Deciphering the Autotransporter Pathway of Gram-negative Pathogens; from Regulation to Secretion.

by

Amanda Eve Rossiter

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

College of Medical and Dental Sciences
School of Immunity and Infection
University of Birmingham
September 2011

ABSTRACT

Autotransporters represent a diverse family of virulence effectors that are secreted from Gram-negative bacteria by the Type V Secretion System. Their initial description coined the term 'Autotransporter' to embody the notion that their three-part architecture governs their navigation through the bacterial cell envelope. The Pet cytotoxic autotransporter is secreted by the diarrhoeal pathogen, Enteroaggregative *Escherichia coli* (EAEC) and was used as a model to study autotransporter biogenesis. Following a global transposon mutagenesis of EAEC, novel accessory factors were identified that are required for Pet biogenesis, including the transcription factors CRP and Fis, periplasmic chaperones and components of the β-barrel assembly machinery (BAM) complex. Using both *in vivo* and *in vitro* techniques, we show that the *pet* promoter is co-dependent on CRP and Fis. We present a novel co-activation mechanism whereby CRP is placed at a non-optimal position for transcription initiation, creating dependence on Fis for full activation and show that this co-activation mechanism extends to functionally similar autotransporters. Furthermore, we highlight novel components of the BAM complex required for AT secretion. This work builds on previous studies that, in recent years, have challenged the 'auto' nature of this secretion process causing a paradigm shift towards a much more complex mechanism of AT secretion than initially suggested.

Acknowledgements

First and foremost, I would like to thank my principal supervisor, Prof. Ian Henderson. I am truly indebted and thankful for all his continual support, patience and expert guidance throughout the entirety of my PhD. I am particularly grateful that he established my reputation as 'the most well travelled PhD student', giving me the invaluable opportunities to meet inspirational peers and present my research at international conferences. I could simply not have wished for a better or friendlier supervisor. I would also like to thank my co-supervisor, Prof. Steve Busby. His quirky advice and support have shaped me as a scientist and it has been a true honour to have had his input. Furthermore, my PhD would not have been possible without the relentless support from Dr. Douglas Browning. I am sincerely grateful to him for teaching me his technical expertise in the lab (and being particularly patient with my extreme clumsiness at times!) and for being a great mentor with his unsurpassed knowledge of transcription. I also thank Doug, Cathy Wardius and Rita Godfrey for helping with some of the experiments.

Above all, I would like to thank Matthew. I feel blessed to have found such a companion to share my PhD experience with and for his support, both on a personal and professional level, my expression of thanks does not suffice.

In my daily work, I feel truly privileged to have met such fun and friendly lab members. I thank everyone, past and present, who have shared these days (and sometimes nights) in the lab with me; Cathy, Rebecca, Faye, Tim, Yana, Karina, Tabish, Riddhi, Tara, Olivia, Sophie and particular thanks goes to David 'Buddy', and Denisse. I am very thankful to Buddy for all the meaningful chats and countless drinks that have been invaluable in seeing me through the more taxing times during my PhD. Having the opportunity to attend various conferences and doing a bit of globetrotting with Denisse has been an absolute hoot; I still wait in anticipation for the day our dance moves become a global phenomenon. You have both become true lifelong friends and the experience would not have been the same without you.

I would also like to thank my personal friends, Hayley, Philippa, Laura, Sarah, Lianne, Jessica, Kerry and Matt who for many years, have been there through my ups and downs.

Finally, I thank my family. My 'Maxi' bro, Leon, has been my rock and has always, so selflessly, given me hours of support and encouragement in all my endeavours, even though I am sure he has absolutely no idea what I do with my time. I thank my 'Mini' bro, Adam, for growing up and becoming a great brother who can always make me laugh; I hope his university experience, which he is soon to embark on, is as enjoyable and worthwhile as mine. I thank my Dad for always knowing how to put a smile on my face with his haphazard stories and fabulous sense of humour and lastly, my Mother; for all her support, love and encouragement and for providing me with a home and good old traditional cooking only a Mother knows how, during the writing of this thesis.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Introduction

Secreted virulence determinants are essential to both colonisation and pathogenesis of Gram-negative bacteria. The biogenesis of virulence determinants is a multi-stage process, which has to be tightly controlled in order to ensure optimal protein expression in the correct growth phase or niche. The journey of a typical secreted protein, from its initial transcription to its final destination outside the cell will be outlined below, detailing the mechanism of each major process, whilst also describing specialised systems that have evolved to circumvent obstacles associated with the secretion of virulence factors through the bacterial cell envelope.

1.2 Transcription in *Escherichia coli*

The ability of a bacterium to survive in a range of hosts and environments largely depends on its capacity to sense and respond to a repertoire of environmental cues. These environmental signals can be integrated into a variety of different pathways that affect gene expression and hence, protein synthesis. *Escherichia coli* uses multiple strategies to modulate protein synthesis, however the most economic choice for the bacterium is to regulate protein synthesis at the principal stage of its biogenesis; transcription initiation. The bacterial RNA polymerase (RNAp) holoenzyme is responsible for all the transcription in the cell and its distribution on the chromosome is subject to regulation by transcription factors (TFs). It is these TFs that mediate responses to environmental cues and translate them into modulation of the transcription machinery.

1.2.1 Bacterial RNA polymerase

Synthesis of all RNA in bacterial cells is dependent on the multi-subunit enzyme, RNAp. The dynamic activities of RNAp determine the ability of pathogens to successfully transit from

one niche to another. Bacterial RNAp exists in two forms, which are functionally distinct from each other; the RNA polymerase core enzyme and the RNA polymerase holoenzyme. The core enzyme comprises 5 subunits; α_2, β, β' and ω (Zhang *et al.*, 1999), which forms the catalytically active, elongation-competent complex (Fig. 1.1). This core composition is conserved in sequence, structure and function, from bacteria to humans, with archaeal and eukaryotic polymerases differing only by the presence of additional subunits (Ebright, 2000; Minakhin *et al.*, 2001). Structural studies, elucidating the organisation of the multi-subunit RNAp, observed a structure reminiscent of a 'crab' claw with two 'pinchers' (the β and β' subunits) shaping the central, active site cavity (Zhang *et al.*, 1999), which is occupied by the DNA template during transcription elongation (Naryshkin *et al.*, 2000). The identical α subunits contain two domains; an N-terminal domain (α-NTD) and a C-terminal domain (α-CTD), which are connected via a flexible linker (Blatter *et al.*, 1994). They form a homodimer in the RNAp structure, with the strongest dimerisation determinants located in the α-NTD (Ebright and Busby, 1995). The flexible linker is instrumental in permitting the interaction of α-CTD with upstream regulatory elements and/or proteins, whilst tethering the α-NTD to the rest of RNAp (Fig. 1.1) (Ebright and Busby, 1995). The ω subunit is non-essential for transcription but plays a role in the assembly of the $\alpha_2\beta$ subunits (Minakhin *et al.*, 2001).

Although the core enzyme is proficient in transcription elongation, association with a σ subunit to create the holoenzyme is necessary for specific, promoter-mediated transcription initiation (Burgess *et al.*, 1969; Typas *et al.*, 2007). Most bacteria have a range of σ subunits that can be divided into two classes; one of which has only a single member σ^{54} and the other sigma factors belonging to the σ^{70} class. Each σ subunit is thought to recognise different promoter elements, hence determining promoter selectivity (Typas *et al.*, 2007). Additionally,

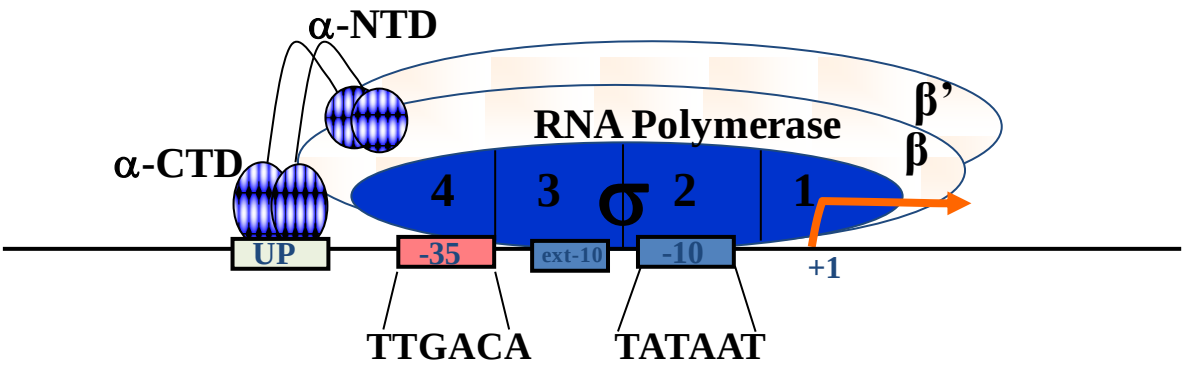


Figure 1.1. ***Schematic representation of the interactions between the RNA polymerase holoenzyme and promoter DNA.*** This figure illustrates the interactions between the different subunits of the RNA polymerase holoenzyme with the core promoter elements. The interactions shown include; the α-CTD subunits with the UP element ('UP'); σ domain 4 with the -35 hexamer element; σ domain 3 with the extended -10 motif ('ext -10'); and σ domain 2 with the -10 hexamer element. Both the -35 and -10 elements are aligned with consensus binding hexamer sequences, shown in bold type. Figure adapted from (Browning and Busby, 2004)

σ is critical for the correct orientation of RNAp at promoters and facilitates subsequent unwinding of duplex DNA, close to the transcript start site (Browning and Busby, 2004; Wosten, 1998). All bacteria contain a major essential σ factor (in the case of *E. coli*, this is σ^{70}), which is responsible for initiating transcription of the housekeeping genes. Other σ subunits are required to co-ordinate gene expression in response to specific environmental stimuli or growth phases. For example, σ^{S} is induced during stationary phase or stress (Typas *et al.,* 2007), whilst σ^{E} responds to extracytoplasmic signals (Helmann, 2002). Furthermore, some σ subunits have now been associated with a specific up-regulation of genes encoding virulence determinants (Kazmierczak *et al.,* 2005) and are therefore indispensable for certain bacterial pathogens such as *Bacillus anthracis* and *Mycobacterium tuberculosis,* to cause disease (Fouet *et al.,* 2000; Geiman *et al.,* 2004). For example, mice infected with a strain of *M. tuberculosis* lacking its alternative σ^{F} sub-unit were attenuated for late-stage disease in comparison to mice infected with the wild type strain (Fouet *et al.,* 2000). Genetic and biochemical analysis of the σ^{70} subunit family revealed four regions of sequence conservation and highlighted the importance of regions 2 and 4 in recognition and binding of the -10 and -35 promoter elements, respectively (Fig. 1.1) (Barne *et al.,* 1997; Fenton *et al.,* 2000). High resolution structural studies confirmed that the σ subunit is comprised of four predominant domains, which assume an entirely α-helical structure; N-terminal domain 1, N-terminal domain 2, 'linker' domain and C-terminal domain (Murakami *et al.,* 2002; Vassylyev *et al.,* 2002).

1.2.2 Transcription initiation

Transcription occurs in three major steps; initiation, elongation and termination. Transcription initiation is further divided into three key stages, involving many structural and conformational changes of RNAp, promoter DNA and regulatory proteins. Firstly, RNAp

holoenzyme binds to the promoter and forms what is referred to as the closed complex (RP_c) (Murakami and Darst, 2003). RP_c then undergoes a conformational change to form the open complex (RP_o), whereby the DNA template is unwound around the transcript start site and brought into close proximity of the active site of RNAp, ready for interaction with the initiating nucleoside triphosphate (NTP) (deHaseth *et al.,* 1998; Murakami *et al.,* 2002). A short, nascent transcript is synthesised by the initial RNAp-promoter transcribing complex, through a scrunching mechanism (Kapanidis *et al.,* 2006). Once the initial transcript reaches the threshold of 8-15 nucleotides (nts) in length, the sigma subunit dissociates from core RNAp (promoter escape) and the core RNAp enters the processive stage of transcription in an RNAp-DNA elongation complex, which is the first irreversible stage of transcription initiation (Browning and Busby, 2004). However, the transition of the initial RNAp-promoter transcribing complex into the RNAp-DNA elongation complex is not a fluent process, as RNAp can go through tens or hundreds of cycles of synthesising an 8-15 nt RNA transcript without successful transition to the elongation complex. This process is known as abortive initiation and has been well characterised as an RNAp phenomenon *in vitro* (Murakami and Darst, 2003) and *in vivo* (Goldman *et al.,* 2009), at many promoters. A role for the aborted 8-15 nt transcripts remains elusive, with theories alluding to its function as a primer for the initiation of RNA synthesis (Goldman *et al.,* 2009; Nickels and Dove, 2011).

1.2.3 Regulation of transcription initiation

Extensive studies on the regulation of transcription initiation have elucidated the many strategies that bacteria employ to modulate initiation of transcription. These include the regulation of chromatin structure (Browning *et al.,* 2010), use of small ligands, such as ppGpp (Srivatsan and Wang, 2008), the exchange of particular σ factors, as described above, promoter DNA sequences and TFs (Browning and Busby, 2004). Whilst all these factors

contribute to the dynamics of gene expression, the latter two have been the subject of most interest and are the most relevant to my work.

1.2.3.1 Promoter DNA sequences

Studies on promoter recognition by RNAp holoenzyme have identified the promoter-specific sequence elements that orchestrate promoter recognition and transcription initiation (Busby and Ebright, 1994; deHaseth *et al.,* 1998; Murakami *et al.,* 2002). Aside from the σ^{54}-dependent promoters, which solely depend on the presence of specific enhancer binding proteins for transcription (Buck *et al.,* 2000), all promoters share a conserved architecture with respect to two predominant features; the -10 and -35 hexamer elements, located 10 and 35 base pairs (bps) upstream of the transcription start site, respectively (Browning and Busby, 2004; Busby and Ebright, 1994). A canonical hexamer binding sequence has been defined for both of these regulatory elements (Fig. 1.1), yet the conservation of each base at each promoter differs, thereby permitting promoter differentiation whilst preventing tight, unfavourable RNAp-promoter interactions. Two other sequence elements that contribute to specific RNAp-promoter interactions are the UP element and the extended -10 motif. The UP element is usually an AT-rich, 20 bp sequence situated upstream of the -35 element and acts as a docking site for the α-CTDs of RNAp (Fig. 1.1) (Ross *et al.,* 2001). The extended -10 element is a 3-4 bp motif located immediately upstream of the -10 element and interacts with domain 3 of the RNAp σ subunit (Fig. 1.1) (Sanderson *et al.,* 2003). All these promoter elements contribute to the specificity and variability of RNAp promoter binding.

1.2.3.2 Transcription factors

The distribution of RNAp at promoters is largely dependent on the activities of TFs, which fluctuate according to external signals or internal metabolic states (Ishihama, 2010). Precise

control of TF activities, in response to such signals, can be achieved in a variety of ways. For example, in the case of the well characterised TF, LacI repressor, the intracellular concentration of a small ligand (allolactose) alters in response to nutrient availability (lactose) and consequently reduces the DNA-binding affinity of LacI, such that the presence of lactose removes LacI-mediated repression, thereby promoting expression of structural genes required for lactose metabolism (Jacob and Monod, 1961; Wilson *et al.,* 2007). Additionally, TF activity can be subject to control via covalent modifications, such as phosphorylation. Phosphorylation of TFs plays an important role in the efficacy of two component systems. Phosphorylation events link the sensing of extracellular stimuli, predominantly by membrane bound sensor-kinases, to a switch that governs the expression of a subset of genes, through phosphorylation of their associated TF (Casino *et al.,* 2010).

The *E. coli* genome contains approximately 300 TFs (Perez-Rueda and Collado-Vides, 2000), which function by binding DNA to activate or repress transcription. Some of these TFs bind and regulate hundreds of genes, whereas others are exclusive to just one or two promoters. In *E. coli*, seven TFs have been identified as 'global' regulators, given that they contribute to the regulation of over 50% of all genes (Grainger and Busby, 2008). One aspect of my work studies the regulation of transcription initiation by two of these global regulators; the cyclic AMP receptor protein (CRP) and the factor for inversion stimulation (Fis), and these are discussed in detail below. It is important to note that the regulation of many pathogen-specific virulence determinants is controlled by dedicated TFs, which have been dubbed 'Master virulence regulators'. For example, in Enteroaggregative *E. coli* 042, the TF AggR is responsible for the up-regulation of key virulence factors, including fimbriae (Nataro *et al.,* 1994) and dispersin (Dudley *et al.,* 2006; Sheikh *et al.,* 2002). Another example is, VirF, the primary regulator of the large *Shigella flexneri* virulence plasmid and its activities are largely

responsible for its characteristic pathogenic attributes (Porter and Dorman, 2002). Global regulators are defined by the pleiotropic phenotype that results from their deletion, and as TFs that regulate a large number of functionally distinct operons (Gottesman, 1984). The seven *E. coli* global TFs identified are listed as CRP, FNR, IHF, Fis, ArcA, H-NS and Lrp (Martinez-Antonio and Collado-Vides, 2003), however there are obvious ommisions from this list, such as HU that could be categorised as a 'global' regulator. Three of these (IHF, Fis and H-NS) are 'nucleoid associated proteins' (NAPs). Although NAPs function as classical TFs at some promoters, their main role is to alter chromosome structure, by bending, bridging or wrapping chromosomal DNA, (Browning *et al.,* 2010; Dillon and Dorman, 2010).

1.2.4 Cyclic AMP receptor protein

The cyclic AMP receptor protein, which was originally identified as the TF essential for the activation of the *E. coli* lactose operon (Emmer *et al.,* 1970; Zubay *et al.,* 1970), can function either as an activator or repressor of transcription initiation. It functions as a homodimer that, when allosterically activated by the small ligand, cyclic adenosine 3',5'-monophosphate (cAMP), binds to specific DNA sequences (Won *et al.,* 2009). DNA targets for CRP contain a 16 bp sequence, consisting of two 8 bp elements organised as an inverted repeat that are each recognised by one of the two CRP monomers (Ebright *et al.,* 1989). Each monomer in the homodimeric complex contains two functional domains connected by a flexible hinge. The smaller C-terminal domain binds to DNA via participation of a helix-turn-helix motif, whilst the larger N-terminal domain is responsible for cAMP binding and also serves as the dimerisation interface (McKay and Steitz, 1981; Weber and Steitz, 1987).

1.2.4.1 CRP-dependent activation

At promoters where CRP activates transcription, it is well established that two surface-exposed determinants, activating region 1 (AR1) and activating region 2 (AR2), mediate interaction with the αCTD and the αNTD of RNAp, respectively (Rhodius *et al.,* 1997). The involvement of these activating regions largely depends on the location of the CRP binding site. There are two major classes of CRP-dependent promoters, known as Class I and Class II (Fig. 1.2). Class I CRP-dependent promoters carry a single CRP binding site positioned at various locations (positions -61.5, -71.5, -82.5 and -92.5) upstream of the -35 element, and CRP interacts with the αCTD of RNAp via AR1 (Fig. 1.2, A). For activation at these promoters CRP and RNAp must bind on the same side of the DNA helix (Gaston *et al.,* 1990). At class II CRP-dependent promoters, the DNA site for CRP binding overlaps the -35 promoter element and is usually centred at position -41.5. At these promoters, both AR1 and AR2 of CRP interact with the αCTD and αNTD of RNAp, respectively (Fig. 1.2, B) (Rhodius *et al.,* 1997).

1.2.4.2 CRP in virulence

CRP was originally identified as an activator of catabolite sensitive genes, in response to glucose starvation (Zubay *et al.,* 1970), hence its alias as catabolite activator protein (CAP). However, it is now apparent that CRP regulates a diverse range of genes including virulence determinants in many Gram-negative pathogens such as *M. tuberculosis, Yersinia pestis, Pseudomonas aeruginosa* and enterotoxigenic *E. coli* (Espert *et al.,* 2011; Fuchs *et al.,* 2010; Stapleton *et al.,* 2010; Zhan *et al.,* 2008). As exemplified in the study of *Y. pestis,* a null deletion of the *crp* gene can lead to an overall attenuation in virulence (Zhan *et al.,* 2008) or it can have a more specific role in the regulation of stationary-phase regulated virulence factors,

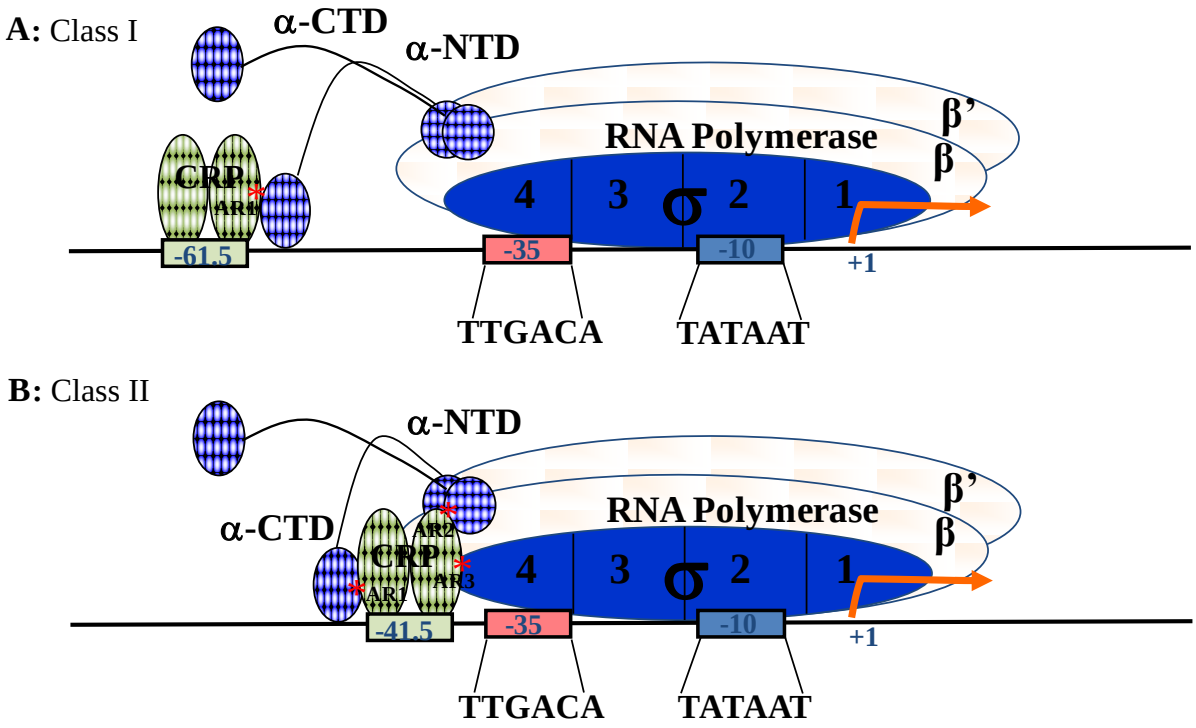


Figure 1.2. ***Interactions at Class I and Class II CRP-dependent promoters.*** **A.** This figure illustrates the organisation of RNA polymerase and CRP at Class I activated promoters. In this situation, CRP binds to a site upstream of the core promoter and binds one α-CTD of RNA polymerase, via Activating region 1 (AR1), mediating its recruitment to the promoter. **B.** This figure illustrates the organisation of RNA polymerase and CRP at class II activated promoters. Here, CRP binds to a site that overlaps the -35 hexamer element which facilitates many CRP interactions with RNA polymerase; AR1 binds one of the α-CTDs, Activating region 2 (AR2) interacts with one of the α-NTDs. AR3 is a non-native region that binds to σ domain 4 (Bell *et al.*, 1990; Rhodius and Busby, 2000). Red asteriks represent the site of protein-protein interactions. All the above interactions mediate recruitment of RNA polymerase to the promoter. Figure adapted from (Busby and Ebright, 1999).

such as the Type 1 Fimbriae of uropathogenic *E. coli* (Muller *et al.,* 2009), which is essentially tailored to the expression pattern of CRP late in *E. coli* growth.

1.2.5 Nucleoid associated proteins

The compaction of DNA into confined subcellular compartments is seen in cells of all kingdoms of life (Luijsterburg *et al.,* 2008). In eukaryotes, the well characterised histone proteins are responsible for forming higher order chromosome structures referred to as chromatin, which play a vital role in all DNA transactions, including transcription and replication (Li and Reinberg, 2011). In bacteria, investigations into functionally similar structures have revealed the presence of a diverse set of structurally unrelated proteins, named 'nucleoid associated proteins' (NAPs) (Azam and Ishihama, 1999). Thus far, in *E. coli* 12 NAPs have been identified (Dillon and Dorman, 2010), yet others are being found with further investigations (Teramoto *et al.,* 2010). A common theme with all characterised NAPs is their ability to bind DNA, yet the subsequent compaction of DNA is mechanistically distinct. Using single-molecule techniques and high-resolution imaging of protein-DNA complexes, observations of H-NS mediated compaction shows the propensity of H-NS to 'bridge' DNA (Dorman and Kane, 2009; Wiggins *et al.,* 2009). In contrast, HU, a constitutively expressed NAP, exhibits dual activities by virtue of DNA bending and condensation (van Noort *et al.,* 2004). Most relevant to my work is the expression and activities of the NAP, Fis (Fig. 1.3). Fis is the most promiscuous NAP in *E. coli,* given that it has the capacity to affect gene expression by at least six different mechanisms (Browning *et al.,* 2010; Opel *et al.,* 2004; Squire *et al.,* 2009) and is discussed in detail below.

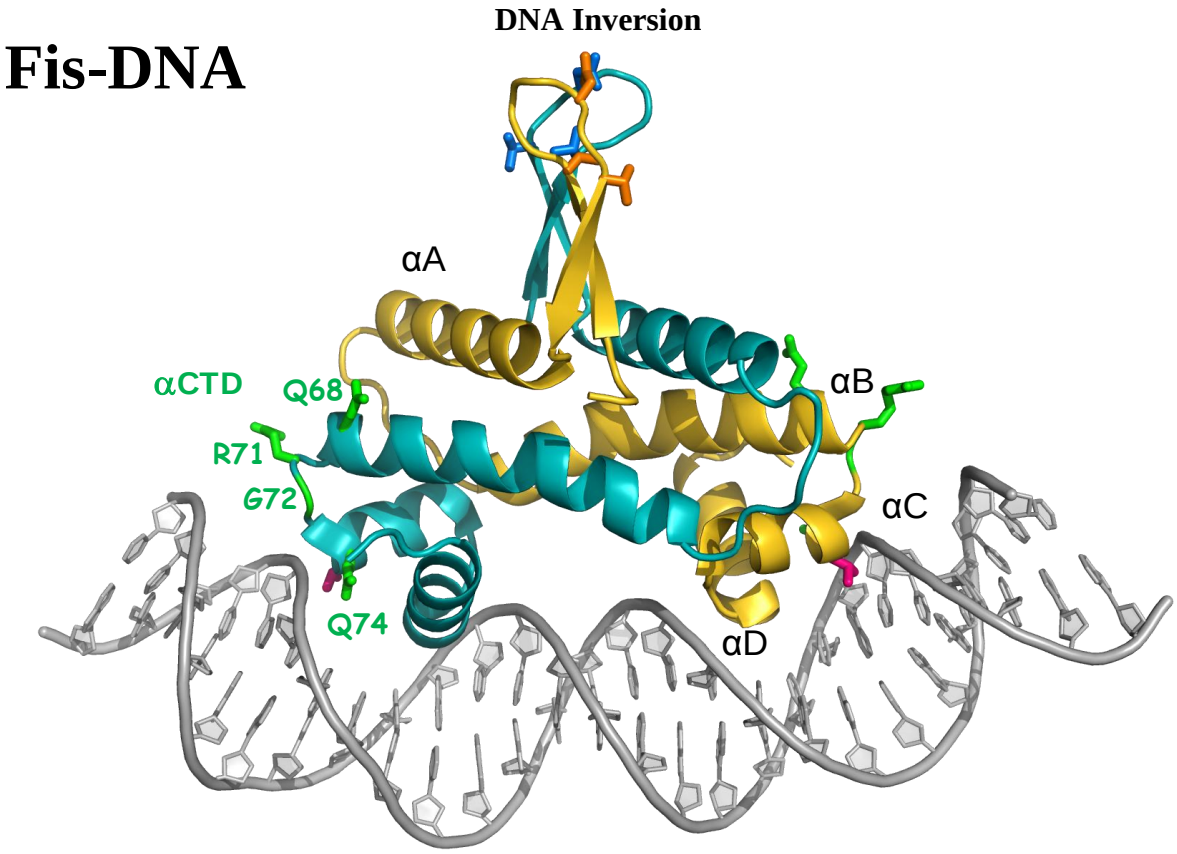


Figure 1.3. ***Fis-DNA interactions.*** This figure shows the residues important for mediating contact between an activation patch on Fis and the α-CTD of RNA polymerase. The residues highlighted in green are those that have been identified, through mutagenesis studies, to contact RNA polymerase and activate transcription at two independently tested promoters (*rrnB1* and *proP2*) (Bokal *et al.*, 1997; McLeod *et al.*, 2002). The locations of these residues are physically and functionally distinct from the region of Fis necessary for its regulation of site-specific DNA inversion, labelled in the figure as 'DNA inversion'. The A-D helices, present in each monomer, are labelled in the figure as αA-αD. Figure provided by Reid Johnson.

1.2.6 Factor for inversion stimulation (Fis)

Fis was originally discovered as an *E. coli* protein essential for the action of a bacteriophage-encoded site specific recombinase and as a co-factor in flagellar phase variation in *Salmonella* (Johnson *et al.*, 1986; Koch and Kahmann, 1986). However, Fis is now known to play many important phage-independent functions by binding and bending DNA (Kostrewa *et al.*, 1991), organising local topology (Travers *et al.*, 2001) and acting as a TF by activating or repressing transcription initiation (McLeod *et al.*, 2002). Similarly to CRP, Fis functions as a homodimer that binds to 15 bp target sequences (Fig. 1.3). Crystallography studies revealed an α-helical core, with each monomer comprising of 4 α helices (A-D) with the latter two helices (C and D) forming the DNA binding helix-turn-helix motif. However, in contrast to CRP, where the DNA recognition sequence is highly specific, the degenerate nature of the DNA recognition sequence for Fis may be due to an unusually high number of positively charged residues in the D helix of the helix-turn-helix motif (Kostrewa *et al.*, 1991). Recent crystallography studies revealed that Fis recognises DNA targets primarily through indirect mechanisms involving the shape of the minor groove, whilst interactions with the adjacent major groove is sequence-dependent (Stella *et al.*, 2010). Additionally, specific residues (Q68, R71, G72 and Q74) have been identified in a surface-exposed loop adjacent to the helix-turn-helix motif of Fis that constitute an activation patch responsible for direct interactions with the α-CTD of RNA polymerase (Fig. 1.3) (Bokal *et al.*, 1997). The activity of Fis is controlled by its concentration within the cell, and its growth-phase dependent regulation restricts its expression to the logarithmic phase of growth (Mallik *et al.*, 2006), consistent with the observation that Fis governs control of rRNA promoters, essential for ribosome synthesis during rapid cell growth (Nilsson *et al.*, 1990). Furthermore, *fis* is part of a bicistronic *dusB-fis* operon, where both *fis* and the preceding gene, *dusB*, which encodes a tRNA

dihydrouridine synthase, are controlled by the integration of IHF, CRP and Fis activites at their promoter region (Nasser *et al.*, 2002).

1.2.6.1 Fis in virulence

Similarly to CRP, the biological role of Fis in virulence has taken precedence over its initially established role as a factor that regulates the activity of a bacteriophage specific recombinase. Many studies have highlighted a role for Fis in the virulence of bacterial pathogens, however it is largely renowned for its prominent role in the pathogenesis of *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium). Fis exerts its effects through the control of one of the major virulence loci, named *Salmonella* pathogenicity island 2 (SPI-2) (Fass and Groisman, 2009; Ó Cróinín *et al.*, 2006). Fis mediates direct activation of the SPI-2 *ssrAB* encoded two-component system, which is responsible for regulating the SPI-2 encoded type III secretion system that delivers effector proteins directly into host cells (Fass and Groisman, 2009).

1.2.7 Models of transcriptional activation and repression

TFs predominantly function as either activators or repressors of transcription, although at some target promoters, TFs can display both activities. Activators use various mechanisms to promote transcription and three functionally distinct classes have been described. In class I activation, activators bind upstream of the transcript start site and typically function by making direct protein-protein interactions with the α-CTD of RNAp, mediating recruitment of RNAp to the promoter (Fig. 1.4, A). In class II activation, activators bind to the promoter region adjacent to, or overlapping, the -35 element and make direct contact with domain 4 of σ^{70}, again mediating recruitment of RNAp to the targeted promoter (Fig. 1.4, B) (Browning and Busby, 2004). In class III activation, a combination of activators is involved, making both

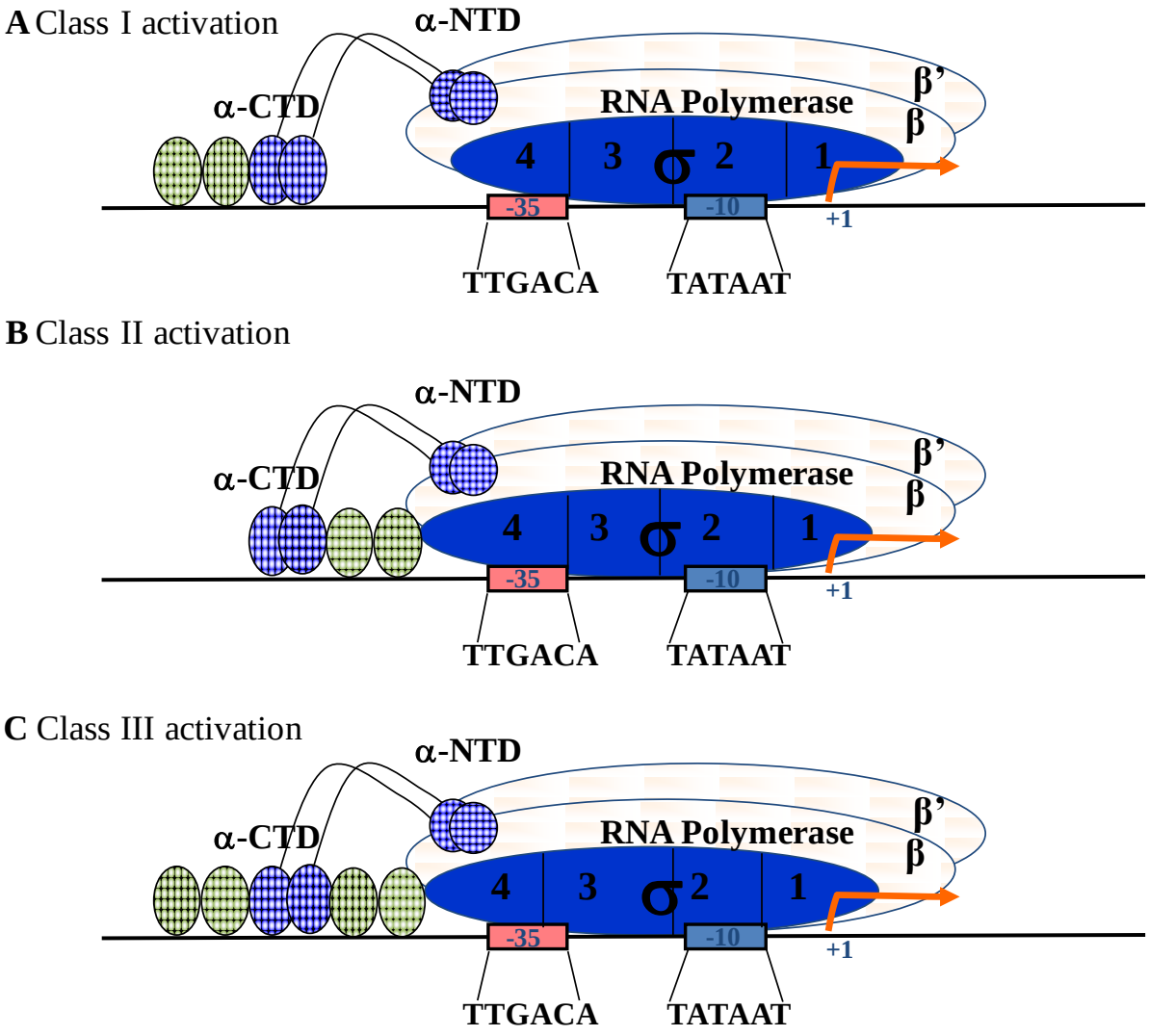


Figure 1.4. ***Models of transcriptional activation.*** This figure illustrates the organisation of RNA polymerase and activators at promoters that are activated in a class I, II and III fashion. In this diagram, activators are represented in green as dimeric proteins. RNA polymerase sub-units are shown in blue. **A.** Class I activation; activator is bound upstream of the transcript start site and typically function by making direct protein-protein interactions with the α-CTD of RNAp, mediating recruitment of RNAp to the promoter. **B.** Class II activation; activators bind to the promoter region adjacent to the -35 element and make direct contact with domain 4 of σ^{70}, again mediating recruitment of RNAp to the target promoter. **C.** Class III activation; a combination of activators are involved, making both class I and class II interactions to achieve full activation. Figure adapted from (Browning and Busby, 2004).

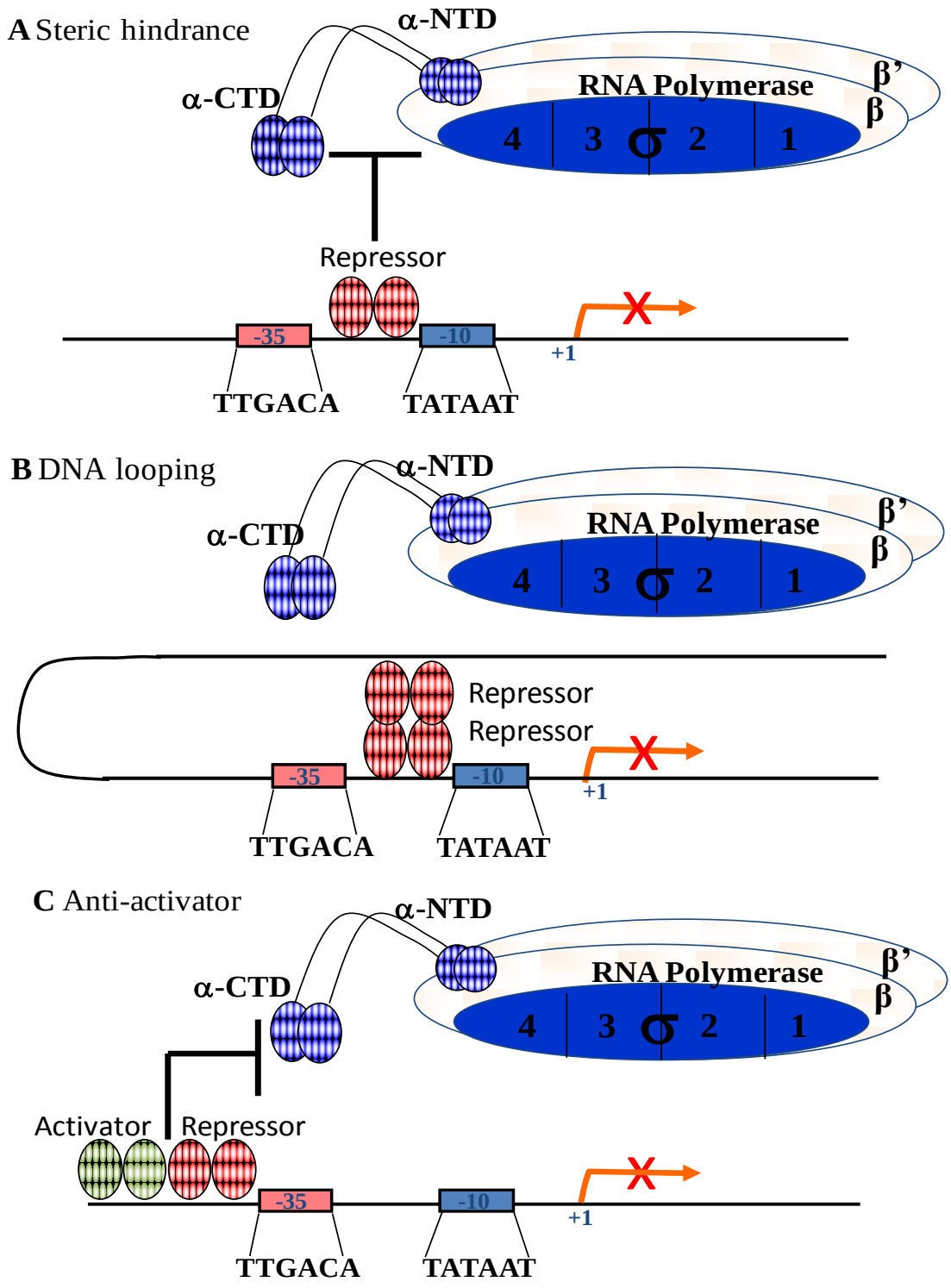


Figure 1.5. ***Models of transcriptional repression.*** This figure illustrates the typical mechanisms used to repress transcription at target promoters. **A.** Repression by steric hindrance; the repressor binds to a site that occludes the core promoter elements from RNA polymerase, thereby abrogating RNA polymerase binding. **B.** DNA looping**;** repressors bind to sites separate from the promoter and act by looping the DNA and blocking access of RNA polymerase to the promoter. **C.** Repression by blocking the activities of an activator; the repressor binds directly to an activator, thereby disrupting interactions with RNA polymerase such that RNA polymerase is no longer recruited to the promoter. Figure adapted from (Browning and Busby, 2004).

class I and class II interactions to achieve full activation (Fig. 1.4, C) (Beatty *et al.,* 2003). Other activators function by causing a conformational change of the DNA topology around the promoter, such that the -10 and -35 elements are more accessible to RNAp (Brown *et al.,* 2003).

Repression, in the simplest form, occurs when a repressor is bound at or near the promoter such that RNAp is unable to bind to the target DNA (Fig. 1.5, A) (Wilson *et al.,* 2007). More complex mechanisms exist whereby repressors cause looping of the DNA, thereby abrogating the ability of RNAp to initiate transcription (Fig. 1.5, B) (Choy and Adhya, 1992). Repressors have also been observed to act as 'anti-activators', whereby interaction with a bound activator disrupts its activity, leading to an indirect repression (Fig. 1.5, C) (Shin *et al.,* 2001).

1.2.8 Co-dependence at promoters

For bacterial pathogenesis, integration of a multitude of signals, specific to various niches within their infected host, is crucial to co-ordinating an appropriate gene expression response. In the case of *S.* Typhimurium infection, passage through the acidic stomach, survival of bile, attachment to intestinal lining, subsequent invasion into mammalian cells and evasion of the immune system all require differential expression of a wide array of virulence genes (Lucas and Lee, 2000). Intriguingly, *S.* Typhimurium *crp* mutants have proven to be effective in the preparation of a live oral vaccine (Curtiss *et al.,* 1988). Thus, in contrast to the simple models of activation and repression outlined in section 1.2.7, the complexity of bacterial pathogenesis is reflected in the organization of TFs at certain promoters, where co-dependence on two or more TFs is often required for transcription initiation. The mechanistic details of co-dependence can vary greatly between promoters (Barnard *et al.,* 2004). TFs can often make independent contacts with RNAp and result in the dependence of two bound activators for full

activation of the promoter (McLeod *et al.,* 2002). Co-dependence is occasionally observed in the initial binding of TFs to the promoter. For example, in *Vibrio cholerae*, co-operative binding is seen with the TFs AphA and AphB at the *tcpPH* promoter, which regulates the cholera toxin and toxin-co-regulated pilus (Kovacikova and Skorupski, 2001). Additionally, the integration of multiple transcription factors is required in scenarios whereby 'anti-repression' is used as a mechanism for activation at a given promoter. This occurs when a secondary activator inhibits the activities of a repressor, thereby allowing the primary activator to function (Browning and Busby, 2004).

1.2.9 Post-transcriptional control of gene expression in *E. coli*

Although transcription initiation is the favoured mode of regulation, it is important to note that many post-transcriptional regulatory processes have been observed in prokaryotes. These include small RNA (sRNA) mediated regulation of mRNA, mRNA secondary structure, mRNA decay, ribosome synthesis and activity, modulation of translation efficiency, protein stability and the recently characterised importance of mRNA leader sequences (Caron *et al.*, 2010; Park *et al.*, 2010; Picard *et al.*, 2009; Raghavan *et al.*, 2011). Although all these processes play a vital role in the genetic output of the cell, the huge pleiotropic effects observed in an RNA chaperone (Hfq) mutant is a good example of the prolific post-transcriptional regulation that is essential for the virulence of many pathogens, including *V. cholerae*, *Listeria monocytogenes*, *P. aeruginosa* and *S.* Typhimurium (Christiansen *et al.*, 2004; Ding *et al.*, 2004; Sonnleitner *et al.,* 2003).

Once transcription has been activated and the barriers to post-transcriptional regulation have been overcome, the mRNA is translated and protein synthesis occurs. Inherent to all virulence effector proteins is the necessity to reach their surface-exposed or extracellular target. In the

case of Gram-negative pathogens, virulence factors must traverse both layers of the cell envelope. The ways in which virulence factors undergo such directive processes to overcome this additional obstacle are discussed below.

1.3 Gram-negative cell envelope

1.3.1 Inner membrane composition

In contrast to Gram-positive bacteria, where a thick peptidoglycan layer coats a single plasma membrane, Gram-negative bacteria are enclosed by two physically distinct lipid bilayers; with an intervening peptidoglycan-containing periplasmic space. They are termed the inner and outer membrane (OM) and are characterised by distinctly different protein and lipid compositions (Fig. 1.6) (Silhavy *et al.,* 2010). The inner membrane (IM) is a phospholipid bilayer, primarily consisting of phosphatidyl ethanolamine, phosphatidyl glycerol and cardiolipin with the protein composition being mainly α-helical in structure (Raetz and Dowhan, 1990; Silhavy *et al.,* 2010). In comparison to the OM, the IM is structurally less complex, but, even so, it has remarkably diverse functions. It is also the site for the synthesis and assembly of structural wall components including lipopolysaccahrides (LPS), phospholipids and the basic chemical units of peptidoglycan (Costerton *et al.*, 1974). The IM poses the first barrier to protein secretion. Two systems that permit translocation of proteins across the IM have been described. These are named the Sec and the Tat systems and are responsible for the translocation of unfolded and folded species, respectively (Cristobal *et al.*, 1999).

1.3.2 The periplasm

The next cell wall layer is the peptidoglycan layer within the periplasmic space, which serves

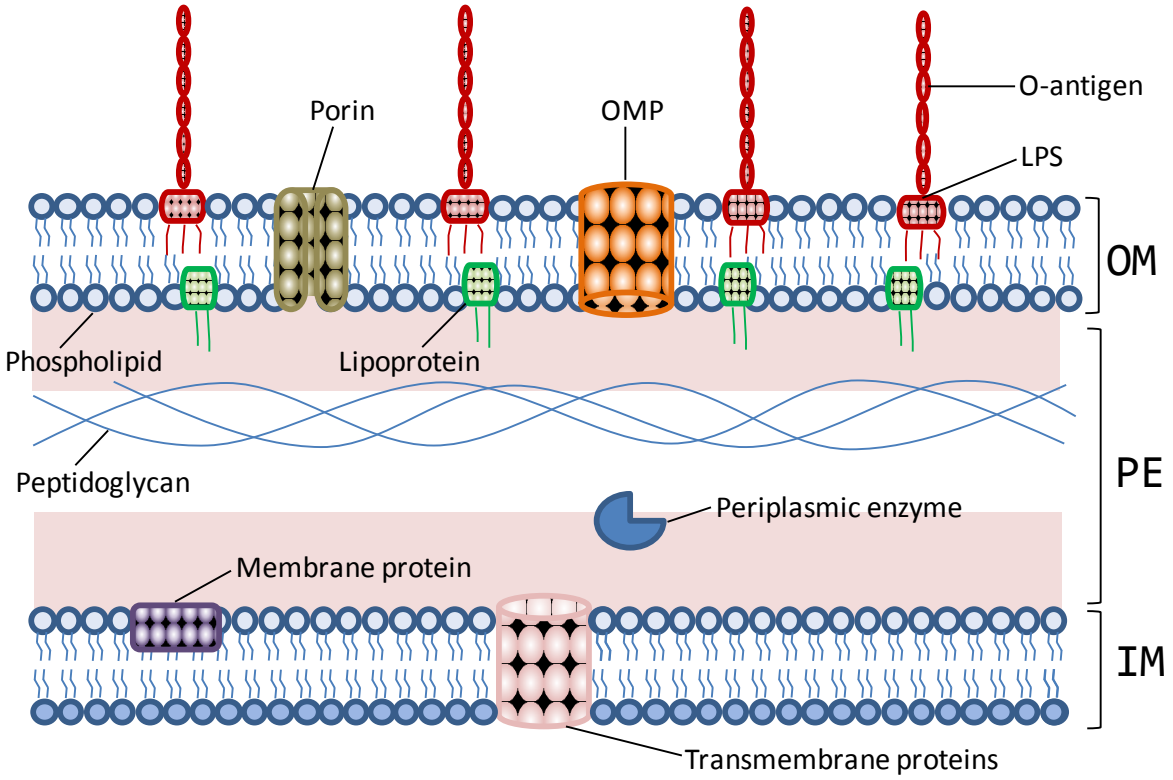


Figure 1.6. ***Gram-negative cell envelope.*** Depicted in this figure is the organisation of the main components of the Gram-negative cell envelope; OMP, outer membrane protein; LPS, lipopolysaccharide; OM, outer membrane; PE, periplasm; IM, inner membrane.

a structural role providing the bacterium with its shape and rigidity (Fig. 1.6) (Silhavy *et al.*, 2010). The peptidoglycan is a complex structure which remains relatively invariant across most bacteria. The anchoring of the most abundant protein in *E. coli*, Lpp (or Braun lipoprotein), to peptidoglycan acts as the foundations for attachment of the OM (Braun, 1975). The surrounding periplasmic space is an extremely viscous sub-cellular compartment, found to contain functionally diverse proteins (Mullineaux *et al.*, 2006). One group of proteins exhibit protective functions, acting to modify toxic compounds such as heavy metals or antibiotics, for example, the degradation of β-lactam antibiotics by β-lactamase. Another group of periplasmic proteins comprise substrate binding proteins that are indispensable to two fundamental processes; chaperoning and folding (Kadokura and Beckwith, 2010; Sklar *et al.*, 2007b). The latter of these processes has now been well characterised and is found to be under the control of the 'Disulphide bond forming' proteins, DsbA-DsbD and DsbG. These proteins form a network of both periplasmically and IM located enzymes that catalyse the formation of disulphide bonds. DsbA is the primary oxidant of periplasmically located proteins and catalyses the formation of disulphide bonds between cysteine residues by virtue of donating its own disulphide bond to its substrate. DsbB-D and DsbG act in concert to stabilise the redox potential necessary to ensure optimal processivity of substrates (Kadokura and Beckwith, 2010).

In Gram-negative bacteria, export of proteins outside the bacterial cell can be achieved using molecular architecture that spans their inherent double membrane, thus averting exposure of their secreted substrates to the periplasm (Henderson *et al.*, 2004). However, many virulence determinants need to passage through the periplasm in order to reach their extracellular or membrane bound target. The chaperones that enable transit of proteins across the periplasm constitute three distinct groups; those that catalyse the formation of disulphide bonds, as

described above, peptidyl-prolyl cis/trans isomerases (PPIases) and those that display general chaperone activities (Sklar *et al.*, 2007b). The latter group are periplasmic proteins that play a major role in the delivery of outer membrane proteins (OMPs) to the OM. Three main chaperones have been described for this purpose; DegP, Skp and SurA. Unique to DegP, is the ability to exhibit both protease and chaperone activities in a temperature-dependent manner (Spiess *et al.*, 1999). Given that DegP chaperone activity is more efficient at lower temperatures, higher temperatures govern the switch to protease activity, which subsequently degrades misfolded protein species (Spiess *et al.*, 1999). Skp was characterised as a chaperone with substrate specificity restricted to OMPs (Chen and Henning, 1996), whereby a *skp* null mutant has decreased levels of the OMPs LamB, OmpA, OmpC, and OmpF (Chen and Henning, 1996). SurA encompasses two of the periplasmic chaperone groups given that it displays both peptidyl-prolyl cis/trans isomerase activity and general chaperone activity (Behrens *et al.*, 2001). Given its broad activities, a *surA* null mutant exhibits the most dramatic phenotype, with a compromised OM structure due to the absence of many OMPs (Rouviere and Gross, 1996). These three periplasmic proteins constitute the major chaperones involved in maintaining OM integrity. One particular study proposed that they form two parallel pathways with functional redundancy (Rizzitello *et al.*, 2001). A subsequent study investigated this redundancy demonstrating that SurA forms the primary pathway responsible for the delivery of most OMPs to the OM, whereas DegP and Skp together form a parallel pathway that only plays a minor role in OMP delivery by rescuing proteins that have fallen off the SurA pathway (Sklar *et al.*, 2007b).

1.3.3 Outer membrane composition

The OM, being the terminal layer of the Gram-negative cell envelope, possesses a more diverse range of components than its IM counterpart (Fig. 1.6). The more abundant

components of the OM include phospholipids, LPS, OMPs and lipoproteins (Silhavy *et al.*, 2010). Lipoproteins are involved in a wide range of cellular functions, from substrate transport to antibiotic resistance. More than 90 species of lipoproteins are found in *E. coli*, most of which are membrane attached via an N-terminal N-acyl-diacylglycerylcysteine and with only a few exceptions, such as SphB1 where the functional domain is surface localised, are orientated facing the periplasm (Bos and Tommassen, 2004). Localization of lipoproteins to the OM has been found to require a lipoprotein-specific sorting machinery, the Lol system, which is composed of five proteins, LolABCDE (Fig. 1.7). LolCDE is responsible for detachment of the lipoproteins from the IM allowing the association of the lipoprotein with the periplasmic chaperone LolA, which delivers the lipoprotein to LolB, the dedicated receptor in the OM (Narita and Tokuda, 2006). LPS is made up of a three-part architecture and is essential for most Gram-negative bacteria. This common structure protrudes from the bacterial cell surface consisting of a distal O-antigen, linked to an oligosaccharide core structure, which is then anchored to the membrane by lipid A. LPS is responsible for causing characteristic endotoxic shock associated with the septicaemia caused by Gram-negative bacteria (Raetz and Whitfield, 2002). Analogous to the Lol system for lipoprotein sorting, delivery of LPS to the OM involves a similar network of IM, periplasmic and OMPs named LptA-F (Fig. 1.7). After synthesis of LPS in the inner leaflet of the IM followed by flipping of LPS to the outer leaflet of the IM by the ABC transporter, MsbA, LptA-F work in a similar mode to Lol, to deliver LPS to the OM. LptB, C, F and G form the IM and cytosolic bound proteins to detach LPS from the IM and deliver it to the periplasmic chaperone LptA. LptA then docks with the OM assembly site comprising of LptD and E (Ruiz *et al.*, 2008).

OMPs are largely transported across the IM via the Sec system. Once they reach the periplasm, the periplasmic chaperones outlined above deliver them to the OM, prior to

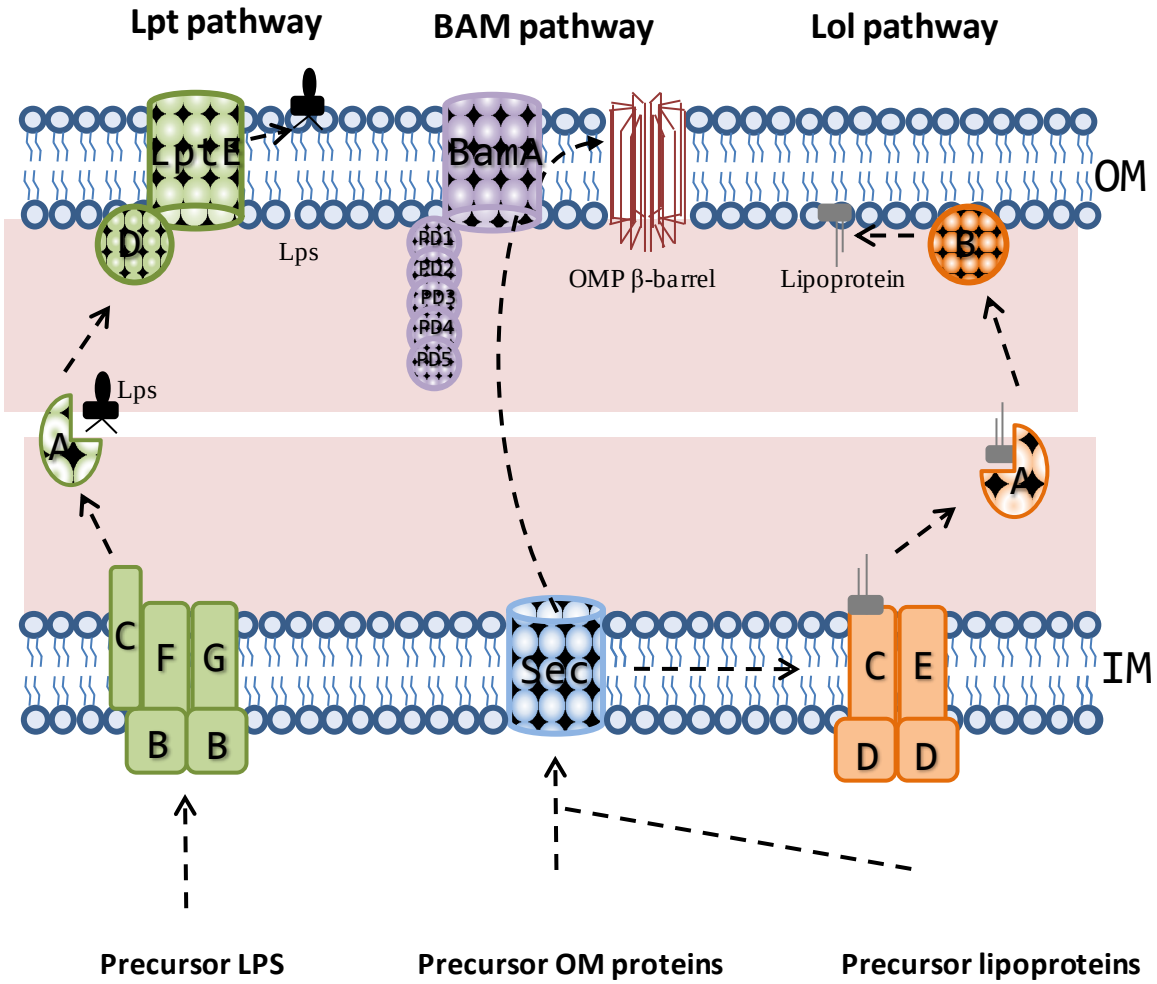


Figure 1.7. ***Outer membrane biogenesis.*** This figure illustrates the cellular machineries necessary for the biogenesis of the major OM constituents; LPS, OMPs and Lipoprotein. LPS is transported from its site of synthesis in the cytoplasm to the outer membrane by the Lpt pathway. OMPs and Lipoproteins are all synthesised with an N-terminal signal sequence that directs transport across the inner membrane via the Sec apparatus. OMPs are then chaperoned to the OM via the major periplasmic chaperones (SurA, DegP and Skp) where they dock with the BAM complex for insertion into the OM in their native β-barrel structure. After transport across the IM and cleavage of their signal sequence, lipoproteins are transported to the inner leaflet of the OM via the Lol system. Note that the mechanism used for incorporation of phospholipids into the outer membrane remains uncharacterised. Figure adapted from (Silhavy *et al.*, 2010).

insertion into the OM as a folded species. Studies of folded OMPs have identified structural conservation, thus the majority of OMPs possess an even number of antiparallel β-strands which assume a cylindrical β-barrel conformation (Wimley, 2003). These embedded proteins serve many functions vital to the bacterial cell, including signal transduction and solute and protein translocation (Koebnik *et al.,* 2000). The insertion and/or folding of all β-barrel OMPs studied thus far is found to be dependent on the recently described β-barrel Assembly Machinery (BAM) complex (Doerrler and Raetz, 2005; Werner and Misra, 2005; Wu *et al.,* 2005) (Fig. 1.7). Initially identified as Omp85 from *N. meningitidis* (Voulhoux *et al.,* 2003), BamA was thought to play a role in LPS biogenesis and phospholipid incorporation into the OM (Genevrois *et al.,* 2003). However, it is now widely accepted that BamA has a pivotal role in OMP assembly and that its effects on phospholipid incorporation and LPS biogenesis is indirect. BamA is part of a super-family of proteins that include transporters found in the membranes of chloroplasts and mitochondria and in all Gram-negative bacteria (Paschen *et al.,* 2003). BamA is comprised of two distinct structures; a C-terminal β-barrel that assumes a cylindrical pore in the OM with a diameter of approximately 2.5 nm (Robert *et al.,* 2006) and a set of five POTRA domains that are anchored to the β-barrel and protrude into the periplasm (Sanchez-Pulido *et al.,* 2003). High resolution structural studies have revealed unique folds and interactions of the POTRA domains from *E. coli,* demonstrating intervening flexible linkers which could generate a large amount of conformational freedom between the domains; thus conferring the ability to bind and create folding intermediates with the associated OMP substrates (Kim *et al.,* 2007; Knowles *et al.,* 2008). BamA is the central component of the BAM complex. BAM has now been well characterised as a hetero-oligomeric complex composed of five proteins; the integral β-barrel membrane protein, BamA, and four OM-associated lipoproteins, BamB-E (Fig. 1.7) (Knowles *et al.,* 2009). A functional role for

BamB-E in OMP assembly has been confirmed given that deletion of any one of these lipoproteins leads to reduced OMP profiles and cell death (Malinverni *et al.,* 2006; Wu *et al.,* 2005). Furthermore, based on co-purification, binding and mutagenesis studies it has been demonstrated that BamC, D and E interact with BamA via its POTRA domain 5, whilst BamB interacts directly with BamA, independent of any other BAM components (Malinverni *et al.,* 2006; Sklar *et al.,* 2007a; Vuong *et al.,* 2008). Although these lipoprotein components are crucial to the functional integrity of the complex (Charlson *et al.,* 2006; Knowles *et al.,* 2011) (Sklar *et al.,* 2007a) only BamA and BamD are universally essential for both bacterial cell viability and OMP biogenesis (Malinverni *et al.,* 2006).

1.4 Protein translocation across the cell envelope

1.4.1 Inner membrane translocation

Roughly 25-30% of bacterial proteins serve structural and functional roles in the cell envelope, or outside the bacterial cell (Driessen and Nouwen, 2008). Bacteria have therefore evolved dedicated systems for the efficient transport and/or insertion of proteins across or into the appropriate cell membrane. Extensive studies in this area have elucidated the presence of a protein translocation pathway, common to both Gram-positive and Gram-negative bacteria that mediate translocation of newly synthesised proteins, across the cytoplasmic membrane before they acquire their final folded structure (Breyton *et al.,* 2002; Wickner *et al.,* 1991). This system has been named the Sec pathway, where proteins destined for translocation across the cytoplasmic or IM converge on the Sec translocase. The Sec translocase comprises a range of functionally diverse proteins; the Sec translocon consists of the integral membrane proteins, SecY, SecE and SecG, which forms a multimeric membrane bound complex, creating a channel for the passage of polypeptides (Fig. 1.7) (Breyton *et al.,* 2002). Complete

with the peripheral subunit, SecA, the Sec translocase machinery is responsible for driving the transport of proteins across the IM (de Gier and Luirink, 2001). Many essential membrane events, including IM translocation, are coupled with ATP hydrolysis to generate the energy required to drive these processes. Thus, these processes are all contingent on the proton motive force to drive ATP synthesis via the membrane-bound ATP synthase (Weber, 1991). Additional factors interact to form a sub-complex named SecDFYajC. This sub-complex has been suggested to play a diverse role in IM translocation with roles ranging from SecA functioning, release of translocated proteins and maintenance of the proton motive force across the membrane (de Gier and Luirink, 2001). Although most proteins do converge at the Sec translocase, the mode of targeting to this pathway differs. The SecB chaperone is involved in the targeting of proteins post translationally, where SecB binds the protein after it is released from the ribosome, keeping it in an unfolded state before delivering it to the SecYEG translocon (Valent *et al.,* 1998). In contrast, the signal recognition particle (SRP) pathway is employed for co-translational targeting of IM proteins (Froderberg *et al.,* 2004). The SRP is composed of the Ffh protein and 4.5S RNA. It acts by binding the first transmembrane segment of the protein upon exposure from the ribosome. The binding of SRP with its IM receptor, FtsY, stimulates the dissociation of the protein from the SRP, such that the protein is translocated across the IM as it emerges from the ribosome (Froderberg *et al.,* 2004). Both targeting mechanisms depend on the presence of a signal sequence at the N-terminus of the secreted proteins. Although variations in both sequence and length of these signal sequences have been observed, attributable functions, additional to targeting, remain contentious (Leyton *et al.,* 2010; Szabady *et al.,* 2005).

An additional pathway for IM translocation was later discovered that, in contrast to the Sec translocon, governs the translocation of fully folded proteins (Santini *et al.,* 1998). Proteins

exported by this Tat (twin-arginine transport) pathway are synthesised as precursors with N-terminal signal peptides bearing conserved SRRXFLK twin-arginine motifs within their amino acid sequence (Cristobal *et al.,* 1999). All precursors exhibiting this sequence are ultimately transported across the IM by the twin-arginine translocation apparatus (Sargent, 2007).

1.4.2 Outer membrane translocation

The OM, along with the periplasm and IM, constitutes the Gram-negative bacterial cell envelope. Although their unifying structure lends protection to the contents of the cell, it also leaves the bacterium with a formidable obstacle for proteins that are destined for the external milieu. Therefore, bacteria have evolved many different secretion systems (Type 1–Type 6) for the translocation of proteins to the cell surface and extracellular space (Fig. 1.8), which can be categorised into systems that contain their own dedicated machinery for IM translocation or those that are dependent on the Sec apparatus.

1.5 Secretion pathways of Gram-negative bacteria

1.5.1 Sec-independent translocation

1.5.1.1 Type 1 secretion

Proteins targeted to the type 1 secretion system (T1SS) secretion system can be assigned to a range of distinct families; proteases, lipases, haemophores, S-layer proteins and repeats in toxin proteins (RTX). The latter of these families contains haemolysin A (HlyA), which best exemplifies the model of the Type 1 secretion system (TISS) (Gentschev *et al.,* 2002). Secreted by uropathogenic *E. coli* (UPEC), HlyA is a virulence-associated protein owing to its cytolytic and cytotoxic activity on certain mammalian cell types such as erythrocytes,

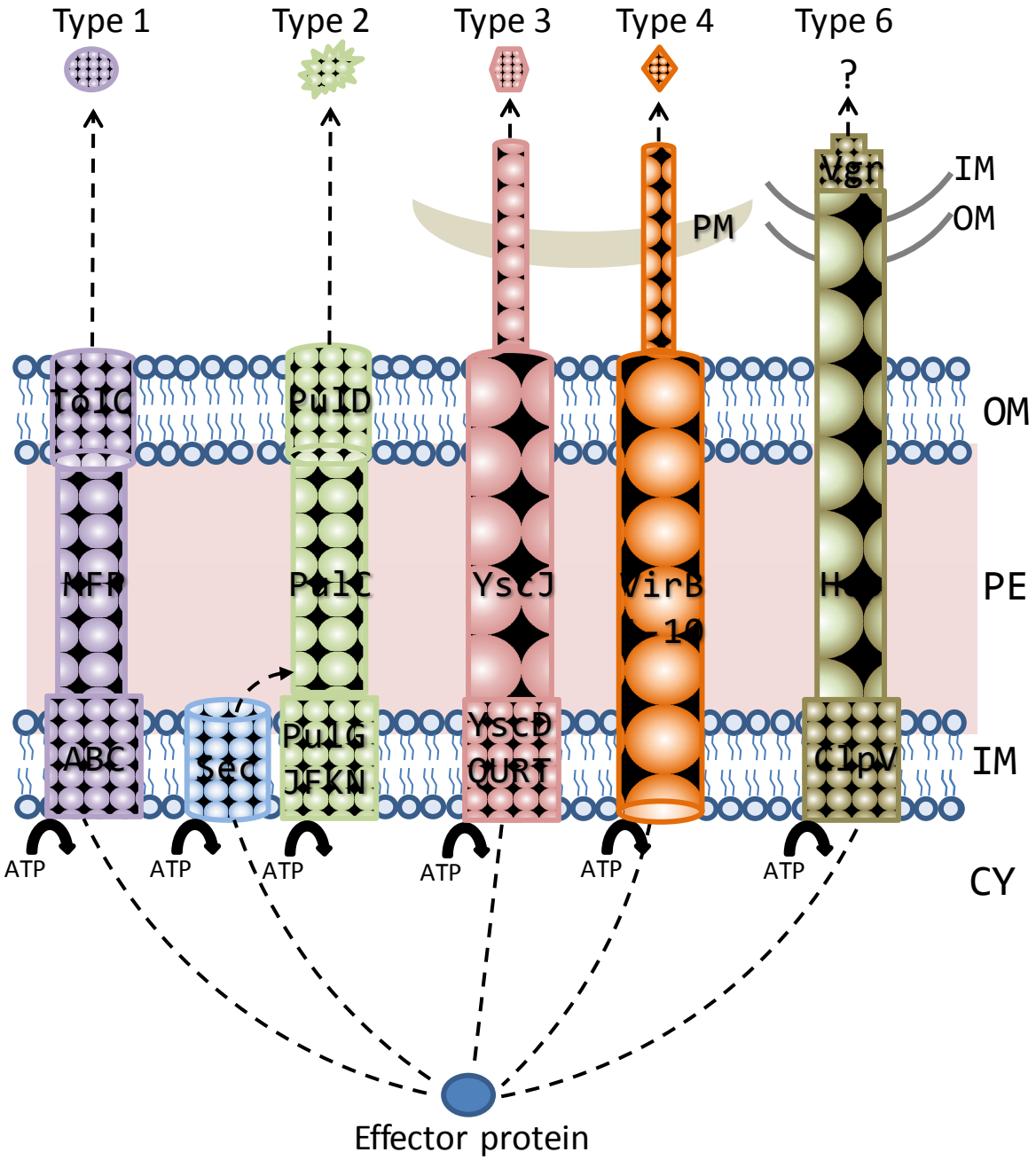


Figure 1.8. ***Type 1, 2, 3, 4 and 6 Secretion systems of Gram-negative bacteria.*** This figure illustrates a simplified view of the molecular structures found in the protein secretion systems of Gram-negative bacteria. Depicted from left to right, the Type 1 secretion system is exemplified by the secretion of HlyA from *E. coli*. The Type 2 secretion system is exemplified by the secretion of pullulanase from *K. oxytoca*. The Type 3 secretion system is represented by the secretion of Yop proteins, by *Y. pestis* and the Type 4 secretion system is exemplified by the VirB system in *A. tumefaciens*. Both the Type 3 and Type 4 secretion systems have the ability to inject their associated effectors directly into the host cytoplasm and unique to the type IV system is the ability to traffic DNA directly into target host cells (Hayes *et al.*, 2010). Little is known about the organisation, function and mechanism of the recently described Type 6 secretion system. However, the conserved T6SS proteins Hcp, Vgr and ClpV are found in many pathogenic bacteria and are thought to assemble as shown in the diagram to inject effectors directly into neighbouring bacteria (Hood *et al.*, 2010; Kanamaru, 2009). Effector proteins delivered through the Type 6 machinery are yet to be fully characterised. CY; cytoplasm, IM; inner membrane, PE; periplasm, OM; outer membrane; PM; host cell plasma membrane. Figure adapted from (Filloux *et al.*, 2008; Henderson *et al.*, 2004).

granulocytes, monocytes and endothelial cells (Gentschev *et al.,* 2002).

The T1SS secretes HlyA, and similar proteins, in a Sec-independent manner, in the absence of any periplasmic intermediates. This is achieved by integration of an IM ATP binding cassette (ABC), a membrane fusion protein (MFP) and a pore-forming OMP. In the case of HlyA, these are represented as HlyB, HlyD and TolC (Tolerance to colicins) respectively. HlyA binds to the cytoplasmic side of the HlyB-D complex by virtue of its C-terminal secretion signal. This interaction stimulates the extension of HlyD to associate with the trimeric TolC located in the OM, which initiates a conformational change in TolC permitting the formation of a trans-periplasmic export channel (Koronakis *et al.,* 2000). This channel then permits secretion of HlyA to the external milieu, therefore eliminating the necessity for any distinct periplasmic intermediates.

1.5.1.2 Type 3 Secretion

The Type 3 secretion system (T3SS) was first discovered to play a pivotal role in the secretion of *Yersinia* outer proteins (Yops) from *Yersinia* spp. (Michiels *et al.,* 1990). Since its initial discovery, the T3SS machinery has been shown to be crucial for the pathogenesis of many bacterial species including *E. coli*, *S.* Typhimurium and *S. flexneri* (Cheng and Schneewind, 2000). The T3SS is defined by the prototypical example of *Yersinia* spp. where the T3SS depends on the transcription of 24 genes. Eleven of these are conserved genes and encode proteins that span the double membrane envelope of Gram-negative bacteria forming a protein conducting channel that enables protein secretion in one step. At the base of this protein conducting channel is an ATPase, which is oriented towards the bacterial cytoplasm and is extremely similar to the flagellar ATPase, FliI (Imada *et al.,* 2007). This machinery, the secretion substrates (*yop* genes) and the chaperones (*syc* genes) are encoded on a 70-kb

virulence plasmid. Contact with the host cells, for example macrophages, triggers *Yersinia* spp. to activate their T3SS and inject the Yop proteins directly into the host cell, hence, for this system, the term 'contact dependent' was coined (Cheng and Schneewind, 2000).

Similar to the T1SS, translocation of effector molecules via the T3SS navigates proteins across the IM in a Sec-independent fashion. Furthermore, these effector proteins are devoid of a recognisable signal sequence, as such the targeting mechanism by which proteins are directed to the T3SS still remains elusive but is thought to involve the first 20 amino acids (Mattei *et al.,* 2011). As a recurring theme with substrates exported via secretion systems, the functions of Type 3 virulence effectors differ remarkably within and between microorganisms and include cytotoxins, tyrosine phosphatases and kinases (Hueck, 1998; Pallen *et al.,* 2003).

1.5.1.3 Type 6 secretion

A relatively recently described secretion system is the Type 6 secretion system (T6SS) that, like others, constitute many proteins that assemble to form a transenvelope delivery apparatus. Interestingly, T6SS associated structures bear resemblance to bacteriophage structures such as the tail and syringe and, on this basis, the Type 6 assembled machinery is now thought of as an upside down bacteriophage (Kanamaru, 2009; Leiman *et al.,* 2009). Thus far, the T6SS is thought to secrete two types of proteins; namely the Hcp and VgrG proteins. Hcp is a small protein thought to hexamerise, forming nanotubes that allow the delivery of other T6SS associated substrates (Ballister *et al.,* 2008). A role for the VgrG proteins remains elusive, however VgrG proteins are thought to trimerise and puncture holes into bound membranes and allow passage of further T6SS substrates. Structural insights into the Hcp and VgrG proteins reveal similarities to bacteriophage tube and tail spike proteins, respectively (Leiman *et al.,* 2009). Although there is still a paucity of information regarding the substrates and

targets of the T6SS, many investigations have suggested that opposed to simply delivering effector proteins, the T6SS may be involved in divergent processes such as biofilm formation, sensing the environment or responding to stress (Bernard *et al.*, 2010). The most intriguing notion came from a recent study of *P. aeruginosa*, with the proposition that the T6SS is involved in the secretion of a bacterial toxin that kills neighbouring bacteria through targeting and hydrolysing peptidoglycan (Hood *et al.*, 2010; Russell *et al.*, 2011).

1.5.1.4 Type 4 secretion system

The first Type 4 secretion system (T4SS) identified was the VirB system of *Agrobacterium tumefaciens*. *A. tumefaciens* is a plant pathogen that causes crown gall disease in dicotyledonous plants (Escobar and Dandekar, 2003). The VirB T4SS translocation machinery consists of at least 12 different Vir proteins, some of which code for ATPases (VirB4, VirB11 and VirD4) and some of which code for proteins that form core complexes in the periplasm and/or membrane (VirB6, VirB7, VirB8, VirB9, and VirB10). These Vir proteins work co-ordinately to deliver T-DNA (transfer DNA which has oncogenic potential) directly into plant cells. The proteins VirD2 and VirE2 contain plant-active nuclear localisation signals (NLS) that allows targeting to the nucleus and subsequent integration of T-DNA into the plant cells genome (Yeo and Waksman, 2004).

Although the majority of effector molecules transported by the T4SS are targeted directly into the host cell, pertussis toxin is secreted into the external milieu. The pertussis toxin is a member of the AB_5 toxins whose action upon G proteins leads to the attenuation of the immune system (Burns, 2003). Furthermore, pertussis toxin, unlike other T4SS effector proteins, is synthesised with a signal sequence causing it to be directed and translocated across the IM via the Sec apparatus in an unfolded state before interacting with the T4SS

(Farizo *et al.,* 2000). However, the mechanism behind translocation across the OM remains obscure.

1.5.2 Sec-dependent translocation

1.5.2.1 Type 2 secretion

The Type 2 secretion system (T2SS) was initially identified by the study of the secretion of the starch-hydrolyzing lipoprotein pullulanase, from *Klebsiella oxytoca*. The model of the T2SS follows two well defined steps. The initial step, involving translocation across the IM, is mediated by the previously described Sec apparatus. Upon exposure to the periplasm, the signal peptide is removed and consequently the mature protein is released into the periplasmic space, ready to be secreted across the OM. Secretion across the OM employs a highly specific multiprotein complex, consisting of 12 to 15 proteins, mostly transcribed in a single operon (Sandkvist, 2001a). In *E. coli*, these have been designated GspA to GspO (Desvaux *et al.,* 2004). The proteins transcribed from the operon generate the accessory proteins responsible for assembling the multi-protein complex, the secreton (Planet *et al.,* 2001; Sauvonnet *et al.,* 2000). The proteins in this complex are assigned very specific roles. Protein C defines the substrate for export, Protein D has been suggested to oligomerise, forming the OM channel, and Protein E has been shown to act as the kinase promoting the assembly of the pilin-like subunits proteins G to K (Planet *et al.,* 2001). In unison, the activities of both the Sec apparatus and the multi-protein complex ensure the efficient translocation of these extracellular proteins. Since its initial description, the T2SS has been assigned to the secretion of a diverse array of proteins from a number of pathogenic bacteria, including cholera toxin from *V. cholerae* and exotoxin A, elastase and phospholipase C from *P. aeruginosa* (Sandkvist, 2001b).

1.5.2.2 Type 5 secretion: The Autotransporter pathway

In comparison to the complexity of the specialised secretion systems outlined above, the molecular machinery and mode of secretion that constitutes the autotransporter (AT) pathway is remarkably simple. Pohlner *et al.* (1987) were the first to describe the primary structure and functional moieties of an AT through studies on the IgA1 protease of *Neisseria gonorrhoeae*. The term 'Autotransporter' was coined through their observations that the AT contained a three-part architecture that was sufficient to drive secretion of the functional extracellular moiety, without the participation of dedicated accessory factors. This three part architecture is comprised of an N-terminal signal sequence that directs IM translocation via the Sec translocon, a central passenger domain that represents the secreted virulence determinant and a C-terminal domain that inserts into the OM in a β-barrel conformation, which is essential for translocation of the passenger domain to the cell surface. Since this initial description, several ATs, all of which exhibit structures reminiscent of the IgA1 protease, have been identified and now constitute the largest family of secreted proteins in Gram-negative bacteria (Fig. 1.9) (Dautin and Bernstein, 2007). Furthermore, the AT pathway has been sub-divided into three categories; the classical AT pathway (type Va), the two-partner secretion system (TPS or type Vb) and the trimeric AT adhesion system (TAA or type Vc) (Fig. 1.9). The two-partner secretion system differs from the classical AT pathway in that the signal peptide and 'passenger' effector domain (TpsA) are synthesised as a separate polypeptide to the C-terminal β-barrel domain (TpsB). Additionally, its C-terminal β-barrel bears most resemblance to the BamA superfamily of OMPs rather than the β-barrels found in the classical ATs, given that in addition to forming a pore in the OM, contains two periplasmically located POTRA domains (Fig. 1.9). The Tps pathway FHA/FhaC (filamentous hemaglutinin/filamentous hemaglutinin transporter) from *B. pertussis* represents

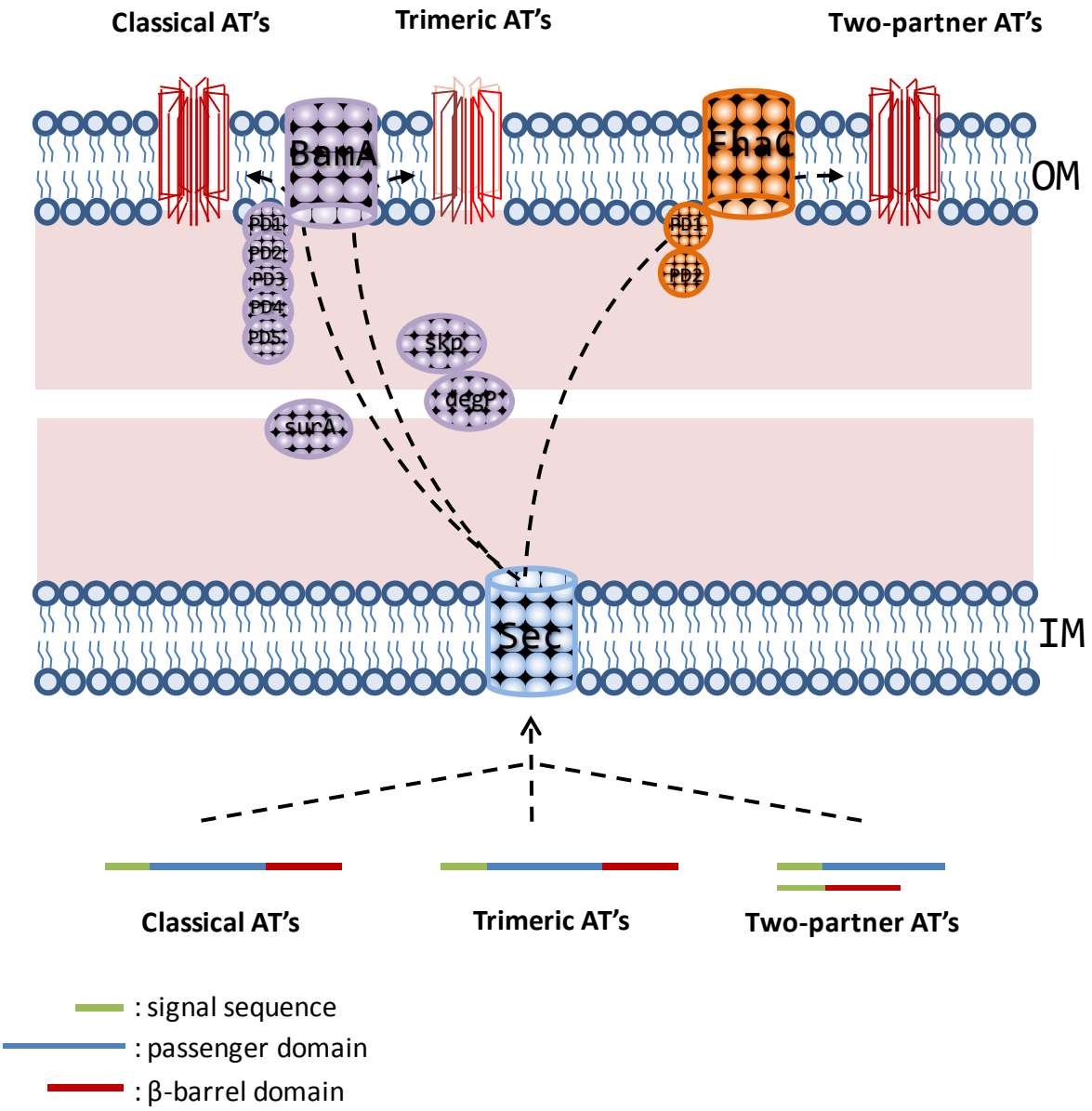


Figure 1.9. ***The autotransporter pathway.*** This figure illustrates the sub-families of the AT pathway; classical AT pathway (Va) is shown on the left, the trimeric AT adhesion system (Vc) is shown in the centre and the two-partner secretion system (Vb) is shown on the right. In the simplest form, the tripartite domain organisation is shown on the bottom; the N-terminal signal sequence (green) mediates inner membrane translocation via the Sec machinery, the passenger domain (blue) represents the effector protein and the C-terminal domain (red) forms a β-barrel in the OM necessary for passenger domain translocation to the cell surface. Shown are the periplasmic chaperones tha aid AT transit across the periplasm and deliver ATs to the depicted BAM complex. Figure adapted from (Henderson *et al.*, 2004).

the model Tps pathway where FhaC constitutes the transporter domain allowing FHA, the main adhesin of the pathogen, passage to the cell surface (Delattre *et al.,* 2010). The TAA family of ATs differ from the classical situation in that biogenesis and function depends on trimerisation (Cotter *et al.,* 2006). Furthermore, functions of the TAAs are largely limited to adhesive activity given that they remain anchored to their OM bound β-barrel.

1.6 Autotransporter biogenesis

Despite exhaustive research into the area of AT biogenesis, the mode of secretion at a molecular level has not been firmly established, with an increasing number of studies alluding to the necessity for accessory factors to facilitate protein secretion (Ruiz-Perez *et al.,* 2009). The biogenesis of ATs, with respect to how their three part architecture governs their translocation across the double membrane, is discussed below, with particular reference to recent literature surrounding regulation and interaction with accessory proteins.

1.6.1 Autotransporter gene regulation

In contrast to the vast amount of functional studies, the regulation of AT expression has received very little attention. The first report in 1996 was a study on the phase-variable OMP of *E. coli*, Antigen 43 (Ag43) (Owen *et al.,* 1996). Since this initial description, many studies have focussed attention on the regulation of Ag43, elucidating the nature of its phase-variability. Ag43 switches between an ‘on’ or ‘off’ phase through competitive interactions of OxyR and Dam-mediated methylation at the *agn43* promoter. In addition to the core -35 and -10 promoter elements, the *agn43* promoter contains three GATC motifs, directly downstream of the +1 transcription start site and an OxyR binding site that overlaps the -10 element (Wallecha *et al.,* 2002). Binding of OxyR to its recognition site represses transcription of *agn43* by blocking access of RNA polymerase to the core promoter. Dam-mediated

methylation of the adenine base within the GATC motifs, acts as an anti-repressor of transcription by abrogating OxyR binding and thus alleviating OxyR-dependent repression (Haagmans and van der Woude, 2000; Waldron *et al.,* 2002). Furthermore, methylation of the GATC motifs, independent of OxyR, is necessary for full activation of *agn43*, in an as yet uncharacterized mechanism (Wallecha *et al.,* 2002). It is this interplay between Dam methylation and OxyR binding that sets *agn43* expression as 'on' or 'off', respectively. However, the molecular basis of switching between the 'on' or 'off' phase still remains unclear. The only other relatively well characterised regulation of an AT concerns IcsA regulation. In *S. flexneri*, IcsA, is an essential virulence factor and is subject to control via a novel mechanism involving the first documented regulatory RNA, RnaG, from *S. flexneri*'s virulence plasmid (Giangrossi *et al.,* 2010). IcsA and RnaG are co-transcribed to generate both the IcsA mRNA and the small RNA (sRNA), RnaG. Subsequently, RnaG acts as a direct repressor of *icsA* transcription, primarily by targeting the 5' end of *icsA* messenger RNA and causing transcription termination resulting in an aborted *icsA* transcript of 100 nt (Giangrossi *et al.,* 2010). Although the expression of *agn43* and *icsA* are relatively well understood, the regulation of ATs on a broader scale has received very little attention, with most studies being restricted to phenomenological aspects of AT gene expression i.e. the change in accumulation of protein in various cellular compartments in response to environmental stimuli.

1.6.2 Inner membrane translocation

After transcription and translation of genes encoding ATs, the next step is translocation across the IM. The energy-driven translocation of ATs across the IM into the periplasm is mediated by the N-terminal signal sequence. All ATs are synthesised with a Sec-dependent signal sequence that exhibit a tripartite organisation; a positively charged N-domain, a hydrophobic H-domain and a C-domain which serves as the recognition site for signal peptidase (Hegde

and Bernstein, 2006). A subset of ATs possesses an unusually long signal sequence, now termed the extended signal peptide region (ESPR). The ESPR comprises an additional region comprising a charged and hydrophobic domain such that the ESPR organisation consists of five regions; N1 charged domain, H1 hydrophobic domain, N2, H2 and C-domain, whereby the primary structure of the N2, H2 and C-domains are reminiscent of the classical signal sequence seen in all Sec-dependent proteins. In contrast to the classical signal sequence, which demonstrates significant sequence variability, the ESPR is largely conserved (Desvaux *et al.,* 2007). Given the conservation of the ESPR, a cohort of studies investigated the notion that the ESPR imparts additional functions upon IM translocation. Initial studies, based on the supposition that signal peptides have the capacity to direct surface proteins to specific subcellular locations, were deemed erroneous given that polar localisation of specific ATs is independent of the ESPR (Charles *et al.,* 2001). Other studies suggested the ESPR might alter the mechanism of targeting to the Sec translocon; although these studies were subsequently discounted (Chevalier *et al.*, 2004). Although roles for the ESPR still remain contentious, recent investigations have led to the putative model that the ESPR slows the translocation of ATs across the IM to prevent accumulation of misfolded species in the periplasm (Desvaux *et al.*, 2007; Jong and Luirink, 2008; Peterson *et al.*, 2006).

1.6.3 Periplasmic transit

Upon exposure to the periplasm, the signal sequence is cleaved by signal peptidase, at the recognition sequence contained within its C-domain (Henderson *et al.,* 2004). Controversy has surrounded how ATs, in a largely unfolded state, avoid degradation by periplasmic proteases and remain in a structurally competent state for outer membrane (OM) translocation. Developments in our understanding of AT biogenesis came through observations that surface presentation of the AT IcsA were diminished in *E. coli* mutants

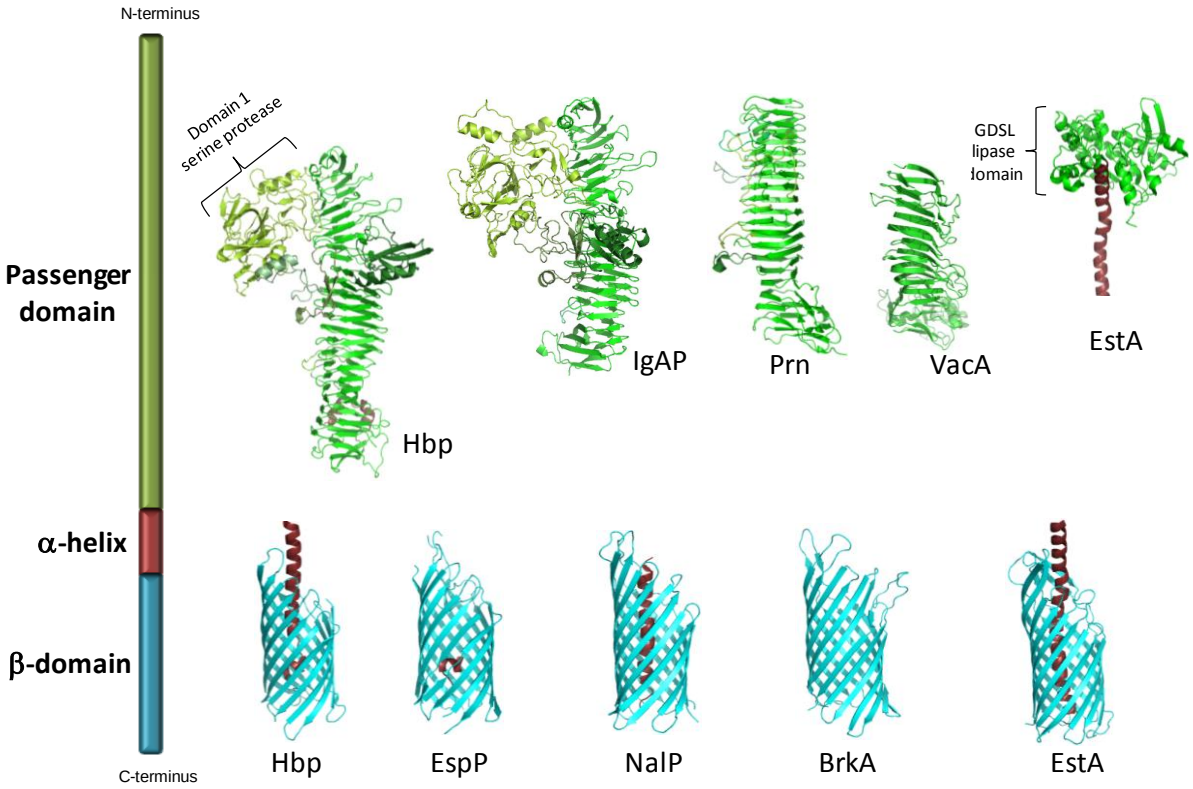


Figure 1.10. Classical AT structure. Depicted in this figure, from top to bottom; X-ray crystallography structures of all solved AT passenger domains, shown in green; Hbp: *E. coli* Hemoglobin protease (PDB entry 1WXR) (Otto *et al.*, 2005), Prn: *B. pertussis* Pertactin (PDB entry 1DAB) (Emsley *et al.*, 1996), VacA: *H. pylori* vacuolating toxin p55 fragment (PDB entry 2QV3) (Gangwer *et al.*, 2007), and IgaP: *H. influenzae* immunoglobulin A1 protease (PDB entry 3H09) (Johnson *et al.*, 2009). α-helical plugs within the β-barrel pore are shown in red. X-ray crystallography structures of all solved AT β-domains are shown in light blue; *E. coli* EspP (PDB entry 2QOM) (Barnard *et al.*, 2007), *N. meningitidis* NalP (PDB entry 1UYN) (Oomen *et al.*, 2004), *P. aeruginosa* EstA (PDB entry 3KVN) (van den Berg, 2010), *E. coli* Hbp (PDB entry 3AEH) (Tajima *et al.*, 2010) and *B. pertussis* BrkA (PDB entry 3QQ2) (Zhai *et al.*, 2011).

lacking the periplasmic chaperones SurA, Skp and DegP (Purdy *et al.,* 2007). These observations raised the question of whether there was a direct protein-protein interaction of ATs with these chaperones. Through both *in vivo* and *in vitro* techniques, later studies were able to demonstrate direct binding of SurA, Skp and DegP with the ATs EspP and Hbp (Ieva and Bernstein, 2009; Ruiz-Perez *et al.,* 2009; Sauri *et al.,* 2009), with the assumption that these interactions extend to many other ATs. Although these interactions have been established, the exact property that they confer on the participating ATs remains enigmatic.

1.6.4 Outer membrane translocation

1.6.4.1 Autotransporter β-domain incorporation into the OM; structure and function

Assisted by the aforementioned periplasmic chaperones, ATs are delivered to the OM and dock with the BAM complex. Thereafter, passenger domain translocation to the cell surface is dependent on the BamA component of the BAM complex. Recent crystallography studies have elucidated the β-domain structures of many ATs (Fig. 1.10) including NalP, EspP, BrkA, Hbp and EstA (Barnard *et al.,* 2007; Oomen *et al.,* 2004; Tajima *et al.,* 2010; van den Berg, 2010; Zhai *et al.,* 2011). Despite the sequence variation of the β-domains, the structures reveal a largely uniform architecture, where each β-domain, comprises twelve strands with interconnecting extracellular loops and periplasmic turns, adopts a β-barrel conformation within the OM (Fig. 1.10). An additional conserved structural feature is the presence of an α-helix within the barrel lumen. A multitude of studies including antibiotic sensitivity studies, molecular dynamics and electrophysiological experiments deduced that the role of the α-helix, together with certain extracellular loops of the β-barrel, functions in both plugging the pore, such that the permeability barrier of the OM is not compromised, and maintaining the width of the β-barrel pore (Clantin *et al.,* 2007; De *et al.,* 2008; Khalid and Sansom, 2006). As well

as highlighting important structural features of AT β-domains, structural and mutagenesis studies have provided insight into the function of the β-domain in the context of passenger domain translocation, dismissing the notion that the β-domain may act as a passive pore, with only subtle functional roles such as targeting to the OM. For example, mutagenesis of residues within strand 6 of both the BrkA and Tsh β-barrels significantly reduced passenger domain translocation to the cell surface illustrating the importance of a conserved hydrophobic cavity within the region of these β-barrels (Yen *et al.,* 2010; Zhai *et al.,* 2011). Consequently, this area of research is currently elucidating the precise molecular interactions that govern passenger domain translocation and establishing that the β-barrel does play an active part in this process.

1.6.4.2 Passenger domain translocation across the OM

After insertion and the correct assembly of the AT β-barrel in the OM, the passenger domain must traverse the OM to reach its extracellular target. Until recently, four models of passenger domain translocation across the OM have been discussed. These were the 'threading', 'hairpin', 'oligomeric' and the 'OMP 85' (now referred to as BamA) models. The 'threading' and the 'oligomeric' models have essentially been discounted by a large body of evidence and will not be discussed further.

1.6.4.2.1 Hairpin model

The classical hairpin model favours secretion of largely unfolded ATs through the lumen of the monomeric β-barrel. This model rationalises that the passenger domain is translocated in a C to N-terminal fashion, through the formation of a hairpin-like structure at the passenger domain to β-domain interface termed the Autochaperone (AC) domain. Experimentally, the engineering of disulphide bonds within the passenger domain has verified this model by

detection of a translocation intermediate whereby the C-terminal portion of the AT is susceptible to proteases/antibodies at the cell surface, whilst the N-terminal portion remains periplasmically orientated (Renn and Clark, 2011). The initial appearance of the extremely stable AC domain at the bacterial cell surface serves a functional role by driving the folding of the subsequent β-helical passenger domain, prior to its release (Junker et al., 2009; Soprova et al., 2010). However, based on the solved dimensions of the β-barrel lumen where the ellipsoidal cross section extends from 1.9 nm to 2.7 nm at its narrowest and widest, respectively, there are size limitations that bring this hairpin model into question. Evidence whereby disulphide bonded or glycosylated regions of the passenger domain are efficiently translocated to the cell surface (Lindenthal and Elsinghorst, 1999; Sherlock et al., 2006; Skillman et al., 2005) highlighted the need for alternative models to explain the contradictions between these observed structures and the inherently narrow β-barrel pore.

1.6.4.2.2 BamA model

Due to the intrinsic problems of the hairpin model, a large body of evidence has now been generated to support an essential role of BamA, a part of the BAM complex, in AT secretion. Using various experimental strategies, the intimate association of ATs with BamA, and the necessity for a functional BamA for AT secretion, have been established (Jain and Goldberg, 2007; Lehr et al., 2010; Ruiz-Perez et al., 2009; Sauri et al., 2009). Currently, these studies lack the molecular insights to substantiate a mechanism for AT biogenesis involving BamA. However, speculations based on these observations have theorised two predominant models. Firstly, a model integrating both the hairpin and BamA model proposes that the β-domain inserts into the pore of BamA whilst the β-barrel is held ‘open’ by BamA allowing the passenger domain to be translocated to the exterior of the cell. However, size constraints relating to the size of the BamA pore again cast doubt on the plausibility of this model. In

another model, the formation of a β-barrel is suspended whilst the β-domain merely docks with BamA and is held in a loosely folded conformation whereby the passenger domain favours the BamA pore as a translocation conduit, after which the BAM complex would disassemble triggering sealing of the AT β-domain and its folding into a stable barrel (Ieva *et al.*, 2011; Sauri *et al.*, 2009). Evidently, significant advances in our understanding of AT biogenesis have been made since their initial description in 1987, calling into question the 'auto' nature of their secretion. However, there is still a distinct lack of mechanistic detail regarding their OM translocation that future studies are still at liberty to resolve.

1.6.5 Passenger domain structure

Once the passenger domain has reached the outside of the bacterium, it has to fold into its native form in order to manifest its function. The passenger domain, which comprises the effector moiety of the AT polypeptide, often exceeds 100 kDa, although they do range from <20 to >400 kDa. Although there is no structural motif common to all ATs, crystal structures, bioinformatics, and biophysical analyses suggest that most passenger domains are variations on a common structural theme. Crystal structures solved for the *Bordetella pertussis* Pertactin (Emsley *et al.*, 1996) and the Haemoglobin protease, Hbp (Otto *et al.*, 2005) from *E. coli* show a characteristic right-handed β-helical stalk-like structure adjoined by loops of varying length and structure (Fig. 1.10). Intriguingly, the striking β-helical stalk of Pertactin and Hbp are two of only a dozen or so right-handed parallel β-helix protein structures solved; with the 24 turns of right-handed β-helix noted in the Hbp structure, compared with 16 in Pertactin, being the longest parallel β-helix known to date. Furthermore, a conserved domain at the extreme C-terminus of the passenger domain, named the autochaperone (AC) domain, appears to be decisive for the correct folding of the β-helical passenger domain on the cell surface (Fig. 1.10) (Oliver *et al.*, 2003; Velarde and Nataro, 2004). Consistent with the notion that

secretion proceeds in a C-to-N terminal direction, this model would assume that the AC domain appears on the cell surface first and initiates subsequent folding of the passenger domain. Additional studies have resolved the structures of the IgAP and VacA ATs, elucidating further structures that decorate the surface of the core β-helix. In the case of Hbp and IgAP, these structures constitute domains independent of the β-helical stem with distinct functional features. For example, Hbp, a member of the Serine Protease Autotransporters of the Enterobacteriaceae (SPATE) subfamily of ATs, has a large globular domain that is fixed at the N-terminal base of the β-helix where the signature serine protease motif lies (Domain 1) (Fig. 1.10). Sequence alignments of Hbp with the other known SPATEs (Pet, Sat, EspC, SepA and Pic) suggest high conservation within this region, with minor residue changes in and around the serine protease motif, assumed to confer substrate specificity on the functionally diverse SPATEs (Otto *et al.*, 2005). Additional solved domains consist of Domain 2, which resides at the base of the β-helix on the opposing face to Domain 1 (Fig. 1.10) and is capable of independent movement and folding (Johnson *et al.*, 2009). However, the revealed structure for EstA challenges this conservation where the characteristic β-helix is replaced by a large globular fold that is dominated by alpha-helices and loops (Fig. 1.10) (van den Berg, 2010).

1.6.6 Extracellular release of the passenger domain

Many ATs are released into the external environment following a proteolytic event, cleaving the passenger domain from its β-domain in the OM. However, there are exceptions whereby ATs remain attached to the β-domain (EstA) or are processed to remain in close association with their cognate β-barrels (Ag43) (Henderson *et al.*, 2004). For those ATs that are cleaved, the mechanisms of AT proteolysis are exceptionally diverse. Dedicated OM proteases are often employed for passenger domain release. In the case of IcsA from *Shigella*, an integral OM protease, IcsP, is responsible for specific cleavage of IcsA (Shere *et al.*, 1997).

Additionally, the AT, NalP, is responsible for the cleavage of associated ATs from *N. meningitidis* such as IgAP and App (van Ulsen *et al.,* 2003). However, cleavage of many ATs is known to undergo auto-proteolysis, due to either endogenous serine protease activity (Hendrixson *et al.,* 1997; Serruto *et al.,* 2003) or an interaction independent of their serine protease, involving residues within the β-barrel lumen (Dautin *et al.,* 2007; Dautin and Bernstein, 2011; Velarde and Nataro, 2004).

1.7 Autotransporter function

Given the prevalence and functional diversity of ATs, only the functions of a few ATs that are relevant to my work will be discussed. However, it is important to note that ATs encompass virulence factors that serve many diverse functions including cytotoxicity, enterotoxicity, invasion and adhesion. Work in our laboratory has focused mainly on the archetypal AT, Plasmid encoded toxin (Pet) from Enteroaggregative *E. coli* (EAEC) 042 and Antigen 43 (Ag43), a phase-variable AT found in many *E. coli* strains. Ag43 is one of the most abundant phase-variable OMPs in *E. coli*. Its presence on the cell surface mediates auto-aggregation, however its function and the phase-variable nature of its expression is not clearly defined (van der Woude and Henderson, 2008). In light of the fact that exhaustive studies have not been able to specify a role for Ag43, it is thought that this AT may play a role in the persistence of *E. coli*, both as a pathogen and a commensal.

1.7.1 Plasmid-encoded toxin

Pet is a 104 kDa protein encoded on the 65-MDa adherence-related plasmid of the prototypical EAEC strain 042. EAEC 042 represents an emerging pathogen that causes enteric and food-borne infectious diseases. EAEC pathogenesis comprises four major features; abundant adherence to the intestinal mucosa, formation of a thick mucus-containing biofilm,

induction of mucosal inflammation and elaboration of enterotoxins and cytotoxins (Harrington *et al.*, 2006). One of these toxins, Pet, is an archetypal SPATE which is cleaved at the bacterial cell surface and released into the extracellular milieu. To cause the characteristic toxic effects on eukaryotic cells, Pet must undergo endocytosis to reach its intracellular cytosolic target, Fodrin. Pet binds to the epithelial cell surface of the human intestinal mucosa, and is internalised by clathrin-coated vesicles (Navarro-Garcia *et al.*, 2007b). In a manner analogous to that of the well characterised AB_5 toxins, such as Cholera and Diptheria toxin, Pet is the first non-AB_5 toxin that exploits the vesicular trafficking pathways of the target cell in order to pass from the cell surface through the golgi apparatus and ER onto the cytosol (Navarro-Garcia *et al.*, 2007a). Pet utilises its serine protease activity to cleave the cytoskeletal protein fodrin within the cytosol (Canizalez-Roman and Navarro-Garcia, 2003), which consequently disrupts the organisation of the actin cytoskeleton leading to loss of actin stress fibers and release of focal contacts in HEp-2 and HT29/C1 cell monolayers (Navarro-Garcia *et al.*, 1999). These cytotoxic effects lead to cell rounding and detachment from the substratum.

1.8 Project aims

The basis for this project began by the identification of novel factors involved in AT biogenesis, using a previously generated transposon library of EAEC 042 (Sheikh *et al.*, 2002) to screen for the secretion of the AT, Pet. This screen identified over 70 transposon mutants that resulted in reduced or abolished secretion of Pet. This project is an extension of this preliminary data and the specific aims are as follows;

1. To confirm a role for the factors identified by the initial transposon screen in the biogenesis of Pet, using defined *E. coli* mutants.

2. To establish a role for the factors in Pet biogenesis, initially by outlining the stage that Pet biogenesis is affected.

3. To focus on a subset of the identified factors and gain insight into the mechanistic details of how these factors affect Pet biogenesis.

4. Extend observations seen with Pet to other ATs.

CHAPTER 2

Materials and Methods

2.1 Suppliers

Unless otherwise stated, culture medium components were purchased from Oxoid Limited and all chemicals were purchased from Sigma Aldrich. *Thermus aquaticus* (Taq) polymerase was supplied by Thermo Scientific. Phusion High Fidelity DNA polymerase, Calf alkaline phosphatase, T4 polynucleotide kinase and 100bp DNA ladder were purchased from New England Biolabs. T4 DNA ligase was supplied by Invitrogen. Unless otherwise specified, all enzymes were used according to manufacturer's instructions and used in the supplied buffers. dNTP mixes and Hyperladder I were purchased from Bioline. Oligonucleotides were purchased from either Alta Bioscience (University of Birmingham), or Eurogentec (Hampshire, UK). All radionucleotides were purchased from Perkin Elmer or MP Biomedicals and ribonucleotides and DNase I was supplied by Roche. *E. coli* RNA polymerase holoenzyme containing σ^{70} was purchased from Epicentre Technologies (Madison, WI). Purified CRP and Fis protein was kindly donated by David Lee and David Grainger (University of Birmingham).

2.2 Bacterial strains and culture conditions

2.2.1 Bacterial strains

The bacterial strains used in this study are listed in Table 2.1. All strains were stored as glycerol stocks at -80°C.

2.2.2 Bacterial culture conditions

Bacteria were routinely maintained using nutrient broth as follows; Luria-Bertani (LB) medium contained 1% (w/v) tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl (pH 7.2).

Table 2.1 Strains used in this study

Strain	***Relevant characteristic***	***Source***
Escherichia coli 042	Wild type prototype strain, Sm^r, Tet^r, Cm^r. Diarrhoeagenic in volunteers; expresses AAF/II; biofilm positive; harbours pAA2.	(Nataro et al., 1995)
Escherichia coli CFT073	Wild type Uropathogenic *Escherichia coli* (UPEC) isolate	(Mobley et al., 1990)
UPEC CFT073 *bamB::aph*	Wild type Uropathogenic *Escherichia coli* (UPEC) isolate, *bamB::aph*	This study
UPEC CFT073 *bamC::aph*	Wild type Uropathogenic *Escherichia coli* (UPEC) isolate, *bamC::aph*	This study
UPEC CFT073 *bamE::aph*	Wild type Uropathogenic *Escherichia coli* (UPEC) isolate, *bamE::aph*	This study
Escherichia coli HB101	F–, *thi*-1, *hsd*S20 (r_B^-, m_B^-), *sup*E44, *rec*A13, *ara*-14, *leu*B6, *pro*A2, *lac*Y1, *gal*K2, *rps*L20 (str^r), *xyl*-5, *mtl*-1	Promega
HB101 *bamB::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *bamB::aph*	This study
HB101 *bamC::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *bamC::aph*	This study
HB101 *bamE::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *bamE::aph*	This study
HB101 *crp::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *crp::aph*	This study
HB101 *degP::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *degP::aph*	This study
HB101 *fis::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *fis::aph*	This study
HB101 *ppiA::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *ppiA::aph*	This study
HB101 *ppiD::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *ppiD::aph*	This study
HB101 *skp::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *skp::aph*	This study
HB101 *surA::aph*	HB101 derivative with Datsenko & Wanner replacement of designated gene with kanamycin resistance cassette, *surA::aph*	This study

Table 2.1 Strains used in this study (continued)

Strain	***Relevant characteristic***	***Source***
Escherichia coli BW25113	*lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ*araBAD*$_{AH33}$ Δ *rhaBAD*$_{LD78}$	(Baba *et al.*, 2006)
BW25113 *ackA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ackA::aph*	(Baba *et al.*, 2006)
BW25113 *adiY::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *adiY::aph*	(Baba *et al.*, 2006)
BW25113 *ais::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ais::aph*	(Baba *et al.*, 2006)
BW25113 *bamB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *bamB::aph*	(Baba *et al.*, 2006)
BW25113 *bamC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *bamC::aph*	(Baba *et al.*, 2006)
BW25113 *bamE::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *bamE::aph*	(Baba *et al.*, 2006)
BW25113 *cadA ::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *cadA::aph*	(Baba *et al.*, 2006)
BW25113 *celB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *celb::aph*	(Baba *et al.*, 2006)
BW25113 *clpB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *clpB::aph*	(Baba *et al.*, 2006)
BW25113 *crp::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *crp::aph*	(Baba *et al.*, 2006)
BW25113 *crr::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *crr::aph*	(Baba *et al.*, 2006)
BW25113 *cspG::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *cspG::aph*	(Baba *et al.*, 2006)
BW25113 *cysK::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *cysK::aph*	(Baba *et al.*, 2006)
BW25113 *degP::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *degP::aph*	(Baba *et al.*, 2006)
BW25113 *dnaG::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *dnaG::aph*	(Baba *et al.*, 2006)

Table 2.1 Strains used in this study (continued)

Strain	***Relevant characteristic***	***Source***
BW25113 *fis::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *fis::aph*	(Baba *et al.*, 2006)
BW25113 *fkpA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *fkpA::aph*	(Baba *et al.*, 2006)
BW25113 *folX::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *folX::aph*	(Baba *et al.*, 2006)
BW25113 *lon::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *lon::aph*	(Baba *et al.*, 2006)
BW25113 *mhpC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *mhpC::aph*	(Baba *et al.*, 2006)
BW25113 *nei::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *nei::aph*	(Baba *et al.*, 2006)
BW25113 *nuoF::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *nuoF::aph*	(Baba *et al.*, 2006)
BW25113 *pal::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *pal::aph*	(Baba *et al.*, 2006)
BW25113 *pitA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *pitA::aph*	(Baba *et al.*, 2006)
BW25113 *ppiA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ppiA::aph*	(Baba *et al.*, 2006)
BW25113 *ppiD::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ppiD::aph*	(Baba *et al.*, 2006)
BW25113 *pqiB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *pqiB::aph*	(Baba *et al.*, 2006)
BW25113 *prmA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *prmA::aph*	(Baba *et al.*, 2006)
BW25113 *proP::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *proP::aph*	(Baba *et al.*, 2006)
BW25113 *proW::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *proW::aph*	(Baba *et al.*, 2006)
BW25113 *rfaJ::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *rfaJ::aph*	(Baba *et al.*, 2006)

Table 2.1 Strains used in this study (continued)

Strain	***Relevant characteristic***	***Source***
BW25113 *rlpA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *rlpA::aph*	(Baba *et al.*, 2006)
BW25113 *skp::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *skp::aph*	(Baba *et al.*, 2006)
BW25113 *slyB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *slyB::aph*	(Baba *et al.*, 2006)
BW25113 *surA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *surA::aph*	(Baba *et al.*, 2006)
BW25113 *thiC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *thiC::aph*	(Baba *et al.*, 2006)
BW25113 *uvrD::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *uvrD::aph*	(Baba *et al.*, 2006)
BW25113 *uxuR::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *uxuR::aph*	(Baba *et al.*, 2006)
BW25113 *yafO::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yafO::aph*	(Baba *et al.*, 2006)
BW25113 *ybbB::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ybbB::aph*	(Baba *et al.*, 2006)
BW25113 *ybeL::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ybeL::aph*	(Baba *et al.*, 2006)
BW25113 *ybhC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ybhC::aph*	(Baba *et al.*, 2006)
BW25113 *ydiM::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ydiM::aph*	(Baba *et al.*, 2006)
BW25113 *ydiN::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ydiN::aph*	(Baba *et al.*, 2006)
BW25113 *yeiC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yeiC::aph*	(Baba *et al.*, 2006)
BW25113 *yeiL::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yeiL::aph*	(Baba *et al.*, 2006)
BW25113 *yeiM::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yeiM::aph*	(Baba *et al.*, 2006)

Table 2.1 Strains used in this study (continued)

Strain	***Relevant characteristic***	***Source***
BW25113 *yfbV::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yfbV::aph*	(Baba *et al.*, 2006)
BW25113 *yfcI::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yfcI::aph*	(Baba *et al.*, 2006)
BW25113 *yfeK::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yfeK::aph*	(Baba *et al.*, 2006)
BW25113 *ygeV::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ygeV::aph*	(Baba *et al.*, 2006)
BW25113 *ygjI::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ygjI::aph*	(Baba *et al.*, 2006)
BW25113 *yhiL::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yhiL::aph*	(Baba *et al.*, 2006)
BW25113 *yhiN::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yhiN::aph*	(Baba *et al.*, 2006)
BW25113 *yiaG::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yiaG::aph*	(Baba *et al.*, 2006)
BW25113 *yihN::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yihN::aph*	(Baba *et al.*, 2006)
BW25113 *yihR::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yihR::aph*	(Baba *et al.*, 2006)
BW25113 *yjdL::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yjdL::aph*	(Baba *et al.*, 2006)
BW25113 *yjiC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yjiC::aph*	(Baba *et al.*, 2006)
BW25113 *ykgG::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ykgG::aph*	(Baba *et al.*, 2006)
BW25113 *ynbC::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *ynbC::aph*	(Baba *et al.*, 2006)
BW25113 *yohM::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yohM::aph*	(Baba *et al.*, 2006)
BW25113 *yphF::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yphF::aph*	(Baba *et al.*, 2006)

Table 2.1 Strains used in this study (continued).

Strain	***Relevant characteristic***	***Source***
BW25113 *yqiH::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yqiH::aph*	(Baba *et al.*, 2006)
BW25113 *yshA::aph*	BW25113 derivative with designated gene disrupted with a kanamycin resistance cassette, *yshA::aph*	(Baba *et al.*, 2006)

solid LB medium included 1.5% (w/v) bacteriological agar (Agar No. 1). To screen for β-galactosidase activity, 5% (w/v) MacConkey lactose agar was used. All media was autoclaved at 121°C for 20 min and allowed to cool to approximately 60°C, prior to use. When appropriate, media was supplemented with antibiotic(s) to a final concentration of 100 µg/ml ampicillin, 50 µg/ml kanamycin and 10 µg/ml tetracycline. Tetracycline was made up with 100% methanol and all others were made up with sterile distilled water (SDW). All antibiotics were made at 1000 x their working concentration, filtered through a 0.2 µM pore and stored at -20°C, until required. Unless otherwise specified, liquid cultures were incubated at 37°C with shaking at 180 rpm for aerobic growth, or in a static 37°C incubator for anaerobic growth. Growth on agar plates was routinely done in a static incubator, overnight at 37°C. Bacterial growth, in liquid cultures, was assessed by reading the optical density of samples at 600 nm (OD_{600}) using an Ultra Spec 2100 spectrophotometer.

2.3 Plasmids

All plasmids used in this study are listed in Table 2.2. Plasmid maps are shown in Figure 2.1.

2.4 DNA Manipulations

2.4.1 Polymerase chain reaction (PCR)

Standard PCR was carried out using Phusion High fidelity DNA polymerase (NEB) in the provided buffer, according to manufacturer's instructions. Between 1-5 ng of plasmid miniprep DNA or purified PCR product would be used as a template in a 50 µl reaction. Oligonucleotides were used at a final concentration of 0.5 µM and dNTP's were used at a final concentration of 200 µM each. Reactions were made up to 50 µl with SDW. Unless stated otherwise, thermal cycling was carried out as follows; 98°C/10 s, (98°C/10 s, X°C/30 s,

Table 2.2 Plasmids used in this study

Plasmids	***Relevant characteristic***	***Source***
pRW50	*oriV, lacZYA*, Tet^r. Broad-host range, low-copy-number *lacZ* expression vector used for cloning *Eco*RI-*Hin*dIII promoter fragments.	(Lodge *et al.*, 1992)
pCEFN1	3.9-kb PCR-derived fragment expressing Pet protein under the control of its native promoter cloned into pSPORT1, Amp^r.	(Eslava *et al.*, 1998)
pSPORT1	High copy number cloning plasmid, Amp^r.	Invitrogen
pSR	pBR322 derivative containing transcription terminator, λoop.	(Kolb *et al.*, 1995)
pDCRP	pBR322 derivative carrying *crp* gene.	(West *et al.*, 1993)
pDU9	Derivative of pDCRP with *crp* deleted.	(Bell *et al.*, 1990)
pDCRP HL159	pDCRP derivative with defective AR1.	(West *et al.*, 1993)
pDCRP KE101	pDCRP derivative with defective AR2.	(Rhodius *et al.*, 1997)
pDCRP HL159 KE101	pDCRP derivative with defective AR1 and AR2.	(Rhodius *et al.*, 1997)
pQuantagen(kx)	*phoA*, p15A origin of replication, promoter p*tac*; Amp^R.	Quantum, Appligene
pQMDSS*pet*	Pet signal sequence cloned into pQUANTagen(kx)	(Desvaux *et al.*, 2007)
pKD46	Red recombinase expression plasmids, Amp^r.	(Datsenko and Wanner, 2000)
pC02	*agn43* gene from MG1655 in pBADMycHisA-kan, Amp^r Km^r.	(Ulett *et al.*, 2007)
pACYC184/*pet*	PCR-derived fragment expressing Pet protein under the control of its native promoter cloned into pACYC184.	(Rossiter *et al.*, 2011b)

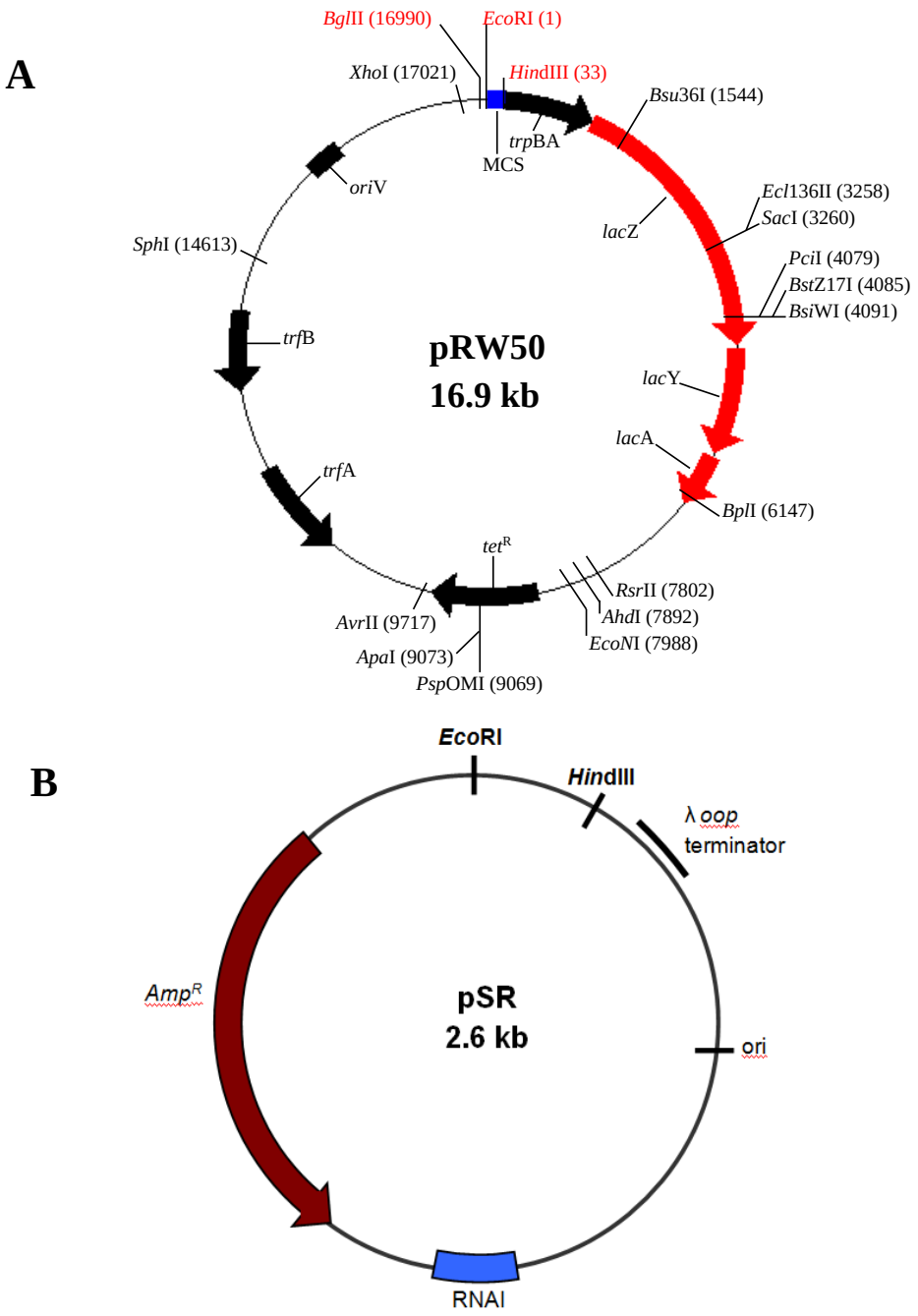


Figure 2.1. ***Plasmid maps.*** **A.** ***Map of the lac fusion vector, pRW50.*** *Eco*RI-*Hin*dIII fragments carrying the promoter of interest were cloned between the *Eco*RI and *Hin*dIII sites, contained within the multiple cloning site (MCS, a blue-filled box), to give *Eco*RI-promoter-*Hin*dIII::*lac* fusions, such that expression of the *lac* genes is under control of the cloned promoter. Also shown are the tetracycline resistance gene (tet^R), the origin of replication (*oriV)*, the plasmid replication genes *trfA* and *trfB*, and *trpBA*, which forms an operon with *lacZYA*. **B.** ***Map of pSR.*** *Eco*RI-*Hin*dIII promoter fragments are cloned upstream of the λ *oop* terminator and used for *in vitro* transcription reactions. Transcription initiates at the cloned promoter and terminates at the λ terminator, to produce a discrete transcript of a defined length. Also shown are the *RNA1* gene, which produces a control transcript during *in vitro* transcription, the *bla* gene, which encodes resitance to ampicillin (Amp^R), and the origin of replication (ori).

72°C/[1 min/kb]) x 30, 72°C/10 min. X refers to the annealing temperature. Annealing temperatures were used at 5°C's lower than that of the lowest given melting temperature of the primers used in the reaction. If not used immediately, PCR products were stored at -20°C.

Colony PCR was used routinely to check for chromosomal insertions or deletions and to check for the presence of cloned inserts within plasmids. Template DNA was prepared by re-suspending a colony in 30 µl of SDW and boiling at 100°C for 5 min. 2.5 µl of the sample was used in a total volume of 25 µl, made up with ReddyMix PCR (Thermo scientific).

2.4.2 Gel electrophoresis of DNA

All DNA samples above 500 bp were analysed using 0.8-1.5% agarose dissolved in 1x Tris-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.4). To stain the DNA, Syber green was added whilst the agarose was still molten, according to manufacturer's instructions. The gel was then poured onto a casting plate and left for approximately 10 min to set. When necessary, DNA samples were mixed in a 5:1 ratio with 6 x DNA loading dye (2.5% Ficoll 400, 11 mM EDTA, 3.3 mM Tris-HCl, 0.017% SDS 0.015% Bromophenol Blue, pH 8.0). Gels were run in 1 x TAE at 80-120 V (constant voltage) for 40-60 min, alongside Hyperladder I DNA marker (Bioline). Gels were visualised using a UV transilluminator with the gel-documentation system (Bio-Rad).

All DNA samples below 1 kb were analysed using 7.5% polyacrylamide gels. All gels were made using clean 1.5 mm thick Bio-Rad Minigel glass plates. Polyacrylamide gels (7.5% polyacrylamide, 4% glycerol and 1 x TBE [90 mM Tris, 90 mM Boric acid, 2 mM EDTA, pH 8.3]) were prepared using the appropriate amount of 30% (w/v) acrylamide, 0.8% (w/v) bis-acrylamide stock solution (Protogel, National diagnostics) and polymerised using 0.01 volumes of 10% (w/v) ammonium persulphate and 0.001 volumes TEMED (N,N,N′,N′-

Tetramethylethylenediamine). DNA samples were mixed in a 5:1 ratio with 6 x DNA loading dye, as above. Gels were run in 1 x TBE at 130 V (constant voltage) alongside 100 bp DNA ladder (NEB), for approximately 2 h, then post-stained in 50 mL de-ionised H_20 with the addition of 1 µl Sybr green (Applied Biosystems) and visualised using a UV transilluminator, as above. If gel bands were excised, they were either used immediately or stored at -20°C.

To analyse samples from DNA footprinting and electrophoretic mobility shift assay (EMSA) reactions, denaturing 6% polyacrylamide gels were made in large glass plates using the SequaGel sequencing system, as described in the manufacturer's instructions. Gels were pre-run in 1 x TBE buffer at 60 W (constant wattage) for approximately 1 h, prior to loading. Samples from footprinting reactions were in loading buffer and subjected to heating at 90°C for 2 min before loading. EMSA samples were loaded directly onto the gel in the absence of loading buffer, alongside a lane containing loading buffer, as a point of reference. Gels were run in 1 x TBE buffer at 60 W for 1-2.5 h, carefully removed from glass plates, then fixed for 10 min in a 10% methanol/10% acetic acid solution. Gels were placed on filter paper and dried under vacuum at 80°C for 30-40 min. Dried gels were exposed to a Fuji Imaging Phosphor screen for 1 to 16 h, the phosphor screen was scanned using a Bio-Rad Molecular Imager FX and images were analysed using QuantityOne software (BioRad).

2.4.3 Extraction and purification of DNA

2.4.3.1 Extraction of plasmid DNA

For small scale preparation, extraction of plasmid DNA from overnight cultures, supplemented with the appropriate antibiotic, was done using a QIAprep Spin Miniprep kit (QIAGEN), as per manufacturer's instructions. For cultures containing high copy number plasmids (10-30 copies per cell), 5 ml of culture was applied to one column, for cultures

containing low copy plasmids (1-5 copies per cell), 10 ml of culture was applied to one column. Plasmid DNA was eluted in 30-50 µl of provided elution buffer.

For the large scale preparation of concentrated plasmid DNA, for *in vitro* assays, caesium chloride density gradient ultracentrifugation was used. To do this, 500 ml LB was supplemented with the appropriate antibiotic(s), inoculated with an overnight culture of a strain harbouring the required plasmid and incubated at 37°C with aeration until an OD_{650} = 0.8-1.0 was reached. 50 mg chloramphenicol was then added to the culture to inhibit protein synthesis, whilst maintaining the ability of the bacterium to replicate plasmid DNA, and then incubated overnight. The following day, cells were harvested by centrifugation at 4,000 rpm for 30 min. The cell pellet was then re-suspended in 12 ml ice-cold TES buffer (50 mM NaCl, 60 mM Tris, 5 mM EDTA, pH 8.0). 2 ml of ice-cold 250 mM EDTA and 3 ml cold TES buffer containing 10 mg/ml lysozyme was then added to the cell suspension and left to incubate on ice for 15 min. To lyse the cells, 2 ml 10% SDS was added, mixed and incubated on ice for a further 10 min. 5 ml ice cold 5 M sodium chloride was added to the solution, mixed and incubated on ice overnight. To remove cell debris, the cell lysate was centrifuged for 2 h at 20,000 rpm at 4°C. The supernatant was then decanted into a fresh tube and allowed to warm to room temperature. DNA was precipitated by adding 15 ml isopropanol, incubating for 15 min at room temperature, and centrifuging for 15 min at 15,000 rpm at 20°C. The supernatant was then carefully discarded and the DNA pellet was dried under vacuum for 10 min to remove excess isopropanol, and re-suspended in 5 ml TES buffer by vigorous vortexing. 5.16 g caesium chloride, 200 µl ethidium bromide (10 mg/ml, Biorad) and 5.16 ml DNA solution were mixed, vortexed and transferred to a 10 ml Beckman quick-seal centrifuge tube. The tubes were sealed, using a soldering iron and then centrifuged for 22 h at 50,000 rpm at 20°C. After centrifugation, two ethidium bromide-stained bands were visible;

an upper band containing linear, chromosomal and nicked circular plasmid DNA, and a lower band containing closed circular plasmid DNA. The lower band was removed by insertion of a needle and a syringe slightly below the band and suction until the band was extracted. The resulting solution was extracted five times with an equal volume of water-saturated butanol to remove ethidium bromide, then dialysed overnight at 4°C against 4 l TE buffer (10 mM Tris, 0.5 mM EDTA, pH 8.0). The DNA solution was then extracted using phenol/chloroform and ethanol precipitated, as described in sections 2.4.3.2 and 2.4.3.3, and re-suspended in 600 µl TE buffer. The purity of the plasmid DNA was checked by agarose gel electrophoresis, and the DNA concentration was measured using an Eppendorf Biophotomoter.

2.4.3.2 Phenol/chloroform extraction of DNA

For the removal of contaminating proteins from enzymatic reactions, DNA samples were mixed in an equal volume of phenol/chloroform, vortexed for 10 s and then centrifuged for 2 min at 14,000 rpm to separate the aqueous and organic phase. The DNA-containing upper aqueous layer was then extracted carefully, without disturbing the interface between the two layers, and added to a fresh tube.

2.4.3.3 Ethanol precipitation of DNA

To precipitate DNA from solutions, 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of ice-cold 100% ethanol were added and the resulting solution was left to precipitate for 20 min at -20°C, or overnight at -80°C. The sample was then centrifuged at 14,000 rpm for 15 min, the supernatant was discarded and the pellet was washed with 1 ml 70% ice-cold ethanol. The sample was then centrifuged again at 14,000 rpm for 10 min, supernatant discarded and the pellet was vacuum dried for 10-15 min. The pellet was then re-suspended in either SDW or DNA loading dye, where appropriate.

2.4.3.4 Purification of PCR products

For purification of PCR products of 100 bp and more, a QIAquick PCR purification kit (QIAGEN) was used, according to manufacturer's instructions. PCR products were eluted from the column using 30-50 µl SDW, depending on concentration requirements.

2.4.3.5 Electroelution of DNA fragments from polyacrylamide gels

Excised gel bands, from polyacrylamide gels, were placed in 6.3 mm dialysis tubing with 200 µl 0.1 x TBE and clipped at both sides, ensuring no air bubbles were present. The dialysis bags were then placed in an electroelution tank filled with 0.1 x TBE and ran at 40 mA (constant amps) for 40-50 min. The buffer contained within the dialysis tubing was removed and put into a fresh tube, the tubing was then rinsed with 200 µl SDW and this was added to the collected buffer to make a total volume of 400 µl. The DNA contained within the buffer was then extracted, purified and concentrated by phenol/chloroform extraction and ethanol precipitation as described in 2.4.3.2 and 2.4.3.3, respectively and finally re-suspended in 20-25 µl SDW.

2.4.4 Restriction digestion of DNA

For cloning purposes, digestion of plasmid DNA was performed by adding 340 µl small-scale prepared plasmid DNA (section 2.4.3.1) to a total of 20 µl restriction enzyme (10 µl of each enzyme if double digestion needed), 40 µl of the appropriate 10 x restriction digest buffer as determined using the New England Biolabs double digest finder (http://www.neb.com/nebecomm/ DoubleDigestCalculator.asp) and 4 µl 100 x bovine serum albumin (BSA) where needed, giving a total reaction volume of 400 µl. For the preparation of DNA inserts, 10 µl purified PCR product was digested with a total of 6 µl of restriction enzymes (3µl of each enzyme for a double digest), 6 µl of the appropriate 10 x restriction

digest buffer and made up to a total reaction volume of 60 µl with SDW. All digests were incubated at 37°C for 3 h. For the preparation of plasmid DNA to be used in downstream cloning procedures, 4 µl calf alkaline phosphatase was added to the restriction digest mix after the 3 h incubation and incubated for a further 60 min at 37°C. This was done in order to remove terminal 5' phosphate groups, to prevent re-ligation of vector DNA. Digested DNA was purified either by phenol/chloroform extraction and ethanol precipitation or electroelution from 7.5% polyacrylamide excised gel bands, as in section 2.4.

2.4.5 DNA ligations

For ligation of digested insert DNA into digested vector, 10 µl purified insert DNA was added to 5 µl purified vector DNA, with 1 µl T4 DNA ligase (Invitrogen) and 4 µl supplied 5 x ligase buffer, in a total reaction volume of 20 µl. Ligations were incubated for 3 h at room temperature and the total ligation mix was then transformed into RLG221 chemically competent cells by the heat shock method, as described in section 2.4.8.

2.4.6 DNA sequencing

For sequencing of plasmid DNA, 5 µl low-copy plasmid, or 3 µl high-copy DNA was added to 3 µl of the appropriate primer (1µM) and made up to 10 µl with SDW. The samples were then loaded onto a ‘Plasmid to profile’ plate in Functional Genomics and Proteomics Laboratory, University of Birmingham and the sequencing results were ready for analysis the next day.

2.4.7 Cloning of promoter fragments

2.4.7.1 Preparation of promoter fragments by PCR

All the primers used for cloning and manipulation of promoter fragments in this study are listed in Table 2.3 and a list of all the promoter fragments are found in Table 2.4. In all cases, forward primers introduced a DNA site for the restriction enzyme of choice to bind and cut just upstream of the 5' upstream region of the desired promoter. The reverse primer introduced a DNA site for the restriction enzyme of choice to bind and cut just downstream of the 3' region of the promoter fragment. Once the promoter of interest had been amplified, it was treated as in section 2.4.7 for cloning into the appropriately digested vector. For more details on specific cloning strategies for cloning into pRW50 and pSR, please refer to sections 3, 4 and 5, respectively.

2.4.7.2 Introduction of point mutations into promoter fragments by mega primer PCR

Point mutations were introduced into *Eco*RI-*Hin*dIII promoter fragments, by mega primer PCR. All primers used to introduce point mutations into promoter fragments are shown in Table 2.3.

The PCR product was used as the ‘mega primer’ in the second round of PCR in conjunction with the upstream primer, D10520, which anneals approximately 80 bp upstream of the *Eco*RI site in pRW50. The PCR product of the correct size was then excised from a 7.5% polyacrylamide gel, electroeluted and purified by phenol/chloroform extraction and ethanol precipitation, as described in section 2.4. The purified fragments were *Eco*RI-*Hin*dIII digested, as described in section 2.4.4, and ligated into the required *Eco*RI-*Hin*dIII digested vector.

Table 2.3 Primers used in this study for amplification of *Eco*RI-*Hin*dIII fragments and for mutagenesis of promoter fragments

Name	***Sequence (5'→3')***	***Use***
Pet_F	ggggaattcttccatgtgacatctcag	Amplifies *pet* regulatory region, to create AER1 fragment with 5' *Eco*R1 site
Pet_R	gggaagcttgtattttattcatatattctctca actcatttattg	Amplifies *pet* regulatory region, to create AER1 fragment with 3' *Hin*dIII site
PetCRPIC[a]	gacatcgtatcattaacgasagcattgtsa cacattaacaatatag	Used in Mega primer PCR to introduce mutations within CRP binding site in AER1
PetFISIC[a]	gcaaaaaaaaacaaaaatsattaatttatg atsaatcagcattgg	Used in Mega primer PCR to introduce mutations within Fis I binding site in AER1
PetFISIIC	ggaacctgataatgctcctaaataggagc aaaaaaaaac	Used in Mega primer PCR to introduce mutations within Fis II binding site in AER1
Pet+1	cacacattaacaata**g**tagaactgttacttttt tacggg	Used in Mega primer PCR to introduce one base pair insertion in AER1 fragment
Pet-10	cattaacaatatagaactgttccttttacgg g	Used in Mega primer PCR to introduce mutations within -10 hexamer element in AER1
F_{45}CC(-40.5)	ggagagctccatggatatcggtacccgg ggatc	Used to introduce one helical turn in *F_{35}CC(-40.5)* promoter derivative
F_{40}CC(-40.5)	ggagagctccatggggtacccggggatc	Used to introduce half a helical turn in *F_{35}CC(-40.5)* promoter derivative
F_{35}CC(-40.5)	ggagaattcaaaaatgattaatttatgatca atcattc gagctcggtacccgggg	Used to introduce Fis binding site 35 bp upstream of *CC(-40.5)* promoter derivative
F_{30}CC(-40.5)	ggagagctcccggggatcaggta	Used to remove half a helical turn in *F_{35}CC(-40.5)* promoter derivative
F_{25}CC(-40.5)	ggagagctcgatcaggtaaatg	Used to remove one helical turn in *F_{35}CC(-40.5)* promoter derivative
Sat_F	ggggaattctgaacgatgttccatgcgaa cag	Amplifies *sat* regulatory region, to create SAT fragment with 5' *Eco*R1 site
Sat_R	gggaagcttgtattttattcatatattctctca actca tttattgaatgaacaaag	Amplifies *sat* regulatory region, to create SAT fragment with 3' *Hin*dIII site
SatCRPI	catcattaatgacaacattgtgacatatttac aatatag	Used in Mega primer PCR to introduce mutations within CRP binding site in SAT
SatFISI	caacccaccatcaattaacgatgaattag	Used in Mega primer PCR to introduce mutations within Fis binding site in SAT
Sat-10	caatatagaactgttccttttagcaagctg	Used in Mega primer PCR to introduce mutations within -10 hexamer element in AER1
Sat+1	cacacattaacaatagtagaactgttacttttt tacggg	Used in Mega primer PCR to introduce one base pair insertion in SAT fragment
D10520	ccctgcggtgcccctcaag	Anneals upstream of *Eco*RI site in pRW50.Used for sequencing of inserts.
D5431	acctgacgtctaagaaacc	Anneals upstream of *Eco*RI site in pSR.Used for sequencing of inserts.
Pet_1	ggggaattcgatggaagctcagatgga	Used for nested deletion analysis to successively remove 25 bp from AER1

[a]Bases labelled 'S' represent a G or a C. The G insertion in primer 'Pet+1' is in bold font.

Table 2.3 Primers used in this study for amplification of *Eco*RI-*Hin*dIII fragments and for mutagenesis of promoter fragments (continued)

Name	***Sequence (5'→ 3')***	***Use***
Pet_2	ggggaattcatatgagcaaaaaaaaac	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_3	ggggaattcgattaatttatgatcaatc	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_4	ggggaattctggtttacatgacatcg	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_5	ggggaattctaacgagagcattgtcac	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_6	ggggaattcaacaatatagaactgttact	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_7	ggggaattccgggatattagtaacaaac	Used for nested deletion analysis to successively remove 25 bp from AER1
Pet_8	ggggaattcctaatggttttaattcttaatc	Used for nested deletion analysis to successively remove 25 bp from AER1

Table 2.4 Promoter fragments used in this study

Promoter fragments	***Relevant characteristic***	***Source***
AER1	*Eco*RI-*Hin*dIII fragment carrying *pet* regulatory region	This study
AER1-crp1	Derivative of AER1 with a C to G and G to C substitution at positions -36 & -45, respectively	This study
AER1-fis1	Derivative of AER1 with a C to G and G to C substitution at positions -84 & -98, respectively	This study
AER1-fisII	Derivative of AER1 with a T to G and G to C substitution at positions -119 & -126, respectively	This study
AER1-10	Derivative of AER1 with an A to C substitution at position -11 in the promoter -10 element	This study
AER1+1	Derivative of AER1 with a G:C insertion between base pairs -21 and -22	This study
Pet_1	Derivative of AER1 with 17 bp removed from the 5' end	This study
Pet_2	Derivative of AER1 with 53 bp removed from the 5' end	This study
Pet_3	Derivative of AER1 with 77 bp removed from the 5' end	This study
Pet_4	Derivative of AER1 with 101 bp removed from the 5' end	This study
Pet_5	Derivative of AER1 with 124 bp removed from the 5' end	This study
Pet_6	Derivative of AER1 with 147 bp removed from the 5' end	This study
Pet_7	Derivative of AER1 with 173 bp removed from the 5' end	This study
Pet_8	Derivative of AER1 with 197 bp removed from the 5' end	This study
CC(-41.5)	Derivative of *E.coli melR* promoter with consensus CRP binding site centred at –41.5	(Gaston et al., 1990)
CC(-40.5)	CC(-41.5) with one base pair deleted between the DNA site for CRP and the -10 element	(West et al., 1993)
CC (-39.5)	CC(-41.5) with two base pairs deleted between the DNA site for CRP and the -10 element	(West et al., 1993)
FCC(-40.5)	CC(-40.5) derivative with the Fis I site, from the AER1 promoter fragment, inserted 35 base pairs upstream of the DNA site for CRP	This study
FCC(-39.5)	CC(-39.5) derivative with the Fis I site from AER1 promoter fragment inserted 35 base pairs upstream of the DNA site for CRP	This study
SAT	*Eco*RI-*Hin*dIII fragment carrying *sat* regulatory region	This study
SAT-crp1	Derivative of SAT with a C to G and G to C substitution at positions -36 & -45, respectively	This study
SAT-fis1	Derivative of SAT with a C to G and G to C substitution at positions -84 & -98, respectively	This study
SAT-10	Derivative of SAT with an A to C substitution at position -11 in the promoter -10 element	This study
SAT+1	Derivative of SAT with a G:C insertion between base pairs -21 and -22	This study

2.4.8 Bacterial transformations of DNA

2.4.8.1 Preparation of chemically competent bacterial cells

To prepare highly competent cells, for transformation by the heat shock method, 100 ml LB broth was inoculated with 1 ml overnight culture of the required strain. The culture was then grown at 37°C with aeration until an OD_{600} of 0.4-0.6 was reached. The culture was then placed on ice for 10 min before spinning at 4,000 rpm at 4°C for 10 min. The supernatant was then discarded and the cell pellet was re-suspended in 30 ml sterile ice-cold TFB1 (30 mM KOAc, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 15% glycerol) and incubated on ice for 90 min. The cell suspension was then centrifuged at 4,000 rpm for 10 min, supernatant discarded and the cell pellet was re-suspended in 2 ml sterile ice-cold TFB2 (0.1 M $CaCl_2$, 15% glycerol). The cells were either used straight away or aliquoted into 100 µl batches and stored at -80°C for future use.

2.4.8.2 Preparation of electro-competent bacterial cells

To prepare competent cells for transformation by electroporation, cultures were prepared and harvested as above, centrifuged at 4,000 rpm for 10 min at 4°C, the supernatant was discarded and the cell pellet was washed three times with 30 ml sterile ice-cold 10% glycerol. On the final spin, the cell pellet was re-suspended in 1 ml 10% glycerol and either used straight away or aliquoted as above and stored at -80°C.

2.4.8.3 Transformation by heat shock method

50 µl chemically competent cells were mixed with 1-3 µl plasmid DNA and incubated on ice for 30 min. Cells were then heat shocked at 42°C for 45 s and 0.5 ml LB medium was added to the heat shocked cells. These cells were then incubated at 37°C for 60 min with aeration.

Cells were then centrifuged at 4,000 rpm for 2 min, the pellet was re-suspended in approximately 100 µl of the supernatant, plated onto agar, supplemented with appropriate antibiotic(s), and incubated overnight at 37°C.

2.4.8.4 Transformation by electroporation

50 µl electro-competent cells were mixed with 1-3 µl plasmid DNA and added straight to an ice-cold 1 mm electroporation cuvette (Invitrogen). Cells were then electroporated at 1.8 kV using a gene pulser (BioRad). 1 ml LB medium was added straight to the electroporation cuvette and the solution was then transferred to a sterile tube and incubated at 37°C for 60 min with aeration. Cells were then centrifuged as above and transformants were selected on solid LB media supplemented with the appropriate antibiotic(s).

2.5 Screening KEIO collection mutants by PCR

The KEIO collection is a library of defined, single-gene deletions of all non-essential genes in *E.* coli K-12. This collection of mutants was made by replacing chromosomal genes with a 1.5 kb kanamycin resistance cassette flanked by FLP recognition target (FRT) sites (Baba *et al.*, 2006). Prior to use in this study, the veracity of the mutations were checked by colony PCR, as detailed in section 2.4.1. Primers that annealed to regions approximately 200 bp upstream and downstream of the first and last codon of the gene of interest, respectively, were used to analyse the size of the loci in question. All primers used for this screening are listed in Appendix 1. PCR products amplified from the candidate mutant and the parental strain, for comparison, were separated on a 1% agarose gel. In cases of a correct mutation, amplification of the candidate loci would result in a 1.5 kb PCR product, representing the size of the inserted kanamycin cassette, whereas amplification of the loci in the parental strain would result in a PCR product representing the size of the native gene. In cases where the size of the

native gene equalled that of the kanamycin resistant cassette, the primers flanking the gene of interest were used in conjunction with primers that annealed to inside the kanamycin cassette, and the presence of the correct size PCR product from the candidate loci was taken as a positive result. Although mutants from the KEIO collection represent insertions of the kanamycin cassette into the target locus, for simplicity, mutants are denoted as '*Δx*', where x represents the gene name. Note that, the result of this screen indicated that the KEIO library is 10-15% inaccurate and thus a thorough screening procedure is required to ensure the veracity of mutants before used for further work.

2.6 Constructions of mutants using the λ-RED system

For construction of mutants containing chromosomal insertions, a method based on the λ-red system was used (Datsenko and Wanner, 2000). Briefly, the allele containing a kanamycin insertion was amplified from the BW25113 strain of interest from the KEIO collection (Baba *et al.,* 2006). The PCR product was amplified using the primers listed in the Appendix 1, which were designed to anneal approximately 200 bp upstream and downstream of the allele of interest, resulting in ~ 400 bp regions of homology to the chromosomal target. The PCR product was electroporated into the desired strain, containing the pKD46 plasmid which harbours the inducible red recombinase system (Datsenko and Wanner, 2000). Kanamycin resistant colonies were screened by two independent colony PCR's for verification of targeted chromosomal gene disruption. For screening of candidate colonies, primers listed in the Appendix 2 were used that annealed to the kanamycin cassette in conjunction with primers that annealed 100 bp upstream of the inserted locus.

2.7 Protein preparation, detection and analysis

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

All SDS-PAGE protein separation was carried out using 12.5% polyacrylamide gels, prepared using Bio-Rad Minigel glass frames of 1.5 mm thickness. Each resolving gel contained the appropriate amount of 30% (w/v) acrylamide, 0.8% (w/v) bis-acrylamide stock solution (ProtoGel, National Diagnostics), 375 mM Tris-HCl (pH 8.8) and 0.1% (w/v) SDS. 4% stacking gels were made with the appropriate amount of acrylamide solution above, 125 mM Tris-HCl (pH 6.8) and 0.1% (w/v) SDS. Polymerisation of the acrylamide was achieved by adding 0.01 volumes of 10% (w/v) ammonium persulphate and 0.001 volumes TEMED (N,N,N′,N′-Tetramethylethylenediamine) to the resolving and stacking gel. Protein samples were mixed in equal volumes of 2 x Laemmli stock solution (Invitrogen). The gels were electrophoresed in 1 x SDS-PAGE running buffer (25 mM Tris-HCl, 0.1% SDS, 190 mM glycine) at 120 V (constant voltage) for 1.5-2 h.

2.7.2 Protein detection

For detection of all separated proteins, 0.25% (w/v) Coomassie Blue-R250 stain (BDH Laboratory Supplies) was dissolved in 10% (v/v) glacial acetic acid and 50% (v/v) methanol. For staining, gels were placed in approximately 25 ml of Coomassie Blue and incubated with shaking at room temperature for 1 h. Gels were then placed in destain (10% glacial acetic acid, 50% methanol) for approximately 1 h or until proteins of interest were visible.

Western immunoblotting was carried out for specific detection of proteins using purified antibodies, raised against the protein of interest. Once protein samples had been separated by SDS-PAGE, gels were placed in 1 x transfer buffer (25 mM Tris, 192 mM Glycine, 20%

Methanol) for 5 min. Nitrocellulose membrane (GE Healthcare) was soaked in 1 x TB for 5 min and placed on top of the gel within a western blot apparatus from Bio-Rad, according to manufacturer's instructions. Proteins were then transferred to the nitrocellulose membrane for 1 h at 100 V (constant voltage). Following transfer, the membrane was placed in 1 x PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4), 0.05% Tween-20 for 5 min. The membrane was then placed in 25 ml blotting solution (5% semi skimmed milk powder, dissolved in 1 x PBS-0.05% Tween-20) and incubated at room temperature for 1 h, with shaking. The membrane was then washed 3 times with 25 ml 1 x PBS, 0.05% Tween-20 for 10 min, then incubated with 25 ml primary antibody, at the required dilution, dissolved in 1 x PBS, 0.05% Tween-20 for 3.5 h at room temperature, with shaking. The membrane was then washed 3 times with 25 ml 1 x PBS, 0.05% Tween-20 for 10 min, then placed in secondary antibody (alkaline phosphatase-labelled goat anti-rabbit IgG antibodies [Sigma Aldrich, UK]), diluted 1:10,000 in 1 x PBS, 0.05% Tween-20. The membrane was then washed 3 times with 25 ml 1 x PBS, 0.05% Tween-20 for 10 min. Antigen–antibody reactions were visualized using 1 ml alkaline phosphatase substrate NBT/BCIP, supplied by Sigma Aldrich.

2.7.3 Inner membrane translocation assay

To determine the ability of the Pet signal sequence (Petss) to mediate translocation across the IM, the Petss was fused to alkaline phosphatase (PhoA) and the subsequent periplasmic PhoA activity was measured. To do this, the desired strains were transformed with the plasmids pQMDSS*pet* and pQuantagen(kx), containing PhoA with and without Pet signal sequence, respectively (Desvaux *et al.*, 2007). To determine the level of translocation across the IM, liquid assays of PhoA activity were performed. Cultures were grown overnight at 37°C in LB broth, supplemented with ampicillin. The following day, 20 ml LB broth was inoculated with

the overnight cultures at a 1:100 dilution and incubated at 37°C till an OD_{600}=3.0 was reached. IPTG was then added to a final concentration of 1 mM and the samples were incubated for a further 2 h at 37 °C. 1 ml volumes were then removed and the OD_{595} recorded. To assay PhoA activity, 1 ml volumes were pelleted and re-suspended in 800 µl SDW. Cells were permeabilised by adding 20 µL 0.1% sodium dodecyl sulphate (SDS) and 20 µL chloroform and vortexed for 1 min. To begin the reaction, 100 µL 10 mg/ml of para-nitrophenolphosphate (dissolved in 1 M Tris-HCl, pH 9.0) was added, and incubated at 37°C until a straw-yellow colour had developed or 1 h had elapsed. The reaction was stopped by adding 100 µl 10 M NaOH. Specific activity was calculated according to the following formula: 1,000 x (A_{420}/[reaction time (min) x volume of cells (ml) x A_{595}])

2.7.4 Preparation of whole cell extracts

To analyse proteins from whole cell extracts, whole-cell lysate of the desired strain was prepared by spinning down 1 ml of an overnight culture, de-canting the supernatant and re-suspending the cell pellet in 100 µl Laemmli buffer before boiling the resulting solution for 5 min at 90˚C. Samples were separated by SDS-PAGE and visualised using the desired technique.

2.7.5 Preparation of supernatant proteins

To prepare culture supernatant fractions, strains were grown overnight at 37°C in 5 ml LB broth, supplemented with the appropriate antibiotic. The following day, overnight cultures were added to 20 ml fresh LB medium at a 1:20 dilution. The new cultures were then incubated at 37°C, with shaking, and grown until OD_{600}=1.0. Cultures were normalised, cells were pelleted by centrifugation and the resultant supernatants were filtered through a 0.22 µm filter and the proteins in the supernatant were precipitated with 10% trichloroacetic acid (v/v).

Samples were incubated on ice for 1 h or at 4°C overnight, then centrifuged at 14,000 x g for 45 min and washed with 100% methanol. The pellets were collected by centrifugation at 14,000 x g for 15 min, dried, and re-suspended in 40-100 µL 2 x Laemmli buffer (Invitrogen). Samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualised using the desired technique.

2.7.6 Release of the α-domain of Ag43 and immunodetection.

For the analysis of Ag43 biogenesis, harvested cells were heat treated to release Ag43 from the surface of the cell by breaking the non-covalent interactions between the alpha-domain of Ag43 and the OM. To do this, cells from overnight LB cultures were normalised, harvested by centrifugation, washed in 0.9% (w/v) NaCl, and re-suspended in 1 ml of heat release buffer (75 mM NaCl–0.5 mM Tris, pH 7.4). The alpha-domain of Ag43 was released from the surface of the cells by heating at 60°C for 3 min. The cells were immediately removed by centrifugation, and the resultant protein in the supernatant was precipitated overnight at 4°C with 10% (w/v) TCA. The pellet was washed with 1 ml of 100% methanol and dried. Samples were then subjected to SDS-PAGE followed by western immunoblotting for detection of Ag43, as described in sections 2.7.1 and 2.7.2, respectively. Primary antibody, kindly donated by Mark Schembri, was used at a 1:1000 dilution.

2.7.7 Ag43 mediated cell-cell aggregation

Overnight cultures (containing 10 ml LB broth supplemented with 0.2% L-arabinose) harbouring strains carrying the pC02 plasmid were normalised to the same OD_{600} and transferred to 15 ml sterile falcon tubes. Prior to initiating the experiment, all cultures were vigorously shaken for 5 s and left in a rack with the lid off. Every 20 minutes, 100 µl of each sample was taken from each tube, approximately 1 cm from the top, and added to a 96-well

polypropylene micro-titre plate. The OD_{600} was then measured using a Bio-systems plate reader (Amersham Biosciences, UK).

2.8 *In vivo* promoter analysis

2.8.1 β-galactosidase assays

β-galactosidase assays were carried out in order to measure the activity of promoter fragments fused to the *lacZ* gene in pRW50, under various conditions and in different genetic backgrounds (Miller, 1972). 5 ml LB broth, supplemented with appropriate antibiotics, was inoculated with a freshly transformed colony, containing pRW50 with the promoter of interest, and incubated overnight at 37°C. The following day, 5 ml LB broth was inoculated with the overnight culture at a 1:100 dilution for measurements of β-galactosidase activity taken during aerobic growth or a 1:50 dilution for measurements taken during anaerobic growth. Once cells had reached the required OD_{650} = 0.2, for aerobic cultures and 0.5, for anaerobic cultures, they were immediately placed on ice and the exact OD_{650} reading was recorded. To lyse cells, 2 ml samples were added to 30 µl toluene and 30 µl 0.1% sodium deoxycholate and incubated for 30 min at 37°C, with aeration. For measurements of β-galactosidase activity, 100 µl of cell lysate was added to 2 ml Z buffer (0.75 g/l KCl, 0.25 g/l $MgSO_4.7H_2O$, 8.53 g/l Na_2HPO_4, 4.87 g/l $NaH_2PO_4.2H_20$, pH 7.0). Note that, β-mercaptoethanol was added to Z buffer prior to use (0.027 ml/l). The assay was initiated by the addition of 0.5 ml o-nitrophenyl-β-D-galactopyranoside (ONPG) solution, dissolved in Z buffer (3.92 g/l). Once the solution had changed from a clear colour to a straw yellow colour, or 1 h elapsed, 1 ml 1 M sodium carbonate was added to stop the reaction and the exact reaction time recorded. The absorbance at 420 nm of the resultant solution was measured and the calculation of β-galactosidase activity was calculated as follows:

$$\beta\text{-galactosidase activity} = \frac{1000 \times 2.5 \times 3.6 \times OD_{420nm}}{OD_{650nm} \times 4.5 \times t \times v} \text{ nmol/min/mg bacterial mass}$$

Where: 2.5 = factor for conversion of OD_{650} into bacterial mass, based on OD_{650} of 1.0 being equivalent to 0.4 mg/ml bacteria (dry weight).

3.6 = final assay volume (ml)

1000/4.5 = factor for conversion of OD_{420} into nmol o-nitrophenyl (ONP), based on 1 nmol ml^{-1} ONP having an OD_{420} of 0.0045

t = incubation time (min)

v = volume of lysate added (in ml)

Each experiment was done in triplicate and the mean and standard deviation calculated accordingly.

2.9 *In vitro* promoter analysis

2.9.1 Electrophoretic mobility shift assay (EMSA)

For preparation of promoter fragments to be used in EMSA, vectors containing cloned *Eco*RI-*Hin*dIII promoter fragments were digested with *Eco*RI and *Hin*dIII and treated with calf intestinal phosphatase, as described in section 2.4.4. Promoter fragments were then purified and concentrated by electroelution from a 7.5% polyacrylamide gel, phenol/chloroform extraction and ethanol precipitation, as described in section 2.4. Promoter fragments were then end-labelled with [γ-^{32}P]-ATP using T4 polynucleotide kinase (PNK), which catalyses the transfer of the γ-phosphate from ATP to 5' terminal hydroxyl groups in DNA. To do this, 16 µl of purified promoter fragment was added to 1 µl PNK, 2 µl supplied PNK buffer and

1.5 µl [γ-^{32}P]-ATP giving a total reaction volume of 20 µl. Samples were then incubated at 37°C for 30 min. Unincorporated [γ-^{32}P]-ATP was then removed by passing the sample down a Sephadex G-50 column. These columns were made by the addition of 400 µl 50% Sephadex-G50 suspension (5 g sephadex G-50 [Pharmacia Biotech] autoclaved in 100 ml TE buffer, washed three times in 150 ml TE buffer and finally suspended in 50 ml TE buffer) onto a Micro Bio-Spin column (Bio-Rad) and centrifugation at 3,500 rpm for 2 min. The column was then placed into a fresh collection tube, and the 20 µl labelling reaction was loaded carefully onto the column and centrifuged for a further 2 min at 3,500 rpm.

For analysis of CRP and Fis binding to promoter fragments, purified CRP and Fis protein was used, which was kindly donated by David Lee and David Grainger and purified as described by (Ghosaini *et al.*, 1988) and (Pan *et al.*, 1996), respectively. CRP and Fis were diluted to 10 x the desired final concentration in protein dilution mix (1 x HEPES buffer [10 x HEPES; 200 mM HEPES (pH8.0), 50 mM $MgCl_2$, 500 mM potassium glutamate, 10 mM dithiothreitol], 1 mg/ml bovine serum albumin [BSA]). The total reaction volume was 10 µl, containing 1 x HEPES, 0.2 mM cAMP, 0.5 mg/ml BSA, 25 ng/µl herring sperm DNA, 0.2-0.4 [γ ^{32}P]-labelled *Eco*RI-*Hin*dIII promoter fragment, 5% glycerol. Samples were incubated for 30 min at 37°C and loaded directly onto a 6% polyacrylamide gel containing 0.2 mM cAMP, which had been pre-run at 166 V for 30-60 min. The total 10 µl reaction was loaded onto the gel alongside 2 µl of loading dye, as a point of reference. The samples were then separated by electrophoresis at 166 V for approximately 2 h in 0.5 x TBE. The gels were then removed from the glass plates and fixed in 10% methanol/10% glacial acetic acid for 10-20 min, placed onto filter paper and dried under vacuum for 20-30 min. The gels were then exposed to a Fuji Imaging Phosphor screen for 1-3 h or overnight. The phosphor screen was scanned using a

Bio-Rad Molecular Imager FX and images were analysed using QuantityOne software (BioRad).

2.9.2 *In vitro* DNA footprinting

2.9.2.1 Preparation of radio-labelled promoter fragments for footprinting

*Aat*II-*Hin*dIII promoter fragments were digested from large-scale prepared (section 2.4.3.1) pSR and used for DNA footprinting reactions. To ensure labelling of only one strand of the promoter fragment, approximately 100 µg of plasmid DNA was incubated with 8 µl *Hin*dIII and 12 µl of supplied buffer 2 (NEB) in a total reaction volume of 120 µl for 3 h at 37°C, then treated with 5 µl CIP for 1 h, to remove 5' hydroxyl groups. The digested plasmid was then purified by phenol/chloroform extraction and ethanol precipitation and the resulting 50 µl of purified DNA was digested with 8 µl *Aat*II in a total of 80 µl 1 x buffer 4 (NEB). After 3 h digestion at 37°C, the reaction was loaded onto a 7.5% polyacrylamide and the *Aat*II-*Hin*dIII digested fragment was purified by electroelution followed by phenol/chloroform extraction and ethanol precipitation. Finally, the fragments were re-suspended in 50 µl TE and checked for purity and concentration on a 7.5% polyacrylamide gel. The *Aat*II-*Hin*dIII promoter fragments were then end-labelled with [γ^{32}P] on the *Hin*dIII end as described in section 2.9.1.

2.9.2.2 Preparation of G+A ladder

*Aat*II-*Hin*dIII radio-labelled promoter fragments, as prepared above, were treated with formic acid followed by piperidine cleavage, to generate Maxam-Gilbert G+A ladders, for use in calibrating DNA footprinting gels. To do this, 3-4 µl labelled promoter fragments were made up to a final volume of 12 µl in SDW, mixed with 50 µl 100% formic acid and incubated at room temperature for 1.5 min. The reaction was stopped by the addition of 700 µl 100% ice-cold ethanol and 200 µl 0.3 M sodium acetate (pH 7.0) and DNA was precipitated as in

section 2.4.3.3. The DNA pellet was re-suspended in 100 µl piperidine and this cleavage reaction was incubated at 90°C for 30 min. Samples were then ethanol precipitated and re-suspended in 20 µl denaturing gel loading buffer (40% deionised formamide; 5 M urea, 5 mM sodium hydroxide, 1 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanole FF). Prior to loading on a gel, the G+A ladder was heated at 90°C for 2 min and 0.5-2 µl loaded onto a footprinting gel.

2.9.2.3 DNase I footprinting

DNase I footprinting was carried out in order to map the positions of CRP and Fis binding at various promoters. For each reaction, 0.1-0.3 end-labelled *Aat*II-*Hin*dIII promoter fragments were mixed with a final concentration of 0-400 nM CRP and 0-1600 nM Fis, as indicated, in a final volume of 20 µl containing 1 x HEPES buffer, 0.2 mM cAMP, 0.5 mg/ml BSA and 30 µg/ml Herring sperm DNA. Proteins were diluted to 10 x required concentration in protein dilution mix (1 x HEPES, 1 mg/ml BSA), prior to addition to the reaction mixture. Samples were incubated at 37°C for 20 min, to allow proteins to bind DNA, prior to DNase I treatment. Following incubation, 2 µl of a 1:1320 dilution of DNase I (10 U/µl, Roche Applied Science) in SDW was added to each reaction and incubated at 37°C for 40 s. To stop the reaction, 200 µl of DNase I stop solution (0.3 M sodium acetate [pH 7.0], 10 mM EDTA) was added. DNA samples were purified and concentrated by phenol/chloroform extraction and ethanol precipitation as described in section 2.4 and re-suspended in 7 µl denaturing gel loading buffer. All samples were then heated to 90°C for 2 min, 3.5 µl was loaded onto a denaturing 6 % acrylamide sequencing gel alongside 'G+A' ladder and separated by electrophoresis in 1 x TBE for approximately 1.5-2 h. Gels were then dried and analysed as described in section 2.9.1.

2.9.2.4 Potassium permanganate footprinting

Potassium permanganate footprinting was carried out in order to examine open complex formation at various promoters. Potassium permanganate modifies unpaired thymine residues in single stranded regions of DNA as found in open complexes at transcription start sites. Subsequent treatment with piperidine cleaves the DNA backbone on the 3' side of the modified thymine residues. For each reaction, 0.2-0.4 µl end-labelled *Aat*II-*Hin*dIII promoter fragments were mixed with a final concentration of 50 nM *E. coli* RNA polymerase (RNAP) σ^{70} holoenzyme, 50 nM CRP and 0-1600 nM Fis, as indicated, in a total volume of 20 µl 1 x HEPES, 0.2 mM cAMP and 0.5 mg/ml BSA. Proteins were diluted to 10 x the required concentration in protein dilution mix (1 x HEPES, 1 mg/ml BSA), prior to addition to the reaction mixture. Samples were then incubated at 37°C for 30 min, to allow proteins to bind DNA. Following incubation, 1 µl of freshly prepared 200 mM potassium permanganate was added to the reaction, incubated for 4 min at 37°C and 50 µl potassium permanganate stop solution (3 M ammonium acetate, 0.1 mM EDTA, 1.5 M β-mercaptoethanol) was then added to stop the reaction. 150 µl SDW was added to the reactions to make the volume up to 200 µl and DNA was then purified by phenol/chloroform extraction and ethanol precipitation, with re-suspension of the final DNA pellet in 40 µl 1 M piperidine. Samples were incubated with piperidine for 90°C for 30 min, to allow cleavage of the DNA to occur. Again, following incubation, the DNA was purified by phenol/chloroform extraction and ethanol precipitation, with the final dried DNA pellet being re-suspended in denaturing gel loading buffer. 3.5 µl of each sample was then heated to 90°C for 2 min and loaded onto a 6% denaturing polyacrylamide sequencing gel alongside 'G+A' ladder and separated by electrophoreses for approximately 1.5-2 h. Gels were then dried and analysed as described in section 2.9.1.

2.9.3 *In vitro* transcription assays

Using purified RNAP, CRP and Fis, multiple-round *in vitro* transcription assays were performed, using caesium chloride preparations of pSR plasmid, carrying various *Eco*RI-*Hin*dIII promoter fragments, as template. CRP, Fis and RNAP were diluted to 10 x the desired final concentration, as indicated, in protein dilution mix (1 x transcription buffer [10 x transcription buffer; 400 mM Tris-Cl [pH 7.9], 100 mM MgCl2, 500 mM KCl, 1mM DTT], 1mg/ml BSA), and added to a final volume of 20 µl containing 1 x transcription buffer, 0.1 mg/ml BSA, 500 µM ATP, 500 µM CTP, 500 µM GTP, 50 µM UTP, 5 µCi [α^{32}P]-UTP, 0.2 mM cAMP and 80 nM pSR plasmid DNA, carrying the promoter of interest. All samples were incubated at 37°C for 30 min, prior to adding RNAP, to allow proteins to bind DNA. RNAP was then added at a final concentration of 50 nM and incubated at 37°C for 20 min before adding 10 µl denaturing loading buffer (40% deionised formamide, 5 M urea, 5 mM sodium hydroxide, 1 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanole FF) to stop the reaction. 4 µl was loaded directly onto a 6% denaturing polyacrylamide gel, alongside 'G+A' sequence ladder, as prepared above, and gels run in 1 x TBE for approximately 1.5-2 h before drying and analysing the gels as described in section 2.9.1.

CHAPTER 3

The role of accessory factors in autotransporter biogenesis

3.1 Introduction

The moniker "Autotransporter" was originally coined by Meyer and colleagues to describe the mechanism of IgA1 protease secretion from *N. gonorrhoea* (Pohlner *et al.*, 1987) (Section 1.6). These investigators proposed that the three domain architecture of IgA1 protease, and other ATs, was sufficient to mediate translocation of the functional protease to the exterior of the cell. Early on, this supposition was questioned as no rational explanation was provided for how these large proteins could be translocated across the cell envelope without degradation, how they might fold on the cell surface, how the passenger domain could be released from the cell or how some ATs with unusual signal sequences were targeted to the inner membrane (Henderson *et al.*, 1998). As a result of these contentions, several groups sought accessory factors which might aid AT biogenesis. Jose *et al.* (1996) were the first to demonstrate that the ATs formed a periplasmic intermediate and as such it was hypothesised that the ATs might interact with periplasmic chaperones during transit through the periplasm (Henderson *et al.*, 1998). In the intervening years, several groups have demonstrated that the periplasmic chaperones SurA, DegP and Skp play a role in AT biogenesis, functioning to protect ATs from degradation by periplasmic proteases and to maintain the ATs in a translocation competent state (Purdy *et al.*, 2007) (see Section 1.6.3). Several of the periplasmic chaperones interact with the β-barrel Assembly Machinery (BAM) which is composed of five proteins termed BamA-E (Section 1.3.3). The necessity of BamA for AT biogenesis was first demonstrated by Tommassen and co-workers (Voulhoux *et al.*, 2003), with later studies showing a direct interaction between BamA and ATs during AT biogenesis (Ieva and Bernstein, 2009; Ruiz-Perez *et al.*, 2009; Sauri *et al.*, 2009). In addition to the demonstration that several proteins are required for periplasmic transit and subsequent OM translocation of AT passenger domains, several studies identified a variety of proteases that are essential for

cleavage of ATs from their β-barrel translocators (Dautin and Bernstein, 2007). This accumulated body of evidence indicates that the AT pathway is more complex than initially suggested.

On the basis of the controversial nature of AT biogenesis, our laboratory adopted a non-biased random-transposon-mutagenesis strategy to identify novel factors required for AT biogenesis. A bank of previously generated random-transposon mutants of Enteroaggregative *E. coli* strain 042 (EAEC 042) was screened for the secretion of Pet, an archetypal autotransporter (Sheikh *et al.*, 2001). Western immunoblotting of supernatant fractions from ca. 11,000 mutants (~ 2-fold coverage) and subsequent sequencing of the transposon insertions of those mutants deficient in Pet secretion, identified over 60 genes that were observed to have a significant reduction in Pet secretion and thus were taken for further study (Table 3.1). To determine how the gene products of these loci are integrated into the mechanism of Pet biogenesis, studies were undertaken to investigate the impact that loss of these genes had on Pet translocation across the cell envelope and on *pet* transcription. Note that this screen would not enable the identification of any essential genes that are required for Pet biogenesis.

3.2 RESULTS

3.2.1 Veracity of defined *E. coli* BW25113 K-12 mutants by PCR

Using single primer PCR to identify the sites of transposon insertion in EAEC 042, which abolished Pet secretion, it was noted that ca. 15% of the mutants had more than one transposon insertion. Furthermore, it is well recognised that such mutagenesis strategies can result in second site mutations elsewhere in the genome that might obscure the true phenotype of the mutant. Based on these observations it was deemed necessary to utilise strains with defined mutations to investigate the impact that loss of the genes listed in Table 3.1 had on

Table 3.1 List of genes and gene functions used in this study

Gene name	*Relevant function*
ackA	Acetate kinase A and propionate kinase 2
adiY	DNA-binding transcriptional activator putative ARAC-type regulatory protein
ais	Conserved protein induced by aluminium
bamB	Outer membrane lipoprotein component of the BAM complex, involved in outer membrane protein incorporation into the outer membrane
bamC	Outer membrane lipoprotein component of the BAM complex, involved in outer membrane protein incorporation into the outer membrane
bamE	Outer membrane lipoprotein component of the BAM complex, involved in outer membrane protein incorporation into the outer membrane
cadA	Lysine decarboxylase 1
celB	Unkown function
clpB	Protein disaggregation chaperone heat shock protein
crp	DNA-binding transcriptional dual regulator, cyclic AMP receptor protein
crr	Glucose-specific enzyme IIA component of PTS
cspG	DNA-binding transcriptional regulator homolog of *Salmonella* cold shock protein
cysK	Cysteine synthase A, O-acetylserine sulfhydrolase A subunit
degP	Periplasmic chaperone/protease
dnaG	DNA biosynthesis; DNA primase
fkpA	FKBP-type peptidyl-prolyl cis-trans isomerase (rotamase)
fis	Site-specific DNA inversion stimulation factor; DNA-binding protein
folX	D-erythro-7,8-dihydroneopterin triphosphate 2'-epimerase
lon	DNA-binding, ATP-dependent protease La; heat shock K-protein
mhpC	2-hydroxy-6-ketonona-2,4-dienedioic acid hydrolase
nei	Endonuclease VIII
nuoF	NADH:ubiquinone oxidoreductase, chain F NADH dehydrogenase

Table 3.1 List of genes and gene functions used in this study (continued)

Gene name	**Relevant function**
pal	Peptidoglycan-associated outer membrane lipoprotein
pitA	Phosphate transporter
pqiB	Paraquat-inducible protein B
prmA	Ribosomal protein L11 methyltransferase
proP	Proline/glycine betaine transporter
proW	Glycine betaine transporter subunit high-affinity transport system for glycine betaine and proline; membrane component of ABC superfamily
rfaJ	UDP-D glucose:(galactosyl)lipopolysaccharide glucosyltransferase
rlpA	Minor lipoprotein
skp	Periplasmic chaperone
slyB	Putative outer membrane protein
surA	Periplasmic chaperone (peptidyl-prolyl cis-trans isomerase)
thiC	Thiamin (pyrimidine moiety) biosynthesis protein
uvrD	DNA-dependent ATPase I and helicase II
uxuR	DNA-binding transcriptional repressor regulator for uxu operon
yafO	Predicted antitoxin of the YafO-YafN toxin-antitoxin system
ybbB	tRNA 2-selenouridine synthase
ybeL	Conserved protein putative alpha helical protein
ybhC	Predicted pectinesterase putative pectinesterase
ydiM	Predicted transporterputative transport system permease protein
ydiN	Predicted transporter putative transport system permease protein
yeiC	Predicted pseudouridine kinase
yeiL	DNA-binding transcriptional activator of stationary phase nitrogen survival stationary phase nitrogen starvation regulator

Table 3.1 List of genes and gene functions used in this study (continued)

Gene name	**Relevant functin**
yeiM	NUP transporter
yfbV	Conserved inner membrane protein
yfcI	Putative transposase, YhgA-like
yfeK	Unknown function
ygeV	Predicted DNA-binding transcriptional regulator putative transcriptional regulator
ygjI	Transmembrane amino acid transporter protein
yhdG(dusB)	tRNA-dihydrouridine synthase B
yhiL	Unknown function
yhiN	Predicted oxidoreductase with FAD/NAD(P)-binding domain
yiaG	Helix-turn-helix XRE-family like proteins. Prokaryotic DNA binding proteins belonging to the xenobiotic response element family of transcriptional regulators
yihN	Predicted transporter putative resistance protein
yihR	Predicted aldose-1-epimerase
yjdL	Predicted DNA-binding transcriptional regulator putative ARAC-type regulatory protein
yjiC	Hypothetical protein
ykgG	Uncharacterised ACR, YkgG family
ynbC	Lysophospholipase
yohM	Membrane protein conferring nickel and cobalt resistance
yphF	Predicted sugar transporter subunit: periplasmic-binding component of ABC superfamily putative LACI-type transcriptional regulator
yqiH	Predicted periplasmic pilin chaperone
yshA(ompL)	Predicted outer membrane porin L

AT biogenesis. It was decided to utilise mutants created in a laboratory strain of *E. coli* since it has consistently been demonstrated that the biogenesis of ATs is largely unaffected by expression in such strains. To achieve these ends we utilised the *E. coli* BW25113 KEIO mutant library. The KEIO library is a collection of in-frame, single-gene knockout mutants, with independently made duplicate copies where the gene of interest has been replaced with a cassette encoding kanamycin resistance (Baba *et al.*, 2006). During the course of this project we noted inconsistent phenotypes attributable to strains apparently harbouring the same mutations. As a result, all mutants required for this study were checked for their veracity. For each proposed mutant, six independent colonies were obtained after growth on LB-agar containing kanamycin. Each colony was screened by PCR as outlined in Chapter 2, section 2.5, using primers shown in the Appendix, Table 1. Gel electrophoresis of the resulting products was used to confirm the presence or absence of the kanamycin cassette in the target gene. A representative agarose gel is shown in Figure 3.1. Mutants without a verifiable mutation were removed from further study. Those mutants with a verified genetic lesion were used for further study and their name and putative function are listed in Table 3.1.

In addition, to the ~70 genes identified through the random mutagenesis approach we selected a number of additional mutants to screen. These additional mutants were based on previously known functional interactions between their gene products and the proteins encoded by the genes identified through transposons mutagenesis. Thus, as we had identified the periplasmic chaperones DegP and SurA through the transposons screening process, we added mutants lacking the genes encoding other known periplasmic chaperones i.e. Skp, PpiA, PpiD and FkpA. Similarly, we identified a mutation in *bamB*, so strains carrying mutations in *bamC* and *bamE* were added to the collection.

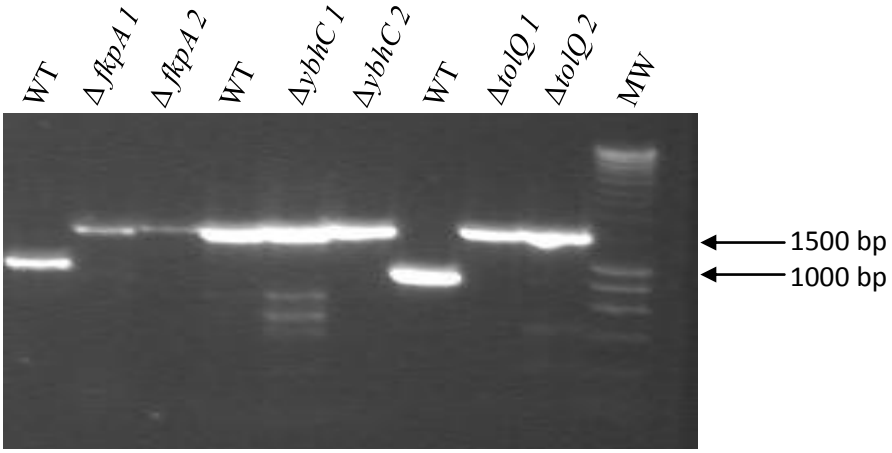


Figure 3.1. ***Gel electrophoresis analysis of PCR products.*** This figure shows the products of a PCR reaction using primers that flanked loci of interest in the *E. coli* BW25113 strain (WT) and the indicated *E. coli* BW25113 mutant strains. The mutant strain was verified when the PCR product equalled that of the kanamycin cassette (~1500 bp) and was different to that of the native size gene in WT. In instances where the native size of the gene equalled that of the kanamycin cassette (as shown here by Δ*ybhC*), the PCR was repeated using primers that annealed to the kanamycin cassette. PCR products were resolved on a 1% agarose gel, pre-stained with Sybr-Green and visualised using a UV transilluminator. MW; molecular weight marker, Hyperladder 1 (Bioline).

3.2.2 Impact of mutations on *pet* transcription

We hypothesised that some of the loci identified through the transposon screen may be involved in the initial stage of Pet biogenesis i.e. transcription. In order to test this hypothesis, the well studied β-galactosidase assay was employed to measure *pet* promoter activity. Thus, the minimal region of the *pet* promoter required for Pet secretion (Desvaux *et al.*, 2006), herein termed the AER1 fragment (Fig. 3.2), was amplified by PCR from the recombinant plasmid pCEFN1 (Chapter 2, Table 2.2) using primers Pet_F and Pet_R (Chapter 2, Table 2.3). The fragment was cloned into the pRW50 low copy number *lac* expression vector, as described in Chapter 2, section 2.4.7. The resulting recombinant (AER1/pRW50) was transformed into the *E. coli* K-12 Δ*lac* strain BW25113 and its isogenic derivatives harbouring the defined mutations listed in Table 3.1. Measurements of β-galactosidase activities in the different strains revealed four categories of mutants (Fig. 3.3). In 59 cases, transcription was unaffected e.g. *ybbB*. For 19 mutants, transcription from the *pet* promoter was increased e.g. *surA*. For 15 genes, transcription was decreased in the respective mutant, including the genes encoding the well-known transcriptional regulators Fis and CRP. Finally, in 20 instances the duplicate mutants gave discrepant results e.g. transcription from the *pet* promoter in the *E. coli* Δ*yjhR*1 is indistinguishable from the wild-type strain *E. coli* BW25113, however transcription from *E. coli* Δ*yjhR*2 is significantly increased when compared to wild-type.

3.2.3 Creation of defined *E. coli* HB101 mutant strains.

Having established that some mutants affected transcription, we wished to test whether loss of other genes would affect later stages of Pet biogenesis. However, inconsistencies between the promoter activities in the independent copies of the *E. coli* BW25113 KEIO mutant library

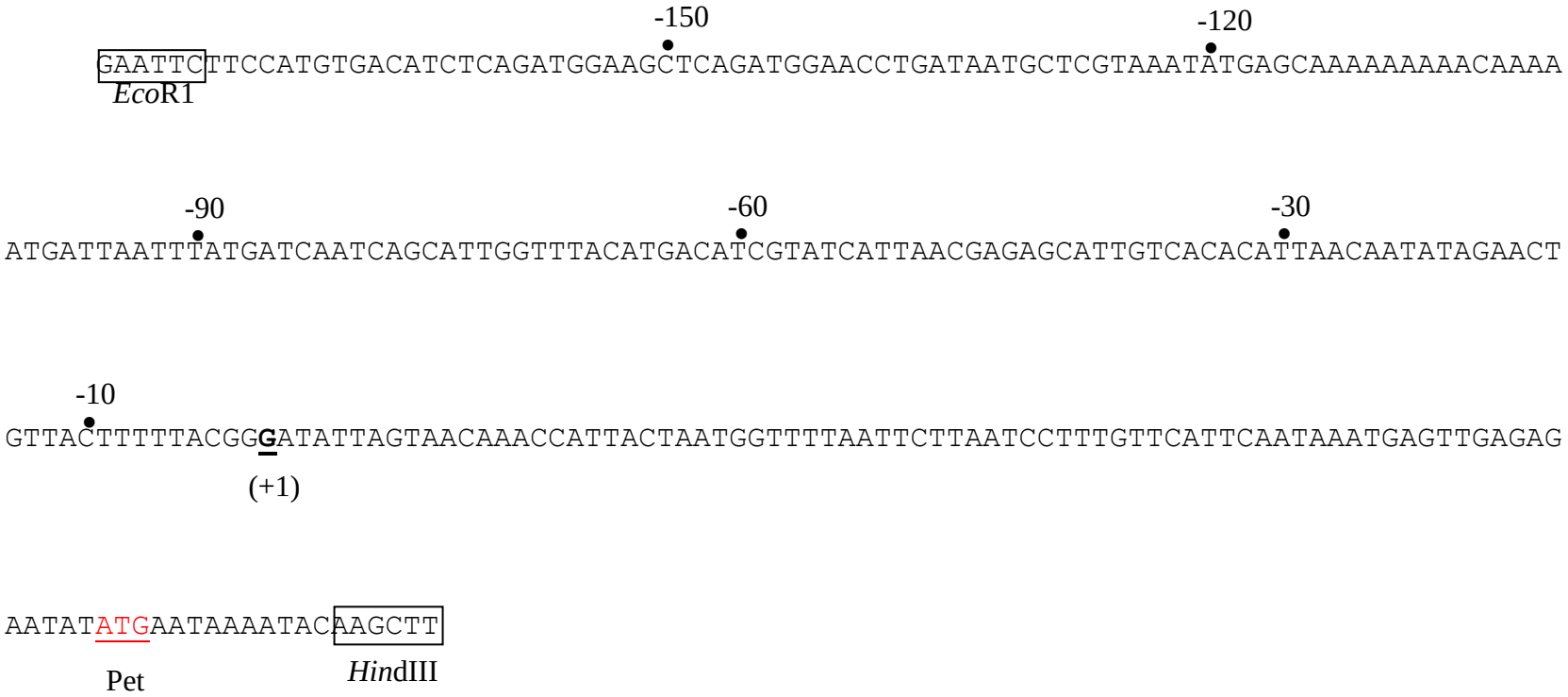
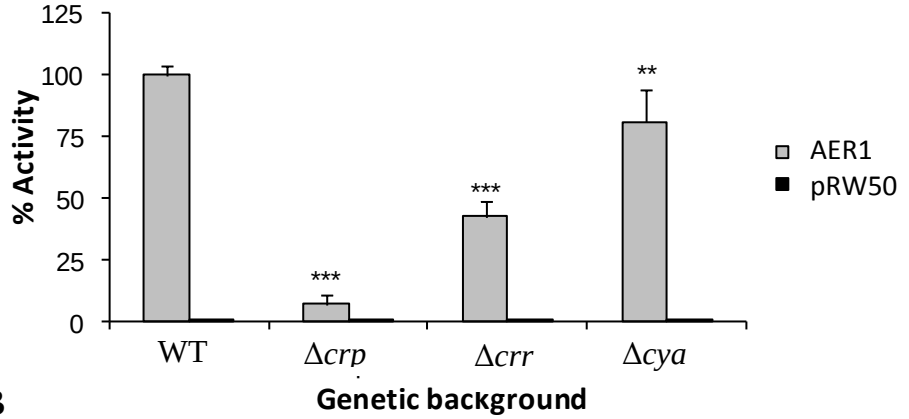
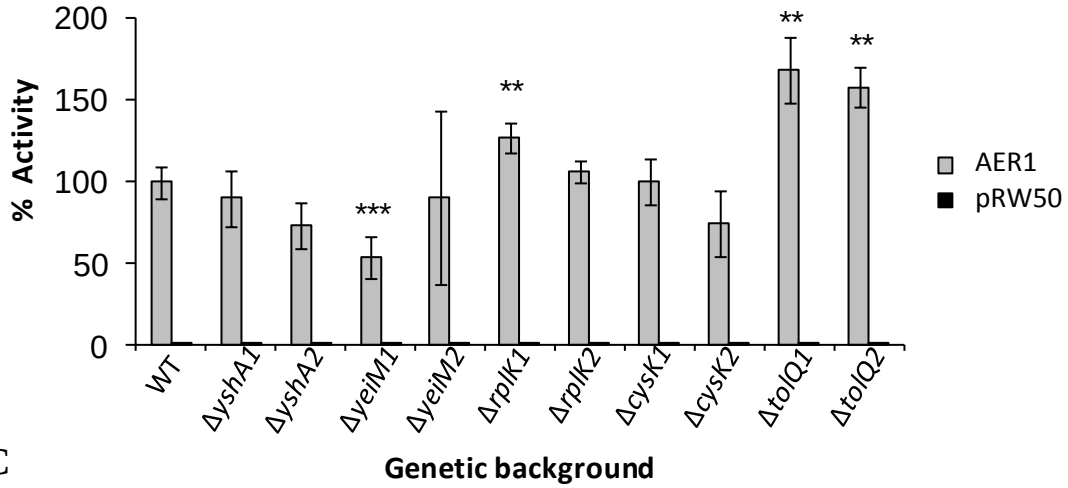


Figure 3.2. ***Base sequence of the AER1 fragment containing the pet gene regulatory region.*** This figure shows the *Eco*R1-*Hin*dIII AER1 fragment, cloned into the promoter-less *lac* expression vector, pRW50. The Pet ATG start codon is underlined and indicated with a horizontal arrow. The predicted *pet* transcript start site is labelled +1.

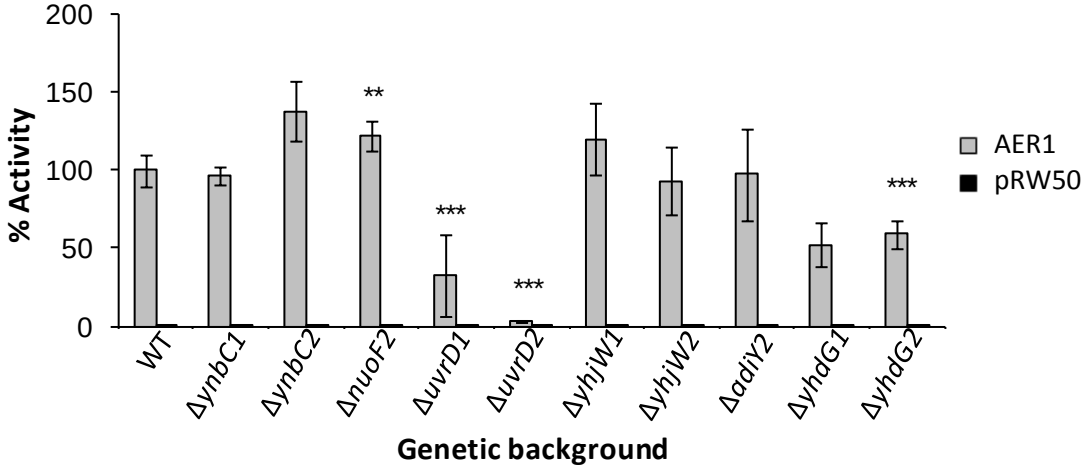
A


125
100
75
50
25
0
% Activity
**
***
***
AER1
pRW50
WT
Δcrp
Δcrr
Δcya
Genetic background

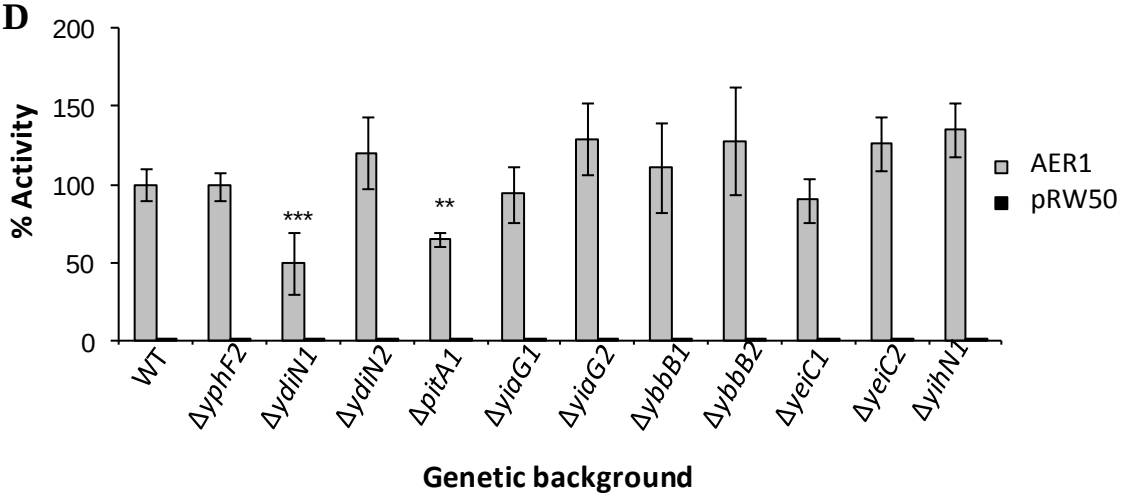
B


200
150
100
50
0
% Activity
**
***
**
**
AER1
pRW50
WT
ΔyshA1
ΔyshA2
ΔyeiM1
ΔyeiM2
ΔrplK1
ΔrplK2
ΔcysK1
ΔcysK2
ΔtolQ1
ΔtolQ2
Genetic background

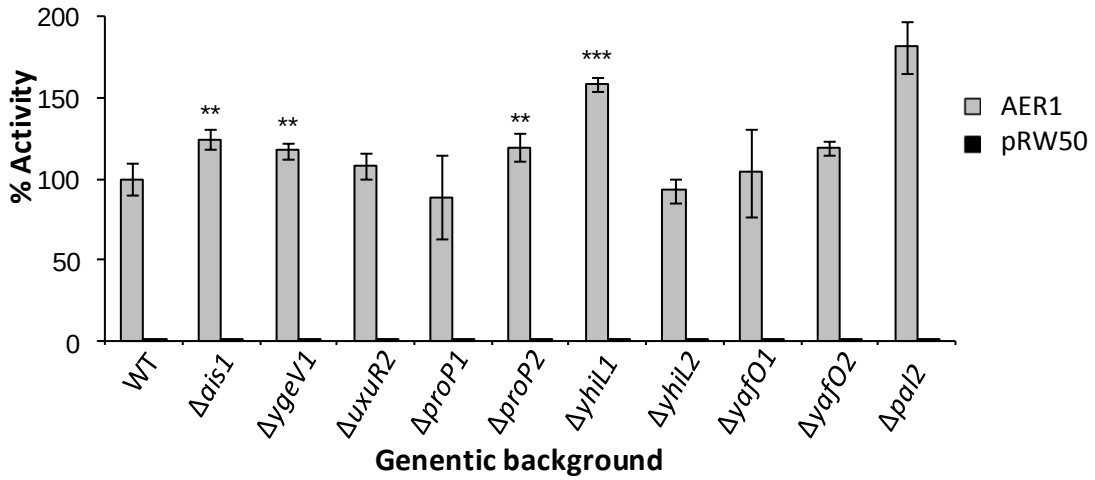
C


200
150
100
50
0
% Activity
**
***
***
***
AER1
pRW50
WT
ΔynbC1
ΔynbC2
ΔnuoF2
ΔuvrD1
ΔuvrD2
ΔyhjW1
ΔyhjW2
ΔadiY2
ΔyhdG1
ΔyhdG2
Genetic background

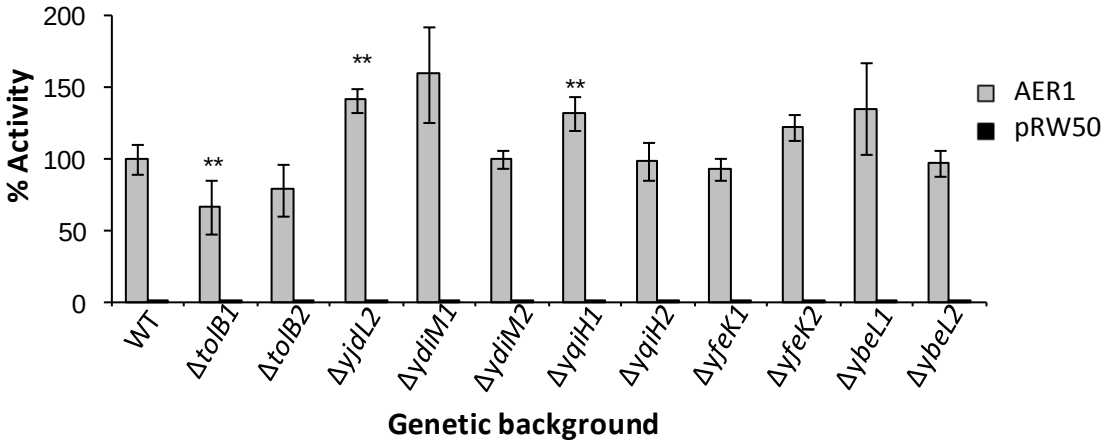
D


200
150
100
50
0
% Activity
AER1
pRW50
***
**
WT
ΔyphF2
ΔydiN1
ΔydiN2
ΔpitA1
ΔyiaG1
ΔyiaG2
ΔybbB1
ΔybbB2
ΔyeiC1
ΔyeiC2
ΔyihN1
Genetic background

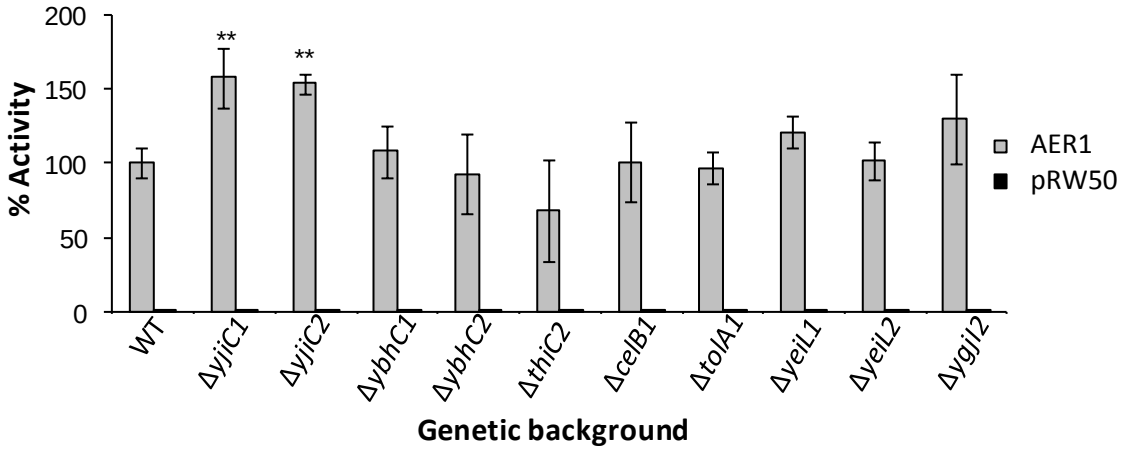
E


200
150
100
50
0
% Activity
AER1
pRW50
**
**
**
***
WT
Δais1
ΔygeV1
ΔuxuR2
ΔproP1
ΔproP2
ΔyhiL1
ΔyhiL2
ΔyafO1
ΔyqfO2
Δpal2
Genentic background

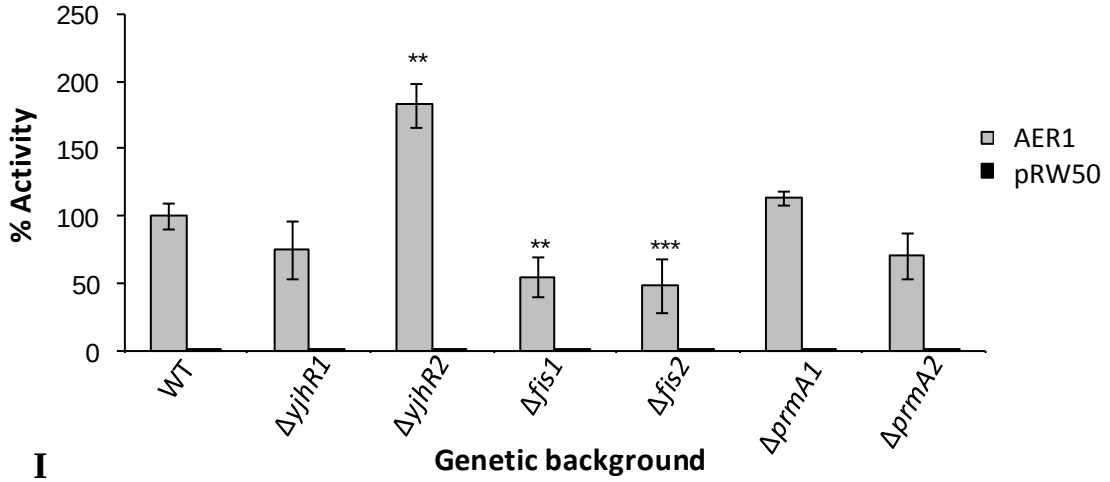
F


200
150
100
50
0
% Activity
AER1
pRW50
**
**
**
WT
ΔtolB1
ΔtolB2
ΔyjdL2
ΔydiM1
ΔydiM2
ΔyqiH1
ΔyqiH2
ΔyfeK1
ΔyfeK2
ΔybeL1
ΔybeL2
Genetic background

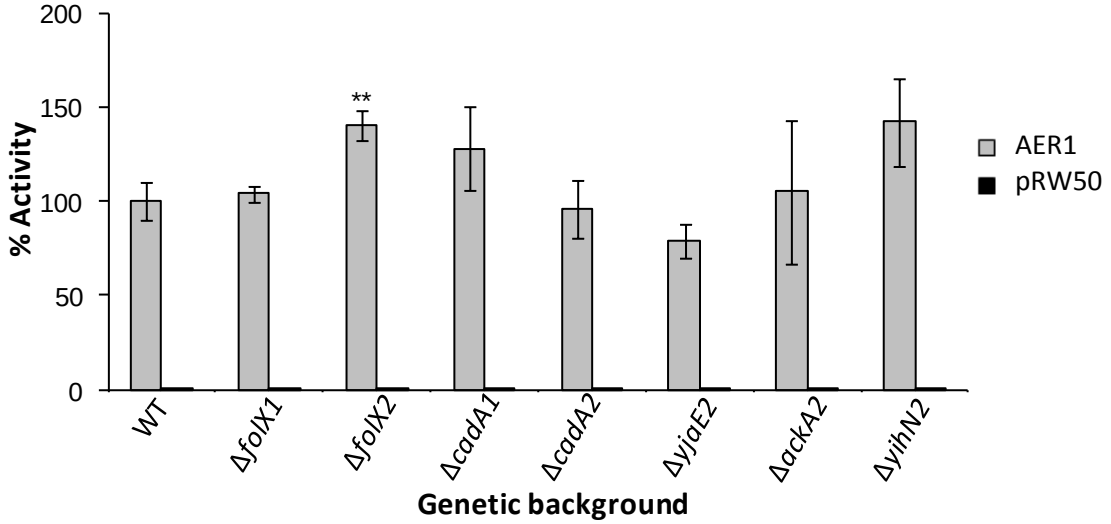
G


200
150
100
50
0
% Activity
**
**
AER1
pRW50
WT
ΔyjiC1
ΔyjiC2
ΔybhC1
ΔybhC2
ΔthiC2
ΔcelB1
ΔtolA1
ΔyeiL1
ΔyeiL2
Δygjl2
Genetic background

H


250
200
150
100
50
0
% Activity
**
**
***
AER1
pRW50
WT
ΔyjhR1
ΔyjhR2
Δfis1
Δfis2
ΔprmA1
ΔprmA2
Genetic background

I


200
150
100
50
0
% Activity
**
AER1
pRW50
WT
ΔfolX1
ΔfolX2
ΔcadA1
ΔcadA2
ΔyjaE2
ΔackA2
ΔyihN2
Genetic background

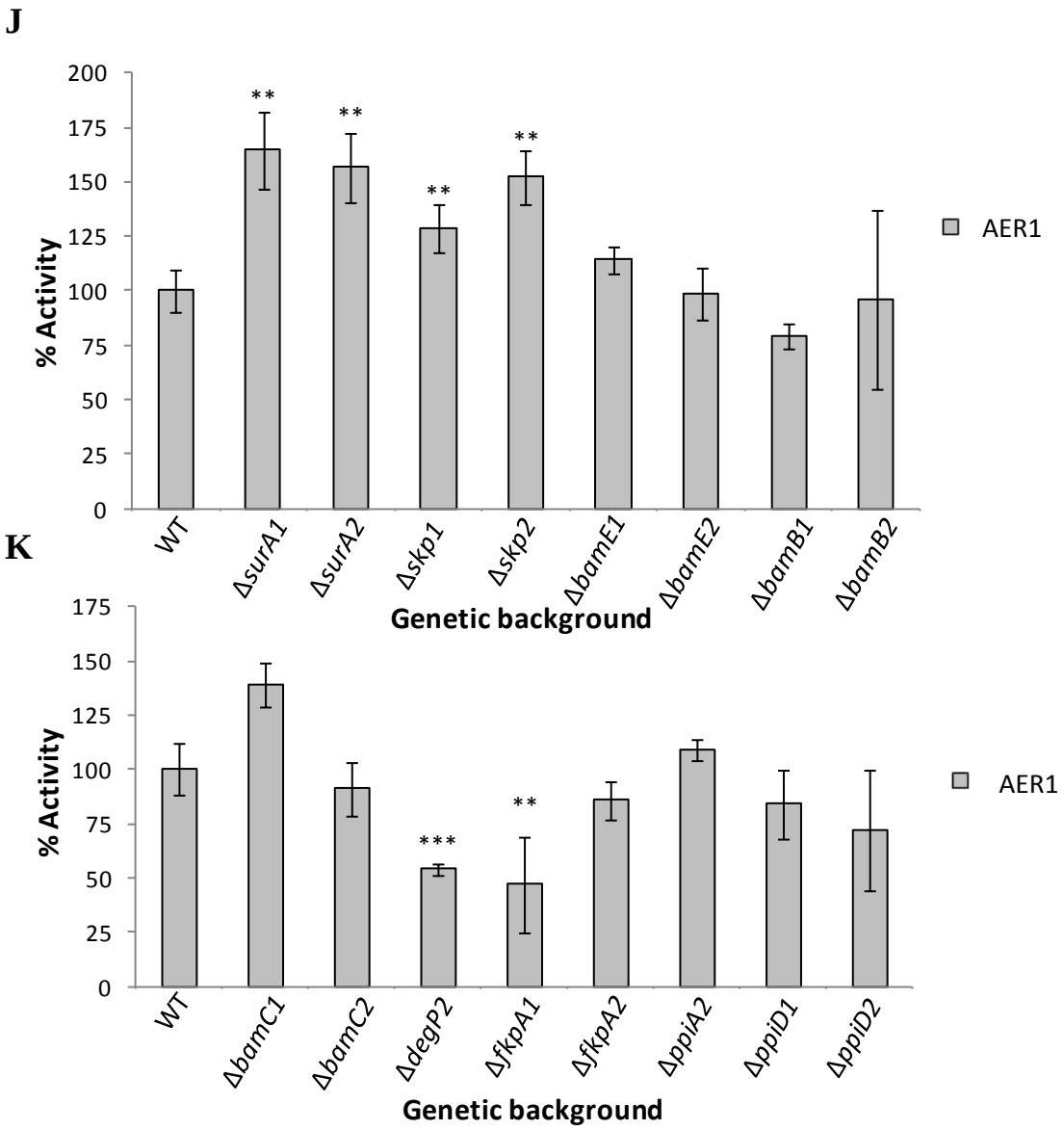


Figure 3.3. ***pet promoter activities in defined E. coli mutant backgrounds.*** **A-K.** This figure illustrates measured β-galactosidase activities in *E. coli* BW25113 and mutant backgrounds containing pRW50 with the *pet* promoter fragment, AER1. Cells were grown aerobically in Luria–Bertani broth at 37°C and harvested during early log phase. β-galactosidase activities are expressed as a percentage of the activity of the *pet* promoter in the *E. coli* BW25113 wild type strain (WT). Each activity is the average of three independent experiments. The unpaired *t*-test was used to compare differences between the promoter activities of mutant backgrounds versus that of the parental strain, ** = $P < 0.05$ & *** = $P < 0.005$.

strain led us to believe that there may be gene duplications or undetected mutations present in some of these strains. Therefore, defined mutations were constructed in the lab strain *E. coli* HB101 by a method based on the λ-Red system (Chapter 2, section 2.6). However, given the large number of loci identified in the original transposon screen and time constraints, we selected only a defined set of mutants for use in the further studies. From the results of the transcription screen shown in Figure 3.3, and the original transposon screen, we selected mutants that we thought to be involved in *pet* transcription; CRP and Fis, mutants that may be involved in the transit of Pet across the periplasm; SurA, Skp, DegP, PpiA, PpiD, FkpA and mutants that may be involved in outer membrane insertion and/or translocation, which included mutants in the BAM complex, BamA - BamE. All mutants were screened by PCR to verify the mutation was correct.

3.2.2 Inner and Outer membrane translocation assays.

Having demonstrated that each of the HB101 mutants was correct, we sought to determine whether inner membrane translocation was affected in these mutant strains. To do this, a construct containing a previously constructed Pet-alkaline phosphatase (PhoA) reporter fusions (pQMDSS*pet*) was used (Desvaux *et al.,* 2007). This reporter construct utilises an inducible promoter and a translational fusion of PhoA to the N-terminal portion of Pet that contains the native Pet signal sequence; thus the efficiency of the Pet signal sequence in directing inner membrane translocation is determined by the accumulation of active PhoA enzyme in the periplasm at a specified time post-induction. To assay for inner membrane translocation the *E. coli* HB101 mutants mentioned above were transformed with pQMDSS*pet* and the equivalent empty vector pQuantagen(kx) and the levels of PhoA activity were measured (Desvaux *et al.*, 2007). The data in Figure 3.4 demonstrate that for all the mutants tested the activity of PhoA in the periplasm was similar for all mutants indicating that

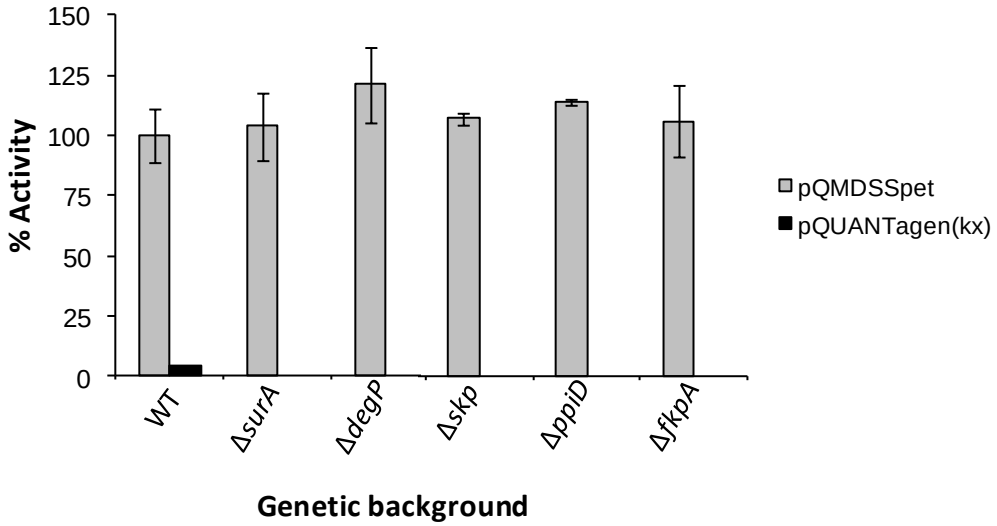


Figure 3.4. ***Efficiency of the Pet signal sequence in mediating inner membrane translocation.*** This figure illustrates the specific activity of periplasmic PhoA expressed from the Pet signal sequence-PhoA fusion (pQMDSS*pet*) or PhoA in the absence of any signal sequence (pQUANTagen(kx), measured in the indicated mutant backgrounds. Each activity is the average from three independent experiments. The efficiency of the Pet signal sequence in mediating inner membrane translocation remains unaffected in the absence of the indicated proteins.

inner membrane translocation was not compromised. In all cases, statistically insignificant PhoA activity was observed for strains harbouring pQuantagen(kx).

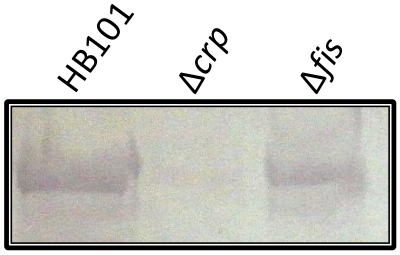
Having established that the *E. coli* HB101 mutants described above lacked defects in inner membrane secretion we sought to confirm that there was a defect in secretion of Pet as indicated by the transposon mutagenesis experiments. Thus, each mutant was transformed with pCEFN1 and after growth to mid-log phase supernatants fractions were harvested, normalised for bacterial numbers and analysed by SDS-PAGE for the presence of Pet. Mutants defective for production of CRP and Fis show significantly diminished levels of Pet accumulating in the extracellular milieu (Figure 3.5). Similarly, *degP* and *surA* null mutants showed significant abrogation of Pet secretion. Notably, the *fkpA*, *ppiA* and *ppiD* mutants produced Pet at levels similar to the wild-type strain, an observation consistent with the fact these loci were not identified via the transposon mutagenesis strategy. Interestingly, the *skp* mutant revealed diminished levels of Pet production. The BAM complex mutants are considered below.

3.2.4 Role of the BAM complex in AT biogenesis

3.2.4.1 BamA and BamD are required for Pet biogenesis

Given that SurA has recently been shown to deliver unfolded OMPs to the BAM complex (Hagan *et al.*, 2010) and that there is a body of evidence supporting a role for BamA in the secretion of ATs, we sought to characterise the precise role of the other BAM components in AT biogenesis. To do this, we first examined the role of the essential BamA and BamD proteins in Pet biogenesis. Full-length Pet protein was expressed under the control of its native promoter in BamA (Lehr *et al.*, 2010) and BamD (Malinverni *et al.*, 2006) depletion strains by transforming them with the *pet*-containing recombinant plasmids pCEFN1 and

A



B

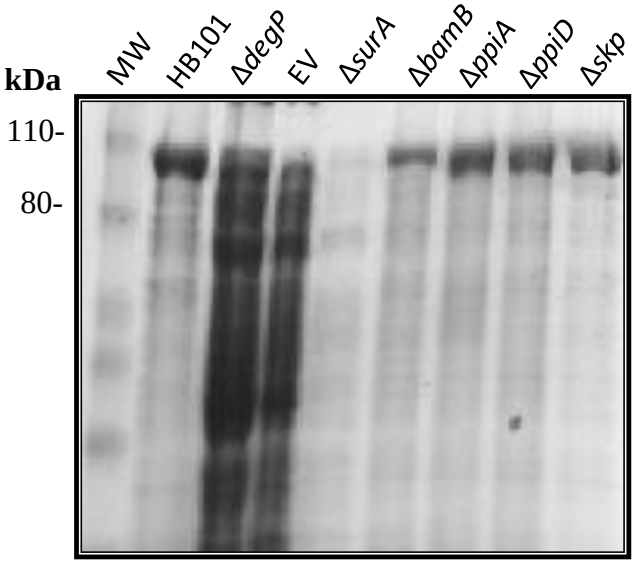


Figure 3.5. ***Secretion of Pet in defined E. coli HB101 mutants.*** SDS-PAGE analysis of trichloroacetic acid-precipitated culture supernatant fractions harvested after the growth of *E. coli* HB101 and indicated mutants, containing pCEFN1 or empty vector (EV). **A.** Pet was detected by SDS-PAGE followed by western immunoblotting using a polyclonal rabbit antiserum generated toward the Pet passenger domain (Eslava *et al.*, 1998). **B.** Pet was detected by SDS-PAGE followed by coomassie blue staining. The sizes of the molecular weight markers are indicated in kDa.

pACYC184/*pet*, respectively (Plasmid details listed in Chapter 2, Table 2..2). In the respective depletion strains, expression of *bamA* or *bamD* is under the control of the pBAD promoter**,** such that expression is induced in the presence of arabinose but repressed in its absence. After growth under both replete (with arabinose) and depletion (without arabinose) conditions, the culture supernatant fractions were harvested from the BamA and BamD depletion strains and supernatant proteins were TCA precipitated. Precipitated proteins were resolved by SDS-PAGE, and detected by Coomassie staining. Results in Figure 3.6 illustrate that when both BamA and BamD are depleted from the cells, secretion of Pet into the culture medium is severely diminished. These results confirm previous observations that BamA is required for AT biogenesis and demonstrate, for the first time, an essential role for BamD in AT biogenesis.

3.2.4.2 BamB, BamC and BamE are not required for AT biogenesis.

While BamA has been implicated in the biogenesis of all examined ATs (Jain and Goldberg, 2007; Sauri *et al.*, 2009) the roles of the non-essential components of the BAM complex (BamB, BamC, and BamE) have not yet been fully investigated. In order to determine the role of these accessory factors in AT biogenesis, *bamB, bamC*, and *bamE* deletion mutants were constructed in laboratory strain *E. coli* HB101 and wild-type UPEC strain CFT073 using the λ-Red system, described in Chapter 2, section 2.6 (Datsenko and Wanner, 2000). In each mutant, the absence of the relevant endogenous protein in whole-cell lysates was confirmed by Western immunoblotting (Fig. 3.7, A). Anti-BamE antibodies were previously described (Knowles *et al.*, 2011); Anti-BamB and BamC antibodies were raised and purified as previously described (Rossiter *et al.*, 2011b). For analysis of Pet biogenesis, *E. coli* HB101 and its mutant derivatives were transformed with pCEFN1 and the empty vector, pSPORT1. Precipitation and analysis of supernatant proteins was performed as described in Chapter 2,

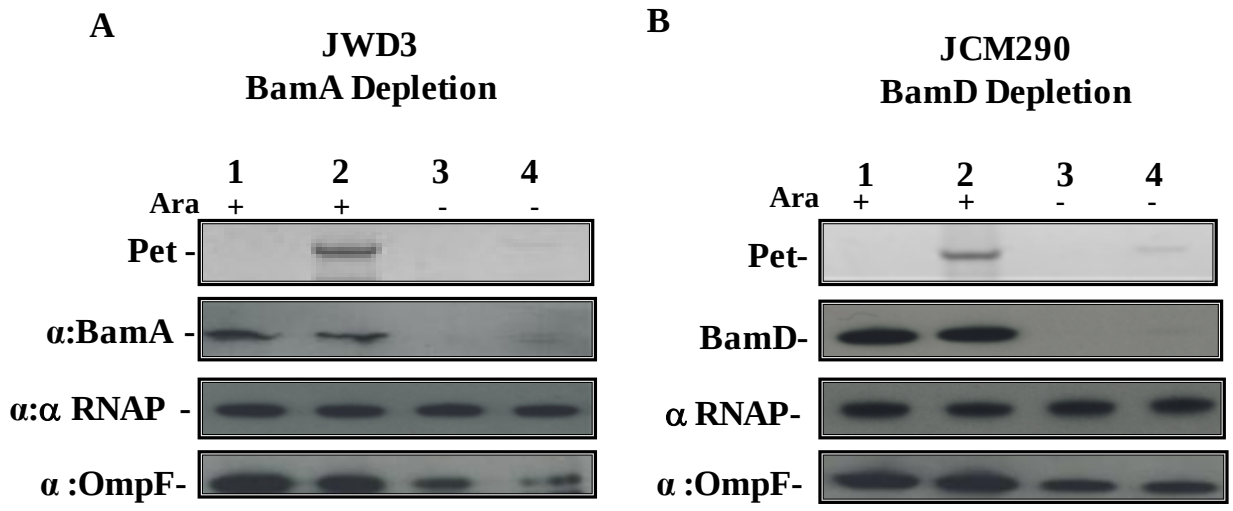


Figure 3.6. ***BamA and BamD are required for Pet secretion.*** BamA and BamD are required for Pet secretion. Shown are Coomassie blue-stained SDS-PAGE of trichloroacetic acid-precipitated supernatant proteins and Western blot analysis of whole-cell extracts from cultures of BamA depletion strain E. coli JWD3 (A) and BamD depletion strain E. coli JCM290 (B). **A.** *E. coli* JWD3, carrying either the empty pSPORT1 vector (lanes 1 and 3) or pCEFN1 (lanes 2 and 4), was grown in LB medium containing 100 μg/ml ampicillin supplemented with either 0.2% L-arabinose (Ara) (+) or 0.2% D-fructose (-). **B.** *E. coli* JCM290, carrying either pACYC184/*pet* (lanes 2 and 4) or the empty vector (lanes 1 and 3), was grown in LB medium containing 30 μg/ml chloramphenicol supplemented with either 0.05% L-arabinose (+) or 0.05% D-fructose (-). For both panels A and B, overnight cultures were diluted into a volume of 50 ml (OD_{600}=0.025) and grown at 37°C with shaking for 3 h. Cultures were harvested, and supernatant proteins were precipitated with 10% trichloroacetic acid. BamA, BamD, and OmpF were detected using antiserum raised in a rabbit, and the α subunit of RNA polymerase (αRNAP) was detected using mouse monoclonal antibodies (Neoclone). Blots were developed using the ECL Plus Western blotting detection system (GE Healthcare). Like OmpF levels, under BamA and BamD depletion conditions, Pet levels are severely diminished whereas the levels of RNAP remain constant.

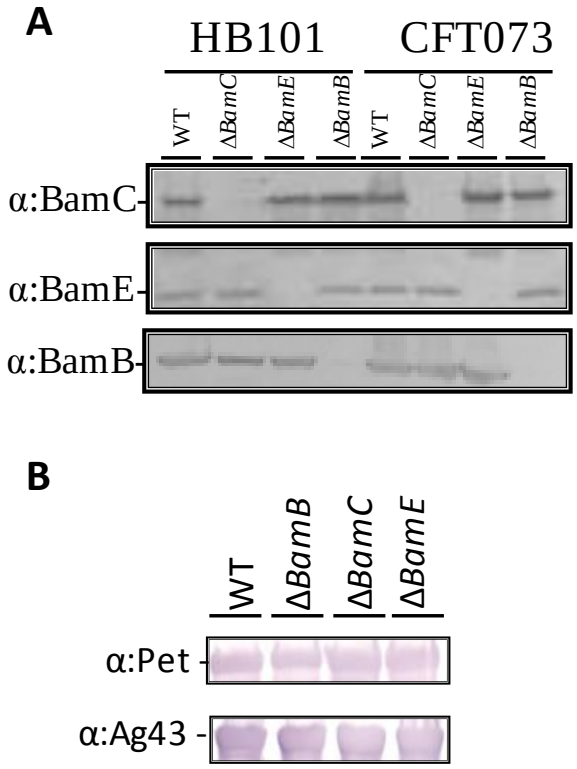


Figure 3.7. ***BamB, BamC and BamE are not required for the biogenesis of Pet or Ag43.*** SDS-PAGE followed by western immunoblotting blot was used to analyse Pet and Ag43 biogenesis in Δ*bamB*, Δ*bamC* and Δ*bamE* cells. **A.** Construction of *bamB*, *bamC*, and *bamE* null mutations in *E. coli* HB101 and *E. coli* CFT073. Whole-cell lysate of each mutant was prepared by spinning down 1 ml of an overnight culture and re-suspending it in Laemmli buffer. Samples were separated by SDS-PAGE and western immunoblotted with antibodies raised to BamB, BamC, and BamE, confirming the absence of the relevant protein from each mutant. **B.** For analysis of Pet biogenesis, cells harbouring either pCEFN1 (Pet +) or pSPORT1 (Pet -) were grown to an OD_{600}=1.0 and pelleted. The supernatant fractions were then filter sterilised and a final volume of 10% trichloroacetic acid was used to precipitate proteins. Pet was localised by SDS-PAGE followed by western immunoblotting. For analysis of Ag43 biogenesis, a heat release assay was performed on cells, transformed with a plasmid carrying the full length Ag43 under control of an arabinose induced promoter, as described in the text. The heat release assay is used to break the non-covalent interactions between Ag43 and the outer membrane, thereby releasing free Ag43, which is then TCA precipitated, as above.

section 2.7 and Pet was detected by Western immunoblotting using a polyclonal rabbit antiserum generated toward the Pet passenger domain (Eslava *et al.*, 1998). Data in Figure 3.7, B show that accumulation of Pet in the culture medium was unaffected by the absence of BamB, BamC, or BamE. To gain a more complete picture of Pet biogenesis, we also examined the production of the C-terminal β-barrel of Pet. In each case, the levels of the Pet β-barrel were unaffected by the absence of the BAM components, and the β-barrel remained heat modifiable (Rossiter *et al.*, 2011b), indicating that it was inserted in the outer membrane in its native conformation. To ensure that these effects were not specific to Pet, we examined the influence of these mutations on another AT, namely, biofilm-promoting Ag43. *E. coli* HB101 and its knockout derivatives were transformed with *agn43*-containing plasmid pCO2, which contains full length Ag43 under the control of the pBAD promoter. After growth in both the presence and absence of arabinose expression levels were monitored using the previously described heat release assay which detects Ag43 release from the bacterial cell surface. The resulting samples were analyzed by SDS-PAGE and Western immunoblotting with anti-Ag43 antibodies (Fig. 3.7, B). The levels of Ag43 were unaffected, indicating that BamB, BamC, and BamE are not required for the translocation of Ag43 to the exterior of the cell. These observations were confirmed for production of the Sat AT cytotoxin in the wild-type strain and mutant versions of UPEC CFT073 (data not shown). While secretion is unaffected, it remains possible that folding of the passenger domain is abnormal. To test this hypothesis, we quantified the functional activity of Pet and Ag43 as a direct indicator of protein folding. An azocasein assay was used to test the function of Pet as described in (Rossiter *et al.*, 2011b). In all of the mutants, the enzymatic activity of Pet was indistinguishable from those of the wild-type, indicating that BamB, BamC, and BamE are not required for folding of the Pet passenger domain (Rossiter *et al.*, 2011b). The Ag43 cell-

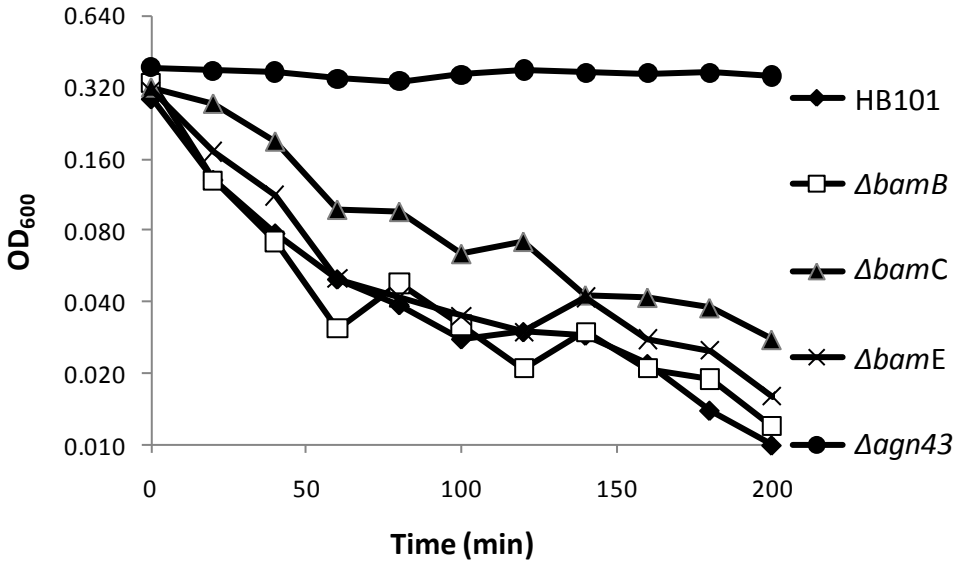


Figure 3.8. ***BamB, BamC, and BamE are not required for folding of Ag43 passenger domains.*** If the Ag43 passenger domain does not fold correctly, it will not be able to perform its extracellular function of mediating cell-cell aggregation. To determine if BamB, BamC, or BamE influences the folding of the Ag43 on the cell surface, the ability of Ag43 to mediate cell-cell aggregation was quantified in wild-type *E. coli* HB101 and mutant derivatives. No difference in the rate of sedimentation between the mutants and the parent strain was observed. In contrast, the *E. coli* HB101 strain lacking the gene encoding Ag43 (Δ*agn43*) did not aggregate at all. These data indicate that Ag43 is not aberrantly folded.

cell aggregation assay was performed as described in Chapter 2, section 2.7.7. Data in Figure 3.8 shows that the ability of Ag43 to mediate cell-cell aggregation remains unaffected in the absence of BamB, BamC and BamE also indicating these lesions do not affect folding of passenger domains.

3.3 DISCUSSION

Whilst the initial transposon screen of EAEC 042 carries with it a large margin of error, the subsequent screen of defined mutations in *E. coli* K-12 has led to the identification of novel factors involved in AT biogenesis. From the study of *pet* promoter activity in these mutants it was found that two of the well characterised 'global' transcription factors of *E. coli*, CRP and Fis, are required for transcription activation of *pet* (Fig. 3.3, A and H). Crr and CyaA are intimately involved in the synthesis of cAMP, which mediates the allosteric activation of CRP. Thus the *crr* and *cyaA* strains are likely to be deficient in cAMP production and thus cause an indirect effect on *pet* transcription due a lack of active CRP. However, at this stage it is a possibility that the deletion of CRP and Fis causes a pleiotropic phenotype, which indirectly causes a reduction in *pet* promoter activity. Further work to confirm a role for CRP and Fis in *pet* transcription is investigated in detail in Chapter 4. Additionally, a significant reduction in *pet* promoter activity was seen in the BW25113 Δ*uvrD* strain. Given that *uvrD* encodes for a DNA helicase involved in nucleotide excision repair, it is difficult to envisage a direct involvement for UvrD in *pet* transcription and it is more likely that this has an effect on the reporter plasmid pRW50, such as replication (Bruand and Ehrlich, 2000). Unexpected changes in *pet* transcription were seen in the mutants encoding for the periplasmic chaperones, SurA, Skp, DegP and FkpA (Fig. 3.3, J and K). In the strains, which carry mutations in SurA and Skp and also the TolQ and Pal mutants, which are components of the

cell envelope spanning Tol-Pal complex, *pet* transcription is significantly increased (Fig. 3.3, B and E). A common phenotype observed in all of these strains is perturbations to the cell envelope. An elaborate response to cell envelope stress is mediated by the CPX system, which is a two component system that is able to sense outer membrane stress and mediate a specific transcriptional response. It has been suggested that the CPX system up-regulates proteases, folding catalysts and even virulence factors to aid survival in adverse conditions (Dorel *et al.*, 2006; Raivio *et al.*, 1999). Thus, it may be a plausible explanation that *pet* expression is up-regulated by the CPX system in response to cell envelope perturbation. In contrast, *pet* promoter activity in the strains carrying mutations in other periplasmic chaperones, DegP and FkpA, was significantly reduced. Discerning a precise reason for this reduction in *pet* promoter activity is unclear, yet it is likely that, again, the phenotype associated with DegP and FkpA null mutants triggers an envelope stress response that indirectly affects *pet* expression.

In accordance with previous studies, we have highlighted a role for SurA and DegP in Pet secretion (Fig. 3.5, B). Our data suggest that SurA plays a role in the periplasmic transit of Pet and/or delivery of Pet to the BAM complex, given that both transcription and inner membrane translocation are not significantly reduced, whereas secretion is severely diminished. An extension of this study came with a collaborative effort to establish direct protein-protein interactions between these proteins using a variety of techniques including Surface Plasmon Resonance. Interactions of DegP and SurA with the homologous AT EspP, from enterohaemorrhagic *E. coli* O157:H7 were observed, where a putative aromatic-polar-aromatic (aro-x-aro) motif was identified within the EspP β-domain as a candidate binding site for these chaperones (Ruiz-Perez *et al.*, 2009). Although, in our hands, we did not see an effect of a *skp* null mutant on Pet biogenesis, our observations were consistent with that seen

for EspP such that a *surA* and a *degP* null mutant caused a reduction in Pet secretion. However, further studies are required to elucidate both specific binding sites and roles for these periplasmic chaperones in AT biogenesis.

Ruiz-Perez and co-workers (2009) were the first to demonstrate a direct interaction of BamA with an AT. Our results confirm that BamA is required for Pet biogenesis. However, this study has highlighted the additional requirement of BamD for AT biogenesis. This observation is consistent with all previous studies which have demonstrated the essential nature of BamD in OMP assembly and reinforce the concept that BamA and BamD work in concert (Hagan *et al.*, 2010; Malinverni *et al.*, 2006). Previous investigations have demonstrated that different OMPs are assembled into the outer membrane via different routes. Thus, porins such as OmpF and LamB, as well as the omptin family member OmpT, have severe biogenic defects in the absence of BamB, whilst, in contrast, levels of TolC increase in a BamB mutant (Charlson *et al.*, 2006; Hagan *et al.*, 2010). Similarly, TolC levels are unaffected by loss of BamE whereas biogenesis of porins is marginally affected (Charlson *et al.*, 2006; Sklar *et al.*, 2007a; Volokhina *et al.*, 2009). We conclude from our data that, like TolC, the AT biogenesis pathway does not require the non-essential lipoproteins BamB, BamC, and BamE. However, in contrast to TolC, AT levels do not increase in the absence of BamB. These studies suggest that there are at least two distinct pathways for OMP assembly in *E. coli*, one which is dependent on BamB, BamC, and BamE (the porins) and one which is independent of these other factors (the ATs and TolC), both of which converge on the core of the BAM complex formed by the BamA and BamD subunits.

Although these studies have made a significant contribution to our understanding of OM translocation by ATs, there is large potential to uncover the involvement of further proteins that are needed for AT biogenesis. On the basis of the preliminary transposon screen and our

transcriptional data outlined in Figure 3.3, further work could be initiated to investigate the role for genes encoding proteins such as PitA, TolQ or Pal in Pet secretion, which may re-define how we perceive AT biogenesis.

CHAPTER 4

Regulation of the *pet* promoter by CRP and Fis

4.1 Introduction

Extensive studies on the regulation of bacterial gene expression have elucidated many models of transcription initiation (see Chapter 1, section 1.2.7). These models highlight the mechanisms used by transcription factors to control the distribution of RNA polymerase between promoters. Although many promoters are regulated by relatively simple mechanisms, additional complexity is often observed when multiple transcription factors converge on a single promoter see Chapter 1, section 1.2.8. Co-regulation of a promoter by two or more transcription factors allows responses to a multitude of environmental cues and thus restricts virulence gene expression to a particular growth phase or niche. Many virulence genes, or pathogenicity islands, are regulated by 'local' transcription factors, as exemplified by the regulation of the *Salmonella* pathogenicity island 1 (SPI-1) by the transcription factor, InvF (Darwin and Miller, 1999). However, the involvement of 'global' regulators, such as CRP and Fis, in the regulation of virulence gene expression is becoming more apparent (See Chapter 1, section 1.2.4.2 and 1.2.6.1). As outlined in Chapter 1, section 1.2.4.1 and 1.2.6, CRP and Fis modulate transcription initiation via highly distinct mechanisms. At promoters where CRP activates transcription, the 16 bp DNA site for CRP is usually located 21 bps upstream from the promoter -10 hexamer element, and activation is dependent upon direct interactions between two surface exposed determinants, activating region 1 (AR1) and activating region 2 (AR2) that interact with the C- and N-terminal domains of the RNA polymerase α subunit (αCTD and αNTD), respectively (Rhodius *et al.,* 1997). In contrast to the rigidity of CRP-dependent regulation, the mechanisms observed for Fis-dependent modulation are rather promiscuous. In order to extend our initial observations that show *pet* promoter activity is dependent on CRP and Fis *in vivo* (Chapter 3, section 3.2.1), we have used a combination of

in vitro and *in vivo* techniques to elucidate the mechanism of transcription initiation, focussing attention on this largely elusive area of AT gene regulation.

4.2 Results

4.2.1 Bioinformatic analyses of the *pet* promoter

The construct AER1/pRW50 outlined in Chapter 3, Figure 3.1 was the starting point for this work and formed the basis for all *in vivo* analyses. On the basis that the *E. coli* BW25113 strains harbouring null mutations in either the *crp* or *fis* genes significantly reduced *pet* promoter activity, the *pet* regulatory region sequence was inspected to identify possible DNA sites for CRP and Fis, and the promoter -10 element. Using bioinformatic approaches and subsequent manual inspection of the nucleotide sequences upstream of the determined translational start site, we identified 5′-TACTTT-3′ as the likely -10 hexamer element, a single DNA site for CRP centred at position -40.5, and tandem DNA sites for Fis, located upstream, centred at positions -91 (Fis I) and -122 (Fis II) (Fig. 4.1). The validity of these predictions was confirmed as described below.

4.2.2 Nested deletion analysis

To investigate the regulatory elements important for modulating the expression of *pet*, a nested deletion analysis of the AER1 fragment in pRW50 was done by sequentially deleting 25 bps of the *pet* regulatory region generating 8 reporter constructs labelled Pet_1 to Pet_8 (Fig. 4.2, A). The primers used for this deletion analysis are shown in Chapter 2, Table 2.3 and the regions that they correspond to in the AER1 fragment are highlighted in bold (Fig. 4.2, A). β-galactosidase activities were determined for the resulting constructs under aerobic conditions at early log phase. The promoter activites shown in Figure 4.2, B indicate the

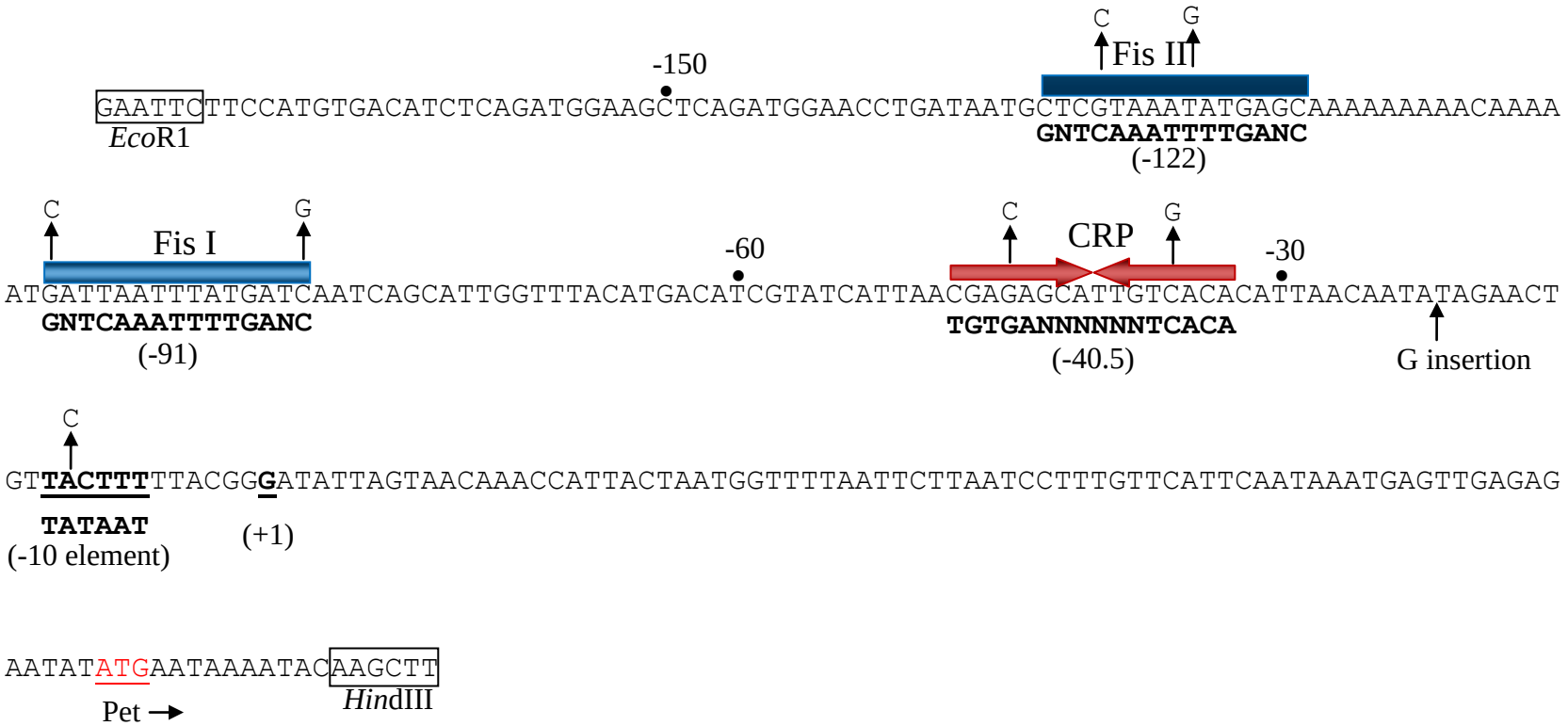


Figure 4.1. ***Base sequence of the AER1 fragment containing the pet gene regulatory region.*** The Pet ATG start codon is underlined and indicated with a horizontal arrow. The predicted transcript start site is labelled +1. Both the +1 start site and the -10 element are highlighted in bold and underlined. Predicted DNA sites for CRP and Fis and the promoter -10 hexamer element are aligned with consensus binding sequences, shown in bold type. The predicted CRP site is indicated with red shaded arrows, the two predicted Fis sites are indicated with blue shaded blocks, and the centre of each site is numbered with respect to the transcript start. Vertical arrows indicate mutations in different targets that are studied in this work.

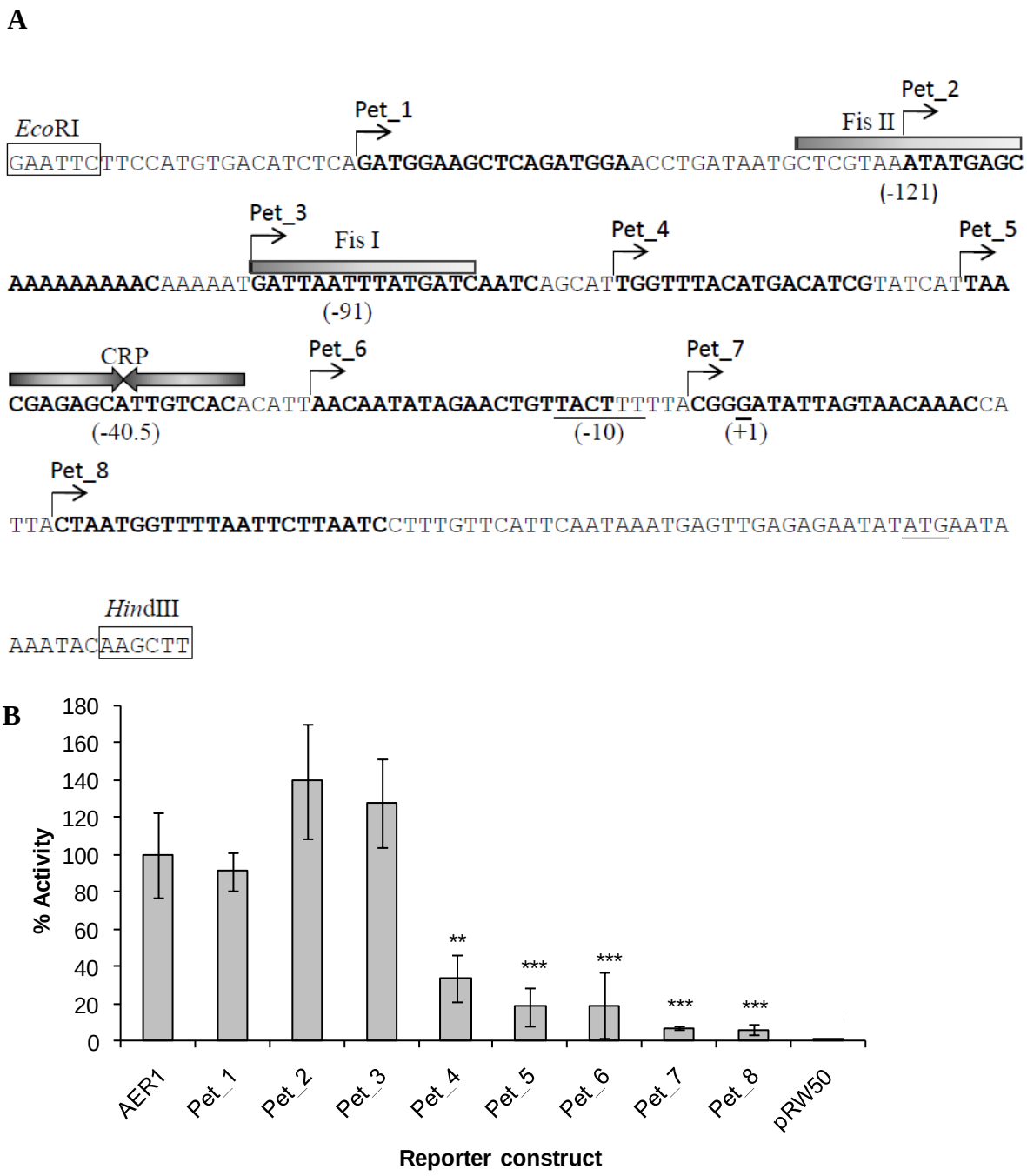


Figure 4.2. ***Nested deletion analysis of the pet promoter.*** **A**. A series of primers were designed, successively deleting 25 bp of the 272 bp region upstream of the +1 transcriptional start sitein the AER1 promoter fragment. These primers are named Pet_1 – Pet_8 and the sequence that they anneal to in the AER1 fragment are indicated in bold. The resulting linear fragments were cloned into pRW50 generating a series of reporter constructs Pet_1-8. **B.** This figure illustrates the β-galactosidase activity of each construct, given as a percentage of the original AER1 reporter construct. The unpaired t-test was used to compare differences between the activities of strains carrying each deleted construct versus that of the original AER1 construct, both in the BW25113 parental strain. ** = $P < 0.05$ and *** = $P < 0.005$.

presence of a mild repressor between the Pet_1 - Pet_2 constructs and the presence of activators between Pet_3 - Pet_4 and Pet_5 - Pet_6, which are consistent with the regulatory elements identified in Figure 4.1.

4.2.3 CRP and Fis act synergistically to activate the *pet* promoter

As CRP and Fis had a demonstrable role in regulating *pet* transcription *in vivo* (Chapter 3 section 3.2.1), a multi-round transcription assay was used to investigate the action of CRP and Fis at the *pet* promoter *in vitro*. For this assay, the AER1 fragment was cloned into plasmid pSR (AER1/pSR), placing the *pet* promoter upstream of the strong factor-independent bacteriophage λ *oop* terminator. Thus, RNA polymerase that initiates transcription at the *pet* promoter runs to the *oop* terminator and generates a discrete transcript that is detected by gel electrophoresis. Figure 4.3, A shows an analysis of transcripts made after incubation of AER1/pSR with purified RNA polymerase holoenzyme containing σ^{70}, in the presence of CRP and Fis. Initiation at the *pet* promoter generates a 182-base transcript that corresponds to a location 77 bps upstream of the *pet* translation initiation ATG that we denote position +1 (Fig. 4.1). Results in Figure 4.3, A demonstrates that optimal activity of the *pet* promoter requires both CRP and Fis. In the absence of Fis, CRP activates transcription poorly, whilst no *pet* transcript is observed in the absence of CRP. Note that in this experiment, the *rna*1 transcript, encoded by the pSR vector, was used as an internal control. This experiment shows that *pet* promoter activity is CRP-dependent but that maximum activity requires Fis to act synergistically with CRP. In a complementary experiment, potassium permanganate footprinting was used to measure CRP- and Fis- dependent open complex formation at the *pet* promoter. Recall that single-stranded DNA, generated by DNA duplex unwinding during transcription initiation, is sensitive to modification by permanganate and modified bases can

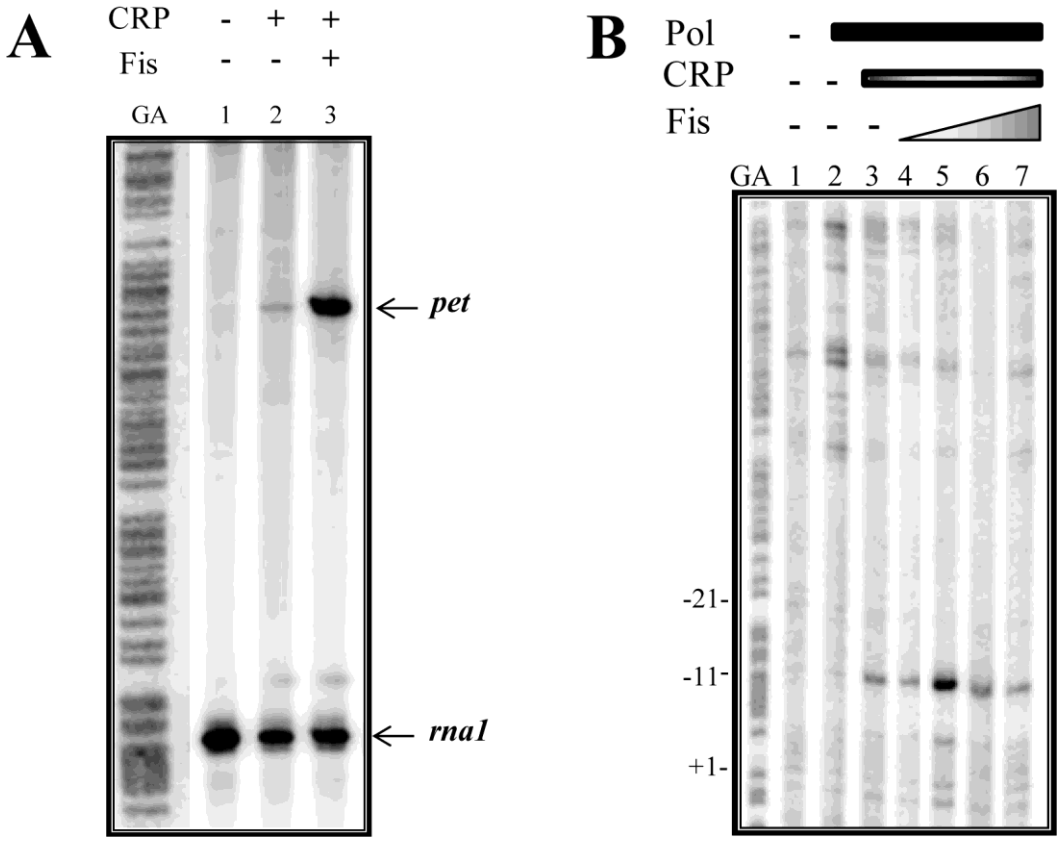


Figure 4.3. ***Transcription activation at the pet promoter*****. A.** The figure shows an autoradiograph of a gel used to analyse transcripts made *in vitro* after plasmid AER1/pSR was incubated with RNA polymerase holoenzyme (50 nM) and 400 nM CRP/200 nM cAMP and 200 nM Fis, as indicated. The transcript generated from the *pet* promoter and the control *rna1* transcripts are indicated by arrows. The activity of the *pet* promoter is optimal when both CRP and Fis are present. **B.** Open complex formation at the *pet* regulatory region. The figure shows the results of *in vitro* potassium permanganate footprinting using the AER1 *Aat*II-*Hin*dIII fragment, end-labelled on the template strand and incubated with a final concentration of 50 nM RNA polymerase holoenzyme (Pol), in the presence or absence of CRP and Fis. Fis concentrations were: lanes 1–3, no protein; lane 4, 100 nM; lane 5, 400 nM; lane 6, 800 nM; lane 7, 1600 nM. CRP concentrations were: lanes 1–2, no protein; lanes 3–7, 50 nM. The gel was calibrated using a Maxam–Gilbert 'G + A' sequencing reaction, and is numbered with respect to the *pet* transcript start site. Open complex formation is present in the presence of Fis, CRP and Pol. Note that, Fis alone is unable to promote open complex formation at the *pet* promoter (data not shown).

be detected by gel electrophoresis (Browning *et al.*, 2009; Sasse-Dwight and Gralla, 1988). Incubation of the AER1 fragment, with purified Fis, CRP and RNA polymerase holoenzyme, causes unwinding of the *pet* promoter at position -11 (Fig. 4.3, B). This unwinding is dependent on CRP and enhanced by Fis, which is consistent with the *in vitro* transcription assay (Fig. 4.3, A). Note that the unwinding is suppressed at higher concentrations of Fis.

4.2.4 Mutational analysis of the *pet* promoter

Given that CRP and Fis affect *pet* promoter activity both *in vitro* and *in vivo*, site-directed mutagenesis was used to confirm the predictions for the DNA sites for CRP and Fis and the assignment of the *pet* promoter -10 element. Derivatives of the AER1 fragment were constructed carrying mutations in the different targets, as illustrated in Figure 4.1. Each mutant derivative was cloned into pRW50, the resultant recombinants were transformed into *E. coli* BW25113, and the activity of each promoter was deduced from the measurement of β-galactosidase expression. Using these constructs we demonstrated that *pet* promoter activity is greatly decreased by mutations in the -10 element or the DNA site for CRP (Fig. 4.4). Similarly, mutation of the Fis I target at position -91 led to a substantial decrease, consistent with the suggestion that *pet* promoter activity is co-dependent on both Fis and CRP. In contrast, mutation of the upstream Fis II target caused an increase in promoter activity, suggesting that upstream-bound Fis may be inhibitory (Fig. 4.4).

4.2.5 CRP and Fis binding at the *pet* promoter

An electrophoretic mobility shift assay was used to investigate direct binding of CRP and Fis to the AER1 promoter fragment and its mutant derivatives, AER1-crp, AER1-fisI and AER1-fisII, carrying mutated DNA sites for CRP, Fis I and Fis II, as shown in Figure 4.1. Results in Figure 4.5, A demonstrate that incubation of labelled AER1 fragment with CRP results in the

AER1:

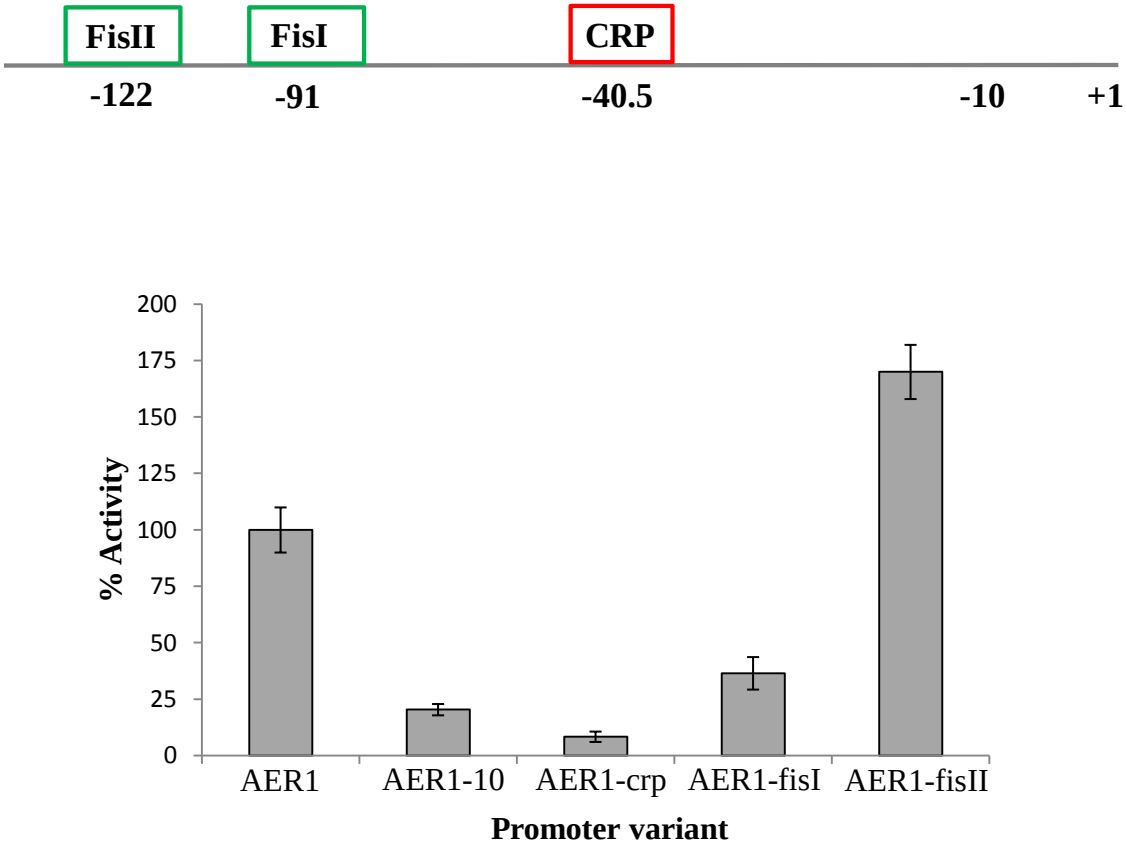


Figure 4.4. ***Mutational analysis of the pet promoter.*** The figure illustrates measured β-galactosidase activities in *E. coli* BW25113, containing a set of pRW50 constructs carrying the *pet* promoter (AER1) and mutant derivatives. The AER1 mutant derivatives were as follows; AER1, starting *pet* promoter; AER1-10, *pet* promoter with mutated -10 element; AER1-crp, *pet* promoter with mutated DNA site for CRP; AER1-fisI, *pet* promoter with mutated Fis I site; AER1-fisII, *pet* promoter with mutated FIS II site (Fig. 4.1). The sites of these regulatory elements are shown in a basic sketch of the AER1 promoter above. Promoter activity is diminished when the -10 element, CRP and Fis I sites are mutated. Promoter activity is increased when the Fis II site is mutated. Cells were grown aerobically in Luria–Bertani broth at 37°C and harvested during early log phase. β-galactosidase activities are expressed as a percentage of the activity with the starting *pet* promoter. Each activity is the average of three independent experiments.

formation of a single shifted species, consistent with a single DNA site for CRP, whilst no shift is seen with the AER1-crp fragment. In contrast, incubation of the labelled AER1 fragment with purified Fis results in the formation of two shifted species, consistent with two DNA sites for Fis (Fig. 4.5, A). Comparison of the shifted species found with AER1 derivatives carrying mutations in either the Fis I site at position -91 or the Fis II site at position -122 (Fig. 4.5, A and B) shows that Fis has a higher affinity for the Fis I site than for the Fis II site.

DNase I footprinting was used to define the DNA sites for CRP and Fis binding. Incubation of the AER1 promoter fragment with CRP resulted in a clear region of protection which maps to around the -40 region (Fig. 4.6, A). As expected, incubation with Fis resulted in two regions of protection which map to around -121, representing Fis II site and -91, representing Fis I site (Fig. 4.6, A). Furthermore, CRP and Fis bind to the *pet* promoter independently of each other, when comparing results shown in Figure 4.6A and B, the pattern of protection in the DNase I footprinting assay is not affected, demonstrating Fis does not alter the binding capacity of CRP.

4.2.6 CRP acts by a class II mechanism at the *pet* promoter

Given the location of the CRP site (centred between bps -40 and -41) we predicted that the *pet* promoter region is a class II CRP-dependent promoter. Transcription activation at class II promoters is dependent upon direct interactions between two activating regions (AR1 & AR2) on CRP with the αCTD and αNTD of RNA polymerase, respectively (Rhodius *et al.,* 1997; West *et al.,* 1993). To test this prediction, we assayed for *pet* promoter activity using our reporter construct in our *crp* strain with plasmids harbouring either native CRP or mutant derivatives, defective in either the AR1, AR2 or both regions simultaneously. Results

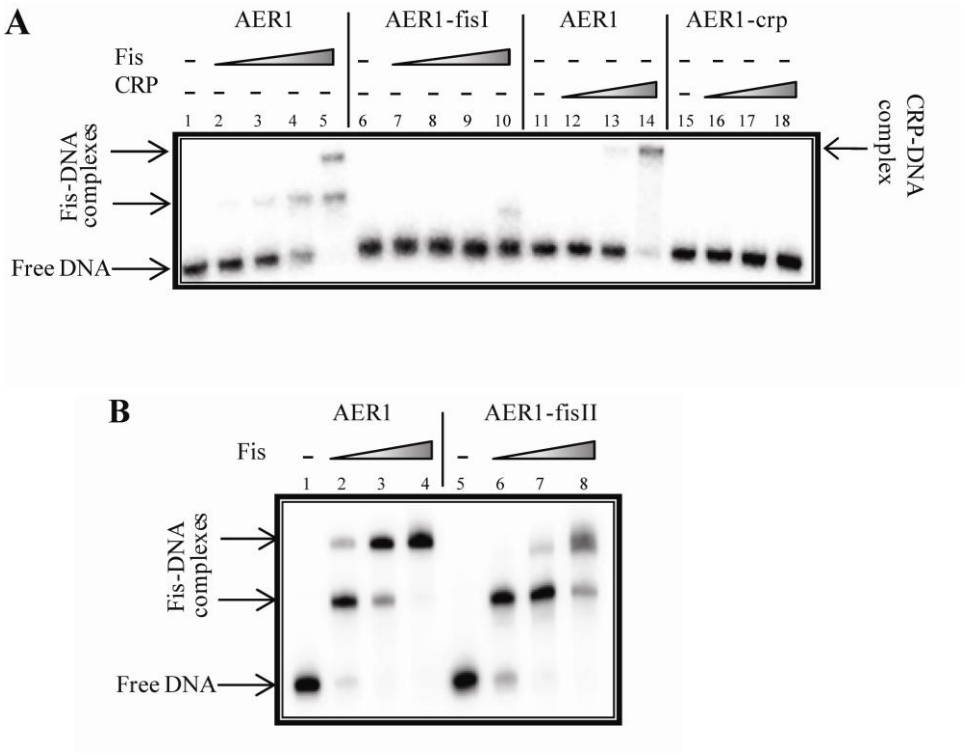


Figure 4.5. ***CRP and Fis binding at the pet promoter.*** **A.** Labelled *Eco*R1-*Hin*dIII promoter fragments (AER1, AER1-fisI or AER1-crp) were incubated with increasing concentrations of purified Fis or CRP protein and subjected to analysis by EMSA. Fis concentrations were: lanes 1, 6 and 11–18, no protein; lanes 2 and 7, 100 nM; lanes 3 and 8, 200 nM; lanes 4 and 9, 400 nM; lanes 5 and 10, 800 nM. CRP concentrations were: lanes 1–10, 11 and 15, no protein; lanes 12 and 16, 50 nM; lanes 13 and 17, 100 nM; lanes 14 and 18, 200 nM. cAMP was included in both the binding reactions and the gel at a final concentration of 200 nM. The positions of the electrophoretically retarded nucleoprotein complexes are indicated by arrows. **B.** Labelled promoter fragments (AER1 or AER1-fisII) were incubated with increasing concentrations of purified Fis and subjected to analysis by EMSA. Fis concentrations were: lanes 1 and 5, no protein; lanes 2 and 6, 200 nM; lanes 3 and 7, 400 nM; lanes 4 and 8, 800 nM. The position of free DNA and nucleoprotein complexes are indicated by arrows.

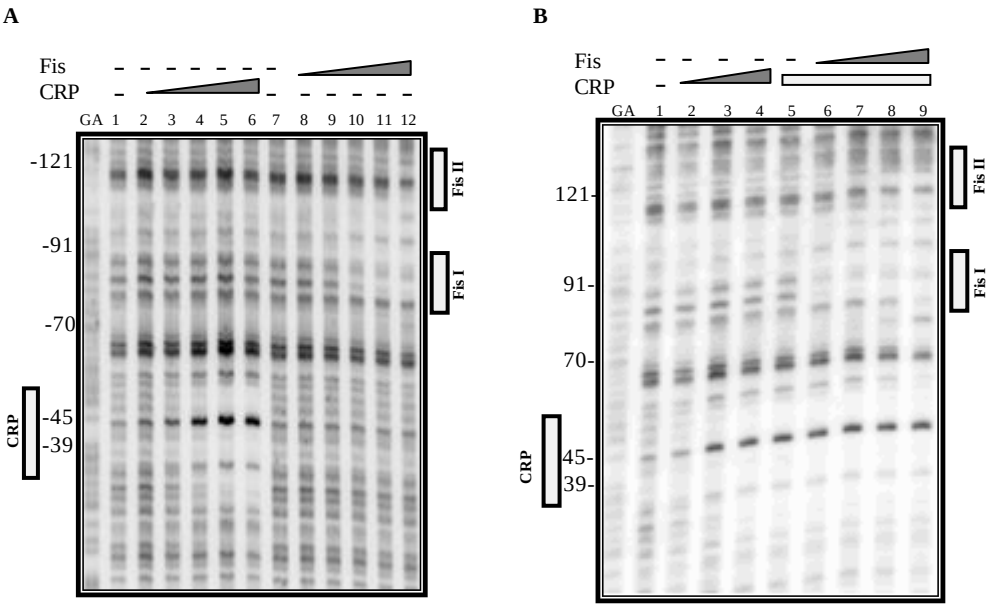


Figure 4.6. ***DNase I footprint analysis of the pet promoter***. **A.** End-labelled *pet Aat*II-*Hin*dIII fragment, on the *Hin*dIII end, was incubated with purified CRP or Fis and subjected to DNase I footprint analysis. Fis concentrations were: lanes 1–7, no protein; lane 8, 100 nM; lane 9, 200 nM; lane 10, 400 nM; lane 11, 800 nM; lane 12, 1600 nM. CRP concentrations were: lanes 1 and 7–12, no protein; lane 2, 25 nM; lane 3, 50 nM; lane 4, 100 nM; lane 5, 200 nM; lane 6, 400 nM. Gels were calibrated using 'G + A' sequencing reactions (indicated by GA). The position of the predicted DNA sites for CRP and Fis are indicated with shaded boxes adjacent to the gel and correspond with regions of protection. **B.** End-labelled *pet Aat*II-*Hin*dIII fragment was incubated with purified CRP and/or Fis and subjected to DNase I footprint analysis. Fis concentrations were: lanes 1–5, no protein; lane 6, 200 nM; lane 7, 400 nM; lane 8, 800 nM; lane 9, 1600 nM. CRP concentrations were: lane 1, no protein; lane 2, 50 nM; lane 3, 100 nM; lane 4, 200 nM; lane 5–9, 400 nM. The position of the predicted DNA sites for CRP and Fis are indicated adjacent to the gel and correspond with regions of protection. Note that the dense bands within the CRP bidning site most likely represent areas of hypersensitivity to DNase I cleavage as a result of CRP bending and exposing specific regions of the DNA.

illustrated in Figure 4.7 indicate that the CRP mutants HL159 and KE101, harbouring defective AR1 or AR2 regions, respectively, had a five-fold decrease in transcription levels in comparison to the plasmid carrying the native CRP. Furthermore, the CRP mutant HL159 KE101, containing both defective AR1 and AR2 regions simultaneously, had a ten-fold decrease in transcription levels.

4.2.7 Fis-dependence of the *pet* promoter is due to non-optimal positioning of CRP.

At most class II CRP-dependent promoters, the DNA site for CRP is located 21 bps upstream from the -10 hexamer element (Rhodius *et al.,* 1997), whilst the spacing is 20 bps at the *pet* promoter (Fig. 4.1). Since location is known to be critical for CRP-dependent activation (Gaston *et al.,* 1990; West *et al.,* 1993), we reasoned that the dependence on Fis for *pet* promoter activity might be due to the location of the DNA binding site for CRP. To address this, a derivative of the AER1 fragment, AER1+1, was constructed with a single G:C bp inserted between position -21 and -22 (Fig. 4.1), and the dependence of the resulting mutant *pet* promoter on CRP and Fis was investigated using *lac* fusions *in vivo* and *in vitro* transcription assays. Results illustrated in Figure 4.8 show that, as expected, the activity of the *pet* promoter in the AER1 fragment is co-dependent on both CRP and Fis. In contrast, whilst the mutant *pet* promoter in the AER1+1 fragment is clearly CRP-dependent, the co-dependence on Fis is lost, both *in vivo* and *in vitro* (Fig. 4.8, A and B).

4.2.8 Co-dependence of promoter activity on CRP and Fis can be generated in a semi-synthetic system

Our data show that the dependence of the wild-type *pet* promoter on Fis is a consequence of the non-optimal positioning of the DNA site for CRP. To investigate whether this is a peculiarity of the *pet* system, we exploited the model *CC(-41.5)* promoter, which is a semi-

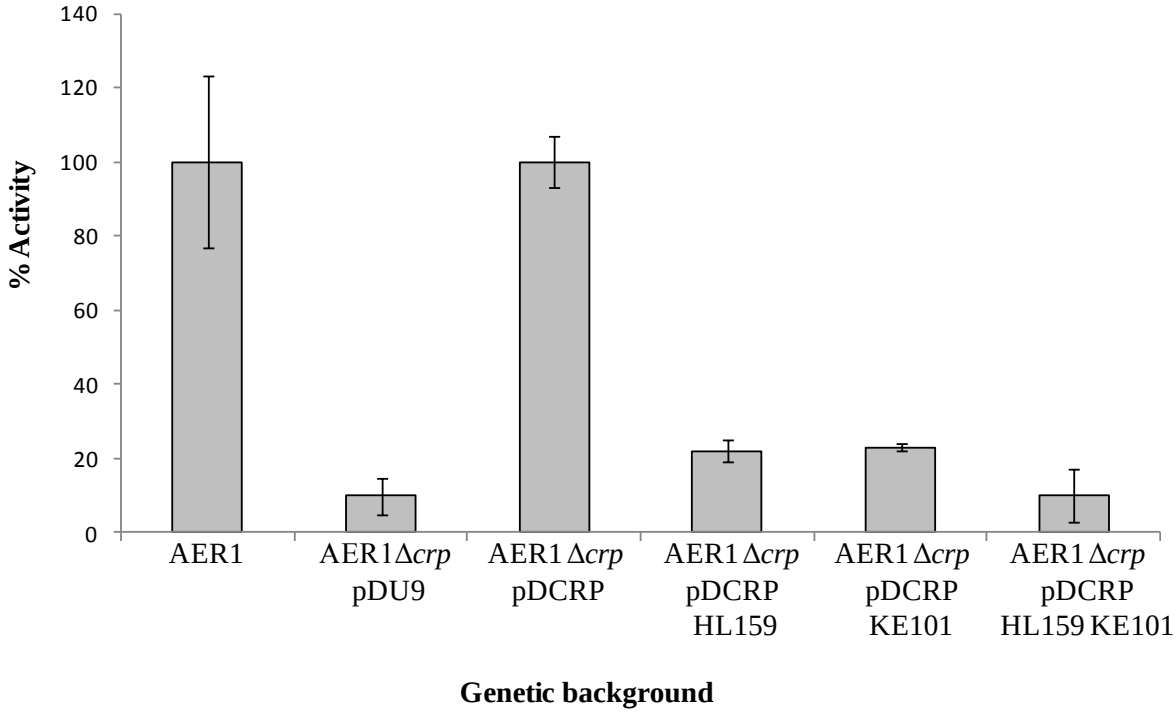


Figure 4.7. ***β-galactosidase activities of the construct AER1/pRW50 measured in the E. coli BW25113 Δcrp strain harbouring plasmids with the native CRP, or its mutant derivatives.*** Plasmids contained in the *E. coli* BW25113 strain included the pDU9, empty vector; pDCRP, native CRP; pDCRP HL159, CRP with a defective AR1 region; pDCRP KE101, CRP with a defective AR2 region; pDCRP HL159 KE101, CRP with defective AR1 and AR2 regions. *E. coli* BW25113 AER1 was included as a positive control. β-galactosidase activities are expressed as a percentage of the AER1/pRW50 construct in the parental strain, *E. coli* BW25113.

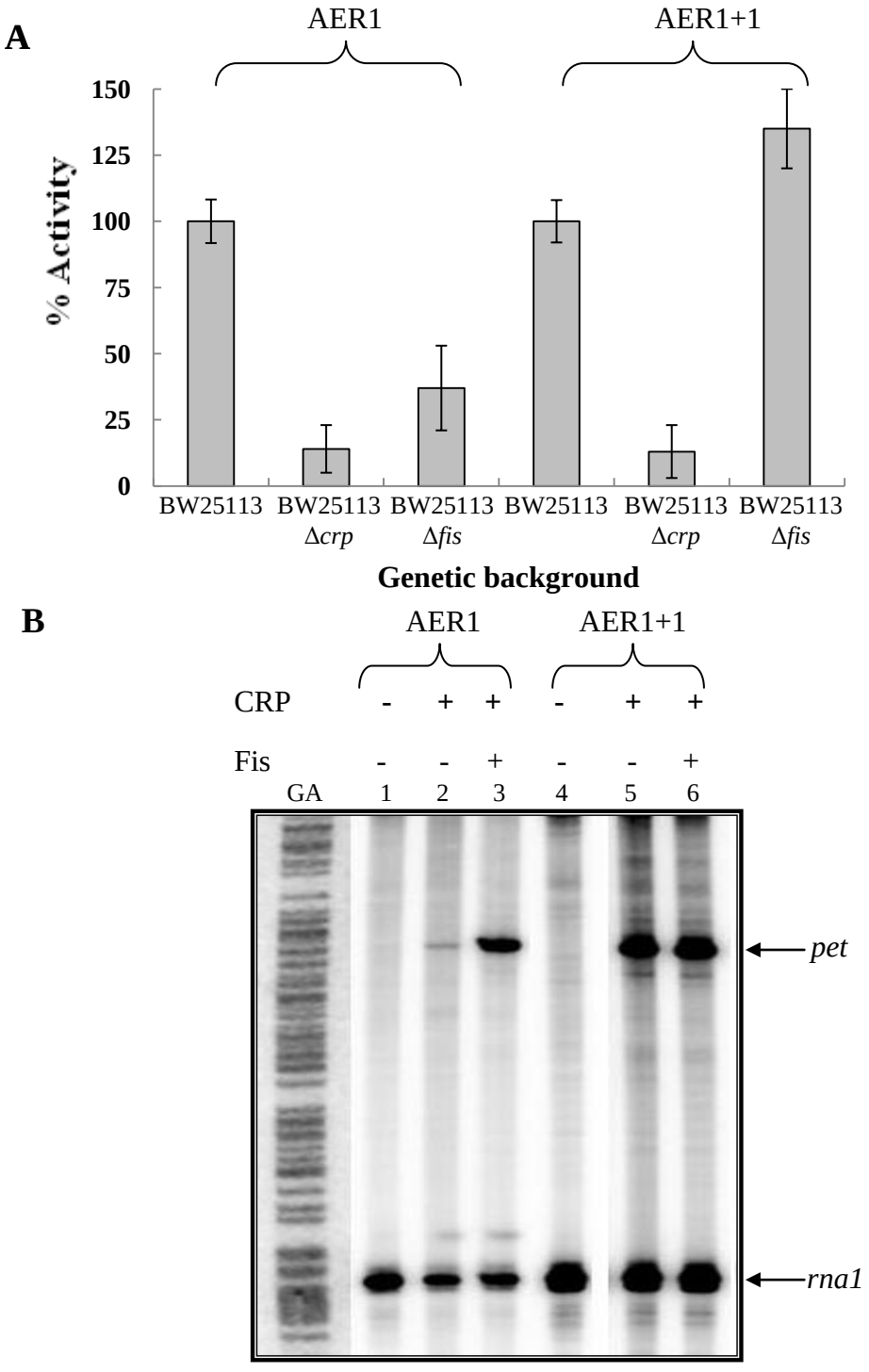


Figure 4.8. ***Optimal spacing between core elements at the pet promoter relieves Fis-dependent CRP activation.*** **A.** The figure shows β-galactosidase activities measured in *E. coli* BW25113, *E. coli* BW25113 *fis* and *E. coli* BW25113 *crp* cells carrying pRW50 containing the *pet* promoter (AER1) or the *pet* promoter derivative with a one base pair insertion (AER1+1). Cells were grown aerobically in Luria–Bertani broth and measurements were taken at early log phase. For each promoter, activities are expressed as a percentage of the activity measured in *E. coli* BW25113 cells and each value derives from three independent experiments. **B.** *In vitro* transcription of the *pet* promoter (AER1) versus the *pet* derivative with a one base pair insertion (AER1+1). A multi-round *in vitro* transcription assay was performed in the presence of 50 nM RNA polymerase holoenzyme with 400 nM CRP and 200 nM Fis, as indicated. The transcript generated by the *pet* promoter and the control *rna1* transcript are indicated by arrows.

synthetic class II CRP-dependent promoter that carries a single consensus DNA site for CRP upstream of a -10 hexamer element (5'-CATAAT-3'), shown in Figure 4.9 (Gaston *et al.*, 1990). It was previously found that CRP-dependent activation is greatly reduced in two derivative promoters *CC(-40.5)* and *CC(-39.5)* that carry one and two bp deletions between the DNA site for CRP and the -10 hexamer element (West *et al.*, 1993). Hence, we constructed the $F_{35}CC(-40.5)$ and $F_{35}CC(-39.5)$ promoters that carry a single DNA site for Fis located 35 bps upstream of the DNA site for CRP, which mimics the organisation of the *pet* promoter (Fig. 4.9, B). Each of the promoter fragments was cloned into pRW50 and promoter activities were measured. Results illustrated in Figure 4.9, C confirm the CRP dependence of each of the promoters and show that, as expected, CRP-dependent transcription is optimal at the *CC(-41.5)* promoter and greatly reduced at the *CC(-40.5)* and *CC(-39.5)* promoters. CRP-dependent activation at *CC(-40.5)* is restored by the introduction of an upstream DNA site for Fis in the $F_{35}CC(-40.5)$ promoter, while the upstream DNA site for Fis has little effect at the $F_{35}CC(-39.5)$ promoter. As the organisation of the -10 hexamers and DNA sites for CRP and Fis at the $F_{35}CC(-40.5)$ promoter and *pet* promoter are identical, we conclude that co-dependence on Fis and CRP is a property of the promoter architecture rather than a peculiarity of the *pet* promoter.

In order to examine the effect of the distance between the upstream Fis and CRP binding sites at the $F_{35}CC(-40.5)$ promoter, the location of the engineered Fis site was altered by increasing or decreasing the space between them, such that the Fis binding site is on the same side or opposite side of the DNA helix (Fig. 4.10, A). This is denoted as $F_nCC(-40.5)$ where 'n' represents the distance between the DNA sites for CRP and Fis binding. These promoter fragments were cloned into pRW50, the resulting recombinants were transformed into *E. coli* BW25113, and the activity of each promoter was deduced from measurement of β-

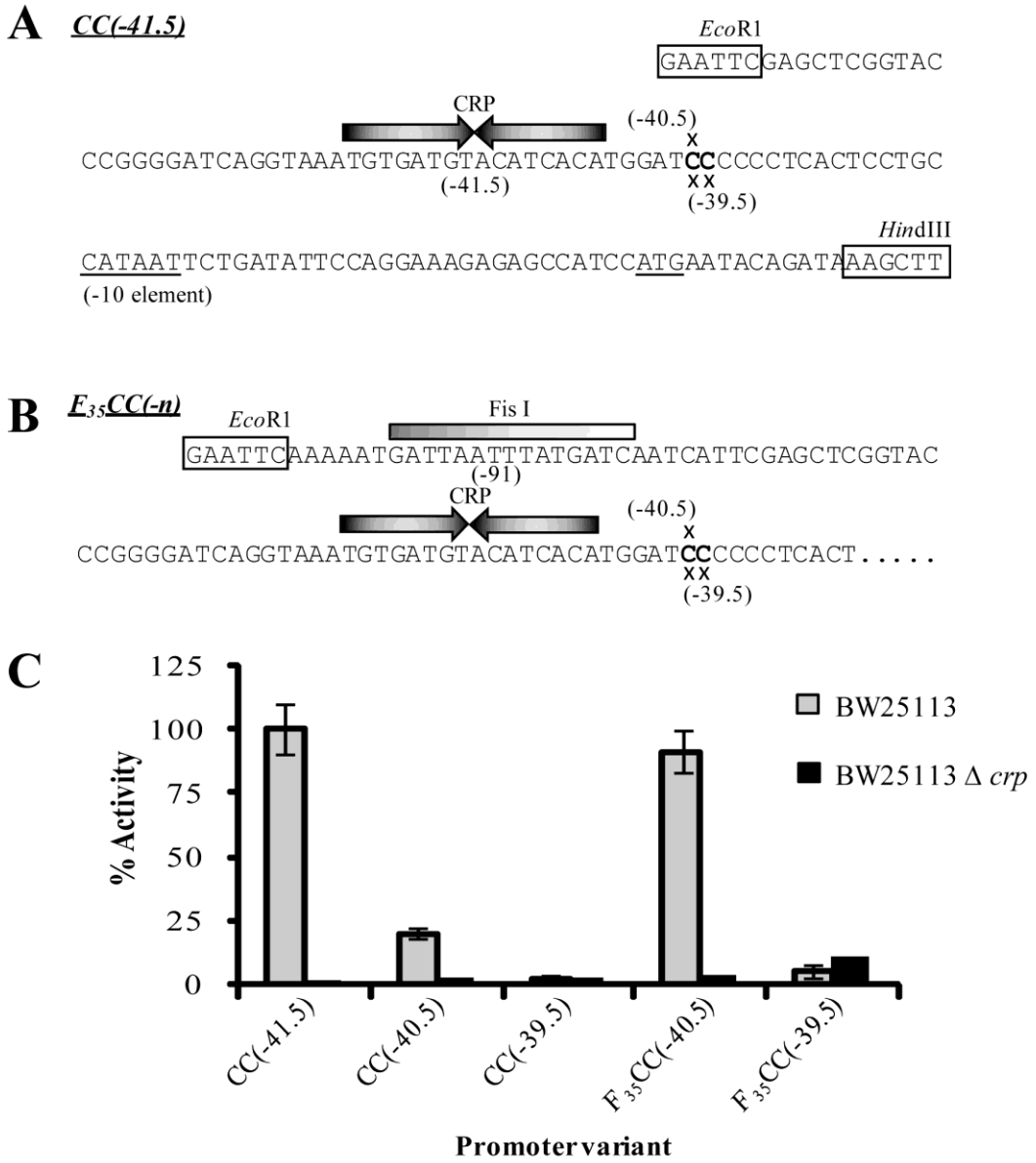


Figure 4.9. ***Non-optimal spacing at the semi-synthetic CC(-41.5) promoter affects CRP-dependent activity, which can be compensated by the insertion of an upstream Fis site.*** **A.** DNA sequence of the *CC(-41.5)* promoter fragment. The DNA site for CRP is indicated by shaded horizontal inverted arrows and the -10 hexamer element is underlined. The crosses (x) denote bases that are deleted in the *CC(-40.5)* and *CC(-39.5)* derivatives. **B.** DNA sequence upstream of the CRP site at the *F35CC(-40.5)* and *F35CC(-39.5)* promoters. **C.** The figure shows β-galactosidase activities measured in *E. coli* BW25113 or *E. coli* BW25113 *crp* cells carrying pRW50 containing the *CC(-41.5)* promoter or the *CC(-40.5), CC(-39.5), F35CC(-40.5)* and *F35CC(-39.5)* derivatives. Cells were grown aerobically in Luria–Bertani broth and measurements were taken at early log phase. β-galactosidase activities are expressed as percentages of the measured activity with the *CC(-41.5)* promoter in *E. coli* BW25113 cells and each value is the average of three independent experiments.

galactosidase expression. Results in Figure 4.10, B demonstrate that promoter constructs *F_{40}CC(-40.5)* and *F_{30}CC(-40.5)*, dramatically reduced promoter activity. However, the deletion of one helical turn, *F_{25}CC(-40.5)*, significantly increased promoter activity. In contrast, the insertion of one helical turn, *F_{45}CC(-40.5)*, did not increase promoter activity (Rossiter *et al.*, 2011a).

4.3 Discussion

The principal finding from this work is that expression of Pet from EAEC strain 042 is co-dependent on Fis and CRP. Fis and CRP are well characterised global transcription factors that regulate hundreds of promoters in *E. coli*. It is well known that CRP activates transcription at many target promoters by binding to a single target that overlaps the promoter -35 element and that activation requires CRP to make two direct contacts with RNA polymerase via two activating regions (AR1 and AR2) (Rhodius *et al.,* 1997). Data in Figure 4.7 confirm that both AR1 and AR2 are essential for CRP-dependent activation at the *pet* promoter. Previous studies (Gaston *et al.,* 1990) have shown the importance of positioning of the DNA site for CRP, whereby optimal activation results when it is located 21 bps upstream from the promoter -10 hexamer element (at position -41.5). Our experiments show that the spacing is 20 bps at the *pet* promoter and that this is the reason why a co-activator is required. It is known that upstream bound transcription factors can co-activate together with CRP by making a direct contact with the RNA polymerase αCTD (Belyaeva *et al.,* 1998; McLeod *et al.,* 2000). Furthermore, the strict geometry of bound regulatory proteins is well recognized as being an important tool for the efficiency of bacterial promoters, whereby the DNA binding sites are arranged in helical register (Gaston *et al.,* 1990; Pul *et al.,* 2008; Ushida and Aiba, 1990). In this case, the upstream bound factor is Fis. Experimental results in Figure 4.10

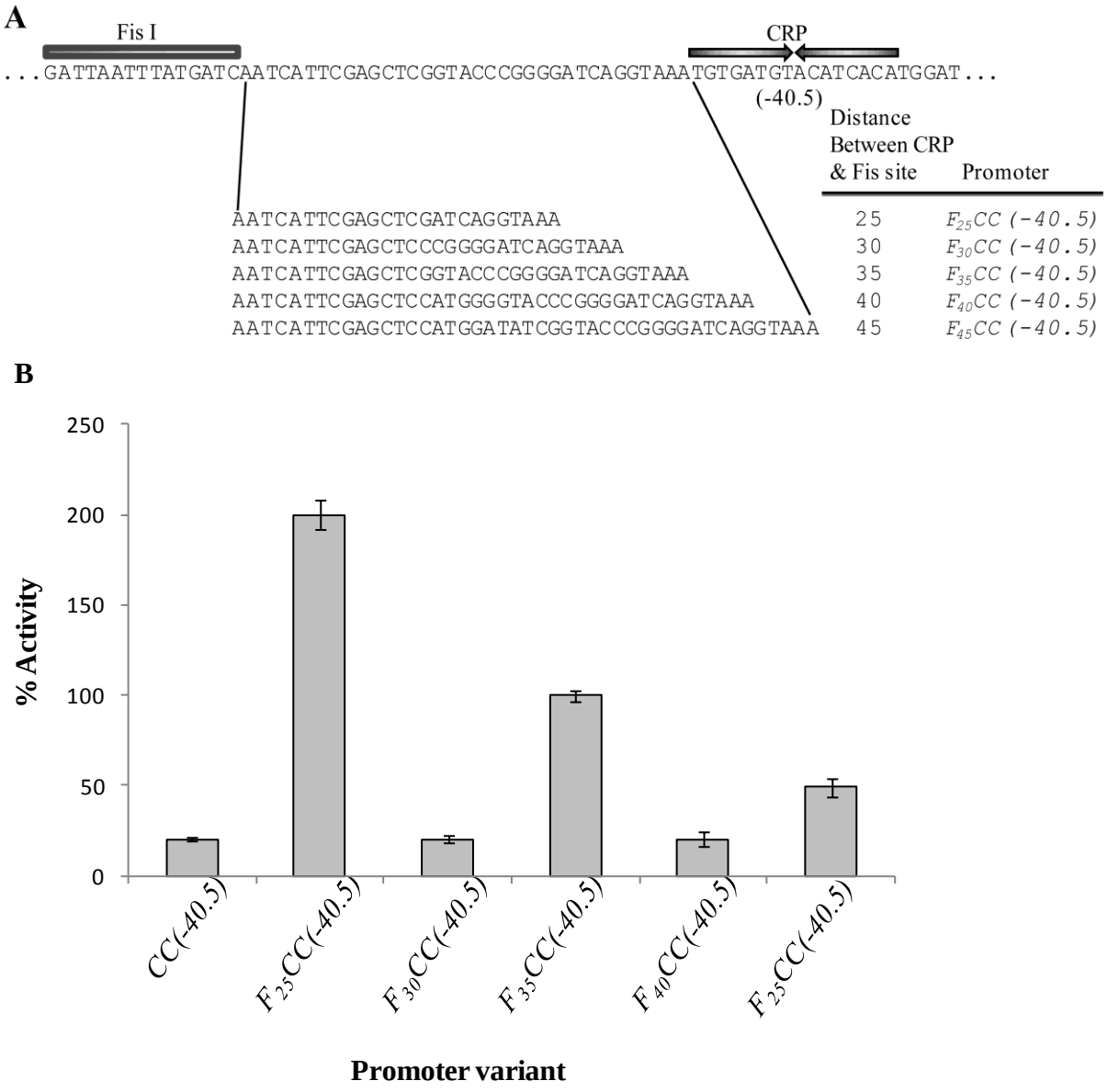


Figure 4.10. ***Helical phasing is crucial for Fis-dependent CRP activation.*** **A.** Base sequence of the *FCC(-40.5)* promoter derivatives where the Fis binding site has been inserted at various position upstream of the CRP binding site. Promoter names are given in the fashion of $F_nCC(-40.5)$, whereby 'n' equals the number of bases between the Fis and CRP binding sites. **B.** The figure shows β-galactosidase activities measured in *E. coli* BW25113 carrying the *CC(-40.5)* promoter and the $F_nCC(-40.5)$ promoter derivatives in pRW50. Cells were grown aerobically in Luria–Bertani broth and measurements were taken at early log phase. β-galactosidase activities are expressed as a percentage of the measured activity of the $F_{35}CC(-40.5)$ promoter. Each value is the average of three independent experiments.

demonstrate the importance of helical phasing in Fis-dependent CRP activation, i.e. that Fis and CRP binding to the same face of the helix is essential for transcriptional activation *in vivo*; the finding that Fis and CRP bind to the same helical face is consistent with the idea that Fis makes a direct interaction with RNA polymerase αCTD and enhances transcription. Our experiments with the semi-synthetic *CC(-41.5)*, *CC(-40.5)* and *CC(-39.5)* promoters (Fig. 4.9) argue that the promoter co-dependence on CRP and Fis is created by the non-optimal positioning of the DNA site for CRP rather than specific sequence features of the *pet* promoter. Individually, Fis and CRP have been implicated in the modulation of many virulence genes in pathogens (Akhter *et al.*, 2008; Muller *et al.*, 2009; Ó'Cróinín *et al.*, 2006; Steen *et al.*, 2010) but, to our knowledge, this is the first example of a virulence gene that is co-regulated by both together. Levels of Fis in *E. coli* vary from very high to very low according to growth conditions and it is these fluctuations that set Fis activity (Mallik *et al.*, 2006; Morin *et al.*, 2010). Thus, highest levels of Fis are found in rapidly growing cells, while Fis is undetectable in non-growing cells. In contrast, CRP activity is controlled by both its expression and the level of the cyclic AMP second messenger (Grainger and Busby, 2008). In *E. coli*, CRP and cyclic AMP levels are known to be raised as cell growth slows, and hence, the Fis and CRP activities are controlled by completely different environmental cues. The result of this is likely to be that the expression of Pet in a mammalian host is restricted to certain phases of growth or certain niches. One particular study, detailing Pet secretion through the growth curve of EAEC 042, demonstrated that Pet secretion spikes early in exponential phase and, thereafter, gradually accumulates (Betancourt-Sanchez and Navarro-Garcia, 2009). This is consistent with our findings that Pet expression is co-regulated by CRP and Fis, suggesting that the growth-phase dependent regulation of Fis expression promotes Pet secretion early on in EAEC 042 growth. Furthermore, Fis is now well documented as a

key player in the pathogenesis of EAEC 042, as its activities are critical for the expression of the EAEC 042 specific AAF I/II fimbriae, the central virulence transcription factor AggR and the formation of biofilms (Morin *et al.*, 2010; Sheikh *et al.*, 2001). It is rather unusual that Fis plays a large role in pathogenesis of EAEC 042, given that expression of fimbriae and formation of biofilms is classically associated with sessile growth, where Fis expression is low (Morin *et al.*, 2010). However, one may envisage a system whereby early Pet secretion, promoted by Fis and CRP, primes the intestine for colonization by cleaving fodrin and rounding the intestinal epithelial cells (Henderson *et al.*, 1999), making the epithelial layer interface more accessible for bacterial adhesion thus, contributing to the establishment of a biofilm. Finally, when looking at the distribution of autotransporters, Pet is confined to EAEC 042. However, functionally similar autotransporters, SigA (Al-Hasani *et al.*, 2009) and Sat (Guyer *et al.*, 2000) are found in *Shigella* spp. and uropathogenic *E. coli* (UPEC) respectively. Interestingly, bioinformatics search patterns suggest that similar promoter architecture, with respect to non-optimal spacing between a CRP binding site and the -10 hexamer element, is apparent in other virulence determinants such as putative type III and type VI secretion system proteins from EAEC 042 and also SigA and Sat. However, an upstream Fis site within these promoters is less evident (Fig. 4.11). This suggests the presence of a novel mechanism whereby CRP placed at a non-optimal position benefits the bacterium by allowing integration of signals, by other transcription factors, to rescue CRP from a non-optimal position and thereby promotes transcription in precise phases of growth/environmental conditions. Therefore, we envisage that through extensive bioinformatic and experimental techniques, we may establish CRP placed at a non-optimal position as a novel mode of virulence gene regulation, applicable to many Gram-negative pathogens.

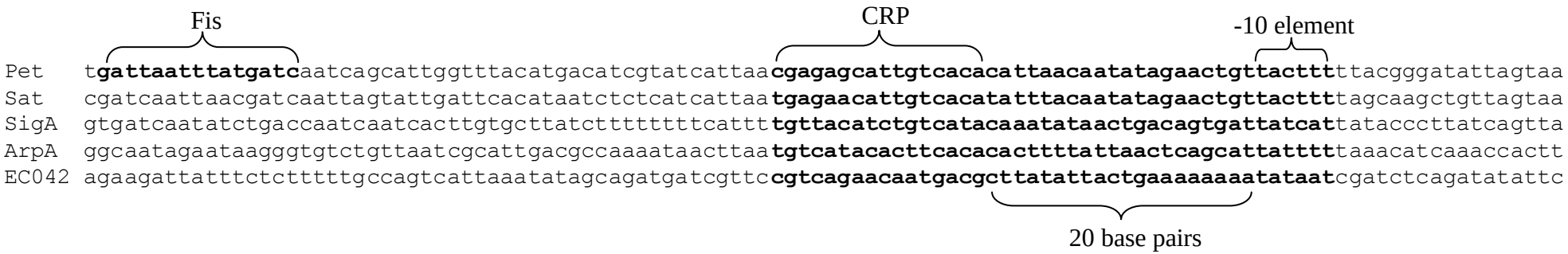


Figure 4.11. ***Identification of other virulence gene promoters, from Shigella spp., UPEC and EAEC, containing similar promoter architecture to pet.*** The putative Type VI secretion system protein is designated EC042-0229. DNA sequences shown in this figure have been retrieved from the open source database Xbase (Chaudhuri *et al.*, 2008).

CHAPTER 5

Regulation of the *sat* promoter by CRP and Fis

5.1 Introduction

Although we were able to re-capitulate the mechanism of co-dependence on CRP and Fis at the *pet* promoter in the semi-synthetically derived CC series of promoters (Chapter 4, section 4.2.8), it is still unclear whether this mechanism is restricted to the *pet* promoter or whether many different naturally occurring promoters employ this strategy. An additional question is raised as to whether the integration of upstream bound factors, other than Fis, can rescue CRP from a non-optimal position. This notion could be tested at the various promoters identified in Chapter 4, Figure 4.11, where similar promoter architecture is apparent with respect to CRP placed at a non-optimal position, but lack an upstream binding site for Fis. However, for this work we have focussed on just one of the promoters identified, which regulates the secreted autotransporter (Sat) gene, encoding a cytotoxin from uropathogenic *E. coli* (UPEC) CFT073. UPEC is the major causative organism of urinary tract infections (UTIs). UTI manifests itself as two clinically recognised conditions; cystitis is the infection of the urethra and bladder, whereas a more severe complication is known as acute pyelonephritis, where the infection ascends to the kidneys (Nielubowicz and Mobley, 2010). UPEC have evolved their characteristic fimbriae (P and type 1 fimbriae) in order to circumvent the obstacles associated with the strong urine flow of the urinary tract (Lane and Mobley, 2007). Uropathogens may also damage their host through the elaboration of toxins. One of these toxins, Sat, was identified from 68% of strains associated with acute pyelonephritis (Guyer *et al.*, 2000) and shares significant homology to Pet, containing the typical tripartite architecture and characteristics of all AT proteins. Sat is also functionally related to Pet in that it internalises and causes cytotoxic effects on bladder and kidney epithelial cells (Guyer *et al.*, 2002). Although ATs encompass a wide array of divergent functions, some ATs evidently display functional similarities and, in the case of the IgA1 proteases and SPATEs, are evolutionary

related (Henderson *et al.*, 2004). Therefore, it could be envisaged that at least some ATs display conserved regulatory mechanisms. In this chapter, the focus will be to study the regulation of *sat*, thus shedding light on the pathogenesis of UPEC, and additionally assess whether the gene encoding this cytotoxin is regulated by CRP and Fis in a manner similar to *pet*.

5.2 Results

5.2.1 *sat* gene regulatory region

Bioinformatic analyses suggested that the *sat* promoter possessed a similar regulatory organisation to *pet*. Manual inspection of the nucleotide sequence confirmed the presence of putative DNA sites for CRP and Fis with similar spacing observed for *pet*. Given the conservation of the DNA sites for Fis and CRP binding between the *pet* and *sat* promoters, we hypothesised that *sat* was regulated in a similar manner, with CRP and Fis being the key activators of transcription initiation. In order to study the regulation of *sat*, we first cloned an *Eco*RI-*Hin*dIII fragment (SAT; Fig. 5.1), which contained the *sat* gene regulatory region, into the pRW50 low copy number *lac* expression vector to create a fusion of the *sat* gene promoter to the *lac*Z gene. To confirm a role for the transcription factors CRP and Fis *in vivo*, the resulting recombinant (SAT/pRW50) was transformed into the *E. coli* K-12 Δ*lac* strain BW25113, and derivatives with disrupted *crp* or *fis* genes. Measurements of β-galactosidase expression in the different strains showed that *sat* promoter activity in the *crp* and *fis* mutants was 15% and 60% of the activity in the *E. coli* BW25113 parental strain, respectively (Fig. 5.2, A). Surprisingly, *sat* promoter activity differs significantly from *pet* promoter activity in the *fis* knockout strain (60% and 25%, respectively). Surprisingly, the promoter activity of *sat* is 2.5-fold lower than that of the *pet* promoter (AER1) (Fig. 5.2, B), which may be due to

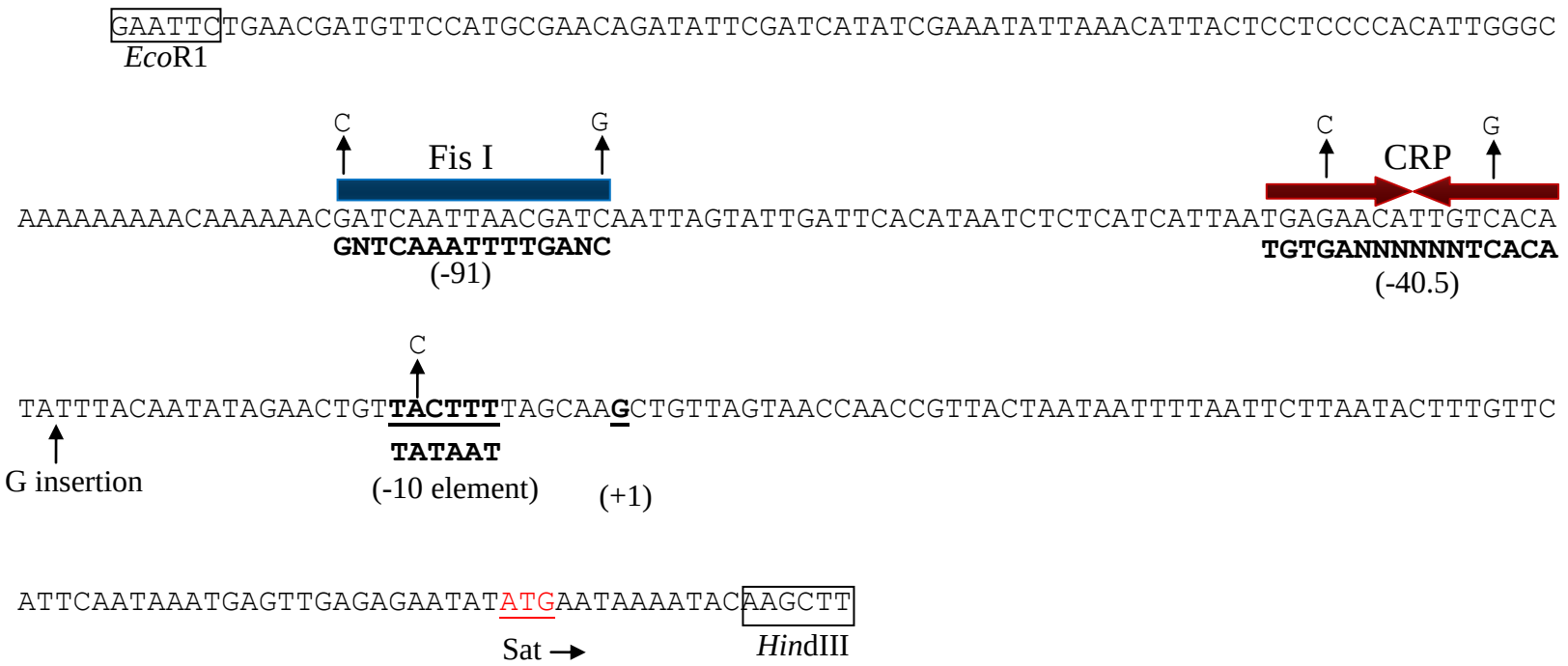
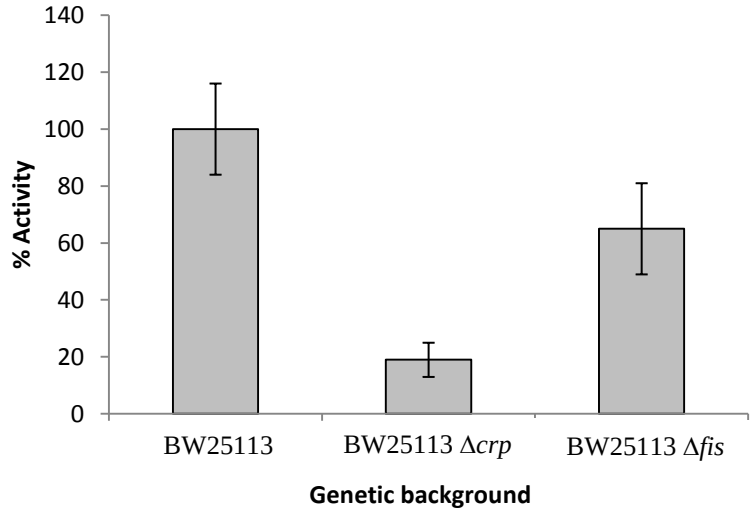


Figure 5.1. ***Nucleotide sequence of the SAT fragment containing the sat gene regulatory region***. The predicted Sat ATG start codon is underlined and indicated with a horizontal arrow. The predicted transcript start site is labelled +1. Both the +1 start site and the -10 element are highlighted in bold and underlined. Predicted DNA sites for CRP and Fis and the promoter -10 hexamer element are aligned with consensus binding sequences, shown in bold type. The predicted CRP site is indicated with red shaded arrows, the predicted Fis sites are indicated with blue shaded blocks, and the centre of each site is numbered with respect to the transcript start. Vertical arrows indicate mutations in different targets that are studied in this work.

A



B

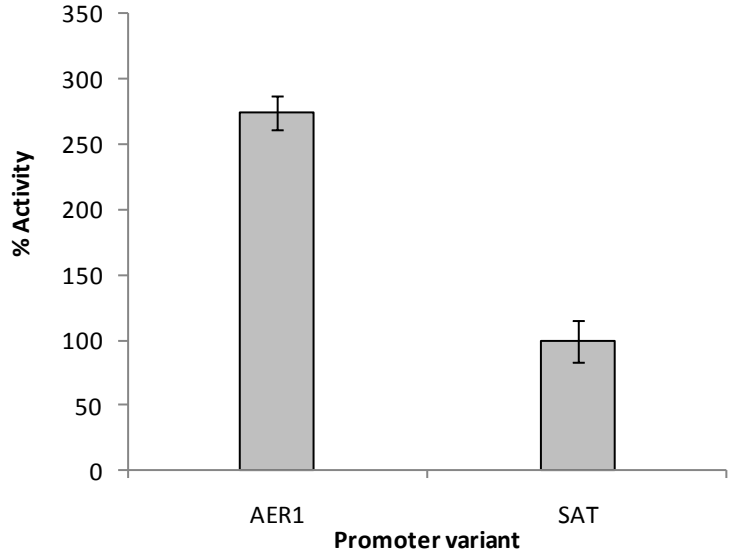


Figure 5.2. ***Promoter activities of pet vs sat.*** **A.** This figure illustrates β-galactosidase activities measured in *E. coli* BW25113 *crp* and *fis* cells carrying pRW50 containing the SAT promoter. **B.** The figure shows β-galactosidase activities measured in *E. coli* BW25113, containing the *pet* (AER1) and *sat* (SAT) promoters in pRW50. Cells were grown aerobically in Luria–Bertani broth at 37°C and harvested during early log phase. β-galactosidase activities are expressed as a percentage of the activity with the *sat* promoter. Each activity is the average of three independent experiments.

differences in promoter sequences between the *pet* and *sat* around the transcript start site (Fig. 4.11).

5.2.2 Mutational analysis of the *sat* promoter

The experiments detailed above confirm a role for CRP and Fis in the regulation of the *sat* promoter. In order to validate the assignment of the CRP and Fis binding sites and the -10 hexamer element identified in Figure 5.1, site-directed mutagenesis was used to target nucleotides within these sites (as detailed in Chapter 2, section 2.4.7.2) and the mutant *Eco*R1-*Hin*dIII SAT promoter derivatives were cloned into pRW50. The resultant recombinants were transformed into *E. coli* BW25113 and the promoter activity of each mutant derivative was deduced from β-galactosidase measurements. The data in Figure 5.3 demonstrate that *sat* promoter activity is significantly reduced by mutations in the -10 element and the CRP and Fis binding sites, suggesting that the *sat* promoter, like *pet*, is co-dependent on CRP and Fis for optimal activity.

5.2.3 CRP and Fis binding at the *sat* promoter

Having established that mutation of the predicted CRP and Fis binding sites have a demonstrable role in *sat* regulation *in vivo*, an electrophoretic mobility shift assay (EMSA) was used to investigate the binding of CRP and Fis at the *sat* promoter. The *Eco*R1-*Hin*dIII fragments carrying the SAT promoter and its mutant derivatives, SAT-fis and SAT-crp, were cloned into pSR, as described in Chapter 2, section 2.4.7. The resultant recombinants were digested and labelled in order to generate fragments suitable for EMSA's. Results in Figure 5.4 show that incubation of labelled SAT fragment with Fis results in the formation of two shifted species, suggestive of two Fis binding sites within the *sat* promoter. Incubation of labelled SAT-fis fragment with Fis results in the loss of one Fis-DNA species, indicating that

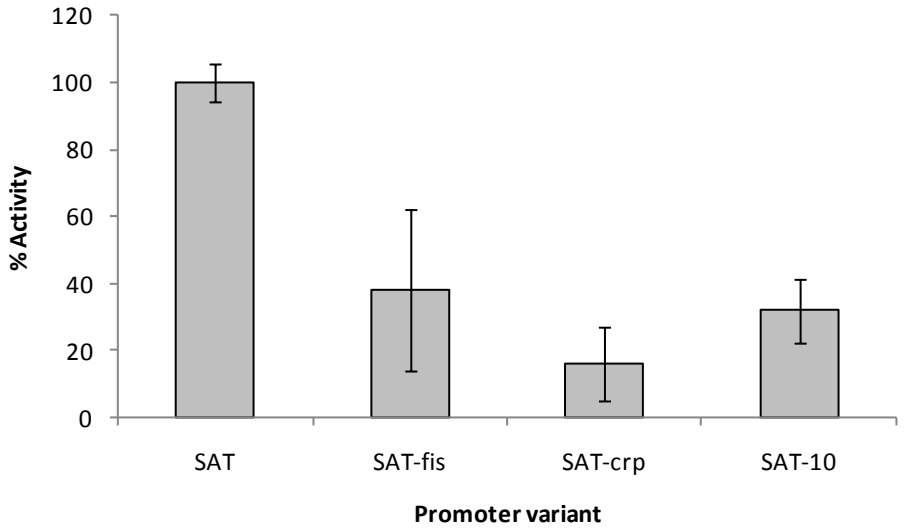


Figure 5.3. ***Mutational analysis of the sat promoter.*** The figure illustrates measured β-galactosidase activities in *E. coli* BW25113, containing a set of pRW50 constructs carrying the *sat* promoter (SAT) and mutant derivatives. The SAT mutant derivatives were as follows; SAT, starting *sat* promoter; SAT-fis, *sat* promoter with mutated Fis site; SAT-crp, *sat* promoter with mutated DNA site for CRP; SAT-10, *sat* promoter with mutated -10 element (Fig. 5.1). Promoter activity is diminished when the Fis, CRP and-10 element are mutated. Cells were grown aerobically in Luria–Bertani broth at 37°C and harvested during early log phase. β-galactosidase activities are expressed as a percentage of the activity with the starting *sat* (SAT) promoter. Each activity is the average of three independent experiments.

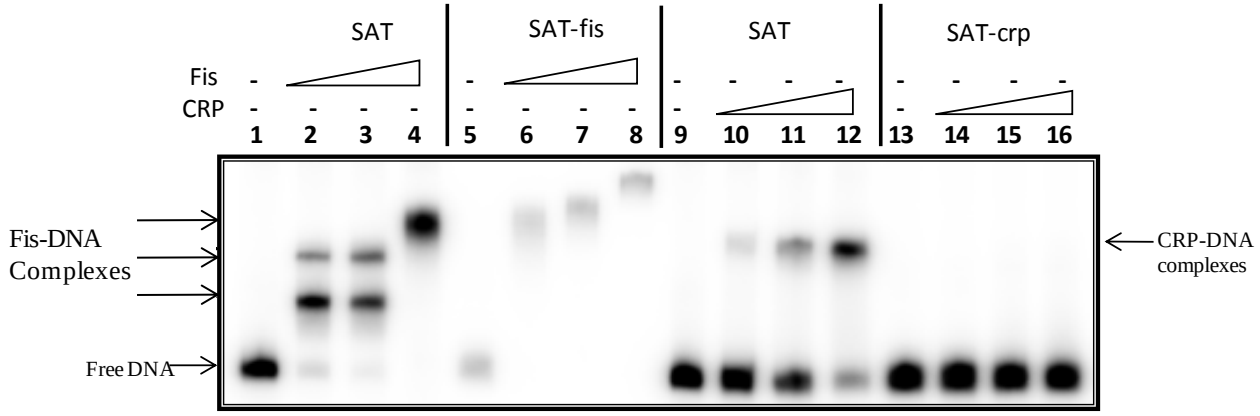


Figure 5.4. ***CRP and Fis binding at the sat promoter.*** Labelled *Eco*R1-*Hin*dIII promoter fragments (SAT, SAT-fis or SAT-crp) were incubated with increasing concentrations of purified Fis or CRP protein and subjected to analysis by EMSA. Fis concentrations were: lanes 1, 5 and 9-16, no protein; lanes 2 and 6, 100 nM; lanes 3 and 7, 200 nM; lanes 4 and 8, 400 nM. CRP concentrations were: lanes 1–9 and 13, no protein; lanes 10 and 14, 50 nM; lanes 11 and 15, 100 nM; lanes 12 and 16, 200 nM. The positions of the electrophoretically retarded nucleoprotein complexes are indicated by arrows. Note that the low intensity bands present in lanes 5-8 are due to inefficient labelling of the SAT-fis promoter fragment. The high molecular weight species in lanes 6-8 most likely represent Fis binding to a second unidentified Fis site within the SAT promoter.

the Fis site identified at position -91 in the *sat* promoter is correct. However, the second Fis binding site within the *sat* promoter remains uncharacterised. Incubation of the labelled SAT fragment with CRP results in the formation of a single shifted species consistent with one CRP binding site. The lack of this single shifted species when labelled SAT-crp fragment is incubated with CRP confirms that CRP binds to the predicted site at -40.5 (Fig. 5.1).

5.2.4 Fis dependence of the *sat* promoter is due to non-optimal positioning of CRP

Consistent with the *pet* promoter, the *sat* promoter has the key features to create co-dependence on both CRP and Fis for transcription initiation; the non-optimal position of the CRP binding site at -40.5 which generates the atypical spacing of 20 bp between the CRP site and the -10 hexamer element, and the upstream Fis binding site at -91. As inferred from the *pet* promoter in Chapter 4, the co-dependence on Fis for full transcription activation of *pet* can be lost by the insertion of one bp between the CRP binding site and the -10 hexamer element (Chapter 4, Fig 4.8). Following this observation, we re-capitulated this mechanism in a semi-synthetic system (Chapter 4, Fig. 4.9). Therefore, on the basis of bioinformatic searches for similar promoter architectures (Chapter 4, Fig. 4.11), we reasoned that this co-dependence could be an evolutionary conserved mechanism apparent in naturally occurring promoters. To test this, we inserted one bp between the CRP binding site and the -10 hexamer element in the *sat* promoter (Fig. 5.1) and the dependence of the SAT+1 promoter derivative on CRP and Fis was measured using *lac* fusions *in vivo* and an *in vitro* transcription assay. Results in Figure 5.5 clearly demonstrate that, as expected, the *sat* promoter in the SAT fragment is co-dependent on CRP and Fis, however the mutant promoter derivative, SAT+1 loses its dependence on Fis both *in vivo* and *in vitro*, given that full transcription activation is seen in the presence of CRP alone (Fig. 5.5, A and B).

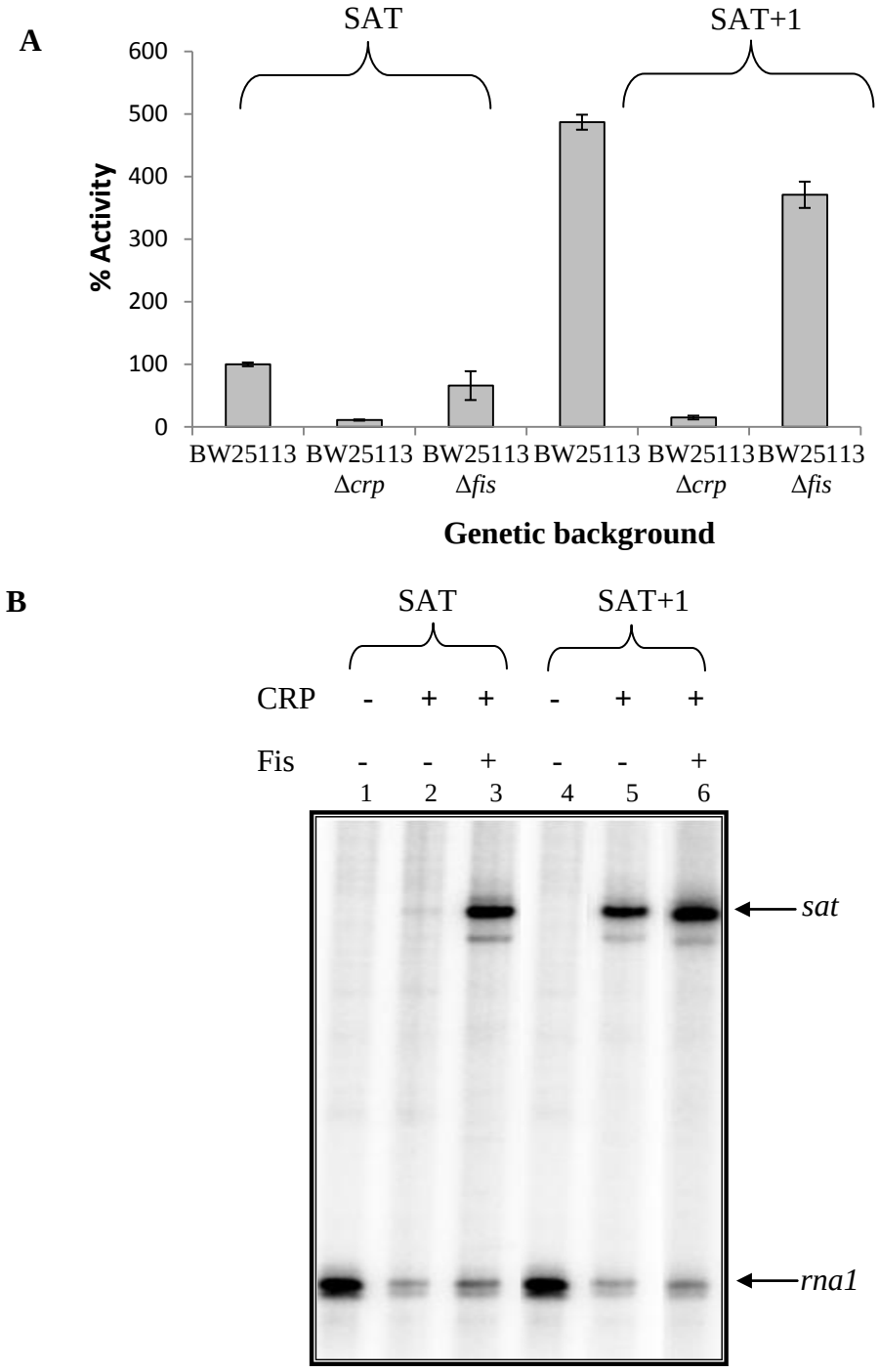


Figure 5.5. ***Optimal spacing between core elements at the sat promoter relieves Fis-dependent CRP activation.*** **A.** The figure shows β-galactosidase activities measured in *E. coli* BW25113, *E. coli* BW25113 *fis* and *E. coli* BW25113 *crp* cells carrying pRW50 containing the *sat* promoter (SAT) or the *sat* promoter derivative with a one base pair insertion (SAT+1). Cells were grown aerobically in Luria–Bertani broth and measurements were taken at early log phase. Activities are expressed as a percentage of the starting SAT promoter activity measured in *E. coli* BW25113 cells and each value derives from three independent experiments. Note that there is a significant increase in SAT+1 promoter activity in *E. coli* BW25113 cells. **B.** *In vitro* transcription of the *sat* promoter (SAT) versus the *sat* derivative with a one base pair insertion (SAT+1). A multi-round *in vitro* transcription assay was performed in the presence of 50 nM RNA polymerase holoenzyme with 400 nM CRP and 200 nM Fis, as indicated. The transcript generated by the *sat* promoter and the control *rna1* transcript are indicated by arrows. Note that the unevenness of *rna1* levels across the gel may be due to limitng RNAp in the presence of the SAT promoter with CRP and/or Fis, such that CRP mediates specific recruitment of RNAp to the SAT promoter, even in the absence of *sat* transcript.

5.3 Discussion

There are two key findings from this work; firstly that the Sat toxin from UPEC CFT073 is co-regulated by CRP and Fis. Secondly, that the mechanism of co-dependence on CRP and Fis is analogous to that described at the *pet* promoter (Chapter 4), whereby the non-optimal positioning of CRP at -40.5 creates co-dependence on upstream bound Fis for full transcription activation. Although this work presents significant findings that this mechanism of co-dependence is evident at two naturally occurring promoters, the homology between Pet and Sat (~90% identical) dictates that this is likely to be the case. However, a feasible extension of these studies would be to study promoters that are distantly related from *pet* and *sat*, to see if CRP misplacement could render other promoters co-dependent on an upstream bound factor, other than Fis. Experiments recently undertaken in our laboratory indicates that replacement of the upstream Fis binding site with an IHF binding site allows for IHF to bind upstream of the non-optimally placed CRP and partially restore promoter activity in the *CC-40.5* promoter of the *CC* promoter derivatives used in Chapter 4, section 4.2.8. This result validates our model suggesting that non-optimally placed CRP may be a strategy to integrate additional upstream bound TFs, other than Fis, for optimal transcription. This model could be investigated further by the study of the promoters identified in Chapter 4, Figure 4.11.

To date, studies have revealed mechanistically distinct gene regulation of ATs (Giangrossi *et al.*, 2010; van der Woude and Henderson, 2008) (Reviewed in Chapter 1, section 1.6.1). This work is the first to describe conservation in the transcriptional regulation of ATs, indicative that this mode of regulation is physiologically significant to both EAEC 042 and UPEC CFT073 pathogenesis, in the intestine and urinary tract, respectively. However, given that these pathogens inhabit remarkably different niches, the clinical significance of this regulation

must be applicable in both settings. As noted in Chapter 4, Fis is a key player in the pathogenesis of EAEC 042, given its demonstrable role in the regulation of *pet* and EAEC specialised fimbriae. This intimate link between *pet* and fimbrial regulation is mimicked in UPEC CFT073, where *sat* and its Type 1 fimbriae are regulated by CRP (Muller *et al.*, 2009). Therefore, further work is needed in the host organism to establish whether additional regulatory mechanisms exist to achieve the systematic expression of fimbriae and toxins that are crucial in establishing infection.

CHAPTER 6

Final Discussion

The findings from this work have provided insights into two broad aspects of microbial biology; namely AT biogenesis and gene regulation. The details of these findings and the impact that they have on our current understanding of these fundamental processes are discussed below.

Autotransporter biogenesis

In recent years, a large number of studies have highlighted the requirement of novel accessory factors in AT biogenesis, challenging the initial supposition that ATs contain all the functional elements necessary for their translocation across the cell envelope. The work described in Chapter 3 builds on these data by the identification of factors involved in Pet biogenesis. Although many of the factors idenitified await further investigation, our work focusses on a subset of these factors and addresses the questions surrounding periplasmic transit and OM translocation of ATs. The supposition that the periplasmic transit of ATs was self-contained was first questioned with the study of IcsA from *S. flexneri* in 2002. This study demonstrated, for the first time, that ATs were accessible to periplasmic chaperones and that IcsA surface presentation was dependent on the periplasmic chaperone, DegP (Purdy *et al.*, 2002). Since this initial description, many reports have detailed the involvement of additional periplasmic enzymes in AT biogenesis, including SurA, Skp and FkpA (Purdy *et al.*, 2007; Ruiz-Perez *et al.*, 2010; Veiga *et al.*, 2004). Our results demonstrate that DegP and SurA are required for Pet secretion and initiated work that provided evidence of putative motifs within ATs that represent binding targets for the chaperones (Ruiz-Perez *et al.*, 2009).

Given that these chaperones DegP, SurA and Skp are intimately involved in the delivery of OMPs to the BAM complex (Hagan *et al.*, 2010), we investigated the role of the BAM complex, in the insertion and/or translocation of ATs into or across the OM. Previous studies

have established a role for the pore-forming component of the complex, BamA, in biogenesis of all examined ATs (Jain and Goldberg, 2007). However, our studies were the first to characterise the role of the other BAM components and demonstrated that BamD, as well as BamA, are essential for the secretion of Pet, whereas the non-essential components BamB, C and E are not required for either secretion or folding of Pet and Ag43 on the cell surface (Rossiter *et al.*, 2011b). Further work is required to identify the exact molecular mechanism of their interactions to address whether the BAM complex is solely required for OM insertion of the AT β-barrel, therefore allowing passenger domain translocation through the cognate β-barrel, or whether BamA is the preferred β-barrel pore for passenger domain translocation to the cell surface. Studies on 'stalled' translocation intermediates (Renn and Clark, 2011) or experiments using the recently developed genetically incorporated *in vivo* cross-linker tool could help to discern a precise role for the periplasmic chaperones and the BAM complex in AT biogenesis (Zhang *et al.*, 2011).

The identification of an increasing number of accessory factors generates more questions regarding their precise molecular involvement in AT biogenesis, highlighting the gaps in our understanding of this secretion system. Thus, the answers to some questions still remain contentious; how exactly is AT gene expression regulated? What are the environmental signals that trigger gene expression within the host organism? What are the precise molecular interactions that govern AT translocation across the periplasm by periplasmic chaperones? How are the cognate β-barrels inserted into the OM? Do passenger domains pass through their own β-barrel or are they translocated through the BamA β-barrel? As time constraints prevented further investigations into the role of the additional factors identified from the transposon screen of EAEC 042, future studies focusing on their role in AT biogenesis could

help to answer these questions and build on the paradigm shift towards a much more complex mechanism of AT secretion than initially suggested.

Global regulators CRP and Fis

The work outlined in Chapter 4 and 5 details the regulation of the *pet* and *sat* promoter by CRP and Fis. The two key findings from this work are; a) that the regulation of at least two functionally similar ATs occurs by a conserved activation mechanism and; b) the way in which non-optimal positioning of CRP generates co-dependence on upstream bound Fis contradicts the typical mechanism seen at all Class II CRP-dependent promoters studied to date (Zheng *et al.,* 2004).

CRP and Fis are perhaps the most well studied bacterial TFs and their role in the modulation of over 200 promoters has earned them the title of 'global' regulators. However, in most instances their effect is seen at promoters that control the expression of metabolic and housekeeping genes, yet in this post-genomic era and with the development of extensive bioinformatics tools, the role of CRP and Fis in the regulation of virulence determinants is becoming ever more prevalent. The observation that *pet* and *sat*, genes encoding cytotoxins from the two distinct pathogens, Enteroaggregative *E. coli* (EAEC) and Uropathogenic *E. coli* (UPEC), share a conserved activation mechanism raises the idea that this regulation must confer some physiological advantage in the infection of their host organism. Fis is regulated in a growth-phase dependent manner, such that the highest levels of Fis are found in rapidly growing cells, while Fis is undetectable in non-growing cells (Mallik *et al.,* 2006; Morin *et al.,* 2010). In contrast, CRP activity is controlled by both its expression and the level of the cyclic AMP second messenger (Grainger and Busby, 2008), thus in *E. coli*, CRP levels are known to be raised as cell growth slows. Given that both *pet* and *sat* are regulated by CRP

and Fis, the result of this is likely to be that the expression in a mammalian host is restricted to certain phases of growth or certain niches. A feasible role for Fis may become relevant to the expression of Pet in the host intestine. Passage through the stomach and proximal small intestine will expose EAEC to favourable conditions high in nutrients causing bacteria to grow exponentially. This will consequently trigger the activities of Fis and stimulate expression of both EAECs characteristic AAF Fimbriae (Sheikh *et al.*, 2001) and Pet. The role of CRP may be to sustain the expression of these toxins once Fis levels are depleted at stationary phase of growth. Further work is required to substantiate the role of CRP and Fis in the natural setting, and to highlight any additional levels of regulation that are pathogen-specific. For example, recent observations have highlighted a possible role for AggR (the master virulence regulator of EAEC) on *pet* expression, given that a putative AggR binding site was found to map within the Fis site I of the AER1 fragment. Thus, this level of regulation could not occur in the *E. coli* K-12 strain used in these studies.

A study that investigated the CRP regulon, identified many new targets for CRP and found that most of these targets were activated by the simple Class II mechanism (Zheng *et al.*, 2004). In chapter 4, using well characterised CRP mutants, we showed that the *pet* promoter is also Class II activated but that it is organised in an atypical manner. All Class II activated promoters contain a single DNA site for CRP centred at -41.5, however at the *pet* promoter the site is centred at -40.5 and our experiments show that this non-optimal spacing creates dependence on upstream bound Fis for full activation of *pet*. An extension of these observations detailed in Chapter 5, show that this non-optimal positioning of CRP also occurs at the *sat* promoter. Therefore, it is possible that the non-optimal positioning of CRP could be a mechanism seen at many other promoters, which awaits investigation by future studies. However, recent studies communicated in our laboratory show that the replacement of the Fis

binding site with an IHF binding site can rescue CRP from its non-optimal position and partially restore promoter activity. Thus, this hints that IHF and Fis are working in a similar fashion to rescue misplaced CRP, perhaps by bending DNA. However, further work is needed to elucidate the precise mechanism that upstream bound Fis uses to co-activate the *pet* promoter. More specifically, investigations should initially determine whether the ability of Fis to interact directly with RNA polymerase or its propensity to bend DNA is the underlying mechanism causing this co-dependence.

In conclusion, our investigations have revealed novel factors that are involved in the secretion of autotransporters by Gram-negative pathogens. However, future studies are still obliged to resolve the precise molecular interactions that underpin AT biogenesis, particularly concerning their translocation across the OM. Furthermore, with the progression of high-throughput technologies, our work highlights the importance of studying single gene promoters to uncover new ways in which TFs operate, which will ultimately re-define the parameters for *in silico* promoter predictions.

APPENDIX 1

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 chromosome, and for generation of mutants by a method based on the λ-red system.

Name	***Sequence (5'→3')***	***Use***
ackA_F	ggtgtcatcatgcgctacgc	Upstream primer for amplification of named chromosomal loci.
ackA_R	cagccgacgctggttccgg	Downstream primer for amplification of named chromosomal loci.
adiY_F	gcgcggccccaccccgcgct	Upstream primer for amplification of named chromosomal loci.
adiY_R	ccgggttatcgctgatgccg	Downstream primer for amplification of named chromosomal loci.
ais_F	gaggcattttgccgacatcg	Upstream primer for amplification of named chromosomal loci.
ais_R	cgttacgtttgaaaggtcttc	Downstream primer for amplification of named chromosomal loci.
cadA_F	ccacaccgcgtctaacgcac	Upstream primer for amplification of named chromosomal loci.
cadA_R	cgcgcgcggctgtgagggtg	Downstream primer for amplification of named chromosomal loci.
celB_F	gtgcttaaggctgcggttgc	Upstream primer for amplification of named chromosomal loci.
celB_R	gcttccgtttgcgtatcggg	Downstream primer for amplification of named chromosomal loci.
clpB_F	cgtcgcgacaagaccaccgg	Upstream primer for amplification of named chromosomal loci.
clpB_R	gggagttattccggcctgac	Downstream primer for amplification of named chromosomal loci.
crp_F	ggacgtcacattaccgtgcag	Upstream primer for amplification of named chromosomal loci.
crp_R	gataaatcagtctgcgccac	Downstream primer for amplification of named chromosomal loci.
crr_F	cgacaacggacgagttaatg	Upstream primer for amplification of named chromosomal loci.
crr_R	gctacacccagcagcatgag	Downstream primer for amplification of named chromosomal loci.
cspG_F	gagtctctctcctgaacac	Upstream primer for amplification of named chromosomal loci.
cspG_R	catcgaatgccgatgtgcgg	Downstream primer for amplification of named chromosomal loci.
cysK_F	ccgcatattctctgagcggg	Upstream primer for amplification of named chromosomal loci.
cysK_R	gcgaaagtttgaagcaggcc	Downstream primer for amplification of named chromosomal loci.
degP_F	ggaacttcaggctataaaac	Upstream primer for amplification of named chromosomal loci.
degP_R	gaagatgtatggagttgtgg	Downstream primer for amplification of named chromosomal loci.

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 and for generation of mutants by a method based on the λ-red system (continued).

Name	***Sequence (5'→3')***	***Use***
dnaG_F	ggctcgcgaaaacgcacgcc	Upstream primer for amplification of named chromosomal loci.
dnaG_R	cggctgtcgggggcttcccg	Downstream primer for amplification of named chromosomal loci.
fis_F	gccagcgaacagctggaggc	Upstream primer for amplification of named chromosomal loci.
fis_R	cctgttctcatggtcactcc	Downstream primer for amplification of named chromosomal loci.
fkpA_F	gacaacgctttatagtaccc	Upstream primer for amplification of named chromosomal loci.
fkpA_R	ctaaattaatacagcggagg	Downstream primer for amplification of named chromosomal loci.
folX_F	gttcgcgccctgccaccggg	Upstream primer for amplification of named chromosomal loci.
folX_R	caaacgcggaatcaaatggc	Downstream primer for amplification of named chromosomal loci.
lon_F	cccatatactgacgtacatg	Upstream primer for amplification of named chromosomal loci.
lon_R	cccgaattagcctgccagccc	Downstream primer for amplification of named chromosomal loci.
mhpC_F	gcaactggcgtagcgaaggg	Upstream primer for amplification of named chromosomal loci.
mhpC_R	gcaatcgcttcgccctgctc	Downstream primer for amplification of named chromosomal loci.
nei_F	cgctctgcatttgccgaaaagg	Upstream primer for amplification of named chromosomal loci.
nei_R	ccgatatggcgctcattatggc	Downstream primer for amplification of named chromosomal loci.
nlpB_F	gggcatgtaaggaactg	Upstream primer for amplification of named chromosomal loci.
nlpB_R	ctaaaatcgcgttttcctgc	Downstream primer for amplification of named chromosomal loci.
nuoF_F	gctgctgccaacttgctgcc	Upstream primer for amplification of named chromosomal loci.
nuoF_R	gtctacatgaattgtagcca	Downstream primer for amplification of named chromosomal loci.
pal_F	cagatgggcgtttcaaagcg	Upstream primer for amplification of named chromosomal loci.
pal_R	cctgagcaaaagcggcccag	Downstream primer for amplification of named chromosomal loci.
pitA_F	gagtgaaatccatacagggg	Upstream primer for amplification of named chromosomal loci.
pitA_R	ggcgttcacgccgcatccgg	Downstream primer for amplification of named chromosomal loci.

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 and for generation of mutants by a method based on the λ-red system (continued)

Name	***Sequence (5'→3')***	***Use***
pqiB_F	gctatgacgtttgacccgcg	Upstream primer for amplification of named chromosomal loci.
pqiB_R	gcatccggccagccacagtg	Downstream primer for amplification of named chromosomal loci.
prmA_F	ggaaaccgtttcggtacatc	Upstream primer for amplification of named chromosomal loci.
prmA_R	gcaaatttgcgtaaataatc	Downstream primer for amplification of named chromosomal loci.
proP_F	ccctttgcggccgtcgcgc	Upstream primer for amplification of named chromosomal loci.
proP_R	ggagagtatgcgcgtcagag	Downstream primer for amplification of named chromosomal loci.
proW_F	cgtggtcgacgaggaccaacag	Upstream primer for amplification of named chromosomal loci.
proW_R	cagtaatgcctttgcccggc	Downstream primer for amplification of named chromosomal loci.
rfaJ_F	gatatagcgcaaaacatatgc	Upstream primer for amplification of named chromosomal loci.
rfaJ_R	cggtaaaaacaaccaagtc	Downstream primer for amplification of named chromosomal loci.
rlpA_F	gtgctgatggctgggttcg	Upstream primer for amplification of named chromosomal loci.
rlpA_R	ccctactatagcaaatgcac	Downstream primer for amplification of named chromosomal loci.
skp_F	ggaatgtagtggtagtgtag	Upstream primer for amplification of named chromosomal loci.
skp_R	ccggtgatgacgatatcgcc	Downstream primer for amplification of named chromosomal loci.
slyB_F	ggattcacatatgccatatac	Upstream primer for amplification of named chromosomal loci.
slyB_R	ctggagcaactgattacgctc	Downstream primer for amplification of named chromosomal loci.
smpA_F	ggagcgagtggtttaccg	Upstream primer for amplification of named chromosomal loci.
smpA_R	ccctctcattgccgatccg	Downstream primer for amplification of named chromosomal loci.
surA_F	gcggtatatgacaacgcaatc	Upstream primer for amplification of named chromosomal loci.
surA_R	gcggagggtgagcggcaaac	Downstream primer for amplification of named chromosomal loci.
thiC_F	cggcgatcgtctcttgcttc	Upstream primer for amplification of named chromosomal loci.
thiC_R	cgtacgcctgcatccaacag	Downstream primer for amplification of named chromosomal loci.

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 and for generation of mutants by a method based on the λ-red system (continued).

Name	***Sequence (5'→3')***	***Use***
uvrD_F	cccggttggcatctctgacc	Upstream primer for amplification of named chromosomal loci.
uvrD_R	cctatccggcctacatgacg	Downstream primer for amplification of named chromosomal loci.
uxuR_F	cgttgtaggccggataaggc	Upstream primer for amplification of named chromosomal loci.
uxuR_R	ccaggaagaatgagtactaac	Downstream primer for amplification of named chromosomal loci.
yafO_F	cgcttccgtccaagtgctgc	Upstream primer for amplification of named chromosomal loci.
yafO_R	ccgctctaatgaaaatagcg	Downstream primer for amplification of named chromosomal loci.
ybbB_F	ggcgcaaatttggcagcggc	Upstream primer for amplification of named chromosomal loci.
ybbB_R	ccggtaactctcaacctttc	Downstream primer for amplification of named chromosomal loci.
ybeL_F	cgcatagcatagcccaaacg	Upstream primer for amplification of named chromosomal loci.
ybeL_R	cggggaatttcattcggccc	Downstream primer for amplification of named chromosomal loci.
ybhC_F	gcaatgccatctggtatcac	Upstream primer for amplification of named chromosomal loci.
ybhC_R	cgaaccgtaggccggataag	Downstream primer for amplification of named chromosomal loci.
yeiC_F	gagcctgctcaaatctgccc	Upstream primer for amplification of named chromosomal loci.
yeiC_R	gattccagcgccacaaccgg	Downstream primer for amplification of named chromosomal loci.
yeiL_F	gaccgggccgatgattgaatc	Upstream primer for amplification of named chromosomal loci.
yeiL_R	gcttgcatccggcaatcgc	Downstream primer for amplification of named chromosomal loci.
yeiM_F	gcgggttaatttctgacggg	Upstream primer for amplification of named chromosomal loci.
yeiM_R	cggagaataaattctccggg	Downstream primer for amplification of named chromosomal loci.
yfbV_F	cccgctgagttgtgaatttg	Upstream primer for amplification of named chromosomal loci.
yfbV_R	gtaacgttcagcatttgccg	Downstream primer for amplification of named chromosomal loci.
yfcI_F	ctgcctgctgcattgggtg	Upstream primer for amplification of named chromosomal loci.
yfcI_R	cgatttagaagaggcgctgg	Downstream primer for amplification of named chromosomal loci.

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 and for generation of mutants by a method based on the λ-red system (continued).

Name	***Sequence (5'→3')***	***Use***
yfeK_F	gtaatccaaccgaaactttac	Upstream primer for amplification of named chromosomal loci.
yfeK_R	gccatagttgacagctaaac	Downstream primer for amplification of named chromosomal loci.
yfgL_F	gcgcgtagtgcatgggaagc	Upstream primer for amplification of named chromosomal loci.
yfgL_R	caacgcacgctatattcgcg	Downstream primer for amplification of named chromosomal loci.
ygeV_F	gatctgtgaaccatcaacgtc	Upstream primer for amplification of named chromosomal loci.
ygeV_R	cacggggtatcagatgattg	Downstream primer for amplification of named chromosomal loci.
ygjI_F	ggttctcaaagtcaccatcg	Upstream primer for amplification of named chromosomal loci.
ygjI_R	ggcagaaactattttccctc	Downstream primer for amplification of named chromosomal loci.
yhiL_F	caattatcttccagggaggg	Upstream primer for amplification of named chromosomal loci.
yhiL_R	catatacaggtgtgggtttc	Downstream primer for amplification of named chromosomal loci.
yhiN_F	ctttcgcggcgcgggcaaag	Upstream primer for amplification of named chromosomal loci.
yhiN_R	caatcacctttccatccacc	Downstream primer for amplification of named chromosomal loci.
yiaG_F	gaatcgaatcatagccagag	Upstream primer for amplification of named chromosomal loci.
yiaG_R	cacgcattggcgggtgatgc	Downstream primer for amplification of named chromosomal loci.
yihN_F	gaatggtcggtatcttattg	Upstream primer for amplification of named chromosomal loci.
yihN_R	gcggcgtaaacgccttatcc	Downstream primer for amplification of named chromosomal loci.
yihR_F	cctaaccctctccccaaagg	Upstream primer for amplification of named chromosomal loci.
yihR_R	ctttagcgagtgaaatgttc	Downstream primer for amplification of named chromosomal loci.
yjdL_F	ggctgatggccgctataccg	Upstream primer for amplification of named chromosomal loci.
yjdL_R	ggtacactgtattatctgcc	Downstream primer for amplification of named chromosomal loci.
yjiC_F	cctacccttaatatcaccg	Upstream primer for amplification of named chromosomal loci.
yjiC_R	catggccgctggataaagtc	Downstream primer for amplification of named chromosomal loci.

Table 1. Primers used for the screening and amplification of loci from the *E.coli* BW25113 and for generation of mutants by a method based on the λ-red system (continued).

Name	***Sequence (5'→3')***	***Use***
ykgG_F	ggccggtgctcatgcggcaagc	Upstream primer for amplification of named chromosomal loci.
ykgG_R	gcataaccatgacggaacgggg	Downstream primer for amplification of named chromosomal loci.
ynbC_F	gcggcgatgtcgtgatgtcag	Upstream primer for amplification of named chromosomal loci.
ynbC_R	gtgaagaagaaaaacggggc	Downstream primer for amplification of named chromosomal loci.
yohM_F	ctactgggggggtagtatcagg	Upstream primer for amplification of named chromosomal loci.
yohM_R	ccttagcgagcagagagtcag	Downstream primer for amplification of named chromosomal loci.
yphF_F	gacggcgattaatccggcttg	Upstream primer for amplification of named chromosomal loci.
yphF_R	cacttttgctaccgggactg	Downstream primer for amplification of named chromosomal loci.
yqiH_F	caacttaaccaccggaacgg	Upstream primer for amplification of named chromosomal loci.
yqiH_R	cattgcccatgcaggtaatg	Downstream primer for amplification of named chromosomal loci.
yshA_F	caatagccgctatttccatc	Upstream primer for amplification of named chromosomal loci.
yshA_R	ccgtatcaaccgtaggccgg	Downstream primer for amplification of named chromosomal loci.
kan_F	cctgcaaagtaaactggatg	Primer that anneals to the internal of the kanamycin cassette on the top strand
kan_R	catgctcttcgtgcagatca	Primer that anneals to the internal of the kanamycin cassette on the bottom strand

APPENDIX 2

Table 2 Primers used for the screening of candidate kanamycin resistant colonies after chromosomal disruption of target genes in *E. coli* HB101 using the λ-red system.

Primer	***Sequence (5' →3')***	***Use***
surA_seq	gccgactctatgttaggtg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
degP_seq	cgcgcttattccacaaactc	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
skp_seq	ccgtatgtctgcgggtatcg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
ppiA_seq	ccagtcagtgcgtaggagag	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
ppiD_seq	gcttccgtaactgaatctctg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
bamB_seq	gccgcccagttacaacaggg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
bamC_seq	gcgatgatgcgagcgcgctg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
bamE_seq	gccacaagtcgcgggatgtg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
crp_seq	cccttcgacccacttcactcg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.
fis_seq	gctttgcgcgcacgttcggg	Primer anneals approximately 100 bp upstream of the loci used to disrupt target chromosomal gene. Used in conjunction with a downstream primer that anneals to the kanamycin cassette.

BIBLIOGRAPHY

Akhter, Y., Yellaboina, S., Farhana, A., Ranjan, A., Ahmed, N. and Hasnain, S. E. (2008) Genome scale portrait of cAMP-receptor protein (CRP) regulons in mycobacteria points to their role in pathogenesis. *Gene* **407**: 148-58.

Al-Hasani, K., Navarro-Garcia, F., Huerta, J., Sakellaris, H. and Adler, B. (2009) The immunogenic SigA enterotoxin of *Shigella flexneri* 2a binds to HEp-2 cells and induces fodrin redistribution in intoxicated epithelial cells. *PLoS One* **4**: e8223.

Azam, T. A. and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**: 33105-13.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006 0008.

Ballister, E. R., Lai, A. H., Zuckermann, R. N., Cheng, Y. and Mougous, J. D. (2008) In vitro self-assembly of tailorable nanotubes from a simple protein building block. *Proc Natl Acad Sci U S A* **105**: 3733-8.

Barnard, A., Wolfe, A. and Busby, S. (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. *Curr Opin Microbiol* **7**: 102-8.

Barnard, T. J., Dautin, N., Lukacik, P., Bernstein, H. D. and Buchanan, S. K. (2007) Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat Struct Mol Biol* **14**: 1214-20.

Barne, K. A., Bown, J. A., Busby, S. J. and Minchin, S. D. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* **16**: 4034-40.

Beatty, C. M., Browning, D. F., Busby, S. J. and Wolfe, A. J. (2003) Cyclic AMP receptor protein-dependent activation of the *Escherichia coli acsP2* promoter by a synergistic class III mechanism. *J Bacteriol* **185**: 5148-57.

Behrens, S., Maier, R., de Cock, H., Schmid, F. X. and Gross, C. A. (2001) The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *EMBO J* **20**: 285-94.

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J. and Busby, S. (1990) Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res* **18**: 7243-50.

Belyaeva, T. A., Rhodius, V. A., Webster, C. L. and Busby, S. J. (1998) Transcription activation at promoters carrying tandem DNA sites for the *Escherichia coli* cyclic AMP receptor protein: organisation of the RNA polymerase alpha subunits. *J Mol Biol* **277**: 789-804.

Bernard, C. S., Brunet, Y. R., Gueguen, E. and Cascales, E. (2010) Nooks and crannies in type VI secretion regulation. *J Bacteriol* **192**: 3850-60.

Betancourt-Sanchez, M. and Navarro-Garcia, F. (2009) Pet secretion, internalization and induction of cell death during infection of epithelial cells by enteroaggregative *Escherichia coli*. *Microbiology* **155**: 2895-906.

Blatter, E. E., Ross, W., Tang, H., Gourse, R. L. and Ebright, R. H. (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**: 889-96.

Bokal, A. J., Ross, W., Gaal, T., Johnson, R. C. and Gourse, R. L. (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the *Escherichia coli rrnB* P1 promoter. *EMBO J* **16**: 154-62.

Bos, M. P. and Tommassen, J. (2004) Biogenesis of the Gram-negative bacterial outer membrane. *Curr Opin Microbiol* **7**: 610-6.

Braun, V. (1975) Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim Biophys Acta* **415**: 335-77.

Breyton, C., Haase, W., Rapoport, T. A., Kuhlbrandt, W. and Collinson, I. (2002) Three-dimensional structure of the bacterial protein-translocation complex SecYEG. *Nature* **418**: 662-5.

Brown, N. L., Stoyanov, J. V., Kidd, S. P. and Hobman, J. L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27**: 145-63.

Browning, D., Savery, N., Kolb, A. and Busby, S. (2009) Assays for transcription factor activity. *Methods Mol Biol* **543**: 369-87.

Browning, D. F. and Busby, S. J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57-65.

Browning, D. F., Grainger, D. C. and Busby, S. J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol* **13**: 773-80.

Bruand, C. and Ehrlich, S. D. (2000) UvrD-dependent replication of rolling-circle plasmids in *Escherichia coli*. *Mol Microbiol* **35**: 204-10.

Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. and Gralla, J. D. (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* **182**: 4129-36.

Burgess, R. R., Travers, A. A., Dunn, J. J. and Bautz, E. K. (1969) Factor stimulating transcription by RNA polymerase. *Nature* **221**: 43-6.

Burns, D. L. (2003) Type IV transporters of pathogenic bacteria. *Curr Opin Microbiol* **6**: 29-34.

Busby, S. and Ebright, R. H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**: 743-6.

Busby, S. and Ebright, R. H. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199-213.

Canizalez-Roman, A. and Navarro-Garcia, F. (2003) Fodrin CaM-binding domain cleavage by Pet from enteroaggregative *Escherichia coli* leads to actin cytoskeletal disruption. *Mol Microbiol* **48**: 947-58.

Caron, M. P., Lafontaine, D. A. and Masse, E. (2010) Small RNA-mediated regulation at the level of transcript stability. *RNA Biol* **7**: 140-4.

Casino, P., Rubio, V. and Marina, A. (2010) The mechanism of signal transduction by two-component systems. *Curr Opin Struct Biol* **20**: 763-71.

Charles, M., Perez, M., Kobil, J. H. and Goldberg, M. B. (2001) Polar targeting of *Shigella* virulence factor IcsA in Enterobacteriacae and *Vibrio*. *Proc Natl Acad Sci U S A* **98**: 9871-6.

Charlson, E. S., Werner, J. N. and Misra, R. (2006) Differential effects of *yfgL* mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *J Bacteriol* **188**: 7186-94.

Chaudhuri, R. R., Loman, N. J., Snyder, L. A., Bailey, C. M., Stekel, D. J. and Pallen, M. J. (2008) xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Res* **36**: D543-6.

Chen, R. and Henning, U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol Microbiol* **19**: 1287-94.

Cheng, L. W. and Schneewind, O. (2000) Type III machines of Gram-negative bacteria: delivering the goods. *Trends Microbiol* **8**: 214-20.

Chevalier, N., Moser, M., Koch, H. G., Schimz, K. L., Willery, E., Locht, C., Jacob-Dubuisson, F. and Muller, M. (2004) Membrane targeting of a bacterial virulence factor harbouring an extended signal peptide. *J Mol Microbiol Biotechnol* **8**: 7-18.

Choy, H. E. and Adhya, S. (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc Natl Acad Sci U S A* **89**: 11264-8.

Christiansen, J. K., Larsen, M. H., Ingmer, H., Sogaard-Andersen, L. and Kallipolitis, B. H. (2004) The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. *J Bacteriol* **186**: 3355-62.

Clantin, B., Delattre, A. S., Rucktooa, P., Saint, N., Meli, A. C., Locht, C., Jacob-Dubuisson, F. and Villeret, V. (2007) Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science* **317**: 957-61.

Costerton, J. W., Ingram, J. M. and Cheng, K. J. (1974) Structure and function of the cell envelope of gram-negative bacteria. *Bacteriol Rev* **38**: 87-110.

Cotter, S. E., Surana, N. K., Grass, S. and St Geme, J. W., 3rd (2006) Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *J Bacteriol* **188**: 5400-7.

Cristobal, S., de Gier, J. W., Nielsen, H. and von Heijne, G. (1999) Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**: 2982-90.

Curtiss, R., 3rd, Goldschmidt, R. M., Fletchall, N. B. and Kelly, S. M. (1988) Avirulent *Salmonella typhimurium* delta *cya* delta *crp* oral vaccine strains expressing a *streptococcal* colonization and virulence antigen. *Vaccine* **6**: 155-60.

Darwin, K. H. and Miller, V. L. (1999) InvF is required for expression of genes encoding proteins secreted by the SPI1 type III secretion apparatus in *Salmonella typhimurium*. *J Bacteriol* **181**: 4949-54.

Datsenko, K. A. and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-5.

Dautin, N., Barnard, T. J., Anderson, D. E. and Bernstein, H. D. (2007) Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J* **26**: 1942-52.

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

Dautin, N. and Bernstein, H. D. (2011) Residues in a conserved alpha-helical segment are required for cleavage but not secretion of an *Escherichia coli* serine protease autotransporter passenger domain. *J Bacteriol* **193**: 3748-56.

De, E., Saint, N., Glinel, K., Meli, A. C., Levy, D. and Jacob-Dubuisson, F. (2008) Influence of the passenger domain of a model autotransporter on the properties of its translocator domain. *Mol Membr Biol* **25**: 192-202.

de Gier, J. W. and Luirink, J. (2001) Biogenesis of inner membrane proteins in *Escherichia coli*. *Mol Microbiol* **40**: 314-22.

deHaseth, P. L., Zupancic, M. L. and Record, M. T., Jr. (1998) RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J Bacteriol* **180**: 3019-25.

Delattre, A. S., Clantin, B., Saint, N., Locht, C., Villeret, V. and Jacob-Dubuisson, F. (2010) Functional importance of a conserved sequence motif in FhaC, a prototypic member of the TpsB/Omp85 superfamily. *FEBS J* **277**: 4755-65.

Desvaux, M., Cooper, L. M., Filenko, N. A., Scott-Tucker, A., Turner, S. M., Cole, J. A. and Henderson, I. R. (2006) The unusual extended signal peptide region of the type V secretion system is phylogenetically restricted. *FEMS Microbiol Lett* **264**: 22-30.

Desvaux, M., Parham, N. J., Scott-Tucker, A. and Henderson, I. R. (2004) The general secretory pathway: a general misnomer? *Trends Microbiol* **12**: 306-9.

Desvaux, M., Scott-Tucker, A., Turner, S. M., Cooper, L. M., Huber, D., Nataro, J. P. and Henderson, I. R. (2007) A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism. *Microbiology* **153**: 59-70.

Dillon, S. C. and Dorman, C. J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8**: 185-95.

Ding, Y., Davis, B. M. and Waldor, M. K. (2004) Hfq is essential for *Vibrio cholerae* virulence and downregulates sigma expression. *Mol Microbiol* **53**: 345-54.

Doerrler, W. T. and Raetz, C. R. (2005) Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J Biol Chem* **280**: 27679-87.

Dorel, C., Lejeune, P. and Rodrigue, A. (2006) The Cpx system of *Escherichia coli*, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? *Res Microbiol* **157**: 306-14.

Dorman, C. J. and Kane, K. A. (2009) DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS Microbiol Rev* **33**: 587-92.

Driessen, A. J. and Nouwen, N. (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* **77**: 643-67.

Dudley, E. G., Thomson, N. R., Parkhill, J., Morin, N. P. and Nataro, J. P. (2006) Proteomic and microarray characterization of the AggR regulon identifies a *pheU* pathogenicity island in enteroaggregative *Escherichia coli*. *Mol Microbiol* **61**: 1267-82.

Ebright, R. H. (2000) RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol* **304**: 687-98.

Ebright, R. H. and Busby, S. (1995) The *Escherichia coli* RNA polymerase alpha subunit: structure and function. *Curr Opin Genet Dev* **5**: 197-203.

Ebright, R. H., Ebright, Y. W. and Gunasekera, A. (1989) Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli lac* DNA site. *Nucleic Acids Res* **17**: 10295-305.

Emmer, M., deCrombrugghe, B., Pastan, I. and Perlman, R. (1970) Cyclic AMP receptor protein of *E. coli*: its role in the synthesis of inducible enzymes. *Proc Natl Acad Sci U S A* **66**: 480-7.

Emsley, P., Charles, I. G., Fairweather, N. F. and Isaacs, N. W. (1996) Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* **381**: 90-2.

Escobar, M. A. and Dandekar, A. M. (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci* **8**: 380-6.

Eslava, C., Navarro-Garcia, F., Czeczulin, J. R., Henderson, I. R., Cravioto, A. and Nataro, J. P. (1998) Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun* **66**: 3155-63.

Espert, S. M., Elsinghorst, E. A. and Munson, G. P. (2011) The *tib* adherence locus of enterotoxigenic *Escherichia coli* is regulated by cyclic AMP receptor protein. *J Bacteriol* **193**: 1369-76.

Farizo, K. M., Huang, T. and Burns, D. L. (2000) Importance of holotoxin assembly in Ptl-mediated secretion of pertussis toxin from *Bordetella pertussis*. *Infect Immun* **68**: 4049-54.

Fass, E. and Groisman, E. A. (2009) Control of *Salmonella* pathogenicity island-2 gene expression. *Curr Opin Microbiol* **12**: 199-204.

Fenton, M. S., Lee, S. J. and Gralla, J. D. (2000) *Escherichia coli* promoter opening and -10 recognition: mutational analysis of sigma70. *EMBO J* **19**: 1130-7.

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

Fouet, A., Namy, O. and Lambert, G. (2000) Characterization of the operon encoding the alternative sigma(B) factor from *Bacillus anthracis* and its role in virulence. *J Bacteriol* **182**: 5036-45.

Froderberg, L., Houben, E. N., Baars, L., Luirink, J. and de Gier, J. W. (2004) Targeting and translocation of two lipoproteins in *Escherichia coli* via the SRP/Sec/YidC pathway. *J Biol Chem* **279**: 31026-32.

Fuchs, E. L., Brutinel, E. D., Jones, A. K., Fulcher, N. B., Urbanowski, M. L., Yahr, T. L. and Wolfgang, M. C. (2010) The *Pseudomonas aeruginosa* Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and -independent mechanisms. *J Bacteriol* **192**: 3553-64.

Gangwer, K. A., Mushrush, D. J., Stauff, D. L., Spiller, B., McClain, M. S., Cover, T. L. and Lacy, D. B. (2007) Crystal structure of the *Helicobacter pylori* vacuolating toxin p55 domain. *Proc Natl Acad Sci U S A* **104**: 16293-8.

Gaston, K., Bell, A., Kolb, A., Buc, H. and Busby, S. (1990) Stringent spacing requirements for transcription activation by CRP. *Cell* **62**: 733-43.

Geiman, D. E., Kaushal, D., Ko, C., Tyagi, S., Manabe, Y. C., Schroeder, B. G., Fleischmann, R. D., Morrison, N. E., Converse, P. J., Chen, P. and Bishai, W. R. (2004) Attenuation of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. *Infect Immun* **72**: 1733-45.

Genevrois, S., Steeghs, L., Roholl, P., Letesson, J. J. and van der Ley, P. (2003) The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *EMBO J* **22**: 1780-9.

Gentschev, I., Dietrich, G. and Goebel, W. (2002) The *E. coli* alpha-hemolysin secretion system and its use in vaccine development. *Trends Microbiol* **10**: 39-45.

Ghosaini, L. R., Brown, A. M. and Sturtevant, J. M. (1988) Scanning calorimetric study of the thermal unfolding of catabolite activator protein from *Escherichia coli* in the absence and presence of cyclic mononucleotides. *Biochemistry* **27**: 5257-61.

Giangrossi, M., Prosseda, G., Tran, C. N., Brandi, A., Colonna, B. and Falconi, M. (2010) A novel antisense RNA regulates at transcriptional level the virulence gene *icsA* of *Shigella flexneri*. *Nucleic Acids Res* **38**: 3362-75.

Goldman, S. R., Ebright, R. H. and Nickels, B. E. (2009) Direct detection of abortive RNA transcripts *in vivo*. *Science* **324**: 927-8.

Gottesman, S. (1984) Bacterial regulation: global regulatory networks. *Annu Rev Genet* **18**: 415-41.

Grainger, D. C. and Busby, S. J. (2008) Global regulators of transcription in *Escherichia coli*: mechanisms of action and methods for study. *Adv Appl Microbiol* **65**: 93-113.

Guyer, D. M., Henderson, I. R., Nataro, J. P. and Mobley, H. L. (2000) Identification of Sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol Microbiol* **38**: 53-66.

Guyer, D. M., Radulovic, S., Jones, F. E. and Mobley, H. L. (2002) Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infect Immun* **70**: 4539-46.

Haagmans, W. and van der Woude, M. (2000) Phase variation of Ag43 in *Escherichia coli*: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol Microbiol* **35**: 877-87.

Hagan, C. L., Kim, S. and Kahne, D. (2010) Reconstitution of outer membrane protein assembly from purified components. *Science* **328**: 890-2.

Harrington, S. M., Dudley, E. G. and Nataro, J. P. (2006) Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiol Lett* **254**: 12-8.

Hayes, C. S., Aoki, S. K. and Low, D. A. (2010) Bacterial contact-dependent delivery systems. *Annu Rev Genet* **44**: 71-90.

Hegde, R. S. and Bernstein, H. D. (2006) The surprising complexity of signal sequences. *Trends Biochem Sci* **31**: 563-71.

Helmann, J. D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* **46**: 47-110.

Henderson, I. R., Hicks, S., Navarro-Garcia, F., Elias, W. P., Philips, A. D. and Nataro, J. P. (1999) Involvement of the enteroaggregative *Escherichia coli* plasmid-encoded toxin in causing human intestinal damage. *Infect Immun* **67**: 5338-44.

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

Henderson, I. R., Navarro-Garcia, F. and Nataro, J. P. (1998) The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6**: 370-8.

Hendrixson, D. R., de la Morena, M. L., Stathopoulos, C. and St Geme, J. W., 3rd (1997) Structural determinants of processing and secretion of the *Haemophilus influenzae* Hap protein. *Mol Microbiol* **26**: 505-18.

Hood, R. D., Singh, P., Hsu, F., Guvener, T., Carl, M. A., Trinidad, R. R., Silverman, J. M., Ohlson, B. B., Hicks, K. G., Plemel, R. L., Li, M., Schwarz, S., Wang, W. Y., Merz, A. J., Goodlett, D. R. and Mougous, J. D. (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* **7**: 25-37.

Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* **62**: 379-433.

Ieva, R. and Bernstein, H. D. (2009) Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc Natl Acad Sci U S A* **106**: 19120-5.

Ieva, R., Tian, P., Peterson, J. H. and Bernstein, H. D. (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc Natl Acad Sci U S A* **108**: E383-91.

Imada, K., Minamino, T., Tahara, A. and Namba, K. (2007) Structural similarity between the flagellar type III ATPase FliI and F1-ATPase subunits. *Proc Natl Acad Sci U S A* **104**: 485-90.

Ishihama, A. (2010) Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol Rev* **34**: 628-45.

Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-56.

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

Johnson, R. C., Bruist, M. F. and Simon, M. I. (1986) Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* **46**: 531-9.

Johnson, T. A., Qiu, J., Plaut, A. G. and Holyoak, T. (2009) Active-site gating regulates substrate selectivity in a chymotrypsin-like serine protease the structure of *Haemophilus influenzae* immunoglobulin A1 protease. *J Mol Biol* **389**: 559-74.

Jong, W. S. and Luirink, J. (2008) The conserved extension of the Hbp autotransporter signal peptide does not determine targeting pathway specificity. *Biochem Biophys Res Commun* **368**: 522-7.

Jose, J., Kramer, J., Klauser, T., Pohlner, J. and Meyer, T. F. (1996) Absence of periplasmic DsbA oxidoreductase facilitates export of cysteine-containing passenger proteins to the *Escherichia coli* cell surface via the Iga beta autotransporter pathway. *Gene* **178**: 107-10.

Junker, M., Besingi, R. N. and Clark, P. L. (2009) Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol Microbiol* **71**: 1323-32.

Kadokura, H. and Beckwith, J. (2010) Mechanisms of oxidative protein folding in the bacterial cell envelope. *Antioxid Redox Signal* **13**: 1231-46.

Kanamaru, S. (2009) Structural similarity of tailed phages and pathogenic bacterial secretion systems. *Proc Natl Acad Sci U S A* **106**: 4067-8.

Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S. and Ebright, R. H. (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**: 1144-7.

Kazmierczak, M. J., Wiedmann, M. and Boor, K. J. (2005) Alternative sigma factors and their roles in bacterial virulence. *Microbiol Mol Biol Rev* **69**: 527-43.

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

Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. and Kahne, D. (2007) Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**: 961-4.

Knowles, T. J., Browning, D. F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D. L., Squire, D., Chaudhuri, R. R., Viant, M. R., Cunningham, A. F., Henderson, I. R. and Overduin, M. (2011) Structure and function of BamE within the outer membrane and the beta-barrel assembly machine. *EMBO Rep* **12**: 123-8.

Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. and Henderson, I. R. (2008) Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol Microbiol* **68**: 1216-27.

Knowles, T. J., Scott-Tucker, A., Overduin, M. and Henderson, I. R. (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**: 206-14.

Koch, C. and Kahmann, R. (1986) Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. *J Biol Chem* **261**: 15673-8.

Koebnik, R., Locher, K. P. and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* **37**: 239-53.

Kolb, A., Kotlarz, D., Kusano, S. and Ishihama, A. (1995) Selectivity of the *Escherichia coli* RNA polymerase E sigma 38 for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res* **23**: 819-26.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. and Hughes, C. (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405**: 914-9.

Kostrewa, D., Granzin, J., Koch, C., Choe, H. W., Raghunathan, S., Wolf, W., Labahn, J., Kahmann, R. and Saenger, W. (1991) Three-dimensional structure of the *E. coli* DNA-binding protein FIS. *Nature* **349**: 178-80.

Kovacikova, G. and Skorupski, K. (2001) Overlapping binding sites for the virulence gene regulators AphA, AphB and cAMP-CRP at the *Vibrio cholerae tcpPH* promoter. *Mol Microbiol* **41**: 393-407.

Lane, M. C. and Mobley, H. L. (2007) Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney Int* **72**: 19-25.

Lehr, U., Schutz, M., Oberhettinger, P., Ruiz-Perez, F., Donald, J. W., Palmer, T., Linke, D., Henderson, I. R. and Autenrieth, I. B. (2010) C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. *Mol Microbiol* **78**: 932-46.

Leiman, P. G., Basler, M., Ramagopal, U. A., Bonanno, J. B., Sauder, J. M., Pukatzki, S., Burley, S. K., Almo, S. C. and Mekalanos, J. J. (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A* **106**: 4154-9.

Leyton, D. L., de Luna, M. G., Sevastsyanovich, Y. R., Tveen Jensen, K., Browning, D. F., Scott-Tucker, A. and Henderson, I. R. (2010) The unusual extended signal peptide region is not required for secretion and function of an *Escherichia coli* autotransporter. *FEMS Microbiol Lett* **311**: 133-9.

Li, G. and Reinberg, D. (2011) Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev* **21**: 175-86.

Lindenthal, C. and Elsinghorst, E. A. (1999) Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli*. *Infect Immun* **67**: 4084-91.

Lodge, J., Fear, J., Busby, S., Gunasekaran, P. and Kamini, N. R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett* **74**: 271-6.

Lucas, R. L. and Lee, C. A. (2000) Unravelling the mysteries of virulence gene regulation in *Salmonella typhimurium*. *Mol Microbiol* **36**: 1024-33.

Luijsterburg, M. S., White, M. F., van Driel, R. and Dame, R. T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* **43**: 393-418.

Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. and Silhavy, T. J. (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* **61**: 151-64.

Mallik, P., Paul, B. J., Rutherford, S. T., Gourse, R. L. and Osuna, R. (2006) DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of *fis* expression in *Escherichia coli*. *J Bacteriol* **188**: 5775-82.

Martinez-Antonio, A. and Collado-Vides, J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* **6**: 482-9.

Mattei, P. J., Faudry, E., Job, V., Izore, T., Attree, I. and Dessen, A. (2011) Membrane targeting and pore formation by the type III secretion system translocon. *FEBS J* **278**: 414-26.

McKay, D. B. and Steitz, T. A. (1981) Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. *Nature* **290**: 744-9.

McLeod, S. M., Aiyar, S. E., Gourse, R. L. and Johnson, R. C. (2002) The C-terminal domains of the RNA polymerase alpha subunits: contact site with Fis and localization during co-activation with CRP at the *Escherichia coli proP* P2 promoter. *J Mol Biol* **316**: 517-29.

McLeod, S. M., Xu, J. and Johnson, R. C. (2000) Coactivation of the RpoS-dependent *proP* P2 promoter by Fis and cyclic AMP receptor protein. *J Bacteriol* **182**: 4180-7.

Michiels, T., Wattiau, P., Brasseur, R., Ruysschaert, J. M. and Cornelis, G. (1990) Secretion of Yop proteins by *Yersiniae*. *Infect Immun* **58**: 2840-9.

Miller, J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H. and Severinov, K. (2001) Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci U S A* **98**: 892-7.

Mobley, H. L., Green, D. M., Trifillis, A. L., Johnson, D. E., Chippendale, G. R., Lockatell, C. V., Jones, B. D. and Warren, J. W. (1990) Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**: 1281-9.

Morin, N., Tirling, C., Ivison, S. M., Kaur, A. P., Nataro, J. P. and Steiner, T. S. (2010) Autoactivation of the AggR regulator of enteroaggregative *Escherichia coli in vitro* and *in vivo*. *FEMS Immunol Med Microbiol* **58**: 344-55.

Muller, C. M., Aberg, A., Straseviciene, J., Emody, L., Uhlin, B. E. and Balsalobre, C. (2009) Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathog* **5**: e1000303.

Mullineaux, C. W., Nenninger, A., Ray, N. and Robinson, C. (2006) Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. *J Bacteriol* **188**: 3442-8.

Murakami, K. S. and Darst, S. A. (2003) Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol* **13**: 31-9.

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. and Darst, S. A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285-90.

Narita, S. and Tokuda, H. (2006) An ABC transporter mediating the membrane detachment of bacterial lipoproteins depending on their sorting signals. *FEBS Lett* **580**: 1164-70.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V. and Ebright, R. H. (2000) Structural organization of the RNA polymerase-promoter open complex. *Cell* **101**: 601-11.

Nasser, W., Rochman, M. and Muskhelishvili, G. (2002) Transcriptional regulation of *fis* operon involves a module of multiple coupled promoters. *EMBO J* **21**: 715-24.

Nataro, J. P., Deng, Y., Cookson, S., Cravioto, A., Savarino, S. J., Guers, L. D., Levine, M. M. and Tacket, C. O. (1995) Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *J Infect Dis* **171**: 465-8.

Nataro, J. P., Yikang, D., Yingkang, D. and Walker, K. (1994) AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J Bacteriol* **176**: 4691-9.

Navarro-Garcia, F., Canizalez-Roman, A., Burlingame, K. E., Teter, K. and Vidal, J. E. (2007a) Pet, a non-AB toxin, is transported and translocated into epithelial cells by a retrograde trafficking pathway. *Infect Immun* **75**: 2101-9.

Navarro-Garcia, F., Canizalez-Roman, A., Vidal, J. E. and Salazar, M. I. (2007b) Intoxication of epithelial cells by plasmid-encoded toxin requires clathrin-mediated endocytosis. *Microbiology* **153**: 2828-38.

Navarro-Garcia, F., Sears, C., Eslava, C., Cravioto, A. and Nataro, J. P. (1999) Cytoskeletal effects induced by Pet, the serine protease enterotoxin of enteroaggregative *Escherichia coli*. *Infect Immun* **67**: 2184-92.

Nickels, B. E. and Dove, S. L. (2011) NanoRNAs: A Class of Small RNAs That Can Prime Transcription Initiation in Bacteria. *J Mol Biol* **412**: 772-81.

Nielubowicz, G. R. and Mobley, H. L. (2010) Host-pathogen interactions in urinary tract infection. *Nat Rev Urol* **7**: 430-41.

Nilsson, L., Vanet, A., Vijgenboom, E. and Bosch, L. (1990) The role of FIS in trans activation of stable RNA operons of *E. coli*. *EMBO J* **9**: 727-34.

Ó Cróinín, T., Carroll, R. K., Kelly, A. and Dorman, C. J. (2006) Roles for DNA supercoiling and the Fis protein in modulating expression of virulence genes during intracellular growth of *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **62**: 869-82.

Oliver, D. C., Huang, G., Nodel, E., Pleasance, S. and Fernandez, R. C. (2003) A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* **47**: 1367-83.

Oomen, C. J., van Ulsen, P., van Gelder, P., Feijen, M., Tommassen, J. and Gros, P. (2004) Structure of the translocator domain of a bacterial autotransporter. *EMBO J* **23**: 1257-66.

Opel, M. L., Aeling, K. A., Holmes, W. M., Johnson, R. C., Benham, C. J. and Hatfield, G. W. (2004) Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol Microbiol* **53**: 665-74.

Otto, B. R., Sijbrandi, R., Luirink, J., Oudega, B., Heddle, J. G., Mizutani, K., Park, S. Y. and Tame, J. R. (2005) Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem* **280**: 17339-45.

Owen, P., Meehan, M., de Loughry-Doherty, H. and Henderson, I. (1996) Phase-variable outer membrane proteins in *Escherichia coli*. *FEMS Immunol Med Microbiol* **16**: 63-76.

Pallen, M. J., Chaudhuri, R. R. and Henderson, I. R. (2003) Genomic analysis of secretion systems. *Curr Opin Microbiol* **6**: 519-27.

Pan, C. Q., Finkel, S. E., Cramton, S. E., Feng, J. A., Sigman, D. S. and Johnson, R. C. (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J Mol Biol* **264**: 675-95.

Park, S. Y., Cromie, M. J., Lee, E. J. and Groisman, E. A. (2010) A bacterial mRNA leader that employs different mechanisms to sense disparate intracellular signals. *Cell* **142**: 737-48.

Paschen, S. A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D. and Neupert, W. (2003) Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* **426**: 862-6.

Perez-Rueda, E. and Collado-Vides, J. (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res* **28**: 1838-47.

Peterson, J. H., Szabady, R. L. and Bernstein, H. D. (2006) An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor, and the SecYEG complex. *J Biol Chem* **281**: 9038-48.

Picard, F., Dressaire, C., Girbal, L. and Cocaign-Bousquet, M. (2009) Examination of post-transcriptional regulations in prokaryotes by integrative biology. *C R Biol* **332**: 958-73.

Planet, P. J., Kachlany, S. C., DeSalle, R. and Figurski, D. H. (2001) Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. *Proc Natl Acad Sci U S A* **98**: 2503-8.

Pohlner, J., Halter, R., Beyreuther, K. and Meyer, T. F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325**: 458-62.

Porter, M. E. and Dorman, C. J. (2002) *In vivo* DNA-binding and oligomerization properties of the *Shigella flexneri* AraC-like transcriptional regulator VirF as identified by random and site-specific mutagenesis. *J Bacteriol* **184**: 531-9.

Pul, U., Lux, B., Wurm, R. and Wagner, R. (2008) Effect of upstream curvature and transcription factors H-NS and LRP on the efficiency of *Escherichia coli* rRNA promoters P1 and P2 - a phasing analysis. *Microbiology* **154**: 2546-58.

Purdy, G. E., Fisher, C. R. and Payne, S. M. (2007) IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J Bacteriol* **189**: 5566-73.

Purdy, G. E., Hong, M. and Payne, S. M. (2002) *Shigella flexneri* DegP facilitates IcsA surface expression and is required for efficient intercellular spread. *Infect Immun* **70**: 6355-64.

Raetz, C. R. and Dowhan, W. (1990) Biosynthesis and function of phospholipids in *Escherichia coli*. *J Biol Chem* **265**: 1235-8.

Raetz, C. R. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**: 635-700.

Raghavan, R., Groisman, E. A. and Ochman, H. (2011) Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res* **21**: 1487-97.

Raivio, T. L., Popkin, D. L. and Silhavy, T. J. (1999) The Cpx envelope stress response is controlled by amplification and feedback inhibition. *J Bacteriol* **181**: 5263-72.

Renn, J. P. and Clark, P. L. (2011) Disulfide bond-mediated passenger domain stalling as a structural probe of autotransporter outer membrane secretion *in vivo*. *Methods Enzymol* **492**: 233-51.

Rhodius, V. A. and Busby, S. J. (2000) Interactions between activating region 3 of the *Escherichia coli* cyclic AMP receptor protein and region 4 of the RNA polymerase sigma(70) subunit: application of suppression genetics. *J Mol Biol* **299**: 311-24.

Rhodius, V. A., West, D. M., Webster, C. L., Busby, S. J. and Savery, N. J. (1997) Transcription activation at class II CRP-dependent promoters: the role of different activating regions. *Nucleic Acids Res* **25**: 326-32.

Rizzitello, A. E., Harper, J. R. and Silhavy, T. J. (2001) Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J Bacteriol* **183**: 6794-800.

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

Ross, W., Ernst, A. and Gourse, R. L. (2001) Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev* **15**: 491-506.

Rossiter, A. E., Browning, D. F., Leyton, D. L., Johnson, M. D., Godfrey, R. E., Wardius, C. A., Desvaux, M., Cunningham, A. F., Ruiz-Perez, F., Nataro, J. P., Busby, S. J. and Henderson, I. R. (2011a) Transcription of the plasmid-encoded toxin gene from enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Mol Microbiol* **81**: 179-91.

Rossiter, A. E., Leyton, D. L., Tveen-Jensen, K., Browning, D. F., Sevastsyanovich, Y., Knowles, T. J., Nichols, K. B., Cunningham, A. F., Overduin, M., Schembri, M. A. and Henderson, I. R. (2011b) The essential {beta}-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J Bacteriol* **193**: 4250-3.

Rouviere, P. E. and Gross, C. A. (1996) SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev* **10**: 3170-82.

Ruiz-Perez, F., Henderson, I. R., Leyton, D. L., Rossiter, A. E., Zhang, Y. and Nataro, J. P. (2009) Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of Enterobacteriaceae. *J Bacteriol* **191**: 6571-83.

Ruiz-Perez, F., Henderson, I. R. and Nataro, J. P. (2010) Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes* **1**: 339-344.

Ruiz, N., Gronenberg, L. S., Kahne, D. and Silhavy, T. J. (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A* **105**: 5537-42.

Russell, A. B., Hood, R. D., Bui, N. K., LeRoux, M., Vollmer, W. and Mougous, J. D. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**: 343-7.

Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M. and Valencia, A. (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem Sci* **28**: 523-6.

Sanderson, A., Mitchell, J. E., Minchin, S. D. and Busby, S. J. (2003) Substitutions in the *Escherichia coli* RNA polymerase sigma70 factor that affect recognition of extended -10 elements at promoters. *FEBS Lett* **544**: 199-205.

Sandkvist, M. (2001a) Biology of type II secretion. *Mol Microbiol* **40**: 271-83.

Sandkvist, M. (2001b) Type II secretion and pathogenesis. *Infect Immun* **69**: 3523-35.

Santini, C. L., Ize, B., Chanal, A., Muller, M., Giordano, G. and Wu, L. F. (1998) A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* **17**: 101-12.

Sargent, F. (2007) The twin-arginine transport system: moving folded proteins across membranes. *Biochem Soc Trans* **35**: 835-47.

Sasse-Dwight, S. and Gralla, J. D. (1988) Probing co-operative DNA-binding *in vivo*. The *lac* O1:O3 interaction. *J Mol Biol* **202**: 107-19.

Sauri, A., Soprova, Z., Wickstrom, D., de Gier, J. W., Van der Schors, R. C., Smit, A. B., Jong, W. S. and Luirink, J. (2009) The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **155**: 3982-91.

Sauvonnet, N., Vignon, G., Pugsley, A. P. and Gounon, P. (2000) Pilus formation and protein secretion by the same machinery in *Escherichia coli*. *EMBO J* **19**: 2221-8.

Serruto, D., Adu-Bobie, J., Scarselli, M., Veggi, D., Pizza, M., Rappuoli, R. and Arico, B. (2003) *Neisseria meningitidis* App, a new adhesin with autocatalytic serine protease activity. *Mol Microbiol* **48**: 323-34.

Sheikh, J., Czeczulin, J. R., Harrington, S., Hicks, S., Henderson, I. R., Le Bouguenec, C., Gounon, P., Phillips, A. and Nataro, J. P. (2002) A novel dispersin protein in enteroaggregative *Escherichia coli*. *J Clin Invest* **110**: 1329-37.

Sheikh, J., Hicks, S., Dall'Agnol, M., Phillips, A. D. and Nataro, J. P. (2001) Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol Microbiol* **41**: 983-97.

Shere, K. D., Sallustio, S., Manessis, A., D'Aversa, T. G. and Goldberg, M. B. (1997) Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. *Mol Microbiol* **25**: 451-62.

Sherlock, O., Dobrindt, U., Jensen, J. B., Munk Vejborg, R. and Klemm, P. (2006) Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. *J Bacteriol* **188**: 1798-807.

Shin, M., Kang, S., Hyun, S. J., Fujita, N., Ishihama, A., Valentin-Hansen, P. and Choy, H. E. (2001) Repression of *deoP2* in *Escherichia coli* by CytR: conversion of a transcription activator into a repressor. *EMBO J* **20**: 5392-9.

Silhavy, T. J., Kahne, D. and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**: a000414.

Skillman, K. M., Barnard, T. J., Peterson, J. H., Ghirlando, R. and Bernstein, H. D. (2005) Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* **58**: 945-58.

Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. and Silhavy, T. J. (2007a) Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc Natl Acad Sci U S A* **104**: 6400-5.

Sklar, J. G., Wu, T., Kahne, D. and Silhavy, T. J. (2007b) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev* **21**: 2473-84.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jager, K. E. and Blasi, U. (2003) Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. *Microb Pathog* **35**: 217-28.

Soprova, Z., Sauri, A., van Ulsen, P., Tame, J. R., den Blaauwen, T., Jong, W. S. and Luirink, J. (2010) A conserved aromatic residue in the autochaperone domain of the autotransporter Hbp is critical for initiation of outer membrane translocation. *J Biol Chem* **285**: 38224-33.

Spiess, C., Beil, A. and Ehrmann, M. (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**: 339-47.

Squire, D. J., Xu, M., Cole, J. A., Busby, S. J. and Browning, D. F. (2009) Competition between NarL-dependent activation and Fis-dependent repression controls expression from the *Escherichia coli yeaR* and *ogt* promoters. *Biochem J* **420**: 249-57.

Srivatsan, A. and Wang, J. D. (2008) Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr Opin Microbiol* **11**: 100-5.

Stapleton, M., Haq, I., Hunt, D. M., Arnvig, K. B., Artymiuk, P. J., Buxton, R. S. and Green, J. (2010) *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J Biol Chem* **285**: 7016-27.

Steen, J. A., Harrison, P., Seemann, T., Wilkie, I., Harper, M., Adler, B. and Boyce, J. D. (2010) Fis is essential for capsule production in *Pasteurella multocida* and regulates expression of other important virulence factors. *PLoS Pathog* **6**: e1000750.

Stella, S., Cascio, D. and Johnson, R. C. (2010) The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev* **24**: 814-26.

Szabady, R. L., Peterson, J. H., Skillman, K. M. and Bernstein, H. D. (2005) An unusual signal peptide facilitates late steps in the biogenesis of a bacterial autotransporter. *Proc Natl Acad Sci U S A* **102**: 221-6.

Tajima, N., Kawai, F., Park, S. Y. and Tame, J. R. (2010) A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J Mol Biol* **402**: 645-56.

Teramoto, J., Yoshimura, S. H., Takeyasu, K. and Ishihama, A. (2010) A novel nucleoid protein of *Escherichia coli* induced under anaerobiotic growth conditions. *Nucleic Acids Res* **38**: 3605-18.

Travers, A., Schneider, R. and Muskhelishvili, G. (2001) DNA supercoiling and transcription in *Escherichia coli*: The FIS connection. *Biochimie* **83**: 213-7.

Typas, A., Becker, G. and Hengge, R. (2007) The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Mol Microbiol* **63**: 1296-306.

Ulett, G. C., Valle, J., Beloin, C., Sherlock, O., Ghigo, J. M. and Schembri, M. A. (2007) Functional analysis of antigen 43 in uropathogenic *Escherichia coli* reveals a role in long-term persistence in the urinary tract. *Infect Immun* **75**: 3233-44.

Ushida, C. and Aiba, H. (1990) Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic Acids Res* **18**: 6325-30.

Valent, Q. A., Scotti, P. A., High, S., de Gier, J. W., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B. and Luirink, J. (1998) The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J* **17**: 2504-12.

van den Berg, B. (2010) Crystal structure of a full-length autotransporter. *J Mol Biol* **396**: 627-33.

van der Woude, M. W. and Henderson, I. R. (2008) Regulation and function of Ag43 (flu). *Annu Rev Microbiol* **62**: 153-69.

van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. and Dame, R. T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc Natl Acad Sci U S A* **101**: 6969-74.

van Ulsen, P., van Alphen, L., ten Hove, J., Fransen, F., van der Ley, P. and Tommassen, J. (2003) A *Neisserial* autotransporter NalP modulating the processing of other autotransporters. *Mol Microbiol* **50**: 1017-30.

Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S. and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**: 712-9.

Veiga, E., de Lorenzo, V. and Fernandez, L. A. (2004) Structural tolerance of bacterial autotransporters for folded passenger protein domains. *Mol Microbiol* **52**: 1069-80.

Velarde, J. J. and Nataro, J. P. (2004) Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger. *J Biol Chem* **279**: 31495-504.

Volokhina, E. B., Beckers, F., Tommassen, J. and Bos, M. P. (2009) The beta-barrel outer membrane protein assembly complex of *Neisseria meningitidis*. *J Bacteriol* **191**: 7074-85.

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. and Tommassen, J. (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**: 262-5.

Vuong, P., Bennion, D., Mantei, J., Frost, D. and Misra, R. (2008) Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *J Bacteriol* **190**: 1507-17.

Waldron, D. E., Owen, P. and Dorman, C. J. (2002) Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli*. *Mol Microbiol* **44**: 509-20.

Wallecha, A., Munster, V., Correnti, J., Chan, T. and van der Woude, M. (2002) Dam- and OxyR-dependent phase variation of *agn43*: essential elements and evidence for a new role of DNA methylation. *J Bacteriol* **184**: 3338-47.

Weber, B. H. (1991) Glynn and the conceptual development of the chemiosmotic theory: a retrospective and prospective view. *Biosci Rep* **11**: 577-617.

Weber, I. T. and Steitz, T. A. (1987) Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 A resolution. *J Mol Biol* **198**: 311-26.

Werner, J. and Misra, R. (2005) YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. *Mol Microbiol* **57**: 1450-9.

West, D., Williams, R., Rhodius, V., Bell, A., Sharma, N., Zou, C., Fujita, N., Ishihama, A. and Busby, S. (1993) Interactions between the *Escherichia coli* cyclic AMP receptor protein and RNA polymerase at class II promoters. *Mol Microbiol* **10**: 789-97.

Wickner, W., Driessen, A. J. and Hartl, F. U. (1991) The enzymology of protein translocation across the *Escherichia coli* plasma membrane. *Annu Rev Biochem* **60**: 101-24.

Wiggins, P. A., Dame, R. T., Noom, M. C. and Wuite, G. J. (2009) Protein-mediated molecular bridging: a key mechanism in biopolymer organization. *Biophys J* **97**: 1997-2003.

Wilson, C. J., Zhan, H., Swint-Kruse, L. and Matthews, K. S. (2007) The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell Mol Life Sci* **64**: 3-16.

Wimley, W. C. (2003) The versatile beta-barrel membrane protein. *Curr Opin Struct Biol* **13**: 404-11.

Won, H. S., Lee, Y. S., Lee, S. H. and Lee, B. J. (2009) Structural overview on the allosteric activation of cyclic AMP receptor protein. *Biochim Biophys Acta* **1794**: 1299-308.

Wosten, M. M. (1998) Eubacterial sigma-factors. *FEMS Microbiol Rev* **22**: 127-50.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. and Kahne, D. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**: 235-45.

Yen, Y. T., Tsang, C., Cameron, T. A., Ankrah, D. O., Rodou, A. and Stathopoulos, C. (2010) Importance of conserved residues of the serine protease autotransporter beta-domain in passenger domain processing and beta-barrel assembly. *Infect Immun* **78**: 3516-28.

Yeo, H. J. and Waksman, G. (2004) Unveiling molecular scaffolds of the type IV secretion system. *J Bacteriol* **186**: 1919-26.

Zhai, Y., Zhang, K., Huo, Y., Zhu, Y., Zhou, Q., Lu, J., Black, I., Pang, X., Roszak, A. W., Zhang, X., Isaacs, N. W. and Sun, F. (2011) Autotransporter passenger domain

secretion requires a hydrophobic cavity at the extracellular entrance of the beta-domain pore. *Biochem J* **435**: 577-87.

Zhan, L., Han, Y., Yang, L., Geng, J., Li, Y., Gao, H., Guo, Z., Fan, W., Li, G., Zhang, L., Qin, C., Zhou, D. and Yang, R. (2008) The cyclic AMP receptor protein, CRP, is required for both virulence and expression of the minimal CRP regulon in *Yersinia pestis* biovar microtus. *Infect Immun* **76**: 5028-37.

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. and Darst, S. A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell* **98**: 811-24.

Zhang, M., Lin, S., Song, X., Liu, J., Fu, Y., Ge, X., Fu, X., Chang, Z. and Chen, P. R. (2011) A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat Chem Biol* **7**: 671-7.

Zheng, D., Constantinidou, C., Hobman, J. L. and Minchin, S. D. (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res* **32**: 5874-93.

Zubay, G., Schwartz, D. and Beckwith, J. (1970) Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc Natl Acad Sci U S A* **66**: 104-10.