UNIVERSITY OF LIVERPOOL

Characterising the adsorption process of the shigatoxigenic bacteriophage, $\phi 24_B$, to its bacterial host via an essential outer-membrane protein, BamA.

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Stuart Alexander David McEwen

March 2018

Table of Contents

Table of Figures

Table of Tables

List of Units and Abbreviations

Symbol	Unit
µ-	micro-
Da	Dalton
g	gram
x *g*	times the force of gravity
L	litre
k-	kilo-
m-	milli-
bp	base pair
U	Units
s	seconds

Abbr.	Abbreviation
AFM	atomic force microscopy
BAM complex	β-barrel assembly machinery complex
BAM model	bacteriophage adhering to mucus model
BSA	bovine serum albumin
CFU	colony forming units
CTD	carboxy terminal domain
DTT	dithiothreitol
EAEC	Enteroaggregative *E. coli*
EAHEC	Enteroaggregative haemorrhagic *E. coli*
EDTA	ethylenediaminetetraacetic acid
EEKS	*Escherichia*, *Erwinia*, Kanamycin, Skp
eGFP	engineered green fluorescent protein
EHEC	Enterohaemorrhagic *E. coli*
EIEC	Enteroinvasive *E. coli*
EPEC	Enteropathogenic *E. coli*
ER	endoplasmic reticulum

ETEC	Enterotoxigenic *E. coli*
Gb_3	globotriaosylceramide
gDNA	genomic DNA
GFP	green fluorescent protein
HGT	horizontal gene transfer
HR	homologous recombination
HUS	haemolytic uraemic syndrome
IM	inner membrane
LEE	locus of enterocyte effacement
MCS	multiple cloning site
MIC	minimum inhibitory concentration
MOI	multiplicity of infection: the ratio of viral particles to potential host cells
MRL	multiple resistance lysogen
MWCO	molecular weight cut off
NTD	amino terminal domain
OM	outer membrane
OMP	outer membrane protein
ORF	open reading frame
PCR	polymerase chain reaction
PFU	plaque forming units
PGM	phage growth medium
RPM	revolutions per minute
STEC (VTEC)	shigatoxigenic *E. coli* (verocytotoxic *E. coli*)
Stx	Shiga toxin
TSP	tail spike protein
TTP	thrombotic thrombocytopaenic purpura
TTSS	type III secretion system
UTR	untranslated region
vpr	verocytotoxin phage recptor (Also referred to as YaeT and BamA)
WT	wild type
λ	lambda
ϕ	bacteriophage

Acknowledgements

The first thank you must go to my PhD supervisors, Heather Allison and Alan McCarthy, for unwavering patience, guidance, faith, and encouragement - I know that I would not have been able to produce this thesis without them. Thank you to the Phage Group of Lab H: Marta Veses Garcia, Mohammed Mohaisen, and Sian Owen, for continuous discussion, troubleshooting, and sharing of success stories! Thank you to Rachel for 4 years of every kind of support possible, and to my family who were always there for me, especially when I needed them most. Of course, the rest of Lab H over the years have been a fantastic bunch of people to work with and/or around, with one exception – they know who they are…

Some people deserve special mentions for the overwhelmingly positive impact they have had on my life in Liverpool:
The Quizteama: Heather, Liam, Jack, and Gabi;
The rest of the Base Room group: Amy, Anshul, Sean, Breno, Ewan, Rhys, and Ric;
The Commune: Alan, Jenny, Jay, Marina, and Tash;
And the Kids: Freddie, Iisakki, and Emily.

For their help with technical aspects of the study; thanks go to Alison Beckett and Alan Roseman for training and support for the TEM work at Universities of Liverpool and Manchester respectively; Marco Marcello and Dave Mason for training and assistance on fluorescence microscopes and finally to Steve Hooton for leading the troubleshooting of the PFGE equipment.

To do science we stand on the shoulders of those who have worked before us: previously, the following were responsible for bringing the project to the exciting stage where I inherited it: Martin Sergeant, Chloe James, Darren Smith, Sybille Schmitter, Shaun Calleja, Jordan Carrigan. Undergraduate students who helped(?) during my time of the study include: Siyuan Zhang, Amy Preistman, Oliver Fletcher, Alexander Poole, and Anna Seager.

"We're still flying. That's not much. [But] it's enough"

For Dorothy and Doreen

They may not see me graduate, but they never doubted that I would.

Abstract

Escherichia coli is a common commensal organism that is present in all humans digestive tracts. When *E. coli* is infected by a bacteriophage containing Shiga toxin genes (Stx phage), the *E. coli* cell is converted into a pathogen that can cause serious harm to humans in the form of Haemolytic Uraemic Syndrome (HUS) or Thrombotic, Thrombocytopaenic Purpura (TTP). It has been shown that ~70% of environmental Stx phages possess a tail that recognises an essential protein in the outer membrane of *E. coli*, BamA.

This study presents the final proof required to demonstrate that BamA is the target of these short-tailed Stx phages and intends to dissect which epitopes of extracellular surface of BamA support phage adsorption by comparing *E. coli* to a phage resistant species, *Pectobacterium atrosepticum*.

This resulted in the production of a library of 32 BamA mutants, which could be compared in an adsorption assay. The physiological effects of these mutants on the adsorption process were then examined in a fluorescence microscopy-based adsorption assay that allows the visualisation of temporal resolution of these processes.

Given that BamA is conserved across all *Enterobacteriaceae*. There is potential for host range expansion, thereby converting more species that are currently harmless into potential human pathogens. The information obtained from this study may help in predicting which species is the next likely to become Stx phage susceptible.

CHAPTER 1: Introduction

1.1 *Escherichia coli*

Escherichia coli is one of the most thoroughly studied organisms on Earth, leading to its current status of "biological rock star" (Blount, 2015). *E. coli* is used in a multitude of applications from being a model organism through to molecular biology, genetic engineering, and being a host for protein overexpression (Blount, 2015).

There are an estimated total of 10^{21} *E. coli* cells in the collective gut of the human population (Conway & Cohen, 2015), the majority of which are harmless and are a lifelong member of the normal human gut flora from just after birth (Palmer *et al.*, 2007). However, there are strains of *E. coli* that have potential to cause disease in humans (Kaper *et al.*, 2004).

1.2 Pathogenic *Escherichia coli*

Every year there are both isolated cases and outbreaks of *E. coli* mediated disease across the globe (Majowicz *et al.*, 2014). *E. coli* can cause a range of symptoms from mild diarrhoea to urinary tract infections (Flores-Mireles *et al.*, 2015) and sepsis or meningitis (Croxen & Finlay, 2009). The array of infections and syndromes caused by all the forms of *E. coli* are beyond the scope of this thesis, which will focus on diarrhoegenic *E. coli* strains. The general infection

strategy for these pathogenic *E. coli* is to colonise a mucosal site somewhere along the gastrointestinal tract, where it evades host defences and multiplies before ultimately distressing the host (Nataro & Kaper, 1998). Among the subset of diarrhoegenic *E. coli*, symptoms can range from mild to potentially life threatening, depending on which strain has caused the infection.

E. coli has been linked with clinical cases of diarrhoea since the 1940s (Bray & Beavan, 1948), but the first confirmed association between *E. coli* and life-threatening symptoms was not made for another 4 decades (Karmali *et al.*, 1983). *E. coli* strains associated with disease have been classified into serotypes. Serotypes are assigned to strains of *E. coli* based on the variety and combination of O (somatic), H (flagellar), and K (capsular) surface antigens that are presented on the cell surface (Lior, 1996). Historically, the serotyping of *E. coli* strains was critical to understanding their virulence, and this classification system is still used today. Although these surface antigens do not confer virulence to strains directly, they reliably correlate with strains that cause specific clinical symptoms (Whittam *et al.*, 1993). There are currently over 500 different serotypes of *E. coli* identified (Allison, 2007) from all the possible combinations of 173 O antigens, 80 K antigens, and 56 H antigens (Orskov & Orskov, 1992; Tenaillon *et al.*, 2010).

Within the umbrella term of diarrhoegenic *E. coli*, there are different classifications based on the mechanism of pathogenicity. For the purpose of this thesis, these subtypes have been separated into those that produce the potentially life-threatening Shiga toxin and those that do not. The Shiga toxin producing *E. coli* (STEC) are the two supervirulent, human-adapted subgroups

the Enterohaemorrhagic *E. coli* (EHEC) and Enteroaggregative haemorrhagic *E. coli* (EAHEC). The subtypes that do not produce Shiga toxin are Enteropathogenic *E. coli* (EPEC), Enterotoxigenic *E. coli* (ETEC), Enteroinvasive *E. coli* (EIEC), and Enteroaggregative *E. coli* (EAEC).

1.2.1 Non-shigatoxigenic diarrheogenic *E. coli*

1.2.1.1 Enteropathogenic *E. coli* (EPEC)

EPEC have been a known causative agent of diarrhoea since the 1940s and remains one in present day developing countries (Frankel *et al.*, 1998; Kotloff *et al.*, 2013; Levine, 1987). EPEC strains are defined by pathogenic mechanics that involve intimate binding with the host intestinal mucosa. EPEC cells do this by forming "attaching and effacing" (A/E) lesions, which are pedestal structures on the surface of the cell caused by remodelling of the cytoskeleton of cells in the intestinal wall, destroying the microvilli structures in the process (Moon *et al.*, 1983) (Polotsky *et al.*, 1977; Wong *et al.*, 2011). Genes that confer this ability on EPEC strains are encoded on a pathogenicity island called the **l**ocus of **e**nterocyte **e**ffacement (LEE), which encodes a type III secretion system (TTSS), secreted proteins, and an adhesin called intimin (Hartland *et al.*, 1999). Translocated intimin receptor, tir, is a secreted protein encoded by *espE* in the LEE, which is translocated into the eukaryotic cell using the TTSS, which is then recognised by intimin on the bacterial surface (Batchelor *et al.*, 2000). Once these proteins bind, the bacterial cell is intimately attached to the intestinal mucosa where it is able to evade phagocytosis (Celli *et al.*, 2001), and cause diarrhoea and localised

inflammation (Ruchaud-Sparagano *et al.*, 2007). Most cases of EPEC are self-limiting and can be successfully treated by drinking lots of fluids.

1.2.1.2 Enterotoxigenic *E. coli* (ETEC)

ETEC are a major cause of infant and traveller's diarrhoea in the developing world (Black *et al.*, 1981; Sack *et al.*, 1978). ETEC are, unsurprisingly, defined by their secretion of toxins. After colonisation of the small-bowel, ETEC secrete either a heat liable (LT) toxin, a heat stable (ST) toxin, or both (Jobling & Holmes, 2012). These toxins are usually plasmid-encoded (Shepard *et al.*, 2012), but act in different ways. LT is an 86 kDa AB_5 toxin, highly similar to the cholera toxin (Dickinson & Clements, 1995), which enhances adherence of bacterial cells to the mammalian intestinal wall and disrupts regulation of adenylate cyclase in the host enterocytes. ST is a small 2 kDa peptide that activates guanylyl cyclase C in the host enterocytes to increase the intracellular concentration of cGMP. This leads to chloride and carbonate ion secretion, which draws water into the lumen of the intestines by osmosis, leading to watery diarrhoea (Taxt *et al.*, 2010).

1.2.1.3 Enteroinvasive *E. coli* (EIEC)

First characterised in 1971, EIEC have a similar pathogenicity mechanism to *Shigella spp.*, whereby the cell enters the colonic epithelial cell where it is able to replicate and eventually kill that host cell (DuPont *et al.*, 1971). EIEC strains harbour a huge 140 MDa plasmid that encodes factors required for invasion of the host epithelium(Hale *et al.*, 1983). EIEC strains impair water and nutrient uptake in the colon to cause symptoms ranging from watery diarrhoea to severe

dysentery, abdominal cramps and stools containing blood and mucus (Schroeder & Hilbi, 2008).

1.2.1.4 Enteroaggregative *E. coli* (EAEC)

EAEC cells adhere to both Hep2 cells and one another in the intestinal lumen (Nataro *et al.*, 1987). EAEC are not associated with invasive characteristics or intimate attachment (Mathewson *et al.*, 1985; 1986), but they do possess a 110 kb plasmid that encodes aggregative adherence fimbrae I-IV (AAF/I-IV), and 3 toxins: Pet, EAST1, and ShET2 (Chaudhuri *et al.*, 2010). The AAF/I-IV factors allow the colonisation of the intestinal mucosa by EAEC where they form a biofilm. The toxins are then released from EAEC causing an inflammatory response and intestinal fluid secretion leading to diarrhoea (Arenas-Hernández *et al.*, 2012).

1.2.2 Shigatoxigenic *E. coli* (STEC)

Shigatoxigenic *E. coli* (STEC) are food-borne bacterial pathogens capable of producing a toxin similar to the one produced by *Shigella dysenteriae* type I. STEC are a significant ongoing health risk in both developed and developing countries (Abbasi *et al.*, 2014; Jenkins *et al.*, 2015). The defining characteristic of STEC is that they produce Shiga toxin, Stx. In humans, Stx causes adverse symptoms including vomiting and diarrhoea, and in disease cases where Stx is produced, 5%-25% of the infected patients from outbreaks go on to develop systemic conditions, such as Haemolytic Uraemic Syndrome (HUS) and Thrombotic Thrombocytopaenic Purpura (TTP), which lead to lifelong morbidity

including hypertension and renal insufficiency (Ethelberg *et al.*, 2004; Griffin *et al.*, 1988; Nataro & Kaper, 1998; Steele *et al.*, 1982). The serotype O157:H7 has the strongest correlation with incidents of HUS worldwide (Tarr *et al.*, 2005).

The genes encoding Shiga toxins, *stx*, are disseminated between host cells by horizontal gene transfer (HGT) via bacterial viruses (bacteriophages) (Khalil *et al.*, 2016). There are examples of *stx* genes in bacterial genome that are transferred vertically, but they form part of phage genomes integrated into the STEC genome (Zhou *et al.*, 2015). When this occurs that phage genome is known as a "prophage" and the host cell has become a "lysogen". The collective term for bacteriophages containing *stx* gene is Stx phages.

STEC have also historically been referred to as VTEC (verocytotoxic *E. coli*) due to the lethal effect the toxin has on cultured Vero cells - a kidney cell line from the African green monkey (Konowalchuk *et al.*, 1977). It was later determined this cytotoxin was identical to the Shiga toxin produced by *E. coli*. After a phase where both were being used in the literature, STEC is now the favoured terminology.

1.2.2.1 Enterohaemorrhagic *E. coli* (EHEC)

EHEC are a subset of STEC that share some of the pathogenicity mechanisms of EPEC in their ability to intimately attach to the intestine epithelium via intimin-tir binding and generate A/E lesions. Despite this similarity, there are 2 key differences that define them: EHEC cells colonise further down the GI tract than EPEC; in the distal ileum or the colon, rather than the small intestine (Griffin *et*

al., 1988; Kelly *et al.*, 1987). EHEC strains also produce Shiga toxin (Stx) which at various times as also known as Shiga-like toxin (SLT) or verocytotoxin (VT).

The first EHEC strain identified was part of an outbreak in the USA that originated from undercooked beef in hamburgers (Riley *et al.*, 1983). It was soon identified that this was due to an O157:H7 strain that had been infected by a Stx phage (O'Brien *et al.*, 1984). *Stx* genes have also spread, via phages, to other members of the *Enterobacteriaceae* family, (Paton & Paton, 1996; Strauch *et al.*, 2001) and even as far as *Acinetobacter haemolyticus* (Grotiuz *et al.*, 2006), crossing into another order – an astonishingly large phylogenetic distance given that phage host ranges are usually limited to a specific species, occasionally even strains of a particular species (Koskella & Meaden, 2013). The continued occurrences of STEC outbreaks up to the present day and emergence of new host species indicates the rapid dissemination of *stx* by phages is being observed in real time, and the next severe outbreak is probably inevitable.

All EHEC strains are STEC and so recently, the acronyms EHEC and STEC have been used interchangeably in the literature e.g. (Yang *et al.*, 2017). However, this is incorrect as not all STEC strains are EHEC (Nataro & Kaper, 1998). Some STEC isolates e.g. from cattle (Callaway *et al.*, 2013), may not be associated with human disease and others are enteroaggregative haemorrhagic *E. coli*.

1.2.2.2 Enteroaggregative haemorrhagic *E. coli* (EAHEC)

The largest outbreak of Shiga toxin-producing *E. coli* in recent history was in Germany over the summer of 2011. This was caused by a strain of

enteroaggregative haemorrhagic *E. coli* (EAHEC) evolved from EAEC after being infected by an Stx phage (Beutin & Martin, 2012). This EAHEC strain combined the shigatoxigenic potential of EHEC with the self-aggregative characteristics of EAEC. This combination of virulence factors leads to a larger number of *E. coli* cells per colony than that associated with EHEC infections, and when combined with Stx phage carriage, more cells expressing Shiga toxin will potentially increase the overall amount of Shiga toxin release. Increasing numbers of colonising cell and increased Stx production may explain why the outbreak of EAHEC in Germany resulted in such a high fatality rate. Recorded numbers vary by source but from 3816-4075 cases of food poisoning were recorded, and there were around 810 – 908 cases of HUS (21% of cases) and 46 – 54 deaths (1.1% – 1.3% of cases) (Bielaszewska *et al.*, 2011; Blaser, 2011; Frank *et al.*, 2011; Mellmann *et al.*, 2011; WHO, 2011). Throughout the outbreak, vegetables were blamed for spreading the infection across Europe (Buchholz *et al.*, 2011) with cucumbers and then beansprouts receiving most of the media attention. Eventually, the infection was traced back to the storage of fenugreek seeds, and could have spread further via mishandling this stock at any stage of crop growth, harvest, or handling (Aldabe *et al.*, 2011; King *et al.*, 2012).

1.3 Shiga toxin

The Shiga toxin (Stx) protein complex was first characterised from *Shigella dysenteriae* serotype 1 after it was identified in 1896 as the cause of a dysentery outbreak in Japan (Shiga, 1898) . The illness causing bacteria was originally

named *Bacillus dysenteryie*, which was associated with a mortality rate reaching >20% in infants (Shiga, 1906; 1936).

During the early- and mid-20th century, repeated efforts were made to isolate the Shiga toxin protein (Stx), as well as to define its role in disease (Keusch *et al.*, 1981). Stx expression is not essential for *S. dysenteriae* to cause dysentery, but loss of Stx expression in naturally-occurring *S. dysenteriae* strains or *Stx* knock-out mutants has been associated with decreased severity of symptoms of dysentery in human or rhesus monkey subjects, respectively (Fontaine *et al.*, 1988; Levine, 1987). Homogeneous purification of Stx was finally achieved in 1980 (O'Brien *et al.*, 1980), and the purified toxin was confirmed to cause HUS or HUS-like symptoms in mice (Lindgren *et al.*, 1994), rabbits (Eiklid & Olsnes, 1983), dogs (Raife *et al.*, 2004), and baboons (Siegler *et al.*, 2003). Shiga toxin can also be the aetiological agent in infection-associated cases of bloody diarrhoea due to its ability to bind to human intestinal mucosa despite the absence of Gb_3 (Hodges & Gill, 2010; Schüller *et al.*, 2007).

Evidence for the production of Stx by certain diarrhoegenic strains of *E. coli* developed from observations that protein extracts of these strains and *S. dysenteriae* had similar biological and antigenic properties (O'Brien *et al.*, 1982) and from the structural similarity of their purified toxins (O'Brien & LaVeck, 1983). Two antigenically-distinct Shiga-like toxins were identified in different strains of *E. coli* EDL933 of serotype O157:H7, and were designated as Stx1 and Stx2 (O'Brien *et al.*, 1984; Strockbine *et al.*, 1986). Stx1 was found to be neutralised by antiserum to Stx produced by *S. dysenteriae*, whereas Stx2 was not

(Strockbine *et al.*, 1986). The primary structure of Stx1 was identical to Stx of *S. dysenteriae*, except for a single amino acid substitution, whereas Stx2 was only ~55% identical at the amino acid level (Jackson *et al.*, 1987; Johannes & Römer, 2010; Strockbine *et al.*, 1988). Stx1 and Stx2 toxins have both been sub-divided into more than one subtype (described in section 1.3.1). Irrespective of sub-type, Shiga toxin production occurs post-induction of the prophage (Wagner *et al.*, 2001) and is released upon cell lysis (Wagner *et al.*, 2002).

Stx toxins have an 'AB_5' quaternary structure (Fig. 1.3.1), similar to other toxins, such as Cholera (Ctx) and Pertussis (Pt) toxins, which are also carried by phages (Beddoe *et al.*, 2010). Subunit A of Stx possesses N-glycosidase activity responsible for their toxicity (Endo *et al.*, 1988), while the five B subunits mediate host-cell surface binding and initiate endocytosis (Donohue-Rolfe *et al.*, 1989; Ling *et al.*, 1998). With the exception of Stx2e, Stx toxins bind the glycolipid globotriaosylceramide (Gb_3) as their receptor on the surface of target cells (Jacewicz *et al.*, 1986). The presence and prevalence of Gb_3 within the tissues of the host organism is directly correlated to severity of the disease, with Gb_3 deficient cell lines and Gb_3 deficient animals e.g. deer and cattle, being insensitive to Shiga toxin (Acheson *et al.*, 1996; Okuda *et al.*, 2006). Treatments that protect against development of HUS and TPP include chloroquine and glucose analogues, which competitively inhibit the interaction between Gb_3 the B subunits of Shiga toxin (Kavaliauskiene *et al.*, 2017).

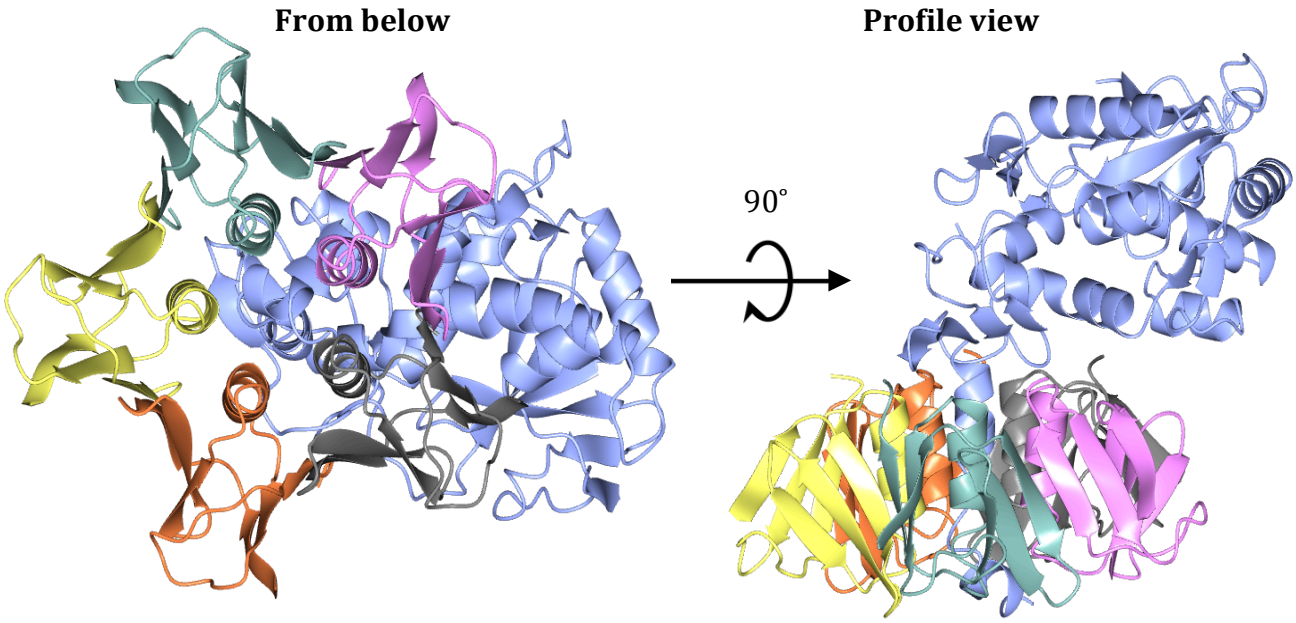

Figure 1.3.1: A schematic of the AB_5 structure of Shiga toxin (PDB 1R4P).

This representation of Stx2 was generated from data published by (Fraser *et al.*, 2004). The subunits of the toxin are coloured by chain – StxA, the toxic subunit, is blue with the 5 identical, non-toxic StxB subunits coloured yellow, green, pink, grey and orange. The underside view (left) shows the 5 fold rotational symmetry of the StxB pentamer that encapsulates an α-helix at the C-terminal end of StxA. The profile view (right) shows that the pentamer only needs to bind a small region of StxA in order to protect the A subunit from premature activation.

Gb_3 is comprised of a ceramide and a trisaccharide moiety, *i.e.* Galα1-4Galβ1-4Glc-ceramide (Bergan *et al.*, 2012). The ceramide moiety anchors Gb_3 to the outer leaflet of the plasma membrane through hydrophobic interactions with its pair of hydrocarbon chains. The Gb_3 trisaccharide moiety is exposed on the exterior surface of the plasma membrane for Stx to bind (Chan & Ng, 2016; Ling *et al.*, 1998). During HUS development, the presence of Gb_3 on the plasma membrane external leaflet of certain cell types in the kidney and central nervous system (CNS) is thought to result in Stx-mediated cytotoxicity and damage to these tissues (Trachtman *et al.*, 2012).

X-ray crystallography of Stx has shown that the B subunits form a homopentameric ring with five-fold rotational symmetry (Figure 1.3.1). Each B subunit contributes a single α-helix to the inner surface of the ring, and an anti-parallel β-sheet to its outer surface (Fraser *et al.*, 1994; 2004). Stx has three Gb_3 binding sites for each B subunit - 15 in total - in contrast to other AB_5 toxins such as Ctx, which has five host binding sites in total (Beddoe *et al.*, 2010; Ling *et al.*, 1998). A mutagenesis study of Stx1 suggested that its three quintets of identical Gb_3 binding sites had different affinities for Gb_3, and found that the lower affinity of mutated individual binding sites was correlated with the impairment of cytotoxicity to Vero cells (Bast *et al.*, 1999).

The A subunit of Shiga toxins is divided into two domains. The N-terminal A_1 domain, containing its N-glycosidase active site, and the C-terminal A_2 domain that is inserted into the centre of the B subunit ring to form the AB_5 structure. The A subunit contains a single disulphide bond that bridges the A_1 and A_2

domains (Fraser *et al.*, 1994; 2004) and defines the polypeptide loop that is cleaved by the host cell enzyme furin to activate the enzymatic activity of the toxin (described in 1.3.2).

1.3.1 Stx variants

Following the discovery of Stx1 and Stx2, additional subtypes of both Stx forms were discovered that had divergent sequences from their "parent" Stx toxins. Nomenclature for the Stx subtypes was eventually formalised by Scheutz *et al.* (Scheutz *et al.*, 2012) in order to develop a protocol for assigning a subtype to each newly-isolated STEC strain. Stx1 toxins were divided into subtypes Stx1a, and Stx1c-e. Stx2 toxins were divided into subtypes Stx2a-g (Johannes & Römer, 2010; Scheutz *et al.*, 2012).

In comparison to Stx2-producing STEC strains, Stx1-producing isolates have been associated with uncomplicated diarrhoea without HUS, or with a lack of symptoms altogether (Boerlin *et al.*, 1999; Chui *et al.*, 2015). *S. dysenteriae* Stx (Strockbine *et al.*, 1988) and the prototypical Stx1 toxin of *E. coli* strain O157:H7 EDL933 (Jackson *et al.*, 1987) were placed in the Stx1a subtype, and the prototypical Stx2 toxin from O157:H7 strain 933 (Jackson *et al.*, 1987) was placed in the Stx2a subtype (Scheutz *et al.*, 2012).

Stx1c toxins were discovered in an STEC strain isolated from sheep (Paton *et al.*, 1995) and, independently, from a longitudinal study of STEC isolates from human patients (Zhang *et al.*, 2002). Stx1c has been found to be associated with

less severe STEC infections, without complications from HUS (Friedrich *et al.*, 2003; Orth *et al.*, 2007).

Stx1d was discovered in an STEC isolated from bovine faeces that gave a positive result from an ELISA Stx screen but a negative result from a PCR-based screen (Bürk *et al.*, 2003). Stx1d-producing STEC have subsequently been isolated from human stool samples, but this toxin subtype is not thought to be associated with a severe disease course (Kumar *et al.*, 2012; Melton-Celsa, 2014).

Subsequently, an additional Stx1 subtype, Stx1e, was discovered in a clinical isolate (*Enterobacter cloaceae*) from a human patient suffering from non-bloody diarrhoea, which had not progressed to HUS (Probert *et al.*, 2014).

Stx2c was originally isolated from the E32511 strain of the O157:H7^{-} ETEC serotype, which was found to be associated with HUS in human patients and produced Stx2a in addition to Stx2c (Schmitt *et al.*, 1991). Stx2b was later identified as a distinct subtype based on sequence divergence of its members from Stx2c (Persson *et al.*, 2007). Stx2a and Stx2c are associated with more severe symptoms of gastrointestinal infection by STEC, in particular HUS (Chui *et al.*, 2015; Friedrich *et al.*, 2002; Orth *et al.*, 2007).

Stx2d was originally identified in a human patient whose STEC infection had progressed to HUS (Ito *et al.*, 1990). Cytotoxicity of Stx2d is activated upon exposure of these toxins to intestinal mucus (Melton-Celsa *et al.*, 1996; Scheutz *et al.*, 2012). Elastase, a protease that is a component of intestinal mucus, activates

Stx2d by cleavage of a glycine-glutamate dipeptide from the C-terminus of the A subunit (Kokai-Kun *et al.*, 2000; Melton-Celsa *et al.*, 2002). Higher cytotoxicity of Stx2d following elastase-catalysed activation has been attributed to higher affinity of binding to the Gb_3 receptor (Bunger *et al.*, 2013). In STEC strains that lack the virulence gene *eae*, which therefore do not intimately adhere or form pedestals, Stx2d is associated with severe symptoms including bloody diarrhoea and HUS (Bielaszewska *et al.*, 2006).

Stx2e was originally identified in *E. coli* isolated from pigs affected by edema disease (ED) (Weinstein *et al.*, 1988). Uniquely among the Stx subtypes, Stx2e toxins preferentially bind globotetraosylceramide (Gb_4) over Gb_3 as their cell-surface receptor (DeGrandis *et al.*, 1989; Samuel *et al.*, 1990). Gb_4 is equivalent to Gb_3 with an additional N-acetylglucosamine β1,3-linked to the terminal galactose residue of its glucan chain; *i.e.* GalNAcβ1-3Galα1-4Galβ1-4Glc-ceramide (DeGrandis *et al.*, 1989). Stx2e-producing STEC have been isolated from human patients with diarrhoea, but this subtype has been found to be associated with relatively mild progression of the symptoms, without HUS (Friedrich *et al.*, 2002).

Stx2f was initially found in the droppings of feral pigeons (Schmidt *et al.*, 2000), while Stx2g was discovered in the faeces of cattle (Leung *et al.*, 2003). Until recently, Stx2f and Stx2g were not thought to be associated with severe disease in humans (Melton-Celsa, 2014). However, Stx2f-producing STEC strains have been isolated from human patients suffering from HUS, and it has been suggested that *E. coli* lysogens of Stx2f-encoding prophages can be transmitted to

humans and cause disease when they have acquired additional non-Stx virulence genes that have been functionally characterised in non-Stx2f-producing STEC (Grande *et al.*, 2016).

1.3.2 Mechanism of Action of Shiga Toxin (Stx) on Human Cells

Stx cytotoxicity begins with the binding of the AB_5 Stx holotoxin to Gb_3 on the cell surface (Fig. 1.3.2.1A). Correct targeting of Stx to the cytosol depends on incorporation of Gb_3 into detergent-resistant membranes (DRMs), also known as lipid rafts or microdomains, which are heterogenous regions of the plasma membrane that are enriched in cholesterol. In Gb_3-expressing cells without Gb_3-containing DRMs, Stx is translocated to lysosomes and degraded (Falguières *et al.*, 2001). Following binding of StxB subunit to Gb_3 in a DRM, endocytosis occurs via the formation of self-contained vesicles with Gb_3-Stx complexes on their internal face (Bergan *et al.*, 2012) (Fig. 1.3.2.1B). Polymerisation of the protein clathrin on the inner surface of the plasma membrane can deform it to form a pit in the membrane, and then narrow the opening of the pit to a bottle-neck, before the GTPase dynamin severs the neck and releases a clathrin-coated vesicle into the cell (Doherty & McMahon, 2009). Clathrin-dependent endocytosis is an important route for Stx entry into the cell, but loss-of-function mutations of proteins involved in this pathway, including dynamin and clathrin, in cultured cell lines have shown that Stx can exploit at least one clathrin-independent mode of endocytosis (Bergan *et al.*, 2012; Lauvrak *et al.*, 2004).

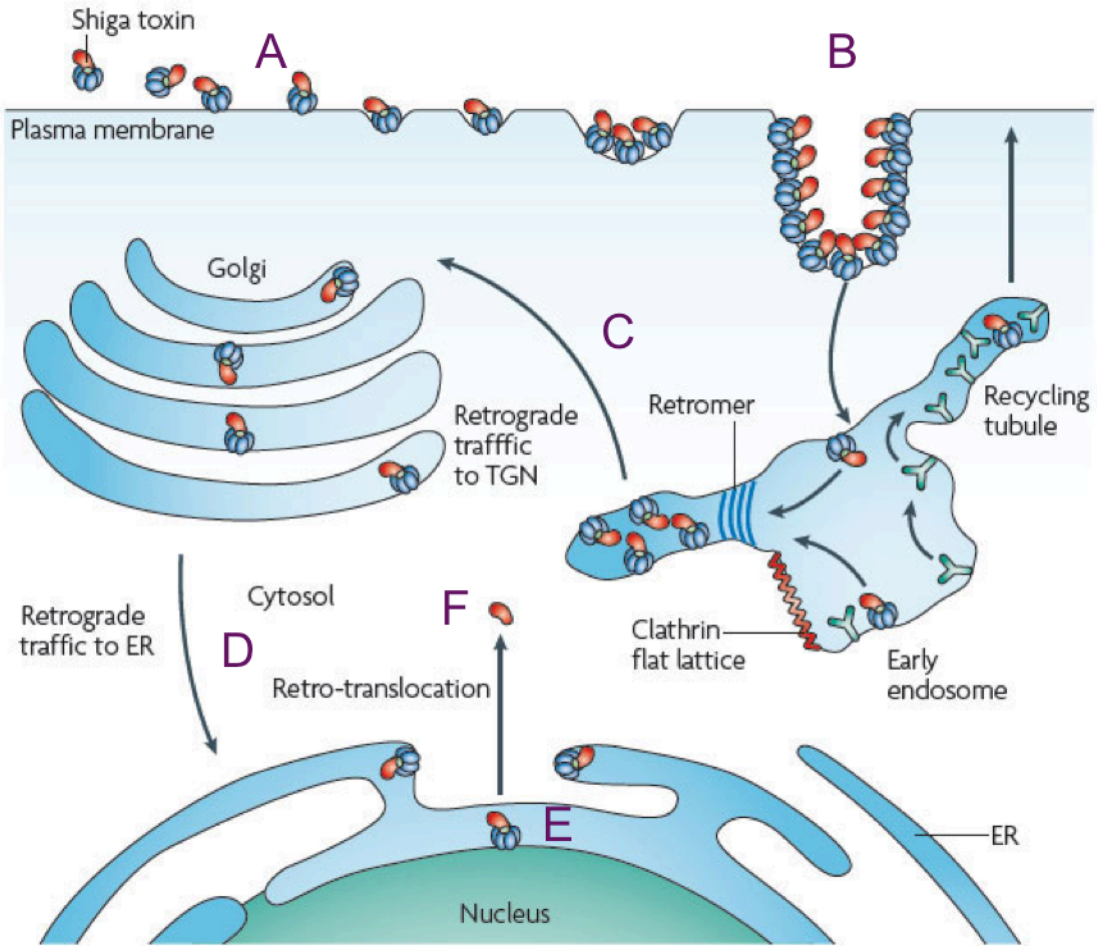

Figure 1.3.2.1: The journey of Shiga toxin from binding to the host cell to retrotranslocation.

(A) Shiga toxin recognises the cell via Gb_3 on the surface of the plasma membrane. (B) Shiga toxin is internalised into an early endosome via a clathrin-coated vesicle. (C-D) Shiga toxin is trafficked to the Golgi body where it undergoes retrograde transport into the ER where the toxin is cleaved by a furinase to generate the activated Stx toxin (E-F). Activated Stx toxin is retrotranslocated into the cytosol where it inhibits tRNA binding to the ribosome, inducing the ribotoxic response and apoptosis. Figure adapted from (Johannes & Römer, 2010)

Following endocytosis, Stx is transported to the endoplasmic reticulum (ER) (Fig. 1.3.2.1C-D) via the endosome and the Golgi apparatus (McKenzie *et al.*, 2009; Sandvig *et al.*, 1992; Warnier *et al.*, 2006). In the process of retrograde transport, the A_1 and A_2 domains of the A subunit become separated, with the former becoming the active Stx, or activated cytotoxin. This is achieved by cleavage of the A subunit polypeptide chain at an Arg-Val-Ala-Arg motif, catalysed by furin, a membrane-bound endopeptidase principally localised to the trans-Golgi network. Reduction of the disulfide bond is also required to allow the A_1 and A_2 domains to separate (Garred *et al.*, 1995; 1997; Samuel & Gordon, 1994)

The holotoxin is retrotranslocated from the ER lumen to the cytosol (Fig. 1.3.2.1E-F). Stx exploits the ER-associated degradation (ERAD) mechanism, which functions to translocate misfolded proteins from the ER lumen to the cytosol so that they can be proteolytically degraded. Retrotranslocation of activated StxA is dependent on the Sec61 translocon, as revealed by siRNA-mediated knock-down of its subunits in HeLa cells (Moreau *et al.*, 2011). Cleavage of the A subunit exposes a C-terminal hydrophobic patch in the A_1 domain that is thought to mimic the surface hydrophobicity of misfolded proteins and trigger retrotranslocation (LaPointe *et al.*, 2005). Stx1 can be co-immunoprecipitated with Sec61 and the ER lumen chaperone HEDJ/ERdj3, which is thought to indicate that HEDJ recruits the A_1 holotoxin to the translocon at the ER membrane (Yu & Haslam, 2005).

Stx holotoxin catalyses the removal of a conserved adenine residue from the 28S rRNA of eukaryotic ribosomes using its N-glycosidase activity, and this results in

the inactivation of ribosomal binding of aminoacyl-tRNAs and thus of protein synthesis (Endo *et al.*, 1988). Stx intoxication has been found to induce apoptosis in multiple cell lines (Tesh, 2012). Inactivation of protein synthesis is thought to induce apotosis via two endogenous signalling mechanisms: the ribotoxic response, which is activated by inhibition of ribosomal tRNA binding and peptidyl transferase activity, and the unfolded protein response, which is activated by accumulation of improperly-folded proteins in the ER lumen (Tesh, 2012).

1.4 Shiga toxin producing *E. coli* infections

The disease characteristics of Shiga toxin-based illnesses result from *E. coli* infections where Shiga toxin is produced in the gut (Brigotti *et al.*, 2006). The onset of symptoms can be caused by the ingestion of less than 10 organisms (CFU mL^{-1}) (Lindgren *et al.*, 1994). This makes STEC one of the most notorious causes of foodborne outbreaks; for example, in the USA, 55% (89 of 163) of all reported foodborne outbreaks between 2007-2012 were associated with shigatoxigenic *E. coli* (Robertson *et al.*, 2016). Case numbers are comparable but the rate of outbreaks involving *E. coli* in England and Wales from 2014 to 2018 is much lower (12% - 11 of 89). This may be due to the lack of reporting for minor food poisoning symptoms not providing enough cases to be classified as outbreaks. It may also be that other foodborne pathogens are more prevalent in the UK, with *Salmonella* (26%), *Campylobacter* (16%), and *Clostridium* (15%) taking a significant share of reported outbreaks (PHE, 2014-2018).

Responses to STEC epidemics using molecular biological and microbiological approaches are thorough, but can be very laborious and take multiple days (Bielaszewska *et al.*, 2011). The best molecular biological approach involved a complete sequencing and assembly of a draft genome of the outbreak strain that was completed within 62 hours (Mellmann *et al.*, 2011). The amount of information generated from this approach was considerable and allowed the identification of EAHEC, a new pathotype of diarrheogenic *E. coli*. However, at over 2.5 days, this method of pathogen detection is still extremely slow to consider using it during an epidemic response time, as it would be unable to inform a rapid response to a developing situation.

1.4.1 Detection of STEC

There used to be three methods of diagnosing an STEC infection: isolation of viable STEC cells from stool samples, detection of Stx in stool sample, or detection of elevated levels of antibodies to STEC antigens in serum (Nataro & Kaper, 1998). Patients presenting with bloody diarrhoea or HUS would have stool samples analysed on sorbitol-MacConkey (SMAC) agar plates to select for enteric Gram negative bacteria and the presence of EHEC strain O157:H7 is detected by its inability to ferment the sorbitol, producing colourless colonies (Boyce *et al.*, 1995). SMAC agar is only able to reliably identify O157:H7; other culturing methods are available for other strains (reviewed by Nataro & Kaper, 1998), but there is no catch-all culturing solution that will detect all known EHEC or STEC strains.

PCR based methods can provide results in just hours as opposed to overnight timescales for bacterial growth on agar plates. Gene targets for PCR include *stx* (EHEC and EAHEC), *eae* (EHEC and EPEC), or the pO157 plasmid (O157:H7 only). Both Stx1 and Stx2 can be detected directly from stool samples by PCR (Ramotar *et al.*, 1995). Multiplexing of targets yields a more reliable method of diagnosis with combinations of type III effector genes producing the highest level of reliability for detecting EHEC (Delannoy *et al.*, 2013).

Looking beyond standardised detection methods, STEC can also be identified by fluorescent-tagged phages or endolysins that bind specifically to pathogens (Bai *et al.*, 2016), but an effective, scalable assay for this would need to be developed for it to be a seful diagnostic tool.

1.4.2 Medical treatments/Preventative measures

1.4.2.1 Antibiotics:

Antibiotics are routinely used to treat non-STEC *E. coli* infections. However, STEC strains present a dilemma to the physician considering using antibiotics, as they tend to induce the STEC's SOS response, which leads to activation of promoters that lead to both the production of Stx and another generation of Stx phages (Yang *et al.*, 2017; Zhang *et al.*, 2000). When using most antibiotics, there is a significant correlation between STEC infections and an increased rate of HUS development (Freedman *et al.*, 2016). For this reason, the currently accepted treatment strategy by physicians is to not treat probable STEC infections with any antibiotics (Freedman *et al.*, 2016; McGannon *et al.*, 2010; Szych *et al.*, 2014).

Exceptions to this have been reported; for example, telithromycin, a ketolide antibiotic that inhibits protein production by binding the 50S ribosomal subunit, can treat STEC strains without inducing the SOS and subsequent Stx production (Nakagawa *et al.*, 2003). This backs up the lack of progression from STEC infection to HUS post-antibiotic treatment seen previously (Takeda, 1998), however, there are no studies to suggest that antibiotics which lead to DNA damage are capable of killing cells without inducing Stx production.

Using antibiotics also targets other gut microflora. This collateral damage reduces biodiversity within the gut, which is linked to an increased likelihood of STEC colonisation (Mir *et al.*, 2016), so they may make the situation worse in the long term. Furthermore, between 1988 and 2007, the carriage of at least one antibiotic resistance gene increased from 2.9% to 79.8% in O157:H7 strains isolated from a variety of sources: bovine and human faeces, bovine milk products, ground beef, and cider (Colavecchio *et al.*, 2017). This limits the options of antibiotics available to use for a STEC infection but as phages can be responsible for the dissemination of antibiotic resistance genes (Modi *et al.*, 2013), it is probably best not to treat STEC infections with antibiotics, thereby denying further opportunity to spread these resistance genes to currently susceptible populations.

Alternatives to small molecule antibiotics have started to arise as potential antibiotic agents. Peptides that bind essential factors on the surface of *E. coli* cells are now being investigated as potential antimicrobial candidates (Hagan *et al.*, 2015; Vimala *et al.*, 2015).

1.4.2.2 Inhibitors

Shiga toxins can be directly inhibited by binding carbohydrates, which recognise the StxB subunits of the AB_5 toxin (Mulvey *et al.*, 2003; Nishikawa *et al.*, 2002; 2005). For example, "SUPER TWIGs" act as immune-response aids that mimic the sugar-rich glycosylated extracellular surface of eukaryotic cells (Nishikawa *et al.*, 2005). They were specifically developed as Stx neutralisers, which also utilise deliberately oriented saccharides to bind to the StxB subunit. This action blocks the binding of StxB to Gb3 while also enabling efficient phagocytosis of toxin molecules by macrophage cells in the circulatory system.

Inhibitors of cell-cell signalling have also demonstrated reduction in colonisation efficiency. Disruption of inter-cell communication impedes the ability of the community to form effective biofilms, making bacterial cells more easily targeted by the immune system as they are not camouflaged by factors secreted by biofilms (Watters *et al.*, 2016). This could be a way to limit the pathogenicity (but may increase the shedding) of STEC (Desin *et al.*, 2015; Vikram *et al.*, 2012).

1.4.2.4 Diet

A high fibre diet was shown to increase O157:H7 colonisation in a mouse model by up to 100x compared to a low fibre diet (Zumbrun *et al.*, 2013). So eating less fibre could reduce the impact of an STEC infection on an individual human. Feeding hay to cattle instead of grain increases shedding duration (Hovde *et al.*, 1999) meaning that the diet of ruminants may increase the spread of STEC and therefore have an influence on human health.

The use of probiotic bacteria in foodstuffs to maintain healthy competition for gut resources has shown promise as a preventative strategy against O157:H7-induced death in mouse model (Takahashi *et al.*, 2004). Probiotics produced by *Lactobacillus acidophilus* La-5 can even down-regulate certain EHEC virulence factors, including Stx2, when fed to mice (Zeinhom *et al.*, 2012).

1.4.2.5 Vaccines:

DNA- and protein-based vaccines have been developed in mouse models to protect against STEC and Shiga toxin, respectively (Ren *et al.*, 2013; Sato *et al.*, 2013). Vaccinating cows with type III secretion proteins reduced both the duration of shedding and numbers of pathogenic bacteria in faeces which reduces the chance of contamination in the food supply chain (Potter *et al.*, 2004). As is always the case with vaccines, their effectiveness would be dependant on predictions of cases/outbreaks, which has not been possible so far.

1.4.2.6 Groundwater monitoring:

Identical phage structural genes have been recorded in metavirome data worldwide (Short & Suttle, 2005) demonstrating the ubiquitous nature of phages. However, high levels of genetic heterogeneity between Stx phages in the environment mean that only Stx genes can be used as targets in attempts to monitor potential STEC outbreaks (García-Aljaro *et al.*, 2004; Muniesa & Jofre, 2004). This limitation means that potential bacterial hosts represent a more stable target. One concern regarding the efficacy of this strategy would be freshwater-dwelling amoebae. *Acanthamoeba spp.* are present in air, soil, and

water sources such as taps, reservoirs, and swimming pools (Rowbotham, 1980; Thom *et al.*, 1992). They are known to feed on Gram-negative bacteria but some pathogenic bacteria, including *E. coli*, are known to be able to survive and multiply post-internalisation, providing survival reservoir for STEC (Chekabab *et al.*, 2013; Ly & Müller, 1990; Winiecka-Krusnell *et al.*, 2002). Similar observations have been made in *Tetrahymena pyriformis* where prophage carriage increases the fitness of both K-12 and O157 *E. coli* strains, promoting their survival within the food vacuoles. This potential to act as a vector for bacterial pathogens is troubling from a health perspective, however, it also raises the possibility that any detection strategy targetting on surface antigens of *E. coli* could fail to detect viable pathogenic *E. coli* if the issue is not accounted for. Production of Stx while the EHEC are internalized increases amoebic mortality and therefore reduces the capacity of the EHEC reservoir (Chekabab *et al.*, 2013).

1.5 The role of bacteriophages in disease

Bacteriophages (phages) are known to be the primary vectors for the horizontal gene transfer of Shiga toxin-encoding genes between bacteria (O'Brien *et al.*, 1984). Such Shiga toxin-encoding bacteriophages (Stx phages) are lambdoid phages, which are temperate, thus capable of entering either the lytic or lysogenic replication cycles (Campbell, 1994). Stx expression is associated with induction of lytic replication irrespective of which life cycle is taken (Tyler *et al.*, 2013). The pathogenic potential of the *stx* genes is enhanced if the infecting phage enters the lysogenic cycle to become a prophage. The prophage is amplified across the lysogen population during host cell growth and division,

while the host becomes a more significant member of the gut flora. In this event, when prophages are induced into the lytic replication cycle by the activation of the bacterial host's SOS stress response, there could be many more *E. coli* cells containing the Stx prophage and therefore there is a greater potential for Stx expression (Fogg *et al.*, 2012).

1.5.1 Stx phages

Stx phages are highly mosaic (Gamage *et al.*, 2004; Hendrix *et al.*, 2003; Johansen *et al.*, 2001; Smith *et al.*, 2012), which is thought to be due to *in-situ* recombination events within lysogens. Therefore, the presence of an *Stx* operon encoding at least one of the two variants (Stx1 or Stx2) is the only way to classify a bacteriophage as a Stx phage (Allison, 2007; Muniesa *et al.*, 2004a, b). Any "chromosomal" versions of Stx genes that have been described in the literature have always been ultimately found to be part of a defective prophage, which has presumably undergone recombination in the host and lost vital phage replication or assembly genes and is no longer able to replicate (Mizutani *et al.*, 1999).

A model Stx phage, $\phi 24_B$ is used in this thesis as it is stable on storage and possesses a short tail that confers a large host range (Smith *et al.*, 2007). Furthermore, $\phi 24_B$ has undergone more genetic characterisation than other Stx phages (Holt *et al.*, 2017; Smith *et al.*, 2007; 2012). $\phi 24_B$ possesses an inverted integrase gene, which allows superinfection events - the presence of more than one prophage in a lysogen genome (Allison *et al.*, 2003; Fogg *et al.*, 2007; 2011). This is extremely rare and goes against the typical immunity paradigm

developed while studying lambda, λ, and lambdoid phages in general which states that only one copy of a lambdoid prophage can be integrated into the host cell genome at any given time (Fogg *et al.*, 2010). Superinfecting λ phage has been studied before, and was only possible using mutants in immunity or excisionase genes (Calef *et al.*, 1965; Schneider & Strack, 1977). This increased copy number of *stx* genes can facilitate a larger or quicker build-up of toxin, and the amount of toxin produced during an infection is directly proportional to the severity of disease phenotype (Fogg *et al.*, 2012). Greater potential toxin release is likely to be due to the phage-infected cells being capable of producing and releasing infective toxigenic phages capable of converting gut microbiota residents into Shiga toxin production factories (Gamage *et al.*, 2003). Uninfected but phage-susceptible commensal bacteria can be powerful contributors to the overall production of Shiga toxin as they have been shown to generate up to 40-fold more toxin than lysogens alone (Gamage *et al.*, 2003; 2006). This may also explain the 3-log increase in Stx2 levels compared to *in vitro* expectations during some phage infections (Tyler *et al.*, 2013).

1.5.2 Effect of prophage carriage on host fitness

The effects of prophage carriage (also referred to as lysogenic conversion) on the host bacterium are many-fold and detailed descriptions of them all are beyond the scope of this thesis. Examples include but are not limited to superinfection exclusion, cell envelope modification for increased pathogenicity, and defence systems to prevent the CRISPR-Cas from cleaving phage DNA. These aspects are reviewed elsewhere (Bondy-Denomy & Davidson, 2014).

Specifically in the $\phi 24_B$ genome, there are several examples of host benefits derived from prophage carriage. The most obvious change during the infection is the conversion of the *E. coli* host to a pathogen is the introduction of *stx* into the cell, allowing the production of Shiga toxin in the host (Schmidt, 2001). The benefits to host fitness are often ways to facilitate or increase the effectiveness of the toxin release.

The upregulation of glutamic acid carboxylase (GAD) operon confers additional acid resistance to the lysogen compared to a naïve MC1061 cell (Veses-Garcia *et al.*, 2015), which increases survival rates of the host cell travelling through the stomach to get to the lower GI tract where it can colonise.

Other examples in the $\phi 24_B$ genome are genes 13b, which increases motility of the host (Veses-Garcia *et al.*, unpublished data), and gene 48, which is a huge 8kb gene that has domains homologous to antibiotic resistance genes. The extent of any increased antibiotic resistance of MC1061 containing an isolated gene 48 have yet to be tested.

1.6 Identification of the receptor for $\phi 24_B$

When the Stx phage $\phi 24_B$ was first identified, an effort was made to determine its adsorption target on the surface of *E. coli*. At this time, most experts felt that Stx phages could only infect rough strains as the smooth strains prevented access to receptors on the cell surface (van der Ley *et al.*, 1986). It was later

demonstrated that Stx phage could infect a smooth serotype O107 *E. coli* strain (James *et al.*, 2001).

The strategy used to identify which antigen $\phi 24_B$ adsorbed to on the host used an adsorption resistant mutant, which would be complemented with an expression library generated from the genomic DNA (gDNA) phage susceptible host (Sergeant, 1998). The Stx phage $\phi 24_B$ was originally identified on the *E. coli* indicator strain MC1061 (Allison *et al.*, 2003). In order to create the necessary phage-resistant mutant, a culture of MC1061 was infected with a cocktail of five different Stx phages with different immunity profiles (Sergeant, 1998). Surviving cells from the infection were identified by spreading onto non-selective agar. Resultant colonies were screened for their ability to produce phage upon induction. Four colonies were identified that failed to produce infective phages. These four clones contained prophages of a subset of the phages in the cocktail, and lost their ability to support $\phi 24_B$ adsorption. The clones were named MC1061-MRL1-4 (MRL = Multiple Resistant Lysogen). One of the MC1061-MRL strains was then used to screen a pUC19-based expression library created by partial digestion of wild type MC1061 genomic DNA with *Sau*3A I (Allison *et al.*, 2003; Sergeant, 1998). Four independent clones from the expression library were identified. The plasmids from these four clones all restored sensitivity to $\phi 24_B$ upon the phage resistant mutant, and all four plasmid clones carried DNA inserts that shared a region of DNA sequence identity. Subcloning of one particular gene within this shared region was necessary and sufficient to restore

phage sensitivity to MC1061-MRL1. This gene was named *vpr* for **v**erocytotoxin **p**hage **r**eceptor (Sergeant, 1998).

The ORF, *vpr*, shared identity with a few other hypothetical proteins from a variety of other Gram-negative organisms, but in 1998, bioinformatic analysis had a more limited ability to predict potential cellular function from DNA sequence data. Around this time the first *E. coli* K12 genome sequence was published (Blattner *et al.*, 1997). Blattner and colleagues had annotated hypothetical genes in the genome location of *vpr* with the three letter code *yae*, with *vpr* being assigned the designation *yaeT* (Blattner *et al.*, 1997).

Although a phage resistant mutant (MC1061-MRL) was isolated by Sergeant (Sergeant, 1998) for the complementation library for phage sensitivity, this mutant was lost because it could not be resuscitated from -80°C stocks. Unfortunately, there are questions raised here that will never be answered: what and where was the exact mutation that resulted in the creation of MC1061-MRL strains?; how did MC1061-MRL cells survive if the receptor protein was mutated?; and was the fragility and short-lived nature of MC1061-MRL due to a mutation in the receptor?

Many efforts were subsequently made to knock out the receptor in MC1061 using a number of strategies. These utilised suicide vectors and attempts to interrupt *yaeT* through a single recombination event through to the complete replacement of *yaeT* by allelic exchange through double homologous

recombination events. However, no knock out mutant was ever generated using these methods (James, 2002; Sergeant, 1998; Smith, 2004). After these efforts, it was discovered that *yaeT* is considered to be one of the ten most essential genes in the *E. coli* genome (Gerdes *et al.*, 2003). It follows that the essential nature of YaeT makes its disruption impossible, regardless of the strategy employed to attempt it and that if natural point mutations arise those strains may be less fit.

The expression of YaeT is regulated by the σE regulon (Dartigalongue *et al.*, 2001b). In 2007 it was demonstrated that the adsorption potential of $\phi 24_B$ was directly proportional to the expression of YaeT following perturbations to growth conditions that up- or down-regulated YaeT (Smith *et al.*, 2007). It was also demonstrated that rabbit antibodies raised to histidine-tagged, denatured YaeT blocked phage adsorption in a dose dependent manner (Smith *et al.*, 2007). These data supported the idea that YaeT was involved in supporting the adsorption of $\phi 24_B$, but failed to definitely prove that YaeT is the adsorption target for $\phi 24_B$. This study also implied that a disproportionately high number (70%) of short-tailed Stx phages from a collection of phages generated over 5 years of sampling various water sources on local farmland (Robinson *et al.*, 2004) would utilise YaeT as a host recognition molecule to initiate the infection process.

In the 10 years between 1997 and 2007, 397 more bacterial genome sequences were uploaded to the Genbank database, revealing that YaeT was conserved across all Gram negative bacterial species and is very highly conserved (>90%

amino acid identity) across the members of the *Enterobacteriacea* family (Ruhe *et al.*, 2013b; Smith *et al.*, 2007). In 2008, YaeT was renamed for what is likely to be the final time after the Silhavy group's work that determined its function to be instrumental in inserting β-barrelled OMPs into the outer membrane (Aoki *et al.*, 2008). As the major constituent of the β-**b**arrel **a**ssembly **m**achinery, it was named BamA.

1.7 Structure and Function of BamA

1.7.1 Discovery

BamA is now quite well known and has managed to overcome the identity crisis it had at the turn of the 21st century; it was previously identified as YaeT in *E. coli* strain MG1655 (Werner & Misra, 2005), EcfK in strains MC4100 and ST131 (Dartigalongue *et al.*, 2001a), YzzN in strain W3110 (Alba *et al.*, 2002), and Vpr in strain MC1061 (Sergeant, 1998). In 2003, the gene function of a BamA homologue, Omp85, was identified in *Neisseria meningitidis* as having a role in outer-membrane protein assembly (Voulhoux *et al.*, 2003). This was possible only because *N. meningitidis* is an unusual Gram-negative organism that is able to survive without a capsule or LPS (Bos & Tommassen, 2005).

1.7.2 BamA function

BamA is a member of the Omp85 superfamily, which has evolved to have functions in protein translocation and assembly (Heinz & Lithgow, 2014). BamA is the principle component in the β-**b**arrel **a**ssembly **m**achinery (BAM) complex. The BAM complex is a heteropentameric complex composed of proteins BamA-BamE, which is responsible for inserting nascent β-barrel shaped outer membrane proteins (OMPs), including itself, into the outer membrane of Gram negative bacteria (Jain & Goldberg, 2007). β-barrelled structures form a cylinder of antiparallel beta sheets that usually create a pore through the centre and have a hydrophobic outer surface that faces the fatty acid domains of the phospholipids in the bilayer. These proteins facilitate selective permeability across the outer membrane for a broad range of metabolites; therefore the BAM complex, which folds these β-barrel shaped proteins, is essential for cell viability and must be present and functional in the outer membrane at all times. The essentiality of BamA was determined by Gerdes *et al.* (2003) before any functional data had been collected (Werner & Misra, 2005; Wu *et al.*, 2005). This makes the BamA and BAM complex difficult to study, as both deletion mutants and loss-of-function mutants are lethal.

1.7.3 BamA Structure

BamA is an autotransporter protein inserted into the outer membrane that contains two distinct modules: one is a membrane-bound β-barrel domain with 16 trans-membrane regions, and the other consists of five soluble **po**lypeptide-

transport-**a**ssociated (POTRA) domains that extend into the periplasm to bind nascent OMPs. These OMPs are chaperoned across the periplasm to the BAM complex by the chaperone protein, SurA (Malinverni & Silhavy, 2011) where they are transferred to the POTRA domains. The original model for the structure of BamA published by Smith *et al.* (2007) was based on FhaC, a closely related β-barrelled protein that had already been crystallised (Voulhoux *et al.*, 2003). This initial prediction misidentified the position where the POTRA domains end and the β-barrel begins in BamA. Crystal structure evidence of BamA now shows us that the first four antiparallel β-sheets are much shorter than the others and only span ~60% of the membrane (Noinaj *et al.*, 2013). This discrepancy in expected length, which is not present in FhaC, combined with the difficulty in predicting β-barrel structures (Bigelow *et al.*, 2004) probably led to the initial mistake. The literature was corrected after the notation used in this study was employed (Volokhina *et al.*, 2013) and subsequently, the notation for loop numbers in this study were "mis-numbered" i.e. "loop II" in this study is actually referring to loop 4, "Loop IV" refers to loop 6, and "loop V" refers to loop 7. References to these loops in this thesis match those in the current literature with the exception of primer names, which have not been updated to the new notation.

Earlier crystal and solution NMR structures were only acquired for combinations of the five soluble POTRA domains of BamA (Gatzeva-Topalova *et al.*, 2008; 2010; Kim *et al.*, 2007; Knowles *et al.*, 2008; Zhang *et al.*, 2011). As these are all periplasmic domains, they can not offer insight into potential recognition sites/surfaces for the phage tail spike protein to recognise the surface of the *E. coli* cell. A series of crystal structures have been obtained in recent years,

including some of the whole BamA protein (Albrecht *et al.*, 2014; Ni *et al.*, 2014; Noinaj *et al.*, 2013), BamA binding to the BAM complex lipoproteins (Bakelar *et al.*, 2016; Bergal *et al.*, 2016; Chen *et al.*, 2016; Jansen *et al.*, 2015), and even the entire BAM complex (Gu *et al.*, 2016; Han *et al.*, 2016; Iadanza *et al.*, 2016).

The first BamA barrels to be crystallised were orthologues from *Neisseria gonorrhoeae* (Fig. 1.7.3.1) and *Haemophilus ducreyi*. These were reported together by Noinaj *et al.* (2013) where the presumptive open and closed conformations of the BamA pore were revealed. The L6 loop is the major difference between the open and closed conformations on the extracellular side being very accessible in the open conformation but is found within the pore in the closed state. The L6 loop provides a key role in supporting the protein folding and/or membrane insertion function of BamA (Leonard-Rivera & Misra, 2012).

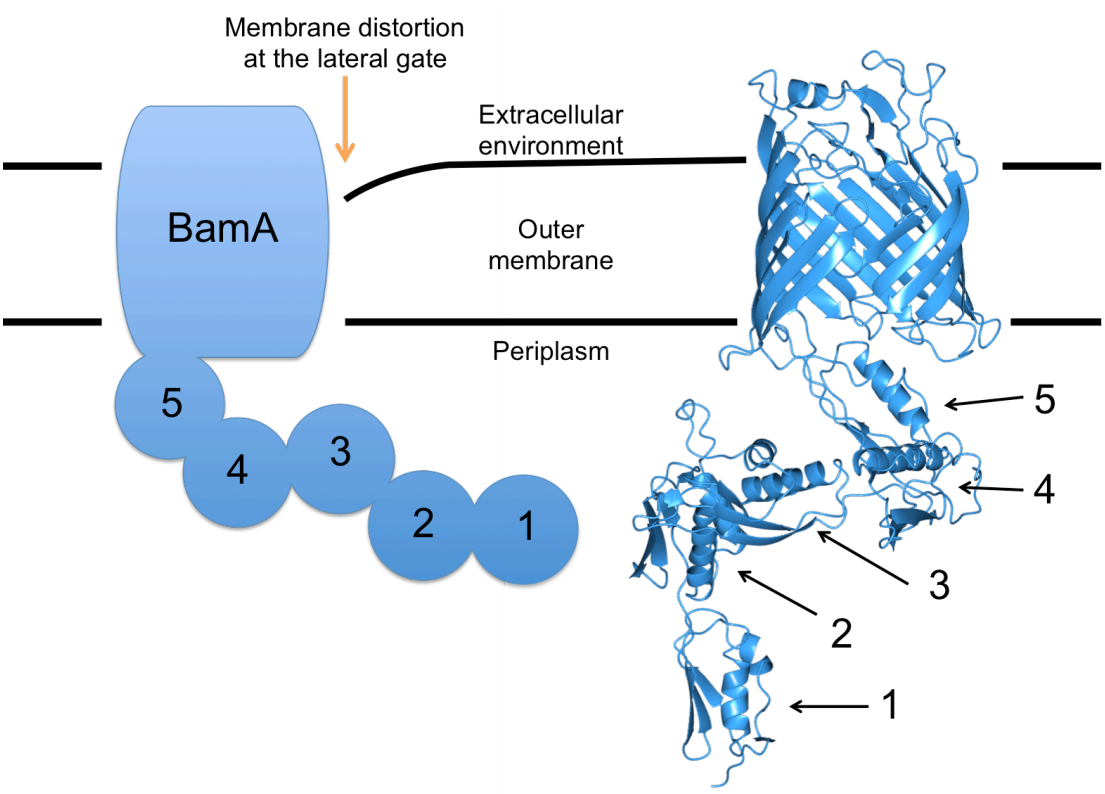

Figure 1.7.3.1: Schematic and crystal structure representations of BamA.

(Left) The schematic version of BamA shows overall structure of the 16-sheet β-barrel transmembrane domain with POTRA domains extending into the periplasm. POTRA domains are numbered 1-5 on both. (Right) 3D structure of BamA orthologue encoded by *N. gonorrheae* (PDB entry: 4K3B (Noinaj *et al.*, 2013)) shows the secondary and tertiary structures present in BamA. Some secondary structures can be seen in the extracellular loops between the linking the sheets in the β- barrel.

Although new data is pointing to a model whereby the BAM complex undergoes a series of "spring-loaded" conformational changes when binding to incoming OMPs, the exact mechanism by which OMPs are manipulated by BamA is currently unknown, and the control mechanism between the open and closed states of the protein is also unclear. Mechanisms have been predicted by modelling structural data to produce sensible theories such as the one put forward by Heuck *et al*. (Heuck *et al.*, 2011), which suggests that nascent OMPs interact with the lipoproteins of the BAM complex, BamB-E. However, these ideas can be comprehensively refuted with additional biochemical data showing that nascent OMPs interact directly with the POTRA domains in BamA while being guided by the lipoproteins (Jansen *et al.*, 2012). Computer models of potential mechanisms have also been put forward (Noinaj *et al.*, 2013). With only a small amount of evidence regarding the flexibility and range of conformational changes of the β-barrel (Doerner & Sousa, 2017), a comprehensive mechanism for BamA still remains elusive.

It has been suggested that the overall mechanism is cyclical in that β-sheets that make up the nascent OMPs are pulled through the BamA pore and inserted into the membrane two at a time (Rigel *et al.*, 2013), but specific details have not yet been confirmed. These data models open up the possibility that Stx phage binding is probably conformation-dependant; however, it cannot be determined which conformation it is dependant on without acquiring further structural or biochemical data regarding the interaction between the BamA β-barrel and the phage.

Crystal structures show evidence that the membrane would be thinner at the lateral gate compared to the rest of the bilayer (Noinaj *et al.*, 2013) introducing a tension into this microenvironment (Arrow in Fig. 1.7.3.1). As the protein insertion process does not require energy (Hagan & Kahne, 2011) and membrane proteins have a high thermodynamic stability, the tension introduced to the membrane may act as the initial "spring", providing free energy to allow the OMP to overcome the kinetic barrier that phosphoethanolamine head groups impose on transmembrane moieties in lipid bilayers (Gessmann *et al.*, 2014). This forced reduction of the width of the membrane to a defined level increases disruption proportionally with membrane thickness (Schiffrin *et al.*, 2017).

1.8 The Role of BamA within the BAM complex

It has been demonstrated that OMP "rafts" are generated at mid-cell during outer membrane biogenesis (Kleanthous *et al.*, 2015) – probably centred around the BamA protein (Rassam *et al.*, 2015). *In vitro* experiments have shown that BamA is able to insert OMPs into the outer membrane alone but takes around 2 hours to accomplish what the BAM complex can do in 5 minutes *in vitro* (Hagan *et al.*, 2013). In *E. coli*, the BAM complex includes four lipoproteins, BamB-E (previously identified as YfgL, NlpB, YfiO and SmpA, respectively (Sklar *et al.*, 2007; Wu *et al.*, 2005)), which are post-translationally palmitoylated to allow association with the inner leaflet of the outer membrane (Fig. 1.8.1). BamA orthologues have been found in the sequenced genomes of all families of Gram-negative bacteria so far, supporting evidence not only for the essentiality of the complex but also its crucial role in the proteination of the LPS membrane.

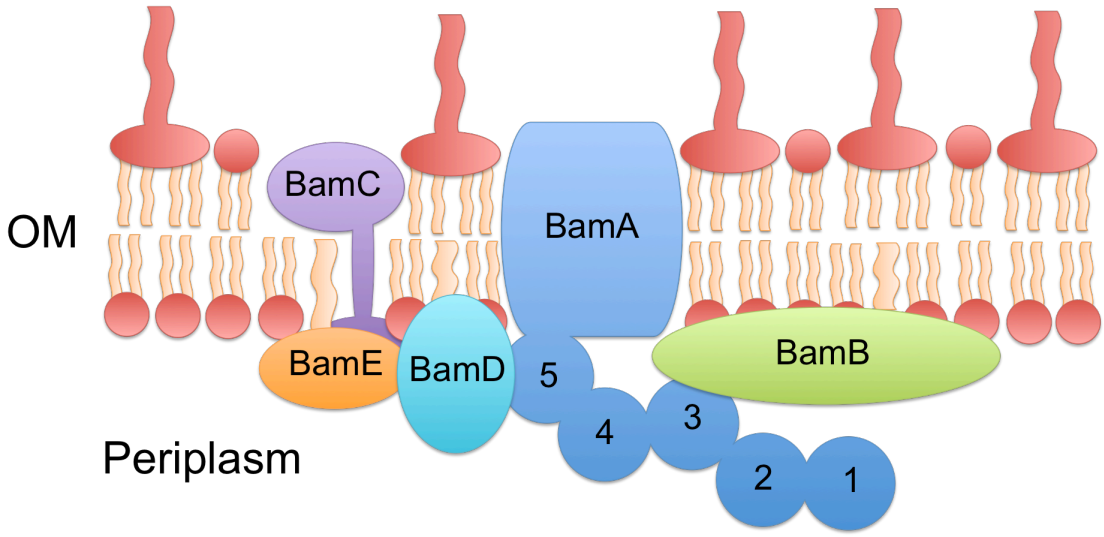

Figure 1.8.1: A schematic representation the BAM complex.

The production of this figure relied on consolidating information from all sources regarding BAM complex interactions with a few key contributors (Kleanthous *et al.*, 2015; Noinaj *et al.*, 2011; Rassam *et al.*, 2015). The BAM complex can be divided into 2 sub-complexes BamAB and BamACDE. BamACDE are the facilitators of OMP insertion while the role of BamAB is not required for this process. BamC and BamE bind to BamA via BamD, perhaps demonstrating why BamD is essential. The supporting role of BamB is yet to be determined. Protein sizes are approximately to scale (by mass) with respect to the width of the membrane.

The BamA-E proteins have a stoichiometry of 1:1:1:1:1 within the BAM complex, but there has been speculation that the complex may oligomerise to a dimer or a tetramer *in vivo* (Robert *et al.*, 2006). The BAM complex has been successfully reconstituted *in vitro* to show that the OMP insertion function does not require energy and that the OMP is presented to the BAM complex by a chaperone, SurA (Hagan & Kahne, 2011). It is unknown how many copies of BamA are required by a single phage to facilitate an adsorption event. As the structure of the $\phi 24_B$ tail protein has been predicted to be a homotrimer (Prof. Olga Mayans, personal communication), it is logical to assume the structure of the tail exhibits 3-fold rotational symmetry, however, there is no obvious symmetry in the extracellular domain of a single molecule of BamA. When also considering that $\phi 24_B$ is a lambdoid phage and that the receptor for λ phage, LamB, is also a homotrimer (Vos-Scheperkeuter & Witholt, 1984), it is plausible that three copies of BamA are required, in close proximity to each other, to allow adsorption. The diameter of the BamA cylinder is the same as the dimensions of the tail (~10nm). But there are no data of models suggesting a docking mechanism.

1.8.1 Lipoproteins of the BAM complex

BamB binds to BamA POTRA domain number 3 (Heuck *et al.*, 2011; Jansen *et al.*, 2015). BamB has been implicated in binding POTRA domains to guide the folding of nascent OMPs by the BAM complex and chaperones (Jansen *et al.*, 2012; Noinaj *et al.*, 2011). While not essential to the protein folding process, it is critical for anti-phagocytosis of *Klebsiella pneumoniae* (Hsieh *et al.*, 2016). Similar, but less pronounced immune-evasion effects are seen in *Yersinia enterolitica* (Weirich *et*

al., 2017) and *Pseudomonas aeruginosa* (Lee *et al.*, 2017) leading to speculation that BamB could become an antibiotic target in the future; however, as BamB is not surface exposed, it would require a treatment delivery system that allows passage through the outer membrane (or the pore of BamA).

Not much is known about BamC. This lipoprotein spans the outer membrane with the N-terminal domain binding to BamD on the periplasmic side and the C-terminal domain being surface exposed. This exposure has been shown experimentally as antibodies have been used to visualise BamC localisation in live cells (Rassam *et al.*, 2015; Webb *et al.*, 2012).

BamD has been shown to be the only lipoprotein of the BAM complex that is essential (Kim *et al.*, 2007; Malinverni *et al.*, 2006). BamD has been shown to interact with BamA in such a way that an OMP binding-induced conformational change in BamD leads to a further conformational change in BamA to insert a new β-barrelled protein into the outer membrane (Lee *et al.*, 2018).

The primary function of BamE appears to be in detecting stress within the outer membrane (Konovalova *et al.*, 2016). This is achieved by threading RcsF (regulator of capsule synthesis) through the pore of β-barrelled proteins to allow "sensing" of the LPS through the detection of cationic antimicrobial peptides, or via the loss of negatively charged phosphates from the LPS (Konovalova *et al.*, 2014). This threading is performed whilst the OMP is being folded by the BAM complex. BamE knockouts produce OMP assembly phenotypes that do not vary

much from WT cells despite a secondary role in modulating BamA activity (Rigel *et al.*, 2012; 2013; Sklar *et al.*, 2007).

While the BAM complex can still function with the loss of BamB and BamE, a double knockout was found to be lethal (Sklar *et al.*, 2007). Interestingly, single residue mutations in the beta-barrel domain of BamA are able to rescue this phenotype meaning that substitution mutations in the β-barrel are able to compensate for non- or mis-functioning subunits and/or POTRA domains (Tellez & Misra, 2012). Furthermore, mutations in residues effecting electrostatic interactions between BamA and BamD disrupt the ability of the BAM complex to begin the process of folding nascent OMPs into the outer membrane (McCabe *et al.*, 2017). The impact of any mutations on the resultant folding and conformation of BamA, and therefore tail spike recognition, has not been explored. It is reasonable to hypothesise that these kind of mutations could have allowed the MC1061-MRL strains to survive despite a lack of phage adsorption. The compromised BAM complex may also explain their inability to survive long term at -80°C.

Functionally significant mutations in BamA have been predicted (Dwyer *et al.*, 2013). But, like all research into BamA function, this analysis neglects the phage receptor role in favour of interactions between component parts of the BAM complex and other intracellular proteins. The extracellular loops that are implicated in supporting phage adsorption have only been mentioned in reference to rescuing a BamD mutant (Lee *et al.*, 2018).

1.9 Aims of This Study

There are no other phages have been identified to exploit an essential protein as a recognition target on their hosts as this is a good way to guarantee a supply of hosts. How many other essential proteins are actually localised to the cell surface and allow phage binding is unknown. The presence of this tail spike gene in 70% of Stx phage genomes (Smith *et al.*, 2007), and the continued observation of *E. coli* outbreaks around the world suggests that Stx phages remain an immediate risk to human health because of the potential host range shift and consequent emergence of new pathogens. Outbreaks are usually caused by a failure of the food supply chain to ensure food is safe and robust control methods exist to monitor this supply worldwide, however, as new Stx-phage hosts are unable to be predicted, it is more likely a case of where and when, rather than if, the next outbreak occurs.

As Stx phages are the driving force behind the dissemination of *stx* into new host cells (Allison, 2007), the interaction between host and the phage tail should be characterised further with an ultimate view to understanding the molecular mechanisms involved. This could potentially determine if $\phi 24_B$ has the ability to efficiently broaden its host range for a wider and more rapid dispersal of converting phages in the environment, and may also inform of the potential species or strains that are likely to facilitate future STEC outbreaks. If the tail of $\phi 24_B$ were to spread beyond the Stx phages, more species could present a greater threat to human health than they currently do.

In order to achieve this, the primary goals of the thesis were to:

- Complete the generation and validation of an incomplete library of BamA mutants (McEwen, 2013).
- Use this library to produce *E. coli* clones that present different mutations to the extracellular surface and determine which mutants are incapable of supporting phage adsorption.
- Produce a chimaeric BamA mutant that would eliminate $\phi 24_B$ adsorption to *E. coli* MC1061 without impairing the viability of the cell.
- Develop a phage adsorption assay that is less prone to variability than standard adsorption assays, which can generate noisy data from a significant amount of sample handling.

CHAPTER 2. Materials and Methods

2.1 Strains, Bacteriophages and Plasmids

The bacterial strains, bacteriophages (Table 2.1.1) and plasmids (Table 2.1.2) used in this study are listed and described with relevant genotypes.

Table 2.1.1: The bacterial and bacteriophage strains used in this study.

Bacterial and bacteriophage strains	Description (incl. relevant genotype info)	Source
***E. coli* strains**		
MC1061	str. K-12 F⁻ λ⁻ Δ*(ara-leu)7697* [*araD139*]B/r Δ*(codB-lacI)3 galK16 galE15* e14⁻ *mcrA0 relA1 rpsL150*(Str^R) *spoT1 mcrB1 hsdR2*(r^-m^+)	(Casadaban & Cohen, 1980)
MC1061 $\phi24_B$::Kan	MC1061 containing an integrated $\phi24_B$::Kan prophage	(James *et al.*, 2001)
MC1061 Bam0	Abbreviated to Bam0. MC1061 with a chimaeric BamA and a Kan^R cassette added between *bamA* and *skp*.	This Study
DM1187-Rif	*His4 leuB ,proA/B thr1 ara-14 galK xyl-5 str31 ilv+s recA441 sfiA lexA51*. Bacteriophage host that can only enter lytic life cycle. Rif^R	(James *et al.*, 2001)
JM109 λ pir	*endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB⁺ Δ(lac-proAB) e14- [F' traD36 proAB⁺ lacI^q lacZΔM15] hsdR17($r_K^-m_K^+$)* and the Lambda phage *pir* gene that allows propagation of Lambda pir–dependent plasmid replicons like R6K in pKNG101	(James *et al.*, 2001)
Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ⁻⁻ Transformation host	Invitrogen™
***Pectobacterium* strain**		
Pectobacterium atrosepticum	The Sanger Institute sequenced the genome of this strain.	ATCC BAA-672
Bacteriophages		
$\phi24_B$::Kan	Mutant ϕE86654-VT2 phage. The toxin-encoding *Stx* operon replaced with Kan^R cassette.	(James *et al.*, 2001)

Table 2.1.2: The plasmids used or created during this study.

Plasmids		
pCR™-Blunt	pCR-series cloning vector (4600 bp) with lac promoter, T7, M13R and M13F primer sites; kanamycin and zeocin resistance; cloning for blunt-ended PCR products containing a l*acZ*-alpha *ccdB* death cassette (interrupted by insert).	Invitrogen, UK
pUC19	2,686 bp expression vector possessing a *lac* promoter. Amp^R, high copy number vector with MCS flanked by M13-F20 and M13R primer binding sites.	(Yanisch-Perron *et al.*, 1985)
pKNG101	Marker exchange suicide vector, *sacBR*, mobRK2, oriR6K, Sm^R	(Kaniga *et al.*, 1991)
pKD46	Red recombinase expression plasmid. Contains 2,154 nt (31088–33241) of phage λ (GenBank accession no. J02459).	(Datsenko & Wanner, 2000)
pUC4K	Used as a PCR template for Kan^R	(Roeder & Collmer, 1985)
pUC19-GFP-MUT2	A mutant of GFP (eGFP) engineered to use common *E. coli* codons cloned into pUC19.	(Cormack *et al.*, 1996)
pUC19-ϕR1D	pUC19 containing an insert of MC1061 containing the coding region of BamA. Amp^R	(Sergeant, 1998)
pBAD/myc-His C	Expression vector under araC promoter. Encodes a myc/His-tag on the CTD of the insert. Amp^R.	(Guzman *et al.*, 1995)
pCE1a	Kan^R. pCR-Blunt plasmid with the C-terminal domain of *P. atrosepticum bamA* insert containing the following *E. coli* epitopes substituted into the β-barrel domain: **4a‡**	(McEwen, 2013)
pCE1b - fixed	As pCE1a except with epitope: **4b**	This study
pCE1c	As pCE1a except with epitope: **6**	(McEwen, 2013)
pCE1d	As pCE1a except with epitope: **7a**	(McEwen, 2013)
pCE1e	As pCE1a except with epitope: **7b**	(McEwen, 2013)
pCE2a - fixed	As pCE1a except with epitopes: **4a 4b**	This study
pCE2b	As pCE1a except with epitopes: **4a 6**	(McEwen, 2013)
pCE2c	As pCE1a except with epitopes: **4a 7a**	(McEwen, 2013)
pCE2d	As pCE1a except with epitopes: **4a 7b**	(McEwen, 2013)
pCE2e - fixed	As pCE1a except with epitopes: **4b 6**	This study
pCE2f - fixed	As pCE1a except with epitopes: **4b 7a**	This study
pCE2g - fixed	As pCE1a except with epitopes: **4b 7b**	This study
pCE2h	As pCE1a except with epitopes: **6 7a**	(McEwen, 2013)
pCE2i	As pCE1a except with epitopes: **6 7b**	(McEwen, 2013)
pCE2k	As pCE1a except with epitopes: **7a 7b**	(McEwen, 2013)
pCE3a - fixed	As pCE1a except with epitopes: **4a 4b 6**	This study
pCE3b - fixed	As pCE1a except with epitopes: **4a 4b 7a**	This study
pCE3c - fixed	As pCE1a except with epitopes: **4a 4b 7b**	This study
pCE3d	As pCE1a except with epitopes: **4a 6 7a**	(McEwen, 2013)
pCE3e	As pCE1a except with epitopes: **4a 6 7b**	(McEwen, 2013)
pCE3f	As pCE1a except with epitopes: **4a 7a 7b**	(McEwen, 2013)
pCE3g - fixed	As pCE1a except with epitopes: **4b 6 7a**	This study
pCE3h - fixed	As pCE1a except with epitopes: **4b 6 7b**	This study

pCE3i - fixed	As pCE1a except with epitopes: **4b 7a 7b**	This study
pCE3k	As pCE1a except with epitopes: **6 7a 7b**	(McEwen, 2013)
pCE4a - fixed	As pCE1a except with epitopes: **4a 4b 6 7a**	This study
pCE4b - fixed	As pCE1a except with epitopes: **4a 4b 6 7b**	This study
pCE4c - fixed	As pCE1a except with epitopes: **4a 4b 7a 7b**	This study
pCE4d	As pCE1a except with epitopes: **4a 6 7a 7b**	(McEwen, 2013)
pCE4e - fixed	As pCE1a except with epitopes: **4b 6 7a 7b**	This study
pCE5 - fixed	As pCE1a except with epitopes: **4a 4b 6 7a 7b**	This study
pUC19 - KS	pUC19 containing a concatenation of Kan^R and skp sublconed from a ligation into pCR™-Blunt	This study
pUC19 - EEKS	pUC19 containing the EcEcaKanSkp fusion product	This study
pUE1a	pUC19-EEKS with the insert from the pCE equivalent subcloned into Eca using *Xcm* I and *Sty* I	This study
pUE1b	As pUE1a except with epitope: **4b**	This study
pUE1c	As pUE1a except with epitope: **6**	This study
pUE1d	As pUE1a except with epitope: **7a**	This study
pUE1e	As pUE1a except with epitope: **7b**	This study
pUE4a	As pUE1a except with epitopes: **4a 4b 6 7a**	This study
pUE4b	As pUE1a except with epitopes: **4a 4b 6 7b**	This study
pUE4c	As pUE1a except with epitopes: **4a 4b 7a 7b**	This study
pUE4d	As pUE1a except with epitopes: **4a 6 7a 7b**	This study
pUE4e	As pUE1a except with epitopes: **4b 6 7a 7b**	This study
pUE5	As pUE1a except with epitopes: **4a 4b 6 7a 7b**	This study
pUF5	pUC19 containing the full length chimaeric *bamA* gene with all 5 *E. coli* extracellular epitopes substituted into the β-barrel domain. Amp^R	This study

‡ See figure 3.1.1.1 for details of these epitopes.

2.2 Growth Media, Supplements and Buffers

Composition and sources of media and supplements used in this study are listed in table 2.2.1. Lysogeny Broth (LB) and LB agar (LBA) were prepared according to manufacturer's (Merck) instructions. Phage growth medium (PGM) was LB containing 10 mM $CaCl_2$, which is a required cofactor in the phage adsorption process. Terrific Broth, TB, was used when higher-density cultures of *E. coli* were required for low-yield or large-scale plasmid purifications. Tryptone soy (TS) broth and TS agar plates were used to grow *Pectobacterium atrosepticum*. Antibiotics (Table 2.2.1) were added to media as required.

Table 2.2.1: Media and supplements used in this study

	Composition (final concentrations)
Media	
Lysogeny broth (LB)	2.5% LB broth (Merck, Hertfordshire, UK)
LB agar	LB + 12 g L^{-1} agar-agar (VWR, UK)
Phage Growth Medium (PGM)	LB + 10 mM $CaCl_2$ (BDH Lab supplies, UK)
Salt-free LB	10 g L^{-1} Tryptone (Lab M, UK); 5 g L^{-1}, Yeast extract (Merck, UK) (NaOH to pH 7.0)
Salt free LBA	Salt free LB + 12 g L^{-1} agar-agar (VWR, UK)
Bottom agar	PGM + 15 g L^{-1} Difco, Agar granulated (Difco, UK)
Top agar	PGM + 4 g L^{-1} Difco, Agar granulated (Difco, UK)
Tryptone Soy Broth (TSB)	30 g L^{-1} Tryptone Soy Broth (Lab M, UK)
Tryptone Soy Agar (TSA)	TSB + 12 g L^{-1} agar-agar (VWR, UK)
Terrific Broth	12 g L^{-1} tryptone, 24g L^{-1} yeast extract, 4 mL glycerol, 10% (v/v) phosphate buffer (0.17 M KH_2PO_4, 0.72 M K_2HPO_4)
SOC medium	28 g L^{-1} SOB broth (Melford, UK); 20 mM glucose (VWR, UK)
Media Supplements	**Working Concentrations**
Ampicillin	100 µg mL^{-1} ampicillin sodium salt (Melford, UK)
Chloramphenicol	12.5 µg mL^{-1} chlorampenicol (Sigma, UK)
Kanamycin	50 µg mL^{-1} kanamycin monosulfate (Melford, UK)
Norfloxacin	1 µg mL^{-1} norfloxacin (Melford, UK)
Rifampicin	300 µg mL^{-1} rifampicin (Melford, UK)
Spectinomycin	50 µg mL^{-1} spectinomycin dihydrochloride pentahydrate (Sigma, UK)
Streptomycin	100 µg mL^{-1} streptomycin sulfate (Melford, UK)
Tetracycline	20 µg mL^{-1} tetracycline (Sigma, UK)
Sucrose	5% (w/v) sucrose (Melford, UK)
$CaCl_2$	10 mM $CaCl_2$ (BDH Lab supplies, UK)
ITPG	100 mM IPTG (Melford, UK)
X-Gal	40 µg mL^{-1} X-gal (Melford, UK)
Arabinose	0.0002% (w/v) – 0.2% (w/v) L-arabinose (Sigma, UK)

Table 2.2.2: Buffers used in this study

	Composition (final concentration)
General Buffers	
TE	25 mM Tris-HCl (pH 7.5); 10 mM EDTA
TAE	25 mM Tris (to pH 7.5 with acetic acid); 10 mM EDTA,
TBT	100 mM Tris-HCl (pH 7.5); 100 mM NaCl; 10 mM $MgSO_4$
SM	50 mM Tris-HCl (pH 7.5); 100 mM NaCl; 8 mM $MgSO_4$
SC	50 mM Tris-HCl (pH 7.5); 100 mM NaCl; 10 mM $CaCl_2$
TfBI	30 mM KOAc; 50 mM $MnCl_2$; 100 mM KCl; 10 mM $CaCl_2$; 15% (v/v) glycerol
TfBII	10 mM MOPS (pH 7.0); 75 mM $CaCl_2$; 10 mM KCl, 15% (v/v) glycerol
Restriction enzyme dilution buffer	10 mM Tris-HCl (pH 7.5); 100 mM KCl; 1 mM EDTA; 1 mM DTT; 50% (v/v) glycerol; 0.2 mg mL^{-1} BSA
Plasmid extraction buffers	
P1	50 mM Tris.HCl, pH 8.0; 10 mM EDTA; 100 µg mL^{-1} RNAseA
P2	200 mM NaOH; 1% (w/v) SDS
N3	4 M guanidine hydrochloride; 0.5 M potassium acetate, (to pH 4.2 with acetate)
PE	80% (v/v) ethanol; 10 mM Tris-HCl (pH 7.5)
QBT	50 mM MOPS (pH 7.0); 750 mM NaCl; 0.15% (v/v) Triton X-100; 15% (v/v) isopropanol
Southern Blot buffers	
20x SSC	3 M NaCl; 0.3 M sodium citrate (pH 7.0)
Depurination buffer	0.25 M HCl
Denaturation buffer	0.5 M NaOH, 1.5 NaCl
Neutralisation buffer	0.5 M Tris-HCl (pH 7.5), 3 M NaCl

2.3 Procedures

2.3.1 PCR

PCR amplifications were typically done over 30 cycles with temperatures and duration of melting and elongation steps and the duration of all steps and reagent concentrations were calculated based on the data sheets for the polymerase used: myTaq Red Mix (Bioline, UK), Phusion (Thermo Fischer Scientific, UK), Q5 (NEB, UK), or Vent (NEB, UK), as appropriate. Annealing temperatures were based on the primers listed in Table 2.3.1. Resultant DNA products were analysed on 1.2% (w/v) TAE agarose gels containing 2 µL Midori green (Nippon Genetics, Europe) per 100 mL agarose.

Table 2.3.1: DNA primers used for PCR in this study.

Restriction enzyme recognition sequences are underlined.

Primer	Sequence (5'- 3')	Amplification Template	Calculated Tm (°C)
Used for fusion of products to create Bam0 by homologous recombination			
EcEca 5' X site fwd	GGTTGATGTCGTCTACAAGG	MC1061 gDNA	59.8
EcEca 3' X site rev	CACGTCTGTACCATAACTCTTG	MC1061 gDNA	58.8
EcEca fusion fwd	GTTATGGTACAGACGTGACCTTGGGCTTCCCG	*P. atrosepticum* gDNA	78.9
5' fwd loop Eca	CCAACGTGAGCTATGGCG	*P. atrosepticum* gDNA/BamA	65.5
3' rev loop Eca	CTACCAGGTTTTGCC	*P. atrosepticum* gDNA/BamA	51.4
5' forkaneco	AAGAATTCAGCCACGTTGTGTCT	pUC4K	65.8
3' revkanskp	GATCGCCTAAAGTCAAAGGTGTTCCTG	pUC4K	70.8
Skp 5' for	ATGACTTTAGGCGATC	MC1061 gDNA	49.9
3' rev SkpEcoRI	AAGAATTCGTAAGTCAGGTAG	MC1061 gDNA	58.0

Primer	Sequence (5'- 3')	Amplification Template	Calculated Tm (°C)
Used to introduce *E. coli* sequences to *P. atrosepticum* BamA			
5' rev loop IIa	CTGATCAGAGGTGCTCGGATGAACCCC GACAGAGATCC	*P. atrosepticum* BamA	84.8
3' fwd loop IIa	GAGCACCTCTGATCAGGATAACAGCTT CAAAGCGAACGA	*P. atrosepticum* BamA	81.5
3'rev loop IIb	CGTGAAATCGTTCGCTTTG	*P. atrosepticum* BamA	64.3
5' fwd loop IIb	TTCAAAGCGAACGATTTCACGTTCAAC TATGGATGGTCATAACAACC	*P. atrosepticum* BamA	82.5
5' rev loop IV	GTAATCATAGTCCGGATCGCCAACAAG GTTCGC	*P. atrosepticum* BamA	70.7
3' fwd loop IV	CCGGACTATGATTACGAATGTGCGACT CAGGATTCAACCAATATGG	*P. atrosepticum* BamA	83.1
5' rev loop Va	GCTGGAATCTTCCCAGTTGGTATCCC	*P. atrosepticum* BamA	72.4
3' fwd loop Va	TGGGAAGATTCCAGCCTGAAAGCAGGT GTACC	*P. atrosepticum* BamA	78.9
5' rev loop Vb	GATATTGCTTGGATCGCCGTAATCCGG TACACC	*P. atrosepticum* BamA	78.5
3' fwd loop Vb	GATCCAAGCAATATCCGCGTTTCCAGT GGTATCGC GTTGC	*P. atrosepticum*BamA	85.4
5' fwd IIb fix	GGATGGTCATACAACAACCTG	*P. atrosepticum* BamA	57.9
3' rev IIb fix	CAGGTTGTTGTATGACCATCC	*P. atrosepticum* BamA	57.9
Used to sequence inserts cloned into plasmids with M13 primer sites			
M13 fwd (-20)	GTAAAACGACGGCCAG	pCR™-Blunt/pUC19	58.9
M13 rev	CAGGAAACAGCTATGAC	pCR™-Blunt/pUC19	52.0
Used to detect the suicide vector, pKNG101.			
pKNG101_F	GATGATACGCACTGAGAAGCC	pKNG101	63.8
pKNG101_R	GACGATGCCAAAGCGCTACACC	pKNG101	68.8
Used to clone ϕ24B tail and tail-associated genes			
FOR TSP	TGCGCAGCATTATCAGAG	$\phi 24_B$ phage/Lysogen gDNA	53.6
Rev tsp	CATTCTCATCTGACTCC	$\phi 24_B$ phage/Lysogen gDNA	50.4
For hypotail	GTCCGAAAGGAGATAAC	$\phi 24_B$ phage/Lysogen gDNA	50.4
Rev hypotail	GATTCTTTACTATCCAG	$\phi 24_B$ phage/Lysogen gDNA	45.5
For Gene 39	TTCTCGTTGCTACAGCG	$\phi 24_B$ phage/Lysogen gDNA	52.8
Rev Gene 39	GTCACAAATGCGTGAC	$\phi 24_B$ phage/Lysogen gDNA	49.2
For Gene 40	AAGTGTGGTGAGACTTC	$\phi 24_B$ phage/Lysogen gDNA	50.4
Rev Gene 40	GCTGTAAATCTGTCCAGG	$\phi 24_B$ phage/Lysogen gDNA	53.7

Used to construct GFP gene fusions			
NcoI TSP for	GGACCATGGGTGTTGTTGTTTCGGGGAC	$\phi 24_B$ phage/Lysogen gDNA	69.5
SZB GFP-for	TGACCATGGGAGAAGAACTTTTCACT	pUC19-GFP-MUT2	61.6
SM GFP-TSP for	TGCCGGCGCAGGAAGTGTTGTTGTTTCGG	$\phi 24_B$ phage/Lysogen gDNA	70.9
SM2 GFP-TSP rev	TTCCTGCGCCGGCACCAGCTTTGTACAATTCATC	pUC19-GFP-MUT2	71.9
SZB SalI-TSP-rev	TCACAGTCGACATGCTTTCCGCAACC	$\phi 24_B$ phage/Lysogen gDNA	66.4
NotI-GFP-TSP for	GGTGCGGCCGCCGGCGCAGGAAGTGTTGTTGTTTCGG	$\phi 24_B$ phage/Lysogen gDNA	89.8
NotI-GFP-TSP rev	GCCGGCGGCCGCACCTTTGTACAATTCATCC	pUC19-GFP-MUT2	73.5
BamB for	ATGCAATTGCGTAAATTACTG	MC1061 gDNA	52.0
BamBp for	GGCTTCTTTGGAACTTGCGCAG	MC1061 gDNA	62.1
BamB-GFP for	TGCCGGCGCAGGAAGTCCCGGAGAAGAACTTTTCAC	pUC19-GFP-MUT2	71.7
BamB-GFP rev	TCCTGCGCCGGCACGTGTAATAGAGTACACG	MC1061 gDNA	72.1
NotI-GFP-TSP for	GGTGCGGCCGCCGGCGCAGGAAGTGTTGTTGTTTCGG	$\phi 24_B$ phage/Lysogen gDNA	89.8
NotI-GFP-TSP rev	GCCGGCGGCCGCACCTTTGTACAATTCATCC	pUC19-GFP-MUT2	73.5
SM GFP-rev	CAGTTATTTGTATAGTTCATCCATGC	pUC19-GFP-MUT2	58.5
BamA amplification/modification			
PhiR1D_BamA_fwd	GATCAAGGGCGGACCGGTATCC	pUC19-ϕR1D	65.8
5' NcoI BamA Fwd	CGTCCATGGCGATGAAAAAGTTGC	MC1061/Bam0 gDNA	62.7
5' BamAp2 Fwd	CGAAATAGCAGCCAATTCGATAGCTG	MC1061/Bam0 gDNA	63.2
SM_BamA_for	ATGGCGATGAAAAAGTTGCTC	MC1061/Bam0 gDNA	55.9
SM_BamA_rev	TTACCAGGTTTTACCGATG	MC1061/Bam0 gDNA	52.4
3' SalI BamA Rev	TACGTCGACCCAGGTTTTGCCAATG	MC1061/Bam0 gDNA	64.6
FLbam5 rev	CTACCAGGTTTTGCCAATGTTAAACTGG	MC1061/Bam0 gDNA	63.7
NcoI BamA over	GCTAACAGGAGGAATTAACCATGGCGATGAAAAAGTTG	MC1061/Bam0 gDNA	69.5
BamAp2 over	GCTAACAGGAGGAATTAACGAAATAGCAGCCAATTCGATAGCTG	MC1061/Bam0 gDNA	72.2
SalI BamA over	GGAGACCGTTTAAACTCAGTCGACCCAGGTTTTGCCAATG	MC1061/Bam0 gDNA	73.5

2.3.2 Restriction enzyme digests

DNA (500-1000 ng) was added to appropriate buffers and the stated number of enzyme units in Table 2.3.2.1 were added for the respective restriction enzymes:

Table 2.3.2.1: Restriction enzymes used in this study

Enzyme	Company	Units used per reaction (U)
*Eco*R I (FastDigest)	Thermo Fisher Scientific, UK	0.5 µL †
*Eco*R I	Promega, UK	6
*Eco*R I - HF	NEB, UK	10
Kpn I (FastDigest)	Thermo Fisher Scientific, UK	0.5 µL †
Nco I	Thermo Fisher Scientific, UK	5
Pst I (FastDigest)	Thermo Fisher Scientific, UK	0.5 µL †
Sal I	Thermo Fisher Scientific, UK	5
Sty I - HF	NEB, UK	10
Xcm I	NEB, UK	2.5

† FastDigest™ enzymes are not assigned units, so volume used per reaction is reported instead.

A final volume of 10 µL was used for restriction enzyme digestion except when this was increased to 20 µL to ensure the final glycerol content was no higher than 5% (v/v). Samples were incubated at 37°C for 45 min (15 minutes for FastDigest enzymes) and visualized on a 1.2% (w/v) TAE agarose gel. Partial digests used 1 µg DNA, 10 U of *Eco*R I (Promega, UK) and were left at room temperature for 1 min before running on a 1.2% (w/v) TAE agarose gel without heat inactivation. The desired fragment was excised using the PCR and Gel Kit (Bioline, UK) following manufacturer's instructions but eluting DNA in dH_2O.

2.3.3 Plasmid Purification

Plasmid purification buffers (Table 2.2.2) based on the plasmid purification kits manufactured by QIAGEN (UK) and Bioline (UK) were prepared, and DNA binding columns from Bioline kits were reused after washing twice with dH_2O, once with QBT buffer followed by storage at 4°C in 1 M HCl. Columns were replaced after either 10 uses, 3 months or if visible disruption to the membrane was observed. Column reuse and homemade buffers were validated by comparing plasmid yields to those produced using brand new columns and kit buffers where no difference was observed by either agarose gel analyses or spectrophotometric (Nanodrop 1000) analysis. DNA was eluted in dH_2O.

Plasmids described as "high-copy number" were purified using the Plasmid Mini-Kit (Bioline) following the manufacturer's instructions, but using buffers detailed in Table 2.2.2 (N3 and PE buffers were used instead of P3 and PWB1, respectively) and DNA was eluted in 30 µL dH_2O.

A higher volume protocol was adapted for purification of low copy-number plasmids (e.g pKNG101): Cells from an overnight culture (grown in 100 mL TB) were resuspended in P1 buffer (4 mL). P2 buffer (4 mL) was added and incubated at room temperature (RT) for 5 minutes. N3 buffer (5.5 mL) was added to precipitate proteins that were pelleted with genomic DNA at 20,000 *g* for 10 minutes (Eppendorf 5810R, rotor A-4-81). The supernatant was passed through a mini-kit spin column (Bioline, UK). Samples were washed twice with PE buffer, and eluted in 2 x 25 µL dH_2O for maximum yield.

2.3.4 Plasmid transformation

2.3.4.1 Heat shock transformation

Mid-log phase cells (OD_{600} ≈ 0.6) were made competent following their harvest by centrifugation at 5000 *g* for 5 minutes using TfBI then TfBII buffers (Table 2.2.2). Briefly, the recovered cells were resuspended in ice cold TfBI for 1 hour with regular agitation, recovered by centrifugation at 5000 *g* for 5 minutes, resuspended in ice-cold TfBII for a further 30 minutes and then divided into 100 µL aliquots for use. Plasmid DNA (50 ng – 500 ng) was added to an aliquot of competent cells and kept on ice for 10 minutes. The mixture was heat shocked at 42°C for 45 seconds and returned immediately to ice. Pre-warmed LB (700 µL) was added, and cells were allowed to recover at 37°C with shaking for 1 hour before spreading 80 µL onto selective LBA plates.

2.3.4.2 Transformation by electroporation

Mid-log phase cells (500 mL of OD_{600} ≈ 0.6) were chilled on ice for 20 minutes before being centrifuged at 4000 *g* for 15 minutes at 4°C. The supernatant was removed by pipette and cells were resuspended in 500 mL ice-cold 10% (v/v) glycerol. Cells were centrifuged at 4000 *g* for 15 minutes at 4°C and resuspended a further 3 times in decreasing volumes of ice-cold 10% (v/v) glycerol (250 mL then 20 mL then 2 mL). The final 2 mL of concentrated cells (~1-3 x 10^{10} cells mL^{-1}) was used immediately or frozen in 50 µL aliquots for storage at -80°C. DNA (10 ng – 100 ng) was mixed with a 50 µL aliquot of competent cells on ice for 5 minutes before the mixture was transferred to an ice-cold 0.2 mL electroporation cuvette. The micropulser (Biorad, UK) was set to the automatic

"Ec2" mode (for *E. coli* cells in 0.2 mL cuvettes) and the sample was pulsed. 1 mL of SOC medium was immediately added to the electroporated mixture and was shaken at 37°C for 60 minutes to allow recovery before spreading 500 µL on selective LBA plates.

2.3.5 Cloning PCR products

2.3.5.1 Cloning

PCR products were cloned into either the pCR™-Blunt vector with the Zero Blunt PCR Cloning Kit or the pCR™ 2.1 TA cloning® kit depending on the ends produced by the polymerase used, following the manufacturer's (both Invitrogen, UK) instructions.

2.3.5.2 Sub-cloning

Sub-cloning was accomplished by ligation following restriction digests of both the insert and the plasmid vector. Inserts were separated from cloning plasmids using appropriate restriction enzymes. The insert fragment was recovered from a 0.6% (w/v) agarose gel using the Gel and PCR Kit (Bioline, UK); the manufacturer's instructions were followed but DNA was eluted in dH_2O. Target plasmids were digested using identical restriction enzymes to generate compatible cohesive ends. The plasmid backbone was recovered from a 0.6% (w/v) agarose gel using the same method as the insert. If only one enzyme was used to cut the plasmid then it was treated with Antarctic phosphatase (NEB, UK) to reduce the likelihood of vector re-ligation. The purified insert was ligated into the prepared vector using T4 DNA ligase (Fermentas – now Thermo Fisher

Scientific, UK). Ligated vector was transformed into competent Top10 *E. coli* and spread onto LBA containing ampicillin and kanamycin. Colonies were screened for inserts using appropriate primers (Table 2.3.1).

2.3.6 Southern Blot Analyses

Genomic DNA (5 µg), extracted from WT MC1061 cells and putative chromosomal BamA mutants of MC1061 using a Genomic DNA kit (Bioline. UK), was digested with *Kpn* I or both *Kpn* I and *Eco*R I according to manufacturer's (Thermo Fischer Scientific, UK) instructions in a final volume of 20 µL. Digestions were run through a 1% (w/v) TAE agarose gel at 100 V for 45 minutes alongside Hyperladder 1kb (Bioline, UK). The gel was imaged under UV light next to a fluorescent ruler to allow determination of the apparent molecular weight of bands. The gel was subjected to depurination for 30 minutes, denaturation for 2 x 20 minutes, and neutralised for 2 x 20 minutes – buffer compositions for these steps are in table 2.2.2. DNA was transferred to a nylon membrane overnight using the capillary action generated by 20x SSC through a Whatman 3MM paper wick (Southern, 1975). DNA fragments were cross-linked to the membrane with UV light (120 mJ) in a Stratalinker 2400 (Stratagene - now Agilent, California, USA). DIG-labelled probes were synthesised with the PCR DIG probe synthesis Kit (Roche, UK) using Phusion polymerase (NEB, UK). Lanes were probed for *E. coli bamA* and the kanamycin resistance gene (*aphA*) following the DIG DNA labeling and detection kit instructions (Roche, UK). Hybridisation of the probes was detected through the use of a chromogenic substrate, BCIP/NBT (Sigma, UK)

to identify which bands the antibody bound to. Membranes were imaged in the ImageQuant4000 (GE Healthcare, UK).

2.3.7 Phage stock generation

2.3.7.1 Small scale phage stock generation

The lysogen, MC1061 $\phi 24_B$::Kan, was grown to mid-log phase ($OD_{600} \approx 0.45$) in phage growth medium (PGM). Norfloxacin was added to a final concentration of 1 μg mL^{-1} and the culture was incubated at 37°C for 1 hour to induce the SOS stress response in the lysogen, causing the autocatalysis of the repressor protein and induction of the prophage into the lytic replication cycle (James *et al.*, 2001). The induced culture (1 ml) was subcultured into 9 mL fresh PGM, thereby diluting norfloxacin 10-fold. The cells were allowed to recover for 2 hours before being filtered (0.22 μm, Millipore) into a plastic universal and stored at 4°C. The phage titre of each stock was determined by plaque assay (2.3.9).

2.3.7.2 Large scale phage stock generation

Larger phage stocks were generated using either a scaled up version of the small scale phage stock generation method (up to 3.5 L) using the same concentrations and ratios for all reagents or by the plate-scraping method. The plate scraping method used 10 mL phage stock generated by the small scale generation protocol, which was then used to produce 50-100 plates containing semi-confluent plaques following the plaque assay method (2.3.9). After overnight

incubation, PGM (3 mL) was added to the surface of each plate with semi-confluent plaques and a sterile spreader was used to disrupt the top agar layer and facilitate the removal of material to a pooled sample of top agar and PGM, both containing $\phi 24_B$ phages. The pooled material was transferred to 250 mL centrifuge tubs and shaken vigorously to break the agar pieces and release as many phages as possible into solution. The resulting debris was centrifuged at 1000 *g* for 10 minutes to remove the agar and the supernatant was collected. NaCl was added the supernatant to a final concentration of 1 M, into which 10% (w/v) PEG 8000 was dissolved. Phages were pelleted by centrifugation at 10,000 *g* at 4°C for 20 minutes. The pellet was resuspended in SC buffer. Phage stock was cleaned up by harvesting the aqueous phase from a phenol chloroform extraction three times using equal parts phenol:chloroform:isoamyl alcohol (25:24:1) and SC buffer. If required, phages were concentrated further by adding NaCl to 1 M and 10% (w/v) PEG 8000, and centrifugation at 10000 *g* at 4°C for 20 minutes before final resuspension in 100 µL – 500 µL SC buffer. As before, a further 3 rounds of phenol chloroform extraction was performed to clean up the sample. The concentrated phage stock was quantified by plaque assay (2.3.9).

2.3.8 Plaque assay

Phage stock was serially diluted to 10^{-6} in phage growth medium (50 µL into 450 µL). Mid-log phase (OD_{600} ≈ 0.45) DM1187 cells (100 µL) were added to each dilution. The total volume (550 µL) of each dilution was transferred to separate 3.5 mL aliquots of top agar and mixed briefly by pipette before being poured onto pre-set bottom agar plates containing the same additives as the top agar.

Plates were ventilated by using a heated wire loop to introduce 2 x 2mm slots on either side of the petri dish before incubation at 37°C overnight, and the titres of plaque forming units (PFU) were quantified by counting the plaques on agar plates possessing 20-200 plaques (James *et al.*, 2001).

2.3.9 Transmission Electron Microscopy – negative staining

In a laminar flow hood, copper grids (300 and 400 hexagonal mesh) were coated with a particle-free 0.3% (w/v) Pioloform (polyvinyl butyral) film. These were desiccated overnight. Phages (5 µL of ~1 x 10^9 stock solution) were added to treated grids for 30 seconds, after which the excess was wicked off using filter paper. Grids with samples were initially stained with 4% (w/v) uranyl acetate (UA) for 10 seconds before washing once by dripping MilliQ H_2O across the grid. Subsequent samples were stained with 2% (w/v) UA as the 4% stain produced overly dark images. Grids were dried completely before being analysed on an FEI Tecnai 12 Bio-twin electron microscope.

2.3.10 Fluorescence microscopy

Phages (500 µL of 1 x 10^9 stock) were incubated overnight at 4°C in PGM containing 1 µL of 10,000x concentrated SYBR-gold (Life Technologies, UK), which bound to phage DNA. Unbound SYBR-gold was removed from the sample by dialysis in 2 x 2 L PGM using 100 kDa MWCO CE membrane to retain SYBR-gold stained phages ready for imaging. Four-compartment glass bottomed dishes (Greiner, UK) were treated with poly-L-lysine for 30 minutes at room

temperature to provide a surface onto which *E. coli* could be attached. Mid-log phase cells (OD_{600} ≈ 0.40-0.50) were concentrated five-fold by centrifugation for 60 s at 11,000 *g* and resuspended in 0.2 volumes of PGM. Concentrated cells (50 µL) were added to the treated glass bottomed plates and allowed to rest at room temperature for 20 minutes to settle onto the glass surface. Samples were washed twice with PGM to remove unattached cells with a final 500 µL PGM added to the glass-bottomed dish before imaging on the fluorescence microscope. Once the dish was loaded onto the microscope and the correct focus was set up, images were taken automatically every 10 s across all required wavelengths (Table 5.2.1.1) for a total of 40 minutes. Phages were added to the medium after the 30 s image was taken at 2 x 10^5 PFU mL^{-1}. Later preparations of cells were treated with the DAPI-based NucBlue® Live ReadyProbes® Reagent according to manufacturer's instructions (Thermo Fisher Scientific, UK) to stain bacterial DNA blue. A propidium iodide stain (Life Technologies, UK) was also added (either at the beginning or the end of the experiment) to detect the dead cells. The method progression is detailed in Chapter 5.

2.3.11 Rendering Protein Images

Co-ordinates for 3D protein images were downloaded using 4 character alphanumeric codes from the Protein Data Bank (PDB) within molecular graphics software, QtMG (a version of CCP4MG released for use with OSX). The model was manipulated to display the desired details and angle. An image was rendered using the "normal" method and the following settings: 6.00 x 4.35 inches, 250 dpi, scaled font and image sizes, set transparent background.

Chapter 3: Generating a Library of Isogenic Mutants to Investigate the BamA-Phage Interaction.

3.1 Restrictions around generating BamA mutants

Recognition targets for other bacteriophages are known to undergo phase or antigenic variation allowing the host to lose support for phage adsorption (van der Woude & Bäumler, 2004). Phase variation is a temporary change in the volume of expression to allow removal of the molecule from the surface e.g. the lambda phage receptor, LamB, is only expressed when the cell is in the presence of maltose (Chapon, 1982). Antigenic variation is a mutation that ultimately leads to a shape change on the surface used by the phage to recognise a host cell (Chaturongakul & Ounjai, 2014) and therefore causes a loss of complementarity. This causes a biological arms race between the phage and the host cell known as the Red Queen hypothesis where the phage must constantly adapt to mutations of the host cell (Paterson *et al.*, 2010).

Interestingly, BamA is the only essential phage target that is currently known. Consequently, the possibility of using phase variation as a defence strategy is unavailable and any antigenic variation is extremely limited – and would be restricted mostly to the extracellular loops (Ruhe *et al.*, 2013b; Smith *et al.*, 2007) These restrictions mean that instead of knocking out the gene, a genetic chimaera would need to be made in order to maintain the protein folding

function while simultaneously removing the support for phage adsorption by substituting epitopes on the extracellular side of the BamA β- barrel.

3.1.1 Differences between *Enterobacteriaceae* BamA β-barrels

Amino acid level sequence comparison of β-barrel domains identified just five specific regions of variation in the primary protein structure of BamA that correlated with a loss of support for phage adsorption (Fig 3.1.1.1). It should be noted that there is no variation elsewhere in the aligned sequences. This conclusion is supported by the analysis of Ruhe *et al.* (2013), which compared 41 *Enterobacteriaceae* BamA sequences to show that genetic variation in BamA sequences occurs almost exclusively within these three extracellular loops (Fig. 3.1.1.2). This shows that, in theory, the only variability that could occur in BamA tertiary structure would be associated with the extracellular surface, which is exactly where that antigenic variation would need to manifest to allow different phage adsorption phenotypes between closely related species. Structural data obtained for BamA of multiple species (Noinaj *et al.*, 2013) enabled mapping of regions expected to support adsorption onto the *E. coli* BamA structure (Fig. 3.1.1.3). The model shows epitopes 4a, 6, and 7a are exposed on the cell surface to the extracellular environment in this conformation. Conversely, loops 4b and 7b are less likely to affect adsorption, as they would be within the transmembrane region, which is inaccessible to the incoming phage tail spike protein. However, it is still possible these epitopes have a role in a second step of adsorption, as lambda phage has (Werts *et al.*, 1994), if BamA undergoes a binding-induced conformational change before DNA ejection occurs.

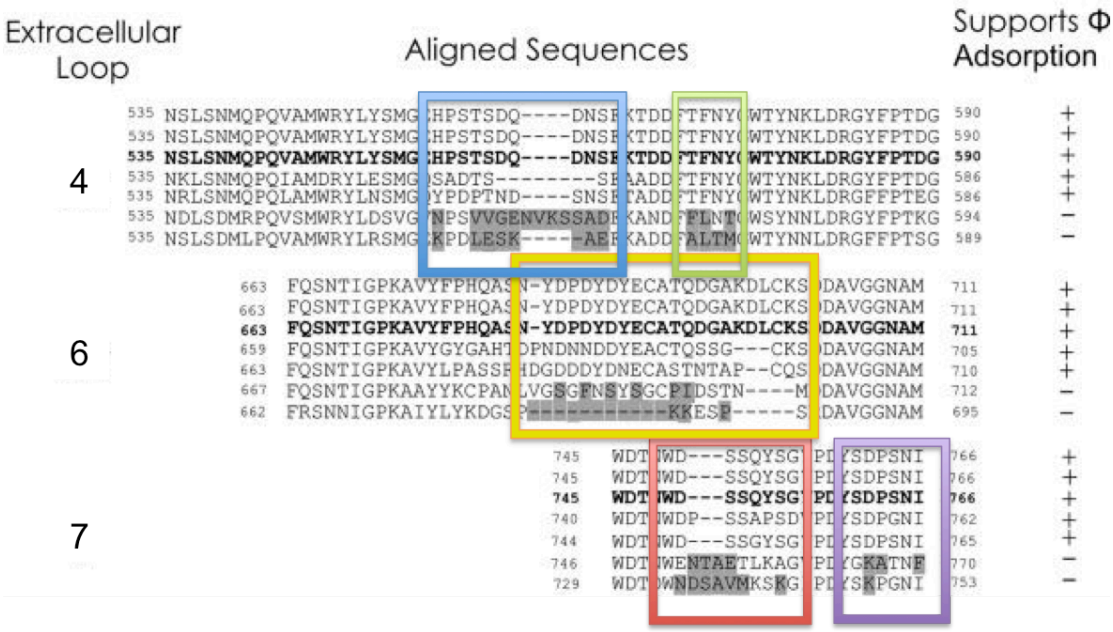

Figure 3.1.1.1: Sequence alignments of BamA orthologues from seven different species.

From top to bottom sequences are from *S. flexneri* PT2a and PT6, *S. sonnei* PT36, ***E. coli* MC1061**, *S. enterica* serovar Cholesaesuis strain SC-B67, *C. rodentium* strain ICC168, *Pectobacterium atrosepticum* and *P. luminescens* subsp. *Laumondii* strain TTO1. Substitutions in type of R-group in amino acid residues (e.g. acidic to basic) between those sequences that support phage adsorption and those that do not are highlighted in grey. Regions of interest from these substitutions have been identified as loop 4a (blue), 4b (green), 6 (yellow), 7a (red) and 7b (purple). Adapted from Smith *et al.*, (2007).

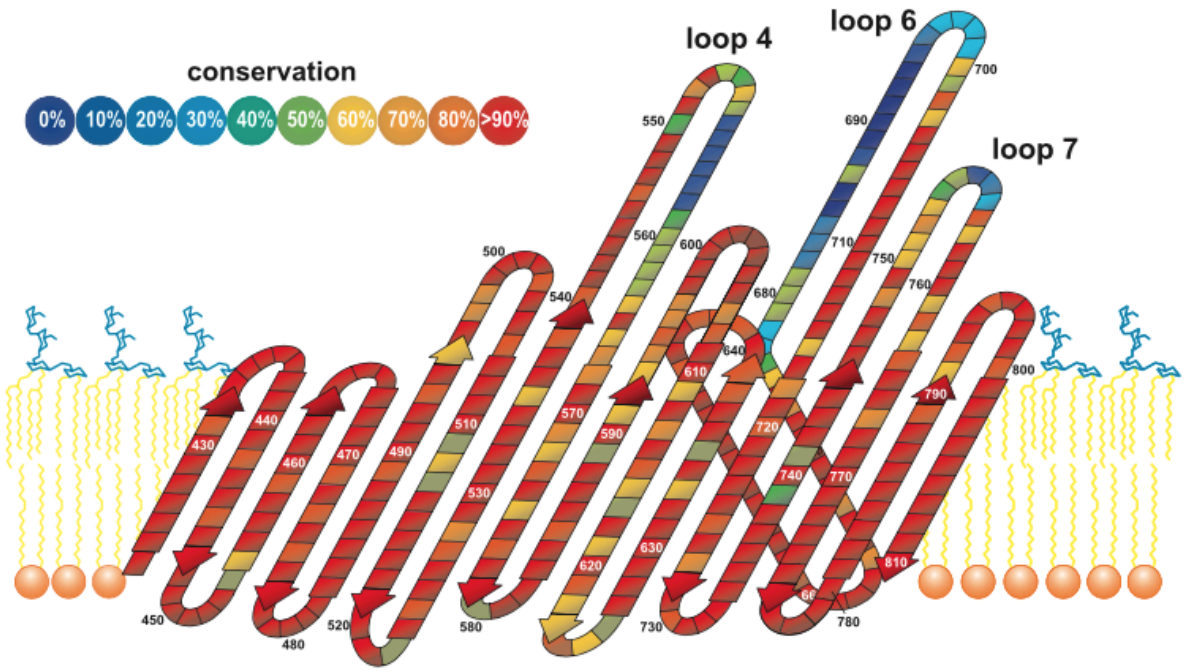

Figure 3.1.1.2: Percentage conservation of the β-barrel domain of BamA among 41 species of the *Enterobacteriaceae* family.

Level of conservation of each residue of the β-barrel is colour coded in ranges from 0%-10% (dark blue) to 90%- 100% (red). The vast majority of the transmembrane β-sheets are >80% conserved. The only residues that are <50% conserved are limited to Loops 4, 6, and 7. Source: (Ruhe *et al.*, 2013b).

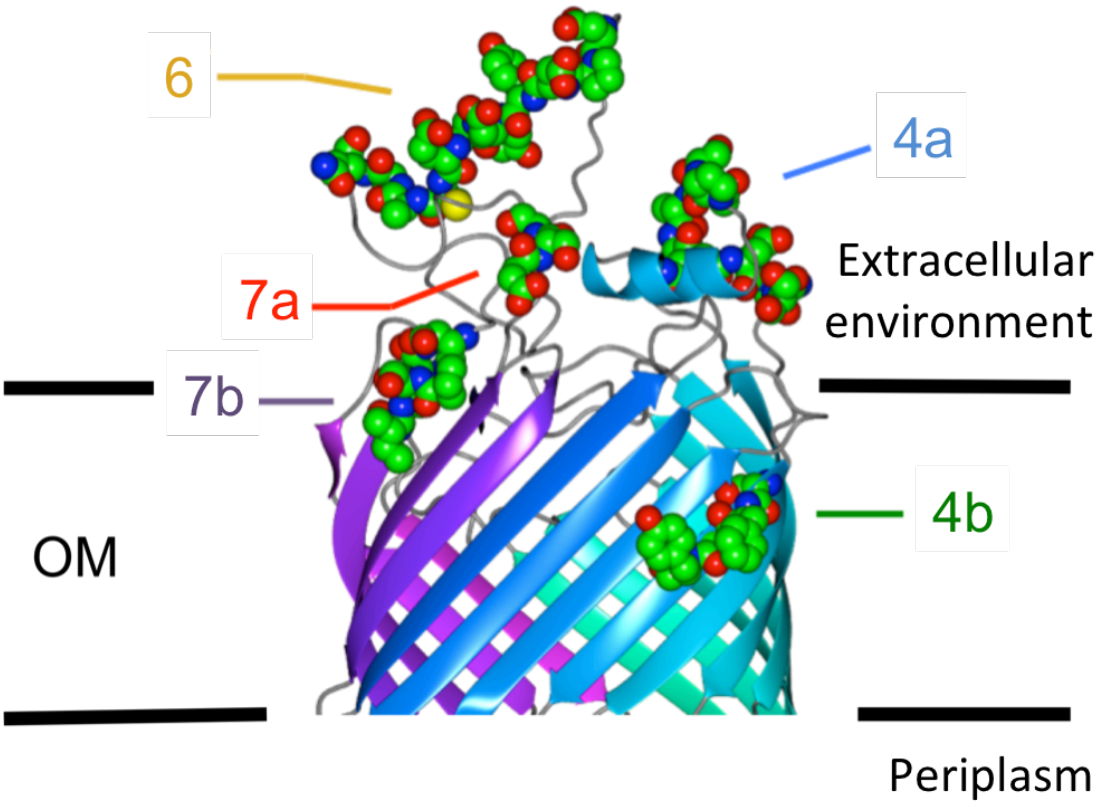

Figure 3.1.1.3: 3D model of *E. coli* BamA with the epitopes to be substituted. Model of *E. coli* BamA presented as ribbons with the amino acids that make up the five epitopes that have been substituted into the *P. atrosepticum* β-barrel represented in spheres generated using supplemental data from Noinaj *et al.* (2013).

Loop 6 has a highly conserved $_{641}$RGF$_{643}$ motif that is critical for maintaining the protein folding function (Leonard-Rivera & Misra, 2012). This epitope was crystallised inside the BamA pore in the closed structure (Noinaj *et al.*, 2013) and Loop 6 must be inside the pore to begin the protein folding and insertion process (Wzorek *et al.*, 2017). Therefore, it is plausible that phage adsorption itself is also conformation-dependent.

3.2 Results

3.2.1 Making a chimaeric BamA protein

In order to definitively prove that BamA supports phage adsorption it was necessary to make a mutant strain that was not susceptible to $\phi24_B$ adsorption. However, the essential nature of BamA means that it cannot be knocked out, as any loss of function would lead to a non-viable cell. Instead, to maintain function while still creating a mutant phenotype, a chimaeric *bamA* was constructed by combining POTRA domains from the BamA orthologue of *E. coli* and the β-barrel of *Pectobacterium atrosepticum* (formerly known as *Erwinia carotovora* subsp. *atroseptica* SCRI 1043).

The β-barrel from *P. atrosepticum* was chosen because it has been shown that this exact strain is unable to support phage adsorption (Smith *et al.*, 2007). In addition, there have also been several examples of expressing functional periplasmic and outer membrane proteins from *P. atrosepticum* in *E. coli* due to

similarities in secretion systems and envelope physiology that makes *P. atrosepticum* a sensible choice for this experiment (Barabote *et al.*, 2003; Hamel *et al.*, 2001; Shevchik *et al.*, 1994).

The chimaeric version of BamA was created by gene fusion and allelic exchange. The entire *P. atrosepticum bamA* gene was not substituted into *E. coli* because another research group had trouble with compatibility when expressing BamA in a range of other species (Volokhina *et al.*, 2013). Given that the homologue from *P. atrosepticum* contains only four POTRA domains compared to the five in *E. coli*, there were potential undesired side effects resulting from a direct swap. The *E. coli bamA* was cloned into pKT230 and transformed into *P. atrosepticum* where it was shown to fold correctly and localise to the outer membrane (Smith *et al.*, 2007). Since this strain still had a functional native BamA no conclusions can be drawn on any functional complementation of the POTRA domains. So to avoid any complications from the lack of POTRA domains or other potential different components of the periplasmic domains of BamA, the POTRA domains of *E. coli* were joined to the β-barrel domain of *P. atrosepticum* by fusion PCR (Fig, 3.2.1). Thereby maintaining all the essential (and possible non-essential) function of the POTRA domains and their interactions with the rest of the *E. coli* BAM complex and nascent OMPs while swapping the β-barrel component for a phage resistant variant. Use of this construct would be able to show that the β-barrel domain of BamA is the only factor playing a role in providing an adsorption target for $\phi 24_B$.

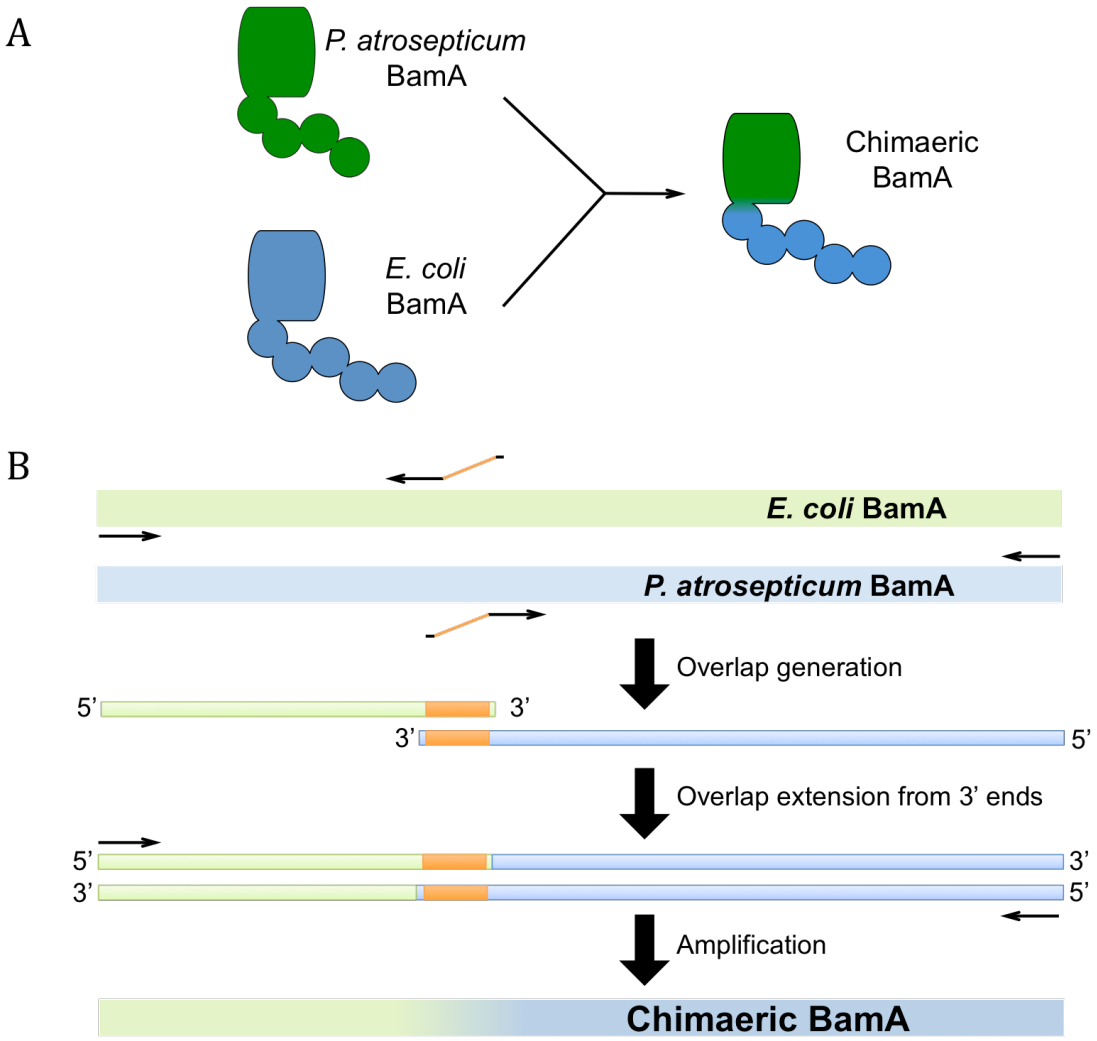

Figure 3.2.1: Fusion PCR strategy to generate a chimaeric BamA.

(A) Schematic showing the strategy of combining BamA homologues to create a functional but phage resistant chimaera. (B) Blue – *P. atrosepticum* DNA, Green – *E. coli* DNA, Orange – 17 bp overlap in PCR primers for fusion. PCR amplification of the POTRA domains of *E. coli* and the β-barrelled domain of *P. atrosepticum* produced fragments, which have sequences that overlap at each end so each acts as the primer for the other to extend from the 3' end of the fragments. The fused product is a genetic hybrid of *E. coli* and *P. atrosepticum* BamA with the reading frame conserved throughout.

3.2.2 Allelic exchange with WT BamA

In order to create a strain of *E. coli* that is resistant to $\phi 24_B$ adsorption, the chimaeric BamA construct was introduced into the MC1061 genome via allelic exchange (homologous recombination). To accomplish this, the β-barrel sequence of *P. atrosepticum* was flanked by *E. coli bamA* sequences upstream and the *skp* gene downstream. A kanamycin resistance cassette (Kan^R) was also included between the *P. atrosepticum bamA* sequence and the *E. coli skp* gene to allow for positive selection of successful exchanges. This construct should be able to recombine with the wild type (WT) *E. coli* genome resulting in the replacement of the WT *bamA* gene with the chimaeric *bamA* and kanamycin resistance gene (Fig 3.2.2.1).

This construction was done in two steps. First the Kan^R cassette was fused to the *skp* gene by fusion PCR to create the "KS fragment", which was then cloned into pUC19 via compatible *Eco*R I restriction sites (Fig. 3.2.2.2A). Then the *E. coli* and "*Erwinia*" *bamA* fusion sequences were also created by fusion PCR to create the *E. coli* and *Erwinia bamA* gene fusion (EE fragment). The EE fragment was then cloned into pUC19-KS upstream of the Kan^R cassette and *skp* gene, also using *Eco*R I restriction sites to create pUC19-EEKS (Fig 3.2.2.2B). As this had no *E. coli* sequences substituted back into the β-barrel, the plasmid was designated as pUC19-EEKS(0) where (0) indicates no sequence substitutions. The order and orientations of inserts were confirmed by PCR.

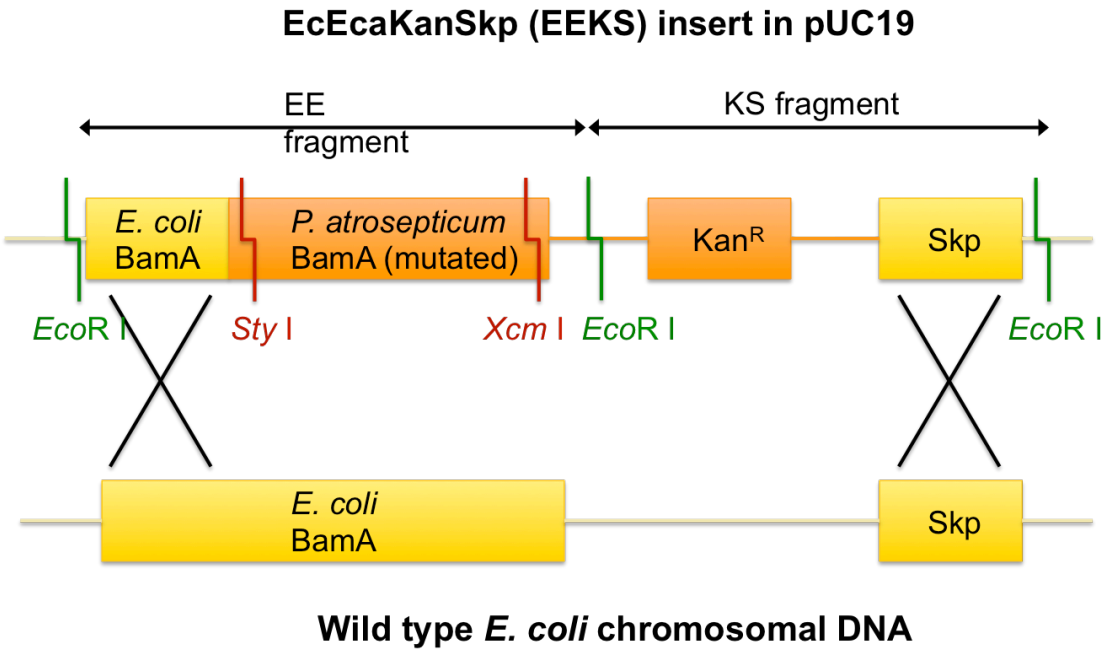

Figure 3.2.2.1: A schematic of the EcEcaKanSkp (EEKS) construct used for homologous recombination.

EEKS was created by sequentially subcloning the fusions of the chimaeric BamA (EE) and *kan*^{R} and *skp* (KS) into pUC19 with *Eco*R I. DNA matching WT *E. coli* chromosomal DNA is coloured yellow with DNA to be introduced in orange. Black crosses indicate regions of homology between the EEKS construct and *E. coli* chromosomal DNA where recombination is expected to occur. Permutations of the *E. coli* epitopes (Table 3.3.1) can be introduced into the *P. atrosepticum* section of BamA by subcloning using *Sty* I and *Xcm* I.

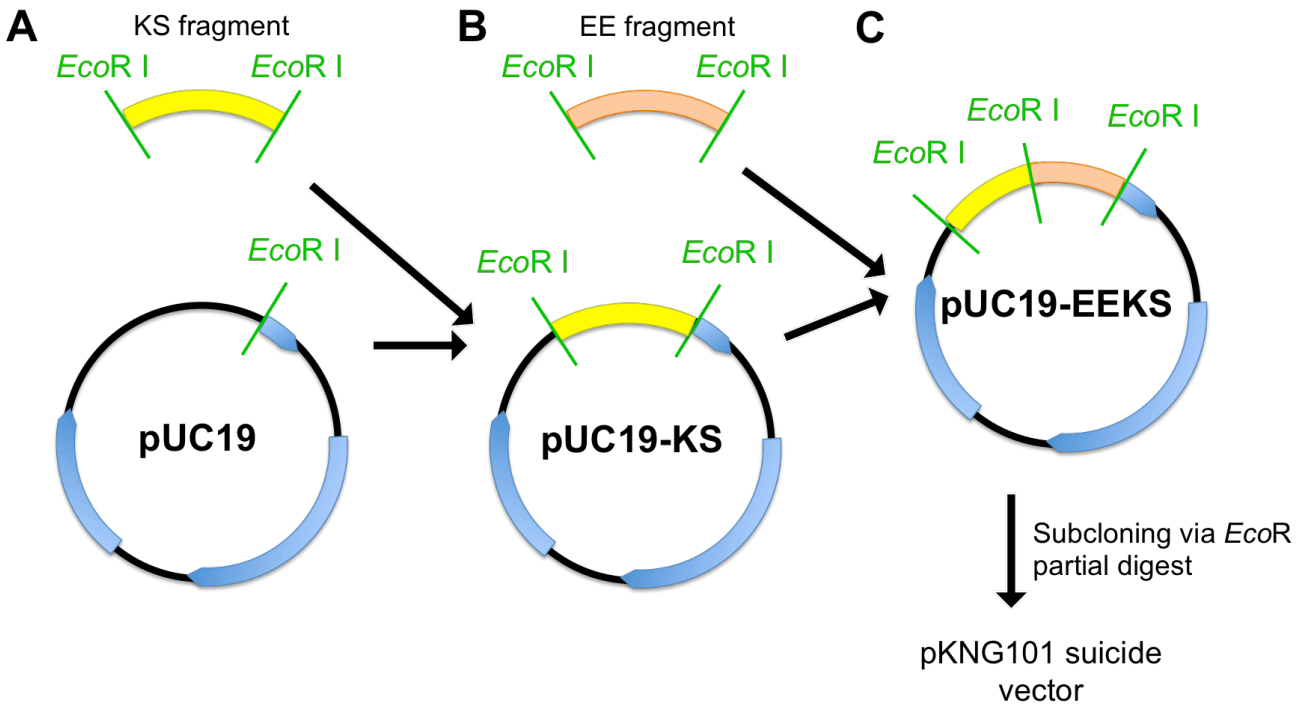

Figure 3.2.2.2: Stepwise process of constructing pUC19-EEKS.

(A) The DNA fusion of *Kan*^{R} and *skp* (KS fragment) was created by fusion PCR and cloned into empty pUC19 vector via compatible cohesive ends generated by digestion with *Eco*R I to create pUC19-KS. (B) Gene fusion of the *E. coli* "X site" and the C-terminal domain of *P. atrosepticum* cloned into a partially digested pUC19-KS with *Eco*R I. Sequence and orientation of the ligated EEKS constructs were confirmed by sequencing. (C) The plasmid, pUC19-EEKS, was partially digested to remove the EEKS fragment, which was subcloned into the pKNG101 suicide vector for recombination.

Even with the production of allelic exchange construct, detecting recombination reactions *in vivo* proved to be very time-consuming. The EEKS(0) construct was subcloned using *Eco*R I into a suicide vector, pKNG101 (Fig 3.2.2.2C). The pKNG101 plasmid contains the *sacB* gene that makes Gram-negative bacteria sensitive to sucrose in low osmolaric media *i.e.* no salt, allowing for positive selection of double recombination events (Kaniga *et al.*, 1991). The pKNG101-EEKS(0) plasmid was heat shock transformed into competent MC1061 cells. Colonies were obtained over multiple attempt on these plates showing that the first single crossover event was successful. On each attempt, twenty colonies were subcultured both in broth and on plates; however, these screenings for the second recombination event returned no growth in salt-free LB broth with added 50 μg mL^{-1} kanamycin and 5% (w/v) sucrose or on plates of the same.

Another mutant generation strategy employing the Red recombinase system on the pKD46 plasmid (Datsenko & Wanner, 2000) was used in parallel to the pKNG101 suicide vector strategy to enhance the likelihood of successfully generating the desired chimaeric mutant. This strategy, called recombineering, uses the pKD46 plasmid, which encodes the entire recombinase system from lambda phage controlled by an arabinose-inducible promoter leading to vastly increased expression of recombinases increased the likelihood of recombination occuring *in vivo*. Indeed, using pKD46 resulted in the first successful allelic exchange (Fig. 3.2.3). The strain was named Bam0 as a reference to this strain harbouring a BamA gene with zero *E. coli* epitopes substituted into the BamA chimaera. Subsequent generations of this strain were stable and could survive a

freeze thaw cycle to -80°C, unlike the MC1061-MRL strains which were unable to be resuscitated. Confirmation of the recombination was screened first using PCR using primers “5’Eca Loop Fwd” and “3’Rev skpEcoRI” (Fig. 3.2.3); However, PCR can not be used to confirm this because the absence of the 3 kb band (expected fragment size without *Kan*^{R} inserted between *bamA* and *skp*) alongside the presence of a 3.8 kb band (with *Kan*^{R}) is not proof that such a sequence no longer exists somewhere else in the cell. The creation of the Bam0 mutant was confirmed by Southern blot, which was the only way to rule out the production of a merodiploid cell, where partial chromosome duplication can result in the presence of both a WT and mutant copy of the *bamA* gene. The Southern blot shows that there is only a single band of BamA in both of the digested genomic DNA preparations of Bam0 and MC1061 (Fig 3.2.4). The difference in sizes of the fragments between the strains is consistent with the expected recombination of the EEKS construct from figure 3.2.2. The presence of a kanamycin resistance cassette in Bam0 also indicates that this recombination occurred.

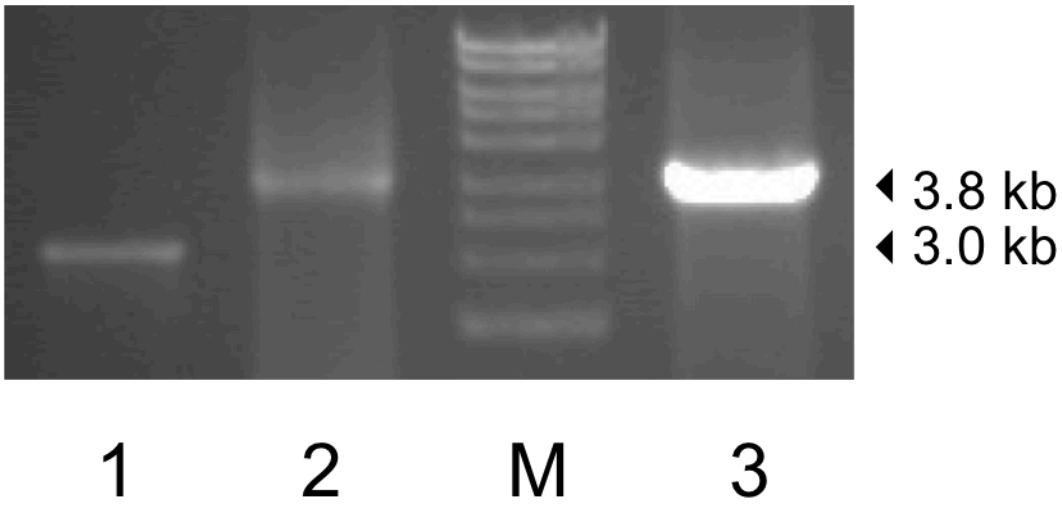

Figure 3.2.3: Screening colonies acquired from recombineering protocol.

Transformation of pUC19-EEKS(0) into MC1061 to introduce genetic constructs of the modified BamA gene. PCR amplifications were performed using "EcEca 5' X sitefwd" and "3' rev Skp EcoRI" for the following templates: (1) An MC1061 *E. coli* colony as a control for lack of recombination. (2) pUC19-EEKS(0) for a positive control indicating recombination of the EEKS(0) construct with genomic DNA. (3) A colony acquired using the Red recombinase method. PCR products were sequenced to confirm they represented the expected BamA sequence. M = Hyperladder 1kb (Bioline).

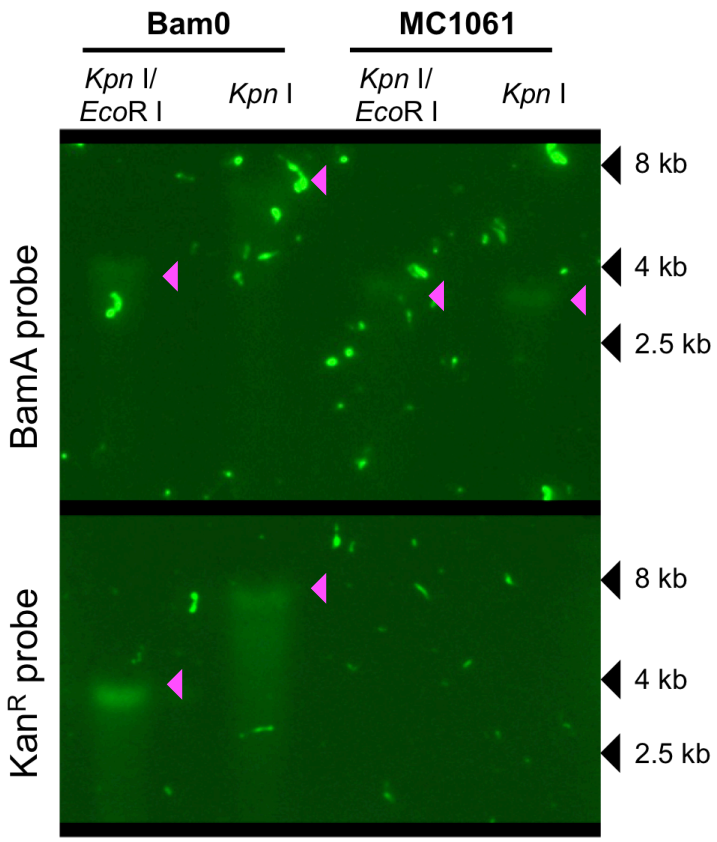

Figure 3.2.4: Southern blots to confirm single copy of BamA in Bam0.

Genomic DNA of MC1061 and Bam0 was digested with the enzymes indicated and probed for BamA gene (top) and the kanamycin resistance gene Kan^R (bottom) using DIG-labelled DNA probes. Expected sizes (indicated by the magenta arrows) for the BamA probe: (left-right) 3885, 7351, 3665, and 3665. Kan^R expected sizes: (left-right) 3466, 7351, n/a, and n/a. A single band in the lanes with the BamA probe shows there is only a single copy of BamA in Bam0 and MC1061.

3.3 Creating a library of isogenic BamA mutants

In order to get a deeper understanding into a potential mechanism for *E. coli* BamA acting as the receptor for $\phi 24_B$, the amino acid sequences of BamA from fully sequenced members of the *Enterobacteriaceae* were aligned. The analysis of the alignment revealed 5 small sequences within the species that supported $\phi 24_B$ adsorption and those that were resistant to $\phi 24_B$ adsorption (Fig 3.1.1.1). A comprehensive analysis of all 32 possible combinations of the different epitopes (Table 3.3.1) was planned in which the minimal required for phage adsorption support could be discovered.

3.3.1 Previous construction of chimaeric *P. atrosepticum* BamA genes

The most recent completed work resulted in the construction of chimaeric genes possessing 25 out of the 32 possible combinations of *E. coli* epitopes (McEwen, 2013). Table 3.3.1 shows which constructs are cloned into pCR-Blunt and sequenced. Two constructs contained a nonsense point mutation within the gene and five were yet to be made as the set of primers used to introduce loop 4b ("5' fwd loop IIb" and "3' rev loop IIb") required a slightly modified PCR protocol that required a higher melting temperature and the addition of 3% (v/v) DMSO to the PCR reaction before the products would consistently amplify using the same due to non-specific binding interactions of the much longer 'overlapping regions' incorporated onto the 5' ends of the primers.

Table 3.3.1: Pre-isogenic BamA mutated chimaeric genes created by McEwen (2013).

Number of swapped epitopes	Code	Epitope Present					Cloned in pCR-Blunt	Sequence confirmed ‡
		4a	4b	6	7a	7b		
0	0						x	
1	1a						x	
	1b						x	
	1c						x	
	1d						x	
	1e						x	
2	2a						x	
	2b						x	
	2c						x	
	2d						x	
	2e						x	
	2f							
	2g							
	2h						x	
	2i						x	
	2k						x	
3	3a						x	
	3b						x	
	3c						x	
	3d						x	
	3e						x	
	3f						x	
	3g							
	3h							
	3i							
	3k						x	
4	4a						x	
	4b						x	
	4c						x	
	4d						x	
	4e						x	
5	5						x	

‡ Green indicates that they were sequenced and their construction confirmed.
Red represents clones with a nonsense mutation.
White indicates that no sequence was obtained.

3.3 Completing the library of chimaeric BamA genes

The work from the previous study (McEwen, 2013) was continued and the library of epitope substitutions in pCR-Blunt was completed to include the constructs not completed (Table 3.3.1) and sequenced (Appendix 8.1).

Upon completion of the library of chimaeric genes, multiple attempts were made to generate Bam5 via allelic exchange. Bam5 would be Bam0 but the chimeric BamA gene would have substituted all five *E. coli* epitopes in for the *P. atrosepticum* epitopes. Yet despite the immediate previous success with the Red recombinase system encoded on pKD46 to create Bam0, this repeatedly failed when attempting to do the same thing to create what would be Bam5. This was the case using either pKNG101 or pKD46 with either circularised or linear DNA. It was later discovered that the root cause was an incorrectly engineered primer ordered in 2007, which was missing two bases of the codon encoding Y578. Not only is this tyrosine residue conserved in BamA across all members of the *Enterobactericeae*, thus being potentially functionally important, but the 2 missing bases introduced a 2 bp deletion in the gene and therefore a -2 reading frame shift. The shift immediately introduces a stop codon after S577, truncating the essential BamA. This explains the lack of positive recombinants using pKD46 and the total lack of colonies for the second recombination event when using pKNG101. With pKD46, the cells in which recombination occurs will not be viable with a truncated BamA and, with pKNG101, the cells are effectively being given the choice between either a suicide gene, sacB, or a non-functional

essential gene, BamA. Either of these results in a non-viable cell and so, in retrospect, the lack of recombinants indicates that the systems were working and confirms that BamA cannot be deleted. Unfortunately this lethal mutation was found in all the constructs containing the 4b epitope (16 of the 32 combinations). New primers (5' fwd IIb fix & 3' rev IIb fix) were used to redo the fusion PCRs for the 16 mutants to fix the reading frame shift and reinstate Y578 and the rest of the gene. The fixed constructs were re-sequenced so the mutant library is still ready to be used to generate the isogenic BamA mutants in *E. coli*. These cloned constructs have been labelled with the suffix "fixed" *e.g.* "pCR-Blunt-1b fixed".

3.4 Testing the infection phenotype of Bam0

The ability of the chimaeric protein to support adsorption in an *E. coli* background was tested using a spot assay: a visible spot (lack of bacterial growth on an indicator lawn) where 10 μL of a phage suspension (6 x 10^7 PFU mL^{-1}) was dropped onto an indicator host lawn would indicate whether a successful infection has occurred.

Adsorption is the first step in infection, so if infection occurs then $\phi 24_B$ adsorption was supported. Both the lack of spot on the *P. atrosepticum* indicator lawn and the presence of one using *E. coli* (Fig. 3.4.1A-B) is consistent with previous findings using $\phi 24_B$ (Smith *et al.*, 2007). The Bam0 strain was unable to support phage infection (Fig. 3.4.1C). Reintroducing a copy of WT *E. coli* BamA

using the pUC19-ϕR1D construct (Sergeant, 1998; Smith *et al.*, 2007) was able to restore the phage susceptible phenotype (Fig. 3.4.1D). This showed that the β-barrel domain of *E. coli* BamA can be substituted with the homologous domain from *P. atrosepticum* to remove adsorption support while complementing the protein folding and insertion functions without impacting cell viability. Growth curve data was not collected but Bam0 appeared to grow at the same rate as MC1061 both in liquid (judging by time taken to reach OD_{600} = 0.45 for the spot assay) and colony appearance on agar plates.

Bam0 remains the only isogenic mutant created so far as the same recombineering strategy was not successful when attempted with EEKS(5) – the mutant of the chimeric BamA which has all five *E. coli* epitopes substituted in for the corresponding residues of the BamA ortholog in *P. atrosepticum* as shown in figure 3.1.1.1. Ideally, EEKS(5) would have been used to create the Bam5 strain, which was expected to reintroduce adsorption. However, with the ϕR1D complementation (Fig. 3.4.1D), the data presented here show a definitive proof that BamA is the adsorption target for $\phi 24_B$ and, by extension, any other phage with the same tail structure.

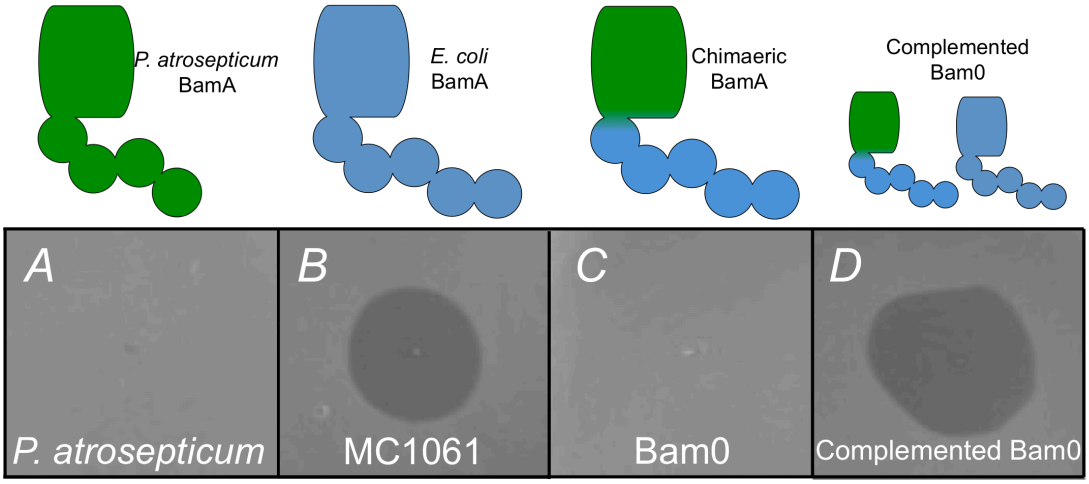

Figure 3.4.1: Spot assay for phage infection of cells expressing variants of BamA.

Schematic representations of the BamA variant encoded by the cell is shown above the assay result. Phage stock (10 µL of a 6 x 10^7 titre) was spotted onto top agar containing the strains indicated (A) *P. atrosepticum*, a negative control for the spot assay; (B) WT *E. coli* strain MC1061, a positive control with an infection phenotype; (C) Allelic exchange of the WT BamA for the chimeric BamA, loss of the infection phenotype from MC1061; (D) Complementing Bam0 with WT *E. coli* BamA (pUC19-ϕR1D) from MC1061 restores the positive infection phenotype.

3.5 Conclusions

3.5.1 Fusion PCR can create viable mutants in BamA

The sequential addition of specific mutations into the *P. atrosepticum* β-barrel of the chimaeric *bamA* gene required the use of a non-random, site directed mutation methodology. This is especially important when making mutations in essential genes where any mutation in a highly conserved area is likely to disrupt an essential function. Fusion PCR using specially designed primers to introduce these multi-nucleotide mutations is a slow but precise method to make a large number (32) of specific mutants.

3.5.2 Creating a $\phi 24_B$ resistant strain, MC1061-Bam0.

Substituting *E. coli* BamA for a chimaeric BamA in MC1061 by introducing the EEKS(0) construct via allelic exchange maintains the essential LPS membrane protein insertion function. This is evidenced by the viability of cells in which this recombination has been confirmed through Southern blot analysis. This mutant strain, Bam0, was then unable to support $\phi 24_B$ infection in a spot assay. This is a crucial step in definitive proof of the function of BamA as the phage receptor.

3.5.3 Restoring $\phi 24_B$ infection support to MC1061-Bam0

Complementing the Bam0 strain with a plasmid carrying and expressing WT BamA, specifically pUC19-ϕR1D, was enough to completely restore an infection susceptible phenotype. This plasmid, pUC19-ϕR1D, also restored phage susceptibility to the original MRL mutants made by Sergeant (1997). Though we will never know what kind of natural mutation was generated in the MRL strain, it seems very likely that it lay somewhere within BamA.

3.6 Discussion of results

The creation of a BamA chimaera by fusion PCR was the first major step towards performing an assay for phage adsorption via an essential protein and subsequent infection of the host cell. It has been shown that the combination of β-barrel and POTRA domains is important as, despite their essentiality, BamA proteins seem to diverge quickly enough in evolutionary time to be species specific (Volokhina *et al.*, 2013). The alignment of the POTRA domains with respect to the β-barrel is the most important factor when considering whether these two regions will be functionally compatible (Browning *et al.*, 2015). When the fusion PCR experiment was designed there was no data to suggest whether swapping the entire BamA protein or if creating a chimaera of *E. coli* and *P. atrosepticum* BamA proteins would impact the function of the BAM complex with respect to lipoprotein and/or substrate binding.

After allelic exchange of the fused chimaeric *bamA* with the genomic copy of *bamA* in *E. coli* MC1061, it was discovered that the POTRA and β-barrel domains of the BamA homologues from *E. coli* and *P. atrosepticum* are able to combine into functional chimaeras and create viable cells. This compatibility may be possible as they have a more conserved structure due to minimal phylogenetic divergence. The BAM complex of the newly created Bam0 strain was able to function with no observable difference to WT MC1061, further implying a strong inter-species compatibility and no loss of functional efficiency.

It was decided that a library of specific site-directed mutants should be made to investigate which of the extracellular epitope are required to support phage adsorption. Non-specific mutants could have been made following a random mutagenesis method, either by chemical (Bose, 2016) or error-prone PCR-based (McCullum *et al.*, 2010) methods, which may have allowed the identification of residues involved in the adsorption process. This high-throughput approach would have yielded results more quickly; however, it probably would not have answered why the difference between the extracellular surface of the BamA homologues in *E. coli* and *P. atrosepticum* result in different phenotypes for $\phi 24_B$ adsorption. Synthetic biology could also have been used to generate these specific constructs in a more time-effective manner and would have avoided the creation of constructs with a reading frame shift.

The heavy time commitment required to generate the library of 32 very specific isogenic BamA mutants was a necessary step to allow dissection of these critical

differences between the BamA proteins of *E. coli* and *P. atrosepticum*, which we already know support or are resistant to phage infection respectively.

An effort was made to provide definitive evidence that the parent strain of Bam0 is MC1061 using pulsed field gel electrophoresis (PFGE). The idea was that digesting both Bam0 genomic DNA (gDNA) and MC1061 gDNA with *Eco*R I would produce identical restriction fragment length polymorphism (RFLP) data, except for 1 band which would be cleaved in two by the *Eco*R I sites introduced in the flanking the Kan^R and *skp* genes (this fragment would also create a new band on the gel). However, PFGE was abandoned as the high and low molecular weight ladders ran as smears through the gels and after exhaustive troubleshooting, it was suspected that the PFGE equipment was not working correctly.

The successful complementation of the Bam0 strain by pUC19-ϕR1D is a significant development and provides the definitive proof that BamA acts as the receptor for $\phi24_B$. This evidence was the final missing piece from the analysis by Smith *et al.* (2007), which was only went as far as introducing $\phi24_B$ susceptibility to *P. atrosepticum* by heterologous expression of *E. coli* BamA without knocking out WT BamA first. The data collected now satisfy molecular Koch's postulates (Falkow, 1988). This allows the study to move forward and plan a more informative assay for expressing and screening the library of mutants for the effect that they have on $\phi24_B$ adsorption. That assay is yet to be established. For most options, additional DNA constructs and molecular tools would need to be made before performing the assays. Potential approaches to answering this question are discussed in Chapter 4.

CHAPTER 4: Expressing BamA Mutants and Generating Molecular Tools to Investigate Adsorption

4.1 Introduction

Having created a library of 32 isogenic *bamA* mutants and a phage resistant strain of *E. coli*, Bam0, through allelic exchange, the next step was to try to express the mutated forms of BamA from the library and determine how the defined mutations impact upon phage adsorption to BamA. There are multiple ways to test this effect and each method has its own advantages and disadvantages.

4.1.1 Infection assay

The simplest method was to ascertain whether $\phi 24_B$ was able to infect cells exclusively expressing one of the engineered BamA mutants, as mutations causing adsorption impairment should be unable to support phage infection as the virus will have no or limited ability to recognise potential host cells. This method was used in Chapter 3 to confirm that Bam0 was unable to support $\phi 24_B$ infection (Fig. 3.4.1). A limitation of this method is that, rather than just measuring adsorption, it examines the ability of the mutant to support the entire infection process. Therefore further experiments would be needed to confirm that any infection resistant phenotype is caused by a failure at the adsorption

step. The complementation of Bam0 with WT *E. coli* BamA provides good evidence that the loss of an infection phenotype is due to loss of adsorption.

4.1.2 Adsorption assay

An adsorption assay uncouples adsorption from the rest of the infection process. The principle of the adsorption assay for $\phi 24_B$, detailed in Smith *et al.* (2007), involves adding a set number of phages to a known number of host cells, allowing them to interact for a prescribed period of time and removing the cells (along with the adsorbed phages) by centrifugation. The remaining phage particles that have not adsorbed are then quantified by plaque assay. This methodology involves lots of sample handling, which can lead to noisy data, and so accurate measurements can be difficult to acquire, but general trends can be seen.

4.1.3 *In vitro* molecular interaction assays

Studying the protein-protein interaction *in vitro* would be the most quantifiable method to determine binding affinities. As examples, surface plasmon resonance (SPR) or isothermal titration calorimetry (ITC) are both biophysical methods that detect rates of molecular interactions. SPR measures small changes in the angle of refraction of electromagnetic waves as they are affected by molecular interactions in the sample, and ITC measures interaction rates by comparing the heat generated when molecules interact when mixing known concentrations of each interaction partner (Pierce *et al.*, 1999). These methods could be used to

determine binding affinities of the phage tail and BamA. The presence of hydrophobic transmembrane regions would present a practical issue for these methods as introducing detergents make data more complicated to interpret (Rajarathnam & Rösgen, 2014). The use of peptides that mimic the extracellular epitopes of *E. coli* would circumvent the need to solubilise the whole β-barrel structure including the hydrophobic transmembrane domains. However it would be impossible to ensure that the peptides would be presented to the phage in a conformationally appropriate manner. Additionally, if the phage binding to BamA results in a conformational change in BamA the use of peptides may not provide the phage with the structures necessary to induce BamA to open up a surface for a secondary binding event, should this be part of the adsorption process. The failure of these assays to work, though, would not be meaningful and very difficult to troubleshoot as protein conformation and protein sequence importance in the support of phage binding are likely to be impossible to unpick.

4.1.4 Fluorescent microscopy assay

The final option considered here was a fluorescence microscopy-based assay. SYBR gold is a fluorescent stain that has been used to label phage DNA in order to monitor phage infection of host cells (Mosier-Boss *et al.*, 2003). This type of assay allows the visualisation of the cells that are susceptible to adsorption, and the release of fluorescent phage DNA into the bacterial host's cytoplasm resulting in the bacterial cell fluorescing. Cells that are resistant to phage adsorption will not lead to fluorescent cells.

4.1.5 Method evaluation

The infection assay has been performed for Bam0 and will continue to serve as a valuable first test to determine infectivity of $\phi 24_B$ on the isogenic strains of *E. coli*. To gain further understanding of the adsorption process another method will need to be used. The molecular interaction characterisation assays would provide quantifiable data but are further removed from a biologically relevant scenario than either the fluorescence microscopy or the adsorption assay. The removal of the phage and BamA (or peptides) from the *in vivo* environment may necessitate more assumptions about exactly what is going on during the interaction. This is not true of the fluorescence microscopy assay, which can visualise phage adsorption and DNA ejection facilitating the troubleshooting process should the assay fail. Microscopy data can be analysed in a quantifiable way on a single-cell level, this will eliminate the noise generated during the adsorption assay. For these reasons, the fluorescence microscopy assay was developed specifically for $\phi 24_B$ in Chapter 5.

All of these assays have different requirements in terms of DNA constructs, protein purification, and molecular tools. The most useful construct to extend the complementation of Bam0 would be the creation of the "Bam5" strain of *E. coli*, which would express the chimaeric BamA with all five *E. coli* epitopes substituted into the β-barrel structure of the *P. atrosepticum* BamA as this would provide a cleaner proof of adsorption restoration. As EEKS(5) is the construct that differs the most from EEKS(0). It would also suggest that there would be no

problems when introducing any of the 32 constructs to MC1061 via allelic exchange, as they are effectively a continuum between the two.

4.2 Re-attempting creation of Bam5 strain with "fixed" EEKS(5)

As mentioned in section 3.4.1, the plasmid pUC19-EEKS(5) (pUE5) did not undergo allelic exchange when introduced into MC1061 containing pKD46, a plasmid encoding the Red recombinase system. As stated in Chpater 3, this failure can be explained by the frame shift mutation within the coding region that would have immediately introduced a stop codon in the essential BamA gene. "Fixing" this frame shift mutation (as described in Chapter 3.3) should allow allelic exchange to occur and produce the Bam5 strain.

4.2.1 Fixing the EEKS(5) construct

In order to reattempt this allelic exchange, the mutation in the EEKS(5) insert was fixed by fusion PCR using two pairs of primers to generate separate overlapping fragments (Fig 4.2.1A-B). These fragments were then fused by PCR to produce the ~3200 bp band of the EEKS insert. The gel extraction of this band only yielded 26 ng μL^{-1}. In order to generate a higher concentration of linear DNA for transformation, this PCR fragment was cloned into pUC19 using *Eco*R I to make pUE5, and this plasmid was used as a template to re-amplify the EEKS(5) construct using "5' EcEca X site" and "3' rev SkpEcoRI" primers (Fig. 4.2.1C).

4.2.2 Allelic exchange using EEKS(5)

Separately, both the linear amplification product of EEKS(5) and the circularised pUE5 plasmid were transformed by electroporation into mid-log phase MC1061 harbouring the pKD46 plasmid using a micropulser (BioRad, UK). Cells were recovered at 30°C and spread on selective plates and incubated at 30°C overnight to maintain the pKD46 plasmid, which possesses a temperature sensitive origin of replication. If no colonies were observed, plates were incubated for a further overnight step in case allelic exchange did occur and the cells exhibited slower growth as a result. Colonies were replica plated onto LB plates containing kanamycin and kanamycin & ampicillin. These colonies were also used to directly inoculate aliquots of liquid LB-kan. These broth cultures were incubated at 43°C overnight to cure the cells of pKD46, as the plasmid is unable to replicate at this temperature. Daughter cells lose the plasmid and the only way the cell can survive is if a recombination reaction occurred to insert the Kan^R cassette into the genomic DNA.

Using "5' EcEca X site" and "3' rev SkpEcoRI" primers, colonies were screened for the presence of a successful exchange event resulting from the insertion of a kanamycin resistance gene, *aphA*, between *bamA* and *skp* in the MC1061 genome. Despite being picked from kanamycin plates, no colonies displayed further growth on the kanamycin replica plate nor was there any growth in the liquid cultures at 43°C. The PCRs resulted in 2.4 kb bands (Fig. 4.2.1D), which is the expected size of bands in the absence of recombination and the size of the bands in all samples were identical to the negative control (naïve MC1061 cells).

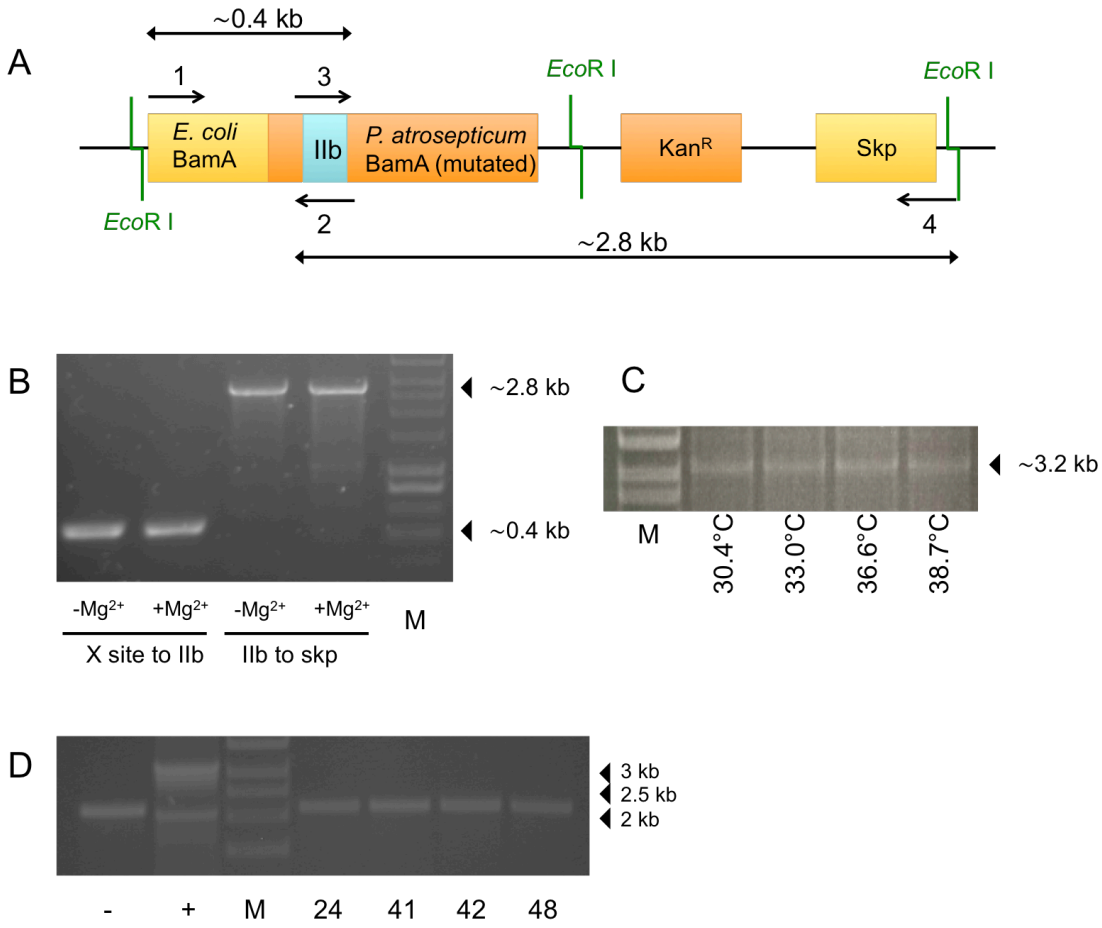

Figure 4.2.1: Fixing 4b loop in EEKS and attempted allelic exchange in MC1061 pKD46.

(A) Fixing the EEKS(5) construct containing the reading frame error was done by fusion PCR. Primers 1-4 are "EcEca 5' X site fwd", "3' rev IIb fix", "5' fwd IIb fix", and "3' rev SkpEcoRI", respectively. (B) Amplification using primer sets 1 & 2 and 3 & 4 of overlapping fragments from the *E. coli* recombination site (X site) to the *skp* gene. Overlapping sequences covering the 4b epitope facilitated gene fusion while also repairing the frame shift mutation. (C) Amplification of the EEKS(5) created by fusion PCR over a range of annealing temperatures. (D) Screening for successful recombinations with templates: (-), MC1061 cells; (+), pUC19-EEKS(5); 24, 41, 42, and 48 are selected clones that were screened with all being negative for recombination. M = Hyperladder 1kb (Bioline,UK)

Altogether, this demonstrates that the kanamycin resistance gene, and therefore the β-barrel from *P. atrosepticum*, did not transfer to bacterial genomic DNA during any of the attempts.

4.3 Subcloning full length BamA in pUC19

Up to this point, most of the genetic manipulation work on BamA has focussed exclusively on the β-barrel domain and has ignored the amino terminus of the protein including the POTRA domains located within the periplasmic space. Any complementation assay or *in vitro* work would require the full length chimaeric BamA (FLbamA) to be expressed. The EEKS constructs used so far did not include the first 1232 bases of the BamA gene, so the full-length chimaeric gene needed to be cloned. Figure 4.3.1 shows the strategy used to achieve this, Bam0 gDNA was extracted using the ISOLATE II Genomic DNA Kit (Bioline, UK) and the POTRA domains of the chimaeric BamA gene was amplified from this template using primers "SM_BamA_for" and "EcEca 5' X site rev". The β-barrel of BamA was amplified from the "fixed" pUE5 plasmid using primers "EcEca 5' X site fwd" and "3' rev loop Eca". These two fragments were then fused together by PCR with the recombination site (X site) on each fragment acting as primer for the other fragment (Fig. 4.3.1A).

This FLbam5 construct was cloned into pCR-Blunt (Invitrogen, UK) following manufacturer's instructions and the construction confirmed by PCR and *Eco*R I restriction endonuclease digestion mapping (Fig. 4.3.1B). The plasmid, pCR-Blunt, encodes *Eco*R I restriction sites flanking both sides of the PCR fragment

insertion site within its MCS. The *Eco*R I released fragment was gel purified and subcloned into an empty pUC19 vector, which had been digested and treated with Antarctic phosphatase (NEB, UK). The ligation reaction had an insert to vector ratio of 2:1 and the ligation reaction was left at 4°C overnight. The ligation reaction, 2.5 μL, was transformed into Top10 cells using heat shock and the mixture was plated onto LB ampicillin plates spread with IPTG and X-gal. All 63 colonies from the first attempt did not contain inserts and were most likely just vector re-ligations (Fig 4.3.2). From the 76 picked colonies on the second attempt, 48 of these were screened using M13 primers. A prominent band at around 1 kb appeared in all lanes, including the negative template control (data not shown). A subset of 6 samples showed bands elsewhere so these were screened again with fresh PCR grade water and freshly diluted M13 primers. Only one sample showed a band at ~2400 bp using both sets of primers (Fig. 4.3.1C). This sample was sequenced (GATC, UK) and it was shown to have no mutations and was cloned in matching orientation to the lacZα and vector promoter. The pUC19-FLbam5 plasmid was heat shock transformed into Bam0, which was subjected to the same infection assay as described in section 3.4. The plasmid did not restore an infection phenotype onto Bam0.

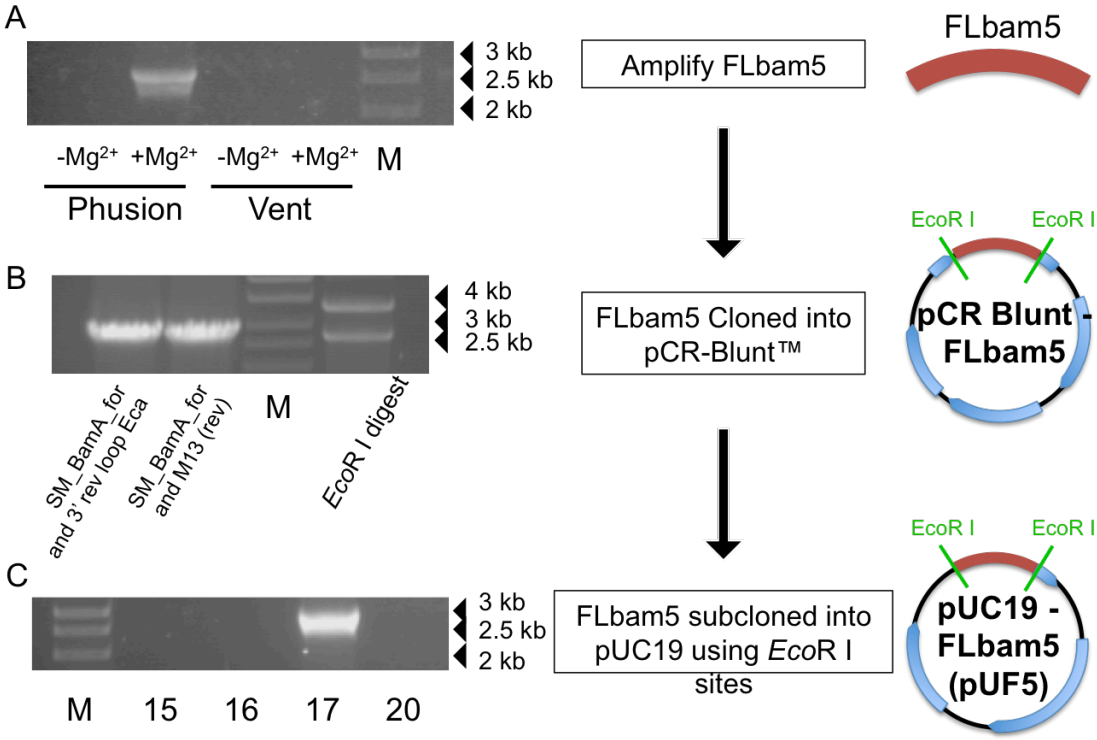

Figure 4.3.1: Creation of pUC19-FLbam5.

(A) A combination of polymerase and PCR conditions was found which allowed the gene fusion of the NTD and CTD to produce a full length Bam5 (FLbam5). (B) FLbam5 was cloned into pCR-Blunt following the manufacturer's instructions (Invitrogen, UK). (C) Screening of colonies transformed with FLbam5 subcloned into pUC19 revealed only 1 successful ligation. M = Hyperladder 1kb (Bioline,UK)

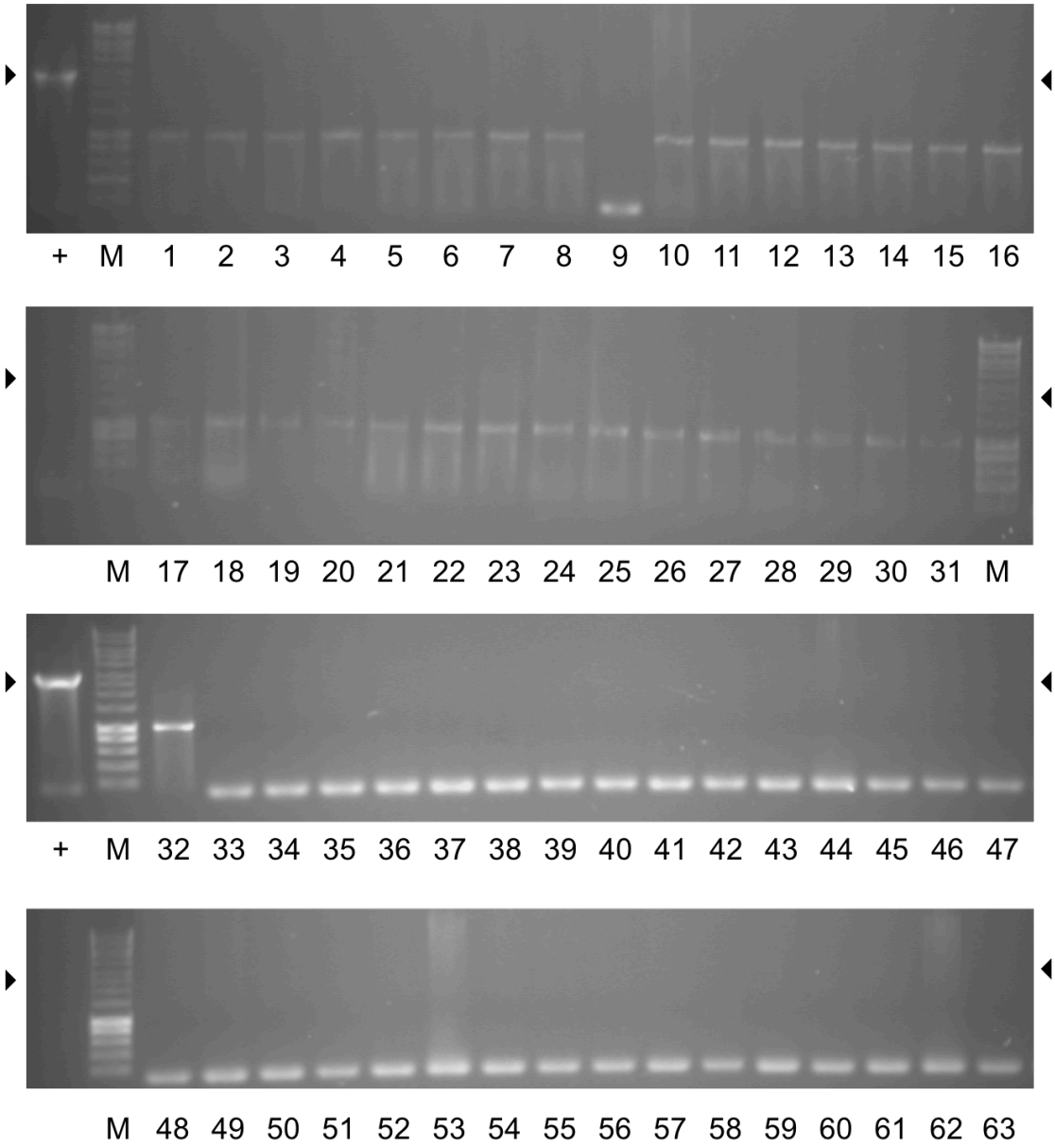

Figure 4.3.2: Screening of the first attempt at subcloning FLbam5 into *Eco*R I- digested pUC19.

All 63 colonies picked and screened using M13 primers for the first screening. No bands at ~2.5 kb (marked by arrows: 2.4 kb gene + 100 bp in flanking regions up to the M13 primer binding sites) demonstrates there was not a successful clone of on this first attempt. (+) = FLbam5 cloned into pCR-Blunt. M = Hyperladder 1kb (Bioline,UK).

It was assumed that FLbam5 being cloned in the same direction would aid expression and therefore facilitate the restoration of ϕ24B adsorption in an infection assay. However, in both the presence and absence of IPTG in the medium, there was no restoration of the phage adsorption. This is contrary to the pUC19-ϕR1D plasmid, which contains the *E. coli* BamA homologue, which does restore adsorption (Fig. 3.4.1). The vector promoter for pUC19-ϕR1D is oriented in the opposite direction to the BamA insert, which expresses BamA without IPTG induction (Sergeant, 1998; Smith *et al.*, 2007). It was not determined whether any mRNA of the Bam5 gene was transcribed from this construct so a definitive conclusion on transcription and/or expression levels cannot be made at this time.

4.4 Cloning FLbam5 into an expression vector

While it is concerning that pUC19-FLbam5 (pUF5) did not restore infection, the plasmid pUC19-ϕR1D created by Sergeant (1998) did restore infection (Fig 3.4.1). Given that the primary protein structures of the intended products of 2 constructs are functionally identical, the more likely conclusion is that there is an issue with expression of the Bam5 construct from the lac promoter of the pUF5 plasmid, as opposed to the FLbam5 being incapable of reintroducing adsorption support.

4.4.1 Altering the expression system

In order to determine whether the expression system was the cause of the lack of adsorption support by Bam5, there were two modifications to the vector that could be made. One option included the use of an alternative expression system. Previously, BamA was expressed in pUC19 (Sergeant, 1998; Smith *et al.*, 2007) and overexpressed in pQE-32 (Smith *et al.*, 2007). Over expression of BamA in pQE-32 required the presence of the pRep4 plasmid, which constitutively expresses the *lacI* repressor to prevent leaky expression from IPTG-inducible promoters, as even low-level expression of BamA beyond that normally produced by the endogenous *bamA* gene was lethal and prevented successful cloning (Smith *et al.*, 2007). The pBAD myc-His C plasmid was chosen as a vector for potential isogenic BamA expression because inserts are controlled under the araC promoter, which is completely silent until expression is induced by L-arabinose (Miyada *et al.*, 1984). This will ensure that BamA expression is turned off and will not hinder the cloning process. When cloning into pBAD myc His C, inserts are routinely designed to be cloned in using restriction digests of *Nco* I at the 5' end (using the start codon in the recognition sequence: CC**ATG**G) to *Sal* I at the 3' end, which lines up the reading frame to omit the myc tag and only adds the His-tag to the C-terminal domain (CTD) of the protein. This strategy was used here.

The other option would be to include the native promoter region of BamA in the insert so that Bam5 would be under the influence of "normal" expression patterns as it was in pUC19-ϕR1D. There is experimental evidence (transcription

initiation mapping) for two separate BamA promoters (Salgado *et al.*, 2013), which are both under the control of the σE regulon: the BamAp promoter begins 107 bp upstream of the start codon and controls the expression of BamA alone (Dartigalongue *et al.*, 2001b), while BamAp2 is 902 bp from the start codon and promotes the operon containing the following genes: BamA-skp-lpxD-fabZ-lpxAB-rnhB-dnaE (Rhodius *et al.*, 2006) (Fig. 4.4.1.1A).

Constructs for cloning BamA into pBAD were designed to use either the araC promoter in the vector or the BamAp2 (which would also include BamAp) that would mimic native expression of BamA (Fig. 4.4.1B). It was assumed that expression of BamA under native conditions would be non-lethal to any host cell as it would be regulated by the standard *trans* acting soluble regulators. Up to this point, the chimaeric BamA clones have not included any native promoter regions. The Bam0 strain was created by allelic exchange of EEKS(0) and MC1061 genomic DNA (gDNA). Therefore, gDNA from this strain was the only source of a chimaeric BamA with the BamAp2 promoter located upstream, so this strain was used as the template for the insert. The *E. coli* epitopes could be introduced afterwards by substituting the *Xcm* I to *Sty* I fragment (sites in Fig. 3.2.2) by restriction digestion for the Bam5 fragment in a single extra step once the full sequence of promoter and Bam0 gene had been cloned into a plasmid.

Attempts to clone the amplified insert were made using five different polymerases. Screening colonies repeatedly failed to identify plasmid-containing colonies when using the Zero Blunt™ PCR cloning kit (Invitrogen, UK) with blunt-ended PCR products generated using Q5, Vent (both NEB, UK) or Phusion

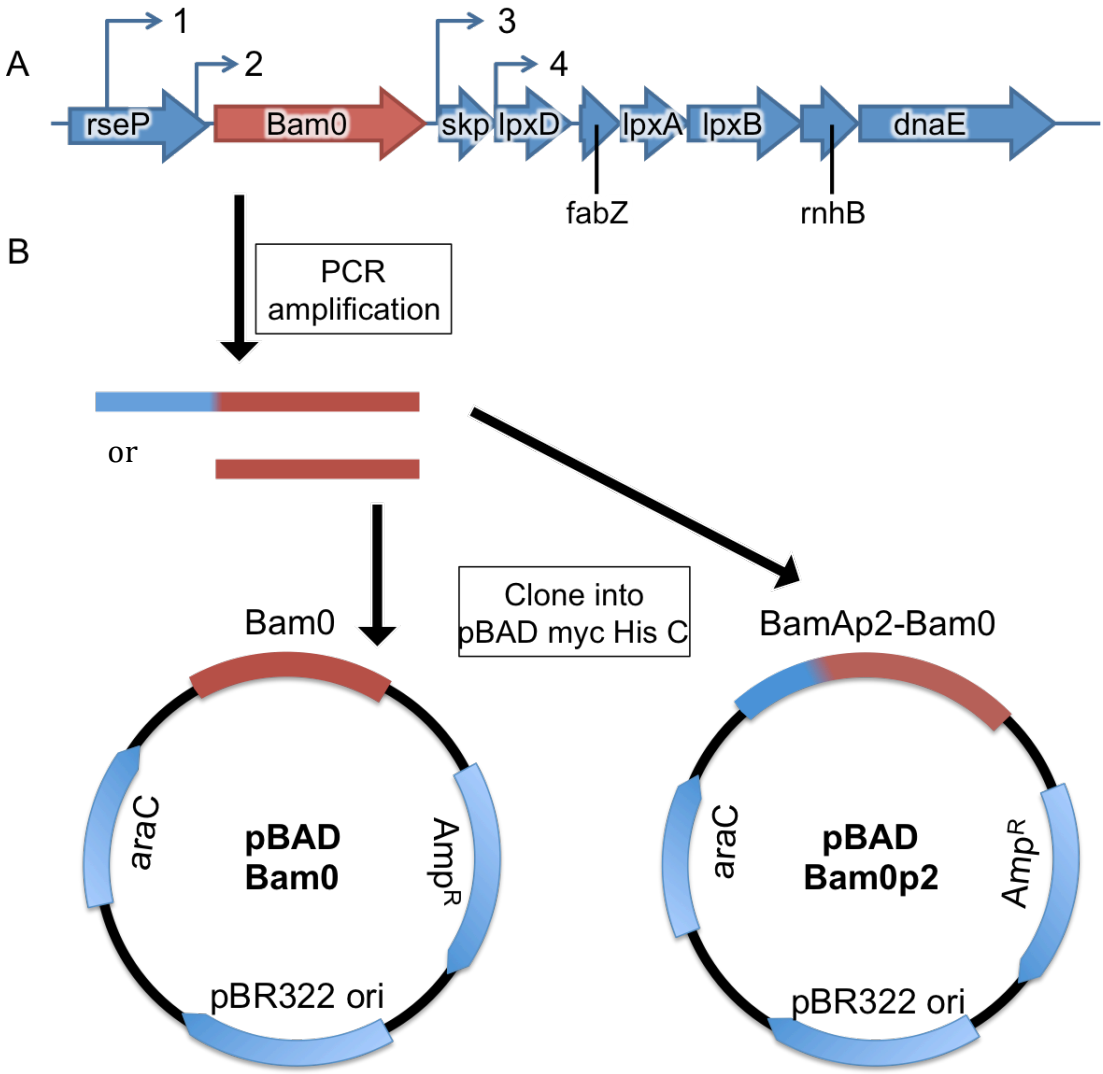

Figure 4.4.1.1: Promoters of the σE regulon, which promotes BamA, in *E. coli*.

(A) Gene sequence in the Bam0 strain. The numbered promoters control the following genes/operons: (1) BamAp2: BamA-skp-lpxD-fabZ-lpxAB-rnhB-dnaE. (2) BamAp: BamA only. (3) hlpAp: skp-lpxD-fabZ-lpxA. (4) lpxDp: lpxD. Promoters were found using RegulonDB (Gama-Castro *et al.*, 2016). (B) The plan involved amplifying and cloning two version of Bam0 into pBAD myc His C. One insert would be under control of the araC promoter (left) and the other would include the native BamAp2 promoter (right).

(Thermo fisher, UK) polymerases. The same lack of colonies resulted when the TA Cloning™ Kit (Invitrogen, UK) using BIO-X-ACT (long) and Taq (both Bioline, UK) polymerases were used, which led to trying an alternative cloning strategy.

4.4.2 Cloning directly by PCR

This final strategy used to clone FLbam5 into the pBAD expression system was to use PCR primers designed to amplify the insert directly into pBAD through a further PCR amplification. This was achieved by using primers for the insert that have extensions on their 5' ends that complement the vector. Thereby, after amplification, the insert contains vector sequence (green and red in Fig. 4.4.2.1A) that act as primers when the insert and vector are mixed for a second PCR (Fig. 4.4.2.1B). The extension step will amplify the entire backbone to generate a nicked plasmid that is stable enough to be transformed without needing to be ligated. The empty vectors used as templates in the second PCR must be destroyed by *Dpn* I digestion (destroys methylated sequences, i.e. DNA that was synthesised *in situ* is methylated while DNA synthesised *in vitro* is not methylated) before transformation to ensure that any colonies contain the insert (Fig 4.4.2.1C-D).

There was still an effort to clone BamA with both the native and araC promoters. This required two different forward primers, ("NcoI BamA over" and "BamAp2 over") with the same reverse primer, "SalI BamA over", again, using Bam0 as a template for the initial amplification of BamA with pBAD vector extensions. These reactions were transformed by heat shock into Top10 cells, and the

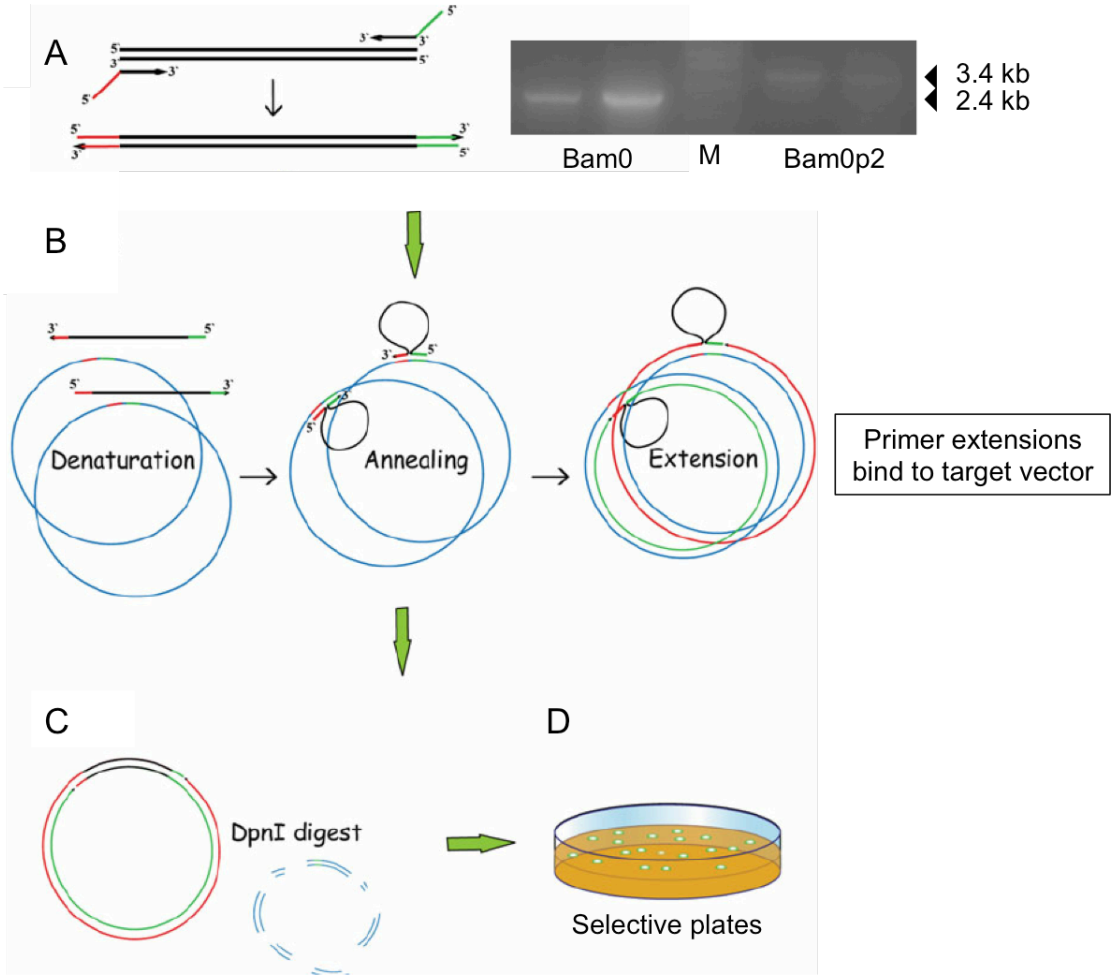

Figure 4.4.2.1: Cloning inserts via overlap extension PCR.

(A) Primers with vector-specific extensions are used to amplify the insert resulting products are Bam0 at 2.4 kb and Bam0p2 at 3.4 kb (B) in a second reaction, the insert and vector are mixed so that the insert acts as a primer to extend around the vector backbone to produce a nicked plasmid (C) empty vector is destroyed by *Dpn* I so that only clones get transformed. (D) Plating the transformation on selective medium allows successful clones to produce colonies. M = HyperLadder 1kb (Bioline, UK). Image adapted from (Bryksin & Matsumura, 2010)

desired constructs were selected by spreading the transformation mix onto selective plates. After overnight incubation at 37°C, the results continued with the theme of no colonies. There were still no colonies even after 36 hours at 37°C. Due to time constraints this was only attempted once.

4.5 Molecular tools using the $\phi 24_B$ tail spike protein

Work so far has described attempts to gather more information about BamA and how altering the host side of the interaction leads to changes in support for adsorption and infection. On the phage side of the interaction we also need more information on the tail as this is the structure that recognises the bacterial host and mediates adsorption. Currently all information on the phage tail and other structural genes are very limited.

The $\phi 24_B$ particle is categorised as a member of the *podoviridae*, phages characterised by a very short tail (~10 nm in this case), and so is a relatively small phage. Smaller phages generally form larger plaques, as the particle is able to diffuse through the agar more readily. However, plaque size for $\phi 24_B$ is very small, which is at odds with this conventional wisdom regarding smaller phages. A potential explanation for this is that the interaction between the phage tail and BamA is stronger than typical phage-receptor interactions. In combination with the presence of BamA (often upregulated in conditions that induce the lytic cycle (Veses-Garcia *et al.*, 2015)), newly formed phages bursting from a host cell may associate with BamA molecules in the debris from that host cell and therefore be

unable to infect a new cell and increase the area of infection thereby not allowing the plaque size to increase at the faster rate.

It is established that the presence of tail associated fibres increases the rate of adsorption (Hendrix & Duda, 1992) as they allow an extra step in host recognition by the phage particle, which tethers the phage to the cell, while the tail/base plate finds the specific adsorption target (Rakhuba *et al.*, 2010). Despite not possessing any such tail fibres, $\phi 24_B$ adsorption appears to occur more quickly than expected with just a tail spike protein. Once again, $\phi 24_B$ displays evidence of a strong association between the tail and BamA.

As there are very few structural genes in the $\phi 24_B$ genome (Fig. 4.5.1) and none of these genes seem large enough to provide enough of a surface to facilitate interaction, it is predicted that only the tail spike protein is involved in host recognition. There are no published data regarding the structure of this phage tail, despite its prevalence in Stx phages and so information is limited to an inspection under the transmission electron microscope which revealed a tail structure of ~10 nm x 10 nm and a bioinformatic approach to estimating 3D structure by Prof. Olga Mayans (personal communication, 2014), which revealed a collagen-like triple helix with a globular domain on each end.

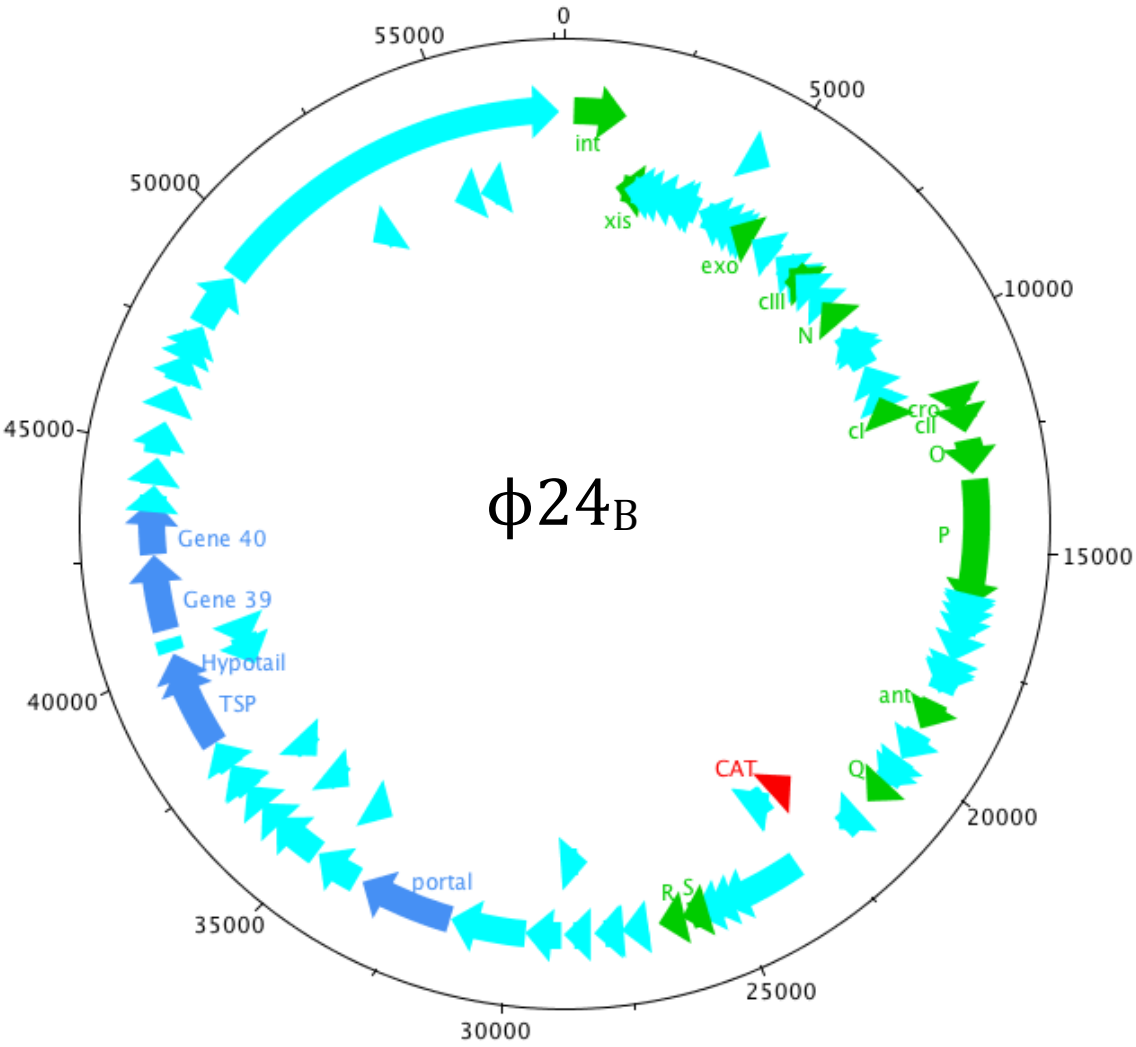

Figure 4.5.1: representation of the ϕ24$_B$ genome (57677 bp).

Genes encoded on the forward strand are on the outer track and genes on the reverse strand are on the inner track. Genes typical of lambdoid phages are coloured and labelled in green, the chloramphenicol resistance gene (CAT) is in red, and those structural genes suspected to form the tail structure are dark blue. All other genes are in light blue. Diagram created using Artemis (Rutherford *et al.*, 2000) and DNAplotter (Carver *et al.*, 2009) with the genome sequence published by Smith *et al.*. (2012).

Attempts to improve on this information with the use of negative staining with uranyl acetate under a Transmission Electron Microscope (TEM) were of a low resolution and were therefore unable to clearly see any tail structures (Fig. 4.5.2A-C). These images were not an improvement on those previously collected (Allison *et al.*, 2003). The tail has been visualised more clearly recently in the lab of Andrey Letarov using cryo-EM (Fig. 4.5.2D-E), but images have not been collated and analysed yet. Given the small size and limited decoration of the tail, no further insight on adsorption can been drawn before these data are analysed.

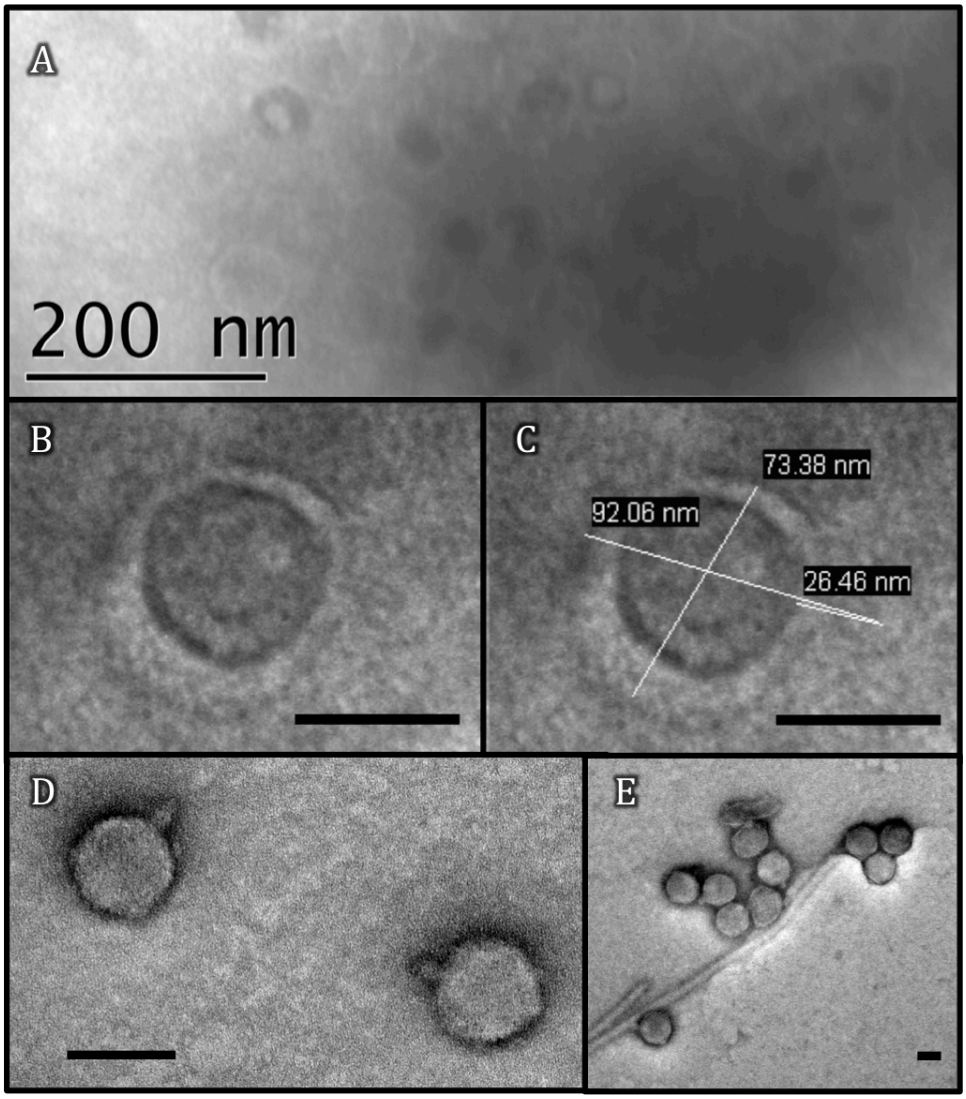

Figure 4.5.2: Negatively stained EM images of ϕ24B

(A) TEM image of a cluster of $\phi 24_B$ stained with 4% (w/v) uranyl acetate (B-C) identical images with measurements overlaid on (C) to show key dimensions of the phage capsid and tail. (D-E) Cryo-electron micrograph images donated by Andrey Letarov (Winogradsky Institute of Microbiology) of negatively stained frozen $\phi 24_B$ preparations ready for cryo EM analysis. Scale bars are (A) 200 nm and (B-E) 50 nm.

4.5.1 Creation of a GFP-TSP fusion gene

A useful molecular tool to investigate adsorption would be a GFP-tagged tail spike protein (TSP). This would allow visualisation of interactions between BamA and TSP under the fluorescent microscope. As such, it would also provide evidence as to whether the tail spike is able to bind to the host cell unassisted by any as yet unidentified accessory proteins. The fusion of two soluble proteins usually requires a linker region to allow space between the proteins for efficient folding (Klein *et al.*, 2014) and to prevent the inadvertent blockage of any functions {Chen:2013fd}. As such, a short flexible alanine-glycine bridge (AGAGAG) was designed into this gene fusion to be introduced via primer extensions between GFP and TSP.

TSP proved to be a troublesome amplification target. There are sequence repeats, which is common for tail spike genes that produce a collagen-like tertiary structure. These repeats are likely to be the reason for every attempted PCR amplification resulting in at least a partial smear (see Fig. 4.5.1.1A for an example). There is evidence to suggest that amplification of TSP could be stabilised by magnesium concentration (Fig 4.5.1.1B). However, this stabilisation effect disappears when attempting to fuse eGFP upstream of TSP (Fig. 4.5.1.1C). Altering annealing temperatures (35°C to 50 °C), DMSO concentration (0% (v/v) to 5%(v/v)), and added Mg^{2+} volume (0 µL to 5 µL) yielded the same smeared products. Ultimately, the ratios of the GFP and TSP fragments and only running 25 cycles instead of 30 were the only factors which helped to generate a putative fusion product with minimal smearing (Fig. 4.5.1.1D). A GFP to TSP ratio of 2:7

was the highest that could be achieved with the yield of TSP PCR product on that occasion (only 7 ng μL^{-1}). This reaction produced a band at around 2.7 kb, the expected size of a GFP-TSP gene fusion. The remainder of this PCR reaction was separated on another agarose gel but the ~2.7 kb band was not visible under blue light. The presence of the band was confirmed using UV light but the potential damage to the DNA caused by this imaging led to abandoning this sample and cloning was not attempted. Instead, the ~2.7 kb band was generated again using the same GFP and TSP PCR products as templates and extracted using the PCR and Gel kit (Bioline, UK). The DNA from this extraction produced a low yield (3.3 ng μL^{-1}), which did not successfully clone using the pCR-blunt kit (Invitrogen, UK). The putative gene fusion PCR product was not observed again despite multiple attempts using the same conditions. The only difference between this and the reactions that did produce the 2.7 kb band was that a different TSP PCR fragment was used as a template. Further re-amplifications of the TSP fragment did not facilitate the appearance of a band for the TSP-GFP fusion.

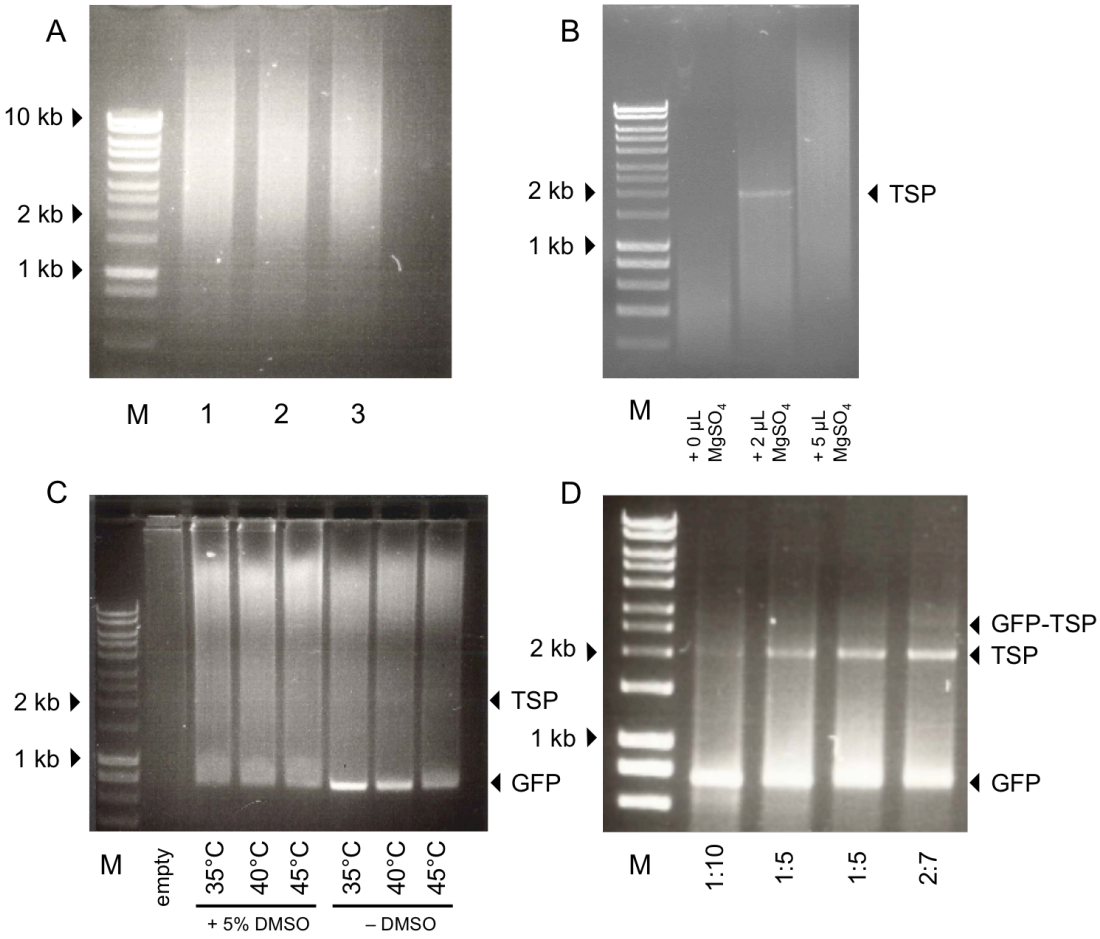

Figure 4.5.1.1: Amplifying tail spike protein (TSP) by PCR.

(A) Initial attempts at amplifying TSP by PCR led to smeared products (reaction was set up in triplicate) (B) Increasing the Mg^{2+} concentration in the reaction by adding 0 µL, 2 µL or 5 µL $MgSO_4$ to the reaction mixture produced a band at 2 kb. (C) Temperature gradient of the fusion reaction in the presence and absence of DMSO using the same additional 2 µL of $MgSO_4$ concentration as used in (B). (D) Increasing the ratios of TSP to GFP molecules in a reaction by using 75 ng of GFP loaded with 17.5 ng, 35 ng, 35 ng, and 52.5 ng of TSP respectively. M = HyperLadder 1 kb (Bioline, UK).

The smearing problem with TSP has been observed in the lab previously, although a cause for it had not been explored experimentally. Upon checking the TSP gene for regions of similarity/identity, there is a 45 bp tandem repeat starting at 1570 bp and 1615 bp in the TSP gene (found using a tandem repeats finder (Benson, 1999)), which could contribute to the appearance of the smears. These sequences of DNA could re-anneal to each other incorrectly and incrementally increase or decrease the size of PCR products by ~45 bp at a time.

A further explanation was raised when a sample was taken immediately after the primers were added at the mid-point of the fusion reaction and run on an agarose gel alongside the final PCR product. Comparison of the two stages of the same reaction revealed that the smearing only seemed to occur during the second stage of the fusion reaction (Fig. 4.5.1.2), indicating the addition of the primers as a potential cause as they are the only component absent during the first stage of 15 cycles.

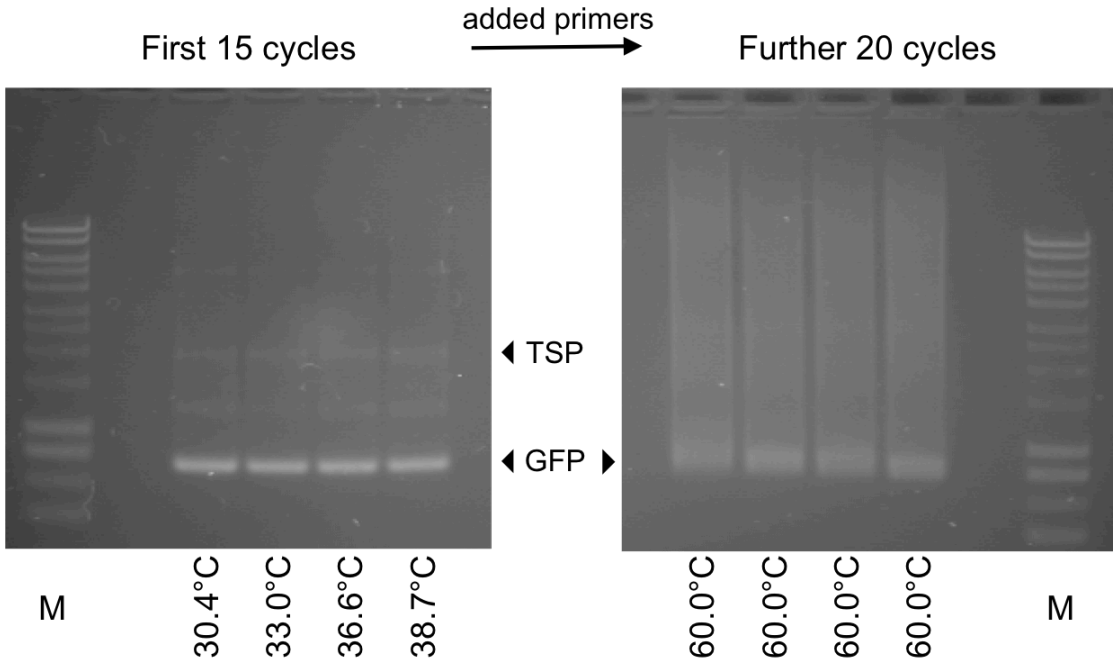

Figure 4.5.1.2: Comparison of DNA product mid- and post-primer addition in a fusion PCR amplification.

Samples (5 µL) of a 50 µL PCR amplification were taken after the primers were added at 15 cycles (left) and from the final PCR product (right). DNA after 15 cycles still shows as bands (bands at sizes that are not labelled are unidentified products), yet the band for a fused GFP-TSP was unable to be reproduced. All of the DNA appears as a large smear after the second stage and a total of 35 cycles. M = HyperLadder 1 kb (Bioline, UK)

4.5.2 Fusing GFP and TSP by introducing a *Not* I site

As attempts to reform the GFP-TSP band by PCR only yielded smears, it was decided that the second PCR step of the fusion strategy was too problematic to efficiently create a GFP-TSP gene fusion. As the initial TSP PCR was capable of producing a band, an alternative strategy was developed that added a *Not* I recognition site to the alanine-glycine bridge between GFP and TSP during that initial fragment creation. This was done using primers that substituted 2 base pairs of the bridge linker changing one of the glycine residues to alanine (Fig. 4.5.2.1A). After the PCR products have been digested using *Not* I the GFP and TSP genes can be joined together by a DNA ligase-mediated reaction.

The resulting gene fusion would encode the GFP-tagged TSP protein. However, the *Not* I-TSP fragment could not be generated as the reaction only produced smears under a range of temperatures and concentrations of DMSO (Fig. 4.5.2.1B-C). GFP also appears to run at 1.3 kb which is too high. The reason for this was not pursued. Producing the GFP-*Not* I and *Not* I-TSP fragments was only attempted twice at this range of temperatures (41°C to 55°C) and DMSO concentrations (0% (v/v) - 10% (v/v)) before focus was shifted to other avenues of research.

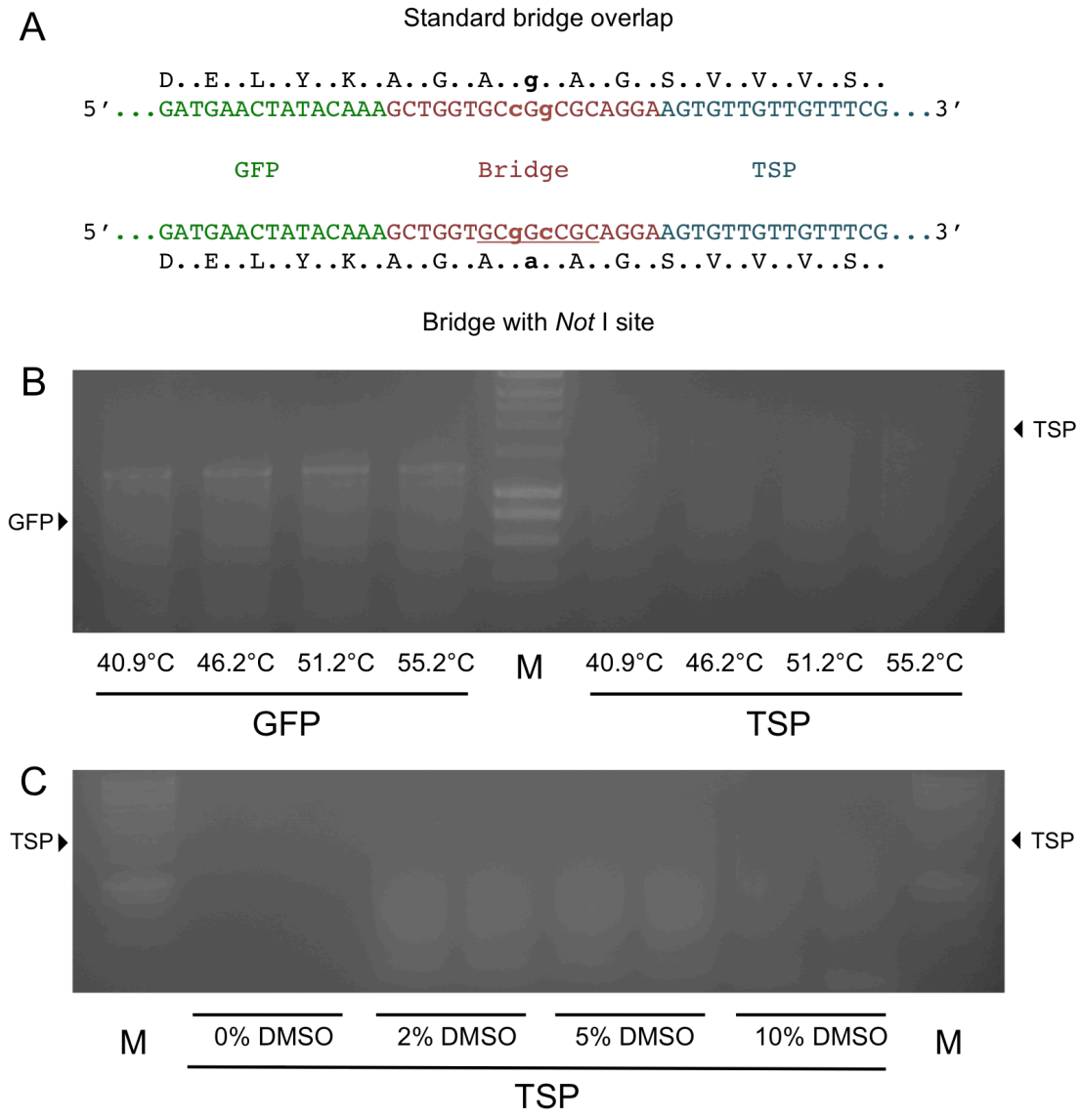

Figure 4.5.2.1: Attempting to create fragments for fusion using *Not* I.

(A) The DNA and encoded residues of the bridge linker and flanking regions of the planned GFP-TSP construct with the bases and residue (lower case and bold) to be changed to generate a *Not* I restriction site (underlined). (B) Temperature gradients for both overlapping fragments with TSP amplifications producing smears. (C) Increasing DMSO concentration does not facilitate the creation of stable TSP bands rather than smears in the amplification reaction. M = Hyperladder 1 kb (Bioline,UK)

.

4.6 Discussion of results

While the decision to go ahead with the fluorescence microscopy assay had been made, the creation of molecular tools for the other assays considered in this thesis are still important for future work.

The creation of the MC1061-bam5 strain was attempted using the same recombineering method used to create Bam0, yet it was unsuccessful. Also, the complementation of the Bam0 using the pUC19-FLbam5 plasmid (pUF5) did not restore phage adsorption. Both of these results do not support the hypothesis that substituting *E. coli* extracellular epitopes into the β-barrel of *P. atrosepticum* BamA would restore phage adsorption. The hypothesis is based on the fact that the amino acid sequence of the *P. atrosepticum* β-barrel would then be identical to the *E. coli* β-barrel domain. Together, the failed production of bam5 and the lack of complementation with FLbam5 suggest that there is a potentially unforeseen issue with regards to, for example, protein expression or correct folding. It is likely that the issue lies in an area that prevents the protein from being expressed appropriately from these constructs because there is no theoretical reason for all *E. coli* extracellular epitope substitutions to cause the BamA protein to fail to function, especially considering that MC1061 underwent the allelic exchange of EEKS(0) to create Bam0 with no observable detrimental effects. It may be that the protein is being expressed and folding correctly and another residue of BamA is required to maintain the essential protein insertion function, as there will be some minor differences in primary structure between the homologues (Fig. 4.6.1).

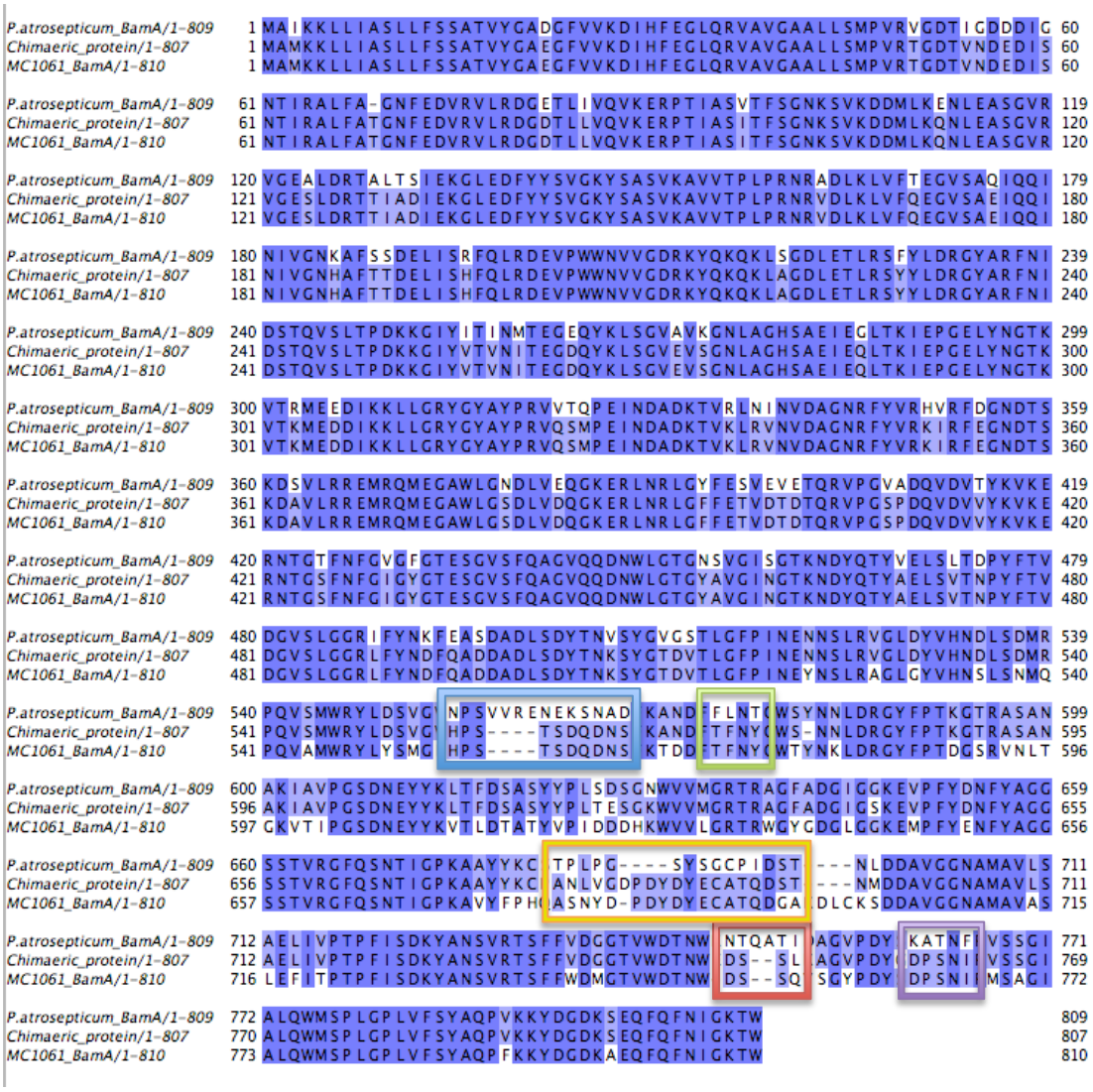

Figure 4.6.1: Sequence alignment of BamA proteins

Sequence alignment of BamA homologues from *E. coli* and *P. atrosepticum* compared to the Chimaeric BamA. Identical residues are shaded in dark blue and residues that differ are not shaded. Loops are coloured as follows: 4a (blue), 4b (green), 6 (yellow), 7a (red) and 7b (purple). Alignment was performed using MUSCLE and visualised in Jalview.

When deciding where to go next with the native promoters (section 4.4), an RNA extraction should have been done to determine if any FLbam5 transcripts were present. However, with other tasks on the go, I tried to go ahead with a workaround without dissecting the problem properly first. The thinking behind this at the time was that lack of functional rescue in the infection assays meant that seeing no (or very low) expression of Bam5 was the most likely outcome. If that were the case then FLbam5 could be subcloned into pBAD to determine expression from the arabinose-inducible araC promoters instead. Alternatively, if expression was occurring, but at lower levels than native expression (assuming that native levels of expression would rescue the infection phenotype), then there was no way to increase induced expression from lac promoters beyond the levels already attempted. FLbam5 could be subcloned into a dose-response system i.e. the araC promoters of the pBAD system, so checking mRNA production levels was omitted due to time restrictions. Ultimately, FLbam5 was not cloned into the pBAD myc His C plasmid so whether this system would have circumvented the issue or not cannot be concluded at this time. The pKT230 plasmid may be a better expression vector to try to being a low copy number and having been successfully used in the past (Smith *et al.*, 2007).

Further characterisation of the genes associated with the tail structure is still required. Ammonium molybdate has been used as a gentler contrast stain for electron microscopy to highlight fine strucutral details in some phages (Pickard *et al.*, 2010). This may have allowed a higher resolution image to be generated when using the TEM on our non-frozen samples. Whichever stain Andrey Letarov is using (unknown at time of writing) is clearly working very well to give

superior detail on single images. The result of averaging all the images from this cryo EM work will be very interesting to see.

Due to the difficulties in amplifying clean PCR products containing TSP further information regarding the $\phi 24_B$ tail remains elusive. The cloning failures with TSP derived constructs could be rooted in the smearing effect. If there is mismatched annealing of the DNA strands there would be 3' overhangs for which the complement strand could not be synthesised. Consequently, the compatible ends required for either of the pCR Blunt and TA cloning kits (both Invitrogen, UK) would not be present and the cloning would always fail, resulting in the characteristic lack of colonies for this chapter.

While the difficulty in developing molecular tools for other lines of experimental investigation were ultimately disappointing, a fluorescence microscopy assay was developed in an effort to examine $\phi 24_B$ adsorption and DNA ejection. As the components required for beginning the development of the assay had already been made, the work could continue without needing to accommodate any plans for future work e.g. other BamA mutants, during the process.

Chapter 5: Developing a fluorescence assay for detecting adsorption and DNA ejection stages of $\phi 24_B$ infection.

5.1 Introduction

This chapter follows the development of a fluorescence microscopy-based adsorption assay for the purpose of determining which *E. coli* BamA epitopes are required to support $\phi 24_B$ adsorption. All data acquired during this study can be found in Appendix 8.2 via this link to a space on the university website: https://pcwww.liv.ac.uk/~hallison/Stuart_Thesis/index.html

5.1.1 Assays for investigating adsorption

Previous chapters have explored the avenues of discovering the key differences between BamA produced by *E. coli*, which is capable of supporting phage adsorption, and the *P. atrosepticum* homologue, which is resistant to adsorption. This chapter details the initial development of a fluorescence-based assay that is designed to detect positive adsorption and/or infection phenotypes at the resolution of a single cell, eliminating the need for adsorption assays, which generate noisy data. The noisy data could be caused by a number of factors including that the adsorption assay only measures the phages that failed to adsorb and so there is a possibility that free phage particles get trapped and pelleted during the centrifugation process, decreasing the phages detected and increasing the perceived adsorption efficiency. Also, the 15 minute incubation

time to allow for adsorption to occur could allow enough time for the first phages to bind to initiate replication, increasing the number of phage particles and thereby decreasing the perceived adsorption efficiency.

The aim of the work detailed here was to detect differences in the adsorption and genome ejection phenotypes displayed by the library of mutants from Chapter 3. This assay should be able to corroborate conclusions from simple plaque-based experiments while also providing a greater resolution of what happens at the single cell level of binding, phage genome ejection and then infection for various BamA mutants from the library.

5.2 Pilot Study

Initially, a pilot study was performed in order to determine whether the cells infected by $\phi 24_B$ labelled with SYBR gold could be visualised and, if they could, which of the available fluorescence microscopes (the Zeiss Axio Examiner Z1 (upright) or the Zeiss Axio Observer Z1 (inverted)) was the most appropriate to use for answering the following:

- Can SYBR gold stained phages be visualised under the microscope?
- Can individual adsorption events be seen and how quickly do they occur after adding phages?
- Can genome ejection be visualised and how long does that take?

5.2.1 Materials and methods used for data collection and analysis

The microscopes used for the pilot study were the Zeiss Axio Observer Z1 (inverted) epifluorescent microscope and the Zeiss Axio Examiner Z1 (upright) LSM 800 multiphoton microscope. Images were captured by an iXon Ultra 897 high sensitivity camera. Data collection was performed using the multi-D acquisition tool in Micro-Manager version 1.4.14 with the settings in Table 5.2.1.1 and all data for analysis were extracted from stacked .tif files using FIJI (Schindelin *et al.*, 2012)

Table 5.2.1.1: The relevant fluorescent channels available during the pilot study

Channel	Channel "ID"	Wavelengths of excitation (nm)	Wavelengths of emission (nm)	Exposure time (ms)
Light	DIC	n/a	n/a	50
DAPI	DAPI	355-375	420-470	200
SYBR gold	GFP	450-490	500-550	400
PI	dsRED	538-562	570-640	200

5.2.2 Results of the pilot study

To label $\phi 24_B$ particles, 10 µL of SYBR gold stain (Thermo Fisher Scientific) was added to a 1 mL sample of phages (2.8×10^9 PFU mL^{-1}). This sample was dialysed against 2 L phage growth medium (PGM) using dialysis tubing with a 12 kDa – 14 kDa cut off (Medicell International, UK) for 2 hours at room temperature to remove the free SYBR gold stain. DM1187 cells were pre-infected with SYBR gold

stained phages at an MOI = 1. These cells were observed using both the Examiner Z1 (upright) and the Observer Z1 (inverted) epifluorescence microscope to compare the sample handling, data collection procedures, and image quality associated with each (Fig. 5.3.1).

Live motile cells were observed tumbling in solution (Fig. 5.3.1A). A fluorescent cell was tracked on the upright microscope (Axio Examiner Z1) for at least 60 s (5.3.1C) and although the sample had dried out by the time the sample was observed on the inverted (Axio Observer Z1), fluorescent cells could be seen trapped within crystal structures, presumably salt crystals, that had formed on the surface of the slide (Fig. 5.3.1B). All raw data for all imaging runs can be found in Appendix 8.2.

5.2.3 Conclusions of the pilot study

As observed (Fig. 5.3.1E) highly fluorescent *E. coli* cells were detected, the pilot was deemed to be successful. Additionally, the inverted epifluorescent Axio Observer Z1 was determined to be the best instrument to use because there was a high sensitivity camera already attached. There was no difference in the image quality or usability of the software packages. In order to get enough magnification (2500x), the 100x oil-immersion objective was used in conjunction with the 2.5x optovar, an optional additional lens that further increases the magnification.

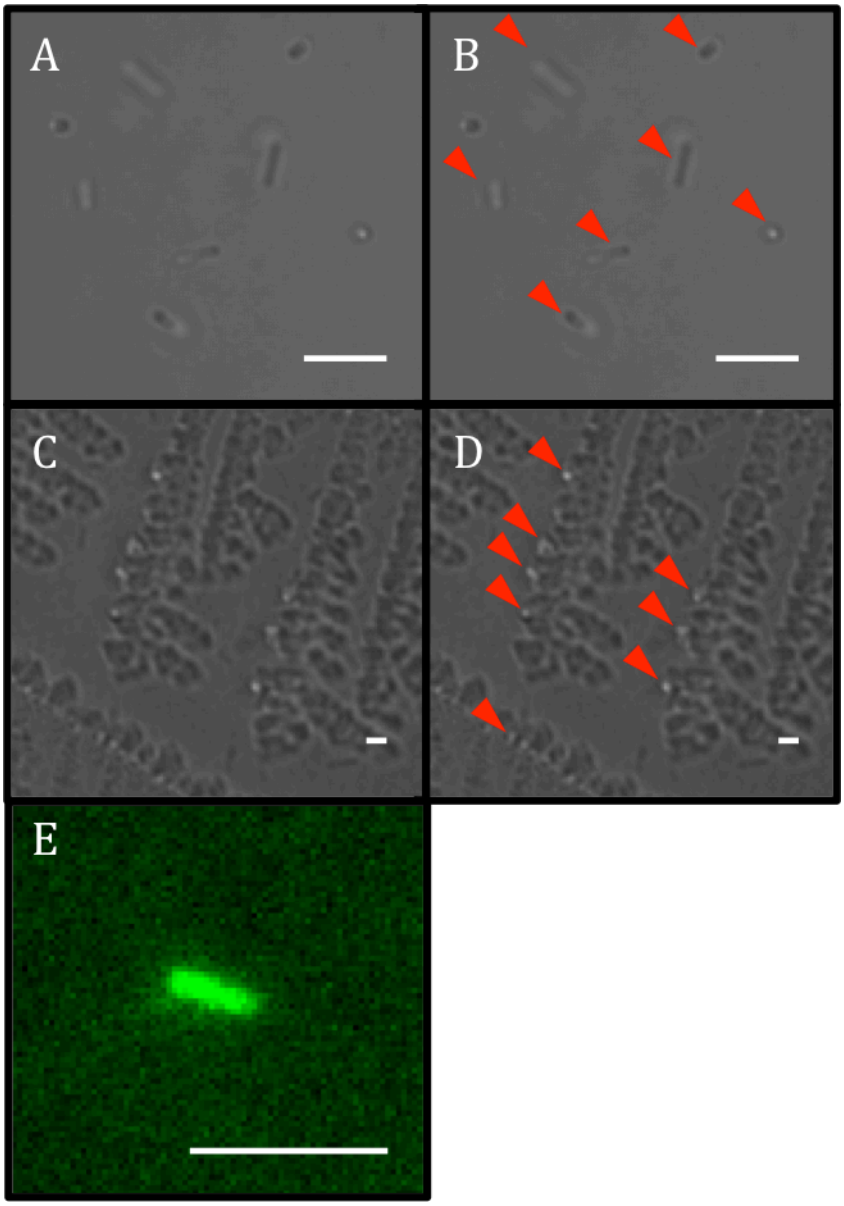

Figure 5.3.1: Images from the pilot study for the fluorescence-based assay. (A & C) *E. coli* cells tumbling in variable orientations/focal planes and trapped in crystal formations in dried out samples, respectively. (B & D) Duplicate image of (A & C), respectively, with cells indicated with red arrows. (E) Image taken whilst tracking a single fluorescent cell under the upright fluorescence microscope. Scale bars = 5 μm.

However, some issues were raised during the imaging process that needed to be addressed before running further experiments. An observation during the pilot study was that tumbling bacterium in solution (Fig. 5.2.1A) could not be tracked easily for the intended 40-minute duration of the experiment. Moreover, successfully tracking a single motile bacterium in 3 dimensions for the extended timeframe adds the disadvantage of ensuring only one cell will be in focus for the duration and, therefore, severely limits potential analysis. To limit the impact of tumbling and motility, the decision was made to fix the cells to the surface of the microscope slide to enable the viewing and analysis of multiple cells simultaneously for the entire 40 minutes. Two methods were explored to attach the cells to the surface: using agar or poly-L-lysine.

For the agar method, 50 µL of 0.4% (w/v) LB agar was dropped into concave wells (volume = 15 µL) within a microscope slide (Cat №: N/A144, Academy Science, UK) resting on a heat block at 45°C. The heat ensured the agar did not set prematurely or unevenly across the slide. The application and setting of the agar was not easy to replicate as some agar drops were so thin that they ruptured after setting and some even slid out of the concave wells. Furthermore, the domed nature of the drops meant that only a few cells remained on the agar and flattening the drop with a cover slip prior to setting led to an increased likelihood of the agar rupturing. The next treatment, the use of poly-L lysine avoided the use of agar.

Poly-L-lysine (PLL) treatment of the glass slides was based on a strategy used for atomic force microscopy (Lonergan *et al.*, 2014). The treatment consisted of

dropping 150 µL of 0.01% (w/v) PLL, pH 7.0 (Sigma, UK) onto glass-bottomed dishes (Cat No: 89125-438, VWR, UK) or microscope slides that were incubated at room temperature for 30 minutes before removing the remaining liquid with a pipette. As the cells were successfully attached to the surface of the PLL treated surfaces when viewed under the light microscope, and the treatment is more consistent and easier to apply than agar, the PLL treatment was used for further experiments.

Having found a good method to immobilise the cells, further parameter optimisation was undertaken. In the initial pilot study, the PGM containing the pre-infected cells was associated with high levels of what was thought to be background autofluorescence, which was thought to be caused by the high concentration of the unbound (or dissociating) SYBR gold stain that had not been removed from the medium before imaging. Improving the dialysis step could solve both of these possibilities simultaneously. The removal of PEG 8000 would have been heavily restricted using the 12 kDa -14 kDa MWCO membrane as the pore size was not 4x larger (the minimum recommendation) than the PEG polymer that needed to be removed, which had an average molecular weight of 8 kDa. The pore size of the membrane used was increased to 100 kDa MWCO (SpectraPor® CE tubing from Spectrum Labs, UK), which would easily accommodate release of the both the PEG 8000 (average mass of 8 kDa) and the free SYBR gold while retaining $\phi 24_B$ phages. The estimated radius of a globular protein of 100 kDa (the MWCO for the dialysis tubing) is 3.05 nm (Erickson, 2009). Since the $\phi 24_B$ particle has a radius of ~25 nm, it would easily be retained in this setup.

A practical limitation to using a microscope slide and cover slip was that phages and cell stains could not be added in a reliable and reproducible manner part-way through an experiment. Glass-bottomed dishes (Greiner, UK) eliminated this problem as any liquid could be added directly to the medium with the freedom to diffuse rapidly across the entire sample. This allowed phages to be added into solution after beginning to record the video without losing position or focus of the sample. Propidium iodide (PI) crosses compromised membranes and could be used to assay for dead cells after the assay had been performed. It can also be added before in order to monitor the timing of cell death.

The issue of the constant slow shift in focal plane away from the cells was caused by the slow drying out of the sample during the 40-minute imaging time. This drying out was countered by increasing the volume of PGM covering the surface of the glass-bottomed dish from 100 μL to 500 μL during the fluorescence infection assay. This resulted in a stable focal plane for much longer (up to 2.5 hours).

5.3 Visualising $\phi 24_B$ phages under the microscope

5.3.1 $\phi 24_B$ phages in solution

Staining $\phi 24_B$ phages with SYBR gold allows them to be seen in solution. They move in and out of the focal plane very quickly so they appear as "flashes", which the software is unable to track. Figure 5.3.1.1(A-C) shows phages (indicated by red arrows) that happen to be in the focal plane when snapshots are taken.

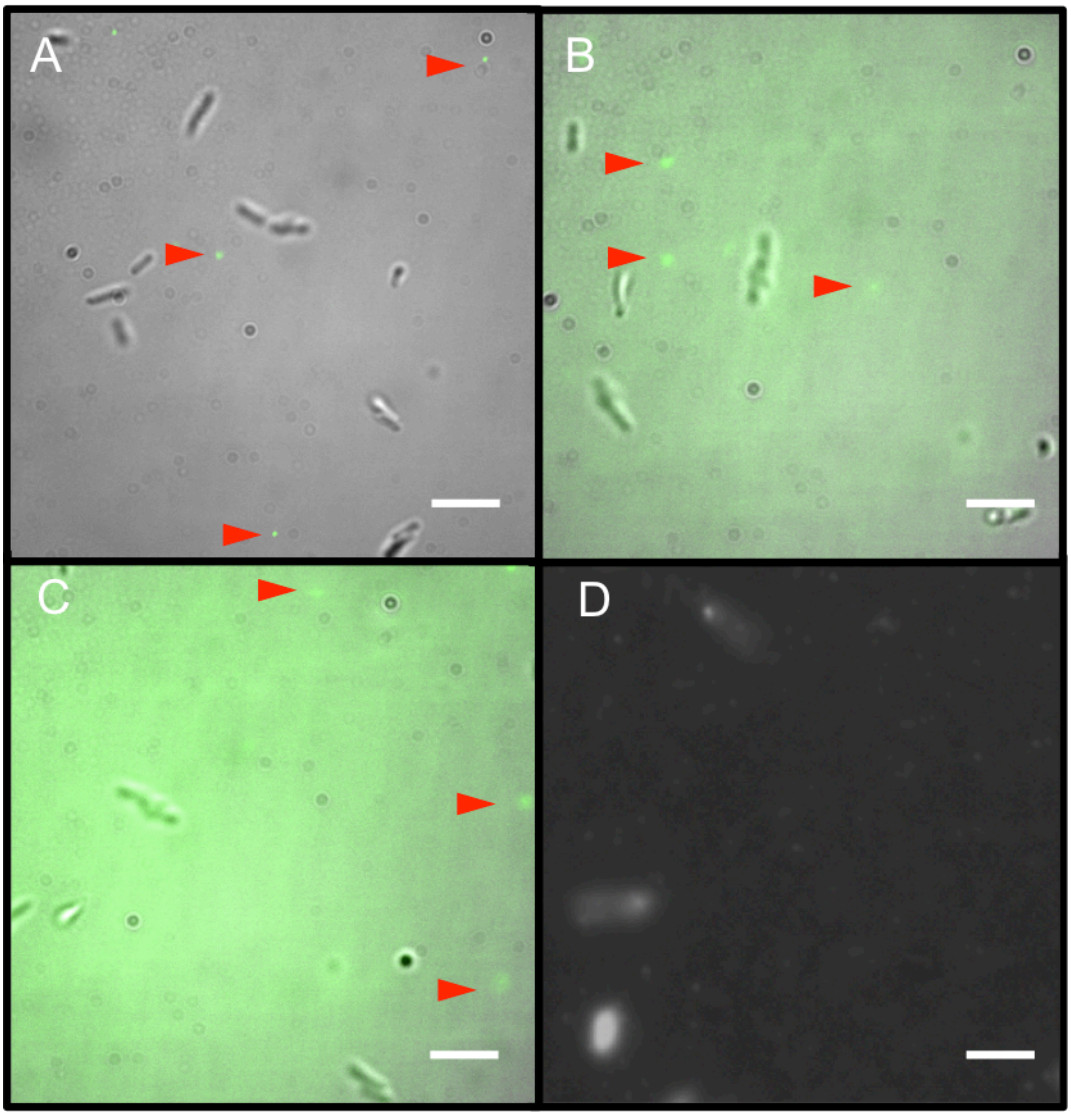

Figure 5.3.1.1: The presence of phages in solution during the experiments

(A-C) Snapshots of SYBR gold stained phages (indicated by red arrows) tumbling in solution at 140 s, 260 s, and 330 s, respectively. (D) The signal produced using Channel ID "GFP" 2 hours after the addition of phages. The larger white areas are bacteria and the pinpricks of white are stained phages. Scale bars = 5 µm.

It is not only cells that adhere to PLL treated surfaces; phages also stick. This can be seen in figure 5.3.1.1D taken 120 minutes after the addition of phages. This decreases the number of potential infective particles in the solution by an indeterminate proportion. It is not clear how comparable this is to the reduction in the number of phages that would have diffused into an agar preparation, as a similar image was not taken on agar. After seeing the effect of the adhesive surface, and after increasing the volume in the dish to 500μL, the number of phages added per experiment was increased to a final concentration of 2×10^5 PFU mL^{-1}. The final MOI cannot be calculated accurately as the number of bacteria that were successfully immobilised cannot be determined.

5.3.2 Phages adsorbing to cells

Phages (added at 30s), seen as fluorescent dots, were observed associating, presumably adsorbing, to DM1187 cells by 60 seconds into the imaging run (Fig. 5.3.2.1). These dots remained associated with the outside of the cell until 450 s and 550 s where the fluorescent dots become indistinguishable from the SYBR gold signal accumulating inside the cell. This accumulation continues to increase further until 900 s when the imaging stopped.

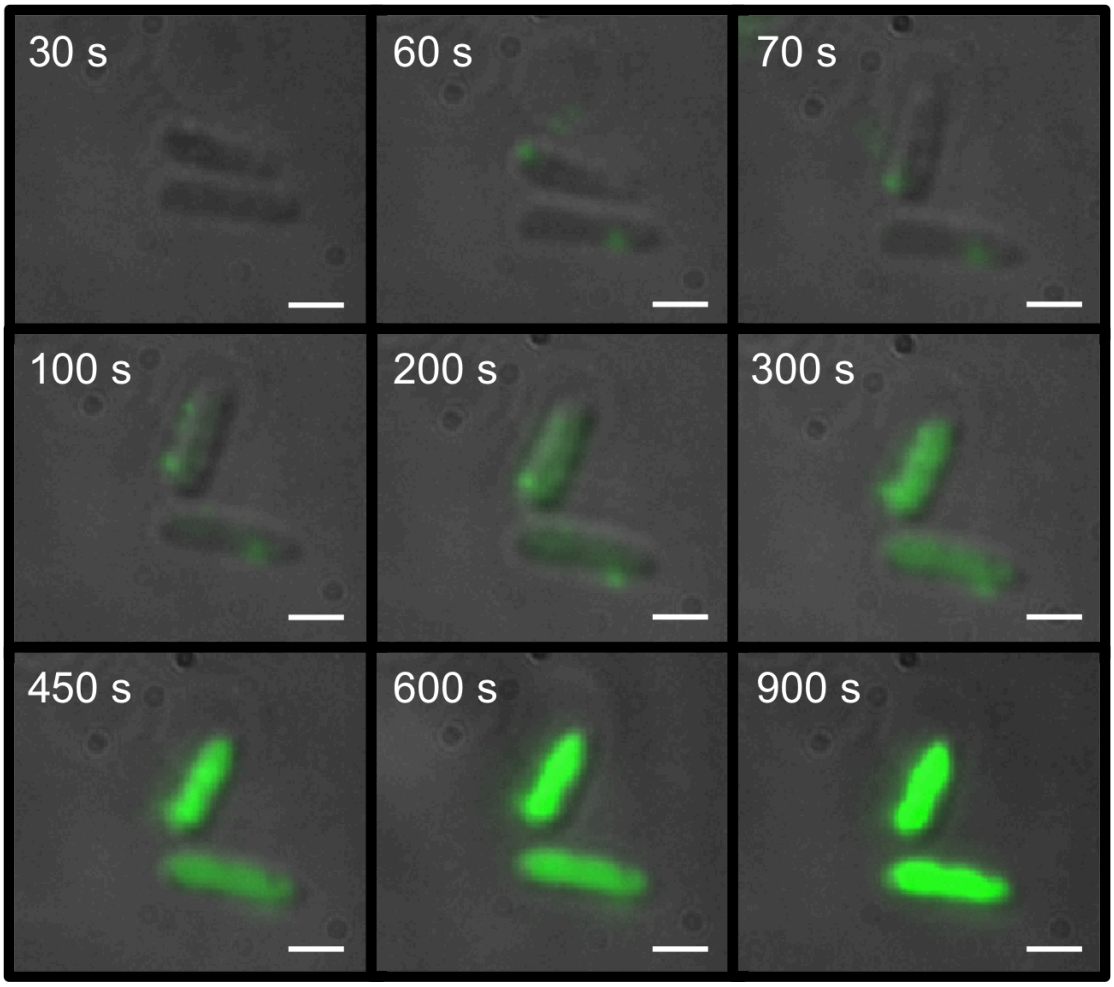

Figure 5.3.2.1: Visualising adsorption of SYBR Gold stained phages to DM1187.

10^5 Phages were added after 30 s. Fluorescent dots can be observed associating with the cells by 60 s. At 70 s one cell alters orientation but still shows an associated phage. By 200s the cell begins to accumulate fluorescence indicating a transfer of SYBR gold labelled DNA from the phage particles into the cell. Over the course of 900 s (15 minutes), the fluorescence accumulates into a strong fluorescent signal from the cell. Scale bars = 1 μm.

5.3.3 Phages ejecting DNA

Breaking down the intervals in figure 5.3.2.1 further shows a probable DNA ejection event. In figure 5.3.3.1, the fluorescence from the phage adsorbed at the pole of the DM1187 cell appears to be transferred into the intracellular environment, which becomes increasingly bright over the course of the experiment (Fig. 5.3.2.1). This method combined with phage host specificity has been exploited to detect specific species in a mixture (Mosier-Boss *et al.*, 2003) so this transfer of SYBR gold fluorescence was the expectation of what would indicate a DNA ejection event.

The diffuse nature of the fluorescent signal after ejection could be the SYBR gold stain dissociating from the phage DNA and associating with cellular DNA (or RNA), which is present in greater abundance and so has the potential to dilute the signal. There are data to suggest that the SYBR gold label is predominantly retained by the phage DNA post-ejection (Kunisaki & Tanji, 2010), which would imply that there were other infection events by a greater number of phages not visualised during the assay.

When quantifying the fluorescence of a labelled phage it was important to avoid fluorescence signals located within the cell as much as possible. Therefore, the area analysed was limited to only include a phage-sized area at the edge of the cell, which then needed to be adjusted for any slight movement under the microscope. The circular area analysing the fluorescent profile of the phage on the top cell needed to re-centred at frames 6, 7, 8, 9, 12, 14, 21, 28, and 37 to

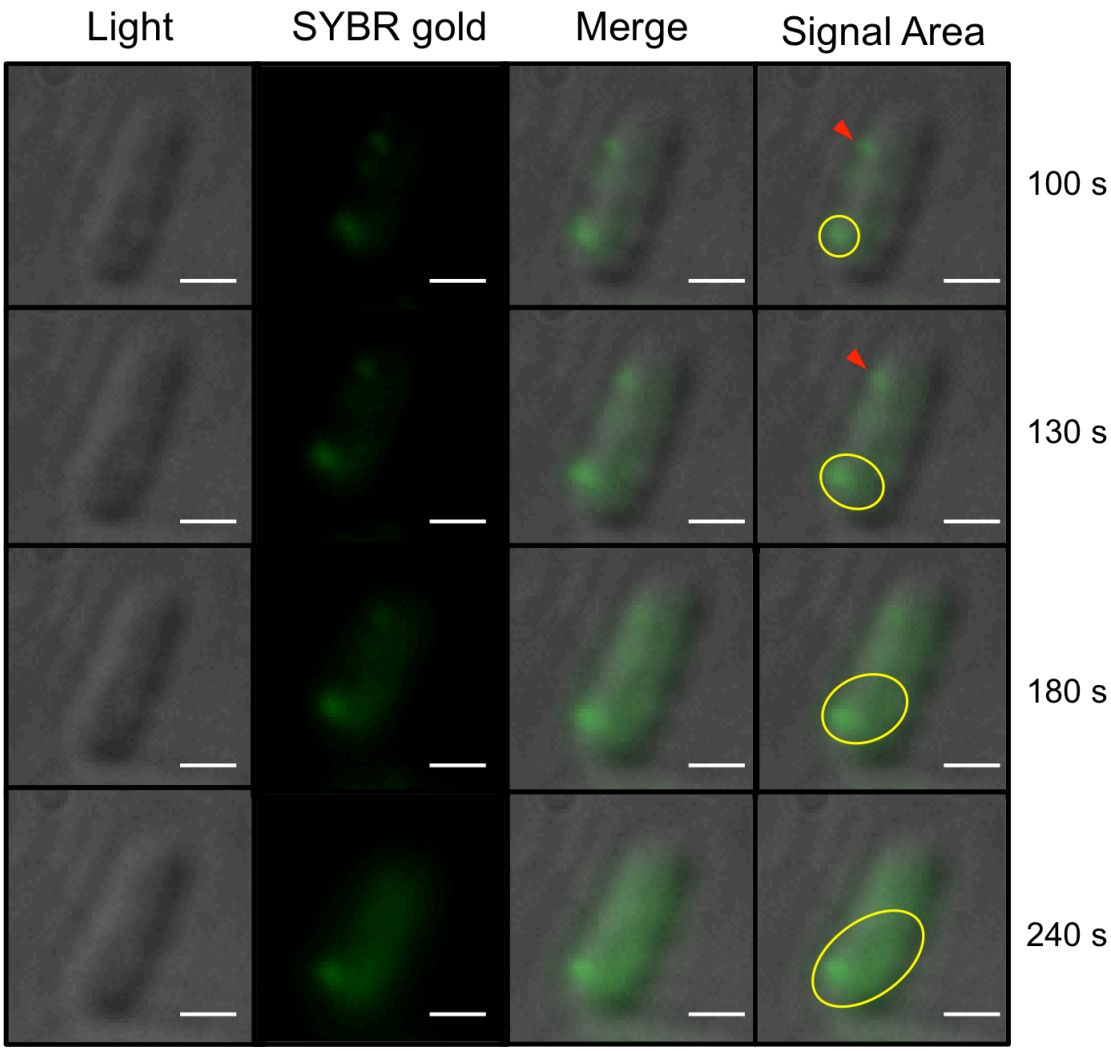

Figure 5.3.3.1. Visualisation of phage DNA labelling with SYBR gold transferring from ϕ24$_B$ particles to a DM1187 cell.

Fluorescence signal from the SYBR gold label can be seen accumulating inside this cell between 100 s and 240 s of the imaging collection. The red arrow shows another phage adsorption event that is not analysed here. For illustration purposes, the approximate area of SYBR gold penetration from "phage 1" into the cell is circled in yellow in the final column. Scale bars = 1 μm.

ensure the phage signal was within the area being analysed. The same area covering the phage on the bottom cell was more stable and only needed re-centring at frames 13, 14, 15, 19, and 25. The fluorescence of either phage 1 or 2 (Fig. 5.3.3.2A) could not be distinguished from cellular fluorescence after frame 55 and frame 45, respectively. Phage-associated fluorescence remains constant relative to cellular fluorescence for the first 200 s (approximately 3 minutes) before decreasing at a steady rate until the phage fluorescence can no longer be distinguished from the constantly increasing cell-associated fluorescence (Fig. 5.3.3.2B). The phage fluorescence represents the contribution of a single stained particle yet the increase in cellular fluorescence is most likely to be an ensemble effect of multiple infection events. This complicates any analysis as the total fluorescence amounts to a much greater signal than an individual phage, yet the presence of additional phages cannot be determined from the data collected. There is an additional phage visible in the same focal plane as "phage 1" in Figure 5.3.3.1 (red arrow - visible at 100s and 130s). Therefore it is likely that other phages have also absorbed to this cell in other focal planes, where they are not visible but do contribute to the overall cellular fluorescence over time. In this sample it would seem that ϕ24B phage are more likely to bind at the poles of the cell or diffusion into midcell is hindered in some way. This could be by a partially formed septum or the SYBR gold stain associating with peripheral genomic DNA, thereby slowing the progress of the label into the centre of the cell. BamA is present across the whole outer membrane with new BAM complexes assembling at mid-cell (Rassam *et al.*, 2015). More data collection is required before any conclusions can be made as to whether this is generally true of a $\phi 24_B$ infection or whether it was just these cells.

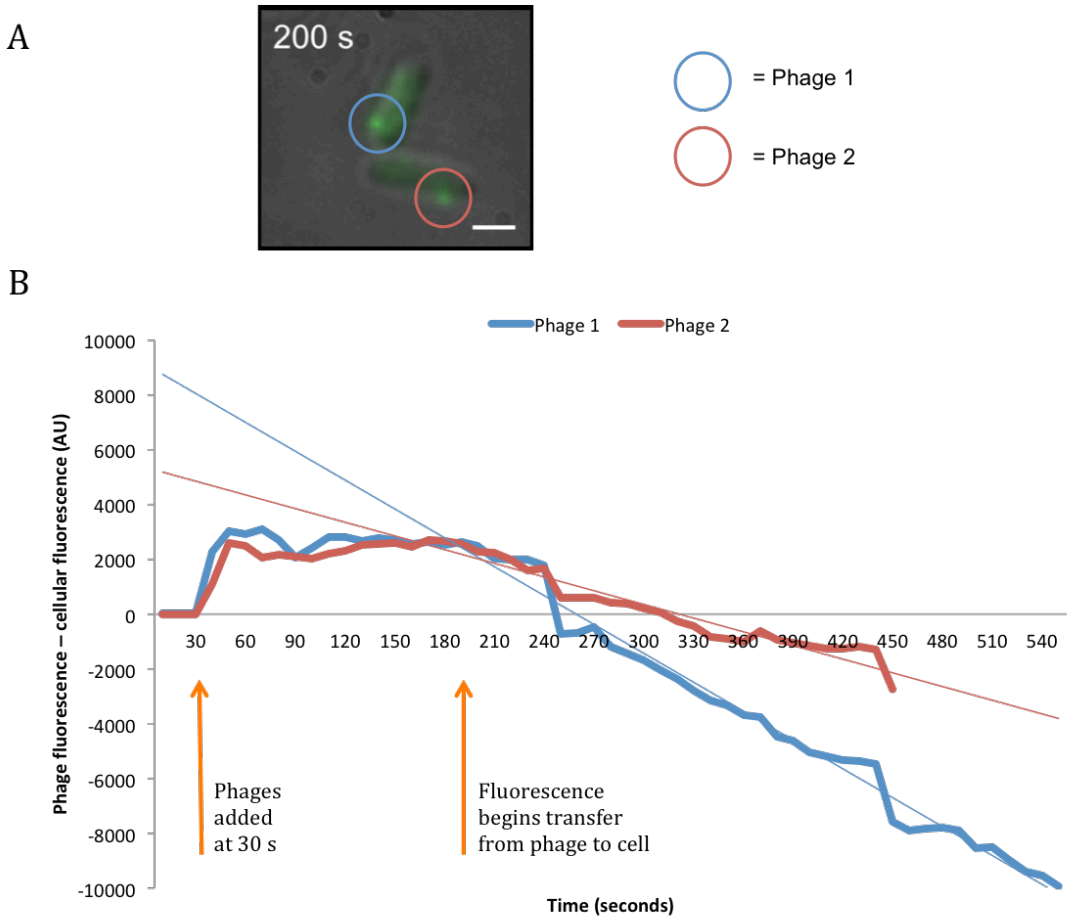

Figure 5.3.3.2 Quantification of figure 5.3.2.1 showing fluorescence associated with two individual phages adsorbed to DM1187 cells.

(A) An excerpt from figure 5.3.2.1 showing data sources for (B). (B) Graph showing the phage-associated fluorescence minus cellular fluorescence. Phages were added after 30s causes a signal spike, which is maintained until 180 s when fluorescence begins to transfer from the phage into the cell. Trend lines are plotted using data at 180 s and beyond only.

5.4 Examining the Bam0 strain

Initial attempts at measuring phage infection in DM1187 were normalised against background fluorescence to eliminate the contribution that any build up of background fluorescence (as seen in figure 5.3.1.1) would have to the total signal. A dilemma was encountered during the analysis when a decrease in SYBR gold fluorescence was detected; the decrease could not be definitively attributed to either the loss of fluorescence from the cell/phage or whether the focal plane had shifted towards the outside of the cell or the signal had dropped for any reason. To control for this, the stain "NucBlue" (Thermo Fisher, UK), which is a DAPI-based DNA stain, was used to identify bacterial DNA as a control for cell-associated fluorescence before imaging started. On the first attempt the NucBlue stain was present but the sample could not be processed as the microscope had the incorrect filter box installed so the light was filtered out before detection. The correct filter box was installed for all subsequent sessions.

5.4.1 Bam0 cells

Bam0 cells were shown to be resistant to $\phi 24_B$ phage infection (Fig. 3.4.1). For this reason, Bam0 should serve as a negative control for phage infection against the DM1187 strain used for the experiments. During these experiments "phage infection" was observed (Fig 5.4.1.1A). Phage infection would present as an increased cellular fluorescence from the SYBR gold stain. However, this fluorescence appeared suspiciously synchronised to all cells at once and, notably, also appears without the addition of any phage particles (data not shown),

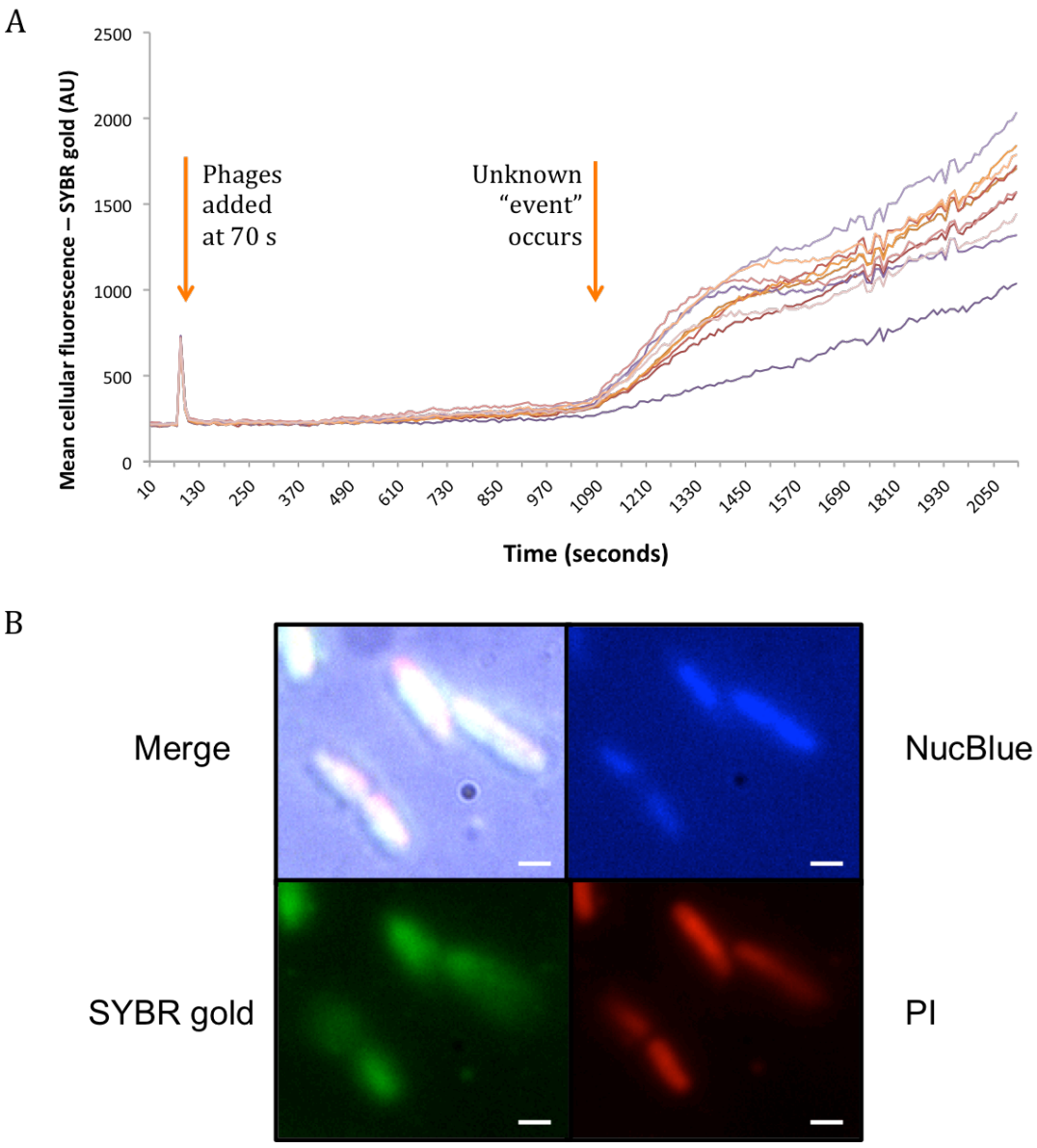

Figure 5.4.1.1: Tracking fluorescence during adsorption assay with Bam0.

(A) Quantification of SYBR gold fluorescence profiles within 10 individual Bam0 cells which demonstrates a synchronised, yet unexplained, "event" at ~1090 seconds. (B) All three labels are present at the end of the experiment. NucBlue shows bacterial DNA stained prior to the imaging collection, and while the SYBR gold would imply phage infection according to the assay design, the build-up of PI shows that the cells are dead and free SYBR gold could have diffused into the dead cell. Scale bars = 1 µm

thereby showing that the fluorescence build up is not caused by phage infection activity in this case. Other instances of the same observation were made using Bam0 cells both with and without the addition of SYBR gold labelled $\phi 24_B$ particles. The build up of SYBR gold fluorescence correlates with the build up of PI fluorescence (Fig 5.4.1.1B) and suggests that the "event" causes cell death.

5.4.2 Bam0 complemented with pUC19-ϕR1D

Bam0 cells containing the pUC19-ϕR1D plasmid were also subjected to the assay, though limited conclusions should be drawn from this as this was only performed with a single replicate. The complemented cells appeared very elongated under the microscope and seemed compartmentalised, where it looked as though the septum formed but cell division had not occured (Fig 5.4.2.1A-B) as the pUC19 plasmid causes a delay in cell division in order to ensure plasmid maintenance in both daughter cells (Ingmer *et al.*, 2001). Staining samples with PI after data collection showed an identical pattern i.e. all cells that are positive for SYBR gold after 20+ minutes are also positive for PI – including individual sections of the compartmentalised cell (Fig. 5.4.2.1C). The same event that affected the Bam0 strain may or may not have occurred during the assay of this sample too; however, it is not possible to confirm this as labelled $\phi 24_B$ particles were added to this single replicate, giving two competing explanations for the observations made. The premature cell death observed was either caused by an unknown event similar to that shown in figure 5.4.1.1. The compromise of Bam0 cells would limit any phage biology analysis, had there been any, to the first 19 minutes, where Bam0 cells are still alive.

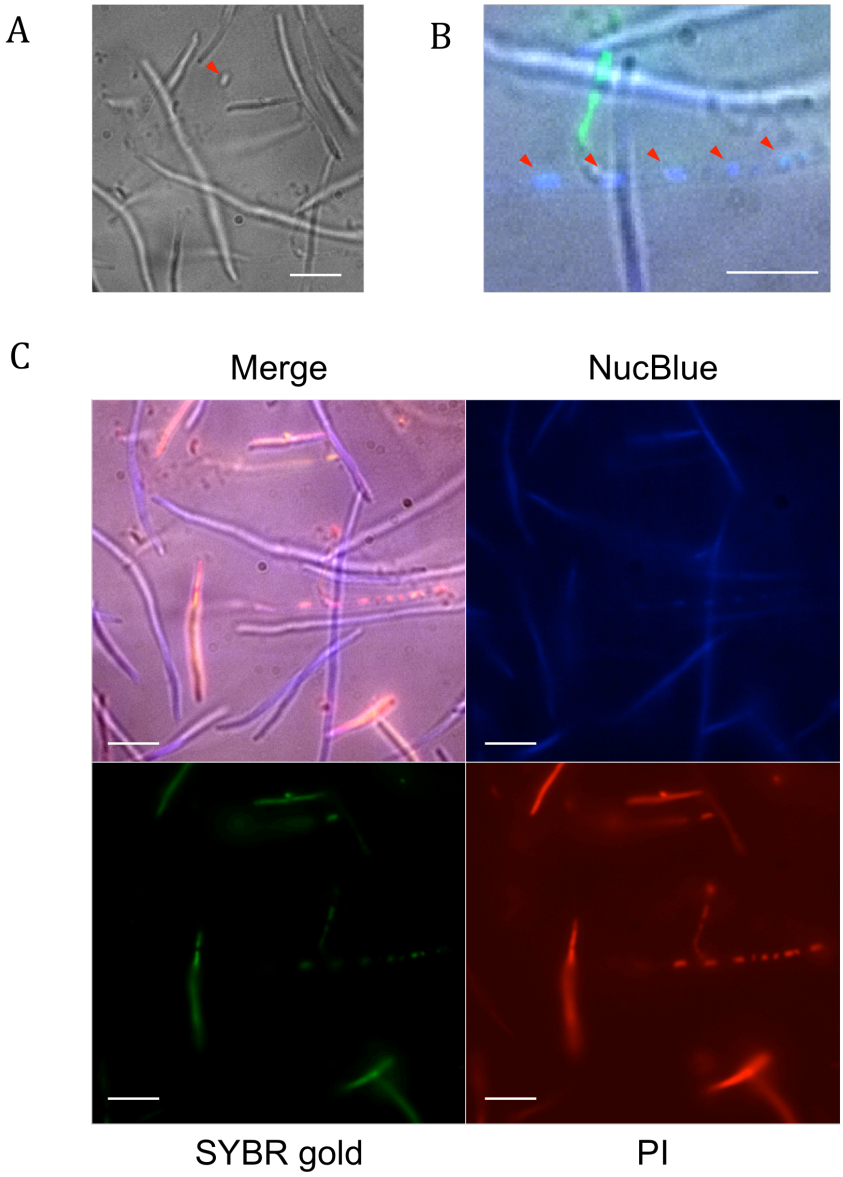

Figure 5.4.2.1: Observing Bam0 strain complemented with pUC19-ϕR1D.

(A) Bam0 pUC19-ϕR1D under a light microscope showing the extent of cell elongation. A cell with proportions matching previous experiments is highlighted by the red arrow. (B) Visualisation of the compartmentalisation of DNA in a an elongated Bam0 pUC19-ϕR1D cell. The red arrows show areas of NucBlue stain that are separated from each other despite being in the same cell. (C) Images of dead Bam0 cells taken at 1170 s on all channels. Not all cells "infected" (SYBR gold channel) but those that are have died (PI channel). Scale bars = 10 μm.

5.5 Conclusions

5.5.1 Visualisation of phages during the assay

The model phage used for the assay, $\phi 24_B$, can be visualised by an established and simple SYBR gold labelling method (Mosier-Boss *et al.*, 2003). Phages, in the form of fluorescent dots, were seen both in solution and adhered to the PLL-coated surface of the glass bottomed dish after the experiment had concluded

5.5.2 Visualisation of phage DNA ejection

There were examples of what appeared to be an infection in progress when tracking fluorescence transfer from the phage to the cell. One such event is presented in figure 5.3.3.1, where the build-up of cellular fluorescence is visible over time. Quantitative analysis of this event showed that fluorescence transfer into the cell has a lag phase of 3 minutes before beginning. This lag phase assumes instant adsorption, which is reasonable to make with the number of phages added to the experiment. An exact MOI is not calculable due to an unknown number of cells adhered to the dish.

5.5.3 Unexplained cell death

The synchronised accumulation of SYBR gold fluorescence in Bam0 cells suggests that the membranes lose integrity due to conditions in the assay causing cell death. While the reason for this compromised state is not

immediately apparent; although, as the no-phage control (PGM containing SYBR gold only) also showed cell death it is likely caused by the experimental setup, the death of Bam0 cells limits any potential conclusions to only the first 20 minutes after addition of phages in all experiments. Once the assay does not display any cell death, the adsorption phenotype of Bam0 can be determined with greater confidence and then, if negative, this can act as a negative control for testing other isogenic mutants.

5.6 Discussion of results

The simple infection assay from figure 3.4.1 has also been visualised under the fluorescent microscope in figure 5.3.1E. This successful infection by at least one labelled $\phi 24_B$ particle results in the production of a fluorescent *E. coli* cell, which was motile in the pilot study and immobilised on the surface of a glass bottomed dish in subsequent experiments. The ability to quantify areas of fluorescence allowed the analysis to hone in on the infections of single cells by multiple phages over a period of about 20 minutes. Visualisation of host cells is only possibly after infection by multiple individual phages as the signal to noise ratio is not high enough for a standard fluorescence microscope from a single phage (Mosier-Boss *et al.*, 2003)

The assay development is not finished, especially in light of the unknown conditions that cause bacterial cells to die more quickly than expected. A potential avenue that could be worth exploring include potentially revisiting agar as it is probably more effective at immobilising cells without killing them.

The main drawback of agar was the need to invert the slide and so adding phages during the imaging was not possible. The mid-assay addition of phages was important be able to measure the time taken between inoculation with phage and the start of DNA translocation, which appears to only take around 3 minutes. The upright microscope (Examiner Z1) could be used instead and would circumvent this issue. Even with this, the slide setup would need deeper divots or some way of making a more solid agar base to ensure the agar does not rupture or dry out in the timeframe of the assay.

Removing all of the PGM during the washing procedure caused many cells to group together, making contacting with each other, in the remaining microdroplets of medium. BamA is implicated in a primary role in contact-dependent growth inhibition (CDI) in *Enterobactericeae* (Ruhe *et al.*, 2013a). Currently, there are no data to suggest what impact CDI may have on either the protein insertion function or the $\phi 24_B$ adsorption process. So it was considered prudent at this stage to avoid cell-cell contact during experiments as much as possible. Therefore, the attached cells were washed by tilting the plate and running PGM from top to bottom to remove unattached cells while avoiding cells congregating in microdroplets. If it turns out that CDI does not significantly alter BamA activity, conformation, or interaction with $\phi 24_B$ then pipetting away all PGM would be a good way to group cells for imaging to get a higher density of cells and, therefore, more data per experiment.

5.6.1 DNA ejection from $\phi 24_B$

There are no data regarding the specific DNA ejection of $\phi 24_B$; however, as a lambdoid phage, an approximate comparison can be made to phage lambda, λ. It has been shown that λ has a constant DNA ejection time even in genome length mutants of 37.7 kb to 48.5 kb as the initial ejection rate increases for longer genomes (Löf *et al.*, 2007). The increase in initial ejection rate is caused by the increase in forces associated with DNA bending and inter-strand repulsion (Löf *et al.*, 2007) and so we might expect $\phi 24_B$ to have a similar ejection time to λ, even with a larger genome size of 57.7 (Smith *et al.*, 2012). Other factors such as temperature and presence of ions are also known to impact ejection rates. An increased concentration of Mg^{2+} ions reduces the repulsion and consequently increases the DNA ejection time (Wu *et al.*, 2010), whereas an increase in temperature means there is more energy to both trigger and maintain DNA ejection (Raspaud *et al.*, 2007) and will decrease the overall ejection time.

The total genome ejection time cannot be inferred from figures 5.3.2.1 and 5.3.3.1 because the end point of ejection cannot be established beyond the phage and cellular fluorescence signals becoming indistinguishable. It is possible for ejection rates of fluorescently stained phages genomes to be measured but it requires a more sophisticated experimental setup e.g. microfluidics in hydrodynamic flow (Mangenot *et al.*, 2005). A conclusion that can be made from these figures is that there is a 3 minute delay between the beginning of adsorption and the beginning of DNA ejection. As the phage being analysed bound within 10 s of being added to the solution, it is assumed that initial host

recognition occurred immediately. Linear trend lines assume constant DNA translocation speed – but these data imply that the rate is comparable to λ, in which total ejection time is around 2 mins at 30°C (García & Molineux, 1999).

All of the experiments were done at 20°C. As adsorption and phage genome ejection rates are known to be temperature-dependent (Raspaud *et al.*, 2007), cooling the microscope stage could slow down these processes and allow a higher temporal resolution of the events described here. This may also reduce the impact of some issues experienced such as evaporation of the medium and overgrowth of bacterial cells during the assay. However, stage cooling is not a feature of the Z1 Observer so different equipment would need to be used.

5.6.2 Premature cell death

The major drawback of the experimental set up was that, generally, immobilised bacterial cells died after approximately 20 minutes in this environment. There was no immediate answer to this and further experiments to answer any questions could not be conducted due to time constraints. Cells tumbling in solution appear to replicate as normal and often build up enough to obscure the view after 50 minutes to 60 minutes of the assay. Conversely, the replication of immobilised cells does not appear to proceed across the same timeframe and so a gentler immobilisation strategy may be required. Poly-L-lysine has been used to successfully attach *E. coli* cells to surface to allow imaging over longer time periods (Vadillo-Rodríguez *et al.*, 2004) despite having a bactericidal effect (Colville *et al.*, 2010). Immobilisation strategies that were not considered include

attachment using antibodies where cells remain viable for 6h (when stored in PBS at 4°C) (Suo *et al.*, 2009) and phages (Hosseinidoust *et al.*, 2011), since adding another phage aspect to the experiment may complicate conclusions.

The standard LIVE/DEAD® BacLight™ Bacterial Viability Kit (Life Technologies, UK) could not be used to determine cell viability during the assay as the SYTO® 9 green-fluorescent nucleic acid stain excites and emits at the same wavelengths as SYBR gold. The live/dead stain was initially used in the absence of a propidium iodide (PI) product to get an impression of the timelines for cells dying during the experiment but PI alone was used for images acquired after 25 May 2016.

Figure 5.4.1.1 shows that PI is able to leak into cells with compromised membranes of Bam0 in the absence of phages in solution. This means that the cell death observed is phage-independent. PLL has known antibiotic properties but not over the timescale of the experiment – and probably not such a synchronised effect on all cells. *E. coli* K-12 cells immobilised on PLL have survived for at least 4 hours in other experiments (Rozhok *et al.*, 2005; 2006) including in the experiments by Lonergan *et al.* (2014), which this method was based on. Conversely, PLL failed to immobilise *E. coli* O157 cells, which is thought to be due to the pili, flagella, and other appendages preventing the cell from making a full enough contact with the immobilisation surface (Suo *et al.*, 2008).

No conclusions were made the reason for the cell death seen in the assay. So questions remain:

- What conditions cause the "event" to occur, and can it be prevented?
- Is the cell death caused by something that is already known? For example, the cell death could be due to photo-sensitivity of bacterial cells (Nakahigashi *et al.*, 1991). This could be in isolation from or in combination with the effects of PLL (Colville *et al.*, 2010).

Chapter 6: General Discussion

6.1 Aims of this study

The general aim of this thesis was to continue the work stemming from evidence collected by Smith *et al.* (2007) that the BamA protein is the host recognition target for >70% short tailed Stx phages possessing a highly conserved tail spike protein. BamA is also highly conserved due to its essential function of folding nascent β-barrel membrane proteins in the outer membrane of all Gram-negative bacteria. The phage $\phi 24_B$ infects *E. coli*; however, it is capable of adsorbing to the BamA protein of other members of the *Enterobacteriaceae* (Smith *et al.*, 2007). Due to the high proportion of Shiga toxin phages that encode this particular tail, there is a greater chance of a mutation that could potentially facilitate a host range shift.

This thesis can be summarised briefly by the following accomplishments:

1. The adaptation of an existing method (Datsenko & Wanner, 2000) to create a novel strain of *E. coli*, Bam0, that contains a chimaeric BamA gene, which becomes resistant to $\phi 24_B$ adsorption.

2. Adsorption support was then complemented in Bam0 with WT *E. coli* BamA in an infection assay. Thus, the heretofore-missing proof that

BamA is the receptor molecule for $\phi 24_B$ has been presented in order to satisfy molecular Koch's postulates.

3. The complementation assay shown here can also be used to determine the ability of which BamA mutants support $\phi 24_B$ adsorption from the newly made library of 32 BamA mutants.

4. The development of a fluorescence microscopy assay that is able to monitor adsorption of a single $\phi 24_B$ particle to an adsorption susceptible cell and acquire temporally resolved data of a DNA ejection event.

Overall, the accomplishments largely match the aims of the study expressed in section 1.9 but they are currently limited by the trouble with complementing Bam0 with a chimaeric BamA mutant. So far only a WT *E. coli* BamA has been shown to restore support for phage adsorption.

For the fluorescence microscopy assay, the proof of principle has been shown to work but it still requires some adaptation or optimisation before it is able to reliably collect data for all 32 BamA mutants. Taking all of the data together, this thesis presents crucial groundwork for developing a greater understanding of $\phi 24_B$ adsorption to susceptible hosts.

6.2 Future work

The results of the infection assays and the fluorescence microscopy assays are able to corroborate each other, with the microscopy providing the ability to acquire further information regarding the timings of $\phi 24_B$ adsorption and DNA ejection. Solving the issues discussed in section 4.6 regarding the ability of chimaeric BamA mutants to be expressed and fold correctly *in vivo* would allow both of these assays to generate the data required to assess which *E. coli* epitopes are important for supporting $\phi 24_B$ adsorption.

Any constructs required for future work, especially if that construct contains TSP sequence, may be more time- and cost-effective to buy custom genes instead of relying on PCR products for cloning.

High- speed atomic force microscopy (AFM) could be employed using a functionalised cantilever tip with either whole $\phi 24_B$ particles or just a purified his-tagged TSP preparation. This would mean that tip would encounter resistance when the tail interacts with BamA. The technique could be used to map both the number and distribution of bioavailable BamA molecules that are present on the surface of a cell, and quantify the strength of the interaction by how much extra force is required to separate the cantilever tip from the cell surface again.

Alternatively, the use of a quartz microbalance would be capable of detecting small changes in mass of a sample attached to its resonator arm. The resonator

arm will vibrate at a given frequency until the mass of the sample changes. By immobilising $\phi 24_B$ particles to the resonator arm and adding *E. coli* cells to the solution, data could be collected to determine if the a BamA mutant expressed by those cell that is capable of supporting $\phi 24_B$ adsorption as the cells would bind to the cantilever via the phages and change the vibration pattern. This assay could be semi-quantitative with regard to the strength of the BamA-$\phi 24_B$ tail interaction. If the interaction is weak then the vibrations may be enough to "undo" the reaction before the adsorption process is complete.

The ultimate proof would be to infect a mouse model previously used in Stx studies (Mohawk & O'Brien, 2011) with those strains determined to resist or support $\phi 24_B$ adsorption and challenge those mice with a dose of $\phi 24_B$ in vivo. Monitoring these mice for the presence or lack of symptoms of a STEC infection would corroborate *in vitro* experiments; however, this may not be considered a necessary step by an ethical review panel.

The information collected from the assays mentioned (or similar assays) could be used to form a better picture of the dissemination of Stx genes in the environment. In order to protect the food supply chain against potential future contamination by new Stx phage host cells that we currently do not suspect. It may also improve responses to future outbreaks and even grant the potential to monitor water and food supply for potential sources of contamination.

Chapter 7: References

Abbasi, P., Kargar, M., Doosti, A., Mardaneh, J., Ghorbani-Dalini, S. & Dehyadegari, M. A. (2014). Characterization of Shiga-toxin producing *E.coli* (STEC) and enteropathogenic *E.coli* (EPEC) using multiplex Real-Time PCR assays for *stx1* , *stx2* , *eaeA*. *Iran J Microbiol* **6**, 169–174.

Acheson, D. W., Moore, R., De Breucker, S., Lincicome, L., Jacewicz, M., Skutelsky, E. & Keusch, G. T. (1996). Translocation of Shiga toxin across polarized intestinal cells in tissue culture. *Infect Immun* **64**, 3294–3300.

Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z. & Gross, C. A. (2002). DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. *Genes Dev* **16**, 2156–2168.

Albrecht, R., Schütz, M., Oberhettinger, P., Faulstich, M., Bermejo, I., Rudel, T., Diederichs, K. & Zeth, K. (2014). Structure of BamA, an essential factor in outer membrane protein biogenesis. *Acta Crystallogr D Biol Crystallogr* **70**, 1779–1789. International Union of Crystallography.

Aldabe, B., Delmas, Y., Gault, G., Vendrely, B., Llanas, B., Charron, M., Castor, C., Ong, N., Weill, F. *et al.* (2011). Household transmission of haemolytic uraemic syndrome associated with *Escherichia coli* O104:H4, south-western France, June 2011. *Euro Surveill* **16**.

Allison, H. E., Sergeant, M. J., James, C. E., Saunders, J. R., Smith, D. L., Sharp, R. J., Marks, T. S. & McCarthy, A. J. (2003). Immunity profiles of wild-type and recombinant Shiga-like toxin-encoding bacteriophages and characterization of novel double lysogens. *Infect Immun* **71**, 3409–3418. American Society for Microbiology.

Allison, H. E. (2007). Stx-phages: drivers and mediators of the evolution of STEC and STEC-like pathogens. *Future Microbiol* **2**, 165–174.

Aoki, S. K., Malinverni, J. C., Jacoby, K., Thomas, B., Pamma, R., Trinh, B. N., Remers, S., Webb, J., Braaten, B. A. *et al.* (2008). Contact-dependent growth inhibition requires the essential outer membrane protein BamA (YaeT) as the receptor and the inner membrane transport protein AcrB. *Mol Microbiol* **70**, 323–340.

Arenas-Hernández, M. M. P., Martínez-Laguna, Y. & Torres, A. G. (2012). Clinical implications of enteroadherent *Escherichia coli*. *Curr Gastroenterol Rep* **14**, 386–394.

Bai, J., Kim, Y.-T., Ryu, S. & Lee, J.-H. (2016). Biocontrol and rapid detection of food-borne pathogens using bacteriophages and endolysins. *Front Microbiol* **7**, 474.

Bakelar, J., Buchanan, S. K. & Noinaj, N. (2016). The structure of the β-barrel assembly machinery complex. *Science* **351**, 180–186.

Barabote, R. D., Johnson, O. L., Zetina, E., San Francisco, S. K., Fralick, J. A. & San Francisco, M. J. D. (2003). *Erwinia chrysanthemi* tolC is involved in resistance to antimicrobial plant chemicals and is essential for phytopathogenesis. *J Bacteriol* **185**, 5772–5778.

Bast, D. J., Banerjee, L., Clark, C., Read, R. J. & Brunton, J. L. (1999). The identification of three biologically relevant globotriaosyl ceramide receptor binding sites on the verotoxin 1 B subunit. *Mol Microbiol* **32**, 953–960.

Batchelor, M., Prasannan, S., Daniell, S., Reece, S., Connerton, I., Bloomberg, G., Dougan, G., Frankel, G. & Matthews, S. (2000). Structural basis for recognition of the translocated intimin receptor (Tir) by intimin from enteropathogenic *Escherichia coli*. *EMBO J* **19**, 2452–2464.

Beddoe, T., Paton, A. W., Le Nours, J., Rossjohn, J. & Paton, J. C. (2010). Structure, biological functions and applications of the AB5 toxins. *Trends Biochem Sci* **35**, 411–418.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences.

Nucleic Acids Res **27**, 573–580.

Bergal, H. T., Hopkins, A. H., Metzner, S. I. & Sousa, M. C. (2016). The Structure of a BamA-BamD Fusion Illuminates the Architecture of the β-Barrel Assembly Machine Core. *Structure* **24**, 243–251.

Bergan, J., Dyve Lingelem, A. B., Simm, R., Skotland, T. & Sandvig, K. (2012). Shiga toxins. *Toxicon* **60**, 1085–1107.

Beutin, L. & Martin, A. (2012). Outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O104:H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC strains. *J Food Prot* **75**, 408–418.

Bielaszewska, M., Friedrich, A. W., Aldick, T., Schürk-Bulgrin, R. & Karch, H. (2006). Shiga toxin activatable by intestinal mucus in *Escherichia coli* isolated from humans: predictor for a severe clinical outcome. *Clin Infect Dis* **43**, 1160–1167.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G. & Karch, H. (2011). Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect Dis* **11**, 671–676.

Bigelow, H. R., Petrey, D. S., Liu, J., Przybylski, D. & Rost, B. (2004). Predicting transmembrane beta-barrels in proteomes. *Nucleic Acids Res* **32**, 2566–2577.

Black, R. E., Merson, M. H., Huq, I., Alim, A. R. & Yunus, M. (1981). Incidence and severity of rotavirus and *Escherichia coli* diarrhoea in rural Bangladesh. Implications for vaccine development. *Lancet* **1**, 141–143.

Blaser, M. J. (2011). Deconstructing a lethal foodborne epidemic. *N Engl J Med* **365**, 1835–1836.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., ColladoVides, J., Glasner, J. D., Rode, C. K. & *et al*. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–&.

Blount, Z. D. (2015). The unexhausted potential of *E. coli. eLife* **4**, 96.

Boerlin, P., McEwen, S. A., Boerlin-Petzold, F., Wilson, J. B., Johnson, R. P. & Gyles, C. L. (1999). Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. *J Clin Microbiol* **37**, 497–503.

Bondy-Denomy, J. & Davidson, A. R. (2014). When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *Journal of Microbiology* **52**, 235–242.

Bos, M. P. & Tommassen, J. (2005). Viability of a capsule- and lipopolysaccharide-deficient mutant of *Neisseria meningitidis*. *Infect Immun* **73**, 6194–6197.

Bose, J. L. (2016). Chemical and UV mutagenesis. *Methods Mol Biol* **1373**, 111–115.

Boyce, T. G., Pemberton, A. G., Wells, J. G. & Griffin, P. M. (1995). Screening for *Escherichia coli* O157:H7--a nationwide survey of clinical laboratories. *J Clin Microbiol* **33**, 3275–3277.

Bray, J. & Beavan, T. E. D. (1948). Slide agglutination of *Bacterium coli* var.neapolitanum in summer diarrhæa. *The Journal of Pathology and Bacteriology* **60**, 395–401.

Brigotti, M., Caprioli, A., Tozzi, A. E., Tazzari, P. L., Ricci, F., Conte, R., Carnicelli, D., Procaccino, M. A., Minelli, F. *et al.* (2006). Shiga toxins present in the gut and in the polymorphonuclear leukocytes circulating in the blood of children with hemolytic-uremic syndrome. *J Clin Microbiol* **44**, 313–317.

Browning, D. F., Bavro, V. N., Mason, J. L., Sevastsyanovich, Y. R., Rossiter, A. E., Jeeves, M., Wells, T. J., Knowles, T. J., Cunningham, A. F. *et al.* (2015). Cross-species chimeras reveal BamA POTRA and β-barrel domains must be fine-tuned for efficient OMP insertion. *Mol Microbiol* **97**, 646–659.

Bryksin, A. V. & Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *BioTechniques* **48**, 463–465.

Buchholz, U., Bernard, H., Werber, D., Böhmer, M. M., Remschmidt, C., Wilking, H., Deleré, Y., an der Heiden, M., Adlhoch, C. *et al.* (2011). German outbreak of

Escherichia coli O104:H4 associated with sprouts. *N Engl J Med* **365**, 1763–1770.
Bunger, J. C., Melton-Celsa, A. R. & O'Brien, A. D. (2013). Shiga toxin type 2dact displays increased binding to globotriaosylceramide *in vitro* and increased lethality in mice after activation by elastase. *Toxins (Basel)* **5**, 2074–2092.
Bürk, C., Dietrich, R., Açar, G., Moravek, M., Bülte, M. & Märtlbauer, E. (2003). Identification and characterization of a new variant of Shiga toxin 1 in *Escherichia coli* ONT:H19 of bovine origin. *J Clin Microbiol* **41**, 2106–2112.
Calef, E., Marchelli, C. & Guerrini, F. (1965). The formation of superinfection-double lysogens of phage λ in *Escherichia coli* K12. *Virology* **27**, 1–10.
Callaway, T. R., TS, E., GH, L., MA, C. & DJ, N. (2013). Shiga toxin-producing *Escherichia coli* (STEC) ecology in cattle and management based options for reducing fecal shedding. *Agric Food Anal Bacteriol* **3**, 39–69.
Campbell, A. (1994). Comparative molecular biology of lambdoid phages. *Annu Rev Microbiol* **48**, 193–222.
Carver, T., Thomson, N., Bleasby, A., Berriman, M. & Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* **25**, 119–120.
Casadaban, M. J. & Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol* **138**, 179–207.
Celli, J., Olivier, M. & Finlay, B. B. (2001). Enteropathogenic *Escherichia coli* mediates antiphagocytosis through the inhibition of PI 3-kinase-dependent pathways. *EMBO J* **20**, 1245–1258.
Chan, Y. S. & Ng, T. B. (2016). Shiga toxins: from structure and mechanism to applications. *Appl Microbiol Biotechnol* **100**, 1597–1610.
Chapon, C. (1982). Expression of malT, the regulator gene of the maltose region in *Escherichia coli*, is limited both at transcription and translation. *EMBO J* **1**, 369–374.
Chaturongakul, S. & Ounjai, P. (2014). Phage-host interplay: examples from tailed phages and Gram-negative bacterial pathogens. *Front Microbiol* **5**, 442.
Chaudhuri, R. R., Sebaihia, M., Hobman, J. L., Webber, M. A., Leyton, D. L., Goldberg, M. D., Cunningham, A. F., Scott-Tucker, A., Ferguson, P. R. *et al.*. (2010). Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS ONE* **5**, e8801.
Chekabab, S. M., Daigle, F., Charette, S. J., Dozois, C. M. & Harel, J. (2013). Shiga toxins decrease enterohaemorrhagic *Escherichia coli* survival within *Acanthamoeba castellanii*. *FEMS Microbiol Lett* **344**, 86–93.
Chen, Z., Zhan, L. H., Hou, H.-F., Gao, Z.-Q., Xu, J.-H., Dong, C. & Dong, Y.-H. (2016). Structural basis for the interaction of BamB with the POTRA3-4 domains of BamA. *Acta Crystallogr D Struct Biol* **72**, 236–244.
Chui, L., Li, V., Fach, P., Delannoy, S., Malejczyk, K., Patterson-Fortin, L., Poon, A., King, R., Simmonds, K. *et al.* (2015). Molecular profiling of *Escherichia coli* O157:H7 and non-O157 strains isolated from humans and cattle in Alberta, Canada. *J Clin Microbiol* **53**, 986–990.
Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D. (2017). Bacteriophages Contribute to the Spread of Antibiotic Resistance Genes among Foodborne Pathogens of the *Enterobacteriaceae* Family - A Review. *Front Microbiol* **8**, 1108.
Colville, K., Tompkins, N., Rutenberg, A. D. & Jericho, M. H. (2010). Effects of poly(L-lysine) substrates on attached *Escherichia coli* bacteria. *Langmuir* **26**, 2639–2644.
Conway, T. & Cohen, P. S. (2015). Commensal and Pathogenic *Escherichia coli* Metabolism in the Gut. *Microbiol Spectr* **3**, 343–362.
Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38.
Croxen, M. A. & Finlay, B. B. (2009). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* **2**, 123–38.
Dartigalongue, C., Loferer, H. & Raina, S. (2001a). EcfE, a new essential inner membrane protease: its role in the regulation of heat shock response in *Escherichia*

coli. *EMBO J* **20**, 5908–5918.

Dartigalongue, C., Missiakas, D. & Raina, S. (2001b). Characterization of the *Escherichia coli* sigma E regulon. *J Biol Chem* **276**, 20866–20875.

Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640–6645.

DeGrandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C. A. (1989). Globotetraosylceramide is recognized by the pig edema disease toxin. *J Biol Chem* **264**, 12520–12525.

Delannoy, S., Beutin, L. & Fach, P. (2013). Discrimination of Enterohemorrhagic *Escherichia coli* (EHEC) from Non-EHEC Strains Based on Detection of Various Combinations of Type III Effector Genes. *J Clin Microbiol* **51**, 3257–3262.

Desin, T. S., Townsend, H. G. & Potter, A. A. (2015). Antibodies Directed against Shiga-Toxin Producing *Escherichia coli* Serotype O103 Type III Secreted Proteins Block Adherence of Heterologous STEC Serotypes to HEp-2 Cells. *PLoS ONE* **10**, e0139803.

Dickinson, B. L. & Clements, J. D. (1995). Dissociation of *Escherichia coli* heat-labile enterotoxin adjuvanticity from ADP-ribosyltransferase activity. *Infect Immun* **63**, 1617–1623.

Doerner, P. A. & Sousa, M. C. (2017). Extreme Dynamics in the BamA β-Barrel Seam. *Biochemistry* **56**, 3142–3149.

Doherty, G. J. & McMahon, H. T. (2009). Mechanisms of endocytosis. *Annu Rev Biochem* **78**, 857–902.

Donohue-Rolfe, A., Jacewicz, M. & Keusch, G. T. (1989). Isolation and characterization of functional Shiga toxin subunits and renatured holotoxin. *Mol Microbiol* **3**, 1231–1236.

DuPont, H. L., Formal, S. B., Hornick, R. B., Snyder, M. J., Libonati, J. P., Sheahan, D. G., LaBrec, E. H. & Kalas, J. P. (1971). Pathogenesis of *Escherichia coli* diarrhea. *N Engl J Med* **285**, 1–9.

Dwyer, R. S., Ricci, D. P., Colwell, L. J., Silhavy, T. J. & Wingreen, N. S. (2013). Predicting functionally informative mutations in *Escherichia coli* BamA using evolutionary covariance analysis. *Genetics* **195**, 443–455.

Eiklid, K. & Olsnes, S. (1983). Animal toxicity of *Shigella dysenteriae* cytotoxin: evidence that the neurotoxic, enterotoxic, and cytotoxic activities are due to one toxin. *J Immunol* **130**, 380–384.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T. & Igarashi, K. (1988). Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *Eur J Biochem* **171**, 45–50.

Erickson, H. P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biological Procedures Online* **11**, 32–51.

Ethelberg, S., Olsen, K. E. P., Scheutz, F., Jensen, C., Schiellerup, P., Enberg, J., Petersen, A. M., Olesen, B., Gerner-Smidt, P. & Mølbak, K. (2004). Virulence factors for hemolytic uremic syndrome, Denmark. *Emerging Infect Dis* **10**, 842–847.

Falguières, T., Mallard, F., Baron, C., Hanau, D., Lingwood, C., Goud, B., Salamero, J. & Johannes, L. (2001). Targeting of Shiga toxin B-subunit to retrograde transport route in association with detergent-resistant membranes. *Mol Biol Cell* **12**, 2453–2468.

Falkow, S. (1988). Molecular Koch's postulates applied to microbial pathogenicity. *Rev Infect Dis* **10 Suppl 2**, S274–6.

Flores-Mireles, A. L., Walker, J. N., Caparon, M. & Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol* **13**, 269–284.

Fogg, P. C. M., Rigden, D. J., Saunders, J. R., McCarthy, A. J. & Allison, H. E. (2011). Characterization of the relationship between integrase, excisionase and

antirepressor activities associated with a superinfecting Shiga toxin encoding bacteriophage. *Nucleic Acids Res* **39**, 2116–2129.

Fogg, P. C. M., Allison, H. E., Saunders, J. R. & McCarthy, A. J. (2010). Bacteriophage lambda: a paradigm revisited. *J Virol* **84**, 6876–6879.

Fogg, P. C. M., Gossage, S. M., Smith, D. L., Saunders, J. R., McCarthy, A. J. & Allison, H. E. (2007). Identification of multiple integration sites for Stx-phage Phi24B in the *Escherichia coli* genome, description of a novel integrase and evidence for a functional anti-repressor. *Microbiology (Reading, Engl)* **153**, 4098–4110.

Fogg, P. C. M., Saunders, J. R., McCarthy, A. J. & Allison, H. E. (2012). Cumulative effect of prophage burden on Shiga toxin production in *Escherichia coli*. *Microbiology (Reading, Engl)* **158**, 488–497.

Fontaine, A., Arondel, J. & Sansonetti, P. J. (1988). Role of Shiga toxin in the pathogenesis of bacillary dysentery, studied by using a Tox- mutant of *Shigella dysenteriae* 1. *Infect Immun* **56**, 3099–3109.

Frank, C., Werber, D., Cramer, J. P., Askar, M., Faber, M., an der Heiden, M., Bernard, H., Fruth, A., Prager, R. & *et al.* (2011). Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N Engl J Med* **365**, 1771–1780.

Frankel, G., Phillips, A. D., Rosenshine, I., Dougan, G., Kaper, J. B. & Knutton, S. (1998). Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol* **30**, 911–921.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 A resolution. *Nat Struct Biol* **1**, 59–64.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D. & James, M. N. G. (2004). Structure of Shiga toxin type 2 (Stx2) from *Escherichia coli* O157:H7. *J Biol Chem* **279**, 27511–27517.

Freedman, S. B., Xie, J., Neufeld, M. S., Hamilton, W. L., Hartling, L., Tarr, P. I., Alberta Provincial Pediatric Enteric Infection Team (APPETITE), Nettel-Aguirre, A., Chuck, A. & *et al.* (2016). Shiga Toxin-Producing *Escherichia coli* Infection, Antibiotics, and Risk of Developing Hemolytic Uremic Syndrome: A Meta-analysis. *Clin Infect Dis* **62**, 1251–1258.

Friedrich, A. W., Bielaszewska, M., Zhang, W.-L., Pulz, M., Kuczius, T., Ammon, A. & Karch, H. (2002). *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J Infect Dis* **185**, 74–84.

Friedrich, A. W., Borell, J., Bielaszewska, M., Fruth, A., Tschäpe, H. & Karch, H. (2003). Shiga toxin 1c-producing *Escherichia coli* strains: phenotypic and genetic characterization and association with human disease. *J Clin Microbiol* **41**, 2448–2453.

Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L. & *et al.* (2016). RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res* **44**, D133–43.

Gamage, S. D., Strasser, J. E., Chalk, C. L. & Weiss, A. A. (2003). Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect Immun* **71**, 3107–3115.

Gamage, S. D., Patton, A. K., Hanson, J. F. & Weiss, A. A. (2004). Diversity and host range of Shiga toxin-encoding phage. *Infect Immun* **72**, 7131–7139.

Gamage, S. D., Patton, A. K., Strasser, J. E., Chalk, C. L. & Weiss, A. A. (2006). Commensal bacteria influence *Escherichia coli* O157:H7 persistence and Shiga toxin production in the mouse intestine. *Infect Immun* **74**, 1977–1983.

García, L. R. & Molineux, I. J. (1999). Translocation and specific cleavage of bacteriophage T7 DNA in vivo by EcoKI. *Proceedings of the National Academy of Sciences* **96**, 12430–12435.

García-Aljaro, C., Muniesa, M., Jofre, J. & Blanch, A. R. (2004). Prevalence of the Stx2

gene in coliform populations from aquatic environments. *Appl Environ Microbiol* **70**, 3535–3540.

Garred, O., Dubinina, E., Polesskaya, A., Olsnes, S., Kozlov, J. & Sandvig, K. (1997). Role of the disulfide bond in Shiga toxin A-chain for toxin entry into cells. *J Biol Chem* **272**, 11414–11419.

Garred, O., van Deurs, B. & Sandvig, K. (1995). Furin-induced cleavage and activation of Shiga toxin. *J Biol Chem* **270**, 10817–10821.

Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. (2008). Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure* **16**, 1873–1881.

Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A. & Sousa, M. C. (2010). Structure and flexibility of the complete periplasmic domain of BamA: the protein insertion machine of the outer membrane. *Structure* **18**, 1492–1501.

Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Balázsi, G., Ravasz, E., Daugherty, M. D., Somera, A. L., Kyrpides, N. C., Anderson, I. & *et al.* (2003). Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol* **185**, 5673–5684.

Gessmann, D., Chung, Y. H., Danoff, E. J., Plummer, A. M., Sandlin, C. W., Zaccai, N. R. & Fleming, K. G. (2014). Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc Natl Acad Sci USA* **111**, 5878–5883.

Grande, L., Michelacci, V., Bondì, R., Gigliucci, F., Franz, E., Badouei, M. A., Schlager, S., Minelli, F., Tozzoli, R. & *et al.* (2016). Whole-Genome Characterization and Strain Comparison of VT2f-Producing Escherichia coli Causing Hemolytic Uremic Syndrome. *Emerging Infect Dis* **22**, 2078–2086.

Griffin, P. M., Ostroff, S. M., Tauxe, R. V., Greene, K. D., Wells, J. G., Lewis, J. H. & Blake, P. A. (1988). Illnesses associated with *Escherichia coli* O157:H7 infections. A broad clinical spectrum. *Ann Intern Med* **109**, 705–712.

Grotiuz, G., Sirok, A., Gadea, P., Varela, G. & Schelotto, F. (2006). Shiga toxin 2-producing *Acinetobacter haemolyticus* associated with a case of bloody diarrhea. *J Clin Microbiol* **44**, 3838–3841.

Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y. & *et al.* (2016). Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64–69.

Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121–4130.

Hagan, C. L. & Kahne, D. (2011). The reconstituted *Escherichia coli* Bam complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50**, 7444–7446.

Hagan, C. L., Westwood, D. B. & Kahne, D. (2013). Bam lipoproteins assemble BamA *in vitro*. *Biochemistry* **52**, 6108–6113. American Chemical Society.

Hagan, C. L., Wzorek, J. S. & Kahne, D. (2015). Inhibition of the β-barrel assembly machine by a peptide that binds BamD. *Proc Natl Acad Sci USA* **112**, 2011–2016.

Hale, T. L., Sansonetti, P. J., Schad, P. A., Austin, S. & Formal, S. B. (1983). Characterization of virulence plasmids and plasmid-associated outer membrane proteins in *Shigella flexneri, Shigella sonnei,* and *Escherichia coli*. *Infect Immun* **40**, 340–350.

Hamel, El, C., Chevalier, S., Dé, E., Orange, N. & Molle, G. (2001). Isolation and characterisation of the major outer membrane protein of *Erwinia carotovora*. *Biochim Biophys Acta* **1515**, 12–22.

Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D. & *et al.* (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat Struct Mol Biol* **23**, 192–196.

Hartland, E. L., Batchelor, M., Delahay, R. M., Hale, C., Matthews, S., Dougan, G.,

Knutton, S., Connerton, I. & Frankel, G. (1999). Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Mol Microbiol* **32**, 151–158.

Heinz, E. & Lithgow, T. (2014). A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front Microbiol* **5**, 370.

Hendrix, R. W. & Duda, R. L. (1992). Bacteriophage lambda PaPa: not the mother of all lambda phages. *Science* **258**, 1145–1148.

Hendrix, R. W., Hatfull, G. F. & Smith, M. C. M. (2003). Bacteriophages with tails: chasing their origins and evolution. *Res Microbiol* **154**, 253–257.

Heuck, A., Schleiffer, A. & Clausen, T. (2011). Augmenting β-augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J Mol Biol* **406**, 659–666.

Hodges, K. & Gill, R. (2010). Infectious diarrhea: Cellular and molecular mechanisms. *Gut Microbes* **1**, 4–21.

Holt, G. S., Lodge, J. K., McCarthy, A. J., Graham, A. K., Young, G., Bridge, S. H., Brown, A. K., Veses-Garcia, M., Lanyon, C. V. *et al.* (2017). Shigatoxin encoding Bacteriophage ϕ24B modulates bacterial metabolism to raise antimicrobial tolerance. *Sci Rep* **7**, 40424.

Hosseinidoust, Z., Van de Ven, T. G. M. & Tufenkji, N. (2011). Bacterial capture efficiency and antimicrobial activity of phage-functionalized model surfaces. *Langmuir* **27**, 5472–5480.

Hovde, C. J., Austin, P. R., Cloud, K. A., Williams, C. J. & Hunt, C. W. (1999). Effect of cattle diet on *Escherichia coli* O157:H7 acid resistance. *Appl Environ Microbiol* **65**, 3233–3235.

Hsieh, P.-F., Hsu, C.-R., Chen, C.-T., Lin, T.-L. & Wang, J.-T. (2016). The *Klebsiella pneumoniae* YfgL (BamB) lipoprotein contributes to outer membrane protein biogenesis, type-1 fimbriae expression, anti-phagocytosis, and *in vivo* virulence. *Virulence* **7**, 587–601.

Iadanza, M. G., Higgins, A. J., Schiffrin, B., Calabrese, A. N., Brockwell, D. J., Ashcroft, A. E., Radford, S. E. & Ranson, N. A. (2016). Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat Commun* **7**, 12865.

Ingmer, H., Miller, C. & Cohen, S. N. (2001). The RepA protein of plasmid pSC101 controls *Escherichia coli* cell division through the SOS response. *Mol Microbiol* **42**, 519–526.

Ito, H., Terai, A., Kurazono, H., Takeda, Y. & Nishibuchi, M. (1990). Cloning and nucleotide sequencing of Vero toxin 2 variant genes from *Escherichia coli* O91:H21 isolated from a patient with the hemolytic uremic syndrome. *Microb Pathog* **8**, 47–60.

Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. (1986). Pathogenesis of *Shigella* diarrhea. XI. Isolation of a *Shigella* toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J Exp Med* **163**, 1391–1404.

Jackson, M. P., Newland, J. W., Holmes, R. K. & O'Brien, A. D. (1987). Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from *Escherichia coli*. *Microb Pathog* **2**, 147–153.

Jain, S. & Goldberg, M. B. (2007). Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol* **189**, 5393–5398.

James, C. E., Stanley, K. N., Allison, H. E., Flint, H. J., Stewart, C. S., Sharp, R. J., Saunders, J. R. & McCarthy, A. J. (2001). Lytic and lysogenic infection of diverse *Escherichia coli* and *Shigella* strains with a verocytotoxigenic bacteriophage. *Appl Environ Microbiol* **67**, 4335–4337.

James, C. E. (2002). *Use of a recombinant verocytotoxigenic bacteriophage to investigate infection of bacterial hosts*. PhD Thesis. University of Liverpool.

Jansen, K. B., Baker, S. L. & Sousa, M. C. (2012). Crystal structure of BamB from Pseudomonas aeruginosa and functional evaluation of its conserved structural features. *PLoS ONE* **7**, e49749

Jansen, K. B., Baker, S. L. & Sousa, M. C. (2015). Crystal structure of BamB bound to a periplasmic domain fragment of BamA, the central component of the β-barrel assembly machine. *J Biol Chem* **290**, 2126–2136.

Jenkins, C., Dallman, T. J., Launders, N., Willis, C., Byrne, L., Jorgensen, F., Eppinger, M., Adak, G. K., Aird, H. *et al.* (2015). Public Health Investigation of Two Outbreaks of Shiga Toxin-Producing *Escherichia coli* O157 Associated with Consumption of Watercress. *Appl Environ Microbiol* **81**, 3946–3952

Jobling, M. G. & Holmes, R. K. (2012). Type II heat-labile enterotoxins from 50 diverse *Escherichia coli* isolates belong almost exclusively to the LT-IIc family and may be prophage encoded. *PLoS ONE* **7**, e29898

Johannes, L. & Römer, W. (2010). Shiga toxins--from cell biology to biomedical applications. *Nat Rev Microbiol* **8**, 105–116.

Johansen, B. K., Wasteson, Y., Granum, P. E. & Brynestad, S. (2001). Mosaic structure of Shiga-toxin-2-encoding phages isolated from *Escherichia coli* O157:H7 indicates frequent gene exchange between lambdoid phage genomes. *Microbiology (Reading, Engl)* **147**, 1929–1936.

Kaniga, K., Delor, I. & Cornelis, G. R. (1991). A wide-host-range suicide vector for improving reverse genetics in gram-negative bacteria: inactivation of the blaA gene of *Yersinia enterocolitica. Gene* **109**, 137–141.

Kaper, J. B., Nataro, J. P. & Mobley, H. L. (2004). Pathogenic *Escherichia coli. Nat Rev Microbiol* **2**, 123–140.

Karmali, M. A., Steele, B. T., Petric, M. & Lim, C. (1983). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* **1**, 619–620.

Kavaliauskiene, S., Dyve Lingelem, A. B., Skotland, T. & Sandvig, K. (2017). Protection against Shiga Toxins. *Toxins (Basel)* **9**, 44.

Kelly, J. K., Pai, C. H., Jadusingh, I. H., Macinnis, M. L., Shaffer, E. A. & Hershfield, N. B. (1987). The histopathology of rectosigmoid biopsies from adults with bloody diarrhea due to verotoxin-producing *Escherichia coli*. *Am J Clin Pathol* **88**, 78–82.

Keusch, G. T., Donohue-Rolfe, A. & Jacewicz, M. (1981). Shigella toxin(s) : Description and role in diarrhea and dysentery. *Pharmacology & Therapeutics* **15**, 403–438.

Khalil, R. K. S., Skinner, C., Patfield, S. & He, X. (2016). Phage-mediated Shiga toxin (Stx) horizontal gene transfer and expression in non-Shiga toxigenic *Enterobacter* and *Escherichia coli* strains. *Pathog Dis* **74**.

Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961–964.

King, L. A., Nogareda, F., Weill, F.-X., Mariani-Kurkdjian, P., Loukiadis, E., Gault, G., Jourdan-DaSilva, N., Bingen, E., Macé, M. *et al.* (2012). Outbreak of Shiga toxin-producing *Escherichia coli* O104:H4 associated with organic fenugreek sprouts, France, June 2011. *Clin Infect Dis* **54**, 1588–1594.

Kleanthous, C., Rassam, P. & Baumann, C. G. (2015). Protein-protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins. *Curr Opin Struct Biol* **35**, 109–115.

Klein, J. S., Jiang, S., Galimidi, R. P., Keeffe, J. R. & Bjorkman, P. J. (2014). Design and characterization of structured protein linkers with differing flexibilities. *Protein Eng Des Sel* **27**, 325–330.

Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. & Henderson, I. R. (2008). Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol Microbiol* **68**, 1216–1227.

Kokai-Kun, J. F., Melton-Celsa, A. R. & O'Brien, A. D. (2000). Elastase in intestinal mucus enhances the cytotoxicity of Shiga toxin type 2d. *J Biol Chem* **275**, 3713–3721.

Konovalova, A., Mitchell, A. M. & Silhavy, T. J. (2016). A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. *eLife* **5**, 5312.

Konovalova, A., Perlman, D. H., Cowles, C. E. & Silhavy, T. J. (2014). Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins. *Proc Natl Acad Sci USA* **111**, E4350–8.

Konowalchuk, J., Speirs, J. I. & Stavric, S. (1977). Vero response to a cytotoxin of *Escherichia coli*. *Infect Immun* **18**, 775–779.

Koskella, B. & Meaden, S. (2013). Understanding bacteriophage specificity in natural microbial communities. *Viruses* **5**, 806–823.

Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., Wu, Y., Sow, S. O., Sur, D. *et al.* (2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* **382**, 209–222.

Kumar, A., Taneja, N., Kumar, Y. & Sharma, M. (2012). Detection of Shiga toxin variants among Shiga toxin-forming *Escherichia coli* isolates from animal stool, meat and human stool samples in India. *J Appl Microbiol* **113**, 1208–1216.

Kunisaki, H. & Tanji, Y. (2010). Intercrossing of phage genomes in a phage cocktail and stable coexistence with *Escherichia coli* O157:H7 in anaerobic continuous culture. *Appl Microbiol Biotechnol* **85**, 1533–1540.

LaPointe, P., Wei, X. & Gariépy, J. (2005). A role for the protease-sensitive loop region of Shiga-like toxin 1 in the retrotranslocation of its A1 domain from the endoplasmic reticulum lumen. *J Biol Chem* **280**, 23310–23318.

Lauvrak, S. U., Torgersen, M. L. & Sandvig, K. (2004). Efficient endosome-to-Golgi transport of Shiga toxin is dependent on dynamin and clathrin. *J Cell Sci* **117**, 2321–2331.

Lee, J., Sutterlin, H. A., Wzorek, J. S., Mandler, M. D., Hagan, C. L., Grabowicz, M., Tomasek, D., May, M. D., Hart, E. M. *et al.* (2018). Substrate binding to BamD triggers a conformational change in BamA to control membrane insertion. *Proc Natl Acad Sci USA* **427**, 201711727.

Lee, K.-M., Lee, K., Go, J., Park, I. H., Shin, J.-S., Choi, J. Y., Kim, H. J. & Yoon, S. S. (2017). A genetic screen reveals novel targets to render *Pseudomonas aeruginosa* sensitive to lysozyme and cell wall-targeting antibiotics. *Frontiers in Cellular and Infection Microbiology* **7**, 372.

Leonard-Rivera, M. & Misra, R. (2012). Conserved residues of the putative L6 loop of *Escherichia coli* BamA play a critical role in the assembly of β-barrel outer membrane proteins, including that of BamA itself. *J Bacteriol* **194**, 4662–4668.

Leung, P. H. M., Peiris, J. S. M., Ng, W. W. S., Robins-Browne, R. M., Bettelheim, K. A. & Yam, W. C. (2003). A newly discovered verotoxin variant, VT2g, produced by bovine verocytotoxigenic *Escherichia coli*. *Appl Environ Microbiol* **69**, 7549–7553.

Levine, M. M. (1987). *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J Infect Dis* **155**, 377–389.

Lindgren, S. W., Samuel, J. E., Schmitt, C. K. & O'Brien, A. D. (1994). The specific activities of Shiga-like toxin type II (SLT-II) and SLT-II-related toxins of enterohemorrhagic *Escherichia coli* differ when measured by Vero cell cytotoxicity but not by mouse lethality. *Infect Immun* **62**, 623–631.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37**, 1777–1788.

Lior, H. (1996). *Escherischia coli in domestic animals and humans* (C. L. Gyles, Ed.). CAB International, Wallington, United Kingdom.

Lonergan, N. E., Britt, L. D. & Sullivan, C. J. (2014). Immobilizing live *Escherichia coli* for AFM studies of surface dynamics. *Ultramicroscopy* **137**, 30–39.

Löf, D., Schillén, K., Jönsson, B. & Evilevitch, A. (2007). Forces controlling the rate of DNA ejection from phage lambda. *J Mol Biol* **368**, 55–65.

Ly, T. M. & Müller, H. E. (1990). Ingested *Listeria monocytogenes* survive and multiply in protozoa. *J Med Microbiol* **33**, 51–54.

Majowicz, S. E., Scallan, E., Jones-Bitton, A., Sargeant, J. M., Stapleton, J., Angulo, F. J., Yeung, D. H. & Kirk, M. D. (2014). Global incidence of human Shiga toxin-producing *Escherichia coli* infections and deaths: a systematic review and knowledge synthesis. *Foodborne Pathog Dis* **11**, 447–455.

Malinverni, J. C. & Silhavy, T. J. (2011). Assembly of Outer Membrane β-Barrel Proteins: the Bam Complex. *EcoSal Plus* **4**.

Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* **61**, 151–164.

Mangenot, S., Hochrein, M., Rädler, J. & Letellier, L. (2005). Real-time imaging of DNA ejection from single phage particles. *Curr Biol* **15**, 430–435.

Mathewson, J. J., Johnson, P. C., DuPont, H. L., Morgan, D. R., Thornton, S. A., Wood, L. V. & Ericsson, C. D. (1985). A newly recognized cause of travelers' diarrhea: enteroadherent *Escherichia coli*. *J Infect Dis* **151**, 471–475.

Mathewson, J. J., Johnson, P. C., DuPont, H. L., Satterwhite, T. K. & Winsor, D. K. (1986). Pathogenicity of enteroadherent *Escherichia coli* in adult volunteers. *J Infect Dis* **154**, 524–527.

McCabe, A. L., Ricci, D., Adetunji, M. & Silhavy, T. J. (2017). Conformational changes that coordinate the activity of BamA and BamD allowing β-barrel assembly. *J Bacteriol* JB.00373–17.

McCullum, E. O., Williams, B. A. R., Zhang, J. & Chaput, J. C. (2010). Random mutagenesis by error-prone PCR. *Methods Mol Biol* **634**, 103–109.

McEwen, S. (2013). *Characterisation of Stx-phage adsorption: The essential outer-membrane protein, BamA.* Masters Thesis. Univeristy of Liverpool.

McGannon, C. M., Fuller, C. A. & Weiss, A. A. (2010). Different classes of antibiotics differentially influence shiga toxin production. *Antimicrob Agents Chemother* **54**, 3790–3798.

McKenzie, J., Johannes, L., Taguchi, T. & Sheff, D. (2009). Passage through the Golgi is necessary for Shiga toxin B subunit to reach the endoplasmic reticulum. *FEBS J* **276**, 1581–1595. Blackwell Publishing Ltd.

Mellmann, A., Harmsen, D., Cummings, C. A., Zentz, E. B., Leopold, S. R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y. *et al.* (2011). Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS ONE* **6**, e22751.

Melton-Celsa, A. R., Darnell, S. C. & O'Brien, A. D. (1996). Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic *Escherichia coli* O91:H21 isolates in orally infected, streptomycin-treated mice. *Infect Immun* **64**, 1569–1576.

Melton-Celsa, A. R. (2014). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* **2**, EHEC–0024–2013.

Melton-Celsa, A. R., Kokai-Kun, J. F. & O'Brien, A. D. (2002). Activation of Shiga toxin type 2d (Stx2d) by elastase involves cleavage of the C-terminal two amino acids of the A2 peptide in the context of the appropriate B pentamer. *Mol Microbiol* **43**, 207–215.

Mir, R. A., Weppelmann, T. A., Elzo, M., Ahn, S., Driver, J. D. & Jeong, K. C. (2016). Colonization of Beef Cattle by Shiga Toxin-Producing *Escherichia coli* during the

First Year of Life: A Cohort Study. *PLoS ONE* **11**, e0148518 (Y.-F. Chang, Ed.). Public Library of Science.

Miyada, C. G., Stoltzfus, L. & Wilcox, G. (1984). Regulation of the araC gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on araBAD expression. *Proceedings of the National Academy of Sciences* **81**, 4120–4124.

Mizutani, S., Nakazono, N. & Sugino, Y. (1999). The so-called chromosomal verotoxin genes are actually carried by defective prophages. *DNA Res* **6**, 141–143.

Modi, S. R., Lee, H. H., Spina, C. S. & Collins, J. J. (2013). Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **499**, 219–222.

Mohawk, K. L. & O'Brien, A. D. (2011). Mouse models of *Escherichia coli* O157:H7 infection and shiga toxin injection. *J Biomed Biotechnol* **2011**, 258185–17.

Moon, H. W., Whipp, S. C., Argenzio, R. A., Levine, M. M. & Giannella, R. A. (1983). Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect Immun* **41**, 1340–1351.

Moreau, D., Kumar, P., Wang, S. C., Chaumet, A., Chew, S. Y., Chevalley, H. & Bard, F. (2011). Genome-wide RNAi screens identify genes required for Ricin and PE intoxications. *Dev Cell* **21**, 231–244.

Mosier-Boss, P. A., Lieberman, S. H., Andrews, J. M., Rohwer, F. L., Wegley, L. E. & Breitbart, M. (2003). Use of fluorescently labeled phage in the detection and identification of bacterial species. *Appl Spectrosc* **57**, 1138–1144.

Mulvey, G. L., Marcato, P., Kitov, P. I., Sadowska, J., Bundle, D. R. & Armstrong, G. D. (2003). Assessment in mice of the therapeutic potential of tailored, multivalent Shiga toxin carbohydrate ligands. *J Infect Dis* **187**, 640–649.

Muniesa, M., Serra-Moreno, R. & Jofre, J. (2004a). Free Shiga toxin bacteriophages isolated from sewage showed diversity although the stx genes appeared conserved. *Environ Microbiol* **6**, 716–725.

Muniesa, M., Blanco, J. E., De Simón, M., Serra-Moreno, R., Blanch, A. R. & Jofre, J. (2004b). Diversity of stx2 converting bacteriophages induced from Shiga-toxin-producing *Escherichia coli* strains isolated from cattle. *Microbiology (Reading, Engl)* **150**, 2959–2971.

Muniesa, M. & Jofre, J. (2004). Factors influencing the replication of somatic coliphages in the water environment. *Antonie Van Leeuwenhoek* **86**, 65–76.

Nakagawa, S., Kojio, S., Taneike, I., Iwakura, N., Tamura, Y., Kushiya, K., Gondaira, F. & Yamamoto, T. (2003). Inhibitory action of telithromycin against Shiga toxin and endotoxin. *Biochem Biophys Res Commun* **310**, 1194–1199.

Nakahigashi, K., Nishimura, K., Miyamoto, K. & Inokuchi, H. (1991). Photosensitivity of a protoporphyrin-accumulating, light-sensitive mutant (visA) of *Escherichia coli* K-12. *Proceedings of the National Academy of Sciences* **88**, 10520–10524.

Nataro, J. P. & Kaper, J. B. (1998). Diarrheagenic *Escherichia coli. Clin Microbiol Rev* **11**, 142–201.

Nataro, J. P., Kaper, J. B., Robins-Browne, R., Prado, V., Vial, P. & Levine, M. M. (1987). Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr Infect Dis J* **6**, 829–831.

Ni, D., Wang, Y., Yang, X., Zhou, H., Hou, X., Cao, B., Lu, Z., Zhao, X., Yang, K. & Huang, Y. (2014). Structural and functional analysis of the β-barrel domain of BamA from *Escherichia coli. FASEB J* **28**, 2677–2685.

Nishikawa, K., Matsuoka, K., Kita, E., Okabe, N., Mizuguchi, M., Hino, K., Miyazawa, S., Yamasaki, C., Aoki, J. *et al*.. (2002). A therapeutic agent with oriented carbohydrates for treatment of infections by Shiga toxin-producing *Escherichia coli* O157:H7. *Proceedings of the National Academy of Sciences* **99**, 7669–7674.

Nishikawa, K., Matsuoka, K., Watanabe, M., Igai, K., Hino, K., Hatano, K., Yamada, A., Abe, N., Terunuma, D. *et al*. (2005). Identification of the optimal structure required for a Shiga toxin neutralizer with oriented carbohydrates to function in the

circulation. *J Infect Dis* **191**, 2097–2105.

Noinaj, N., Fairman, J. W. & Buchanan, S. K. (2011). The crystal structure of BamB suggests interactions with BamA and its role within the BAM complex. *J Mol Biol* **407**, 248–260.

Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T. & Buchanan, S. K. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–390.

O'Brien, A. D. & LaVeck, G. D. (1983). Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli. Infect Immun* **40**, 675–683.

O'Brien, A. D., LaVeck, G. D., Griffin, D. E. & Thompson, M. R. (1980). Characterization of *Shigella dysenteriae* 1 (Shiga) toxin purified by anti-Shiga toxin affinity chromatography. *Infect Immun* **30**, 170–179.

O'Brien, A. D., LaVeck, G. D., Thompson, M. R. & Formal, S. B. (1982). Production of *Shigella dysenteriae* type 1-like cytotoxin by *Escherichia coli*. *J Infect Dis* **146**, 763–769.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694–696.

Okuda, T., Tokuda, N., Numata, S.-I., Ito, M., Ohta, M., Kawamura, K., Wiels, J., Urano, T., Tajima, O. *et al.* (2006). Targeted disruption of Gb3/CD77 synthase gene resulted in the complete deletion of globo-series glycosphingolipids and loss of sensitivity to verotoxins. *J Biol Chem* **281**, 10230–10235.

Orskov, F. & Orskov, I. (1992). *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol* **38**, 699–704.

Orth, D., Grif, K., Khan, A. B., Naim, A., Dierich, M. P. & Würzner, R. (2007). The Shiga toxin genotype rather than the amount of Shiga toxin or the cytotoxicity of Shiga toxin in vitro correlates with the appearance of the hemolytic uremic syndrome. *Diagn Microbiol Infect Dis* **59**, 235–242.

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. (2007). Development of the human infant intestinal microbiota. *PLoS Biol* **5**, e177.

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D. & *et al.* (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278.

Paton, A. W., Beutin, L. & Paton, J. C. (1995). Heterogeneity of the amino-acid sequences of *Escherichia coli* Shiga-like toxin type-I operons. *Gene* **153**, 71–74.

Paton, A. W. & Paton, J. C. (1996). *Enterobacter cloacae* producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome. *J Clin Microbiol* **34**, 463–465.

Persson, S., Olsen, K. E. P., Ethelberg, S. & Scheutz, F. (2007). Subtyping method for *Escherichia coli* shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. *J Clin Microbiol* **45**, 2020–2024.

PHE (2014) https://www.gov.uk/government/publications/foodborne-illness-in-humans-general-outbreaks-in-england-and-wales-in-2014 - accessed on 20 Apr 2018

PHE (2015) https://www.gov.uk/government/publications/foodborne-illness-in-humans-general-outbreaks-in-england-and-wales-in-2015 - accessed on 20 Apr 2018

PHE (2016) https://www.gov.uk/government/publications/foodborne-illness-in-humans-general-outbreaks-in-england-and-wales-in-2016 - accessed on 20 Apr 2018

PHE (2017) https://www.gov.uk/government/publications/foodborne-illness-in-humans-general-outbreaks-in-england-and-wales-in-2017 - accessed on 20 Apr 2018

PHE (2018) https://www.gov.uk/government/publications/foodborne-illness-in-

humans-general-outbreaks-in-england-and-wales-in-2018 - accessed on 20 Apr 2018

Pickard, D., Toribio, A. L., Petty, N. K., van Tonder, A., Yu, L., Goulding, D., Barrell, B., Rance, R., Harris, D. & *et al.* (2010). A conserved acetyl esterase domain targets diverse bacteriophages to the Vi capsular receptor of *Salmonella enterica* serovar Typhi. *J Bacteriol* **192**, 5746–5754.

Pierce, M. M., Raman, C. S. & Nall, B. T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213–221.

Polotsky, Y. E., Dragunskaya, E. M., Seliverstova, V. G., Avdeeva, T. A., Chakhutinskaya, M. G., Kétyi, I., Verténsyl, A., Ralovich, B., Emödy, L. *et al.* (1977). Pathogenic effect of enterotoxigenic *Escherichia coli* and *Escherichia coli* causing infantile diarrhoea. *Acta Microbiol Acad Sci Hung* **24**, 221–236.

Potter, A. A., Klashinsky, S., Li, Y., Frey, E., Townsend, H., Rogan, D., Erickson, G., Hinkley, S., Klopfenstein, T. *et al.* (2004). Decreased shedding of *Escherichia coli* O157:H7 by cattle following vaccination with type III secreted proteins. *Vaccine* **22**, 362–369.

Probert, W. S., McQuaid, C. & Schrader, K. (2014). Isolation and identification of an *Enterobacter cloacae* strain producing a novel subtype of Shiga toxin type 1. *J Clin Microbiol* **52**, 2346–2351.

Raife, T., Friedman, K. D. & Fenwick, B. (2004). Lepirudin prevents lethal effects of Shiga toxin in a canine model. *Thromb Haemost* **92**, 387–393. Schattauer Publishers.

Rajarathnam, K. & Rösgen, J. (2014). Isothermal titration calorimetry of membrane proteins - progress and challenges. *Biochim Biophys Acta* **1838**, 69–77.

Rakhuba, D. V., Kolomiets, E. I., Dey, E. S. & Novik, G. I. (2010). Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Pol J Microbiol* **59**, 145–155.

Ramotar, K., Waldhart, B., Church, D., Szumski, R. & Louie, T. J. (1995). Direct detection of verotoxin-producing *Escherichia coli* in stool samples by PCR. *J Clin Microbiol* **33**, 519–524.

Raspaud, E., Forth, T., São-José, C., Tavares, P. & de Frutos, M. (2007). A kinetic analysis of DNA ejection from tailed phages revealing the prerequisite activation energy. *Biophys J* **93**, 3999–4005.

Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R. *et al.* (2015). Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523**, 333–336.

Ren, W., Yu, R., Liu, G., Li, N., Peng, Y., Wu, M., Yin, Y., Li, Y., Fatufe, A. A. & Li, T. (2013). DNA vaccine encoding the major virulence factors of Shiga toxin type 2e (Stx2e)-expressing *Escherichia coli* induces protection in mice. *Vaccine* **31**, 367–372.

Rhodius, V. A., Suh, W. C., Nonaka, G., West, J. & Gross, C. A. (2006). Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol* **4**, e2.

Rigel, N. W., Ricci, D. P. & Silhavy, T. J. (2013). Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for β-barrel assembly in *Escherichia coli. Proc Natl Acad Sci USA* **110**, 5151–5156.

Rigel, N. W., Schwalm, J., Ricci, D. P. & Silhavy, T. J. (2012). BamE modulates the *Escherichia coli* beta-barrel assembly machine component BamA. *J Bacteriol* **194**, 1002–1008.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M. & *et al.* (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N Engl J Med* **308**, 681–685.

Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P. & Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* **4**, e377.

Robertson, K., Green, A., Allen, L., Ihry, T., White, P., Chen, W.-S., Douris, A. & Levine, J. (2016). Foodborne Outbreaks Reported to the U.S. Food Safety and

Inspection Service, Fiscal Years 2007 through 2012. *J Food Prot* **79**, 442–447.

Robinson, S. E., Wright, E. J., Hart, C. A., Bennett, M. & French, N. P. (2004). Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. *J Appl Microbiol* **97**, 1045–1053.

Roeder, D. L. & Collmer, A. (1985). Marker-exchange mutagenesis of a pectate lyase isozyme gene in *Erwinia chrysanthemi*. *J Bacteriol* **164**, 51–56.

Rowbotham, T. J. (1980). Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *J Clin Pathol* **33**, 1179–1183.

Rozhok, S., Fan, Z., Nyamjav, D., Liu, C., Mirkin, C. A. & Holz, R. C. (2006). Attachment of motile bacterial cells to prealigned holed microarrays. *Langmuir* **22**, 11251–11254.

Rozhok, S., Shen, C. K.-F., Littler, P.-L. H., Fan, Z., Liu, C., Mirkin, C. A. & Holz, R. C. (2005). Methods for fabricating microarrays of motile bacteria. *Small* **1**, 445–451.

Ruchaud-Sparagano, M.-H., Maresca, M. & Kenny, B. (2007). Enteropathogenic *Escherichia coli* (EPEC) inactivate innate immune responses prior to compromising epithelial barrier function. *Cell Microbiol* **9**, 1909–1921.

Ruhe, Z. C., Low, D. A. & Hayes, C. S. (2013a). Bacterial contact-dependent growth inhibition. *Trends Microbiol* **21**, 230–237.

Ruhe, Z. C., Wallace, A. B., Low, D. A. & Hayes, C. S. (2013b). Receptor polymorphism restricts contact-dependent growth inhibition to members of the same species. *MBio* **4**, e00480-13.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944–945.

Sack, D. A., Kaminsky, D. C., Sack, R. B., Itotia, J. N., Arthur, R. R., Kapikian, A. Z., Orskov, F. & Orskov, I. (1978). Prophylactic doxycycline for travelers' diarrhea. Results of a prospective double-blind study of Peace Corps volunteers in Kenya. *N Engl J Med* **298**, 758–763.

Salgado, H., Peralta-Gil, M., Gama-Castro, S., Santos-Zavaleta, A., Muñiz-Rascado, L., García-Sotelo, J. S., Weiss, V., Solano-Lira, H., Martínez-Flores, I. *et al.* (2013). RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res* **41**, D203–13.

Samuel, J. E. & Gordon, V. M. (1994). Evidence that proteolytic separation of Shiga-like toxin type IIv A subunit into A1 and A2 subunits is not required for toxin activity. *J Biol Chem* **269**, 4853–4859.

Samuel, J. E., Perera, L. P., Ward, S., O'Brien, A. D., Ginsburg, V. & Krivan, H. C. (1990). Comparison of the glycolipid receptor specificities of Shiga-like toxin type II and Shiga-like toxin type II variants. *Infect Immun* **58**, 611–618.

Sandvig, K., Garred, O., Prydz, K., Kozlov, J. V., Hansen, S. H. & van Deurs, B. (1992). Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum. *Nature* **358**, 510–512.

Sato, T., Matsui, T., Takita, E., Kadoyama, Y., Makino, S.-I., Kato, K., Sawada, K. & Hamabata, T. (2013). Evaluation of recombinant forms of the Shiga toxin variant Stx2eB subunit and non-toxic mutant Stx2e as vaccine candidates against porcine edema disease. *J Vet Med Sci* **75**, 1309–1315.

Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R. & *et al.* (2012). Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* **50**, 2951–2963.

Schiffrin, B., Calabrese, A. N., Higgins, A. J., Humes, J. R., Ashcroft, A. E., Kalli, A. C., Brockwell, D. J. & Radford, S. E. (2017). Effects of periplasmic chaperones and membrane thickness on BamA-catalysed outer membrane protein folding. *J Mol Biol*. **429**, 3776–3792.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,

Preibisch, S., Rueden, C., Saalfeld, S. *et al.* (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682.

Schmidt, H. (2001). Shiga-toxin-converting bacteriophages. *Res Microbiol* **152**, 687–695.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L. H. & Karch, H. (2000). A new Shiga toxin 2 variant (Stx2f) from *Escherichia coli* isolated from pigeons. *Appl Environ Microbiol* **66**, 1205–1208. American Society for Microbiology (ASM).

Schmitt, C. K., McKee, M. L. & O'Brien, A. D. (1991). Two copies of Shiga-like toxin II-related genes common in enterohemorrhagic *Escherichia coli* strains are responsible for the antigenic heterogeneity of the O157:H- strain E32511. *Infect Immun* **59**, 1065–1073.

Schneider, J. & Strack, H. B. (1977). Analysis of the phase variation in lambda reduced immunity lysogens. *Mol Gen Genet* **151**, 181–187.

Schroeder, G. N. & Hilbi, H. (2008). Molecular pathogenesis of *Shigella spp.*: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* **21**, 134–156.

Schüller, S., Heuschkel, R., Torrente, F., Kaper, J. B. & Phillips, A. D. (2007). Shiga toxin binding in normal and inflamed human intestinal mucosa. *Microbes Infect* **9**, 35–39.

Sergeant, M. J. (1998). *Molecular biological characterisation of verotoxigenic bacteriophage in E. coli*. PhD Thesis. University of Liverpool.

Shepard, S. M., Danzeisen, J. L., Isaacson, R. E., Seemann, T., Achtman, M. & Johnson, T. J. (2012). Genome sequences and phylogenetic analysis of K88- and F18-positive porcine enterotoxigenic *Escherichia coli*. *J Bacteriol* **194**, 395–405. American Society for Microbiology.

Shevchik, V. E., Condemine, G. & Robert-Baudouy, J. (1994). Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity. *EMBO J* **13**, 2007–2012.

Shiga, K. (1898). Ueber den Erreger der Dysenterie in Japan. *Zentralbl Bakteriol Mikrobiol Hyg, Vorläufige Mitteilung* **23**, 599–600.

Shiga, K. (1906). Observations on the epidemiology of dysentery in Japan. *Philippine Journal of Science* **1**, 485–500.

Shiga, K. (1936). The trend of prevention, therapy and epidemiology of dysentery since the discovery of its causative organism. *N Engl J Med* **215**, 1205–1211.

Short, C. M. & Suttle, C. A. (2005). Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Appl Environ Microbiol* **71**, 480–486.

Siegler, R. L., Obrig, T. G., Pysher, T. J., Tesh, V. L., Denkers, N. D. & Taylor, F. B. (2003). Response to Shiga toxin 1 and 2 in a baboon model of hemolytic uremic syndrome. *Pediatr Nephrol* **18**, 92–96.

Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli. Proc Natl Acad Sci USA* **104**, 6400–6405.

Smith, D. L. (2004). *Characterisation of the infection of* E.coli *by a Shiga-toxin encoding bacteriophage*. PhD Thesis. University of Liverpool.

Smith, D. L., James, C. E., Sergeant, M. J., Yaxian, Y., Saunders, J. R., McCarthy, A. J. & Allison, H. E. (2007). Short-tailed stx phages exploit the conserved YaeT protein to disseminate Shiga toxin genes among *Enterobacteria*. *J Bacteriol* **189**, 7223–7233.

Smith, D. L., Rooks, D. J., Fogg, P. C. M., Darby, A. C., Thomson, N. R., McCarthy, A. J. & Allison, H. E. (2012). Comparative genomics of Shiga toxin encoding bacteriophages. *BMC Genomics* **13**, 311.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503–517.

Steele, B. T., Murphy, N., Arbus, G. S. & Rance, C. P. (1982). An outbreak of hemolytic uremic syndrome associated with ingestion of fresh apple juice. *J Pediatr* **101**, 963–965.

Strauch, E., Lurz, R. & Beutin, L. (2001). Characterization of a Shiga toxin-encoding temperate bacteriophage of *Shigella sonnei*. *Infect Immun* **69**, 7588–7595.

Strockbine, N. A., Jackson, M. P., Sung, L. M., Holmes, R. K. & O'Brien, A. D. (1988). Cloning and sequencing of the genes for Shiga toxin from *Shigella dysenteriae* type 1. *J Bacteriol* **170**, 1116–1122.

Strockbine, N. A., Marques, L. R., Newland, J. W., Smith, H. W., Holmes, R. K. & O'Brien, A. D. (1986). Two toxin-converting phages from *Escherichia coli* O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. *Infect Immun* **53**, 135–140.

Suo, Z., Avci, R., Yang, X. & Pascual, D. W. (2008). Efficient immobilization and patterning of live bacterial cells. *Langmuir* **24**, 4161–4167.

Suo, Z., Yang, X., Avci, R., Deliorman, M., Rugheimer, P., Pascual, D. W. & Idzerda, Y. (2009). Antibody selection for immobilizing living bacteria. *Anal Chem* **81**, 7571–7578.

Szych, J., Wołkowicz, T., La Ragione, R. & Madajczak, G. (2014). Impact of antibiotics on the intestinal microbiota and on the treatment of Shiga-toxin-producing *Escherichia coli* and *Salmonella* infections. *Curr Pharm Des* **20**, 4535–4548.

Takahashi, M., Taguchi, H., Yamaguchi, H., Osaki, T., Komatsu, A. & Kamiya, S. (2004). The effect of probiotic treatment with *Clostridium butyricum* on enterohemorrhagic *Escherichia coli* O157:H7 infection in mice. *FEMS Immunol Med Microbiol* **41**, 219–226.

Takeda, T. (1998). Strategy to Prevent the Progression of Enterohemorrhagic *Escherichia coli* O157 Infection to Hemolytic Uremic Syndrome. *Jpn J Med Sci Biol* **51**, S124–S128.

Tarr, P. I., Gordon, C. A. & Chandler, W. L. (2005). Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* **365**, 1073–1086.

Taxt, A., Aasland, R., Sommerfelt, H., Nataro, J. & Puntervoll, P. (2010). Heat-stable enterotoxin of enterotoxigenic *Escherichia coli* as a vaccine target. *Infect Immun* **78**, 1824–1831.

Tellez, R. & Misra, R. (2012). Substitutions in the BamA β-barrel domain overcome the conditional lethal phenotype of a ΔbamB ΔbamE strain of *Escherichia coli. J Bacteriol* **194**, 317–324.

Tenaillon, O., Skurnik, D., Picard, B. & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* **8**, 207–217.

Tesh, V. L. (2012). The induction of apoptosis by Shiga toxins and ricin. *Curr Top Microbiol Immunol* **357**, 137–178.

Thom, S., Warhurst, D. & Drasar, B. S. (1992). Association of *Vibrio cholerae* with fresh water amoebae. *J Med Microbiol* **36**, 303–306.

Trachtman, H., Austin, C., Lewinski, M. & Stahl, R. A. K. (2012). Renal and neurological involvement in typical Shiga toxin-associated HUS. *Nat Rev Nephrol* **8**, 658–669.

Tyler, J. S., Beeri, K., Reynolds, J. L., Alteri, C. J., Skinner, K. G., Friedman, J. H., Eaton, K. A. & Friedman, D. I. (2013). Prophage induction is enhanced and required for renal disease and lethality in an EHEC mouse model. *PLoS Pathog* **9**, e1003236.

Vadillo-Rodríguez, V., Busscher, H. J., Norde, W., De Vries, J., Dijkstra, R. J. B., Stokroos, I. & Van Der Mei, H. C. (2004). Comparison of atomic force microscopy interaction forces between bacteria and silicon nitride substrata for three commonly used immobilization methods. *Appl Environ Microbiol* **70**, 5441–5446.

van der Ley, P., de Graaff, P. & Tommassen, J. (1986). Shielding of *Escherichia coli* outer membrane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. *J Bacteriol* **168**, 449–451.

van der Woude, M. W. & Bäumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**, 581–611.

Veses-Garcia, M., Liu, X., Rigden, D. J., Kenny, J. G., McCarthy, A. J. & Allison, H. E. (2015). Transcriptomic Analysis of Shiga-Toxigenic Bacteriophage Carriage Reveals a Profound Regulatory Effect on Acid Resistance in *Escherichia coli. Appl Environ Microbiol* **81**, 8118–8125.

Vikram, A., Jesudhasan, P. R., Pillai, S. D. & Patil, B. S. (2012). Isolimonic acid interferes with *Escherichia coli* O157:H7 biofilm and TTSS in QseBC and QseA dependent fashion. *BMC Microbiol* **12**, 261.

Vimala, A., Ramakrishnan, C. & Gromiha, M. M. (2015). Identifying a potential receptor for the antibacterial peptide of sponge Axinella donnani endosymbiont. *Gene* **566**, 166–174.

Volokhina, E. B., Grijpstra, J., Beckers, F., Lindh, E., Robert, V., Tommassen, J. & Bos, M. P. (2013). Species-specificity of the BamA component of the bacterial outer membrane protein-assembly machinery. *PLoS ONE* **8**, e85799.

Vos-Scheperkeuter, G. H. & Witholt, B. (1984). Assembly pathway of newly synthesized LamB protein an outer membrane protein of *Escherichia coli* K-12. *J Mol Biol* **175**, 511–528.

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262–265.

Wagner, P. L., Neely, M. N., Zhang, X., Acheson, D. W., Waldor, M. K. & Friedman, D. I. (2001). Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *J Bacteriol* **183**, 2081–2085.

Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W. K., Friedman, D. I. & Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol Microbiol* **44**, 957–970.

Warnier, M., Römer, W., Geelen, J., Lesieur, J., Amessou, M., van den Heuvel, L., Monnens, L. & Johannes, L. (2006). Trafficking of Shiga toxin/Shiga-like toxin-1 in human glomerular microvascular endothelial cells and human mesangial cells. *Kidney Int* **70**, 2085–2091.

Watters, C., Fleming, D., Bishop, D. & Rumbaugh, K. P. (2016). Host Responses to Biofilm. *Prog Mol Biol Transl Sci* **142**, 193–239. Elsevier.

Webb, C. T., Selkrig, J., Perry, A. J., Noinaj, N., Buchanan, S. K. & Lithgow, T. (2012). Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J Mol Biol* **422**, 545–555.

Weinstein, D. L., Jackson, M. P., Samuel, J. E., Holmes, R. K. & O'Brien, A. D. (1988). Cloning and sequencing of a Shiga-like toxin type II variant from *Escherichia coli* strain responsible for edema disease of swine. *J Bacteriol* **170**, 4223–4230.

Weirich, J., Bräutigam, C., Mühlenkamp, M., Franz-Wachtel, M., Macek, B., Meuskens, I., Skurnik, M., Leskinen, K., Bohn, E. & *et al.* (2017). Identifying components required for OMP biogenesis as novel targets for antiinfective drugs. *Virulence* **10**, 1–20.

Werner, J. & Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli. Mol Microbiol* **57**, 1450–1459.

Werts, C., Michel, V., Hofnung, M. & Charbit, A. (1994). Adsorption of bacteriophage lambda on the LamB protein of *Escherichia coli* K-12: point mutations in gene J of lambda responsible for extended host range. *J Bacteriol* **176**, 941–947.

Whittam, T. S., Wolfe, M. L., Wachsmuth, I. K., Orskov, F., Orskov, I. & Wilson, R. A. (1993). Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect Immun* **61**, 1619–1629.

WHO (2011). Outbreaks of *E. coli* O104:H4 infection: update 30. World Health Organization.

Winiecka-Krusnell, J., Wreiber, K., Euler, von, A., Engstrand, L. & Linder, E. (2002). Free-living amoebae promote growth and survival of *Helicobacter pylori. Scand J Infect Dis* **34**, 253–256.

Wong, A. R. C., Pearson, J. S., Bright, M. D., Munera, D., Robinson, K. S., Lee, S. F., Frankel, G. & Hartland, E. L. (2011). Enteropathogenic and enterohaemorrhagic *Escherichia coli*: even more subversive elements. *Mol Microbiol* **80**, 1420–1438.

Wu, D., Van Valen, D., Hu, Q. & Phillips, R. (2010). Ion-dependent dynamics of DNA ejections for bacteriophage lambda. *Biophys J* **99**, 1101–1109.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235–245.

Wzorek, J. S., Lee, J., Tomasek, D., Hagan, C. L. & Kahne, D. E. (2017). Membrane integration of an essential β-barrel protein prerequires burial of an extracellular loop. *Proceedings of the National Academy of Sciences* **114**, 2598–2603.

Yang, S.-C., Lin, C.-H., Aljuffali, I. A. & Fang, J.-Y. (2017). Current pathogenic *Escherichia coli* foodborne outbreak cases and therapy development. *Archives of Microbiology* **199**, 811–825. Springer Berlin Heidelberg.

Yanisch-Perron, C., Viera, J. & Messing, J. (1985). Improved M13 Phage Cloning Vectors and Host Strains - Nucleotide-Sequences of the M13mp18 and pUC19 Vectors. *Gene* **33**, 103–119.

Yu, M. & Haslam, D. B. (2005). Shiga toxin is transported from the endoplasmic reticulum following interaction with the luminal chaperone HEDJ/ERdj3. *Infect Immun* **73**, 2524–2532.

Zeinhom, M., Tellez, A. M., Delcenserie, V., El-Kholy, A. M., El-Shinawy, S. H. & Griffiths, M. W. (2012). Yogurt containing bioactive molecules produced by *Lactobacillus acidophilus* La-5 exerts a protective effect against enterohemorrhagic *Escherichia coli* in mice. *J Food Prot* **75**, 1796–1805.

Zhang, H., Gao, Z.-Q., Hou, H.-F., Xu, J.-H., Li, L.-F., Su, X.-D. & Dong, Y.-H. (2011). High-resolution structure of a new crystal form of BamA POTRA4-5 from *Escherichia coli. Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**, 734–738.

Zhang, W., Bielaszewska, M., Kuczius, T. & Karch, H. (2002). Identification, characterization, and distribution of a Shiga toxin 1 gene variant (stx(1c)) in *Escherichia coli* strains isolated from humans. *J Clin Microbiol* **40**, 1441–1446.

Zhang, X., McDaniel, A. D., Wolf, L. E., Keusch, G. T., Waldor, M. K. & Acheson, D. W. (2000). Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J Infect Dis* **181**, 664–670.

Zhou, K., Ferdous, M., de Boer, R. F., Kooistra-Smid, A. M. D., Grundmann, H., Friedrich, A. W. & Rossen, J. W. A. (2015). The mosaic genome structure and phylogeny of Shiga toxin-producing *Escherichia coli* O104:H4 is driven by short-term adaptation. *Clin Microbiol Infect* **21**, 468.e7–18.

Zumbrun, S. D., Melton-Celsa, A. R., Smith, M. A., Gilbreath, J. J., Merrell, D. S. & O'Brien, A. D. (2013). Dietary choice affects Shiga toxin-producing *Escherichia coli* (STEC) O157:H7 colonization and disease. *Proc Natl Acad Sci USA* **110**, E2126–33.

Chapter 8 Appendices

8.1 Sequencing confirmation of constructs made in this study

Initally, sequencing data was saved, then alignment files created by comparing the sequencing data to reference sequences were saved. Files can be found on at https://pcwww.liv.ac.uk/~hallison/Stuart_Thesis/index.html in the section: "Appendix 8.1 Sequencing data".

8.2 Fluorescence microscopy data generating for this study

Videos and images used for making figures in Chapter 5 can be found at https://pcwww.liv.ac.uk/~hallison/Stuart_Thesis/index.html in the section "8.2 Thesis data". Files have been named after and sorted by the figure(s), in which they appear. The instructions followed for generating and analysing data from image analysis are also included.

All data collected for this study, usable or otherwise, has been included at https://pcwww.liv.ac.uk/~hallison/Stuart_Thesis/index.html in the section "8.2 All data".