directed mutagenesis and sequence analysis of mutated *espC* encoded by pLAC1 and pLAC2 transformants.**

With reference to *E. coli* codon usage tables, the primers (I) PheMut\_For and PheMut\_Rev and (II) 3AlaMut\_For and 3AlaMut\_Rev (Table 2.3) were designed to incorporate the least number of DNA nucleotide base changes necessary. The first pair of primers were utilised to generate an F<sup>977</sup>A point mutant and the second pair to replace the entire EspC GFS motif with alanine at each position (producing A<sup>976</sup> A<sup>977</sup> A<sup>978</sup>) in the plasmids pLAC1 and pLAC2, which respectively express EspC either constitutively or by induction with arabinose. The Stratagene Quikchange II XL kit was used successfully to generate specific mutations in the WT *espC* gene sequence. In pLAC1 a single amino acid alteration was successfully achieved (Fig 6.3), which was not used further in this study (pPheMut1A). In pLAC2 (inducible *espC* construct), two single amino acid alterations were successfully introduced. The first amino acid alteration F<sup>977</sup>A, was as expected (pPheMut1). Unexpectedly however, a second single amino acid substitution was also generated in construct pPheMut2. Sequencing of the *espC* fragment in this mutated construct, revealed a F<sup>977</sup>G substitution. A triple amino acid alteration was also

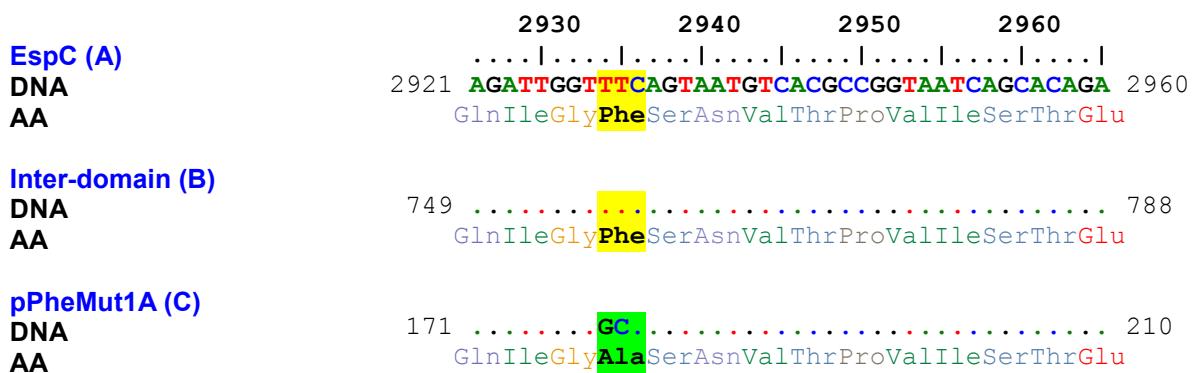


Figure 6.3: Bioedit alignment of DNA and corresponding amino acid (AA) sequence showing successful site-directed mutagenesis of arabinose independent construct pLAC1 (pBAD18::espC) using the Stratagene Quikchange II XL Site-directed mutagenesis kit. Only one successful mutant, Phe977Ala (C) was obtained through site directed mutagenesis, the DNA substitutions and corresponding AA change have been highlighted in green as opposed the unchanged sequence highlighted in yellow. The arabinose-independent construct that was successfully mutated was not used for further studies.

successfully introduced by mutation of the pLAC2 encoded *espC* (p3AlaMut). The final mutated plasmids were designated pPheMut1A ( $F^{977}A$ ), pPheMut1 ( $F^{977}A$ ), pPheMut2 ( $F^{977}G$ ) and finally p3alaMut ( $G^{976} F^{977} S^{978}$  mutated to AAA).

### 6.3.2. Secretion analysis in WT and *espC* mutant EPEC strains transformed with mutated pLAC2 plasmids

Following successful site directed mutagenesis, the mutated pLAC2 constructs, were transformed into electro-competent cells of either WT EPEC (strain E2348/69) or its isogenic *espC* mutant. Intriguingly, Western blots probed with anti-EspC antisera showed that WT EPEC carrying various mutant pLAC2::*espC* constructs (comprising single and triple substitutions) did not accumulate a cell associated EspC pool after overnight incubation (Fig 6.4A, Lanes 5-7). In contrast WT EPEC hosting control plasmid pBAD18 or pLAC2 (Fig 6.4A, Lanes 3 and 4) encoding WT *espC* did accumulate cell associated EspC. However, wild-type EPEC harbouring pPheMut1, pPheMut2, or p3AlaMut grew poorly in media containing 0.2% glucose, apparently undergoing lysis after extended incubation (Fig 6.5). This phenomenon was reflected in the reduced accumulation of EspC observed in supernatants from overnight cultures and likely explains the absence of EspC in lysates (it should also be noted that while equivalent pellets were not obtained, all samples were adjusted by OD<sub>600nm</sub> to enable comparison).

Cell lysis of these strains was apparently initiated by the presence of glucose (present in the medium to repress induction of pLAC2) it may be that (I) lysis was osmotically induced as a consequence of glucose diffusion into the cells and (II) the co-expression of mutant and non-mutant EspC in EPEC is detrimental, enhancing susceptibility to osmotic shock. Controls indicated that cell lysis correlates with co-expression of mutant and non-mutant EspC. Although glucose should tightly suppress induction of the mutant EspC from the pBAD promoter of pLAC2, this was clearly not the case since a low level of EspC secretion occurred when the EPECΔ*espC* mutant carrying pLAC2 derivatives was grown in glucose (Fig 6.7). Cell associated accumulation of EspC was observed in the EPECΔ*espC* mutant complemented with both pLAC2 and its mutant derivatives (Fig 6.4A, Lanes 10-12). Secreted EspC was also observed in the supernatants of wild-type EPEC and the EPECΔ*espC* mutant complemented with pLAC2 and mutant derivatives, although in the latter case (EPECΔ*espC*-p3AlaMut, and EPECΔ*espC*-pPheMut1), EspC secretion appeared reduced (Fig 6.4B, lanes 11 & 12). EspC secretion and

accumulation in supernatants was further assessed in a series of Western immunoblots comparing samples taken at timed (30 min) intervals. These experiments identified that production of detectable quantities of EspC occurred between 0 to 30 min after induction (Fig 6.6). Secretion analyses were further refined, using 10 min intervals between 0 and 30 min, and also at 60 min. As expected, EPEC $\Delta$ espC pBAD18 did not contain cell associated or secreted EspC (Fig 6.7). WT EPEC pBAD18 contained detectable cell associated EspC only after 20 min (Fig 6.7) and in the supernatant from as early as 0 min. Similarly, in EPEC $\Delta$ espC (pLAC2), cell associated EspC could not be detected before 20 min, mirroring detection at the same time point in the supernatant (Fig 6.7). This delay in the appearance of extracellular EspC in the EPEC $\Delta$ espC compared to wild-type is probably linked to the need for the plasmid encoded *espC* to be induced by arabinose.

A similar pattern of EspC production and secretion was demonstrated by EPEC $\Delta$ espC pPheMut2 (Fig 6.8). In contrast, although the EspC secretion profile of EPEC $\Delta$ espC-p3AlaMut was similar to EPEC $\Delta$ espC-pLAC2, including the time it could first be detected inside the cells (Fig 6.7), the EspC AAA triple mutant appeared to be less stable as cell associated EspC was no longer detectable 60 min after induction (Fig 6.7). The EspC<sup>PheMut1</sup> also exhibited a non-WT profile as it could be detected as early as 10 min inside the cells (Fig 6.7) and even earlier (at 0 min) in the supernatant. Like the EspC<sup>3AlaMut</sup>, the EspC<sup>PheMut1</sup> was no longer detectable 60 min after induction in cell lysates (Fig 6.7). Sequencing confirmed the integrity of the mutated *espC* in pLAC2 and the restoration of WT phenotype by complementation of the *espC* mutant EPEC strain with the parent pLAC2 plasmid confirms the integrity of the gene deletion in this strain. Consequently, the difference in secretion profiles between strains complemented with pLAC2 encoding either WT EspC or the F to A EspC mutant can be directly attributed to the mutated GFS (G<sup>976</sup> F<sup>977</sup> S<sup>978</sup>) motif. Mutation of this motif clearly affects both EspC secretion (and post-secretion stability) as assessed by complementation of the *espC* mutant strain, and also profoundly affects the viability of the bacterium itself when both mutant and WT versions of EspC are present in the same cell, as shown by introduction of mutated pLAC2 into WT EPEC.

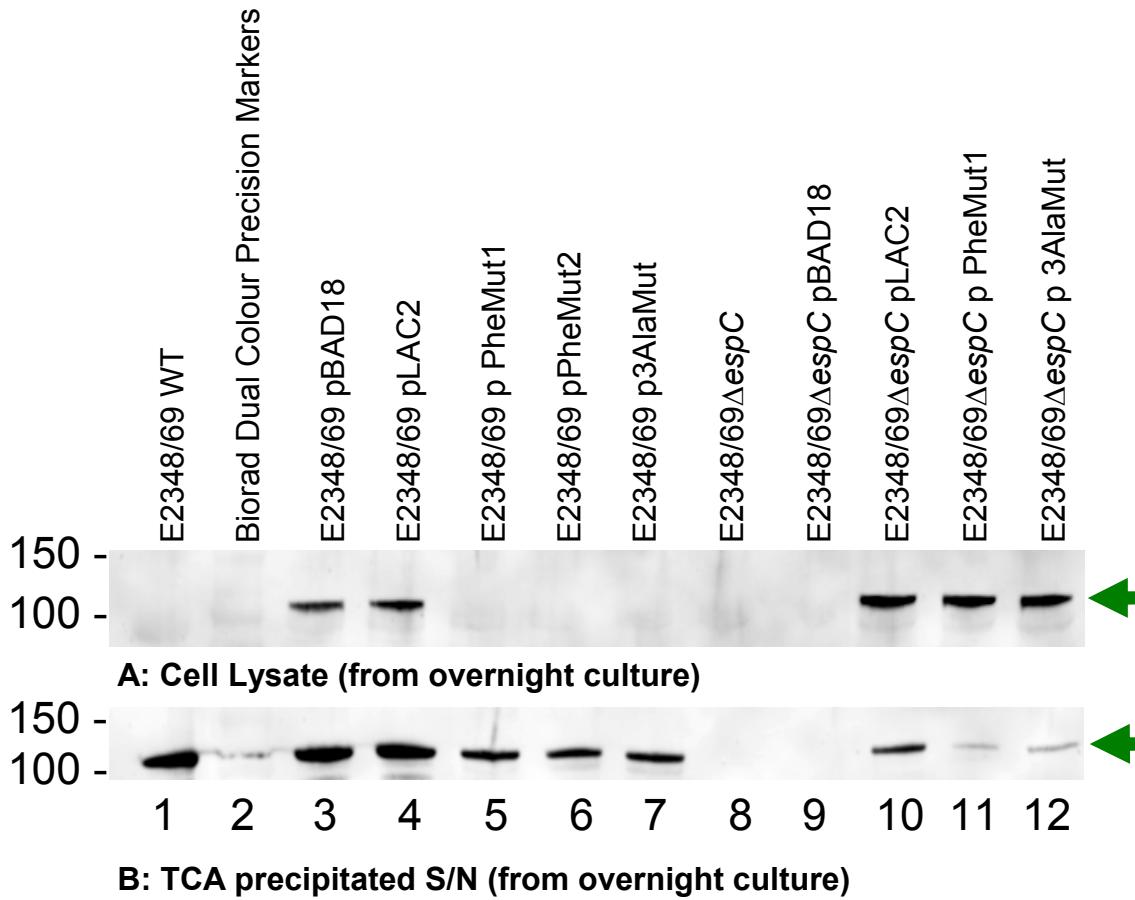


Figure 6.4: Analysis of EspC secretion in WT EPEC and its *espC* deficient mutant hosting mutated and non-mutated *espC* expression construct pLAC2. EspC production was induced and over-expressed in WT EPEC and its *espC* deficient mutant, complemented with arabinose inducible pLAC2:*espC* constructs which had been subjected to site directed mutagenesis and compared against controls pLAC2 (non mutated EspC) and pBAD18 (containing no *espC* insert). Whole cell lysates adjusted for cell density were sonicated, separated by SDS-PAGE, then Western blotted. Supernatants were TCA precipitated, separated by SDS-PAGE gel and Western blotted. No accumulation of EspC is observed in EPEC WT hosting mutated pLAC2 (Lanes A5, A6, A7) in contrast to WT hosting pBAD18 (A3) and pLAC2 (A4). EspC secretion is still observed in the supernatants of WT hosting mutated EspC (B5, B6, B7) however the level of EspC production appears reduced when compared to pLAC2 (B4) or pBAD18 (B3). The *espC* deficient mutant hosting pLAC2 (A10), and two of the mutated variants (A11 and A12) do accumulate EspC in whole cells in contrast to the WT. pLAC2 successfully complements the *espC* deficient mutant demonstrating EspC secretion into the supernatant (B10), though the production of EspC revealed by rabbit anti-EspC polyclonal indicates a reduced level of secretion of the mutated variants in supernatant (B11, and B12) taken from overnight cultures. EspC expression indicated by green arrow.

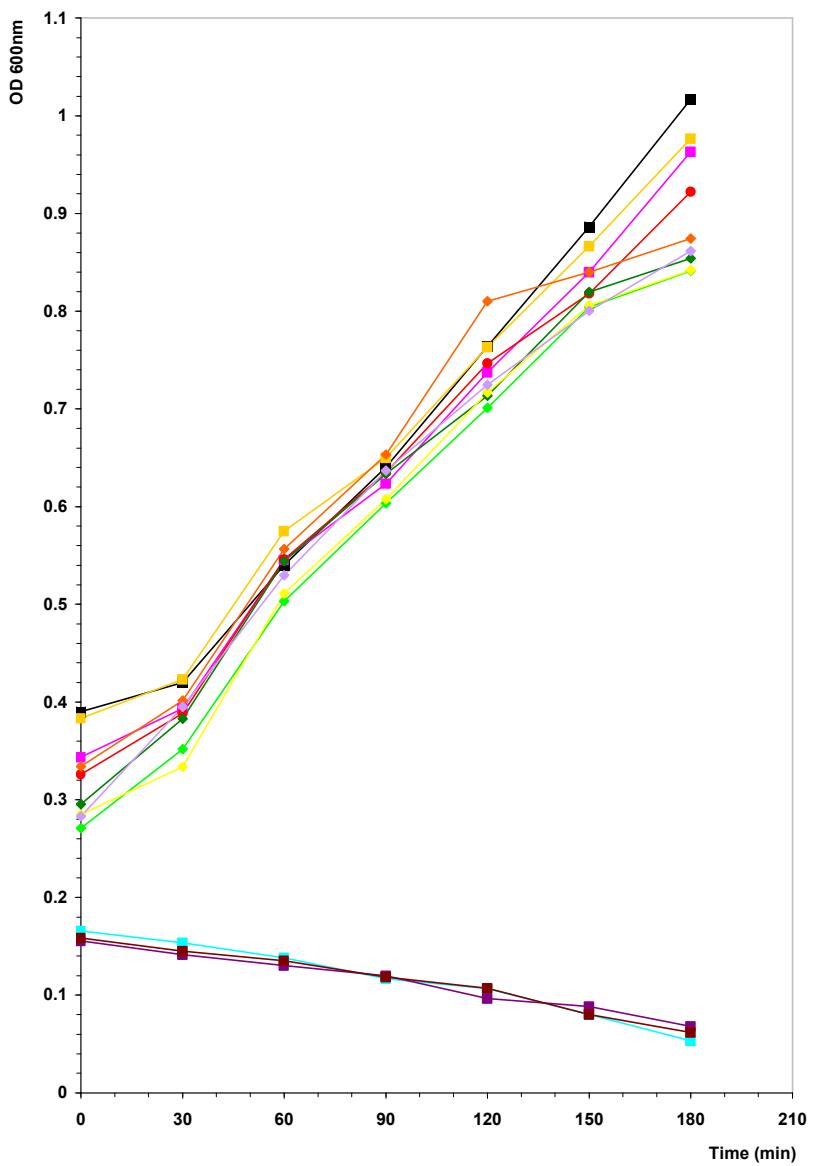


Figure 6.5: Analysis of growth in WT EPEC and its *espC* deficient mutant hosting mutated and non-mutated *espC* expression construct pLAC2 following arabinose induction. WT EPEC, strain E2348/69, has been abbreviated to E69. E69 WT (—■—), E69 pBAD18 (—■—), E69 pLAC2 (—□—), E69 pPheMut1 (—△—), E69 pPheMut2 (—▲—), E69 p3AlaMut (—■—), E69ΔespC (—●—), E69ΔespC pBAD18 (—◆—), E69ΔespC pLAC2 (—◆—), E69ΔespC pPheMut #1 (—◆—), E69ΔespC pPheMut2 (—◆—), and E69ΔespC p3AlaMut (—◆—). Growth of WT EPEC, harboring mutated pLAC2 vector (pPheMut1, 2, or 3AlaMut) was already compromised compared to that of WT EPEC alone or transformed WT harboring either pBAD18 or pLAC2 in L-Broth supplemented with glucose. Following arabinose induction (Time = 0 min) of pBAD18, pLAC2 and mutant pLAC2, expression of mutated and non-mutated EspC in the WT strain continued to prove detrimental to cell viability.

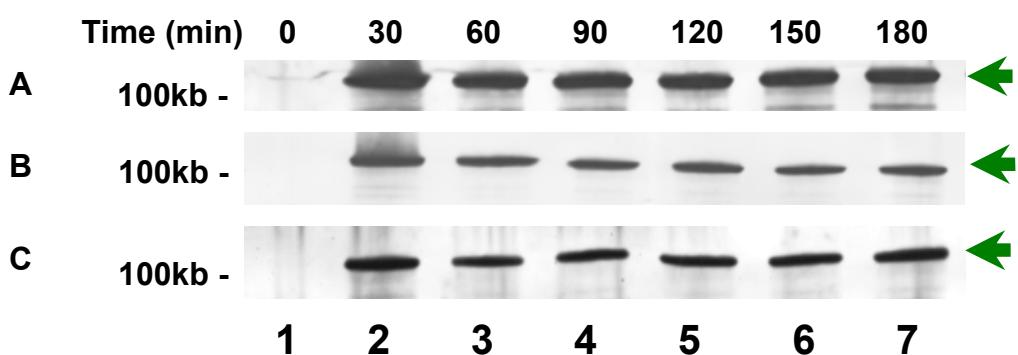


Figure 6.6: Western immunoblot analysis of TCA precipitated secreted proteins from the supernatant cultures of the EPEC *espC* deficient mutant complemented with mutated pLAC2 constructs over 3 h. Following arabinose induction of E2348/69Δ*espC* pPheMut1 (A), E2348/69Δ*espC* pPheMut2 (B), and E2348/69Δ*espC* p3AlaMut (C), supernatant samples were taken at 30 min intervals over a 3 h time course. Supernatants were TCA precipitated, separated by SDS-PAGE gel and Western blotted. EspC was detected using rabbit anti-EspC (indicated by green arrow). EspC production was observed in all samples within the first 30 min following induction and significant cell associated levels were maintained beyond 3 h post-induction (Lane 2 to 7).

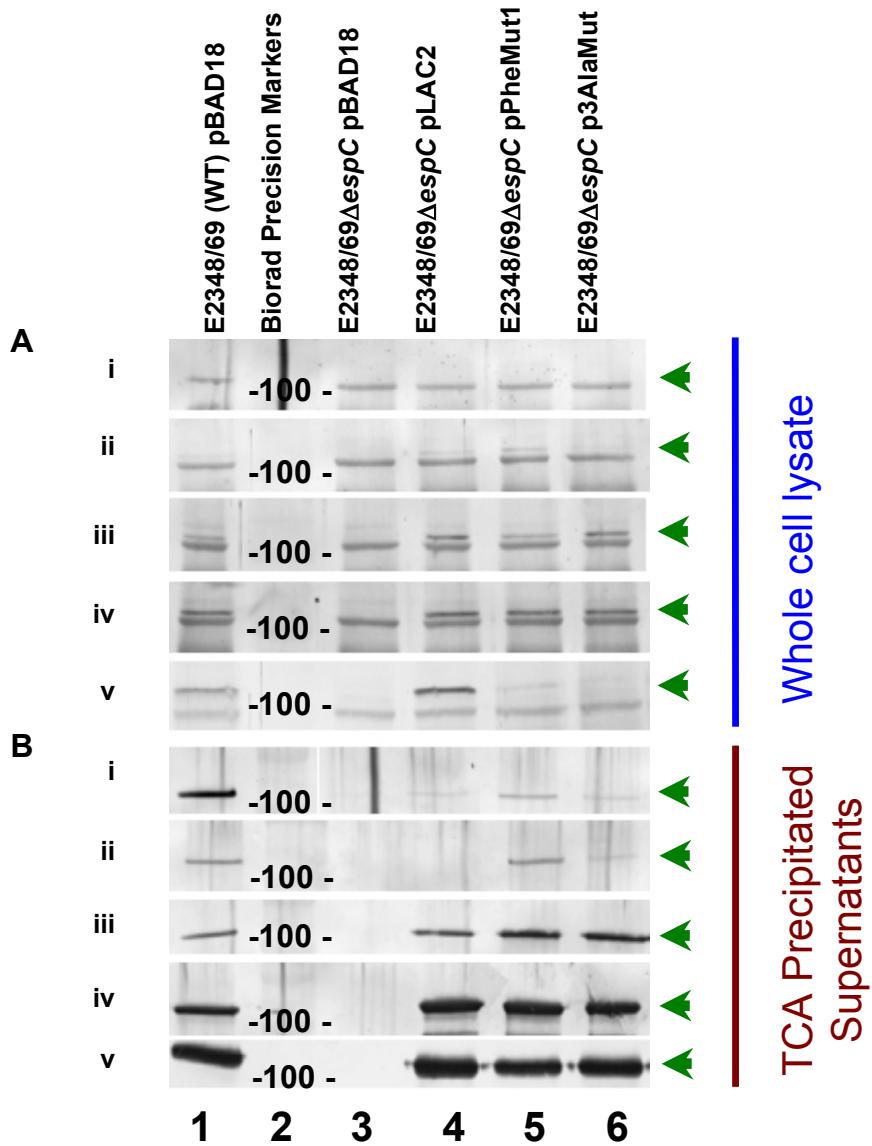


Figure 6.7: Western immunoblot analysis of EPEC *espC* deficient mutant complemented with mutated pLAC2 constructs. Bacterial cultures were initially sampled prior to arabinose induction at (i) 0 min (uninduced), then at (ii) 10 min, (iii) 20 min, (iv) 30 min and finally (v) 60 min after induction. Whole cell lysate (A) and TCA precipitated supernatant (B) samples were separated by SDS-PAGE gel and Western blotted. EspC was localised using rabbit anti-EspC sera (indicated above by green arrow). In whole cell lysates EspC production was observed in the complemented EPEC *espC* deficient mutant hosting pLAC2 (A4), pPheMut1 (A5), and p3AlaMut (A6) after 10 min (A4 and A5), and 20 min (A6). EspC production was also observed in WT EPEC hosting control plasmid pBAD18 (A1) after 10 min. Levels of accumulated EspC decreased in the whole cells between 30 min to 1h in EPEC *espC* deficient mutant hosting mutated pLAC2 (Lane E5 and E6 respectively). In TCA precipitated samples EspC production was detected in both WT EPEC hosting pBAD18 (B1) and *espC* deficient mutant complemented with pLAC2 (non-mutated and mutated) (B4, 5, and 6) after 20 min. Glucose utilised to tightly suppress induction of EspC still permitted some low level EspC secretion to occur without induction (Bi). Following inoculation of media containing 0.02% arabinose, *espC* deficient mutant complemented with mutated pLAC2, production of EspC was observed at 10 min (B5 and B6 respectively). EspC secretion was only detectable in the *espC* deficient strain hosting non-mutated pLAC2 at 20 min (B4).

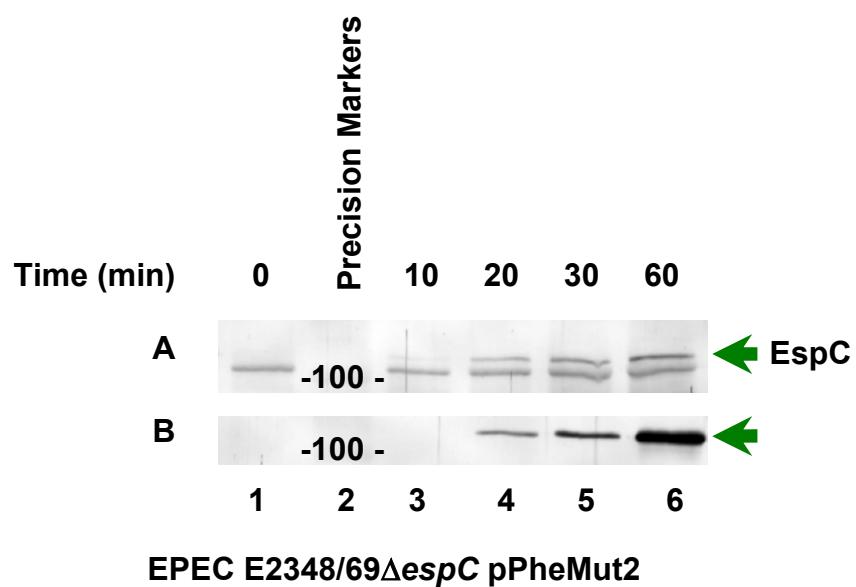


Figure 6.8: Western immunoblot analysis of whole cell lysates and TCA secreted proteins of the EPEC $\Delta$ espC deficient mutant hosting pPheMut2. Physical constraints of SDS-PAGE gels previously prevented inclusion of EPEC $\Delta$ espC deficient mutant hosting pPheMut2, therefore samples of lysates (A) and TCA precipitated supernatants (B) were applied to a separate SDS-PAGE gel and Western blotted. Accumulation of mutated EspC is observed in whole cell lysates (A) from 10 min (A3). Secretion is undetectable in supernatants (B) using chromogenic substrate until 20 min (B4). Increasing amounts of mutated EspC were detected over the course of 1 h (B6). The increased amount of EspC in supernatants at later time points reflects increased accumulation of the continuously secreted protein.

#### **6.4 Discussion.**

ATs are considered to generate pores in the bacterial OM as a function of their C-terminal  $\beta$ -barrel domain which facilitates export of the passenger-domain ([Viega et al, 2002](#); [Henderson et al, 2004](#)). Presumably, the  $\beta$ -barrel pore channel is either too small or has properties which would not normally allow entry of solvent molecules from the exterior of the cell. The lysis observed therefore when mutant EspC is produced in WT EPEC could be a consequence of several possibilities including (I) defective  $\beta$ -barrel pores resulting from association of  $\beta$ -barrels, (II) defective release of the  $\beta$ -barrel into the OM, or (III) misfolding of the autochaperone.

The mutation introduced to the GFS motif in pLAC2 does not appear to block insertion of the  $\beta$ -barrel into the OM, as it is apparently still secreted from the EspC mutant strains, though this may simply result from lysis of the cells expressing the mutant form of EspC. However transport of the passenger-domain across the membrane via BamA and/or its proteolytic processing, maturation, into the OM from BamA, is suggestive of a profound effect resulting from mutation of the GFS motif of the inter-domain leading to decreased cell viability, though it could also be suggested that EspC is aggregating in the IM, or periplasm and impacting detrimentally OM transport and cell viability. When both mutant and non-mutant forms of EspC were produced in WT EPEC, membrane integrity may have been compromised, affecting regulation of osmotic potential leading to cell lysis. The effects observed are thought to be indirect, perhaps resulting from altered conformation of the autochaperone domain. The autochaperone lies in the vicinity of an  $\alpha$ -helix motif common to many ATs, and which crystallography studies have revealed sits within the  $\beta$ -barrel ([Kostakioti et al, 2006](#); [Barnard et al, 2007](#); [leva et al, 2008](#)). The inter-domain is considered to affect folding and consequent secretion of the passenger ([Oliver et al, 2003](#)), therefore mutation of a motif likely responsible for threading it through the  $\beta$ -barrel may affect either/both pore formation and export. Localization of the  $\beta$ -barrel to determine whether perhaps the mutant EspC localizes in the OM, may be crucial to further understanding the precise nature of the effect of the site directed mutation.

As bacterial cell lysis was only demonstrated in the WT EPEC background, it could also be hypothesized that interaction and pore formation involving both WT and mutant EspC generates a defective hybrid  $\beta$ -barrel pore which perhaps has a larger channel or is

'stuck' open allowing influx of solute molecules. Conversely, pores generated purely from either WT or mutant EspC are resistant to solute influx. However, a number of studies clearly indicate that the serine protease ATs are monomeric, Tsh for example has been analyzed by native gel electrophoresis, gel filtration, and cross-linking analyses, providing evidence against a multimeric model with a central secretion pore ([Hritonenko et al, 2006](#)). These findings concur with other studies indicating that the EspP, AIDA, and NalP ATs are monomeric ([Muller et al, 2005](#); [Oomen et al, 2004](#); [Skillman et al, 2005](#)). In addition [Hritonenko et al, \(2006\)](#) point out that the multimeric model poses questions as to how a hydrophilic pore could be formed by the hydrophobic interface of the  $\beta$ -barrel monomers and how the OM lipids could be excluded from such a pore. However these studies are based on the expression of a single AT in a bacterial cell, not WT and mutated AT in the same system. The mutant EspC is suggested to be defective at passenger-domain export since less EspC accumulation is observed in supernatants from mutant strains (Fig 6.4, lanes 11 & 12). The observations regarding bacterial lysis could also imply that when mutant and non-mutant EspC are expressed in the same host bacterium they associate in a manner that would not be possible for either when expressed in the absence of the other, perhaps leading to the formation of multimeric  $\beta$ -barrel complexes which affect membrane integrity and cell viability, or associate in the periplasm preventing the BamA complex from facilitating AT  $\beta$ -barrel insertion into the OM rendering it potentially toxic to the host cell, and increasing its susceptibility to osmotic shock.

These possibilities remain to be explored. In the first instance, diffusion of solute molecules through putative EspC 'hybrid' pores can relatively easily be assessed using different size sugars or fluorescent dyes and will determine if this particular hypothesis is correct. The effect of the F<sup>977</sup> to Ala mutation to the GFS motif can also be assessed in the context of the inter-passenger domain interaction in the Y2H system in future studies.

## 7. General Discussion

Hundreds of AT proteins have been identified amongst diverse Gram-negative bacteria (Kajava and Steven, 2006; Pallen *et al*, 2003; Yen *et al*, 2002). Of the relatively small number of ATs which have been characterised, most are involved in pathogenesis. By implication, the remaining uncharacterised ATs identified by *in silico* analyses are thought likely to be associated with some role in virulence (Henderson *et al*, 2004).

Our initial *in silico* analysis of EspC indicated that it has similar domain structure to other ATs (see Henderson *et al*, 2004; Hritonenko *et al*, 2006; Skillman *et al*, 2005; Khalid and Sansom, 2006; Müller *et al*, 2005). By aligning EspC sequence with related sequences of other characterised ATs, we were able to define approximate EspC domain boundaries, and clearly identify signal-, passenger- and β-domains. Additionally using the known crystal structure of Hbp (Accession: 1WXR), we were able to generate a limited model of EspC based on truncated Hbp (Hbp2). Structural modelling encompassed the passenger domain region immediately prior to the clearly defined EspC β-domain. As this region is evidently part of the EspC passenger-domain, based on the downstream location of the β-domain consensus cleavage site, and because the passenger domains of both EspC and Hbp are considerably different from each other, our ability to model EspC on an equivalent region of Hbp suggests significant structural and functional conservation. Previous research has revealed that this region found in a number of ATs immediately upstream of the β-barrel (and encoded by our inter-domain) is crucial in the secretion process, this includes the ATs BrkA (Oliver *et al*, 2003), Ssp (Ohnishi *et al*, 1994), IgA protease (Klauser *et al*, 1993) and AIDAI (Maurer *et al*, 1999; Konieczny *et al*, 2001). Oomen *et al*, (2004) showed that the C-terminal portion closest to the NalP β-domain was helical and interacts with residues extending into the hydrophilic core of the β-domain. Additionally the inter-domain has also been suggested to be important in stabilization of the β-domain in the OM (Konieczny *et al*, 2001). Finally Oliver *et al*, (2003) and Ohnishi *et al*, (1994) have proposed the N-terminal region of the inter-region upstream of the putative α-helix may function as a chaperone.

In this study, we therefore aimed to characterise the EspC inter-domain and its contribution to EspC export. We used two main approaches. The first was to investigate interactions involving the putative inter-domain. The second was to assess the effect of mutation of key residues in the inter domain on EspC export. The former approach

employed several protein-protein interaction methods, and the latter, *in silico* modelling, mutagenesis and immunological techniques. Both research strands required generation of a considerable range of plasmid constructs, soluble recombinant protein, specific antibody reagents and isogenic EPEC mutant strains, all of which have been well characterised in this study and are available for future work.

To determine whether the putative EspC inter-domain was functionally important, we reasoned that in order to perform a role as chaperone, the inter-domain might be anticipated to directly interact with its passenger-domain ‘substrate’. Consequently, we employed the Y2H system to assess a range of interactions involving passenger-, inter- and β-domains, and also included the putative export accessory factor YbgC. Our data revealed for the first time that the region we defined as the EspC inter-domain was able to interact with the EspC passenger-domain, corroborating observations from research with other ATs, BrkA ([Oliver et al, 2003](#)), and EspP ([Verlade and Nataro, 2004](#); [Skillman et al, 2005](#)). This interaction was largely confirmed using recombinant protein in co-purification assays, although this secondary approach showed weaker interaction. Further definition of the inter-/passenger-domain interacting structures may increase the efficacy of co-purification approaches in future studies, although alternative methods of validation of the Y2H results can also be considered. To this end, several rationally defined truncations of both the passenger- and inter-domains were in the process of construction towards the end of the project in order to better define the inter-passenger interaction by both Y2H and co-purification approaches. However, time constraints prevented generation of the complete complement of required constructs for use in these assays; consequently, these experiments will remain the subject of future work.

Although we were unable to show an interaction between any EspC domain and the putative accessory factor YbgC, we clearly demonstrated that YbgC could dimerise. Whether YbgC is able to interact directly with EspC is unresolved since the propensity to dimerise could effectively preclude easy assessment of heteromeric interactions. If YbgC is required to dimerise prior to EspC interaction, the Y2H system, which detects pairwise interactions is unlikely to be a suitable method.

A number of homologues of YbgC such as 4-hydroxybenzoyl-CoA thioesterase ([Thoden et al, 2003](#)), form macromolecular structures. *In silico* analyses to model the *E. coli* YbgC sequence used to design primers for this study revealed homology to Chain E of

hydrolase (Accession: 1S5U). The hydrolase structure was shown to be composed of eight chains arranged in a homotetrameric structure similar to that described for 4-hydroxybenzoyl-CoA thioesterase *Pseudomonas* sp. (Benning *et al*, 1998). YbgC proteins are widely distributed among Gram-negative bacteria and are part of the Tol-Pal system, machinery crucial to maintenance of cell membrane integrity and cell division. No phenotype has been attributed however to deletion mutants of YbgC in *E.coli*, except for a contribution to EspC secretion (Dr. Louise Arnold PhD thesis). Gully and Bouveret (2006), hypothesised that YbgC could be required for synthesis, termination and the release from acyl-CoA of specific lipids which are then transferred to phospholipid synthetic enzymes i.e. YbgC could exert an indirect effect on EspC secretion through its involvement in lipid/membrane biogenesis/maintenance. This indirect association with EspC would also explain the lack of direct interaction within the Y2H system.

The EspC β-barrel could not be further investigated using the Y2H system as control constructs were found to self activate. As the β-domain is hydrophobic and in its native context, likely to insert into the bacterial OM, an alternative system to the Y2H system which does not require nuclear localisation is likely to be required to assess β-domain interactions. The split Ubiquitin system originally described by Johnsson and Vashavsky (1994) has been adapted to investigate membrane proteins (Stagljar *et al*, 1998) and may be more appropriate for the purpose of studying β-domain interactions.

Having demonstrated an EspC inter-passenger domain interaction, we wished to determine whether the region we defined could influence EspC export. Structural modelling of EspC against Hbp2 revealed several conserved structural motifs. The surface position of several of these motifs raised the possibility that they may have functional significance, perhaps of importance in conformation or contributing to a binding interface. One such prominent motif, evidently in a surface loop structure and unusually comprising the hydrophobic amino acid residue F (Phe), was selected as a candidate for mutagenesis experiments. Using site-directed mutagenesis of constructs pLAC1 and pLAC2 (encoding *espC*), the prominent Phe residue within the inter-domain was successfully mutated. Two single amino acid alterations and a triple amino acid alteration were successfully introduced to the pLAC2 (arabinose inducible) encoded *espC* insert, and a single amino acid alteration was also successfully achieved in pLAC1 (encoding constitutively expressed *espC*). The pLAC2 constructs containing the mutated inter-domain were introduced into both WT EPEC and an *espC* mutant strain generated

as part of this study. The transformed strains were exploited by monitoring the accumulation and production of EspC in both cells and supernatants with anti-sera to EspC (also generated as part of this study). Complementation of the EPEC *espC* mutant strain with EspC inter-domain mutants indicated that discrete mutation of the specifically targeted residues profoundly affects secretion. Furthermore, the expression of mutant EspC in WT EPEC was demonstrated to have a detrimental impact on cell viability, which could conceivably result from the formation of defective mutant export structures, or formation of a mutant and non-mutant complex in the periplasm, possibly preventing  $\beta$ -barrel insertion, and thus translocation.

As previously discussed, the observed lysis of WT EPEC strain producing mutant EspC was speculated to be caused by increased susceptibility to osmotic stress resulting from either unchecked glucose diffusion into the cells or possibly toxic accumulation of untranslocated EspC within the periplasm. The apparent osmotic stress phenomenon was manifested in the diminished accumulation of EspC observed in supernatants from overnight cultures and most likely explains the absence of EspC in lysates. Under normal circumstances, the  $\beta$ -barrel pore channel would presumably be required to regulate entry of solvent molecules from the exterior of the cell, either as a function of restrictive pore size or as a consequence of physical properties which inhibit entry/exit of solutes. Although studies have indicated that ATs are indeed monomeric ([Hritonenko et al, 2006](#); [Skillman et al, 2005](#); [Khalid and Sansom, 2006](#); [Müller et al, 2005](#)) it is possible that the expression of a mutant and non-mutant AT within the same cell associate in a manner that alone they are unable do. This association of mutant and non-mutant EspC could therefore form an oligomeric complex (analogous to that observed by [Viega et al, 2003](#)), either in the OM or the periplasm. Whichever of the two it may be, it infers that the GFS motif (within the inter-domain) of EspC is critical to either passenger-domain translocation, and/or  $\beta$ -barrel insertion into the OM. A possible role for the putative AT inter-domain in  $\beta$ -barrel localization, by interaction with the BamA complex has not been previously reported however, although [Konieczny et al, \(2001\)](#) have speculated that it plays a role in  $\beta$ -domain stabilization. An EspC ‘hybrid’ multimeric complex located in the OM could be relatively easily studied initially, by using diffusion of solute molecules via the use of different size sugars or fluorescent dyes to help determine if this hypothesis is correct. The effect of the F<sup>977</sup> to A<sup>977</sup> mutation to the function of the GFS motif can also be assessed in the context of the inter-passenger domain interaction in the Y2H systems in future studies. It is conceivable that if an association such as that proposed occurs,

then the interaction of mutant and non-mutant forms might be significantly stronger than that observed in this study.

It should be noted that [Verlarde and Nataro \(2004\)](#) have previously demonstrated that translocation of EspP is significantly affected by sequential deletions in the inter-domain region. One such deletion included the GFS motif identified in EspP and was also shown by these authors to significantly affect translocation of EspP. However, in this present study, the GFS motif of EspC was selected independently of this previous research on the basis of its location in the modeled structure of the EspC-inter domain.

To further investigate EspC interactions important for its export, an attempt was made to create a random representative genomic library of enteropathogenic *E. coli* strain E2348/69 in the Y2H plasmids. However, a library of sufficient titre could not be obtained in the time available and was therefore not included in this thesis. However, for the benefit of future studies, it is noted that this approach offers significant scope for investigation into potential interaction partners for the EspC inter- and passenger domains, particularly EspA ([Vidal and Navarro-García, 2008](#)).

Another particularly interesting observation was made during the course of this study during characterisation of the *espC* mutant, in that we observed elevated levels of FliC in the *espC* mutant compared to wildtype following labelling of Western blots with anti-FliC polyclonal sera. This finding was wholly unexpected, as FliC and EspB were anticipated to be produced at similar levels to that found in the WT EPEC strain. The elevated level of FliC produced in the *espC* mutant was in direct contrast to that EspB which was produced at equivalent levels in both EPEC wild type and mutant strains. Regulation of EspC and FliC are known to be linked by BipA, a member of the GTPase superfamily, and demonstrated by [Grant et al, \(2003\)](#) to regulate multiple cell surface and virulent components. We hypothesise that the deletion of *espC* may have inadvertently disrupted this BipA regulatory pathway ([Grant et al, 2003](#)), leading to increased expression of FliC. Alternatively EspC may form a part of the BipA regulatory pathway itself. Conceivably EspC is a protease which acts to degrade flagella and/or bfp which are no longer required once intimate attachment has taken place. Initial future work could address this possibility by using anti-BipA antibodies to investigate the expression of the regulator in both wild type and *espC* deletion strains. The level of EspC could also be investigated in FliC mutants, which are already available from the work of [Ferrández et al, \(2005\)](#).

Complementation studies could also be performed to determine if FliC can be down-regulated again once EspC expression is restored; over-expression plasmids encoding EspC are also already available for this. The converse should also be investigated in a double mutant background – does complementation with a plasmid encoding *fliC* reduce EspC expression or impact on T3SS? Potential interactions could be investigated between the two proteins, as the *fliC* gene has already been cloned in to the pET30a vector, the His-tag used successfully both for purification and to aid detection. In the absence of a direct FliC-EspC interaction, there is a possibility that a common regulatory gene or pathway (discussed earlier) may have been affected instead. Additionally motility assays should be conducted, to determine whether the *espC* mutant is hyper-motile when contrasted with the WT strain.

## 8. References

- Abe A, de Grado M, Pfuetzner RA, Sanchez-SanMartin C, DeVinney R, Puente JL, Strynadka NCJ, and Finlay BB.** Enteropathogenic *Escherichia coli* translocated intimin receptor, Tir, requires a specific chaperone for stable secretion. *Mol Microbiol*. 1999; **33**:1162-1175.
- Akeda Y, and Galán JE.** Chaperone release and unfolding of substrates in type III secretion. *Nature*. 2005; **437**:911-5.
- Akerley BJ, Cotter PA, and Miller JF.** Ectopic expression of the flagellar regulon alters development of the *Bordetella*-host interaction. *Cell*. 1995; **80**:611-620.
- Al-Hasani K, Henderson IR, Sakellaris H, Rajakumar K, Grant T, Nataro JP, Robins-Browne R, and Adler B.** The sigA gene which is borne on the she pathogenicity island of *Shigella flexneri* 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. *Infect Immun*. 2000; **68**:2457-63.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; **25**:3389-402
- Angelini A, Cendron L, Goncalves S, Zanotti G, Terradot L.** Structural and enzymatic characterization of HP0496, a YbgC thioesterase from *Helicobacter pylori*. *Proteins*. 2008; **72**:1212-21.
- Aoyama, K, Haase AM, and Reeves P.** Evidence for effect of random genetic drift on G+C content after lateral transfer of fucose pathway genes to *Escherichia coli*. *Mol. Biol. Evol*. 1994; **11**:829-838.
- Arnold L.** Characterisation of factors involved in the secretion of EspC in enteropathogenic *Escherichia coli*. 2003. PhD Thesis. University Of Nottingham.
- Attwood TK, and Parry-Smith DJ.** Introduction to Bioinformatics. 1999. Prentice Hall, Harlow, England.
- Bally M, Filloux A, Akrim M, Ball G, Lazdunski A, and Tommassen J.** Protein secretion in *Pseudomonas aeruginosa* characterisation of seven xcp genes and processing of secretory apparatus components by prepilin peptidase. *Mol. Microbiol*. 1992; **6**:1121-1131.
- Barenkamp SJ, and Leininger E.** Cloning, expression and DNA sequencing analysis of genes encoding non-typeable *Haemophilus influenzae* high-molecular-weight surface exposed proteins related to filamentous haemagglutinin of *Bordetella pertussis*. *Infect Immun*. 1992; **60**:1302-1313.
- Barnard TJ, Dautin N, Lukacik P, Bernstein HD, and Buchanan SK.** Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat Struct Mol Biol*. 2007; **14**:1214-20.
- Bartel P, Chien C-T, Sternglanz R, Fields S.** Elimination of false positives that arise in using the two-hybrid system. *Biotechniques* 1993; **14**:920-924.
- Bäumler AJ.** The record of horizontal gene transfer in *Salmonella*, *Trends Microbiol*. 1997; **5**: 318-322.
- Begue RE, Metha DI, Blecker U.** *Escherichia coli* and the Hemolytic-uremic Syndrome. *Southern Medical Journal*. 1998; **91**:798 – 804.
- Benjelloun-Touimi Z, Sansonetti PJ, Parsot C.** SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol*. 1995; **17**:123-135.

**Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D, and Holden HM.** The three dimensional structure of 4-Hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. Strain CBS-3. *J. Biol. Chem.* 1998; **50**: 33572-33579.

**Benz I, and Schmidt MA.** Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect. Immun* 1989; **57**:1506-1511.

**Berks BC.** A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* 1996; **22**:393-404.

**Bernstein HD.** Are bacterial 'autotransporters' really transporters? *Trends Microbiol.* 2007; **15**:441-7.

**Bingle LE, Bailey CM, Pallen MJ.** Type VI secretion: a beginner's guide. *Curr Opin Microbiol.* 2008; **11**:3-8.

**Bitter W, Koster M, Latijnhouwers M, de Cock H, and Tommassen J.** Formation of oligomeric rings by XcpQ and PilQ which are involved in protein transport across the outer membrane of *Pseudomonas aeruginosa*. *Mol. Microbiol.* 1998; **27**:209-219.

**Bitter W.** Secretins of *Pseudomonas aeruginosa*: large holes in the outer membrane. *Arch Microbiol.* 2003; **179**:307-314.

**Blattner FR, Plunkett GI, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden M, Rose DJ, Mau B, and Shao Y.** The complete genome sequence of *Escherichia coli* K-12. *Science* 1997; **277**:1453-1462.

**Bleves S, Lazdunski A, and Filloux A.** Membrane topology of three Xcp proteins involved in exoproteins transport by *Pseudomonas aeruginosa*. *J Bacteriol.* 1996; **178**:4297-4300

**Bolhuis A, Mathers JE, Thomas JD, Barrett CM, and Robinson C.** TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *J Biol Chem.* 2001; **276**:20213-20219.

**Bos MP, Robert V, and Tommassen J.** Biogenesis of the Gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* 2007; **61**:191–214.

**Bost S, and Belin D.** *prl* Mutations in the *Escherichia coli secG* Gene. *J. Biol. Chem.* 1997; **272**:4087-4093.

**Bruder W, Schmidt H, Karch H.** EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol Microbiol.* 1997; **24**:767-78.

**Brzuszkiewicz E, Brüggemann H, Liesegang H, Emmerth M, Olschläger T, Nagy G, Albermann K, Wagner C, Buchrieser C, Emody L, Gottschalk G, Hacker J, and Dobrindt U.** How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc. Natl. Acad. Sci. USA.* 2006; **103**:12879-84.

**Bulteris PV, Behrens S, Holst O, and Kleinschmidt JH.** Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharides. *J. Biol. Chem.* 2003; **278**: 9092-9099.

**Burns DL.** Biochemistry of type IV secretion. *Curr Opin Microbiol.* 1999; **2**: 25-29.

**Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadian V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J and Emili A.** Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 2005; **433**: 531–537.

**Cao TB, and Saier MH Jr.** Conjugal type IV macromolecular transfer systems of Gram-negative bacteria: organismal distribution, structural constraints and evolutionary conclusions. *Microbiol.* 2001; **147**:3201-3214.

**Chakrabarti R, and Schutt CE.** The enhancement of PCR amplification by low molecular weight sulphones. *Genes* 2001; **274**:293-298.

**Chen LY, Chen DY, Miaw J, and Hu NT.** XpsD, an outer membrane protein required for protein secretion by *Xanthomonas campestris* pv. *Campestris*, forms a multimer. *J. Biol. Chem.* 1996; **271**: 2703-2708.

**Chen M, Samuelson JC, Jiang F, Muller M, Kuhn A, Dalbey RE.** Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion. *J. Biol. Chem.* 2002; **277**:7670-7675.

**Chien C-T, Bartel PL, Sternblitz, and Fields S.** The two-hybrid system: A method to identify and clone genes for proteins that act with a protein of interest. *Proc. Natl. Acad. Sci. USA.* 1991; **88**:9578-9582

**Christie PJ.** Bacterial type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochem. Biophys. Acta.* 2004; **1694**:219-234.

**Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, and Cascales E.** Biogenesis, Architecture, and Function of Bacterial Type IV Secretion Systems. *Annu. Rev. Microbiol.* 2005; **59**:451-485.

**Clark SC.** Diarrhoeogenic *Escherichia coli* – an emerging problem? *Diag. Micro. Infect. Dis.* 2001; **41**: 93-98.

**Cline K, and Mori H.** Thykaloid ΔpH – dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4 – dependent transport. *J Cell. Biol.* 2001; **154**: 719-729.

**Collin S, Guilvout I, Chami M, and Pugsley AP.** YaeT-independent multimerization and outer membrane association of secretin PufD. *Mol. Microbiol.* 2007; **64**:1350–1357.

**Comanducci M, Bambini S, Brunelli B, Adu-Bobie J, Aricò B, Capecchi B, Giuliani MM, Massignani V, Santini L, Savino S, Granoff DM, Caugant DA, Pizza M, Rappuoli R, and Mora M.** NadA, a novel vaccine candidate of *Neisseria meningitidis*. *J Exp Med.* 2002; **195**:1445-1454.

**Cope LD, Thomas SE, Latimer JL, Slaughter CA, Müller-Eberhard U, and Hansen EJ.** The 100kDa Haem:Haemopexin binding protein of *Haemophilus influenzae*: structure and localisation. *Mol Microbiol.* 1994; **13**: 863-873.

**Cope LD, Thomas SE, Hrkal Z, and Hansen EJ.** Binding of Haem-Haemopexin complexes by soluble HxuA protein allows utilization of this complexed Haeme by *Haemophilus influenzae*. *Infect Immun.* 1998; **66**: 4511-4516.

**Cordes FS, Komoriya K, Larquet E, Yang S, Egelman EH, Blocker A, and Lea SM.** Helical structure of the needle of the type III secretion system of *Shigella flexneri*. *J Biol Chem.* 2003; **278**:17103-7

**Cornelis GR.** The type III secretion injectisome. *Nat. Rev. Microbiol.* 2006; **4**:811-825.

**Cotter SE, Surana NK, St Geme JW 3<sup>rd</sup>.** Trimeric autotransporters: a distinct family of autotransporter proteins. *Trend Microbiol* 2005; **13**:199-205

**Crane JK, McNamara BP, and Donnenberg MS.** Role of EspF in host cell death induced by Enteropathogenic *Escherichia coli*. *Cell. Microbiol.* 2001; **3**:197-211.

**Creasey EA, Friedberg D, Shaw RK, Umanski T, Knutton S, Rosenshine I, and Frankel G.** CesAB is an enteropathogenic *Escherichia coli* chaperone for the type-III translocator proteins EspA and EspB. *Microbiol.* 2003a; **149**:3629-3647.

**Creasey EA, Delahay RM, Bishop AA, Shaw RK, Kenny B, Knutton S, and Frankel G.** CesT is a bivalent enteropathogenic *Escherichia coli* chaperone required for translocation of both Tir and Map. *Mol Microbiol.* 2003b; **47**:209-221.

**Cristobal S, Scotti P, Luirink J, von Heijne G, and de Gier J.-W.** The signal recognition particle – targeting pathway does not necessarily deliver proteins to the Sec-translocase in *Escherichia coli*. *J.Biol. Chem.* 1999; **274**:20068 – 20070.

**Dahan S, Wiles S, La Ragione RM, Best A, Woodward MJ, Stevens MP, Shaw RK, Chong Y, Knutton S, Phillips A, and Frankel G.** EspJ is a prophage-carried type III effector protein of attaching and effacing pathogens that modulates infection dynamics. *Infect Immun.* 2005; **73**:679-686.

**Dalbey RE, Kuhn A.** YidC family members are involved in membrane insertion, lateral integration, folding and assembly of membrane proteins. *J. Cell. Biol.* 2004; **166**:769-74.

**Daniell SJ, Delahay RM, Shaw RK, Hartland CL, Pallen MJ, Booy F, Knutton S, and Frankel G.** Coiled-coil domain of enteropathogenic *Escherichia coli* type III secreted protein EspD is involved in EspA filament mediated cell attachment and hemolysis. *Infection and Immunity.* 2001; **69**:4055-4064.

**Das A, and Xie YH.** Construction of a transposon TnphoA: its application in defining the membrane topology of the *Agrobacterium tumefaciens* DNA transfer proteins. *Mol. Microbiol.* 1998; **27**:405-414.

**Das S, Chakrabortty A, Banerjee R, Roychoudhury S, Chaudhuri K.** Comparison of global transcription responses allows identification of *Vibrio cholerae* genes differentially expressed following infection. *FEMS Microbiol Lett.* 2000; **190**:87-91

**Datsenko KA, and Wanner BL.** One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 2000; **97**:6640-6645.

**Dautin N, and Bernstein HD.** Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol.* 2007; **61**:89-112.

**Davis J, Smith AL, Hughes WR, and Golomb M.** Evolution of an autotransporter: domain shuffling and lateral transfer from pathogenic *Haemophilus* to *Neisseria*. *J. Bact.* 2001; **183**:4626-4635.

**de Gier JW and Luirink J.** The ribosome and YidC: New insights into the biogenesis of *Escherichia coli* inner membrane proteins. *EMBO Rep.* 2003; **4**:939-943.

**Delahay RM, Shaw RK, Elliott SJ, Kaper JB, Knutton S, and Frankel G.** Functional analysis of the enteropathogenic *Escherichia coli* type III secretion system chaperone CesT identifies domains that mediate substrate interactions. *Mol Microbiol.* 2002; **43**:61-73.

**Delepelaire P, and Wanderman C.** The SecB chaperone is involved in the secretion of the *Serratia marcescens* HasA protein through an ABC transporter. *EMBO J.* 1998; **17**:936-944.

**Delepelaire P.** Type I secretion in Gram-negative bacteria. *Biochimica et Biophysica Acta* 2004; **1694**:149-161

**de Leeuw E, Granjon T, Porcelli I, Alami M, Carr SB, Müller M, Sargent F, Palmer T, and Berks BC.** Oligometric properties and signal peptide binding by *Escherichia coli* Tat protein complexes. *J. Mol. Biol.* 2002; **322**: 1135-1146.

**de Lisa MP, Tullman D, Georgiou G.** Folding quality in the export of proteins by the bacterial twinarginine translocation pathway. *Proc. Natl. Acad. Sci USA.* 2003; **100**:6115-6120.

**d'Enfert C, Ryter A, Pugsley AP.** Cloning and expression in *Escherichia coli* of the *Klebsiella pneumoniae* genes for production, surface location and secretion of the lipoprotein pullulanase. *EMBO J.* 1987; **6**:3531-3538.

**Deng W, Puente JL, Gruenheid S, Li Y, Vallance BA, Vazquez A, Barba J, Ibarra JA, O'Donnell P, Metalnikov P, Ashman K, Lee S, Goode D, Pawson T, and Finlay BB.** Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc. Natl. Acad. Sci. USA.* 2004; **101**:3597-3602.

**Desvaux M, Scott-Tucker A, Turner SM, Cooper LM, Huber D, Nataro JP, and Henderson IR.** A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism. *Microbiology.* 2007; **153**:59-70

**Dinh T, Paulsen IT, Saier MH Jr.** A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of Gram negative bacteria. *J Bacteriol* 1994; **176**:3825-3831.

**Doerrler WT, and Raetz CR.** Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J Biol. Chem.* 2005; **280**:27679-27687.

**Dower WJ, Miller JF, Ragsdale CW.** High Efficiency transformation of *E.coli* by high voltage electroporation. *Nucleic acids Res.* 1988; **16**:6127-6145.

**Drago-Serrano ME, Parra SG, Manjarrez-Hernandez HA.** EspC, an autotransporter protein secreted by enteropathogenic *Escherichia coli* (EPEC), displays protease activity on human hemoglobin. *FEMS Microbiol Lett.* 2006; **265**:35-40.

**Driessen AJM, Fekkes P. and van der Wolk JPW.** The Sec system. *Curr Opin Microbiol.* 1998; **1**:216-222.

**Driessen AJ, Nouwen N.** Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem.* 2008; **77**:643-67

**Dubuisson JF, Vianney A, Hugouvieux-Cotte-Pattat N, Lazzaroni JC.** Tol-Pal proteins are critical cell envelope components of *Erwinia chrysanthemi* affecting cell morphology and virulence. *Microbiology.* 2005; **151**:3337-47.

**Dudley EG, Thomson NR, Parkhill J, Morin NP, and Nataro JP.** Proteomic and microarray characterisation of the AggR regulon identifies a *pheU* pathogenicity island in enteroaggregative *Escherichia coli*. *Mol. Micro.* 2006; **61**:1267-1282.

**Duong F, and Wickner W.** The SecDyajC domain of pre-protein translocase controls pre-protein movement by regulating SecA membrane cycling. *EMBO J.* 1997; **16**:4871-4879.

**Duong F, Eichler J, Price A, Leonard MR, Wickner W.** Biogenesis of the gram-negative bacterial envelope. *Cell.* 1997; **91**:567-573.

**Dutta PR, Cappello R, Navarro-Garcia F, Nataro JP.** Functional comparison of serine protease autotransporters of enterobacteriaceae. *Infect Immun.* 2002; **70**:7105-13.

**Economou A, and Wickner W.** SecA promotes preprotein translocation by undergoing ATP driven cycles of membrane insertion and deinsertion. *Cell.* 1994; **78**:835-843.

**Economou A, Pogliano JA, Beckwith J, Oliver DB, and Wickner W.** SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell.* 1995; **83**:1171-1181.

**Economou A.** Bacterial preprotein translocase: mechanism and conformational dynamics of a processive enzyme. *Mol. Microbiol.* 1998; **27**:511-518.

**Economou A.** Following the Leader: bacterial protein export through the sec pathway. *Trends Microbiol.* 1999; **7**:315-320.

**Eichelberg K, Galán JE.** The flagellar sigma factor FliA (sigma(28)) regulates the expression of *Salmonella* genes associated with the centisome 63 type III secretion system. *Infect Immun.* 2000; **68**:2735-43

**Elliott SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng Y, Lai LC, McNamara BP, Donnenberg MS, and Kaper JB.** The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol Microbiol* 1998; **28**:1-4.

**Elliot SJ, Hutcheson SW, Dubois MS, Mellies JL, Wainwright LA, Batchelor M, Frankel G, Knutton S, and Kaper JB.** Identification of CesT, a chaperone for the type III secretion of Tir in enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 1999; **33**:1176-1189.

**Elliot SJ, Sperandio V, Giron JA, Shin S, Mellies JL, Wainwright L, Hutcheson SW, McDaniel TK, and Kaper JB.** The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE and non-LEE encoded virulence factors in Enteropathogenic and enterohaemorrhagic *Escherichia coli*. *Infect. Immun.* 2000; **68**:6115-6126.

**Elliott SJ, Krejany EO, Mellies JL, Robins-Browne RM, Sasakawa C, Kaper JB.** EspG, a novel type III system-secreted protein from enteropathogenic *Escherichia coli* with similarities to VirA of *Shigella flexneri*. *Infect Immun.* 2001; **69**:4027-33.

**Elliot SJ, O'Connell CB, Koutsouris A, Brinkley C, Donnenberg MS, Hecht G, and Kaper JB.** A gene from the locus of enterocyte effacement that is required for enteropathogenic *Escherichia coli* to increase tight junction permeability encodes a chaperone for EspF. *Infect. Immun.* 2002; **70**: 2271-2277.

**Eslava C, Navarro-Garcia F, Czeczulin JR, Henderson IR, Cravioto A, and Nataro JP.** Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun.* 1998; **66**:3155-63.

**Fath MJ, and Kolter R.** ABC Transporters: bacterial exporters. *Microbiol Rev* 1993; **57**:995-1017.

**Fekkes P, van der Does, and Driessens AJ.** The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J.* 1997; **16**:6105-6113.

**Fekkes P, de Wit JG, van der Wolk JP, Kimsey HH, Kumamoto CA, and Driessens AJ.** Preprotein transfer to the *Escherichia coli* translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol Microbiol* 1998; **29**:1179-1190.

**Felmlee T, Pellett S, and Welch RA.** Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. *J Bact* 1985; **163**:94-105.

**Fernandez RC, and Weiss AA.** Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect Immun.* 1994; **62**:4727-38

**Ferrández M-J, Bishop K, Williams P, Withers H.** HosA, a member of the SlyA family, regulates motility in enteropathogenic *Escherichia coli*. *Infect Immun* 2005; **73**: 1684-1694.

**Fields S, and Song O.** A novel genetic system to detect protein-protein interactions. *Nature*. 1989; **340**:245-6.

**Filloux A.** The underlying mechanisms of type II protein secretion. *Biochim Biophys Acta*. 2004; **1694**:163-79.

**Filloux A, Hachani A, Bleves S.** The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology*. 2008; **154**:1570-83.

**Foubister V, Rosenshine I, Donnenberg MS, and Finlay BB.** The eaeB gene of enteropathogenic *Escherichia coli* is necessary for signal transduction in epithelial cells. *Infect Immun* 1994; **62**: 3038-3040.

**Francetic O, and Pugsley AP.** The cryptic general secretory pathway (*gsp*) operon of *Escherichia coli* K-12 encodes functional proteins. *J Bact.* 1996; **178**:3544-3549.

**Frankel G, Phillips AD, Rosenshine I, Dougan G, Kaper JB, and Knutton S.** Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol*. 1998; **30**: 911-921.

**Friedberg D, Umanski T, Fang Y and Rosenshine I.** Hierarchy in the expression of the locus of enterocyte effacement genes of enteropathogenic *Escherichia coli*. *Mol Microbiol*. 1999; **34**: 941-952.

**Frithz-Lindsten E, Du Y, Rosqvist R, and Forsberg A.** Intracellular targeting of exoenzyme S of *P. aeruginosa* via type III-dependant translocation induces phagocytosis resistance, cytotoxicity and disruption of actin microfilaments. *Mol. Microbiol*. 1997; **25**: 1125-1139.

**Fronzes R, Schäfer E, Wang L, Saibil HR, Orlova EV, and Waksman G.** Complex Structure of a Type IV Secretion System Core. *Science*. 2009; **323**:266-268.

**Geer LY, Domrachev M, Lipman DJ, Bryant SH.** CDART: protein homology by domain architecture. *Genome Res.* 2002; **12**: 1619-1623.

**Gerlach RG, and Hensel M.** Protein secretion systems and adhesins: The molecular armoury of Gram-negative pathogens. *Int J. Med. Microbiol.* 2007; **297**:401-415.

**Ghang G.** Structure of MsbA from *Vibrio cholera*: A Multidrug Resistance ABC Transporter Homolog in a Closed Conformation. *J. Mol. Biol.* 2003; **330**:419-430.

**Ghosh P.** Process of protein transport by the Type III Secretion System. *Microbiol Mol. Biol. Revs.* 2004; **68**:771-795.

**Gietz RD, Schiestl RH.** Transforming Yeast with DNA. *Methods Mol Cell Biol.* 1995; **5**:255-269.

**Girón JA, Torres AG, Freer E, and Kaper JB.** The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Mol Microbiol*. 2002; **44**:361-79.

**Gohlke U, Pullan L, McDevitt CA, Porcelli I, de Leeuw, Palmer T, Saibil HR, Berks BC.** The TatA component of the twin-arginine protein transport system forms complexes of variable diameter. *Proc. Natl. Acad. Sci. USA*. 2005; **102**:10482-10486.

**Goldberg MB, Barzu O, Parsot C, and Sansonetti PJ.** Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. *Infect Agents Dis.* 1993; **2**:210-1

**Goodier RI, and Ahmer BM.** SirA orthologs affect both motility and virulence. *J. Bacteriol.* 2001; **183**:2249-2258.

**Grant AJ, Farris M, Alefounder P, Williams PH, Woodward MJ, and O'Connor CD.** Co-ordination of pathogenicity island expression by the BipA GTPase in enteropathogenic *Escherichia coli* (EPEC). *Mol Microbiol*. 2003; **48**:507-21

**Grass S, and St. Geme<sup>3rd</sup> JW.** Maturation and secretion of the non-typable *Haemophilus influenzae* HMW1 adhesin: roles of the N-terminal and C-terminal domains. *Mol Microbiol* 2000; **36**: 55-67.

**Guédin S, Willery E, Tommassen J, Fort E, Drobecq H, Locht C, and Jacob-Dubisson F.** Novel topological features for FhaC, the outer membrane transporter involved in the secretion of *Bordetella pertussis* filamentous haemagglutinin. *J. Biol. Chem.* 2000; **274**:37731-37735.

**Guilvout I, Chami M, Berrier C, Ghazi A, Engel A, Pugsley AP, and Bayan N.** *In vitro* multimerization and membrane insertion of bacterial outer membrane secretin PulD. *J. Mol. Biol.* 2008; **382**: 13–23.

**Gully D, and Bouveret E.** A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in *E. coli*. *Proteomics*. 2006; **6**: 282-293.

**Guyer DM, Henderson IR, Nataro JP, and Mobley HLT.** Identification of Sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol. Microbiol.* 2000; **38**:53-66.

**Guzman L-M, Belin D, Carson MJ, and Beckwith J.** Tight regulation, modulation and High-Level Expression by Vectors Containing the Arabinose P<sub>BAD</sub> Promoter. *J. Bact.* 1995; **177**:4121-4130.

**Hanahan D.** Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol.* 1983; **166**:557-80.

**Hardie KR, Seydel, Guilvout I, and Pugsley AP.** The secretin specific, chaperone-like protein of the general secretory pathway: separation of proteolytic protection and piloting functions. *Mol. Microbiol.* 1996; **22**:967-976.

**Harlow E, and Lane D.** Antibodies: A Laboratory Manual. 1988. Published by Cold Spring Harbour Laboratory, USA.

**Harris RL, Hombs V, and Silverman PM.** Evidence that F-plasmid proteins TraV, TraK and TraB assemble into an envelope-spanning structure in *Escherichia coli*. *Mol. Microbiol.* 2001; **42**: 757-66.

**Harshey RM, and Toguchi A.** Spinning tails: homologies among bacterial flagellar systems. *Trends Microbiol.* 1996; **4**:226-231.

**Hartland EL, Batchelor M, Delahay RM, Hale C, Matthews S, Dougan G, Knutton S, Connerton I, Frankel G.** Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Mol Microbiol.* 1999; **32**:151-8.

**Hauser AR, Kang PJ, Fleiszig SJM, Mostov K, and Engel J..** Defects in type III secretion correlate with internalization of *Pseudomonas aeruginosa* by epithelial cells. *Infect. Immun.* 1998; **66**:1413-1420.

**Heidelberg J, and Sebastian Y.** Annotation of *Vibrio harveyi* HY01 (*Unpublished*). NCBI (Direct submission) 2007.

**Henderson IR, Navarro-Garcia F, and Nataro JP.** The great escape: structure and function of the autotransporter proteins. *Trends Microbiol.* 1998; **6**:370-378.

**Henderson IR, Czeczulin J, Eslava C, Noriega F, and Nataro JP.** Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect. Immun.* 1999; **67**:5587-5596.

**Henderson IR, and Nataro JP.** Virulence functions of autotransporter proteins. *Infect Immun.* 2001; **69**:1231-43.

**Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC and Ala'Aldeen D.** Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev.* 2004; **68**:692-744.

**Henrixson DR, de la Morena ML, Stathopoulos C, St Geme JW 3<sup>rd</sup>.** Structural determinants of processing and secretion of the *Haemophilus influenzae* hap protein. *Mol Microbiol* 1997; **26**:505-518.

**Herskovits AA, Bochkareva ES, and Bi E.** New projects in studying the bacterial signal recognition particle pathway. *Mol. Microbiol.* 2000; **38**:927-939.

**Hertle R.** *Serratia marcescens* haemolysin (ShIA) binds artificial membranes and forms pores in a receptor-independent manner. *J Membr Biol.* 2002; **189**:1-14.

**Hertle R, Hilger M, Weingardt-Kocher S, and Walev I.** Cytotoxic action of *Serratia marcescens* haemolysin on human epithelial cells. *Infect Immun.* 1999; **67**:817-825.

**Hinsley AP, Stanley NR, Palmer T, and Berks BC.** A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. *FEBS Lett.* 2001; **497**:45-9.

**Hirono I, Tange N, and Aoki T.** Iron-regulated haemolysin from *Edwardsiella tarda*. *Mol. Microbiol.* 1997; **24**:851-856.

**Howard SP, Critch J, and Bedi A.** Isolation and analysis of eight exe genes and their involvement in extracellular protein secretion and outer membrane assembly in *Aeromonas hydrophila*. *J. Bacteriol.* 1993; **175**:6695-6703.

**Hritonenko V, Kostakioti M, and Stathopoulos C.** Quaternary structure of a SPATE autotransporter protein. *Mol Membr Biol.* 2006 **23**:466-74

**Hueck CJ.** Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Molec Biol Rev* 1998; **62**: 379-433.

**Ide T, Laarmann S, Greune L, Schillers H, Oberleithner H, and Schmidt MA.** Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic *Escherichia coli*. *Cell Microbiol.* 2001; **3**: 669-79.

**Ieva R, Skillman KM, Bernstein HD.** Incorporation of a polypeptide segment into the beta-domain pore during the assembly of a bacterial autotransporter. *Mol Microbiol.* 2008; **67**:188-201.

**Inoue H, Nojima H, and Okayama H.** High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 1990; **96**:23-28

**Jacob-Dubuisson F, Buisine C, Willery E, Renauld-Mongénie G, and Locht C.** Lack of functional complementation between *Bordetella pertussis* filamentous haemagglutinin and *Proteus mirabilis* HpmA haemolysin secretion machineries. *J Bacteriol.* 1997; **179**:775-83.

**Jacob-Dubuisson F, Locht C, Antoine R.** Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol Microbiol* 2001; **40**: 306–313.

**Jacob-Dubuisson F, Fernandez R, and Coutte L.** Protein secretion through autotransporter and two-partner pathways. *Biochim Biophys Acta.* 2004; **1694**:235-57

**Jain S, and Goldberg MB.** Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol.* 2007; **189**:5393-8.

**Jakubowski S, Krishnamoorthy V, Cascales E, and Christie PJ.** *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system. *J. Mol. Biol.* 2004; **341**: 961-977.

**James P, Halladay J, and Craig EA.** Genomic Libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics.* 1996; **144**:1425-1436.

**Jarvis KG, Giron JA, Jerse AE, McDaniel TK, Donnenberg MS, and Kaper JB.** Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in the attaching and effacing lesion formation. *Proc. Natl. Acad. Sci. USA.* 1995; **92**:7996-8000.

**Jepson MA, Pellegrin S, Peto L, Banbury DN, Leard AD, Mellor H, and Kenny B.** Synergistic roles for the Map and Tir effector molecules in mediating uptake of enteropathogenic *Escherichia coli* (EPEC) into non-phagocytic cells. *Cell Microbiol.* 2003; **5**:773-83.

**Jerse AE, Yu J, Tall BD, and Kaper JB.** A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc. Natl. Acad. Sci. U S A* 1990; **87**, 7839–7843

**Jerse AE, and Kaper JB.** The eae gene of enteropathogenic *Escherichia coli* encodes a 94-kilodalton membrane protein, the expression of which is influenced by the EAF plasmid. *Infect Immun* 1991; **59**: 4302-4309.

**Johnson TJ, Kariyawasam S, Wannemuehler Y, Mangiamele P, Johnson SJ, Doetkott C, Skyberg JA, Lynne AM, Johnson JR, and Nolan LK.** The Genome Sequence of Avian Pathogenic *Escherichia coli* Strain O1:K1:H7 Shares Strong Similarities with Human Extraintestinal Pathogenic *E. coli* Genomes *J. Bacteriol.* 2007; **189**: 3228-3236

**Johnson N, and Varshavsky A.** Split Ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl. Acad. Sci. USA*. 1994; **91**:10340-10344.

**Jong WS, ten Hagen-Jongman CM, den Blaauwen T, Slotboom DJ, Tame JR, Wickström D, de Gier JW, Otto BR, and Luirink J.** Limited tolerance towards folded elements during secretion of the autotransporter Hbp. *Mol Microbiol.* 2007; **63**:1524-36.

**Jongbloed JDH, Martin U, Antelman H, Hecker M, Tjalsma H, Venema G, Bron S, van Dijl JM, and Müller J.** TatC is specificity determinant for protein secretion via the twin-arginine translocation pathway. *J. Biol. Chem.* 2000; **275**:41350-41357

**Jose J, Jahnig F, and Meyer TF.** Common Structural Features of IgA1 protease like outer membrane protein autotransporters. *Mol Micro* 1995; **18**:378-380.

**Journet L, Agrain C, Broz P, and Cornelis GR.** The needle length of bacterial injectisomes is determined by a molecular ruler. *Science*. 2003; **302**:1757-60.

**Kajava AV, and Steven AC.** The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol.* 2006; **155**:306-15

**Kanack KJ, Crawford A, Tatsuno I, Karmali MA, and Kaper JB.** SepZ/EspZ is secreted and translocated into HeLa Cells by the Enteropathogenic *Escherichia coli* Type III Secretion System. *Infect. Immun.* 2005; **73**:4327-4337.

**Kaper JB.** EPEC delivers the goods. *Trends Microbiol.* 1998; **6**: 169-172.

**Karaolis DK, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, and Reeves PR.** A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc. Natl. Acad. Sci. USA* 1998; **95**:3134-3139

**Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JC, and Dorman CJ.** A global role for Fis in the transcriptional control of metabolism and type III secretion in *Salmonella enterica* serovar Typhimurium. *Microbiology*. 2004; **150**:2037-53.

**Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA, and Finlay BB.** Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 1997; **91**: 511-520

**Kenny B.** Mechanism of action of EPEC type III effector molecules. *Int J Med Microbiol.* 2002; **291**:469-77.

**Khalid S, and Sansom MS.** Molecular dynamics simulations of a bacterial autotransporter: NalP from *Neisseria meningitidis*. *Mol Membr Biol.* 2006; **23**:499-508.

**Kim Y, Joachimiak A, Skarina T, Savchenko A and Edwards A.** Crystal Structure Of Hypothetical Protein Ec709 From *Escherichia coli* (*Unpublished*). NCBI (Direct submission) 2004.

**Kjaergaard K, Schembri MA, Hasman H, and Klemm P.** Antigen 43 from *Escherichia coli* induces inter- and intraspecies cell aggregation and changes in colony morphology of *Pseudomonas fluorescens*. *J Bacteriol.* 2000a; **182**:4789-96.

**Kjaergaard K, Schembri MA, Ramos C, Molin S, and Klemm P.** Antigen 43 facilitates formation of multispecies biofilms. *Environ. Microbiol.* 2000b; **2**:695-702.

**Klauser T, Kramer J, Otzelberger K, Pohlner J, and Meyer TF.** Characterisation of the *Neisseria* IgA beta-core. The essential unit for outer membrane targeting and extracellular protein secretion. *J Mol Biol* 1993; **234**:579-593.

**Klemm P, Vejborg RM, and Sherlock O.** Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol.* 2006; **296**:187-95.

**Konieczny MPJ, Benz I, Hollinderbäumer B, Beinke C, Niederweis M, and Schmidt MA.** Modular organization of the AIDA autotransporter translocator: the N-terminal beta1-domain is surface-exposed and stabilizes the transmembrane beta2-domain. *Antonie Van Leeuwenhoek.* 2001; **80**:19-34

**Konkel ME, Klena JD, Rivera-Amill V, Monteville MR, Biswas D, Raphael B, and Mickelson J.** Secretion of virulence proteins from *Campylobacter jejuni* is dependent on a functional flagellar export apparatus. *J Bacteriol.* 2004; **186**:3296-3303.

**Koronakis V, Koronakis E, and Hughes C.** Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* haemolysin protein across both bacterial membranes. *EMBO J.* 1989; **8**:595-605.

**Knowles TJ, Scott-Tucker A, Overduin M, and Henderson IR.** Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nature.* 2009; **7**:206-214

**Knutton S, Baldwin T, Williams PH, and McNeish AS.** Actin accumulation at sites of bacterial adhesion to tissue culture cells: basis of a new diagnostic test for enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun* 1989; **57**:1290-1298.

**Knutton S, Adu-Bobie J, Bain C, Phillips AD, Dougan G, and Frankel G.** Down regulation of intimin expression during attaching and effacing enteropathogenic *Escherichia coli* adhesion. *Infect Immun.* 1997; **65**:1644-52.

**Knutton S, Rosenshine I, Pallen MJ, Nisan I, Neves BC, Bain C, Wolff C, Dougan G, and Frankel G.** A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *EMBO J.* 1998; **17**: 2166-2176.

**Knutton S, Shaw RK, Anantha RP, Donnenberg MS, and Zorgani AA.** The type IV bundle-forming pilus of enteropathogenic *Escherichia coli* undergoes dramatic alterations in structure associated with bacterial adherence, aggregation and dispersal. *Mol Microbiol* 1999; **33**: 499-509.

**Kostakioti M, Stathopoulos C.** Role of the alpha-helical linker of the C-terminal translocator in the biogenesis of the serine protease subfamily of autotransporters. *Infect Immun.* 2006; **74**:4961-9.

**Kresse AU, Rohde M, and Guzman CA.** The EspD protein of enterohemorrhagic *Escherichia coli* is required for the formation of bacterial surface appendages and is incorporated in the cytoplasmic membranes of target cells. *Infect. Immun.* 1999; **67**: 4834-4842.

**Kroll, J. S., K. E. Wilks, J. L. Farrant, and P. R. Langford.** Natural genetic exchange between *Haemophilus* and *Neisseria*: intergeneric transfer of chromosomal genes between major human pathogens. *Proc. Natl. Acad. Sci. USA.* 1998; **95**:12381-12385

**Kushner SH.** An improved method for transformation of *Escherichia coli* with ColE1 derived plasmids. In *Genetic Engineering*. Boyer HW, Nicosia S. (Eds) NY: Elsevier/North Holland Biomedical Press. 1978; p173

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

**Lazar SW, and Kolter R.** SurA assists the folding of *Escherichia coli* outer membrane proteins. *J Bacteriol.* 1996; **178**: 1770-1773.

**Lederberg J, and Lederberg EM.** Replica plating and indirect selection of bacterial mutants. *J Bacteriol.* 1952; **63**:399-406

**Lee PA, Tullman-Ercek D, and Georgiou G.** The bacterial twin-arginine translocation pathway. *Annu. Rev. Microbiol.* 2006; **60**: 373-395.

**Letley DP, Rhead JL, Bishop K, and Atherton JC.** Paired cysteine residues are required for high levels of the *Helicobacter pylori* autotransporter VacA. *Microbiol.* 2006; **152**:1319-1325.

**Letzelter M, Sorg I, Mota LJ, Meyer S, Stalder J, Feldman M, Kuhn M, Callebaut I, and Cornelis GR.** The discovery of SycO highlights a new function for type III secretion effector chaperones. *EMBO J.* 2006; **25**:3223-33.

**Levine MM, Bergquist EJ, Nalin DR, Waterman DH, Hornick RB, Young CR, and Sotman S.** *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet.* 1978; **1**:1119-1122

**Leyton DL, Sloan J, Hill RE, Doughty S, and Hartland EL.** Transfer region of pO113 from enterohaemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect. Immun.* 2003; **71**:6307-6319.

**Llosa M, Zunzunegui S, and de la Cruz F.** Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxomes. *Proc. Natl. Acad. Sci. USA.* 2003; **100**:10465-10470.

**Lloyd SA, Sjostrum M, Anderson S, and Wolf-Watz.** *Yersinia* YopE is targeted for type III secretion by N-terminal, not mRNA signals. *Mol. Microbiol.* 2001; **39**:520-531.

**Lloyd SA, Sjostrum M, Anderson S, and Wolf-Watz.** Molecular characterisation of type III secretion signals via analysis of synthetic N-terminal amino acid sequences. *Mol. Microbiol.* 2002; **43**:51-59.

**Liu Z, and Binns AN.** Functional subsets of the virB type IV transport complex proteins involved in the capacity of *Agrobacterium tumefaciens* to serve as a recipient in virB-mediated conjugal transfer of plasmid RSF1010. *J. Bacteriol.* 2003; **185**:3259-69.

**Luirink J, High S, Wood H, Giner A, Tollervey D, and Dobberstein B.** Signal-sequence recognition by an *Escherichia coli* ribonucleoprotein complex. *Nature.* 1992; **359**:741-743.

**Luirink J, and Dobberstein B.** Mammalian and *Escherichia coli* signal recognition particles. *Mol. Microbiol.* 1994; **11**:9-13.

**Luirink J, von Heijne G, Houben E, and de Gier JW.** Biogenesis of inner membrane proteins in *Escherichia coli*. *Annu Rev Microbiol.* 2005; **59**:329-55

**Ma J, and Ptashne M.** Converting a eukaryotic transcriptional inhibitor into an activator. *Cell.* 1988; **55**:443-6.

**Ma X, and Cline K.** Precursors bind to specific sites on thylakoid membranes prior to transport on the delta pH protein translocation system. *J. Biol Chem.* 2000; **275**:10016-10022.

**Makino K, Yokoyama K, Kubota Y, Yutsudo CH, Kimura S, Kurokawa K, Ishii K, Hattori M, Tatsuno I, Abe H, Iida T, Yamamoto K, Ohnishi M, Hayashi T, Yasunaga T, Honda T, Sasakawa C and Shinagawa H.** Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak *Genes Genet. Syst.* 1999; **74**: 227-239.

**Malinvern JC, Werner J, Seokhee K, Sklar JG, Kahne D, Misra R, and Silhavy TJ.** YfiO stabilises the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol.* 2006; **61**:151-164.

**Manjarrez-Hernandez HA, Gavilanes-Parra S, Chavez-Berrocal E, Navarro-Ocana A, and Cravioto A.** Antigen detection in enteropathogenic *Escherichia coli* using secretory immunoglobulin A antibodies isolated from human breast milk. *Infect Immun* 2000; **68**: 5030-5036.

**Marches O, Ledger TN, Boury M, Ohara M, Tu X, Goffaux F, Mainil J, Rosenshine I, Sugai M, De Rycke J, and Oswald E.** Enteropathogenic and enterohaemorrhagic *Escherichia coli* deliver a novel effector called Cif, which blocks cell cycle G2/M transition. *Mol. Microbiol.* 2003; **50**:1533-1567.

**Marchès O, Covarelli V, Dahan S, Cougoule C, Bhatta P, Frankel G, and Caron E.** EspJ of enteropathogenic and enterohaemorrhagic *Escherichia coli* inhibits opsono-phagocytosis. *Cell Microbiol.* 2008; **10**:1104-15.

**Marvolits TC, Kubori T, Lara-Tejero M, Thomas D, Unger VM, and Galan JE.** Assembly of the inner rod determines needle length in the type-III injectisome. *Nature*. 2006; **441**:637-640.

**Maurer J, Jose J, and Meyer TF.** Characterization of the essential transport function of the AIDA-I autotransporter and evidence supporting structural predictions. *J Bacteriol.* 1999; **181**:7014-20.

**McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E., Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, and Wilson RK.** Complete genome sequence of *Salmonella enterica* serovar *Typhimurium* LT2. *Nature* 2001; **413**:852-856.

**McNamara BP, Koutsouris A, O'Connell CB, Nougayrede JP, Donnenberg MS, and Hecht G.** Translocated EspF protein from Enteropathogenic *Escherichia coli* disrupts host intestinal barrier function. *J. Clin. Investig.* 2001; **107**:621-629

**Mellies JL, Elliott SJ, Sperandio V, Donnenberg MS, and Kaper JB.** The Per regulon of enteropathogenic *Escherichia coli*: identification of a regulatory cascade and a novel transcriptional activator, the locus of enterocyte effacement (LEE)-encoded regulator (Ler). *Mol Microbiol* 1999; **33**: 296–306.

**Mellies JL, Navarro-Garcia F, Okeke I, Frederickson J, Nataro JP, and Kaper JB.** espC pathogenicity island of enteropathogenic *Escherichia coli* encodes an enterotoxin. *Infect Immun.* 2001; **69**:315-24.

**Meng G, Surana NK, St Geme JW, Waksman G.** Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia timeric autotransporter. *EMBO J* 2006; **25**:2297-2304

**Meyer TH, Menetret JF, Breitling R, Miller KR, Akey CW, Rapoport TA.** The bacterial SecY/E translocation complex forms channel-like structures similar to those of the eukaryotic Sec61p complex. *J. Mol. Biol.* 1999; **285**:1789-1800.

**Miller JH.** Experiments in molecular genetics. 1972. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA

**Mogensen JE, Tapadar D, Schmidt MA, and Otzen DE.** Barriers to folding of the transmembrane domain of the *Escherichia coli* autotransporter adhesion involved in diffuse adherence. *Biochemistry* 2005; **44**:4533-4545.

**Mota LJ, Journet L, Sorg I, Agrain C, and Cornelis GR.** Bacterial injectisomes: needle length does matter. *Science*. 2005; **307**:1278.

**Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordóñez CL, Lory S, Walz T, Joachimiak A, and Mekalanos JJ.** A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* 2006; **312**: 1526-1530.

**Mougous JD, Gifford CA, Ramsdell TA, and Mekalanos JJ.** Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat. Cell. Biol.* 2007; **9**:797-803.

**Mulcahy H, O'Callaghan J, O'Grady EP, Adams C, and O'Gara F.** The post-transcriptional regulator RsmA plays a role in the interaction between *Pseudomonas aeruginosa* and human airway epithelial cells by positively regulating the type III secretion system. *Infect. Immun.* 2006; **74**:3012-3015

**Müller D, Benz I, Tapadar D, Buddenborg C, Greune L, and Schmidt MA.** Arrangement of the translocator of the autotransporter adhesin involved in diffuse adherence on the bacterial surface. *Infect Immun.* 2005; **73**:3851-9.

**Müller SA, Pozidis C, Stone R, Meesters C, Chami M, Engel A, Economou A, and Stahlberg H.** Double hexameric ring assembly of the type III protein translocase ATPase HrcN. *Mol. Microbiol.* 2006; **61**:119-125.

**Mundy R, Petrovska L, Smollett K, Simpson N, Wilson RK, Yu J, Tu X, Rosenshine I, Clare S, Dougan G, Frankel G.** Identification of a novel *Citrobacter rodentium* type III secreted protein, EspI, and roles of this and other secreted proteins in infection. *Infect Immun.* 2004; **72**:2288-2302.

**Murphy KC, and Campellone KG.** Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E.coli*. *BMC Mol Biol.* 2003; **4**:11

**Nataro JP, and Kaper JB.** Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 1998; **11**:142-201.  
**Navarro-García F, Eslava C, Villaseca JM, Lopez-Revilla R, Czeczulin JR, Srinivas S, Nataro JP, and Cravioto A.** In vitro effects of a high-molecular-weight heat-labile enterotoxin from enteroaggregative *Escherichia coli*. *Infect. Immun.* 1998; **66**:3149-3154.

**Navarro-García F, Canizalez-Roman A, Sui BQ, Nataro JP, and Azamar Y.** The serine protease motif of EspC from enteropathogenic *Escherichia coli* produces epithelial damage by a mechanism different from that of Pet toxin from enteroaggregative *E. coli*. *Infect Immun.* 2004; **72**:3609-21.

**Neves BC, Mundy R, Petrovska L, Dougan G, Knutton S, and Frankel G.** CesD2 of enteropathogenic *Escherichia coli* is a second chaperone for the type III secretion translocator protein EspD. *Infect. Immun.* 2003; **71**:2130-2141.

**Newman CL, and Stathopoulos C.** Autotransporter and two-partner secretion: delivery of large-size virulence factors by gram-negative bacterial pathogens. *Crit Rev Microbiol.* 2004; **30**:275-86.

**Nouen N, Stahlberg H, Pugsley AP, and Engel A.** Domain structure of secretin PilD revealed by limited proteolysis and electron microscopy. *EMBO J.* 2000; **19**:2229-2236.

**Ochman H, Lawrence JG, and Groisman EA.** Lateral gene transfer and the nature of bacterial innovation. *Nature*. 2000; **405**:299-304.

**Ochsner UA, Snyder A, Vasil AI, and Vasil ML.** Effects of the twin-arginine translocase on secretion of virulence factors, stress response and pathogenesis. *Proc. Natl. Acad. Sci USA*. 2002; **99**:8312-8317.

**O'Connell CB, Creasey EA, Knutton S, Elliott S, Crowther LJ, Luo W, Albert MJ, Kaper JB, Frankel G, and Donnenberg MS.** SepL, a protein required for enteropathogenic *Escherichia coli* type III translocation, interacts with secretion component SepD. *Mol Microbiol*. 2004; **52**:1613-25.

**O'Grady EP, Mulcahy H, O'Callaghan J, Adams C, and O'Gara F.** Kruppel-like factors 2 and 6 via RsmA-mediated regulation of type III exoenzymes S and Y. *Infect. Immun.* 2006; **74**:5893-5902

**Ohnishi Y, Nishiyama M, Horinouchi S, and Beppu T.** Involvement of the COOH-terminal pro-sequence of *Serratia marcescens* serine protease in the folding of the mature enzyme. *J Biol Chem* 1994; **269**:32800-32806.

**Oliver DC, Huang G, and Fernandez RC.** Identification of secretion determinants of the *Bordetella* pertussis BrkA autotransporter. *J Bacteriol* 2003a; **185**:489-495.

**Oliver DC, Huang G, Nodel E, Pleasance S, and Fernandez RC.** A conserved region within the *Bordetella* pertussis autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* 2003b; **47**: 1367-1383.

**Onufryk C, Crouch M-L, Fang F, and Gross CA.** Characterisation of Six Lipoproteins in the σE Regulon. *J Bacteriol*. 2005; **187**:4552-456

**Oomen CJ, van Ulsen P, Van Gelder P, Feijen M, Tommassen J and Gros P.** Structure of the translocator domain of a bacterial autotransporter. *EMBO J*. 2004; **23**:1257-1266.

**Orriss GL, Tarry MJ, Ize B, Sargent F, Lea SM, Palmer T, and Berks BC.** TatBC, TatB, and TatC form structurally autonomous units within the twin arginine protein transport system of *Escherichia coli*. *FEBS Lett*. 2007; **581**:4091-4097.

**Otto BR, van Dooren S, Nuijens JH and Luirink J.** Characterization of a haemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *J. Expt. Med.* 1998; **188**: 1091-1103.

**Otto BR, Sijbrandi R, Luirink J, Oudega B, Heddle JG, Mizutani K, Park SY, and Tame JR.** Crystal structure of haemoglobin protease, a haem-binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem*. 2005; **280**:17339-45.

**Paetzl M, Dalbey RE, and Strynadka NC.** Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature*. 1998; **396**:186-190.

**Pallen MJ, Chaudhuri RR, and Henderson IR.** Genomic analysis of secretion systems. *Curr Opin Microbiol*. 2003; **6**:519-27

**Pallen MJ, Beatson SA, and Bailey CM.** Bioinformatics analysis of the locus for enterocyte effacement provides novel insights into type-III secretion. *BMC Microbiol*. 2005; **5**:9

**Palmer T, and Berks BC.** Moving folded proteins across the bacterial cell membrane. *Microbiol*. 2003; **149**:547-556.

**Parreira VR, and Gyles CL.** A novel pathogenicity island integrated adjacent to the thrW tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. *Infect Immun*. 2003; **71**:5087-96.

**Parsons DA, and Heffron F.** *sciS*, an *icmF* homologue in *Salmonella enterica* serovar *Typhimurium*, limits intracellular replication and decreases virulence. *Infect. Immun.* 2005; **73**: 4338-4345.

**Patel SK, Dotson J, Allen KP, and Fleckenstein JM.** Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infect Immun*. 2004; **72**:1786-94.

**Paul K, Erhart M, Hirano T, Blair DF, and Hughes T.** Energy source of flagellar type III secretion. *Nature*. 2008; **451**:489-492.

**Perna NT, Plunkett G 3rd, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Pósfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, and Blattner FR.** Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*. 2001; **409**:529-33.

**Peterson JH, Szabady RL, and Berstein HD:** An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor and the secYEG signal. *J Biol Chem*. 2006; **281**:9038-48

**Planet PJ, Kachlany SC, DeSalle R, and Figurski DH.** Phylogeny of genes for secretion NTPases: identification of the widespread tada subfamily and development of a diagnostic key for gene classification. *Proc. Natl. Acad. Sci USA*. 2001; **98**:2503-2508.

**Pogliano JA, and Beckwith J.** SecD and SecF facilitate protein export in *Escherichia coli*. *EMBO J*. 1994; **13**:554-561.

**Pohlner J, Halter K, Beyreuther K, and Meyer TF.** Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* 1987; **325**:458-462

**Pohlner J, Langenberg U, Wolk U, Beck SC, and Meyer TF.** Uptake and nuclear transport of *Neisseria* IgA1 protease-associated alpha-proteins in human cells. *Mol Microbiol*. 1995; **17**:1073-83.

**Pohlschröder M, Prinz WA, Hartmann E, and Beckwith J.** Protein translocation in the three domains of life: variations on a theme. *Cell*. 1997; **91**:563-566.

**Pohlschröder M, Haartman E, Hand NJ, Dilks K, and Haddad A.** Diversity and evolution of protein translocation. *Annu. Rev. Microbiol*. 2005; **59**:91-111.

**Possot OM, Gérard-Vincent M, and Pugsley AP.** Membrane association and multimerization of secreton componant PulC. *J. Bacteriol*. 1999; **181**: 4004-4011.

**Possot OM, Vignon G, Bomchil N, Ebel F, and Pugsley AP.** Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *J Bact*. 2000; **182**: 2142-2152.

**Pratt LA, and Kolter R.** Genetic analyses of *Escherichia coli* biofilm formation: roles of flagella, motility and type I pili. *Mol Microbiol*. 1998; **30**: 285-293.

**Provence DL, and Curtiss<sup>3rd</sup> R.** Isolation and characterization of a gene involved in haemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect Immun*. 1994; **62**:1369–1380

**Pugsley AP, and Dupuy B.** An enzyme with type IV prepilin peptidase activity is required to process components of the general secretory pathway of *Klebsiella oxytoca*. *Mol. Microbiol*. 1992; **6**:751-760.

**Pugsley AP.** The complete general secretory pathway in Gram negative bacteria. *Microbiol Rev*. 1993; **57**:50-108.

**Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, and Mekalanos JJ.** Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictostelium* host model system. *Proc. Natl. Acad. Sci. USA*. 2006; **103**: 1528-1533.

**Pukatzki S, Ma AT, Revel AT, Sturtevant D, and Mekalanos JJ.** Type VI secretion translocates a phage like protein into target cells where it cross links actin. *Proc. Natl. Acad. Sci. USA*. 2007; **104**: 15508-15513.

**Pupo GM, Karadis DKR, Lan R, and Reeves PR.** Recent Evolution of EPEC from Commensals. *Infect. Imm*. 1997; **64**: 2685-2692.

**Purdy GE, Hong M, and Payne SM.** *Shigella flexneri* DegP facilitates IcsA surface expression and is required for efficient intercellular spread. *Infect Immun* 2002; **70**: 6355–6364.

**Py B, Loiseau L, and Barras F.** An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Rep.* 2001; **2**:244-248.

**Rain JC, Selig L, De Reuse H, Battaglia V, Reverdy C, Simon S, Lenzen G, Petel F, Wojcik J, Schächter V, Chemama Y, Labigne A, and Legrain P.** The protein-protein interaction map of *Helicobacter pylori*. *Nature*. 2001; **409**:211-5

**Rao PS, Yamada Y, Tan YP, and Leung KY.** Use of proteomics to identify novel virulence determinants that are required for *Edwardsiella tarda* pathogenesis. *Mol. Micro*. 2004; **53**:573-586.

**Ravaud S, Wild K, and Sinning I.** Purification, crystallization and preliminary structural characterization of the periplasmic domain P1 of the *Escherichia coli* membrane-protein insertase YidC. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2008; **64**:144-8

**Ren CP, Chaudhuri RR, Fivian A, Bailey CM, Antonio M, Barnes WM, and Pallen MJ.** The ETT2 gene cluster, encoding a second type III secretion system from *Escherichia coli*, is present in the majority of strains but has undergone widespread mutational attrition. *J Bacteriol*. 2004; **186**:3547-60.

**Reyrat JM, Charrel M, Pagliaccia C, Burroni D, Lupetti P, de Bernard M, Ji X, Norais N, Papini E, Dallai R, Rappuoli R, and Telford JL.** Characterisation of a monoclonal antibody and its use to purify the cytotoxin of *Helicobacter pylori*. *FEMS Microbiol Lett*. 1998; **165**:79-84.

**Rizzitello AE, Harper JR, and Silhavy TJ.** Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J. Bacteriol*. 2001; **183**: 6794–6800.

**Robert V, Volokhina EB, Senf F, Bos MP, van Gelder P, and Tomassen J.** Assembly factor Omp85 recognises its outer membrane protein substrates by a species specific C-terminal motif. *PLOS Biol*. 2006; **4**:e377. doi: 10.1371/journal.pbiol.0030377

**Roggenkamp A, Ackermann N, Jacobi CA, Truelzsch K, Hoffmann H, and Heesemann J.** Molecular analysis of transport and oligomerization of the *Yersinia enterocolitica* adhesin YadA. *J Bacteriol*. 2003; **185**:3735-3744.

**Ruby EG, Urbanowski M, Campbell J, Dunn A, Faini M, Gunsalus R, Lostroh P, Lupp C, McCann J, Millikan D, Schaefer A, Stabb E, Stevens A, Visick K, Whistler C, and Greenberg EP.** Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium with pathogenic congeners. *Proc Natl Acad Sci U S A*. 2005; **102**:3004-3009.

**Rutherford N, and Mourez M.** Surface Display of proteins by Gram-negative bacterial autotransporters. *Microbial Cell Factories* 2006; **5**:22

**Sambrook J, Fritsch EF, and Maniatis T.** Molecular Cloning: A Laboratory Manual. 1989. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA

**Saier MH.** Protein secretion and membrane insertion systems in gram-negative bacteria. *J Membr Biol*. 2006; **214**:75-90.

**Saiki R, Gelfand D, Stoffel S, Scharf S, Higuchi R, Horn G, Mullis K, and Erlich H.** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988; **239**:487-91.

**Sanchez-Pulido L, Devos D, Genevrois S, Vicente M, and Valencia A.** POTRA: a conserved domain in the FtsQ family and a class of β-barrel outer membrane proteins. *Trends Biochem. Sci.* 2003; **28**: 523–526.

**Sandkvist M, Hough LP, Bagdasarian MM, and Bagdasarian M.** Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus of *Vibrio cholerae*. *J Bact*. 1999; **181**:3129-3135.

**Sandkvist M.** Biology of type II secretion. *Mol Microbiol*. 2001; **40**:271-283.

**Sandt CH, and Hill CW.** Four different genes responsible for non-immune immunoglobulin-binding activities within a single strain of *Escherichia coli*. *Infect Immun*. 2000; **68**:2205-2214.

**Sandt CH, and Hill CW.** Non-immune binding of human immunoglobulin A (IgA) and IgG Fc by distinct sequences segments of EibF cell surface protein of *Escherichia coli*. *Infect Immun*. 2001; **69**:7293-7303.

**Sargent F, Gohike U, de Leeuw E, Stanley NR, Palmer T, Saibil HR, and Berks BC.** Purified components of the Tat protein transport system of *Escherichia coli* form a double layered ring structure. *Eur J Biochem*. 2001; **268**:3361-3367.

**Schoenhen IC, Stratilo C, and Howard SP.** An ExeAB complex in the type II secretion pathway of *Aeromonas hydrophila*: effect of ATP binding cassette mutations on complex formation and function. *Mol Microbiol*. 1998; **29**:1237-1247.

**Sekiya K, Ohishi M, Oginio T, Tamano K, Sasakawa C, and Abe A.** Supermolecular structure of the Enteropathogenic *Escherichia coli* type III secretion system and its direct interaction with the EspA sheath like structure. *Proc. Natl. Acad. Sci. USA*. 2001; **98**:11638-11643.

**Settles AM, Yonetani A, Baron A, Bush DR, Cline K, and Martienssen R.** Sec-independent protein translocation by the maize Hcf106 protein. *Science*. 1997; **278**:1467-1470.

**Shrivastava S, and Mande SS.** Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. *PLoS ONE*. 2008; 3:e2955

**Sijbrandi R, Urbanus ML, ten Hagen-Jongman CM, Berstein HD, Oudega B, Otto BR, and Luijink J.** Signal recognition particle (SRP) - mediated targeting and sec-dependant translocation of an extracellular *Escherichia coli* protein. *J.Biol.Chem*. 2003; **278**:4654-4659.

**Skillman KM, Barnard TJ, Peterson JH, Ghirlando R, and Bernstein HD.** Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* 2005; **58**:945-958.

**Sklar JG, Wu T, Gronenberg LS, Malinverni JC, Kahne D, and Silhavy TJ.** Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*. 2007; **104**:6400-6405.

**Skorko-Glonek J, Laskowska E, Sobiecka-Szkatula A, and Lipinska B.** Characterisation of the chaperone-like activity of HtrA (DegP) protein from *Escherichia coli* under the conditions of heat shock. *Arch Biochem Biophys*. 2007; **464**:80-9.

**Soscia C, Hachani A, Bernadac A, Filloux A, Bleves S.** Cross talk between type III secretion and flagellar assembly systems in *Pseudomonas aeruginosa*. *J Bacteriol*. 2007; **189**:3124-32.

**Stagljar I, Korotenny N, Johnsson N, and te Heesen S.** A genetic system based on split-ubiquitin for the analysis of interactions between membrane and normal proteins in vivo. *Proc. Natl. Acad. Sci. USA*. 1998; **95**:5187-5192.

**Stanley P, Koronakis V, and Hughes C.** Mutational analysis supports a role for multiple structural features in the C- terminal secretion signal of *Escherichia coli* haemolysin. *Mol Microbiol*. 1991; **5**:2391-2403.

**Stanley NR, Palmer T, and Berks BC.** The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem*. 2000; **275**:11591-6.

**Stathopoulos C, Provence DL, and Curtiss<sup>3rd</sup>, R.** Characterization of the avian pathogenic *Escherichia coli* hemagglutinin Tsh, a member of the immunoglobulin A protease-type family of autotransporters. *Infect. Immun*. 1999; **67**:772-781.

**Stegmeier JF, and Andersen C.** Characterisation of Pores Formed by YaeT (Omp85) from *Escherichia coli*. *J. Biochem.* 2006; **140**:275:283.

**Stein M, Kenny B, Stein MA, and Finlay BB.** Characterization of EspC, a 110-kilodalton protein secreted by enteropathogenic *Escherichia coli* which is homologous to members of the immunoglobulin A protease-like family of secreted proteins. *J. Bacteriol.* 1996; **178**:6546-6554

**St Geme<sup>3rd</sup> JW, de la Morena ML, and Falkow S.** A *Haemophilus influenzae* IgA protease-like protein promotes intimate interaction with human epithelial cells. *Mol Microbiol.* 1994; **14**:217-33.

**Surana NK, Cutter D, Barenkamp SJ, St Geme<sup>3rd</sup> JW.** The *Haemophilus influenzae* Hia autotransporter contains an unusual short trimeric translocator domain. *J Biol Chem* 2004; **279**: 14679-14685.

**Szabady RL, Peterson JH, Skillman KM, and Bernstein HD.** From The Cover: An unusual signal peptide facilitates late steps in the biogenesis of a bacterial autotransporter. *Proc. Natl Acad Sci USA.* 2005; **102**:221-6.

**Tacket CO, Stein MB, Losonsky G, Abe A, Finlay BB, McNamara BP, Fannery GT, James SP, Nataro JP, Levine MM, and Donnenberg MS.** Role of EspB in experimental human enteropathogenic *Escherichia coli* infection. *Infect Immun* 2000; **68**: 3689-3695.

**Taylor KA, O'Connell CB, Luther PW, Donnenberg MS.** The EspB protein of enteropathogenic *Escherichia coli* is targeted to the cytoplasm of infected HeLa cells. *Infect. Immun.* 1998; **66**: 5501-5507.

**Terradot L, Durnell N, Li M, Li M, Ory J, Labigne A, Legrain P, Colland F, and Waksman G.** Biochemical characterization of protein complexes from the *Helicobacter pylori* protein interaction map: strategies for complex formation and evidence for novel interactions within type IV secretion systems. *Mol Cell Proteomics.* 2004; **3**:809-19.

**Thanabalu T, Koronakis E, Hughes C, and Koronakis V.** Substrate induced assembly of a contiguous channel for protein export from *Escherichia coli* reversible bridging of an inner membrane to outer membrane exit pole. *EMBO J.* 1998; **17**:6487-6496.

**Thanassi DG, Saulino ET, and Hultgren SJ.** The chaperone/usher pathway: a major terminal branch of the general secretory pathway. *Curr Opin Microbiol.* 1998; **1**:223-31.

**Thanassi DG, and Hultgren SJ.** Multiple pathways allow protein secretion across the bacterial cell membrane. *Curr Opin Cell Biol.* 2000; **12**: 420-430.

**Thanassi DG, Stathopoulos C, Dodson K, Geiger D, and Hultgren SJ.** Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. *J Bacteriol.* 2002; **184**:6260-62669.

**Thanassi DG, Stathopoulos C, Karkal A, and Li H.** Protein secretion in the absence of ATP: the autotransporter, two-partner secretion and chaperone/usher pathways of gram-negative bacteria. *Mol Membr Biol.* 2005; **22**:63-72

**Thoden JB, Zhuang Z, Dunaway-Mariano D, and Holden HM.** The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU. *J. Biol. Chem.* 2003; **278**:43709-43716.

**Tobe T, and Sasakawa C.** Role of bundle-forming pilus of enteropathogenic *Escherichia coli* in host cell adherence and in microcolony development. *Cell Microbiol.* 2001; **3**:579-85.

**Tobe T, Beatson SA, Taniguchi H, Abe H, Bailey CM, Fivian A, Younis R, Matthews S, Marches O, Frankel G, Hayashi T, and Pallen MJ.** An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci U S A.* 2006; **103**:14941-14946.

**Tomson FL, Viswanathan VK, Kanack KJ, Kanteti RP, Straub KV, Menet M, Kaper JB, and Hecht G.** Enteropathogenic *Escherichia coli* EspG disrupts microtubules and in conjunction with Orf3 enhances perturbation of the tight junction barrier. *Mol. Microbiol.* 2005; **56**: 447-464.

**Tu X, Nisan I, Yona C, Hanski E, and Rosenshine I.** EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic *Escherichia coli*. *Mol Microbiol.* 2003; **47**:595-606.

**Uchiya KI, Tobe T, Komatsu K, Suzuki T, Watarai M, Fukuda I, Yoshikawa M, and Sasakawa C.** Identification of a novel virulence gene, virA, on the large plasmid of *Shigella*, involved in invasion and intercellular spreading. *Mol. Microbiol.* 1995; **17**:241-250

**Uphoff TS, and Welch RA.** Nucleotide sequencing of the *Proteus mirabilis* calcium-independent haemolysin genes (hpmA and hpmB) reveals sequence homology with the *Serratia marcescens* haemolysin genes (shlA and shlB). *J Bacteriol.* 1990; **172**:1206-1216.

**Valent QA, Scotti PA, High S, de Gier J-WL, von Heijne G, Lentzen, G, Wintermeyer W, Oudega B, and Luijink J.** The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J.* 1998; **17**:2504-2512.

**Vallance BA, and Finlay BB.** Exploitation of host cells by enteropathogenic *Escherichia coli* *Proc. Natl. Acad. Sci. USA.* 2000; **97**:8799-8806

**Vallance BA, Khan MA, Deng W, Gruenheid S, and Finlay BB.** Modelling enteropathogenic and enterohaemorrhagic *E.coli* infections and disease. *Drug Discovery Today: Disease Models.* 2004; **1**:1-7.

**van der Laan M, Bechtuft P, Kol S, Nouven N, and Dreissen AJ.** F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J.Cell. Biol.* 2004; **165**:213-222.

**van der Laan M, Nouwen NP, and Driessens AJ.** YidC--an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr Opin Microbiol.* 2005; **8**:182-7.

**van Klompenburg W, and de Kruijff B.** The role of anionic lipids in protein insertion and translocation in bacterial membranes. *J Membr. Biol.* 1998; **162**:1-7.

**Veiga E, de Lorenzo V, and Fernandez LA.** Probing secretion and translocation of a  $\beta$ -autotransporter using a reporter single-chain Fv as a cognate passenger domain. *Mol Microbiol* 1999; **33**: 1232-1243

**Veiga E, Sugawara E, Nikaido H, de Lorenzo V, and Fernandez LA.** Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. *EMBO J* 2002; **21**: 2122-2131.

**Veiga E, de Lorenzo V, and Fernández LA.** Structural tolerance of bacterial autotransporters for folded passenger protein domains. *Mol. Micro.* 2004; **52**:1069-1080.

**Velarde JJ, and Nataro JP.** Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger. *J Biol Chem* 2004; **279**:31495-31504.

**Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, and Hooykaas PJ.** VirB/D4 – dependent protein translocation from *Agrobacterium* into plant cells. *Science.* 2000; **290**:979-982.

**Vergunst AC, van Lier MC, den Dulk-Ras A, Grosse Sturze TA, Ouwehand A, and Hooykaas PJ.** Positive charge is an important feature of the C-terminal transport signal of the VirB/D4 – translocated proteins of *Agrobacterium*. *Proc. Natl. Acad. Sci. USA.* 2005; **102**: 8323-837.

**Vidal JE, and Navarro-García F.** Efficient translocation of EspC into epithelial cells depends on Enteropathogenic *Escherichia coli* and host cell contact. *Infect. Immun.* 2006; **74**:2293-2303.

**Vidal JE, and Navarro-García F.** EspC translocation into epithelial cells by enteropathogenic *Escherichia coli* requires a concerted participation of type V and III secretion systems. *Cell Microbiol.* 2008; **10**:1975-86.

**Vogel JP, and Isberg RR.** Cell Biology of *Legionella pneumophila*. *Curr. Opin. Microbiol.* 1999; **2**:30-34.

**Voulhoux R, Ball G, Ize B, Vasil ML, Lazdunski A, Wu LF, and Filloux A.** Involvement of the twin-arginine translocation system in protein secretion via the type II pathway, *EMBO J.* 2001; **20**:6735-6741.

**Voulhoux R, Bos MP, Geurtzen J, Mols M, and Tommassen J.** Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 2003; **299**: 262-265

**Voulhoux R, and Tommassen J.** Omp85, an evolutionary conserved bacterial protein involved in the outer-membrane-protein assembly. *Res. Microb.* 2004; **155**: 129-135.

**Wachter C, Beinke C., Matthes M, and Schmidt MA.** Insertion of EspD into epithelial target cell membranes by infecting enteropathogenic *Escherichia coli*. *Mol Microbiol* 1999; **31**:1695-1707.

**Walburger A, Lazdunski C, and Corda Y.** "The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB." *Mol Microbiol* 2002; **44**:695-708. PMID: 11994151

**Wainwright LA, and Kaper JB.** EspB and EspD require a specific chaperone for proper secretion from enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 1998; **27**:1247-1260

**Wandersman C, and Delepelaire P.** TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad USA* 1990; **87**:4776-4780.

**Warawa J, Finlay BB, and Kenny B.** Type III secretion-dependent hemolytic activity of enteropathogenic *Escherichia coli*. *Infect Immun* 1999; **67**: 5538-5540.

**Ward DV, Draper O, Zupan JR, and Zambryski PC.** Peptide linkage mapping of the *Agrobacterium tumefaciens vir*-encoded type IV secretion systems reveals protein sub-assemblies. *Proc. Natl. Acad. Sci. USA.* 2002; **99**:11493-11500.

**Wattiau P, and Cornelis GR.** SycE, a chaperone like protein of *Yersinia enterocolitica* involved in the secretion of YopE. *Mol. Microbiol.* 1993; **8**:123-131.

**Welch RA, Burland V, Plunkett G 3rd, Redford P, Roesch P, Rasko D, Buckles EL, Liou SR, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HL, Donnenberg MS, and Blattner FR.** Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2002; **99**:17020-17024.

**Werner J, and Misra R.** YaeT (Omp85) affects the assembly of lipid-dependent proteins and lipid independent outer membrane proteins of *Escherichia coli*. *Mol. Microbiol* 2005; **67**:1450-1459.

**West NP, Sansonetti P, Mounier J, Exley RM, Parsot C, Guadagnini S, Prévost MC, Prochnicka-Chalufour A, Delepierre M, Tanguy M, and Tang CM.** Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science*. 2005; **307**:1313-7

**Wexler M, Bogsch EG, Klösgen RB, Palmer T, Robinson C, and Berks BC.** Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the delta pH pathway. *FEBS Lett.* 1998; **431**:339-42.

**Whittam TS, Wolfe ML, Wachsmuth IK, Orskov F, Orskov I, and Wilson RA.** Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect Immun.* 1993; **61**:1619-29.

**Wilkins MR, and Williams KL.** Cross-species identification using amino acid composition : a theoretical evaluation. *J. Theor. Biol.* 1997; **186**:7-15.

**Wolff C, Nisan I, Hanski E, Frankel G, and Rosenshine I.** Protein translocation into host epithelial cells by infecting enteropathogenic *Escherichia coli*. *Mol Microbiol.* 1998; **28**:143-55

**Wolff N, Sapriel G, Bodenreider C, Chaffotte A, and Delepe laire P.** Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate. *J Biol Chem.* 2003; **278**:38247-53.

**Wolin SL.** From the elephant to *E. coli*: SRP-dependent protein targeting. *Cell.* 1994; **77**:787-790.

**Wu L-F, Ize B, Chanal A, Quentin Y, and Fichant G.** Bacterial twin arginine signal peptide-dependant protein translocation pathway: evolution and mechanism. *J. Mol Microbiol Technol.* 2000; **2**:179-189.

**Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, and Kahne D.** Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell.* 2005; **121**:235-245.

**Yahr TL, and Wickner WT.** Functional reconstitution of bacterial Tat translocation *in vitro*. *EMBO J.* 2001; **20**:2471-2479.

**Yen MR, Peabody CR, Partovi SM, Zhai Y, Tseng YH, and Saier MH.** Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim Biophys Acta.* 2002; **1562**:6-31

**Yeo HJ, Cotter SE, Laarmann S, Juehne T, St Geme JW, and Waksman G:** Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. *EMBO J.* 2004; **23**: 1245-1256.

**Yi L, Jiang F, Chen M, Cain B, Bolhuis A, and Dalbey RE.** YidC is strictly required for membrane insertion of subunits a and c of the F(1)F(0)ATP synthase and SecE of the SecYEG translocase. *Biochemistry.* 2003; **42**:10537-10544.

**Yip CK, and Strynadka NCJ.** New structural insights into the bacterial type III secretion system. *Trends Biochem. Sci.* 2006; **31**:223-230.

**Zhuang Z, Song F, Martin BM, and Dunaway-Mariano D.** "The YbgC protein encoded by the ybgC gene of the tol-pal gene cluster of *Haemophilus influenzae* catalyzes acyl-coenzyme A thioester hydrolysis." *FEBS Lett* 2002; **516**:161-3.

**Zupan J, Hackworth CA, Arguilar JA, Ward D, and Zambryski P.** VirB1\* promotes T-pilus formation in the vir-type IV secretion system of *Agrobacterium tumefaciens*. *J Bacteriol.* 2007; **189**:6551-63.