REGULATION IN THE *ESCHERICHIA COLI LEE* PATHOGENICITY ISLAND

BY

GHADAH SALEH ALSHARIF

A THESIS

SUBMITTED TO THE UNIVERSITY OF BIRMINGHAM FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

SCHOOL OF BIOSCIENCES

THE UNIVERSITY OF BIRMINGHAM

EDGBASTON

BIRMINGHAM

B15 2TT

UK

DECEMBER 2014

Summary

Enterohaemorrhagic *Escherichia coli* (EHEC) pathogens represent major difficulties to health as they cause infections and complications in the gastrointestinal tracts of animals. Previous studies revealed that the locus of enterocyte effacement (*LEE)* pathogenicity island is the major player in triggering the attaching and effacing lesions that eventually induce death of host cells. The architecture of this island and its organization is understood, and it is clear that its expression is carefully regulated. The first step in its expression is activation of the LEE1 promoter that controls expression of the *LEE* encoded regulator (Ler) that activates LEE expression. Only under suitable conditions is expression of the *LEE* activated. The global regulator of *LEE* activator (GrlA) transcription is considered to be especially important in this process. Previous work showed that it acts by binding at the LEE1 promoter and causes a conformational change that leads to promoter activation. Hence, the LEE1 promoter, though displaying some basal level of activity, requires GrlA for maximal expression.

Although GrlA binds to the LEE1 promoter and activates transcription initiation *in vitro*, the fold of activation was only ~ 1.5 fold (Islam *at al.*, 2011). The aim of this work was to discover the factors that trigger GrlA activity. Thus, in chapter 3, it is shown how bacterial attachment to host cells triggers GrlA activity to an extent not seen during planktonic growth, with up to ~20 fold activation. Data also show that the level of free unbound GrlA defines the activity of the *LEE* promoter by GrlA, and the previously characterised GrlR anti-GrlA protein merely serves to buffer the level of GrlA. Hence, free GrlA plays a major

role in activating transcription initiation at the LEE1 promoter. Results also reveal that the cytoskeleton rearrangements caused by EHEC depend on GrlA acting at the LEE1 promoter. The results point to a critical strategy in bacterial pathogenesis, whereby the microbe must know when it is attached before it turns on its “pathogenic features”. In this case, attachment to host cells is crucial, and it seems that EHEC actually senses and knows when to start pathogenesis. There could also be other trigger(s) or factors responsible for this that could be present in the gut.

Results in Chapter 4 concern the role of leader sequences in the LEE1 transcription unit. These sequences are rich in AT bases and are a target for the histone-like nucleoid structuring protein (H-NS) silencing molecule. Data in chapter 4 show that translation of *ler* could be activated *in cis* by expression of a small open reading frame. Several other aspects of the long leader sequence are investigated. Results suggest that *ler* leader sequences might serve as a target for a silencing molecule(s) as, upon deleting them, Ler expression is enhanced.

The LEE1 promoter is inefficient, and GrlA compensates for this inefficiency by binding between the promoter -35 and -10 elements. It is known that other transcription factors, such as the global regulator, the cAMP receptor protein (CRP), functions by binding to upstream sequences. Experiments described in Chapter 4 investigate whether CRP can activate synergistically with GrlA at the LEE1 promoter. The results show that GrlA and CRP act independently but not synergistically. This result suggests that synergy between activators can be produced only by certain combinations of transcription factors.

For my beloved mommy and daddy for everything they have provided to me

Acknowledgments

This work has been accomplished in Steve Busby's lab. During the years I spent in the lab, I came across a number of people who I will never forget. Most importantly, I enjoyed my studying experience and the easy friendly family-like environment in the lab. Thanks to my mother and my father for their continuous support and for their patience. I am thankful for the King Abdulaziz University in Jeddah, Saudi Arabia for sponsoring my studies in the UK.

I would like to deeply thank my number one support my dearest supervisor Steve Busby. Not to mention how knowledgeable Steve is. He backed up all my lab work, my thesis work, and general support. Nothing of what I have achieved would have been possible without him. I am utterly honoured to be one of his PhD students and to have this rather exceptional opportunity to work in his group. Many thank you Steve.

I would also like to thank Anne-Marie Krachler, a wonderful and smart scientist. It was pleasant and cosy to have worked with Anne-Marie. She immensely helped and supported me.

I am really grateful to my lab-mates. We had a nice and supportive environment. Thanks to Dr David, Dr Douglas, Rita, for their scientific comments. Thanks to Dr MD. Shahidul for teaching me various methods and techniques. Thanks to Dr Jack and Dr Laura for their continuous help on troubleshooting. Thanks to Yasir for being a great friend, thanks to Basema and Bandar.

Contents

List of Figures

List of Tables

List of abbreviations

A: Adenosine
AE: Attaching and Effacing
Amp^R: Ampicillin-resistance
APS: Ammonium persulfate
ATP: Adenosine triphosphate
bp: Base pair
BSA: Bovine serum albumin
C: Cytosine
cAMP: Cyclic adenosine monophosphate
CAP: Calf intestinal alkaline phosphate
Cm^R: Chloramphenicol-resistance
CRP: Cyclic AMP receptor protein
CTD: C-terminal domain
DNA: Deoxyribonucleic acid
DNase: Deoxyribonuclease
dNTP: Deoxyribonucleoside triphosphate (N=A, C, G, or T)
G: Guanine
GTP: Guanosine-5'-triphosphate
H-NS: Histone like-nucleoid structuring protein
HU: Heat unstable protein
IHF: Integration host factor
kDa: Kilo Dalton
LB: Luria broth
NTD: N-terminal domain
OD: Optical density
ONPG: Ortho-nitrophenyl-b-D-galactopyranoside
PCR: Polymerase chain reaction
RNA: Ribonucleic acid
RNAP: RNA polymerase
rRNA: Ribosomal RNA
SDS: Sodium dodecyl sulphate
T: Thymidine
TBE: Tris/ Borate/EDTA buffer
Tet^R: Tetracycline-resistance
TEMED: N,N,N',N'-Tetramethylethylenediamine
tRNA: Transfer RNA
Tris: 2-Amino-2-(hydroxymethyl)-1,3-propanediol
v/v: Volume per volume
w/v: Weight per volume

Chapter one:

Introduction

1.1 *Escherichia coli* as a model organism

It is believed that the name, *Escherichia coli,* is related to the paediatrician Theodor Escherich (Shulman *et al.,* 2007). He was a German scientist born on 29 November 1875 in Ansbach, Germany. He was concerned by high mortality rates in paediatrics and health care among the poor in society. In 1885, Escherich gave a talk about "The intestinal bacteria of neonate and in infant", and in this presentation he explained the shape and characteristics of a bacterial colony to the Society of Morphology and Physiology as "bacterium coli commune" (Hacker, and Blum-Oehler, 2007). Gradually, the science of *E. coli* became a major scientific interest; the Nobel Prize has been awarded to a number of scientists because of their research on various aspects of *E. coli* biology.

1.1.1 Cellular organization

E. coli can be described as a Gram-negative rod-shaped organism with a length of 2.5 micrometres and a diameter of 0.8 micrometres. The end caps of the organism's shape are semi-circular (Berg, 2004). When the bacterial cells enter phases of growth and division, they elongate and divide from the centre of the cell. At the place of division a construct forms to aid in the division called a divisome, which is derived from a huge gathering of proteins that are encoded by filamentous temperature-sensitive genes (*fts*) (Reshes *et al.,* 2008). These proteins of the divisome assemble in a careful manner to form a ring called an FtsZ ring. This ring contains a number of proteins, FtsA, FtsK, FtsQ, FtsB, FtsL, FtsW, FtsI, FtsN, and ZipA with each playing a distinct role in the cell division as well as other functions including

attaching the Z ring to the cytoplasmic membrane, aiding in the production of the new peptidoglycan layer in the newly synthesized cells, and other roles still under investigation. Eventually, the cells discard the Z ring by shrinking until closure and complete division of the cells is achieved. Other cellular structures of *E. coli* include the inner membrane, outer membrane, periplasmic area, peptidoglycan layer, pili and, flagella. *E. coli* does not contain a nucleus, thus the genetic material or DNA is present in the cytoplasm in a structure called the nucleoid. Each structure plays a role in the growth, division, and pathogenicity of the organism. The usual habitat of *E. coli* can be divided into two sections: primary habitat, and secondary habitat (Savageau, 1983) where the primary habitat is the large intestine of animals and the secondary habitat includes water, sediment, and soil. The author believes that there are a number of factors that shape the relationship between the organism and the primary habitat, which include diet, immune system status, the physiological state of the host, and the interactions between the host and other bacteria that are present in the host. For the secondary habitat, the relationship is as complex as in the primary habitat and there are some factors that define *E. coli* living in this secondary habitat, such as nutrients, and climate.

1.1.2 The *E. coli* cell envelope

The *E. coli* cell envelope is composed of the cell wall, cell membrane, and other structures (Fig. 1.1). The cell wall is formed of (from outside to inside) an outer membrane, periplasmic space, and the peptidoglycan layer (Mitchell, 1961; Glauert and Thornley, 1969). These are followed by the cell membrane, which encloses the cell. Other structures include, capsules, flagella, pili, and secretion systems. The outer membrane acts as a barrier for the bacterial cell, protecting it from environmental damage and it also serves as an expellant of toxic

molecules (Nikaido, 2003). It is evident that the outer membrane and the inner membrane are each composed of a phospholipid bilayer (Kellenberger and Ryter, 1958). The Gram-positive cell wall is different from the Gram-negative cell envelope (Fig. 1.1). A thick peptidoglycan layer characterizes the cell envelope of Gram-positive bacteria. This thick layer is thought to be due to the lack of the outer membrane that is present in the Gram-negative bacteria. The Gram-positive cell envelope is also characterized by teichoic acids that are found to be attached to the peptidoglycan layer and are composed of glycerol phosphate, glucosyl phosphate, and ribitol phosphate repeats. There is also another acid that is anchored to the membrane lipids; this is lipoteichoic acid (Neuhaus and Baddiley, 2003). The Gram-positive cell envelope also contains some other structures that resemble proteins present in the Gram-negative periplasm space (Dramsi *et al.*, 2008).

The outer membrane is formed of lipid bilayers, containing phospholipids, glycolipids, and lipopolysaccharides (LPS) (Kamio and Nikaido, 1976). It also contains a number of outer membrane proteins (OMPs), and lipoproteins that are implanted and anchored in the membrane (Silhavy *et al.*, 2010). The LPS is composed of three parts: lipid A, core oligosaccharide, and the O-antigen (Raetz, 1990). Fragmentation of the cell envelope using high levels of salt resulted in two fractions. One was identified as the cytoplasmic membrane, which was small and present in the form of vesicles.

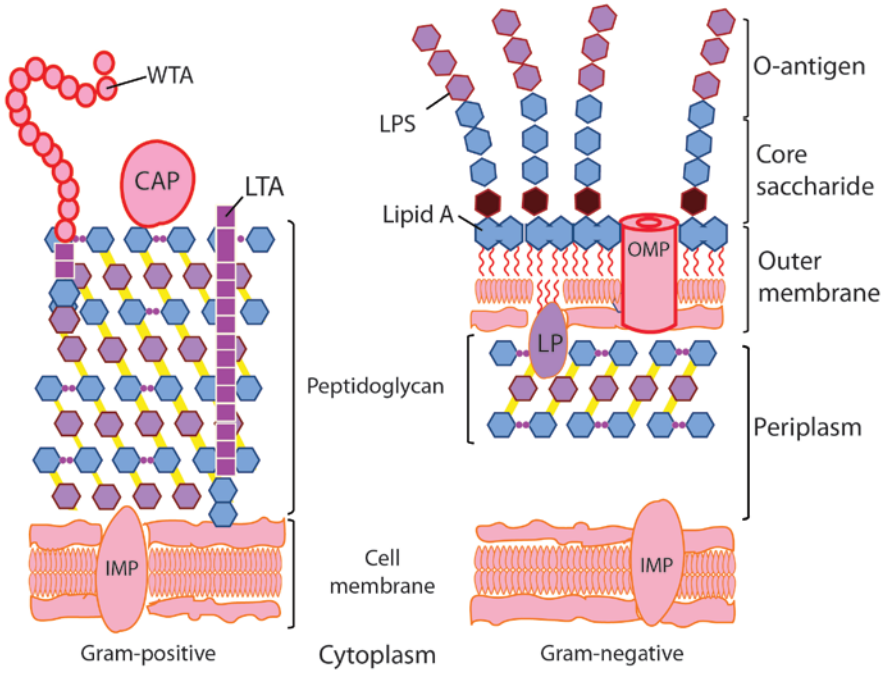


Fig. 1.1: Cell envelope structure of Gram-positive and Gram-negative bacteria. A schematic representation of the bacterium cell envelope to the left showing the cell envelope of Gram-positive bacteria. The thick peptidoglycan layer, characteristic of Gram-positive bacteria, the plasma membrane, and cell envelope proteins are shown. To the right is the cell envelope of the Gram-negative bacterium. The outer and inner membrane, the periplasm space, the peptidoglycan layer, and the lipopolysaccharide components are shown. Acronyms denote: WTA = wall teichonic acid; CAP = covalently attached protein; LTA = lipotechonic acid; IMP = integral membrane protein; LPS = lipopolysaccharide; OMP = outer membrane protein; LP = lipoprotein (Figure redrawn from Silhavy *et al.*, 2010).

The other fragment was large and C-shaped, and which was identified as cell wall, including the outer membrane and the glycoproteins. These C-shaped fragments can be described as rigid open structures. The reason that these structures are rigid was thought to be due to the presence of glycoproteins (murein) layer. Hence, these fragments were unable to vesicluate. The distribution of oxidative enzymes was similar. In addition, the distribution and composition of proteins in the cell wall was also identified by fractionation of the cell envelope on sucrose gradients, SDS-PAGE, etc. (Schnaitman, 1970). The periplasmic space in Gram-negative bacteria is arranged into two layers bordered by the outer and the inner membranes, and wrapping around the peptidoglycan layer (Mullineaux *et al.,* 2006). The periplasmic space is also very dense in proteins that are involved in the movement and transportation of nutrition, or chaperone-like proteins for molecules that form the outer membrane (Wülfing and Plückthun, 1994). The peptidoglycan layer can be defined as a multi-sugar structure that is connected by peptide bonds. The sugar constituent of this molecule is N-acetyl glucosamine-N-acetyl muramic acid (Schleifer and Kandler, 1972; Vollmer *et al.,* 2008), and because of its rigidity, the peptidoglycan layout defines the shape of the cell (Silhavy *et a.,* 2010).

The cell envelope is also composed of the cytoplasmic membrane, or the inner membrane. This membrane is formed of phospholipid bilayers and the main phospholipids here are from the phosphatidyl group (glycerol, serine), cardiolipin and also polyisoprenoid carriers (Raetz and Downhan, 1990). The cell envelope also contains other structures and virulence factors like, TTSS (1.5.3), capsule structures protecting cells from attack by the immune cells (Moxon and Kroll, 1990), and finally, some bacteria are motile due to the presence of flagella structures. In the case of EHEC, the main function of the flagella is to deliver the bacteria to

the site of infection (Silverman and Simon, 1974). The bacterial cells use other cell envelope structures, pili, to adhere to the site of infection (Sauer *et al*., 2000).

1.1.3 The *E. coli* chromosome

The bacterial chromosome, as in eukaryotes, is where the genetic information is stored; its organization, structures, replication, etc. are matters of interest, especially, given the fact that the size of the chromosome is actually thousands of times larger than the cell itself, where cell size 7 μm while the length of the chromosome could extend to 9 Mb. (Fig. 1.2). The methods for determining bacterial chromosome size, and their accuracy, vary. Such methods include, colorimetry, kinetics of renaturation, two-dimensional gel electrophoresis, summation of the sizes of restriction fragments, or using pulsed-field gel electrophoresis. Based on the data gathered from the PFGE and libraries of restriction fragments, the shape of the bacterial chromosome was found to be circular and the size ranges from 1 to 9 Mb (Krawiec and Riley, 1990). The cell must find ways of accommodating this huge mass of genetic material, hence the use of folding, bridging, and supercoiling, and the involvement of enzymes, such as topoisomerase.

The bacterial chromosome was isolated from the cell, and the shape of the chromosome *in vitro* was then identified as a rosette (Kavenoff and Bowen, 1976). The rosette structure is formed of a core structure, held together through DNA-RNA interactions, and about 12-80 topologically independent loops extending outward from the core. However, *in vivo*, the

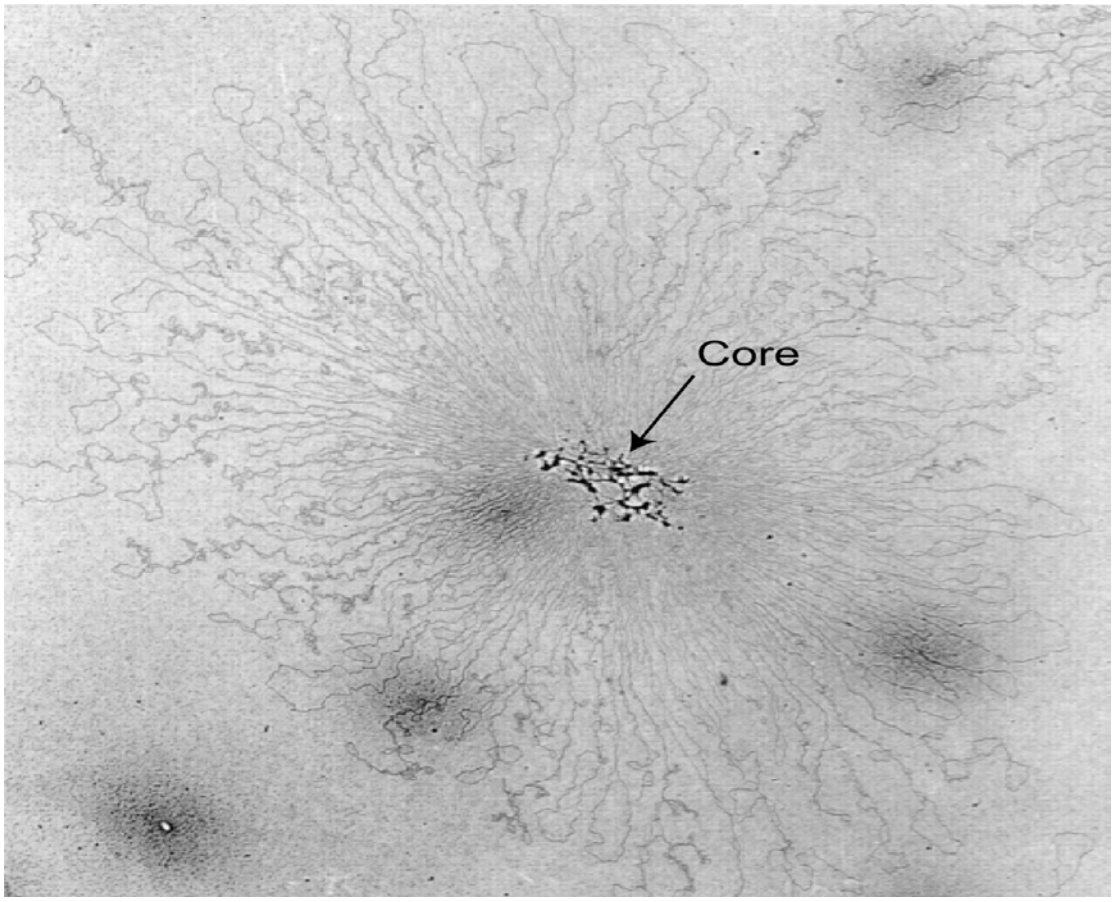


Fig. 1.2: Purified bacterial chromosome *in vitro.*

An electron microscopy picture of the purified *E. coli* chromosome showing the organization of the nucleoid, which is formed of a core structure and DNA loops. (Taken from Toro and Shapiro, 2010, which was originally re-printed from Kavenoff and Bowen, 1976).

chromosome is arranged as follows: the OriC (origin of replication) is located at the mid-cell, (Nielson *et al.,* 2006), the left and right chromosomal arms are positioned toward the cell halves and the terminus is broadly distributed in equal proportion across both halves (Wang *et al.,* 2006). The *E. coli* chromosome is circular and replicates bidirectional from a unique OriC, and genetic material is double both in the direction of OriC and at the replication fork. Thus creating two chromosomes arms or replichores (Kaguni, 2006). Replication is believed to be controlled strictly by limiting initiation by DnaA to once per cell generation. During replication, the replication fork speeds at 600-1000 bp/s (Kornbag, 1988). The leading strand is replicated toward the edge of the cell or the lagging strand replicates toward the edge of the cell. This is to maintain the configuration of the chromosomal organization <L-R-L-R>. However, this may not happen and the aforementioned configuration might be altered when, for example, in one replicating cell, the leading strand replicates to the end, and in the other replicating cell, the lagging strand replicates to the edge of the cell resulting in what is known as a symmetric segregation and the configuration <L-R-R-L> or <R-L-L-R> (White *et al.,* 2008). Termination of replication occurs at broad terminus regions delineated by abundant groups of replication termination sites. Termination sites are usually bound by terminator protein, Tus, which functions to ensure that each replication arm is in ~ 50% of chromosome (Wang *et al.,* 2007).

The mechanism by which segregation of sister chromosomes occurs is uncertain and one suggestion is that the chromosomes are actually anchored to the cell wall (Jacob, 1963), and that the process of segregation is passively achieved by a process of cell elongation. However, for this theory to be valid, first the chromosomes need to be attached to the cell wall (Sueoka and Quinn, 1968), and then the growth of the cell wall must be random (Mobley *et al.,* 1984).

Another hypothesis on chromosome segregation suggests that segregation occurs gradually during the process of replication and that it can be driven by the forces of DNA polymerase and RNA polymerase enzymes (Lemon and Grossman, 2001; Dworkin and Losick, 2002).

1.2 Control of gene expression in *E. coli*

Transcription is the synthesis of a new messenger RNA from a double-stranded DNA template and it involves a number of stages. These are transcription initiation, elongation, and then finally, termination (Browning and Busby, 2004; Robinson and Oijen, 2013) (Fig. 1.3). First, in the initiation stage, a site located upstream of the transcribed gene is recognised by the RNA polymerase (RNAP) enzyme. It is directed by the sigma subunit to the regulatory region, the promoter. There are a number of sigma factors that initiate transcription, and a specific sigma factor is selected to function according to what triggers the gene expression. However, *E. coli* σ^{70} is the main sigma factor that facilitates the initiation of transcription from housekeeping genes in *E. coli* (Lonetto *et al*., 1992). RNAP holoenzyme is formed when the core RNAP is bound to a sigma subunit. Then the RNAP forms a closed complex with the promoter. Holo RNAP looks like a crab claw (Sekine *et al.,* 2012), like the core enzyme (Kapanidis *et al.,* 2006; Revyakin *et al.,* 2006).

Transcription initiation regulation is considered to be the major regulation step in transcription, where a number of regulatory proteins bind to transcription units and either block or stimulate transcription by RNAP enzyme (Müller-Hill, 1998). It is believed that gene

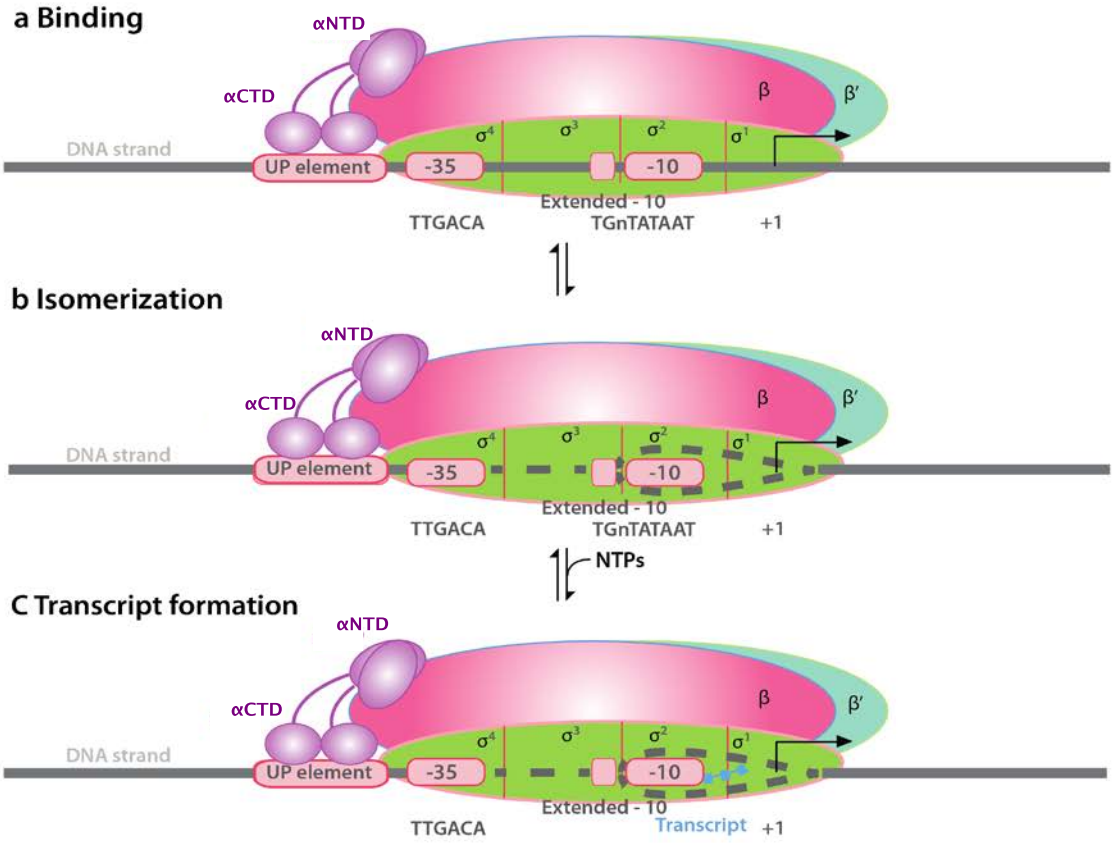


Fig. 1.3: Schematic representation of the transcription initiation stages.

This figure shows the transcription initiation in prokaryotes. (a) The RNAP binds to sigma subunits that direct the enzyme to regulatory elements on the DNA. (b) Once the RNAP-HE is bound, the DNA melts and an open complex is formed. The process is called isomerization. (c) The process of transcription is then initiated (Figure redrawn Lee *et al*, 2012).

expression is generally controlled at various steps from initiation of transcription through to the stability of the transcript. However, from the point view of economy, the key step in gene regulation lies at the point of initiation of mRNA transcript formation (Browning and Busby, 2004).

1.2.1 RNAP subunits and organization

The RNAP core enzyme is formed of α1α2ββ'ω subunits. For bacterial DNA to start transcription into mRNA, an RNAP enzyme binds specifically the ORF that is required for transcription. Each subunit of the enzyme is responsible for binding to specific elements on the DNA, ββ' bind to elements on both the DNA and the mRNA. They form the catalytic centre for RNA synthesis. ββ' also facilitate binding sites for the double-stranded downstream DNA, and DNA/RNA hybrid witch is formed during transcription and RNA synthesis. ββ' are believed to be highly conserved in bacteria. However, sequence insertions were found in these subunits. These insertions indicate evolutionary lineages of bacteria. These insertions can be isolated, and crystallized to determine the X-ray structure. It is believed that the crystal structures of the insertion sequences provided atomic images of the bacterial RNAP. Insertions were found to be located at the peripheral surface of the RNAP (Murakami, 2015).

The two α subunits are formed of carboxyl and amino terminal domains, where the carboxyl-terminal domain is a DNA-binding domain. αCTD domain plays significant roles in regulating transcription by interacting with transcription factors (Jeon *et al.,* 1995). Additionally, it binds to upstream promoter elements of DNA. The structure of αCTD indicates its compact structure and unique protein topology in contrast to other DNA binding

proteins (Ishihama, 1992). The amino terminal domain αNTD, is bound to other RNAP subunits. αNTD domain crustal structure was revealed and showing that the α subunit is homodimer. Homodimer is especially important platform which binds to the largest RNAP subunits, ββ'.

The small 91-residue 11 kDa ω domain has a role in transcription indirectly, in response to stress regulated by ppGpp, and functions as a chaperone to assist in the folding of the β' subunit. ω associates mainly with β' subunit by contacting both the N and C terminal parts of β' subunit (Vassylyev *et al.*, 2002). By binding to the full-length β' subunit, ω therefore helps the final step of RNAP core assembly when β' associates with α1α2β sub-complex (Ghosh *et al.*, 2001). Despite that the role of ω is still not fully understood, it has an important effect on the overall gene expression pattern. Inefficient recruitment of σ factor results in strains lacking ω. This results in down regulation of highly expressed genes and up regulation of a number of low expressed genes (Gunnelius *et al.*, 2014). Therefore, ω is essential in initial association of the primary σ factor with RNAP core enzyme. Hence, ω leads to efficient transcription of highly expressed genes (Browning and Busby, 2004).

Sigma factors play significant roles in transcription initiation stages (Saecker *at al.*, 2011). They can be classified into two groups according to their homology. Primary σ factor σ^{70} are responsible for most transcription in growing cells. The other group is σ^{54}, which direct transcription as a response to environmental signallings and require ATP hydrolysis to melt DNA (Zhang and Buck, 2015). Sigma factors that are related to sigma 70 in turn can be classified into four groups that are both phylogenetically and structurally distinct groups, 1-4

(Perude and Roberts, 2011; Raffaelle *et al.*, 2005). Alternative sigma factors are important for specialised functions and their activity or expression can be controlled at a number of stages but especially, their regulation is performed at the level of post-translation regulation. In this stage anti-σ factor binds to alternative sigma factors, preventing binding to the RNAP.

The four groups within the σ^{70} family vary due to the presence or absence of four conserved regions (σR1.1, σR1.2-2.4, σR3.0-3.2. σR4.1-4.2) reflecting four σ domains (Lonetto *at al.*, 1992). Domain σ_2 makes specific interaction with a single stranded non-template DNA -10 element (σR2.3-2.4) (Feklistov and Darst, 2011). Group 1 and 2 factors σR1.2 consists of two α helices oriented at 90° to one another (Haugen *et al.*, 2006). Domain σ3 is composed of a compact three-helical bundle interacting with the major groove of double-stranded DNA (Mitchell *et al.*, 2003). Domain σ4, is composed of four helices with third and fourth helix-turn helix motif. Domain σ4 forms the second largest interface with RNAP, it also engages with the -35 element, and also acts as contact point for transcription activators that bind upstream -35 elements (Zhang *et al.*, 2012). Domain σ1.1 (σR1.1) is only found in group 1 and promotes a compact form of free σ, this occludes the DNA binding determinants and hence, inhibiting its non-productive interaction with promoter DNA in the absence of core (Schwartz *et al.*, 2008).

Alternative sigma factors are composed of four groups as in σ^{70} family. In group one σ3 is in variable presence, promoter specificity is different and some stages in transcription initiation. Group 2 are closely related to group1. Group 2 is not essential. This group is involved in adaption to stress including nutrients limitations, temperature changes, light regulation, and

sugar metabolism and are associated with stationary phase. Examples of this group include, σ^S and σ^{38} (Buttner and Lewis, 1992). Group three are both structurally and functionally diverse containing σ2, σ3 and σ4 domains. This group usually recognize different -10 and -35 sequence to groups one and two. This group is formed of four-phylogenetically distinct subgroup responsible for: flagellum biosynthesis, sporulation, and general stress. Examples of this group are, σ^{28}, σ^D, σ^{32}, and σ^B (Koo *et al.,* 2009). Finally, group four which is known as extracytoplasmic function group due to their functions in sensing and responding to signals generated outside the cells or in the cell membrane. This group is believed to be the largest and most diverse at the primary sequence level. Hence they consist of 43 major phylogenetically distinct sub groups such as σ^E, σ^W, σ^{FecI}, and σ^R (Staron *et al.,* 1994). Generally, σ^{70} is the main specificity factor that directs RNAP to promoters. However, there are six different additional σ factors encoded by *E. coli* that are activated according to specific stimuli (Ishihama, 2000).

1.2.2 RNAP assembly

In the cell, RNA polymerase is assembled from individual subunits (Fig. 1.4). The first step is the dimerization of the two α subunits (Ishihama and Ito, 1972). It has been shown that the determinant that drives its dimerization is the αNTC. In the next step, free β subunit binds to the α1α2 dimer to give the α1α2β sub assembly (Ito *et al.,* 1975). In this sub assembly, each of the two α subunit is identical chemically, but they can now be distinguished in space. One of the 2 α subunits is associated with β, while the other will be associated with β' after the next assembly step. Prior to association, β' subunit folds together with the ω subunit which functions as chaperone (Matthew and Chaterji, 2006). Thus, the final step of assembly is the binding of β' and ω to

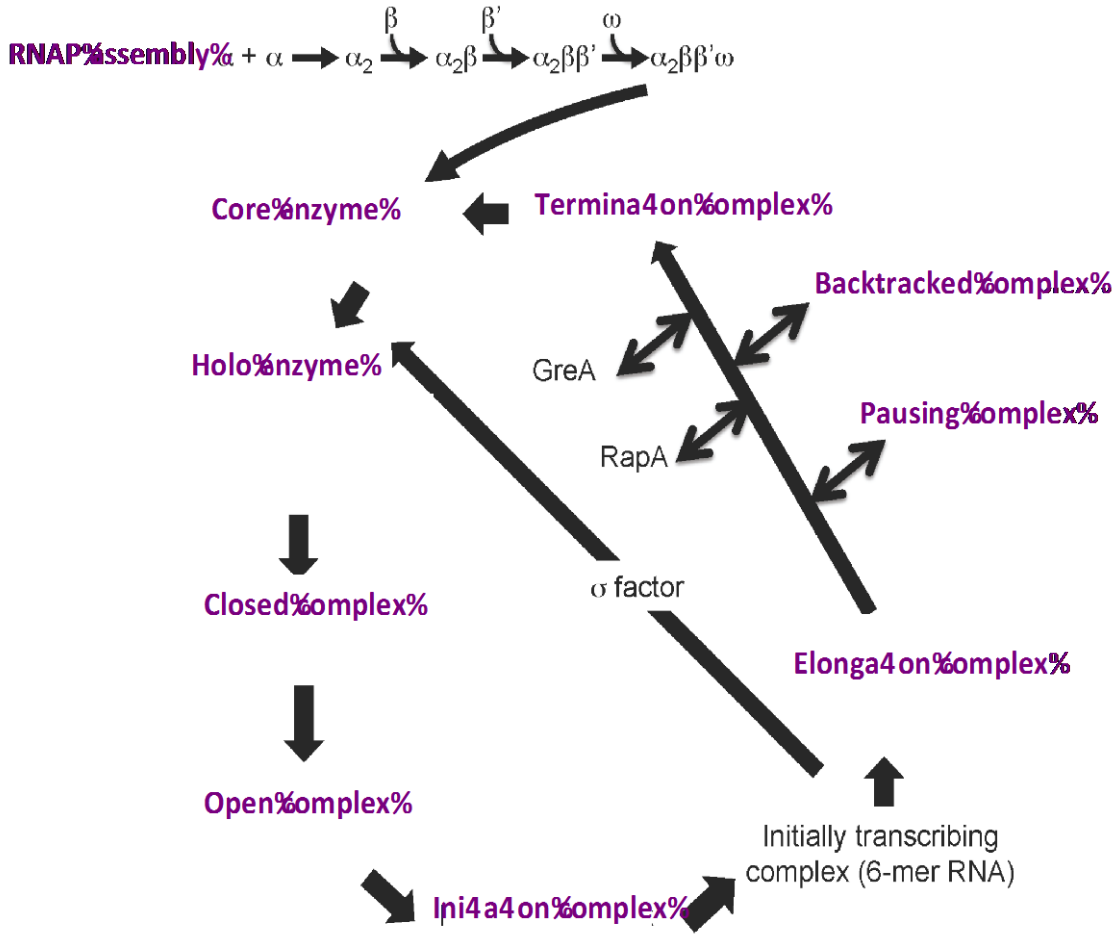


Fig.1.4: Schematic representation of RNAP assembly and transcription cycle.
The figure represents bacterial RNAP assembly and transcription cycle (Figure redrawn from Murakami, 2015)

α1α2β to give α1α2ββ'ω the RNAP core enzyme (Ishihama and Ito, 1972).

The core RNA polymerase can make RNA, but it cannot recognize promoters. In order for it to recognize promoters, it has to associate with a σ factor to form holo enzyme. This occurs by σ making contact with two determinants on the core enzyme. These are the β subunit flap and the β' subunit coiled-coil (Arthur and Burgess, 1998). Interestingly domain 4 recognize the β flab and this interaction is believed to be weak (Kuznedelov et al., 2002), and σ domain 2 recognizes the β' coiled coil (Burgess and Anthony, 2001; Young *et al.*, 2004). The interaction is well conserved. All sigma subunits bind to the same determinant on RNAP core enzyme. The consequence of this is that each core enzyme can only bind to one sigma factor. Amazingly, sigma 54, even thought it is unrelated to sigma 70, binds to the same two determinants on the RNAP (Buck *et al.*, 2015).

Hence RNA polymerase holoenzyme is a multi-subunit enzyme, and catalyses DNA-dependent RNA synthesis from specific promoter (Fig. 1.5). RNAP holoenzyme plays an important role in controlling gene expression. In *E. coli*, transcription of more than ~4000 genes is catalysed by ~2000 RNAP molecules, of which 80% are involved in transcription of only 80-90 transcriptional units, such as rRNA or tRNA genes (Ishihama, 2000). This results in short supply of RNAP for initiating transcription. In this complex, subunits of the RNAP (Fig. 1.5) bind to different promoter elements; the -10 and the -35 elements are recognised by two domains of the RNAP holoenzyme σ2 and σ4. And after the -10 elements upstream, there is a short sequence of 6-8 bases called the extended -10 element, which is recognized by the σ3 domain, and then σ1 overlaps with the transcription start site.

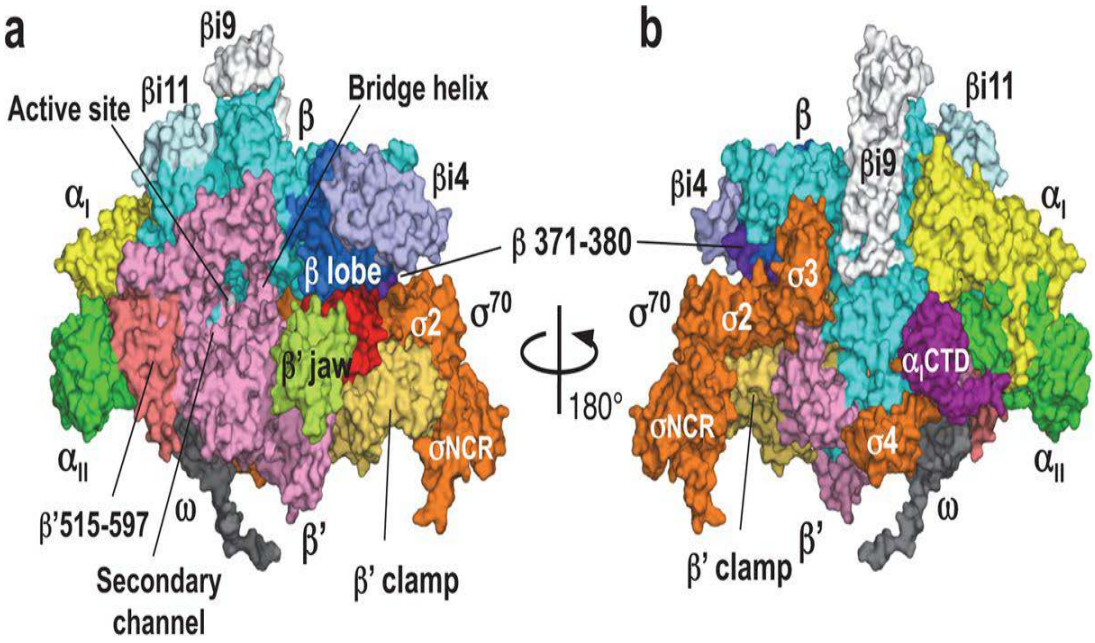


Fig. 1.5: Structure of the RNA holoenzyme.

The structure of the RNAP-HE complex is shown. The light yellow and green shades represent the NTD of the αI and αII subunits of the RNAP that bind to determinant on β subunit. Pink and blue shades are the β and β' subunits of the RNAP. The sigma σ subunit of the RNAP in orange, which bind to the -10, -35 extended, -10 and of the promoter, and omega are in grey (Figure taken from Murakami, 2013).

At some promoters, the sequence upstream of the -35 element is called the UP element which is recognized α subunit the CTD terminal RNAP. So here the sigma subunits act like the bridge between the promoter elements and the RNAP core enzyme. Once the RNAP-promoter complex is formed, the transcription site is melted (Browning and Busby, 2004).

1.2.3 Promoter elements

Five elements on promoter DNA have been found to be the binding sites of the RNAP holoenzyme. For *E. coli* σ^{70} promoters, two of these are the -10 hexamer (consensus 5'-TATAAT-3'), and the -35 hexamer (consensus 5'-TTGACA-3') where σ2 and σ4 domains bind to respectively; both elements are located 10 and 35 bps upstream of the transcript start site, respectively, and are separated from each other by a spacer having a characteristic length of 17 bp. The third is an extended -10 element that is 3-4-bp and located just upstream of the -10 element and recognized by σ3 domain. The fourth element is the UP element which is a ~20 bp sequence and located upstream of the promoter -35 hexamer and is recognized by α subunits. Discriminator elements located -10 to +1 region and rich in GC and are target to ppGpp in starved cells. As the RNAP is in short supply, promoters compete with each other for acquiring it.

1.2.4 Small ligands

Bacterial cells can sense changes in the environment and respond to stimuli such as, mechanical stress, osmolarity change, oxidative stress, or microbial products. Hence, they have developed ways of modulating gene expression using alternative mechanisms other than

σ factors, via for example, guanosine 3’ 5’ biphosphate (ppGpp) or iNTP (initiating nucleotide triphosphate). iNTPs regulate transcription of stable RNA promoters in response to growth rate, and its concentration changes according to growth rate. When cells outgrow from stationary phase, the level of NTP increases. This leads to rapid and direct regulation of rRNA synthesis. High concentration of iNTP are especially required in synthesis of rRNA because NTPs form short lived complexes with the RNAP, so high concentrations of NTPs are required for binding and stabilizing open complexes and hence synthesis of rRNA. However, when cells enter stationary phase, the level of NTPs decrease and consequently, the level of rRNA synthesis. In this case, ppGpp regulate expression of rRNA. This regulation of expression according to growth condition is called “growth-rate-dependent-control”. By contrast, the availability of an intracellular level of ppGpp shuts down the transcription process rapidly by destabilizing open complexes when protein synthesis slows down due to amino acid limitation (Schneider and Gourse, 2004). Clearly, ATP controls expression in response to growth rate, whereas ppGpp controls the expression of translational machinery in response to sudden starvation (stringent response) (Gaal *et al*., 1997).

1.2.5 Transcription factors

For the *E. coli* genome, more than 300 genes encode proteins that bind to promoters, so transcription can be either up regulated or down regulated. These transcription factors regulate expression along with the environmental signals. Their functions can be regulated either by controlling their activity or by regulating their expression. These outcomes are achieved by either modulating the DNA-binding affinity of the transcription factors using small ligands or covalent modifications, by sensing their concentrations in the cells, or lastly

by sequestration by regulatory proteins. The mechanisms for regulating the transcription can be either a simple activation or a simple repression. The simple activation includes three classes of activation, Class I, Class II, and Conformation Change transcription factor (Browning and Busby, 2004) (Fig. 1.6). In Class I the activation of transcription is performed by binding to a target element upstream of the promoter -35 element. This binding recruits RNAP to the promoter by directly interacting with the αCTD domain subunit of the RNAP; cyclic AMP receptor protein functions in this way (Ushida and Aiba, 1990; Reviewed by Ebright, 1993). Class II, however, functions by binding the activator to parts from the RNAP, for example the αNTD part, and a target sequence that overlaps with the -35 promoter element (Ushida and Aiba, 1990; Reviewed by Dove *et al.*, 2003).

The final simple activation type is Conformation Change where the activators actually bind to a target sequence between -10 and -35 promoter elements. In this case, the promoter is not optimal. These transcription factors belong to the MerR family (Reviewed by Brown *et al.*, 2003). The MerR family include a group of transcription factors that respond to transitions in metals and act as transcription factors. MerR TF is an example, which responds to mercury metal. MerR binds to the long spacer region between -10 and -35 elements (usually 19-20 bp) that is greater than the optimal spacer 17 ± 1 bases required for sufficient transcription initiation (Parkhill and Brown, 1990). Hence. MerR causes slight bend in the promoter sequence approaching the -10 and -35 sequences making appropriate for the RNAP to bind and for the transcription to be initiated (Browning and Busby, 2004). Since the spacer region between the -10 and -35 is not optimal, the activator binds there and enhances the promoter activity and binds a σ factor of RNAP (Fig. 1.7).

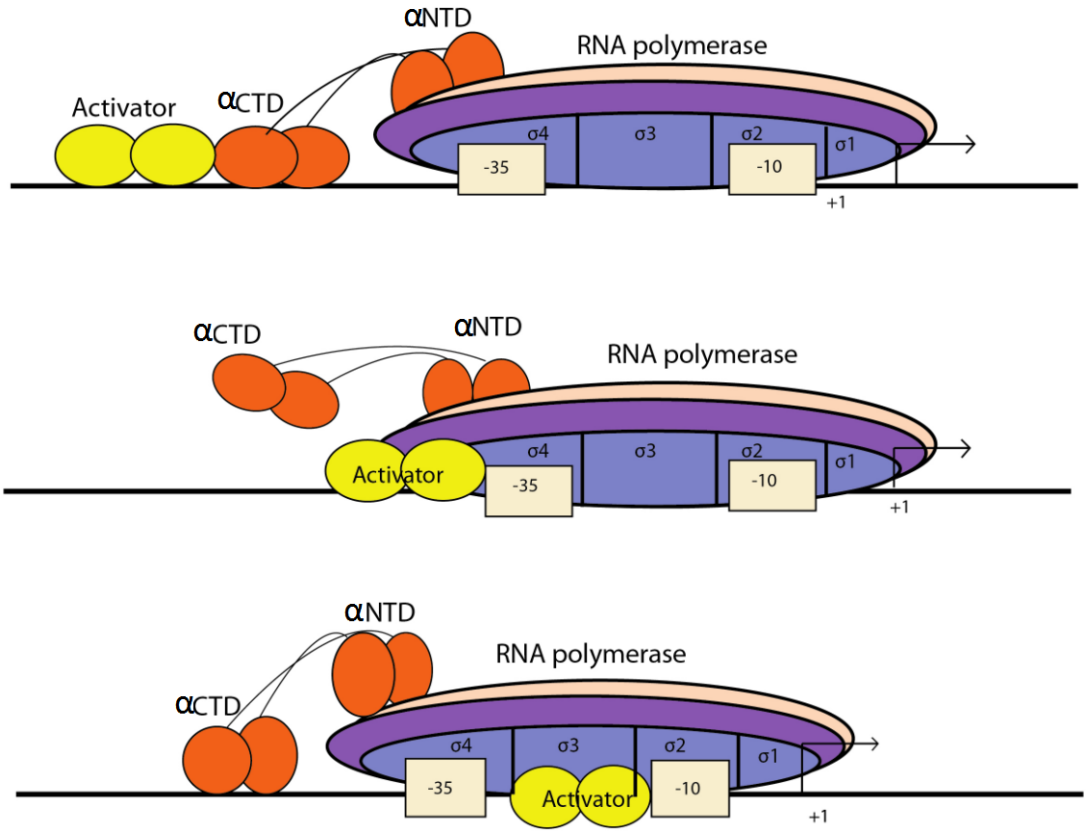


Fig. 1.6: Mechanisms of transcription initiation activation in prokaryotes.

A. Image shows the activator binding upstream from the binding sites of the αCTD terminal of the α subunit of the RNAP, hence recruiting the RNAP to the promoter elements.

B. In this image the activator overlaps with the α-CTD binding site to activate transcription initiation.

C. In this image the transcription factor binds to within the spacer region between the -10 and the -35 elements of the promoter to initiate transcription. (Figure redrawn from Browning and Busby, 2004).

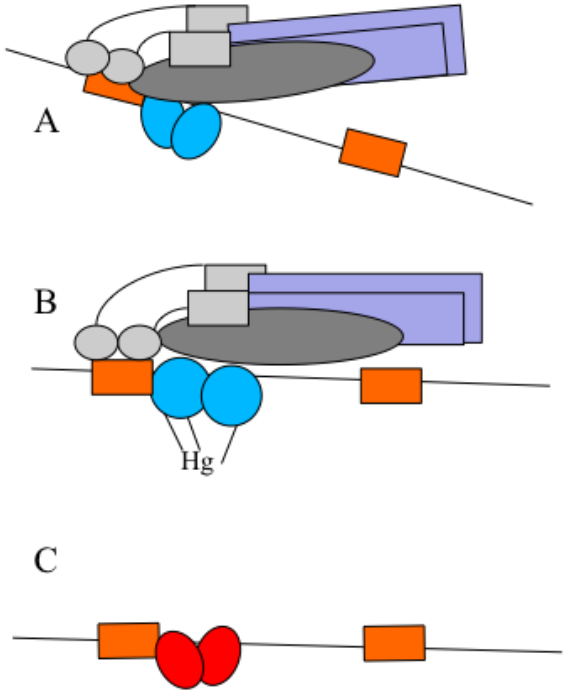


Fig. 1.7: The model for MerR action at *merOP* region. A: MerR binds to promoter elements, recruits RNAP, but open complex not formed. B: Hg(II) induces conformational changes after binding to ternary complex and transcription is initiated. C: promoter repressed by displacement of Hg-$MerR_2$ with MerD at low Hg (II) or non-metallated MerR binds to promoter elements giving structure in A. -10 and -35 shown in orange, RNAP holoenzyme are: grey σ, light grey α, purple, β, β'. MerR in blue, and MerD in red (Figure redrawn from Brown *et al.*, 2003).

Conversely there is simple repression of RNAP, transcription initiation repression can be accomplished by the involvement of a single repressor and there are three known mechanisms of simple repression; steric hindrance of the RNAP binding to promoter DNA is considered to be the simplest mechanism of simple repression (Fig. 1.8). The repressor-binding site is located in or close to the core promoter element. An example of this repression is the binding of the Lac- repressor at the *lac* promoter (Gilbert and Müller-Hill, 1966; Müller-Hill, 1996). The second mechanism of simple repression involves two different repressors binding to the promoter distal site, for instance GalR is a repressor of *gal* promoter (Choy and Adhya, 1996; Lewis and Adhya, 2015). In this case, the repression is caused by looping of the DNA. The third mechanism of simple repression involves inhibition of an activator by a repressor. An example for this mechanism is inhibition of CRP activation by CytR repressor (Shin *et al.*, 2001), where a direct interaction occurs between the activator CRP and the repressor CytR, preventing CRP-dependent activation (SØgaard-Anderson *et al.*, 1990; Browning and Busby, 2004).

1.2.6 Transcription initiation

Once the RNAP-P complex is in bound (close complex formation), the double helix DNA is unwound, forming an open complex (melted, or isomerised). After this stage, abortive RNA synthesis takes place: a certain length displaces the σ2, and 3 of the RNAP from the promoter. This process is believed to occur in a scrunching process where the RNAP unwinds and pulls down the DNA elements to the active site and abortive RNA synthesis takes place. Then the transcription is continued, in which the enzyme moves forward downstream from the transcription start site.

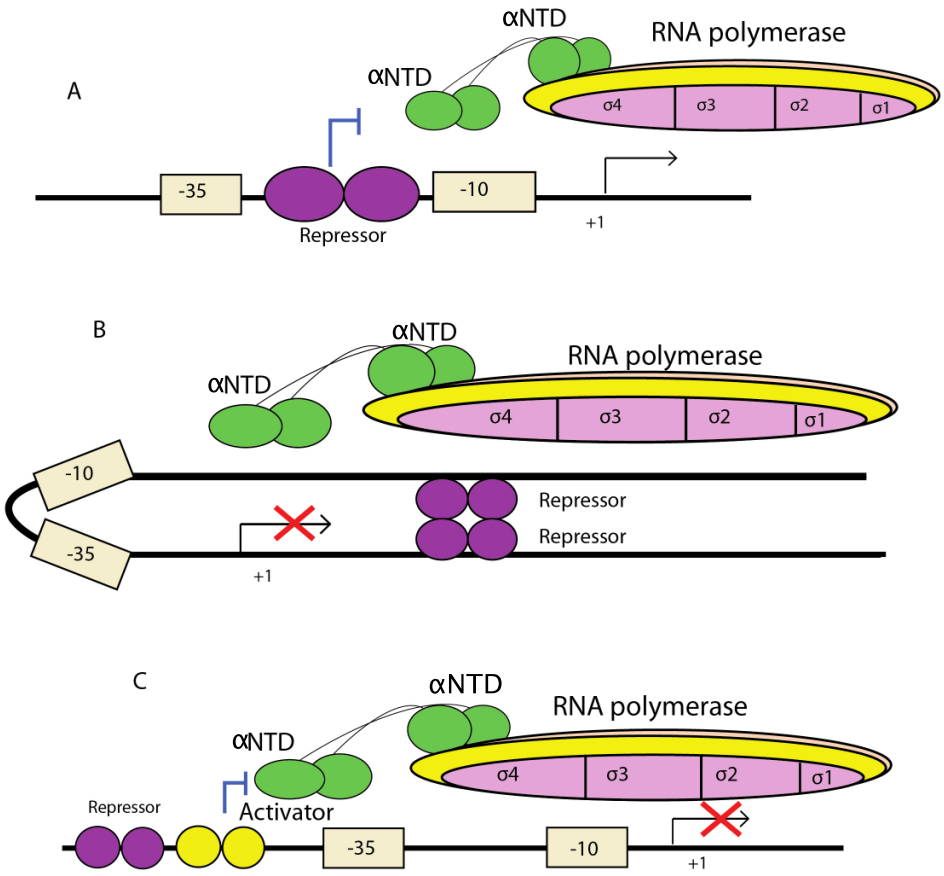


Fig. 1.8: Mechanisms of simple repression of transcription initiation in prokaryotes.

A. In this image repression is achieved by steric hindrance. The repressor-binding site overlaps with the promoter and elements, stopping the RNAP from binding.

B. In this image repression is achieved by the binding of the repressor to distal sites which interact by looping, repressing the intervening promoter.

C. In this image repression is achieved by modulating activator proteins. Here, the activator binding upstream from the promoter elements is trying to recruit the RNAP, but the repressor binds to the activator and prevents the activator from functioning. (Figures were redrawn from Browning and Busby, 2004).

After synthesis of mRNA is initiated, elongation of the transcript occurs, where the RNAP clears the promoter (Frank and Gonzalez, 2010; Yu and Oster, 2012). This process of elongation is believed to be a rapid stage, and usually the sigma factor would dissociate from the RNAP enzyme, but in some cases, the sigma factor remains bound to the enzyme whilst transcription is elongated (Kapanidis *et al.,* 2005). Although elongation of the transcription is smoothly continuous and a quick process, there are moments of transcription pausing. Transcription pausing could be due to the opening of the RNAP enzyme pincer off DNA, which results in pausing. The RNAP complex with the DNA elements is usually in a closed conformation, and the RNAP is functional in this state, but when the RNAP enzyme pincer is in open conformation (Chakraborty *et al.,* 2012), the transcription is paused. There are other reasons for transcription pausing, which include the role of transcription factors. Transcription factors, mentioned below can either lengthen or limit the extent of pausing during transcription elongation (Tagami *et al.,* 2010; Hartzog and Kaplan, 2011).

1.2.7 Distribution of RNAP between promoters

As mentioned earlier, the number of transcription units is more than the number of the available RNAP molecules to promote transcription, hence the need for economy and tight and careful regulation of transcription. A number of studies measured the amount of the RNAP distributed along the bacterial chromosome using ChIP techniques (Grainger *et al.,* 2005; Grainger and Busby, 2008; Sala *et al.,* 2009). The outcome revealed that the distribution of the RNAP along the DNA is variable and related to changes in the environmental signalling or growing conditions. For example, upon treating the cells with rifampicin, transcription by RNAP is blocked, and the measured amount of the RNAP can

reveal the strength of the promoter. So, strong promoters would bind to RNAP. This is termed the enzyme rate and is limited by recruitment. However, a weak promoter would bind less to the RNAP and after rifampicin treatment, only a small amount of the measured enzyme is bound. In these types of promoters, the rate of enzyme binding is limited by escape (Herring *et al.*, 2005; Grainger *et al.*, 2007). The distribution of the RNAP can also be dependent on the DNA elements. AT-rich regions can recruit the RNAP, preventing it from discriminating between the promoter elements and initiating the transcription (Jeong *et al.*, 2012). The H-NS silencing molecules can also influence the distribution of the RNAP. H-NS can silence the transcription by either blocking the RNAP from binding to a promoter, since the H-NS binds to the promoter elements, or the RNAP can be trapped by the H-NS (Lim *et al.*, 2012). Regulation of the RNAP is also performed at the open complex level. When the enzyme is bound, and the open complex is formed RNAP function can be disrupted by the interaction of the H-NS molecules and the αCTD terminal of the enzyme (Shin *et al.*, 2012). The specific interaction between the RNAP and a promoter element was further investigated using the more advanced technique, PING (protein-protein interaction network generation) (Gerber *et al.*, 2009; Meier *et al.*, 2013). This method was used to screen for the interaction of a number of transcription factors with the RNAP subunits. RNAP can be regulated at various stages of the transcription, such as in the process of elongation and, as mentioned in 1.2.6 in the description of transcription pausing, pausing could be undone. This is achieved by specific interaction between the RNAP core enzyme and CRE (core recognition elements) on the DNA, which was revealed by using the technique NET-seq (native elongating transcript sequencing) (Churchman and Weissaman, 2012; Vvedenskaya *et al.*, 2014). This suggested that the CRE are essential in the process of elongation of the transcription, by overcoming pausing as well as their known function in stabilising the formation of the open complex.

1.2.8 Role of nucleoid associated proteins in transcription regulation

Nucleoid associated proteins have DNA-binding abilities which allow them to shape and compact the DNA molecules by bridging, bending, and wrapping the DNA molecules, and they influence gene expression (Dillon and Dorman, 2010). This feature resembles the histones in eukaryotes (Drlica and Rouviere-Yaniv, 1987; Dorman and Deighan, 2003). The compaction of the DNA molecules by NAPs differs according to the growing state of the cells. In the exponential growth phase, the nucleoid is more looped due to the negative supercoiling, and the NAPs are scattered throughout the domains of the nucleoids (Deng *et al.,* 2005). DNA loops become more relaxed in stationary state and the transcription is reduced (Azam *et al.,* 1999). H-NS, 15.6 kDa, is one example of a NAP and has a major role in regulating gene expression, as it is believed to be the silencer of transcription machinery (Lucchini *et al.,* 2006). H-NS binds to AT-rich elements of the DNA, which is where the promoters are likely to be located, and it obstructs the accessibility of the RNAP enzyme to the transcription unit, and hence blocks gene expression. Additionally, H-NS binds to the curved parts of the DNA (Owen-Hughes *et al.,* 1992). The bacterial chromosome is folded and supercoiled in such a way that domains are formed and because the H-NS binds to the DNA domains, it can also be called a domain barrier as it marks the boundaries between the different chromosomal loops (Deng *at al.,* 2005; Postow *et al.,* 2004; Dame, 2005; Dame *et al.,* 2006; Noom *et al.,* 2007). FIS (factor for inversion stimulation) is another member of the NAPs that can also affect gene expression, as it is believed to target the promoters of the rRNA genes and like the H-NS, FIS also binds to AT-rich regions of the DNA but as a homodimer (Dorman, 2009; Cho *et al.,* 2008). Other proteins that function as regulators of DNA looping and compaction and gene expression include IHF, HU, MukB, Lrp, CRP, Dps,

etc. (Swinger and Rice, 2004; Claret and Rouviere-Yaniv, 1997; Niki *et al.,* 1991; Cui *et al.,* 1995; Browning and Busby, 2004; Almiron *et al.,* 1992; Johnson *et al.,* 2005).

1.2.9 Translation

After DNA is transcribed into mRNA, either after mRNA has been produced or during the process of making the mRNA, the process of translation is initiated. Translation is the process of polypeptide chain synthesis from mRNA and it involves a number of steps (Fig.1.8). First the translation initiation which involves the ribosomal subunits 30S and 50S, IF1, IF2, and IF3 (translation initiation factors), the mRNA which contains the ribosomal-binding site (Shine-Dalgarno sequence) which is generally located around 8 bases upstream of the start codon AUG. Translation start-sites, usually AUG, initiator tRNA (fMet-$tRNA_f^{Met}$), which carries N-formylmethionine fMet, which is believed to be the first amino acid of almost all proteins in the bacteria. Firstly, the 30S subunit binds to the elements on the mRNA. After that the first methionated tRNA and binds to the complex at the P-site of the ribosome (Laursen *et al.,* 2005; Moore, 2012), facilitated by IF2 (translation initiation factor 2) and energy source, GTP (guanosine tri phosphate). Then the 50S subunit, which encloses the complex, and the 70S ribosome, is then formed, and the translation initiated. The process of IF2-GTP binding to the 30S ribosome and recruitment of fMet-$tRNA_f^{Met}$ to the 30S ribosome is regulated by IF1 and IF3. These regulators make sure that fMet-$tRNA_f^{Met}$ is correctly positioned on the right start codon on the mRNA (Simonetti *et al.,* 2009), and this happens in 65% of cases. There are, however, other ways of initiating the translation complex where IF2-GTP first binds to the 30S subunit, then the fMet-$tRNA_f^{Met}$ is recruited to bind to the 30S and then finally, the 50S subunit, and this was

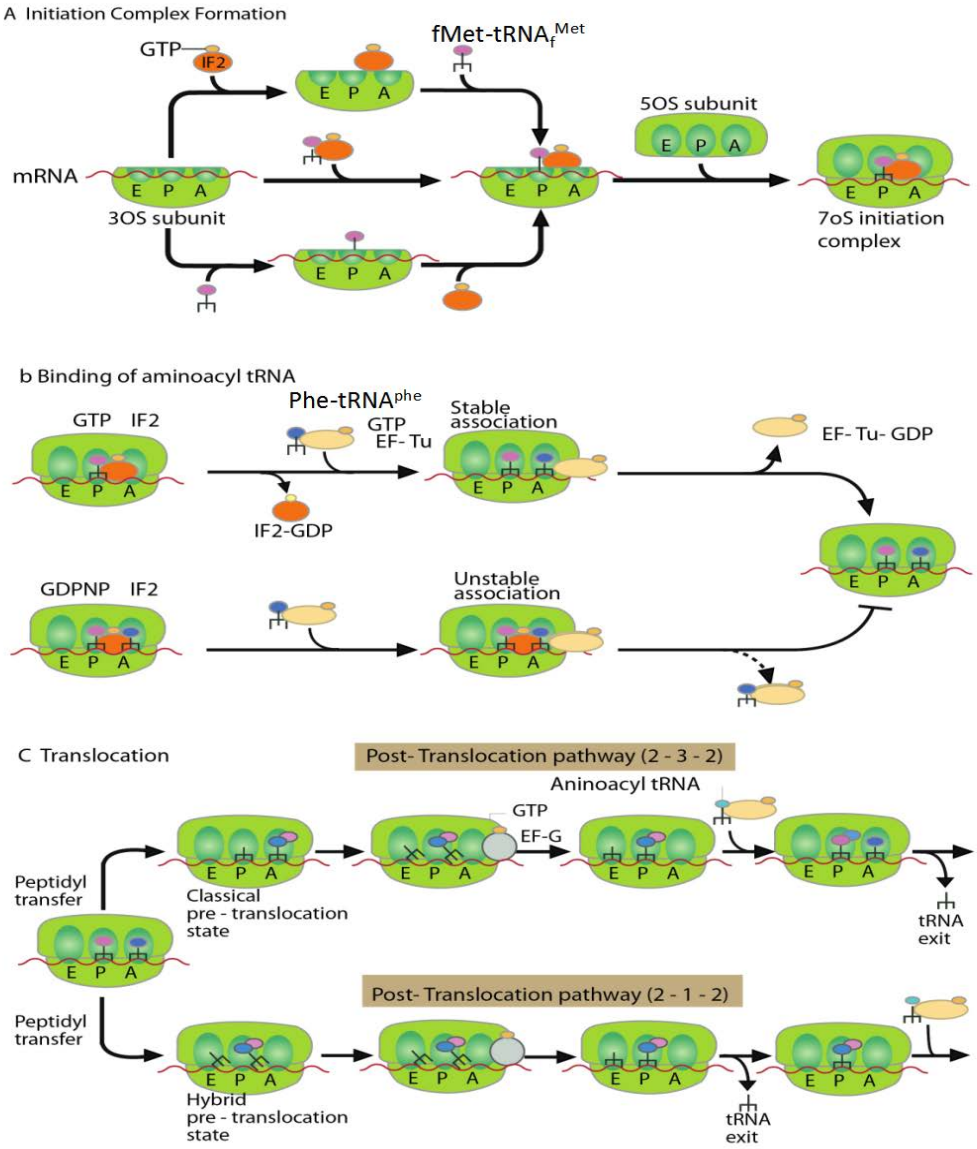


Fig. 1.9: Schematic representation of the translation process in prokaryotes.

A. Figure shows the initiation of the translation complex that involves mRNA, IF2-GTP, the 30S subunit, fMet $tRNA_f^{Met}$, IF1, IF3, and the 50S subunit.

B. Figure shows the recruitment of the first uncharged tRNA to the translational machinery and, for this complex to be stable, EF-Tu-GTP delivers the first uncharged tRNA to the ribosome, which dissociates later, leaving the complex stable. The unstable complex formation when the first uncharged tRNA is delivered by GDPNP-IF2 is also shown.

C. In this figure the process of translocating along the mRNA by the ribosome complex using two pathways, either 2-3-2 or 2-1-2 is shown (figure redrawn from Robinson and Oijen, 2013)

observed in 30% of cases. The other observation was that the IF2-GTP and the fMet-tRNAf^{met} bind at the same time to the 30S subunit, then 50S ribosome binding and this accounts for 5% of the cases (Milon *et al.,* 2012; Tsai *et al.,* 2012).

The process is then elongated and it involves the release of IF2 and GDP, and involvement of the first non-initiator aminocylated tRNA, facilitated by the GTP, and EF-Tu (elongation factor) (Tsai *et al.,* 2012). Here the second tRNA is located at the A position which is the entry position of the tRNA, apart from the first methionated tRNA that enters the ribosome from the P-site (peptidyl site). Now the complex is stable, with the tRNAs located where they should be, facilitated by the GTP energy supply. However, if GDPNP (a non-hydrolysable GTP analogue) were used instead of the GTP energy source to initiate the translation, the complex formed would be unstable. Hence, in the elongation process, when the first non-initiator tRNA binds, facilitated by the GTP and the elongation factor, it dissociates from the complex. Once the complex is stable, the translation is translocated. This is where the methionine charge at the first tRNA moves to bind to the aminocylated tRNA in the A-site (acceptor site), then the energy source GTP with EF (elongation factor) enters the ribosome, moving the first tRNA to the E-site (exit) and the second tRNA to the P site, and another new aminoacylated tRNA is brought in to the ribosome at the A site. This process of amino acid change during the elongation process is called a peptidyltransferase reaction. The choice of the amino-acid-charged tRNA is dependent on the code on the mRNA where every three bases represent a codon for an amino acid, which then specifies the choice of the amino acid. The ribosome complex keeps on translocating along the mRNA until it reaches a stop codon where no charged tRNA matching the codon is available. The peptide chain synthesis is then terminated. The peptide of the protein chain then undergoes posttranslational modification.

The process of exiting the ribosome at the E-site can be by one of two pathways: either 2-1-2 or 2-3-2 (Chen *et al.*, 2011). In 2-1-2 the ribosome is first occupied with two tRNA – one at the P-site and the other at the A-site –, then the tRNA is released at the E-site and only one tRNA is present at the P-site, until another tRNA binds at the A-site. In the other pathway, 2-3-2, two tRNA are at both the P and the E sites when another new tRNA enters at the A-site. Then a new tRNA enters at the A-site while the old tRNA at the E-site is still there. Then a new tRNA comes that pushes the tRNAs and the old tRNA at the E-site is released.

1.2.10 Regulation of translation initiation

Regulation of translation initiation can be performed by both *cis*- and *trans*- acting elements. *Cis*-acting elements such as secondary structures within the 5' untranslated region of the mRNA molecules near the translation initiation contain characteristic genetic elements known as riboswitches that can directly sense various metabolites without involving proteins (Mandal *et al.*, 2003). *Trans*-acting repressor elements can compete with 30S ribosomes for the SD-sequence hence, down-regulating translation. Secondary structures near the initiation codon that are induced or stabilised by repressor proteins can also cause inhibition of translation and result in unproductive pre-initiation complexes (Henkin, 2008). Secondary structures may contain short ORFs in the 5' untranslated long leader region. Translation of these ORFs can cause changes in the structure of the downstream region and eventually, can modulate the translation of downstream genes. Such ORF-dependent regulatory mechanisms depend on the coupling between transcription and translation (Vitreschak *et al.*, 2009).

1.3 Pathogenic *Escherichia coli*

The *E. coli* family is amongst the most studied among bacteria and it can be divided into non-pathogenic and pathogenic *E. coli* (Evans *et al.*, 193). Non-pathogenic *E. coli* include the ones that live in a symbiotic relationship with humans or animals, and are harmless. In fact, they are beneficial in helping with digestion (Saavedra, 2001; Ouwenhand *et al.*, 2002). The pathogenic ones however, can be divided into either intestinal pathogenic *E. coli* (Kaper *et al.*, 2004) or extra-intestinal pathogenic *E. coli* (Russo and Johnson, 2006; Smith *et al.*, 2007). Pathogenic intestinal *E. coli* causes infections in the gut of animals, including humans. Pathogenic *E. coli* can be in a number of types according to the type and site of infection; for example, include EHEC (enterohaemorrhigc *E. coli*), EPEC (enteropathogenic *E. coli*), ETEC (enterotoxigenic *E. coli*), EAEC (enteroaggregative *E. coli*), and (enteroinvasive *E. coli*) EIEC (Kaper *et al.*, 2004). However, *E. coli* can infect not only the gut, but also some other parts of an animal's body, and examples of the extra-intestinal pathogenic *E. coli* include UPEC (Uropathogenic *E. coli*), ECNM (*E. coli* neonatal meningitis), and APEC (Avian pathogenic *E. coli*) (Johnson and Stamm, 1989; Bingen *et al.*, 1998; Blanco *et al.*, 1997). *E. coli* is member of the *Enterobacteriaceae* genera, which include other bacterial species, which share the infection site, the gut. *Enterobacteriaceae* are Gram-negative bacteria. Besides *E. coli* the genera includes, *Klebsiella*, which are also present in the gut as normal flora. However, some species are pathogenic such as *Klebsiella pneumoniae*. *Salmonella* is another member of the *Enterobacteriaceae* genera and includes some serious infection-causing species of *Salmonella*, such as *Salmonella bongori*, *Salmonella enterica*. *Shigella* is also an important pathogen and there are a number of *Shigella* species that are all pathogenic, such as *S. dysenteriae*, *S. flexneri*, *S. sonnei*, and *S. boydii*. There are other members of

Enterobacteriaceae that cause infections in animals such as *Yersinia*, *Proteus*. (Hedegaard *et al.*, 1999; Paradis *et al.*, 2005).

Bacteria were present ages before the presence of humans. Bacteria have been present for over 3000 million years, but humans for only 4 million years. Thus, bacteria have long adapted and modified their ways of living according to their available environments. Bacteria can therefore be seen living in hot environments, cold environments such as those at the earth's poles, and in soil, water, and animals. To live in such environments, the bacteria adapted and evolved by acquiring genes through horizontal gene transfer and these genes are encoded on islands that incorporate themselves precisely within the bacterial chromosome, in such a way that these pathogenicity islands can be expressed. However, not all types of each bacterial strain have managed to acquire pathogenicity islands (Hacker *et al.*, 1990; Blum *et al.*, 1994). A major factor could be, the site of living. For example, *E. coli* can infect anywhere in the human body. In the urinary tract, where urine flow is continuous, the main virulence factor of UPEC is the pili structure that assures its attachment against the flow (Nataro and Kaper, 1998).

The human gut is one of the most complicated and variable ecosystems. *E. coli* organisms, and others, can resist the pH of the stomach with the help of local conditions such as being shielded by food, or because stomach acidity is low because of medication, while some bacteria combat this difficulty by being spore forming bacteria or acid resistant. In these circumstances, having survived the acidity of the stomach, *E. coli* can pass down to the intestine. Once they reach the intestine, adherence becomes a vital factor in ensuring their

survival. In this environment, the small intestine, oxygen availability is restricted, so bacteria that live here are normally facultative anaerobes (Poxton, 2010). *E. coli* is prolific in the environment. It can live in soil, water, or animal hosts and so it has had to adapt to different ranges of temperature. Infection by *E. coli* happens while eating raw or undercooked beef from cows previously infected with *E. coli*. In addition, infection with *E. coli* can be contracted by ingesting contaminated water or contaminated soil on a person's hands. The pathogen passes through the digestive gut; it can resist the strong acidic environment in the stomach but cannot live there, so it passes down to the intestine. It resides there, colonises, and produces pathogenic features. Infection with *E. coli* can be avoided by improving hygiene, eating well-cooked meat, and drinking sterilised water.

EPEC causes diarrhoea, in particular acute infantile diarrhoea in developing countries. It usually inhabits the small intestine (Kaper *et al.,* 2004) and can be a life-threatening organism (Kaper *et al.,* 2004). In addition to EPEC, EHEC, which are human pathogens, and are found to colonize the intestine, are responsible for 17 outbreaks and 75,000 cases of infection annually in the USA (Mead *et al.,* 1999). Infection with these strains results in a range of clinical complications, such as diarrhoea, haemorrhagic colitis in developed nations (Tarr *et al.*, 2005), and haemolytic uremic syndrome (HUS).

1.4 Virulence factors associated with pathogenic *E. coli*

Various pathogenic determinants contribute to the pathogenicity of both EPEC and EHEC; for example, production of Shiga toxin (Kaper *et al.,* 1997), colonization facilitated by flagella, especially chaperone/usher and Type IV fimbriae and outer membrane proteins (Tree *et al*., 2009), and the production of a Type III secretion system that causes histopathological

Table 1.1: Virulence factors responsible for colonization and fitness functions in *E. coli*[1]

E. coli type	Virulence factor	Role
IcsA (VirG)	EIEC	Nucleation of actin filaments
Intimin	EPEC, EHEC	Adhesion, induces TH1 response
Dr adhesins	DAEC, UPEC	Adhesion, binds to decay-accelerating factor (DAF), activates PI-3-kinase, induces MICA
P (Pap) fimbriae	UPEC	Adhesion; induces cytokine expression
CFAs	ETEC	Adhesion
Type-1 fimbriae	All	UPEC adhesion; binds to uroplakin
F1C fimbriae	UPEC	Adhesin
S fimbriae	UPEC, MNEC	Adhesin
Bundle-forming pilus (BFP)	EPEC	Type IV pilus
Aggregative adherence fimbriae	EAEC	Adhesin
Paa	EPEC, EHEC	Adhesin
ToxB	EHEC	Adhesin
Efa-1/LifA	EHEC	Adhesin
Long polar fimbriae (LPF)	EHEC, EPEC	Adhesin
Saa	EHEC	Adhesin
OmpA	MNEC, EHEC	Adhesin
Curli	Various	Adhesin; binds to fibronectin
IbeA, B, C	MNEC	Promotes invasion
AslA	MNEC	Promoters invasion
Dispersin	EAEC	Promotes colonization; aids mucous penetration
K antigen capsules	MNEC	Antiphagocytic

Aerobactin	EIEC	Iron acquisition, siderophore
Yersiniabactin	Various	Iron acquisition, siderophore
IreA	UPEC	
IroN	UPEC	
Chu (Shu)	EIEC, UPEC, MNEC	Iron acquisition, haem transport
Flagellin	All	Motility; induces cytokine expression through TLR5; >50 flagella (H) serotypes
Lipopolysaccharide	All	Induces cytokine expression through TLR4; >180 O types

[1] CFA, colonization factor antigen; CS, coli surface antigen, MICA, MHC class I chain-related gene A; PCF, putative colonization factor; PI-3-kinase, phosphatidylinositol 3-kinase; TLR, Toll-like receptor.

lesions on the intestinal epithelial. These are known as attaching/and effacing lesions (AE). The lining microvilli of the intestinal cells are destroyed and their cytoskeleton is rearranged for the benefit of the bacterial cells (Nataro & Kaper, 1998). The scenario of invading host cells by pathogenic microbes, including *E. coli* is a multistep process. In each step, a number of virulence factors are employed which are either expressed from the bacterial chromosome or are plasmid-encoded (Yamamoto and Yokota, 1983; Turner *et al.,* 2006). It starts with recognizing the site of infection and attaching to it, reproduction within host cells possibly, invasion of the host cells, escaping from the immune system, and initiating apoptosis to the host cells (Mims *et al.*, 2001).

Among these steps, adherence is one of the most important stages in infection. Various virulence factors found in pathogenic *E. coli* are involved in one or more functions to aid in bacterial survival in the infection site, such as the Type I pili virulence factor in EPEC. It was found that this is responsible for the rise in IL-6 (interleukin 6) production from the host cells during infection for a short time of about 2 hours. Other factors were tested, but failed to augment the IL-6 host response in this way (Schilling *et al*., 2000). It was also shown that adherence factors are not enough to ensure invasion of the epithelial lining, so the pathogen requires further factors to invade and in this case using Type I pili. Additionally, Hick *et al* used IVOC cells to test the pathogenesis of EPEC rather than HEp-2 cells because IVOC are freshly harvested cells from children's intestines. It was shown that the EPEC strains that express intimin, a virulence factor responsible for tight attachment to the mucosal cells, were able to adhere to IVOC even though they were deficient in BFP (bundle forming pili) which is expressed from *bfp* on EAF (EPEC adherence factor) plasmid and is responsible for the initial attachment to host cells. It is believed that the role of BFP involves the formation of

complex three-dimensional colonies through bacterial interconnections. This finding demonstrates that the function of BFP is the formation of a three-dimensional colony but it is not responsible for the initial attachment of EPEC to the host cells (Hicks *et al.*, 1998).

1.5 The *LEE* pathogenicity island

As mentioned earlier, pathogenic bacteria have evolved over time by acquiring pathogenicity islands through horizontal gene transfer. This study is interested in the *LEE* pathogenicity island (Fig. 1.9). This is because this pathogenicity island is responsible for the production of Type III secretion machinery that extrudes from the bacterial surface toward the host cells and then injects effector molecules and toxins that modulate the immune response and initiate infection. *LEE* stands for locus of enterocyte effacement. The *LEE* pathogenicity island, 36.5 kb encodes 41 genes which are arranged mainly in five operons: *LEE1* (contains 9 open reading frames, where the first one, *ler*, is believed to be the most important in *LEE* operons); *LEE2* (contains 6 open reading frames); *LEE3* (contains 7 open reading frames); *LEE4* (contains 8 open reading frames); and *LEE5* (contains 3 open reading frames). GrlR and GrlA are encoded from a bisictronic operon located between *LEE1* and *LEE2* operons. The other genes are scattered between the operons and their expression is derived from their own promoter elements (Deng *et al.*, 2004). Each operon encodes proteins that either makes up the Type III secretion system (T3SS) apparatus, or encodes the effector proteins that pass through the apparatus to the host cells (Elliot *et al.*, 1998).

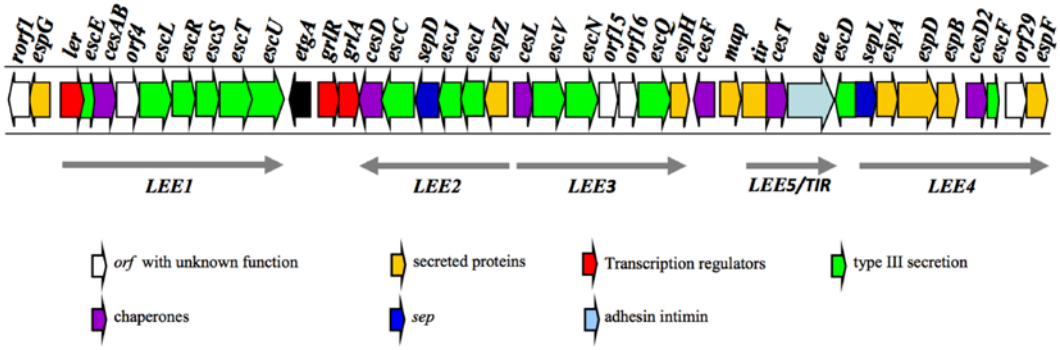


Figure 1.10: Organization of genes in the locus of enterocyte effacement.

The figure shows the arrangement of 41 genes in the *LEE* region. Genes shown are organised in five major polycistronic operons: LEE1, 9 genes; LEE2, 6 genes; LEE3, 7 genes; LEE4, 8 genes and LEE5 or TIR, 3 genes. Orf, open reading frames with unknown functions; ler, encodes LEE-encoded regulator; esc, encodes *E. coli* secretion component (homologous to Yersinia type III secretion, ysc); grlR, encodes global regulator of LEE repressor; grlA, encodes global regulator of LEE activator; ces, encodes chaperone of *E. coli* secretion; sep, encode secretion of *E. coli* proteins (not ysc homolog but involved in type III secretion); esp, encode *E. coli* secreted proteins; map, encodes mitochondrial associated protein; tir, encodes translocated intimin receptor protein; eae (*E. coli* attaching and effacing) encodes intimin; etgA, encodes a lytic transglycosylase. This figure has been adapted from Castillo et al. (2005). Nomenclature of the genes has been given mainly following Pallen et al. (2005a)

Genes in operon *LEE1*, *LEE2*, and *LEE3* encode structural and auxiliary proteins that are vital for composing the T3SS. *LEE4*, on the other hand, encodes effector proteins such as EspA, EspB, EspD, and EspE. Finally, *LEE*5 encodes adhesion factors called intimin and Tir. The *LEE* pathogenicity island is believed to determine the attaching and effacing feature of EPEC, EHEC, and other bacteria such as *C. rodentium*. It was also found that the *LEE* pathogenicity island is highly conserved among pathogens on the basis of linear gene order, nucleotide sequence, and even at the level of the predicted protein sequence, which suggests that the *LEE* pathogenicity island originates from a common ancestor.

Despite the fact that there are similarities between the *LEE* pathogenicity islands among bacteria harbouring it, there are still great differences among them too. These differences are caused by the different bacterial-host interaction. So, this interaction might, for example, determine which genes are needed. The content of the GC in the *LEE* pathogenicity island is less than that of the *E. coli* K-12 strain chromosome (Elliott *et al.,* 1998; Perna *et al.,* 1998; Zhu *et al.,* 2001). In fact, the loss and gain of the pathogenicity island is documented in a number of pathogenic *E. coli* and the gaining of new mobile elements through recombination, in particular, is clearly prevalent in the EHEC strain O157:H7 due its poor “mismatch repair” mechanism (LeClerc *et al.,* 1996). Because the *LEE* pathogenicity island was acquired through horizontal gene transfer (Reid *et al.,* 2000), and believed to be an evolutionary development, the insertion sites of the pathogenicity island within the chromosomes of pathogenic bacteria vary. For example, in one of the EPEC strains, O127:H6, it was inserted at 82 min downstream from the gene encoding the tRNA for selenocysteine (*selC*) (McDaniel *et al.,* 1995). In a different strain, O157:H7, the *LEE* was inserted at 94 min in the gene encoding the tRNA for phenylalanine (*pheU*) (Sperandio *et al.,* 1998).

1.5.1 Regulation of *LEE* transcription unit

Fig. 1.10 shows a complex regulation of the *LEE* pathogenicity island, which contains global and *LEE* specific regulators. The global regulators are IHF (integration host factor), FIS (Factor for inversion stimulation), BipA (Ig heavy chain binding protein), H-NS (histone-like nucleoid structuring), quorum sensing QseA, and GadX (glutamate-dependent acid resistance). The specific regulators, by comparison, are Per (plasmid encoded regulators) protein regulators expressed from EAF (EPEC adherence factor) plasmid, and Ler (locus of enterocyte effacement regulator) as well as other regulators such as GrlR (global regulator of *LEE* repressor), and GrlA (global regulator of *LEE* activator) (Kaper *et al.,* 2004). Despite the complicated and tight regulation made possible by multiple systems, Ler is believed to be the most important regulator. This is because Ler has a global regulatory role, regulating gene expression from all the *LEE* operons either positively or negatively. Details on the role of Ler are found below. H-NS plays an integral role in silencing the *LEE* genes and it responds to a number of environmental signals, mainly temperature, $HCO3^-$ levels (Mellies *et al.,* 2007; Umanski, *at al.,* 2002). H-NS binds to *LEE1, LEE2,* and *LEE3* promoters, which result in direct control of their expression (Mellies *et al.,* 2007). This repression is relieved under some conditions and T3SS is produced (Umanski, *at al.,* 2002). Accordingly, regulation of the *LEE* transcription unit is dependent on environmental signals and on growth phase conditions. So at 37° C the *LEE* genes are expressed, but at lower temperatures like 27° C, *LEE* genes are repressed by H-NS (Umanski *et al.,* 2002). As far as growth phase is concerned, when cells are grown in minimal media, the *LEE* genes are expressed at the late exponential phase and are repressed at the stationary phase (Bergholz *et al.,* 2007).

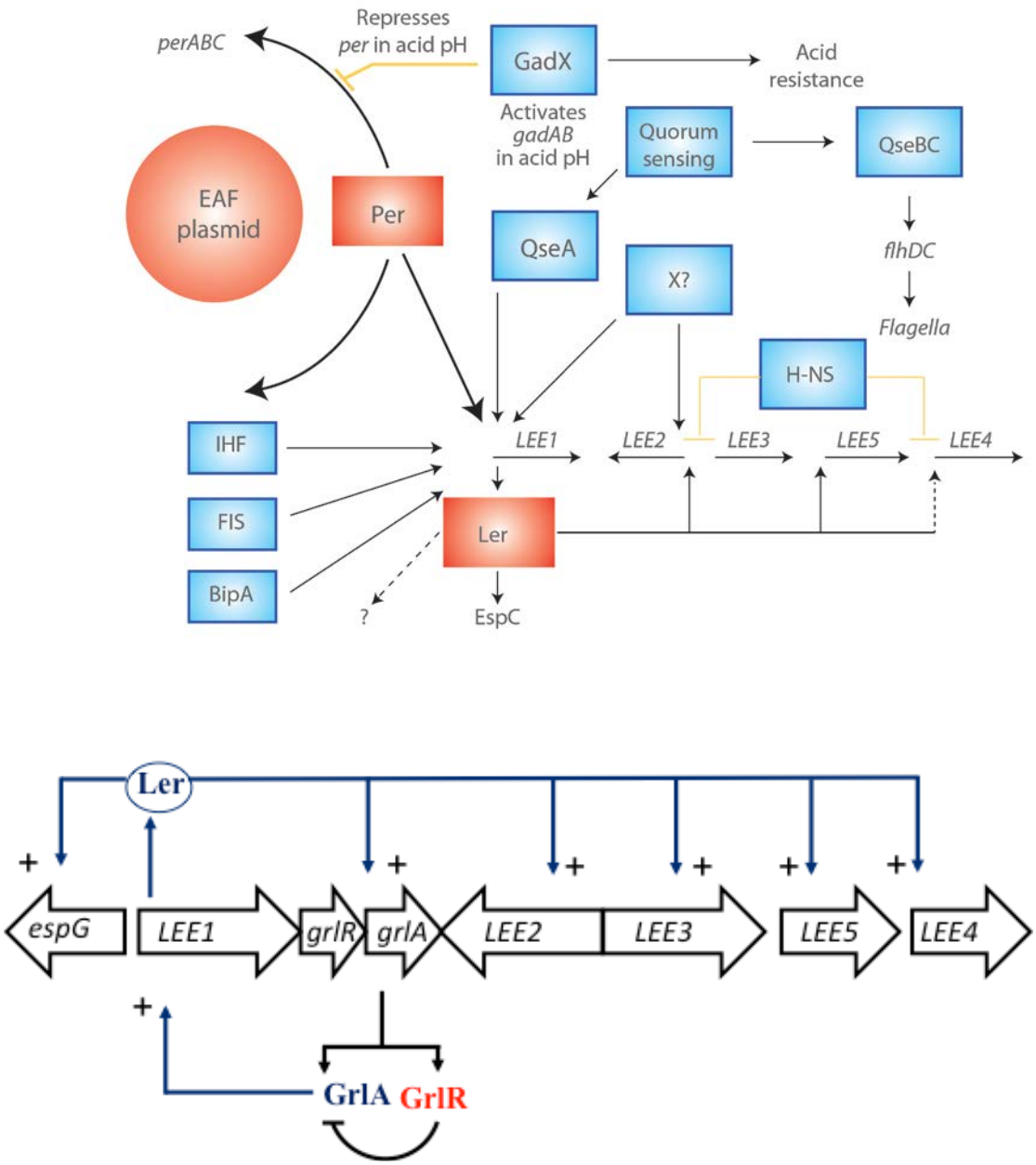


Fig. 1.11: A schematic representation of the regulation of the *LEE* operons. The tope diagram shows complex network regulatory molecules that modulate activity of the *LEE* operons are shown. LEE can be regulated by *LEE*-encoded regulators, mainly Ler, or non-*LEE* encoded regulators, such as Per. *LEEs* are also regulated by global regulators, H-NS, IHF, FIS, BipA, QseA, and by stress response regulators. (Figure redrawn from Kaper *et al.*, 2004). In lower diagram, shown is regulation of expression of Ler by GrlR and GrlR.

However, when using rich growing media like LB, the *LEE* genes are expressed during the transition into the stationary phase (Sperandio *et al.,* 1999; Sharma and Zuerner, 2004). Taking EPEC and EHEC as examples of pathogenic bacteria harbouring the *LEE* pathogenicity island, it is possible to see that the *LEE* genes are expressed differently from each strain depending on growth phase. So in a mid-log phase cell in EPEC, 85 genes are transcriptionally regulated by Ler both inside and outside the *LEE* region. This is mainly activation regulation. But in mid-log phase, one gene is reported to be regulated by Ler. And in the late log-phase, Ler regulates 97 genes. Looking at EHEC however, Ler regulates 39 genes in mid-log phase. 35 of them are located within the *LEE* region itself, while the rest are outside the *LEE* region (Bingle *et al.,* 2014). The stress response regulators noted above (GadX and other regulators) also have a regulatory role in gene expression from the *LEE* (Tatsuno *et al.,* 2003; Vanaja *et al.,* 2009), and it was found that expression of some of the *LEE* genes is enhanced in strains lacking the *gadE*, namely, *ler*, *espA*, and *tir* (Branchu *et al.,* 2014), suggesting that GadE is a repressor of *LEE2*, *LEE4* and *LEE5*. In the *gadX* mutant, expression of *ler* was enhanced, but expression from the *espA* and *tir* were repressed (Branchu *et al.,* 2014). Clearly, regulation of *LEE* regions is dependent on a number of factors that are either *LEE*-encoded or from outside the *LEE* region and LEE-encoded regulators appear to influence genes outside of *LEE*.

1.5.2 Attaching and effacing feature of EPEC and EHEC

Upon formation of the T3SS, effector proteins expressed from the *LEE* 4 operon are injected into the host cells, including translocon *E. coli* secretory protein A (EspA). Exporting EspA protein requires both SepL and SepD; a deletion in *LEE2* and *LEE4* genes causes the injection

of EspA to fail and instead of injecting EspA into the host cells, it is injected in the extracellular space.

SepL and SepD prevent the delivery of effector proteins into host cells while the Type III secretion systems compartments are still in formation, possibly by binding to the Tir protein. After the secretion of translocon, the cells switch to secrete effector proteins such as EspH and NleH that are dependent on CesT chaperones (Tree *et al*., 2009). Protein expressed from *LEE* 5 is crucial for the tight attachment of both EPEC and EHEC to the host cells. This intimate attachment to the host cells is caused by intimin, which is expressed from the *eae* gene in the *LEE*5 operon (Jerse *at al.,* 1990). *Eae* mutant strains failed to adhere to HEp-2 cells, unlike the wild type, which adhered after 6 hours of incubation, and the tight attachment feature was restored when *eae* was supplemented in the bacterial cells by a plasmid (Donnenberg *et al.,* 1993). The ability of both EPEC and EHEC to adhere to the microvilli lining of the intestine is enhanced by some proteins that are expressed from different locations within both strains. In EPEC, EPEC Adherence Factor (EAF) contains plasmid-encoded regulatory genes (*per*), which are called *per*A, *per*B and *per*C, which are responsible for the activation of bundle-forming pili and for the activation of *LEE* genes. In EHEC however, these genes are located on the EHEC chromosome and are called *per*C homologues. They consist of five genes, *pchA, B, C, D* and *E* and encode for 104 aa. A single deletion mutation in *perC* homologues *pchA*, *pchB* or *pchC* reduces the adhesion of the EHEC to the epithelial cells. Double deletion mutation in the *perC* homologues further decreases the adherence of the EHEC to the epithelial cells in contrast to a single deletion. The reduction in adherence was also observed in *ler* (*LEE* encoded regulator) mutant cells since the adherence feature of the EHEC was restored in both single and double *pchA/B/C* mutants with a plasmid carrying

the *ler* gene. These findings show that adhesion of EHEC to the epithelial cells by PchA/B/C is facilitated by Ler expression (Iyoda *et al*., 2004). However, PerA-C are required for activation of *LEE*. In addition to *LEE* encoded regulators that facilitate the attachment to the host cells, and PerC homologues, it was found that curli fibres play a role in the attaching and effacing of both EPEC and EHEC (Saldaña *et al.*, 2009). The curli fibre in both strains was assembled in the same growing conditions as the Type III secretion systems, which suggests that curli fibre could be produced in the host cells. The curli fibres are defined as thin adhesive structures; they are coiled, highly aggregative fibres of variable length that extend from the bacterial surface as amorphous matrices (Olsen *et al.*, 1989; Collinson *et al.*, 1991; Collinson *et al.*, 1992; Provence and Curtiss, 1992). The ability of both EHEC and EPEC to adhere to host cells was not significantly reduced in strains that lack curli fibres, indicating that both EPEC and EHEC rely on other adherence factors for example, intimin which is produced from *LEE5* operon. Moreover, the adhesion of curli-fibre-producing strains is not efficient when cellulose is absent, suggesting some synergy between curli fibre and cellulose in adhering to host cells (Saldaña *et al.*, 2009).

1.5.3 Type III secretion system

T3SS (Type III secretion system) machinery (Fig. 1.11) is employed by a number of plant and animal Gram-negative pathogens to translocate proteins from the bacterial cytoplasm into the host cells (Galán and Collmer, 1999). This machinery delivers proteins across the bacterial inner membrane, the peptidoglycan layer, the outer membrane and the plasma membrane of the host cells in animals as well as across the cell wall in plants (Ghosh, 2004). It is composed of 20 proteins. These proteins can be divided into two groups; one group contains outer

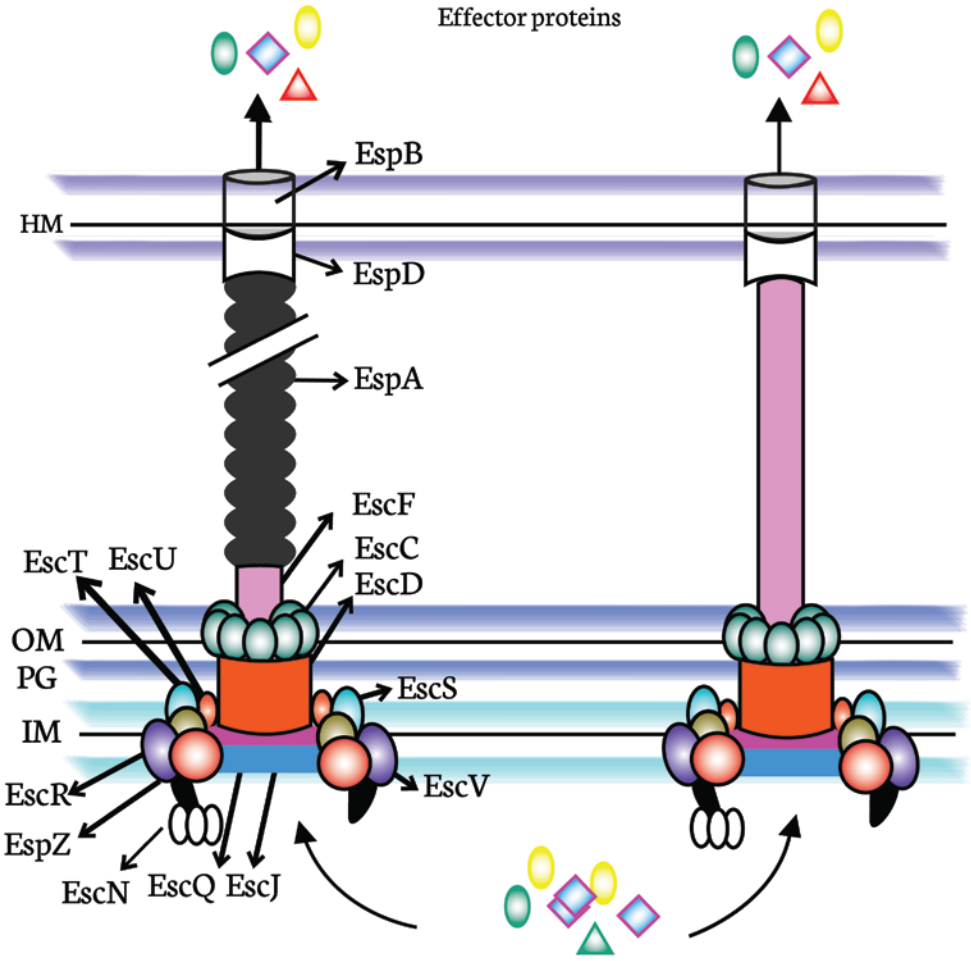


Figure. 1.12: Model for the type III secretion system. The system is composed of a basal body that is a multi-ring structure spanning the inner and the outer membrane, a needle structure extruding outward and used as a channel to deliver effectors molecules, and finally, translocation proteins at the tip of the needle. Shown also are the chaperone-associated proteins in the cytoplasm prior to dissociation and exporting. (Figure redrawn from Izore *et al.*, 2011). HM=host membrane, OM=outer membrane, IN=inner membrane, PG=peptidoglycan.

membrane proteins, which have similar sequences to the secretin proteins and some lipids. The second group contains integral membrane proteins that form the typical needle structure. The needle complex is a structure that spans both the inner and outer membrane of the bacterial envelope and is a hollow structure 120 mm long, which is composed of two domains. The first domain is a needle-like protein projecting outward from the cell surface. The other domain however, is a cylindrical base that anchors the structure to the outer and inner membranes (Galán and Collmer, 1999; Pushar and Sansonetti, 2014). The proteins secreted through this needle require chaperone proteins for protection from premature interaction with other proteins in the system and requires an energy source from the cytoplasm and the membrane to export proteins (Hueck, 1998).

1.5.4 Regulation of *LEE* by Ler

Ler is a 15-kDa protein encoded by the first gene of the *LEE1* operon. The *LEE* encoded regulator (Ler) is a central regulator that is needed for expression of all *LEE* genes. H-NS represses expression of *LEE* genes, whereas Ler activates them by counteracting the action exerted by H-NS (Sperandio *et al.,* 2000). The C-terminal domain of Ler exhibits amino acid homology to H-NS C-terminal domain. It is found that the CT-Ler does not participate in base specific contacts. It recognizes structural features in the DNA minor groove. This mechanism of contact is known as indirect readout. The key residue of the CT-Ler interaction with DNA is found to be Arg90 and found to be crucial. The side chain of Ler is inserted deep down in the minor groove (Tiago *et al.,* 2011). *Ler* regulation (Fig. 1.12) is performed by global and A/E specific regulators. IHF is essential for *ler* activation. IHF binds to a DNA region upstream from the *ler* promoter. Additionally, *ler* expression is up-regulated by other global

regulators, for instance, BipA. Other factors include FIS and QseA (quorum sensing *E. coli* regulator A). H-NS and Hha negatively regulate expression from *ler* by directly binding to its regulatory region. Specific regulators include plasmid encoded regulator C (PerC), which is a product of the third gene of *per* in the EPEC adherence factor plasmid positively regulate expression of *ler*. GadX regulates the expression of the *perABC* operon, so that it regulates expression from *ler* indirectly. Ler binds to its promoter and autorepresses it. Other *LEE* regulators, YhiE, and EtrA (*E. coli* Type III secretion systems 2, regulator A), negatively regulate *LEE* by possibly altering GrlA and GrlR. In addition, GrlA and GrlR positively and negatively regulate *LEE* genes respectively (Barba *et al.*, 2005). IHF, GrlA, and PerC coordinate to activate *ler* expression by counteracting the silencing effect of H-NS (Bustamante *et al.*, 2011), and *ler* activation is optimum when either GrlA or PerC regulate with IHF (Porter *et al.*, 2005). Ler regulates non-LEE genes positively or negatively. In the case of haemolysin genes (*ehx*) located on pO157, it was found that Ler can directly up-regulate expression of EhxC independently of other regulators such as GrlA when GrlR was present; however, in strains lacking *grlR,* coordinated activation of *ehxC* by both Ler and GrlA was detected (Iyoda *et al.*, 2011).

1.5.5 Regulation of *LEE* region by GrlR and GrlA

Orf10 and *orf11* encode regulators that affect transcription of *ler* both negatively and positively respectively. *Orf10* was named *grlR* and *orf11* was named *grlA* (Deng *et al.*, 2004). GrlA regulator is conserved in EPEC, EHEC, and *C.rodentium* (Elliott *et al.*, 1999). Attempts to purify GrlA and crystalize it failed (Islam *et al.*, 2011). However, GrlR was purified and a crystal structure was obtained (Jobichen *et al.*, 2007). GrlR is involved in the

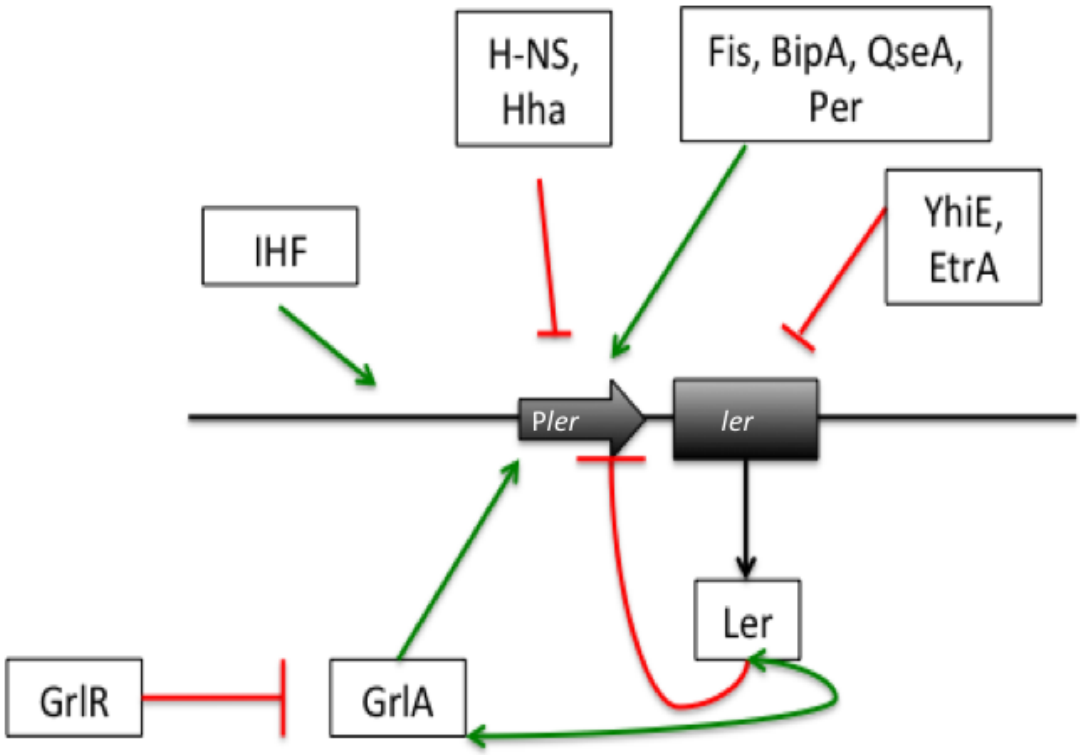


Fig. 1.13: Schematic representation of regulation at *Ler*. Shown are global and specific regulators of *Ler*. Regulators act either upstream promoter, at the promoter, or interact with the Ler protein.

regulation of *LEE* genes. It is expressed from an ORF that is located between *LEE1* and *LEE2* operons along with GrlA regulator. GrlR negatively regulate expression from *ler* by counteracting GrlA preventing it from binding to *ler* promoter and activate transcription.

It was found that GrlRA regulators activate the transcription from *LEE2* and *LEE4* operons independently of Ler, while other operons are not affected (Russell *et al.,* 2007). It was found that GrlA targets P1 in the *LEE1* regulatory region and results in enhanced promoter activity, and thus *ler* expression (Islam *et al.,* 2011). In addition, GrlA activates the LEE1 regulatory region by both counteracting H-NS repression and by acting independently of H-NS (Jimenez *et al.,* 2010; Islam *et al.,* 2011). It was evident that acitvation of thr *LEE1* regulatory region by GrlA is achieved by GrlA binding to the spacer region between -35 and -10 elements, a feature seen in the MerR family (Islam *et al.,* 2011). Also, GrlA activated expression of the *LEE1* regulatory region when the bacterial culture was grown in a static +CO_2 condition. By contrast, when the bacterial cutlure was set to static conditions, a reduction in *ler* expression was obsereved (Bustamante *et al.,* 2011). Expreesion of *LEE4* and *LEE5* were regulated positvely by GrlA in cultures grown under shaking condtions.

1.5.6 Regulation of *LEE* region by quorum sensing molecules

Host hormones such as epinephrine and norepinephrine control expression of the production of T3SS through the quorum-sensing regulator QseA (Sperandio *et al.,* 2002, Sperandio *et al.,* 2003). It was found that inactivation of the *qseA* gene resulted in decreased production of T3SS and transcription of *LEE* operons (Sperandio *et al.,* 2002). This regulator controls the

transcription level of both *ler* and *grlRA.* It does so by responding to bacterial-compound-designated autoinducer-3 (AI3) (Walters and Sperandio, 2006). Other regulators that controls *LEE* include the two-compontent system, QseEF. QseE is a histidine kinase and QseF is a response regulator responding to epinephrine, sulphate, and phosphate. QseEF transcriptionally regulate genes involved in pedestal formation by controlling the expression of genes encoding the T3SS effector EspFu that in turn is vital in actin pedestal formation in EHEC (Mellies *et al.,* 2007). QseG regulator is encoded from *qseG* located between *qseE* and *qseF*. *QseG* encodes an outer membrane protein, QseG, which is proposed to be crucial for the translocation of the Tir effector protein by an unknown mechanism (Tree *et al*., 2009).

1.5.7 Horizontally acquired regulators

Other regulators and virulence outside the *LEE* region control the expression of T3SS. These regulators are believed to be acquired through horizontal gene transfer and to be introduced into the bacterial cells by lysogenic bacteriophages. Sakai, an EHEC strain contains additional regions in the chromosome known as the O island and S-loop. One of these proteins encoded on the OI includes PerC homologues and, as mentioned previously, there are five homologues present in the EHEC chromosome: *pch*A, *pch*B, *pch*C, *pch*D, and *pch*E. *PchA* in particular is negatively regulated by an RcsBCD phospho-relay system, but it is regulated positively by LrhA (LysR-type regulator). LrhA also positively regulates *pchB,* hence *LEE* operons (Tree *et al*., 2009; Honda *et al.,* 2009). This activation was found to be dependent on both Ler and GrlA (Iyoda *et al.,* 2006). PerC can activate expression from *LEE* regulatory regions when the bacterial cultures are grown under static conditions in the presence of $+CO_2$; however, a decrease in activity was seen when the bacterial culture was grown while being shaken

(Bustamante *et al.,* 2011). The same effect of PerC on *LEE4* and *LEE5* was measured and it was concluded that PerC activates these operons when the bacterial culture is grown under static conditions with $+CO_2$. PerC homologues activate *ler* expression during growth in static conditions of $+CO_2$, and when *ler* is expressed, the gene product activates GrlA which results in further activation of *ler* transcription, creating a positive regulatory loop (Barbra *et al.,* 2005).

1.5.8 SOS induction

Prophages respond to threats to the host bacterial cells and they monitor the SOS response in the bacterial cell that is stimulated by DNA damage. In short, the process starts by auto-lytic cleavage of the phage repressor cI by RecA. The results of this cleavage are seen in the transcription of the late gene anti-terminator Q and transcriptional anti-termination of the transcript produced from the $P_{R'}$ promoter. In EPEC, during DNA damage, a DNA-damaging agent enhances the expression of LEE2/3 operons and other non-LEE genes such as *nleA.* This is accomplished through de-repression of the promoter by RecA-mediated autocatalytic cleavage of LexA. As a result, the transcription of the LEE operon is increased (Tree *et al.,* 2009).

1.6 Role of *LEE* encoded protein in regulation of other virulence factors of *E. coli*

The LEE-encoded regulators have a regulatory role outside the LEE pathogenicity island. This role is to ensure concise system production and it further implies the tight and careful regulation of this system. Production of T3SS is dependent on other prior mechanisms such as

movement and, within the *LEE* operons, regulators that can actually regulate the flagella function.

1.6.1 Flagellar operon regulation by GrlRA

In *E. coli* and *Salmonella*, 50 flagellar genes are organized into more than a dozen operons. Flagellar genes are regulated negatively and positively by a number of proteins including ClpXP protease, GrlA, and GrlR. In *Salmonella*, ClpX and ClpP ATP-dependent proteases were found to negatively regulate the flagellar genes. On the other hand, ClpXP positively controls *LEE* expression by negatively regulating GrlR (Fig. 1.13). In *clpXP* mutant strains, FliC protein is expressed in large amount as detected by SDS-PAGE. Therefore, ClpXP negatively regulates FliC expression in EHEC. Moreover, the amount of FliC decreases significantly when *grlR* is inactivated. Thus, the expression of flagellar proteins is positively regulated by GrlR. In addition, the FliC protein level is reduced when cells are growing in DMEM when *LEE* expression is induced, even when the strain is a *clpXP* mutant. Regulatory genes for flagellar regulon also control the expression of Type III secretion systems and subsequent virulence phenotypes in some bacteria. GrlA negatively regulates expression of the flagellar genes in EHEC by reducing transcription of *flhD*. It inversely regulates the expression of flagellar, and also LEE-encoded expression in EHEC, which could be vital for efficient cell adhesion to EHEC. GrlA inhibits FliC expression in *grlR* mutant strains. This repression is not Ler-dependent since FliC is repressed by GrlA in *ler* mutant strains. Another pathogenic feature of EHEC that is under GrlA and GrlR regulation is motility. EHEC motility is reduced in *grlR* mutant strains but not in *grlA* or *grlR grlA* mutant strains. So, GrlA negatively regulates EHEC motility in the absence of GrlR. Constitutively expressed

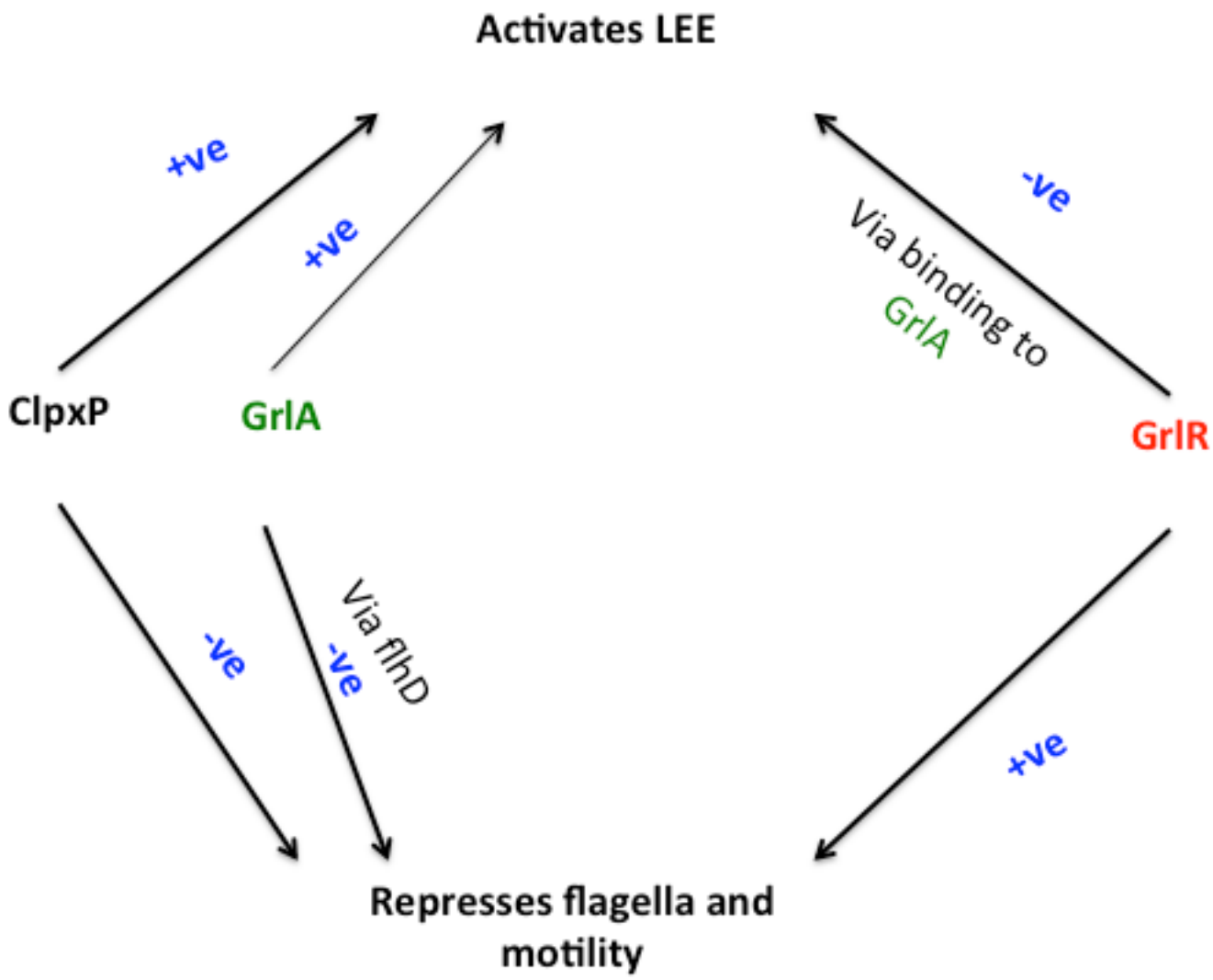


Fig. 1.14: Schematic representation of regulation at *LEE*. Shown is complex interplay between ClpxP protease, GrlA, and GrlR regulation of both LEE and flagella and motility.

FlhD/FlhC inhibits efficient adherence of EHEC to HeLa cells. However, overproduced flagellar elements are not the sole reasons for the deleterious adhesion of EHEC to host cells (Iyoda *et.al,* 2006).

1.6.2 Regulation of haemolytic genes

Enterohaemolysin (Ehx) is encoded by the *ehxA* gene located on pO157, 92-kb virulent plasmid, belongs to the *ehxC* operon. The haemolytic phenotype of EHEC is increased in a *grlR* mutant. GrlA was found to enhance the haemolytic phenotype of the *grlR* mutant strains. A deletion of *ehxA* and *ehxCABD* completely removes the haemolytic phenotype of EHEC even if *grlR* is deleted. The haemolytic activity however is restored with complementation with *ehxCABD,* which indicates that *ehx* is indispensable for the GrlA-dependent haemolytic phenotype. It was found that the expression of Type III secretion system in *grlR* mutant strains caused no change in the haemolytic phenotype of EHEC. Therefore, T3SS is not involved in the haemolytic phenotype of EHEC. In addition, the haemolytic activity facilitated by GrlA is not dependent on *ler* since *ler* mutant strains were able to show a hyper-haemolytic phenotype as well as a *ler grlR* double mutant. Clearly, GrlA up-regulates the expression of both *LEE* and *Ehx,* but it down-regulates flagellar gene expression. GrlA up-regulates *ehxC,* which leads to hyper-haemolytic activity in EHEC, and it is a universal *exh* up-regulator in EHEC O157 strains (Saitoh *et al*., 2008).

1.7 Aims of the project

The aim of this project is to understand what triggers the activation of the expression of *LEE* genes. Since Ler is considered to be the principal regulator of the entire system, this study focused on the regulation of the *Ler* promoter by GrlA. This was achieved my introducing eukaryotes system by using CaCo-2 cells, to mimic the intestinal host environment. Additionally, GFP reporter system and single substrate were used to gain a wider insight into the regulation of expression. As Ler contains a long leader sequence, which could paly a role in translation regulations, its impact on expression was investigated as well. Finally, hybrid promoters were constructed to create synthetic promoter that can be used to study how different transcription factors regulate at a given promoter.

Chapter Two:

Materials and Methods

2.1 Suppliers

All buffers, solutions and reagents used in this study were purchased from Sigma-Aldrich, Bioline, Bio-Rad, National Diagnostics, or New England Biolab (NEB), except where otherwise mentioned. Oligodeoxyribonucleotides were purchased from Alta Bioscience in the University of Birmingham. Restrictions enzymes, CIP, and T4 DNA ligase were purchased from NEB. Biotaq DNA polymerase, dNTP mix, and Red Mix were purchased from Bioline.

2.2 Buffers, solutions, and reagents

Solutions, which were homemade, were all autoclaved prior to use in experiments with bacterial cultures for 15 minutes at 120ºC and the pressure was 15 psi. Other solutions, which were not suitable for autoclaving, were filter sterilized using 0.2 μm filters.

2.2.1 DNA and proteins electrophoresis gel components

6 x gel loading dye: were either purchased from NEB or homemade as follows: 1 x buffer composition; 2.5% Ficoll®-400; 11mM EDTA; 3.3 mM Tris-HCl; 0.017% SDS; 0.015% bromophenol blue; pH 8.0 at 25°C. Loading dye was kept at room temperature.

Ethidium Bromide: was used to stain the DNA gel after electrophoresis. A 10 mg/ml stock was stored in a lightproof container at room temperature. A 0.5 µg/ml working solution was made for staining.

DNA markers: were either 1 kb or 100 bp markers. A 1 kb marker, on the one hand, contains ten fragments ranging from 0.5 kb to 10 kb and the intensity of each band was provided. This marker was used to estimate large fragment sizes. A 100 bp marker, on the other hand, contains twelve bands ranging from 100 bp to 1517 bp, and it was used to estimate the sizes of PCR products. Both markers come with 6 x loading gel and were prepared as follows, prior to loading: 4 µl distilled water, 1 µl 6 x gel loading dye, and 1 µl DNA marker to make the total volume 6 µl. The contents were mixed gently and stored at -20ºC.

Gel electrophoresis buffers: were purchased as either 50 x TAE or 5 x TBE. The composition of 50 x TAE is as follows: 2 M Tris-acetate and 100 mM Na_2EDTA in distilled/deionized water (pH 8.3 at 1 x concentration). TAE Buffer is mainly used for agarose DNA electrophoresis. 5 x TBE were used for polyacrylamide gel electrophoresis and when diluted to 1 x it contains: 0.089 M Tris base, 0.089 M boric acid (pH 8.3) and 2 mM Na_2EDTA.

Agarose gel solutions: were made in 1 x TAE buffer to either 0.8% or 1.3% depending on the fragment or the plasmid size to be run. For 0.8% agarose 0.8g, as an example, was dissolved in 100 ml 1 x TAE buffer and was microwaved for two minutes or until the agarose

was completely dissolved in the buffer and it was left to cool before pouring it into the gel-casting units.

Polyacrylamide gel electrophoresis components: the composition of stock polyacrylamide solutions was as follows: 30% (w/v) acrylamide, 0.8% bisacrylamide stock solution. Stock acrylamide solution was stored in a dark space at room temperature and was suitable for use for 24 months. N, N, N', N' - tetramethylethylene diamine (TEMED) was also kept in the dark at room temperature. 10% ammonium persulphate (APS) was homemade by dissolving 0.1 grams of ammonium persulphate into 1 ml distilled water and was kept at room temperature. Fresh APS was made to ensure the gel polymerized better. Acrylamide gel used in this study was prepared as follows: 7.5% Acrylamide working solution, 125 ml 30% (w/v) acrylamide, 0.8% bisacrylamide stock solution, 100 ml 5 x TBE, 20 ml glycerol, made up to 500 ml with distilled water. To run a small Acrylamide gel 10 ml of 7.5% Acrylamide working solution was added to a small beaker; 100 µl or 10% ammonium persulphate was added, along with 15 µl TEMED. The solution was mixed thoroughly and then poured between two Acrylamide gel-casting glasses.

Protein gel buffers

Resolving gel buffer components: 30% acrylamide/bisacrylamide, 1.5 M Tris-HCl

Containing 0.4% SDS, pH 8.8, 10% APS, TEMED.

Stacking gel buffer components: 30% acrylamide/bisacrylamide, 0.5 M Tris-HCl containing 4.0% SDS, pH 6.8, 10% APS, TEMED.

Sample blue 5 x: 10% W/V SDS, 10 mM Dithiothreitol, or β-mercaptoethanol, 20% V/V Glycerol, 0.2 M Tris-HCL, pH 6.8, 0.05% Bromophenolblue.

Running buffer: 25 mM Tris-HCl, 200 mM Glycine, 0.1% W/V SDS.

Protein ladder: 10-250 kD ladder containing twelve bands to estimate the size of a protein band in the gel. The ladder is kept at -20^{o}C.

Protein staining solutions: 50% Methanol, 10% acetic acid, 0.25% Coomassie Blue

Protein de-staining solutions: 5% Methanol, 7.5% acetic acid, 87.5% distilled water

Shrink solution: 48% (v/v) methanol, 2% (v/v) glycerol

2.2.2 Extraction and purification of DNA fragments

Phenol/Chloroform/Isoamyl alcohol: phenol/chloroform/isoamyl alcohol (25/24/1 v/v), pH 8.0 from Fisher scientific

TEN buffer: 60 mM Tris (pH 8.0), 5 mM EDTA, 50 mM NaCl

TE buffer: 10 mM Tris (pH 8.0), 0.05 mM EDTA

2.2.3 *E. coli* transformation with DNA buffers

Calcium chloride method:

Calcium chloride: 0.1 M $CaCl_2$ (autoclaved) 0.1 M $CaCl_2$ + 10% glycerol (autoclaved)

Freeze-thaw buffer: 100 mM $CaCl_2$, 10 mM Tris-HCl, pH 7.5, 15% glycerol (autoclaved)

Rubidium chloride method:

TFB1 buffer: 30 mM potassium acetate, 10 mM CaCl2, 50 mM MnCl2, 100 mM RbCl, 15% glycerol, pH 5.8 with 1 M acetic acid, filter-sterilized (0.2 µM) and stored at 4ºC.

TFB2: 10 mM PIPES, pH 6.5, 75 mM CaCl2, 10 mM RbCl, 15% glycerol, pH 6.5 with 1M KOH, filter sterilized with (0.2 µM) and stored at 4ºC.

10% glycerol (Autoclaved)

2.2.4 β-galactosidase assay solutions

Z-buffer: 0.75 g KCl, 0.25 g $MgSO_4.7H_2O$, 8.53 g Na_2HPO_4, 4.87 g $NaH_2PO_4.2H_2O$ and 2.70 ml β-mercaptoethanol, made up to 1 L with distilled water and autoclaved.

β-mercaptoethanol was kept in the dark at room temperature and it was only added at the time of the *β*-galactosidase assay and adjusted to pH 7.0.

Ortho-nitrophenol-β-D-galactopyranoside (ONPG): was kept at -20°C in a freezer. 80 mg ONPG was added to 100 ml Z-buffer just prior to the *β*-galactosidase assay. Fresh ONPG was prepared each time.

Sodium deoxycholate ($C_{24}H_{39}O_4Na$): 1 M Sodium dexoycholate was prepared in distilled water and stored at room temperature.

Toluene: was used with Sodium deoxycholate in the *β*-galactosidase assay to lyse bacterial cells.

Sodium Carbonate (Na_2CO_3): 1 M solution was prepared in distilled water and kept at room temperature. Sodium carbonate was used in the *β*-galactosidase assay to stop the reaction.

2.2.5 Attachment experiment solutions

Cell-splitting solutions: Phosphate buffered saline (PBS) purchased from Sigma-Aldrich or homemade as follows: 10 x PBS 1.37 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, and 18 mM KH_2PO_4. The pH was adjusted to 7.4. The solution was then autoclaved. Dulbecco's Modified

Eagle's Medium (DMEM), feto bovine serum, Penicillin, and Streptomycin were added to prevent bacterial and fungal contamination. Trypsin was used to digest cells lines.

Cell attachment solutions: PBS, DMEM without antibiotics, and PBS + 0.1% Triton™ X-100 from Sigma-Aldrich.

Cell staining solutions: Ethanol, to sterilise cover slips, Trypsin, DMEM, to dilute cell lines, 1 x PBS, 10 x PBS: 800 ml dH_2O, 80 g NaCl, 2 g KCl, 7.62 g Na_2HPO_4, 0.77 g KH_2PO_4. Solution was at pH 7.4 and dH2O was added to make 1 L. 3.2% formaldehyde in PBS was prepared as follows: for 50 ml preparation, 5 ml formaldehyde solution (37%) added to 5 ml of 10 x PBS then dH2O to make up 50 ml solution. The prepared solution was then kept in the fridge. Dye solutions: 1.4 µl Hoechst to stain the nucleic acids, and 14 µl phalloidin to label fluorescently the actin filaments in 1400 µl 1 x PBS for a 6-well plate. Mounting solution: 5 ml 2 x PBS, 5 ml glycerol, 0.1 g propyl gallate, aliquoted to 0.5 ml and stored at -20°C.

Single molecules used to trigger GrlA: Poly-L-lysineP4707: no dilution required. Collagen C2124: no dilution required. Fibronectin stock: 0.5 mg/ml diluted 1:5 in 1 x PBS.

2.3 Culturing media for bacteria

2.3.1 Liquid media

E. coli strains were cultured in 2 x Luria-Bertani (LB) medium or DMEM medium in case of the selection of a plasmid-harbouring strain. Antibiotics were added to the growth media in concentrations recorded in 2.3.3.

2 x LB media was composed of 20 grams Tryptone, 10 grams Yeast extract, and 10 grams NaCl solution in 1 L distilled water. Contents were heated and stirred until all ingredients were dissolved, then it was autoclaved.

DMEM (Dulbecco's modified Eagle's medium): pH 7.4, adjusted with HEPES buffer was used ready-made from Sigma-Aldrich.

SOC medium: a rich medium used to maximise the efficiency of transformation, used ready-made from Sigma-Aldrich

2.3.2 Agar media

Various agar media were used for growing *E. coli* strains. Most strains were grown on nutrient agar or LB agar. When lactose fermentation was monitored, strains were grown on MacConkey lactose agar mediums, supplied by Difco Laboratories. Media were made in distilled water according the protocols from the manufacturers, then autoclaved. When needed, antibiotics were added to the media after it had cooled, following autoclaving.

2.3.3 Antibiotics

Ampicillin stock was made to 50 mg/ml by dissolving ampicillin in distilled water then using 2 μm sterilizing filters. It was stored at -20°C in a freezer and the working solution was 200 μg/μl. Tetracycline stock was made to 10 mg/ml by dissolving 1 gram of tetracycline in 100 ml of 70%-solution methanol. It was then stored at -20°C in a freezer. Since tetracycline is sensitive to light, the antibiotic container was wrapped in aluminium foil. The working solution of tetracycline was 35 μg/μl. Chloramphenicol stock was made to 35 mg/ml by dissolving chloramphenicol into pure ethanol and was stored at -20°C in a freezer. The working solution was 35 μg/μl.

2.4 Bacterial strains and plasmids

2.4.1 Bacterial strains and growth conditions

E. coli strains used in this study were mentioned in Table 2.1. Strains were routinely streaked out on MacConkey maltose agar plates each month. At the same time, glycerol stocks of

Table 2.1: *E-coli* strains used in this study.

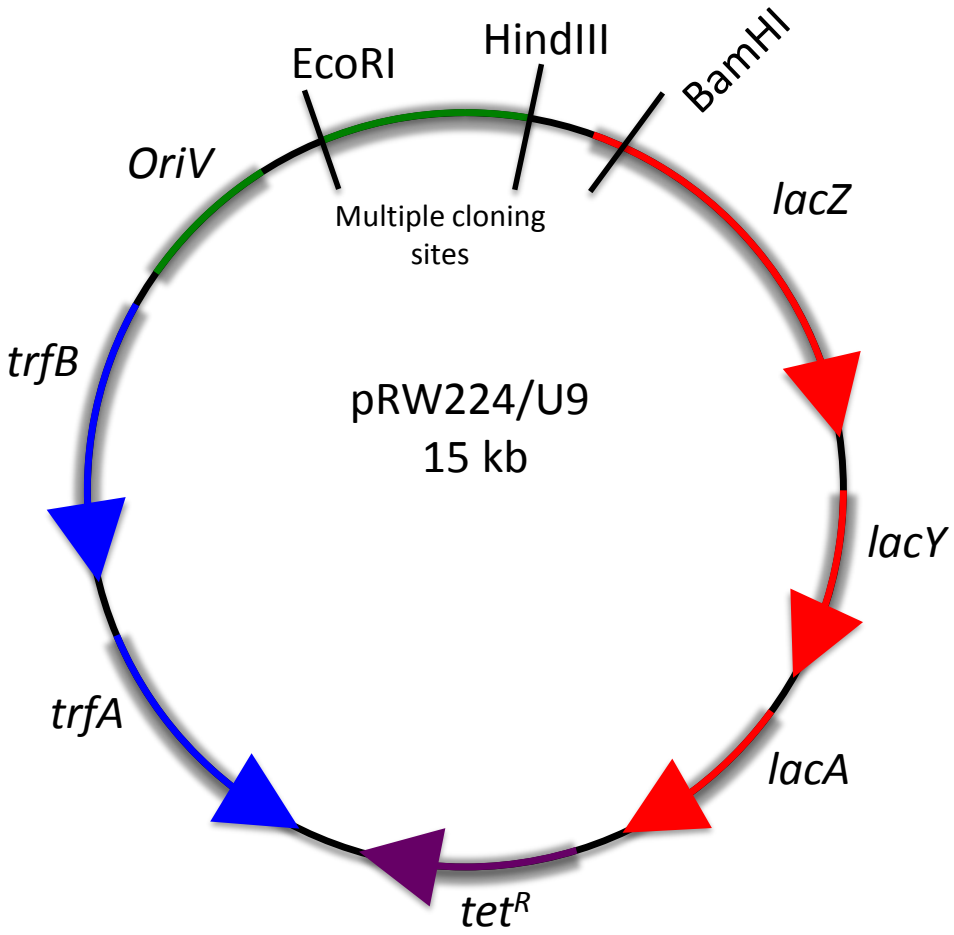
Strain	Genotype/description	Reference
M182	*E. coli* K-12 strain, Δ (*lac*IPOZY) X74, *gal*K, *gal*U, *str*A	Casadaban and Cohen (1980)
M182 *Δcrp*	A derivative of M182 (Δ (*lac*IPOZY) X74, *gal*K, *gal*U, *str*A) where *crp* gene was deleted	Busby *et al.*, (1983)
Sakai 813 *ΔStx*	EHEC O157:H7 Sakai derivative strain, *ΔStx1* and *ΔStx2*	Islam *et al.*, (2011)
Sakai *grlR*	EHEC O157:H7 Sakai strain, a derivative of Sakai 813 *ΔStx* where the alanine codon 37 in *grlR* changed to TAG amber stop codon	Islam *et al.*, (2011)
Sakai *grlA*	EHEC O157:H7 Sakai strain, a derivative of Sakai 813 *ΔStx* where the alanine codon 55 in *grlA* changed to TAG amber stop codon	Islam *et al.*, (2011)

Table 2.2: Plasmids used in this study.

Plasmid	Description	Reference
pRW224	A derivative of pRW50 that allow cloning of promoter fragments as transcription or translation fusions to *lacZ*	Islam *et al.*, (2011)
pRW224/U9	A derivative of pRW50 that allow cloning of promoter fragments as transcription or translation fusions to *lacZ* carrying pUC9 linker between the EcoRI and HindIII sites and encodes for tetracycline resistance (Fig. 2.1)	Islam *et al.*, (2011)
LEE10-568/pRW224	A derivative of pRW224, carrying an EcoRI-HindIII LEE1 promoter from position -568 to position -19 relative to ler translation site as a transcription fusion to *lacZ*	Islam *et al.*, (2011)
LEE20-203/pPW224	A derivative of LEE10-568/pRW224, carrying an EcoRI-HindIII fragment from position -240 to position 123 relative to *ler* translation site as a transcription fusion to *lacZ*	Islam *et al.*, (2011)
GS100/pRW224	A derivative of LEE20-203/pRW224 that contains a CRP binding site at position -61.5 relative to *ler* transcription start site as a transcription fusion to *lacZ*	This study
GS101/pRW224	A derivative of GS100/pRW224 that contains mutations in the CRP binding site	This study
GS102/pRW224	A derivative of GS100/pRW224 that contains a CRP binding site at position -55.5 relative to transcription start site of *ler*	This study
GS103/pRW224	A derivative of GS100/pRW224 that contains a CRP binding site at position -51.5 relative to transcription start site of *ler* promoter	This study
GS103 (9C)/pRW224	A derivative of GS103/pRW224 that contains a point mutation at position -9 relative to transcription start site of *ler*	This study
pRW225	A derivative of pRW224 that allows cloning of promoter fragments in EcoRI-HindIII or EcoRI-BamHI as translational fusion to *lacZ* encodes for tetracycline resistance (Fig. 2.4)	Islam *et al.*, (2011)

LEE30-275/pRW225	A derivative of pRW225 carrying LEE30-275 promoter fragment between EcoRI-HindIII sites as a translation fusion to *lacZ*	Islam *et al.*, (2012)
LEE30-275 (46T)/pRW225	A derivative of LEE30-275/pRW225 a carrying point mutation in position +46 relative to *ler* transcription start site	Islam *et al.*, (2012)
LEE30-275 (28A/46T) pRW225/	A derivative of LEE30-275/pRW225 carrying point mutations in positions +28 and +46 respectively relative to *ler* transcription start site	This study
LEE150/pRW225	A derivative of pRW225 carrying a LEE150 promoter fragment between EcoRI-HindIII sites as a translation fusion to *ler*.	Islam *et al.*, (2012)
LEE150-1/pRW225	A derivative of LEE150/pRW225 with a base pair substitution in the Shine-Dalgarno sequence (GGTGG) ➔(GGAGG) at +26 relative to *ler* transcription start site, as a translation fusion to *ler*	This study
LEE150-2/pRW225	A derivative of LEE150/pRW225 carrying a point mutation in the Shine-Dalgarno sequence fused to *ler* (GGAG)➔(TGAG) at +156 relative to *ler* transcription start site, between EcoRI-HindIII sites as a translation fusion to *ler*	This study
LEE150-3/pRW225	A derivative of LEE150/pRW225 that contains point mutation in both Shine-Dalgarno sequences (GGTGG)➔(GGAGG) and (GGAG)➔(TGAG)	This study
LEE150-4/pRW225	A derivative of LEE150/pRW225 carrying a point mutation in the Shine-Dalgarno sequence fused to *ler* (GGAG)➔(AGAG) between EcoRI-HindIII sites as a translation fusion to *ler*	This study
LEE150-5/pRW225	A derivative of LEE150/pRW225 that contains point mutation in both Shine- Dalgarno sequences (GGTGG)➔(GGAGG) and (GGAG)➔(AGAG)	This study
pRW400/U9	A derivative of pRW224/U9 where *lacZ*, *lacY*, and *lacA* genes were replaced with *gfp* derived from vector pJB (Fig. 2.6) in frame downstream from a promoter-cloning site (Fig. 2.5)	This study
LEE10-568/pRW400	A derivative of pRW400/U9 that carries LEE10-568 promoter between EcoRI- HindIII	This study

	sites as a transcription fusion to *gfp*	
LEE20-203/pRW400	A derivative of LEE10-568/pRW400 that carries LEE20-203 between EcoRI-HindIII sites as a transcription fusion to *gfp*	This study
LEE20-203 99T/pRW400	A derivative of LEE10-568/pRW400 that carries LEE20-20399T between EcoRI-HindIII sites as a transcription fusion to *gfp*.	This study
pACYC184	Cloning vector that is made by ligating restriction fragments from pSC101, Tn9 and p15A. Used for cloning HindIII-SalI DNA fragments encodes tetracycline and chloramphenicol resistance (Fig. 2.3)	Chan and Cohen (1987)
pACYC184/ΔHN	Derivative of pACYC184 in which HindIII-NruI fragment has been deleted. Encodes for chloramphenicol resistance	Mitchell *et al.*, (2007)
pACYC184/GrlR+A	A derivative of pACYC184 carrying *grlR+A+* gene cloned between HindIII and SalI sites	Islam *et al.*, (2011)
pACYC184/GrlA	A derivative of pACYC184/GrlR+A+ carrying *grlR-A+* gene cloned between HindIII and *Sal*I sites	Islam *et al.*, (2011)
pACYC184/GrlR	A derivative of pACYC184/GrlR+A+ carrying *grlR+A-* gene cloned between HindIII and SalI sites	Islam *et al.*, (2011)
pSR	High copy number pBR322derivative, used for cloning EcoRI-HindIII promoter fragments upstream λ terminator coding for ampicillin resistance (Fig. 2.2)	Kolb *et al.*, (1995)
LEE150-1/pSR	A derivative of pSR that carries LEE150-1 fragment between EcoRI-HindIII sites as a transcription fusion	This study
LEE150 *Δl*/pSR	A derivative of pSR that carried *LEE*150 Δl fragment between EcoRI-HindIII sites as a transcription fusion	This study



attcgcgagagccttgagtccacgctagatctgaattcCCGGGGATCCGTCGACCTGCAGCC

aagcttaatggagcgaattatgagagttctggttaccgccaagctccgggatcccgtcgttt

tacaacgtcgtgactgggaaaaccct

Fig. 2.1: Plasmid map of the lac fusion vector pRW224/U9.
This figure shows the plasmid map of pRW224/U9, which is a derivative of pRW50 (Lodge *et al.*, 1992). This plasmid was modified by removing the *trpAB* and the cloning sites were reconstructed to allow cloning of promoter fragments fused to *lacZYA* either transcriptionally or transnationally. It encodes for tetracycline resistance. The cloning site sequence is displayed below the figure. A promoter fragment was cloned between EcoRI-HindIII sites and the translation of the *lacZ* was regulated by the Shine-Dalgarno sequence. To clone a promoter as a translation fusion to *lacZ*, *EcoRI-BamHI* sites were employed for the fragment. The translation of the *lacZ* here was controlled by the Shine-Dalgarno sequence presented in the cloned fragment. The cloning site's base sequences showing the pUC9 linkers in capital letters, restriction sites in grey shades (*EcoRI*, *HindIII*, and *BamHI* respectively), Shine-Dalgarno sequence, the start codon sites shaded brown, and the *lacZ* start site in green are also shown.

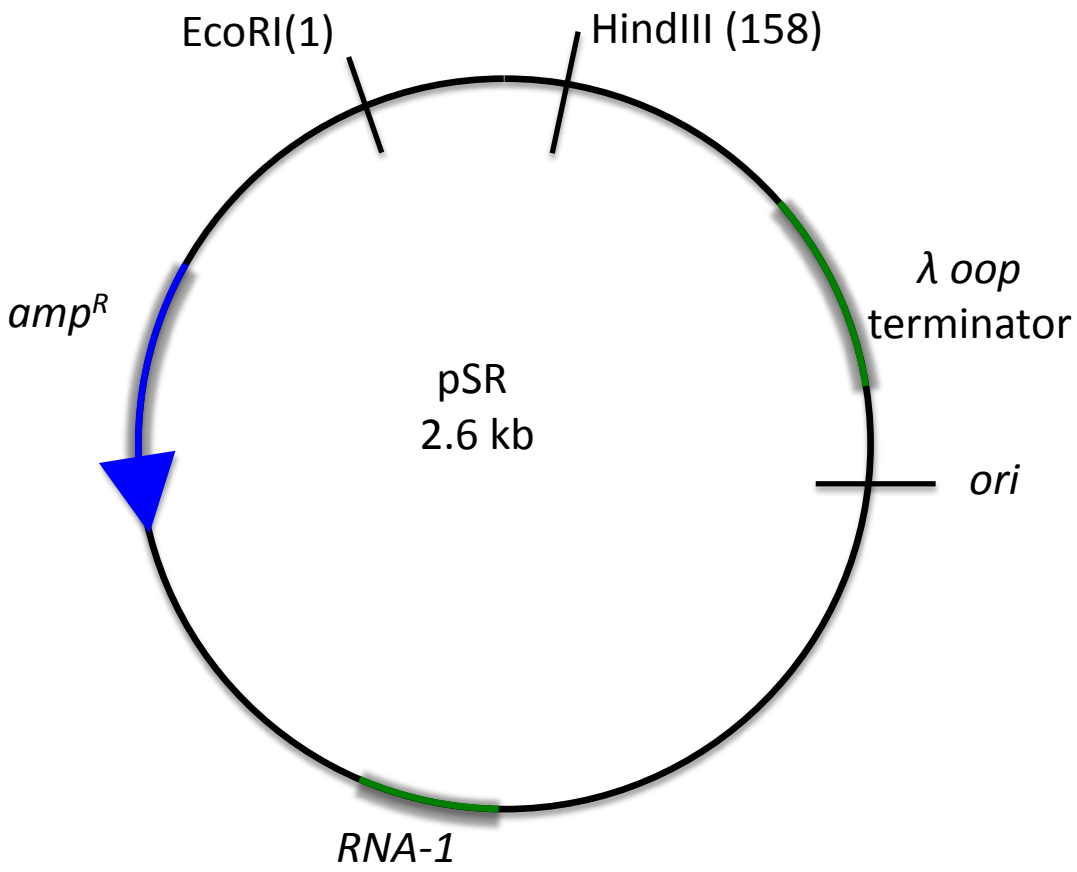


Fig. 2.2: Plasmid map of pSR cloning vector.

The figure represents a plasmid map of a pSR cloning vector that encodes ampicillin resistance. Promoter fragments can be cloned between EcoRI-HindIII sites upstream from the *λ oop* terminator. The origin of replication and the *RNA-I* are also shown.

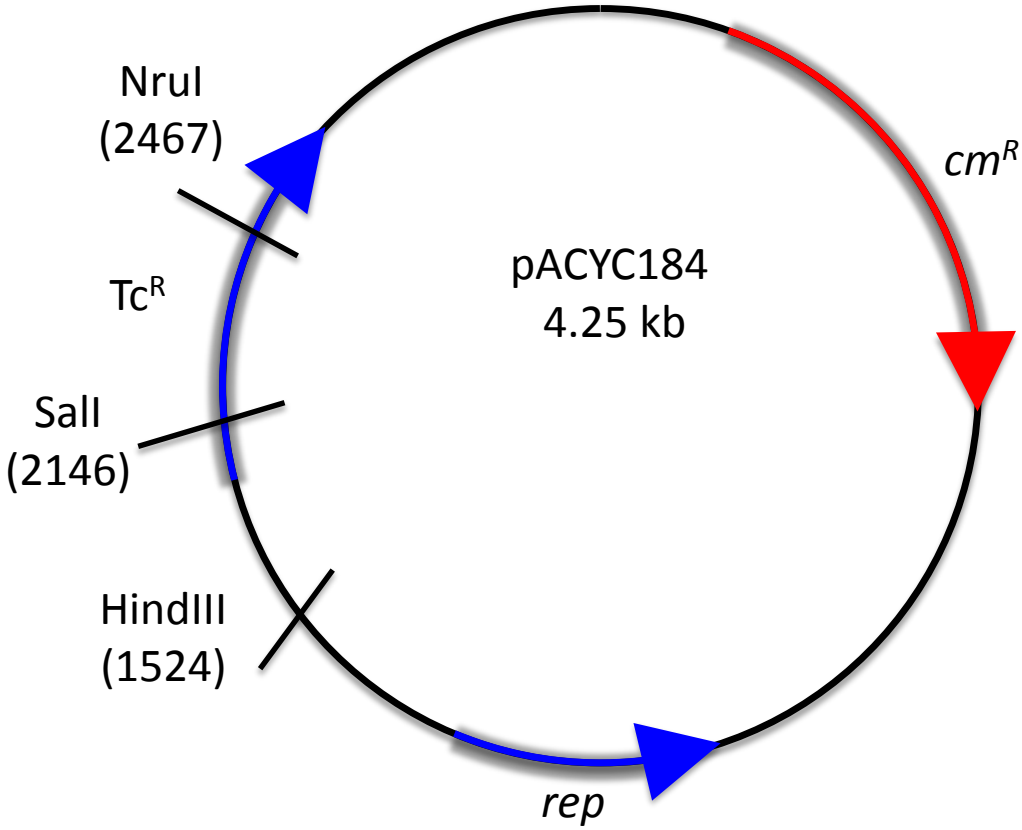
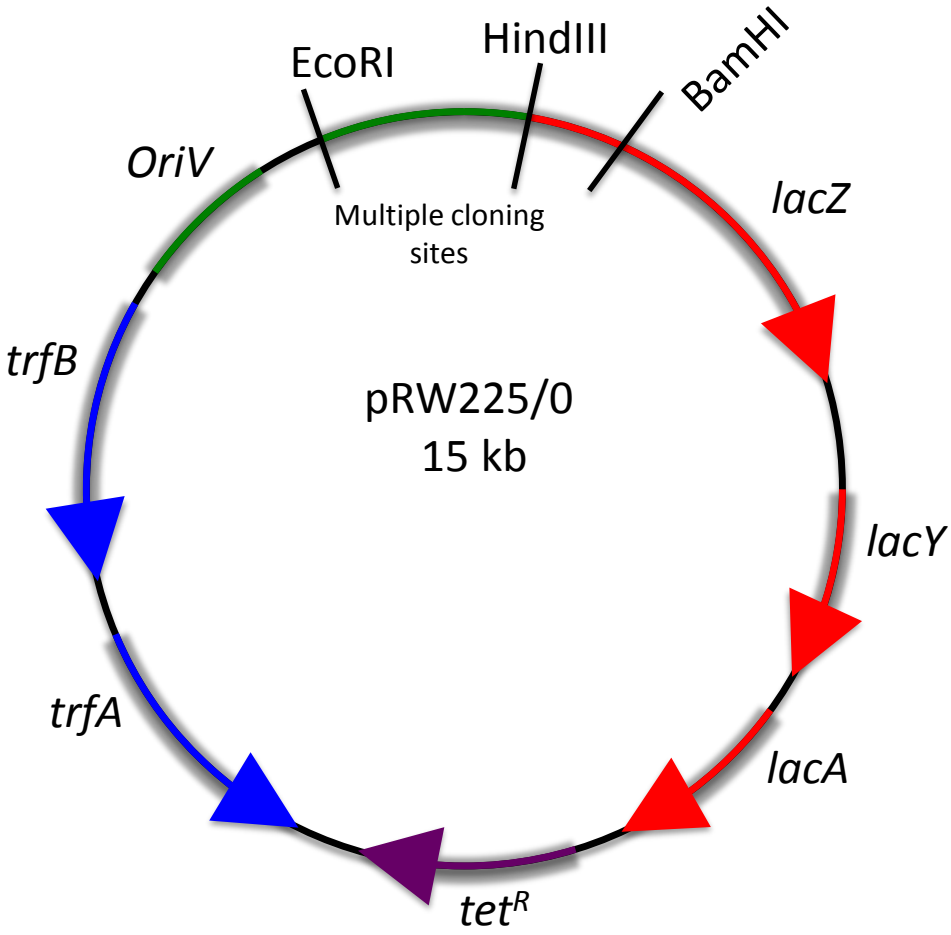


Fig. 2.3: Plasmid map of pACYC184 cloning vector.

This figure represents a plasmid map of the pACYC1842 cloning vector that encodes chloramphenicol resistance. Tetracycline resistance is lost in this vector as it was cloned by restricting a fragment in the NurI-SalI site where tetracycline resistance was encoded. This vector was used to clone a number of fragments controlled by their own promoters.



attcgcgagagccttgagtccacgctagatctgaattcCCGGGGATCCGTCGACCTGCAGCCaagcttccgggat

cccgtcgttttacaacgtcgtgactgggaaaaccct

Fig. 2.4: Plasmid map of pRW225/0 cloning vector.
This figure represents a plasmid map of pRW225/0 (Islam *et al.*, 2011), which is a derivative of pRW224. The plasmid was constructed by cloning a fragment without a promoter, EcoRI-BamHI; upstream from the *rsd* promoter that carries the HindIII site between the EcoRI-BamHI sites in pRW224. This vector can be used to clone promoter fragments as translation fusion to *lacZ* utilizing EcoRI-HindIII cloning sites or EcoRI-BamHI cloning sites, shown below the figure.

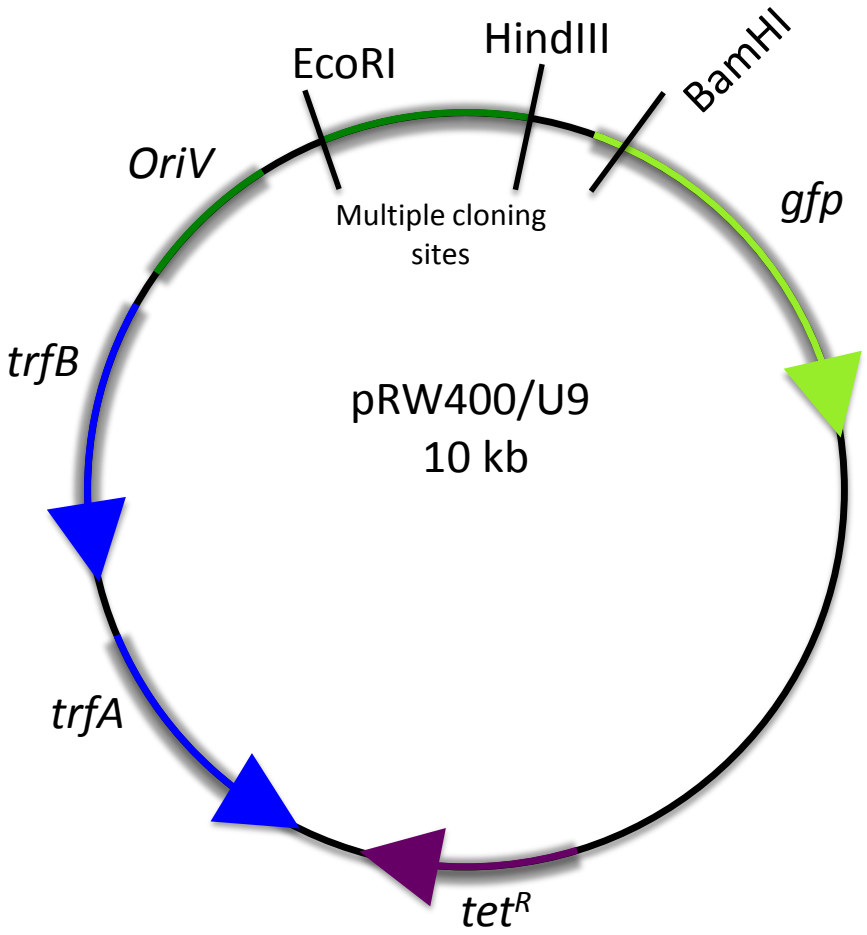


Fig. 2.5: Plasmid map of pRW400 cloning vector.
This figure represents a plasmid map of the pRW400 cloning vector that is a derivative of pRW224 where the *gfp* from pJB (Bryant, 2013) was excised by digesting the vector with BamHI-NheI enzymes. The resultant *gfp* was then cloned into pRW224 after digesting it with the same enzymes to produce pRW400. This vector was used to clone a promoter fragment as transcription fusion to *gfp* either in EcoRI-HindIII sites or EcoRI-BamHI sites.

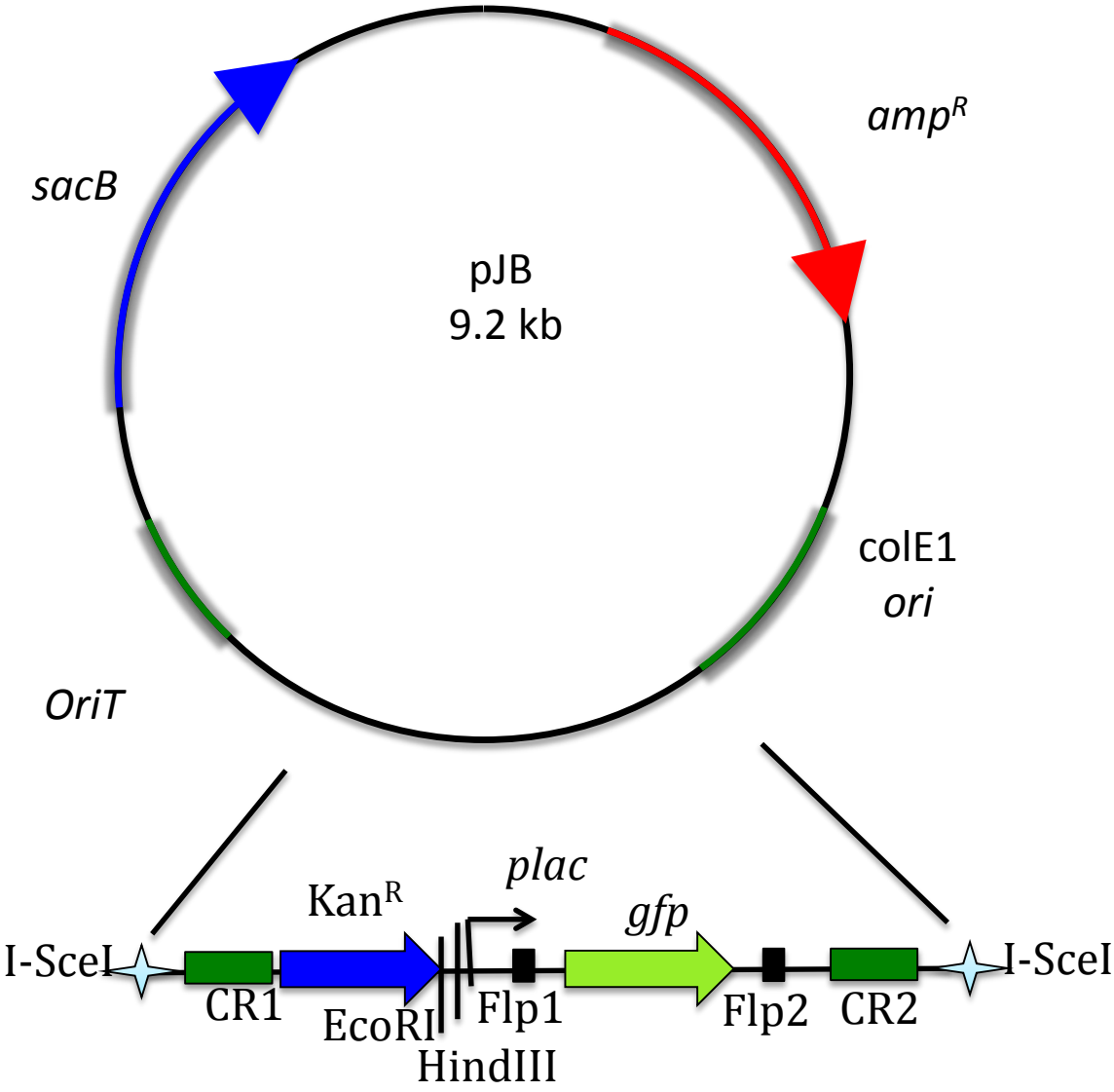


Fig. 2.6: Plasmid map of pJB cloning vector.
This figure represents a plasmid map of the pJB cloning vector that is a derivative of pDOC-C (Lee *et al.*, 2009). The plasmid was used to restrict the *gfp* fragment and clone it into pRW224 after removing the *lac* genes. pJB was supplied by Bryant *et al.*, 2014.

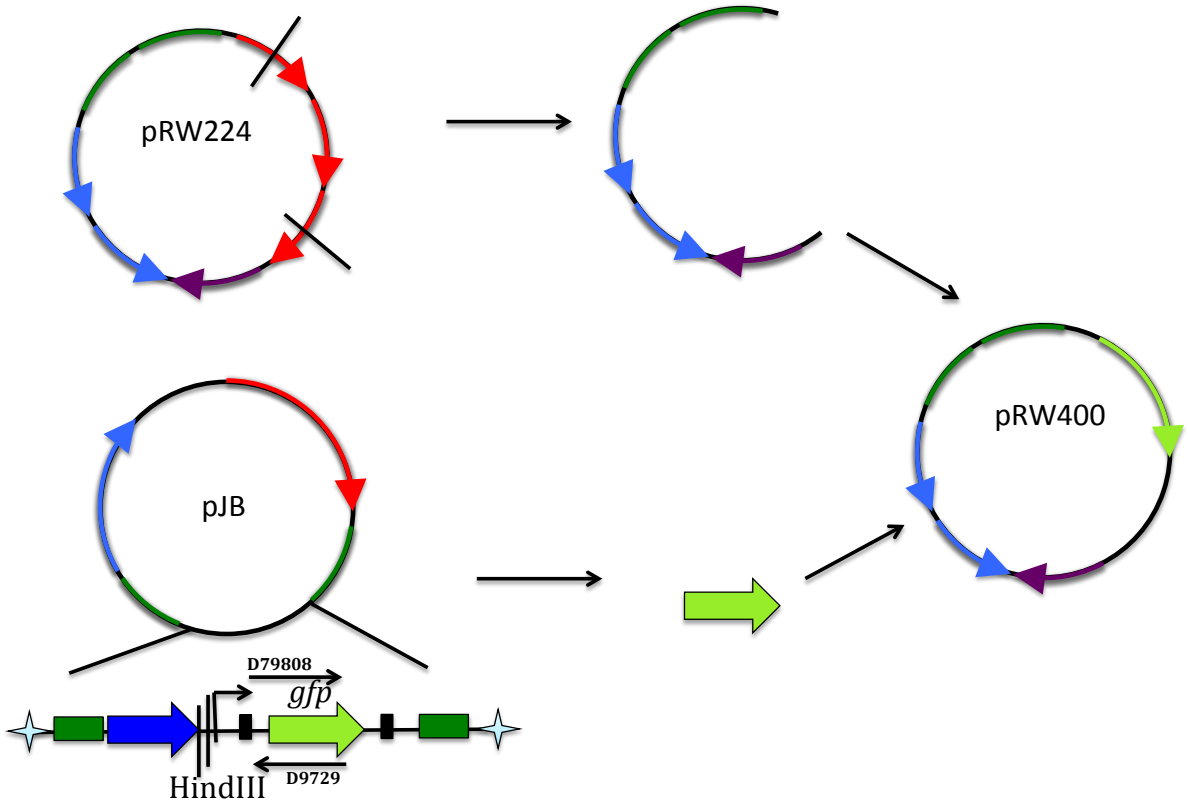


Fig. 2.7: Schematic representation of pRW400 cloning.
This figure represents cloning steps followed to construct a pRW400 cloning vector. The vector was cloned by digesting the *lacZYA* from pRW2240 using HindIII and NheI enzymes and amplifying *gfp* from pJB cloning vector. The resulting fragments were ligated and produced pRW400.

strains were made by inoculating a single colony of a strain to be stored in 5 ml LB media, contained in a 25 ml flask, supplemented with antibiotics if required. The flask was then left on a 37°C shaker overnight (16-18 hours). After that 500 µl from the culture was placed into 1.5 ml microfuge tube and 500 µl of sterile 70% glycerol was added. Finally the tube was kept at -80°C. When subculturing was needed, 100 µl from the overnight culture was diluted into fresh 5 ml LB, with antibiotics if needed. Cells were left to grow until they reached the exponential phase of growth by assessing the optical density of the cultures using a Helios Gamma Spectrophotometer supplied by Thermo Fisher Scientific Inc. The wavelength used to measure the optical density was 650 nm.

2.4.2 Plasmids

Plasmids used in this study are shown in Table 2.2 and their maps are shown in figures 2.1, 2.2, 2.3, etc. In short, pRW224 plasmid was used to clone the LEE1 promoter fragments as a transcriptional fusion to *lacZ* and the promoter strength was assessed by measuring the *β*-galactosidase enzyme that is produced when *lac* genes are expressing under the control of LEE1 promoters.

LEE30-275 and derivatives and LEE150 and derivatives were cloned into pRW225 plasmid as translation fusions to *lacZ* and the strength of *lacZ* expression was measured by performing *β*-galactosidase assays.

GrlAR gene versions were cloned into a pACYC184 cloning vector that is used to clone a gene of interest, in which its own promoter regulates its transcription.

Plasmid pJB was used as a template for *gfp*, which was later, cloned into pRW224 to construct pRW400.

2.5 Gel electrophoresis

2.5.1 Agarose gel electrophoresis

Agarose gels of 0.8% or 1.3%, depending on the fragment size, were employed to visualize a resolved DNA band or for extraction and purification of desired vectors or inserts. Agarose gels were usually used to resolve DNA fragments ranging from 200 bp to 10 kb in size. Firstly, the agarose of a desired percentage was dissolved into 1 x TAE buffer and microwaved, so that the agarose was properly dissolved then it was left aside to cool down to approximately 50°C. It was then poured into a gel-casting unit and a comb was inserted to create wells that can hold samples ranging from 10 µl to 100 µl. Once the gel was settled, which took about 15 minutes depending mainly on room temperature, the comb was removed and the gel was placed in the running tank horizontally and 1 x TAE running buffer was added to the tank. Then, the DNA samples to be analysed were prepared by mixing them with DNA loading dye where five portions of DNA were added to one portion of 6 x DNA loading dye, so that the gel loading dye was diluted to 1 x. A suitable DNA marker was run for fragment size comparison as well, and for molecular weight estimation. The samples were then subjected to electrophoresis at 10 V/cm for 30-45 minutes. The gel was then placed in Ethidium bromide solution at 0.5 µg/ml for 20 minutes. The gel was visualized using a Molecular Imager (Gel DOC XR imaging system Bio-Rad Laboratories). The gels were

viewed under UV light at 360 nm because using this wavelength limits the amount of damage to the DNA. As little as 10 ng DNA can be visualized.

2.5.2 Polyacrylamide gel electrophoresis (PAGE) of DNA

Polyacrylamide gels were used to separate DNA fragments of small sizes that cannot be resolved by agarose gels. It can resolve fragments sizes ranging from 50 to 1000 bp. 10 ml 7.5% polyacrylamide was polymerized by adding 100 µl fresh APS and 15 µl TEMED. This was mixed thoroughly and poured between two glass plates that were separated by 0.15 cm thick spacers, and then a comb was inserted to create samples wells that can hold 10-15 µl samples. The gel was left to polymerize in an upright position then the comb and the spacers were removed and the gel was placed vertically in an electrophoresis apparatus that contained 1 x TBE running buffer. A syringe was used to remove bubbles at the bottom of the gel and to clear un-polymerized gel from the wells if any remained, otherwise the samples would not settle in the well. The samples were prepared by mixing them with DNA loading dye and run on the gel with a suitable DNA marker at 30 mA for one to four hours. The gel was then removed, stained, and viewed under UV light as mentioned in 2.5.1. However, the staining time was reduced to 15 minutes.

2.5.3 Examination of DNA concentration by gel electrophoresis

The concentration of a DNA fragment was estimated by comparing the intensity of the band with the bands of DNA markers of known concentration. For precise determination of a band

size, NanoDrop 1000 Spectrophotometry (Thermo Scientific) was used. 1 µl of a sample was placed on a fibre optic cable (the receiving fibre). A second fibre optic cable (source fibre) was brought into contact with the samples. Xenon light passes through the sample that is detected by the instrument and the molecular weight in nanograms is displayed on the computer screen. Precise DNA size-determination was used to calculate the amount of DNA needed for cloning.

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

Protein gels were used to separate protein bands according to their molecular weight. SDS is added to coat the proteins in negative charge, so separation is then based on molecular weight. In this study, SDS-PAGE was used to monitor *ler* expression after modifying the Shine-Dalgarno sequence. The gel casting units were prepared and tested for water leakage (BioRad) then the gel was prepared as mentioned in 2.2.6, poured into the gel casting unit and left to polymerize for 40 to 60 minutes. Butanol was used as an overlay buffer on top of the resolving gel once the polymerized butanol had been discarded. The stacking gel was prepared as mentioned in 2.2.6 and poured over the resolving gel, a comb was inserted and the gel was left to polymerize for 30 minutes. Once the gel was polymerized, the comb was removed and the gel was placed in an SDS-PAGE running tank filled with running buffer 2.1.6. The protein samples to be run were prepared by mixing them with sample blue and heating at 95°C. They were then vortexed for 10 seconds and centrifuged at 15 seconds at ~18000 x g. The gel was run for 40 minutes to one hour at 140 volts. Then it was stained for 30 minutes in Coomassie blue stain 2.6.1. It was de-stained overnight in de-staining buffer 2.6.1 and the gel was then scanned using a GS-900™ Calibrated Densitometer. It was then

dried between two clade sheets and stored.

2.6 Extraction and purification of nucleic acids

2.6.1 Extraction of DNA fragments from agarose gels

DNA samples were mixed in a 1:1 ratio with gel-loading dye and run on 0.8% or 1.3% agarose gels, as described in section 2.5.1. DNA bands for extraction were excised from the gel and eluted using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's instructions. Extracted DNA fragments were eluted from the QIAquick columns in 50 µl sterile distilled water.

2.6.2 Electroelution of DNA fragments from polyacrylamide gels

DNA samples to be electroeluted were first run on 7.5% Polyacrylamide gel electrophoresis as mentioned in 2.4.2. However, the wider well sizes were made so that large amounts of the samples could be loaded. The DNA band was then excised once the gel had run, stained, and visualized. The band was then placed in a 6.3 mm dialysis tube (Medicell International Ltd.); the tube was filled with 0.1 x TBE buffer and sealed using clips. The dialysis tubes containing the gel were placed in an electroelution running tank and electroelution was run at 30 mV for 30 minutes. At the end of the electroelution run, the electrodes were inversed and the electroelution was run for one minute. The buffer was drained from the tube into a 1.5 ml microfuge tube then 100 µl distilled water was added to the dialysis tube and poured into the

same microfuge tube. The DNA fragment was then treated with phenol/chloroform and ethanol and precipitated as mentioned in 2.6.3 and 2.6.4.

2.6.3 Phenol/chloroform extraction of DNA

Electroeluted DNA was further treated with phenol/chloroform to remove contaminants such as proteins. Electroeloted DNA was mixed with an equal amount of phenol/chloroform solution and vortexed for 15 seconds and centrifuged for 3 minutes at 18000 x g. The sample after centrifugation formed three or two layers. The DNA layer was in the aqueous solution at the top, which was then collected carefully into a new 1.5 ml microfuge tube. The procedure of adding phenol/chloroform was repeated until the final volume of purified DNA was 300 µl. Then 10% 3 M Na acetate pH 5.2 was added to the solution and finally, 2-3 times 100% ice-cold ethanol were added, as mentioned above in 2.6.4.

2.6.4 Ethanol precipitation of DNA

After the addition of 100% ice-cold ethanol the sample was left in -80°C for one hour or at -20°C overnight. Then the sample was centrifuged for 15 minutes at 4°C, then 1 ml 70% ice-cold ethanol was added and the sample was subjected to centrifugation for 10 minutes at 4°C. Then, the supernatant was discarded and the remaining pellet was dried using a speed vec dryer at medium temperature for 10-15 minutes. The pellets were finally re-suspended in 20-50 µl distilled water and the purified DNA sample was kept at -20°C.

2.6.5 Purification of DNA using QIAquick PCR purification kit (Qiagen)

PCR products or the digested DNA fragments were then purified using the QIAquick PCR purification kit according to the protocols provided by Qiagen. Purified DNA was collected in 30-70 µl elution buffers or in distilled water and the sample was stored at 20ºC.

2.7 Isolation of genomic DNA using illustra bacteria genomicPrep Mini Spin kit

2.7.1 Isolation of genomic DNA on a small-scale

Illustra bacteria genomic Prep Mini Spin Kit from GE Healthcare was used to isolate bacterial genomic DNA on a small-scale according to the kit instructions and guidelines. In short, a single colony was inoculated into a 25 ml flask with 5 ml LB media and left overnight in a shaker at 37°C. 1 ml from the overnight culture was centrifuged and the supernatant was decanted. The pellet was then lysed using detergent and salt together with proteinase K so that the genomic DNA was separated from the bacterial cell debris into the solution. To get rid of RNA contamination RNase was used. The solution was then transferred to a silica column with chaotrope, so genomic DNA was bound to the column. Debris, such as proteins, was separated from the membrane-bound genomic DNA using chaotrope salt. To get rid of salts and interfering substances and to dry the silica membrane that contained bound genomic DNA Ethanolic wash buffer was used. Finally, 200 µl low ionic strength buffer was used to elute genomic DNA into a sterile 1.5 ml microfuge tube and then it was stored at -20°C in a freezer.

2.7.2 Small-scale preparation of plasmid DNA using QIAprep Spin Miniprep kit

For rapid isolation of plasmid DNA Qiagen kit was used, which provided spin columns lined with silica membranes, so that plasmid DNA was bound while other contaminants passed through and were discarded. The protocol was made for purification and isolation of high copy number plasmid DNA up to 20 µg. Other plasmids, such as low copy number or large plasmids > 10 kb, were also purified by this kit with little modification to the protocols. Firstly, a single colony was inoculated into a 5 ml LB, containing antibiotics if necessary, and left to grow in a shaker at 37°C overnight (12-16 hours). Later, the culture was transferred into a sterile 1.5 ml microfuge tube and subjected to centrifugation at 14,000 rpm for one minute then the supernatant was discarded. The step was repeated to pellet all 5 ml of bacterial culture. The miniprep kits take advantage of a silica gel membrane to which the plasmid DNA is bound. It is then separated from other cellular components, washed and eluted. The kit enables the purification of up to 20 µg high-copy plasmid DNA from 1-5 ml overnight cultures of *E. coli* grown in LB. A single bacterial colony was picked up from an agar plate and used to inoculate 5 ml of LB, supplemented with appropriate antibiotic, in a 25 ml conical flask. The culture was incubated for at least 8 hours at 37°C with vigorous shaking. From this culture, 1.5 ml of cells were taken to a sterile microfuge tube and harvested by centrifugation at 14,000 rpm for 1 minute. The bacterial pellets were then re-suspended and lysed in a NaOH/SDS buffer. The lysate was then neutralised and set to high salt-binding conditions. The lysing step was implemented to denature proteins, genomic DNA, cellular contaminants, and SDS, but to keep the plasmid DNA maintained in the solution. The lysate was subjected to 15 minutes of centrifugation at 14,000 rpm then the contaminated precipitate at the bottom of the microfuge tube was discarded, leaving the plasmid DNA at the top in a clear solution layer. This top layer was transferred into a spin

column provided with the kit. Several washing steps with various buffers, according to the protocol provided, were applied and the plasmid DNA was bound to the silica gel membrane that was eluted at the final step into a 1.5 ml microfuge tube with 30-50 µl elution buffer that contained 10 mM Tric-Cl at pH 8.5. Finally, the purified plasmid was stored at -20°C in a freezer. For low copy number plasmids and large plasmids > 10 kb, the washing buffer PB was required for all strains used, and the elution buffer at the final step was preheated to 70°C before use.

2.7.3 Large-scale preparation of plasmid DNA using QIAprep Spin Miniprep kit

For large-scale preparation of plasmid DNA, QIAfilter Plasmid Maxi Kit (QIAGEN) was used, from bacterial cells as mentioned in the kit protocol. In short, a single colony was inoculated into 5 ml LB, supplemented with suitable antibiotics, and left to grow for about 8 hours at 37°C in a shaker with vigorous shaking. Then 100 ml LB media, supplemented with appropriate antibiotics, was added to a 500 ml flask and 200 µl from the starting culture was inoculated into it, so the starting culture was 1/500 times diluted. The flask was then incubated at 37°C for 16 hours with vigorous shaking. The overnight cultures were then transferred into 50 ml Falcon tubes and subjected to centrifugation at ~3400 x g for 30 minutes at 4°C. The pellets were then re-suspended in a re-suspension buffer that contained RNaseA, and they were lysed using NaOH/SDS. Potassium acetate was added to neutralize the pH. Contaminants such as cell debris, genomic DNA, proteins, and SDS were filtered using a QIAfilter cartridge and the lysates were passed through an anion-exchange resin by gravity flow under appropriate low salt and pH conditions that help bind the plasmid DNA to the resin. Proteins, RNA, or low molecular weight contaminants were removed from the

plasmid DNA that was bound onto the resin by washing with medium salt. It was then eluted in a high salt buffer, concentrated and the salts were removed using isopropanol and 70% precipitation and it was left to air dry. Finally, the purified plasmid DNA was dissolved in TE buffer and stored at -20°C in a freezer.

2.8 Transformation methods of *E. coli* with DNA

2.8.1 Using $CaCl_2$ method to prepare competent cells

A single colony was inoculated with 5 ml LB, supplemented with antibiotics if necessary, then left to grow overnight at 37°C in a shaker. From the overnight culture, 1 ml was subcultured into fresh 50 ml LB in a 250 ml flask, supplemented with antibiotics if necessary, and left to grow until the mid-logarithmic phase where the OD_{650} = 0.3-0.6 in a shaker at 37°C. Once the cells had grown to a mid-logarithmic phase, they were collected into a 50 ml sterile Falcon tube, and it was placed on ice for ten minutes. After that, cells were subjected to centrifugation at 4000 rpm at 4°C. 25 ml ice-cold 0.1 M $CaCl_2$ was used to re-suspend the pellet. The suspension was kept on ice for ten minutes, and then it was centrifuged for five minutes at 4000 rpm. The supernatant was discarded and the pellet was re-suspended into 3.3 ml ice-cold 0.1 M $CaCl_2$ + 10% glycerol and kept on ice for 24 hours. 200 µl cell suspension was aliquoted into 1.5 sterile microfuge tubes and the cells were stored at -80°C in a freezer.

2.8.2 Using RbCl method to prepare competent cells

A single colony was inoculated into 5 ml LB, supplemented with antibiotics if necessary, and left to grow overnight at 37°C in a shaker. Then 1 ml from the overnight culture was subcultured in 100 ml LB in a 250 ml flask, supplemented with antibiotics if necessary, and incubated at 37ºC in a shaker until the culture reached the mid-logarithmic phase OD_{650} = 0.3-0.6. The culture was then transferred to two sterile 50 ml Falcon tubes and kept on ice for ten minutes. The cells were subjected to centrifugation for five minutes at 4500 rpm at 4ºC. The supernatant was discarded and the pellet from one Falcon tube was re-suspended in ice-cold 40 ml (1/2.5volume) of TFB1 buffer, and then the content was transferred to the second Falcon tube to re-suspend the pellet in it. This was then left on ice for ten minutes. After that, the cells were subjected to centrifugation for five minutes at 4500 rpm. The supernatant was discarded and the pellet was re-suspended in 4 ml (1/25 volume) ice-cold TFB2 buffer and kept on ice for an hour. The cells were ready for transformation or storage in 200 µl aliquot in sterile microfuge tubes and were stored at -80ºC in a freezer.

2.8.3 Chemically competent cell transformation with plasmid DNA

Cells to be transformed were removed from a freezer at a temperature of -80ºC and placed on ice to thaw. 50-100 µl competent cells were then mixed with 1-3 µl plasmid DNA on ice. Cells were incubated on ice for 45-60 minutes. After that, cells were heat-shocked for 90 seconds at 42°C then placed on ice for two minutes. 1 ml LB or SOC media was added to the cells and cells were placed in a shaker at 37°C for one hour. Cells were then centrifuged for two minutes at 9,000 rpm, the supernatant was decanted, and then about 100 µl media was left to re-suspend the pellet which was then plated onto an agar plate, supplemented with appropriate antibiotics, and incubated overnight at 37°C.

2.8.4 Using electroporation method to transform *E. coli* with plasmid DNA

2.8.4.1 Preparation of cells

$CaCl_2$ or RbCl methods of making competent cell strain K-12 were not suitable for transforming other strains of *E. coli* such as EDL933 or Sakai strains, so electroporation was used as an alternative method of transformation. The cells were first prepared by inoculating a single colony into 5 ml LB, supplemented with antibiotics if necessary, and left to grow at 37°C in a shaker overnight. Then 500 µl overnight cultures were subcultured in 50 mL LB medium in a 250 ml flask, supplemented with appropriate antibiotics if necessary, and placed in a shaker at 37°C. When cells reached the mid-exponential growth phase OD_{650} = 0.3-0.6, they were collected in a 50 mL Falcon tube and subjected to centrifugation at 4500 rpm at 4°C for ten minutes. Pellets were re-suspended in 25 ml of ice-cold 10% glycerol and the cells were centrifuged at 4500 rpm for ten minutes at 4°C. Pellets were re-suspended in 1 ml of ice-cold 10% glycerol, transferred to a 1.5 ml Eppendorf tube and centrifuged at 13000 rpm for one minute at 4°C. The supernatant was removed and the pellet was re-suspended in 500 µl ice-cold 10% glycerol and kept on ice.

2.8.4.2 Electroporation

Before electroporating cells, SOC media was placed at 37°C in an incubator, a P1000 pipette was set to 1 ml and a P100 to 40 µl, I mM cuvettes were placed on ice, the electroporator was

Table 2.3: PCR cycle

Step	Temperature	Duration	No. of Cycles
Pre-denaturation	98ºC	30 seconds	1
Denaturation	98ºC	5-10 seconds	25-35
Annealing	45-72ºC	10-30 seconds	
Elongation	72ºC	30 seconds per 1 kb	
Final elongation	72ºC	5-10 minutes	1
Hold	-10ºC	-	-

set to 1.8 kilo-volts, post-electroporation tubes were kept in a rack and a dry orbital shaker was turned to 37°C. Once the items were prepared, 1-5 µl plasmid DNA was transferred into a microfuge tube and placed on ice, and 40 µl cells were added to the DNA on ice. Cells were mixed up and down twice using a pipette. Then the contents were instantly transferred to an electroporation cuvette that was subjected to pulse in the electroporator. This was followed by immediate addition of 1 ml of 37°C SOC, followed by transfer into a 1.5 ml tube. The tube was then placed at 37°C in a heat rack. Once all electroporations were done, the tubes were placed in the dry orbital shaker for vigorous shaking (approximately 250 rpm) for an hour at 37°C. The cells were then subjected to centrifugation at 13,000 rpm for one minute, then the supernatant was decanted and about 100 µl of the supernatant was left to re-suspend the pellet, which was then plated onto a suitable agar, supplemented with appropriate antibiotic(s), and incubated overnight at 37°C.

2.9 Recombinant DNA techniques

2.9.1 Polymerase chain reaction (PCR)

To amplify a gene of interest on a DNA strand, polymerase chain reaction (PCR), was used. The reaction is composed of elements that are essential for a successful yield of PCR products, and these include template DNA, two oligonucleotide primers, DNA polymerase, deoxynucleoside triphosphate (dNTPs), reaction buffer, and $MgCl_2$. The primers were constructed in such a way as to make them anneal to the DNA complementary strands on both sides of the template with their 3' facing each other. The primers' lengths were between 20-25 nucleotides, and enzyme restriction sites were added to both primers, so that amplified

fragments could be cloned into suitable vectors. The PCR reaction amplifies DNA by cycling three different steps. The reaction is composed of three stages; the first stage is denaturation where the temperature of the reaction rises to unfold the double-stranded DNA into two single strands. Second, the reaction temperature drops down so that the primers can anneal to the complementary sequence on the single stranded DNA template. In the third stage the reaction temperature rises again to allow extension of the newly formed DNA strands from the 3' end of each primer. These stages were repeated for a number of cycles to ensure yields, and they were preceded by a pre-denaturation step and followed by a post-extension step as well Table. 2.3.

Q5® high-fidelity DNA polymerase (New England BioLabs® Inc.) is a thermostable enzyme that possesses 3'➔5' exonuclease activity and robust DNA amplification with an error rate > 100-fold lower than that of *Taq* DNA polymerase. This enzyme is provided with an optimized buffer, which facilitates robust amplification regardless of the GC content. The reaction buffer supplied is 5 x Q5 that contains 2 mM $MgCl_2$ at a final concentration of 1 x reaction concentration. Using 5 x Q5 High GC enhancer is recommended when the target is rich in GC content. The reaction was made in either 50 µl or 25 µl total volume, as follows; 10 µl 5 x Q5 reaction buffer, final concentration 1 x; 1 µl 10 mM dNTPs, final concentration 200 µM; 2.5 µl 10 µM forward primer, final concentration 0.5 µM; 2.5 µl 10 µM reverse primer, final concentration µM; template DNA <1,000 ng; Q5 high-fidelity DNA polymerase 0.5 µl, final concentration 0.02 U/ µl; 5 x high GC enhancer (optional) 10 µl, final concentration 1 x; and nuclease free water to make the volume up to 50 µl. DNA amplification was performed in an oil-free thermal cycler (GeneAmp® PCR System, Applied Biosystems). The PCR cycling conditions are shown in table 2. 3. The annealing temperature was calculated using an NEB melting temperature calculator.

Finnzymes' Phusion™ High- Fidelity DNA polymerase (New England BioLabs® Inc.) supplemented with 5 x Phusion HF or GC buffer was used for PCR as a robust method second to the previous one. The enzyme properties of 5→3 DNA polymerase activity and 3→5 exonuclease activity allow it to clone the DNA fragment with enhanced accuracy and in a shorter reaction time. This method utilises small amounts of template DNA so plasmid DNA used as a template was diluted 1:100 and 1:200 in sterile distilled water and 1 µl was used in a 50 µl reaction volume. PCR reactions were performed in a 50 µl reaction mix containing 10 µl of 5 x Phusion HF buffer containing 1.5 mM $MgCl_2$, 1 µl of 10 µM forward and reverse primer (Alta Biosciences, University of Birmingham, UK), 1 µl of 0.8 mM dNTPs (Bioline), 1µl of 10-50 ng of template DNA, 1 unit of Phusion DNA polymerase (New England BioLabs® Inc.) and a suitable amount of sterile distilled water. DNA amplification was performed in an oil-free thermal cycler (GeneAmp® PCR System, Applied Biosystems). PCR cycling conditions are shown in Table 2.3. For primers > 20 nt, annealing temperature (X) was maintained at Tm +3°C of the lower Tm primer, whereas for ≤ 20 nt, Tm of the lower Tm primer was used as the annealing temperature.

Taq DNA Polymerase with Standard *Taq* Buffer (M0273) (New England BioLabs® Inc.) was used in this study to check for successful cloning. The reaction was made in 50 µl total volume as follows: 5 µl 10 x standard *Taq* reaction buffer, final concentration 1 x; 1 µl 10 mM dNTPs, final concentration 200 µM; 1 µl 10 µM forward primer, final concentration 0.2 µM; 1 µl 10 µM reverse primer, final concentration 0.2 µM; template DNA < 1,000 ng; 0.25 µl *Taq* DNA polymerase, final concentration 1.25 units/50 µl PCR; and nuclease-free water to make the volume up to 50 µl. DNA amplification was performed in an oil-free thermal cycler (GeneAmp® PCR System, Applied Biosystems). PCR cycling conditions are shown in Table

2.2, and the annealing temperature was calculated using an NEB annealing temperature calculator.

BioMix™ Red (Bioline) was used to check for successful cloning, as it is a convenient method. It contains pre-mixed, pre-optimized 2 x solution of ultra-stable *Taq* DNA polymerase, can amplify fragments up to 5 kb, the risk of contamination is reduced, the duration of the reaction is dramatically reduced, the results are reproducible, and products can be loaded directly into the gel as the mix contains inert red dye. The reaction mixture was as follows: the reaction took place in 50 µl total volume, of which 25 µl was BioMix Red buffer containing 6 mM Mg^{+2}. The final concentration in the reaction was 3 mM, 1 µM final concentration, of both forward and reverse primers, and template DNA. DNA amplification was performed in an oil-free thermal cycler (GeneAmp® PCR System, Applied Biosystems). PCR cycling conditions are shown in Table 2.3.

2.9.2 Colony PCR

To screen for successful cloning, deletions or insertions in the chromosomal DNA or to check for the presence of a cloned fragment of a plasmid, colony PCR was performed. A single colony was re-suspended into 50 µl distilled water and boiled at 100°C for ten minutes then 1-5 µl was used in the PCR reaction. The PCR method used for screening was BioMix™ Red.

2.9.3 Error-prone PCR

Table 2.4: Primers used in this study

Code	Sequence 5' to 3'	Description
D61221	GCA**GAATTC**TGCACCCGTTCCA GG	Upstream primer containing EcoRI site, used with D68959 to amplify LEE20-203 promoter
D68959	GCA**AAGCTT**TTGACATTTAATG A TAATGT	Downstream primer containing HindIII site, used with D61221 to amplify LEE20-203 promoter
D74087	GCA**GAATTC**AAATGTGATGTA CATCACATGGATCCAGATCAAT TCTTGACATTTAATG	Upstream primer containing EcoRI site, used with D64100 to create CRP site upstream of LEE20-203 promoter at position -61.5
D64100	GCA**AAGCTT**ATTCTCTTTTTTC T AATG	Downstream primer containing HindIII site for amplification of LEE10-568 promoter
D76829	GCA**GAATTC**AAATGTGATGTA CATCACATGGATCCAGATCAAT TCTTGACATTTAATGATAATGT ATTTT**C**CACA	Upstream primer containing EcoRI site, used with D64100 to create point mutations in position 9 of LEE20-203 promoter
D74088	GCA**GAATTC**AAATGTCATGTA **C**AT**G**ACATG	Upstream primer containing EcoRI site, used with D64100 to create point mutations in CRP binding site upstream of LEE20-203 promoter
D77199	GCA**GAATTC**AAATGTGATGTA CATCACATGGATCCAGATCTTG ACATTTAATGATAATGTATTTT ACACA	Upstream primer containing EcoRI site, used with D64100 to create CRP site upstream of LEE20-203 promoter at position -55.5
D77200	GCA**GAATTC**AAATGTGATGTA CATCACATGGATCCTTGACATT TAATGATAATGTAGTTTACACA	Upstream primer containing EcoRI site, used with D64100 to create CRP site upstream of LEE20-203 promoter at position -51.5
D76830	GCA**GAATTC**AAATGTGATGTA CATCACATCAATTCTTGACATT TAATGATAATGTATTTT**C**CACA	Upstream primer containing EcoRI site, used with D64100 to create point mutations in position 9 of LEE20-203 promoter
D53463	GGGGGATGTGCTGCAAGGCG	Downstream primer that anneals within *lacZ* coding sequence downstream from BamHI and HindIII sites in pRW224 and pRW225, used for sequencing of inserts in these vectors.
D63949	GCA**GAATTC**CGTTTGTTAACGA GATGATTTTCTTC	Upstream primer used in PCR reaction with D79264 to amplify LEE150
D65474	GCA**GGATCC**CAAGCTTTAGGA C	Downstream primer used with D63949 to amplify LEE30-275 fragment

	ACATC	
D66870	GAAATATATGTGTCCTAAAGCTTT	Upstream primer used in PCR reaction with D53463 to produce megaprimer to make LEE30-275 46T fragment
D76233	CGT**AAGCTT**TAGGACACATATATTTCATCAAACAACCTCCTTAAAATGT	Downstream primer containing HindIII site, used with D61221 to create point mutation at position 28 in LEE30-275 46T fragment to make LEE30-275 28A 46T
D79263	GCA**CCATGG**CATGCGGAGATTATTTATTATG	Upstream primer containing NcoI site, used with D61221 for amplification of LEE150
D79264	GCA**CCATGG**TCTTTTTTCTAA	Downstream primer containing NcoI site, used to amplify LEE150 with D79263
D65811	GCA**AAGCTT**CATAATAAATAATCTCCG	Downstream primer containing HindIII site, used with D61221 for amplification of LEE150
D75954	GCA**CCATGG**TCTTTTTTCTAA	Downstream primer containing KpnI site, used with D61221 for amplification of LEE150
D75956	GCA**GGTACC**CATGCGGAGATTATTTATTATG	Upstream primer containing KpnI site, used with D75957 for amplification of LEE150
D75957	GCGATTAAGTTGGGTAACGCCAGGGTT	Downstream primer used with D75956 for amplification of LEE150
D58793	CCTTGCAGCACATCC	Downstream primer used with D75956 for amplification of LEE150
D78791	GCA**AAGCTT**CATAATAAATAATCTC**T**G	Downstream primer containing HindIII site and a point mutation in the first base of Shine-Dalgarno at +156 sequence fused to *ler* (**G**GAG)➔(**A**GAG)
D78792	GCA**AAGCTT**CATAATAAATAATCTC**A**G	Downstream primer containing HindIII site and a point mutation in the first base of Shine-Dalgarno at +156 sequence fused to *ler* (**G**GAG)➔(**T**GAG)

LEE10-568 (549 bp)

EcoRI -568

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

-100 -90 -80 P1 UP elements -35 element

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

-10 element

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTGATGAAATAGATGTGTCCTAATTTGATAGATAAACGTTATCTCA

P2

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**AAGCTT**

HindIII

LEE20-203 (46 bp)

18 bp

GAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

-35 element -10 element

LEE20-203 99T(46 bp)

GAATTCTTGACATTTAATGATAATGTATTTTAT**T**ACATTAGAAAAAAGAGAAT**AAGCTT**

Fig. 2.8: DNA sequences of LEE10-568, LEE20-203 and LEE20-203 99T promoters.
DNA sequences of LEE10-568, LEE20-203, and LEE20-203 99T promoter fragments of *LEE*1 regulatory region. EcoRI and HindIII sites are in bold, -10 and -35 elements are underlined for P1. -10 and -35 elemtents for P2 are in grey shades.

GrlRA (1035 bp)

HindIII

AAGCTTTTTTACGTTGTTACTCAATATTATTAATCAGAAATTACATATGTTAACCAGGGAAACAGCAGGTTGAAACATGAGTATATTTAATGATATA

TTACATTGCAATCTGGAGAAAAAGAAAGGTCTCCATTATTCTTGATATTGCTTATGGATAGAACAAATTGAAAGGAGTGAGGTTAGTATGAAACTGA

GTGAGTTATGATTATGAAGGATGGCATCTATAGCATTATATTTATTAGCAATGAAGACTCCTGTGGGGAAGGTATACTGATTAAAAATGGAAATATG

ATCACTGGCGGCGATATTGCTTCTGTGTATCAGGGGGTCCTCTCTGAAGATGAGGACATCATACTTCATGTCCATCGATATAATTACGAAATTCCCT

CGGTGCTAAACATTGAACAAGATTATCAATTAGTTATCCCTAAAAAAGTACTGAGTAATGATAATAATCTCACATTACATTGCCATGTAAGAGGAAA

TGAAAAATTGTTTGTTGATGTTTATGCCAAATTTATAGAACCATTAGTTATTAAAAACACAGGAATGCCACAAGTTTATTTAAAATAATTTTATGTA

TGTTTTTTATGTCGATTTATTTATCAAATAAAAAGAATATGGAAAATGGAATCTAAAAATAAAAATGGCGACTATGTAATTCCTGACTCAGTAAAGA

ATTACGATGGTGAACCTCTGTATATCTTGGTTTCTCTTTGGTGTAAATTGCAGGAGAAATGGATTTCTCGCAATGATATTGCCGAAGCATTCGGTAT

AAACCTGAGGAGAGCATCATTTATTATAACTTATATATCGAGAAGAAAAGAAAAAATTTCATTTCGTGTCAGATATGTTAGTTATGGTAATTTGCAT

TATAAGCGCCTTGAGATTTTCATTTATGATGTTAACCTTGAGGCGGTTCCGATAGAAAGTCCTGGAACAACCGGACCAAAAAGAAAAACCTACCGAG

TTGGTAATGGTATTGTGGGACAGTCTAATATCTGGAACGAAATGATCATGAGGCGGAAAAAGGAGAGTTAG**GTCGAC**

SalI

Fig. 2.9: DNA sequences of *grlRA* region.

The figure shows the DNA sequences of *grlRA* region. HindIII-SalI sites are in bold, -10 and -35 elements are shaded yellow. Start codons of both *grlR* "top" and *grlA* "bottom" are shown in green and stop codons are in red. Genes are separated by a short intergenic sequence shown in blue.

LEE30-275 (162bp)

-35 element

EcoRI

GAATTCCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAATG

TATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTGATGAAATAAATGTGTCCTA**AAGCTT**

-10 element Shine-Dalgarno sequence HindIII

LEE30-275 (46T)

GAATTCCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAATG

TATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTGATGAAATATATGTGTCCTA**AAGCTT**

LEE30-275 (28A 46T)

GAATTCCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAAT

GTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGAGG**TTGTTTGATGAAATATATGTGTCCTA**AAGCTT**

Fig. 2.10: EcoRI-HindIII LEE30-275 promoter fragments cloned into pRW225.
The figure shows the DNA sequences of three LEE1 promoter regulatory regions. EcoRI-HindIII sites are in bold, -10 and -35 elements are underlined. SD-sequence in each fragment is shaded grey and point mutations both in SD-sequence and in the stop codon are indicated by yellow shading.

LEE150 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTGATGAAATAGATGTGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGCATGC**GGAG**ATTATTTATTATG**AAGCCT**

LEE150-1 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGAGG**TTGTTTGATGAAATAGATGTGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGCATGC**GGAG**ATTATTTATTATG**AAGCCT**

LEE150-2 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTG**ATG**AAATAG**ATG**TGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**ATG**C**T****GAG**ATTATTTATT**ATG****AAGCCT**

LEE150-3 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GG****A****GG**TTGTTTG**ATG**AAATAG**ATG**TGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**ATG**C**T****GAG**ATTATTTATT**ATG****AAGCCT**

LEE15-4 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTG**ATG**AAATAG**ATG**TGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**ATG**C**A****GAG**ATTATTTATT**ATG****AAGCCT**

LEE150-5 (571 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GG****A****GG**TTGTTTG**ATG**AAATAG**ATG**TGTCCTAATTTGATAGATAAACGTTATCTCA

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**ATG**C**A****GAG**ATTATTTATT**ATG****AAGCCT**

LEE150 Δ*l* (427 bp)

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

TGTATTTTACACATTAGAAAAAA**GGTACCC**ATGC**GGAG**ATTATTTATTATG**AAGCCT**

Fig. 2.11: DNA sequences of LEE150 and LEE150 Δ*l* fragments of *LEE1* regulatory region.
*Eco*RI and *Hind*III sites are in bold. Predicted proximal and distal promoters are shaded yellow and grey respectively. Nucleotide sequences are numbered relative to the transcription start site of the predicted proximal promoter. Dark grey shades in LEE150 correspond to a transcriptional start site (TSP) that has a ribosome binding site sequence underlined in front of it. The light grey shades represent the predicted ORF in the fragment.

LEE20-203 (46 bp)

*Eco*RI *Hind*III

GAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

-35 element -10 element

GS100 (82 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS100 (9C) (82 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTT**C**CACATTAGAAAAAAGAGAAT**AAGCTT**

GS101 (82 bp)

GAATTCAAA**TGTCA**TGTACA**TGACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS102 (77 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS103 (72 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS103 (9C) (72 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCTTGACATTTAATGATAATGTATTTT**C**CACATTAGAAAAAAGAGAAT**AAGCTT**

Fig. 2.12: DNA sequences of LEE20-203 fragment and semi-synthetic promoters.
The figure shows the DNA sequences of *LEE20-203* fragment of the *LEE*1 regulatory region. *Eco*RI and *Hind*III sites are in bold, -10 and -35 elements are underlined. CRP binding sites in the semi-synthetic promoters are in blue and point mutations in each fragment are in red.

To produce a random mutagenic library of DNA, error-prone PCR was used. This method relies on the low fidelity of the *Taq* polymerase enzyme and the higher error rate of 0.02% (Eckert and Kunkel, 1991). The protocol was modified from the classic PCR protocol where different concentrations of some elements were changed to cause random mutation in the DNA template in addition to the error generated by the enzyme itself. The concentration of dNTPs and Mg_{2+} were increased in the PCR reaction mix (Leung *et al*., 1989).

Mutagenic or error-prone PCR reactions were performed in a 50 µl reaction mix containing various concentrations of 2-5 µl 50 mM $MgCl_2$ (Bioline); 5µl (10 µM) each of the two primers (Alta Biosciences, University of Birmingham, UK); 2 µl (2.5 mM) dNTPs (Bioline); 1 µl of template DNA (1-10 ng); 5 µl of 10 × NH4 buffer (no Mg Cl_2); 1 unit of Biotaq polymerase (New England BioLabs® Inc); and a suitable amount of sterile distilled water. DNA amplification was performed in an oil-free thermal cycler (GeneAmp® PCR system).

The PCR cycles were as follows: pre amplification denaturation at 95°C for 5 minutes, then 10 seconds denaturation at 95°C, 1 minute annealing at 60°C, and a 2 minute extension at 74°C. the cycle was repeated 40 times. The final stage was a 5 minute extension at 74°C to allow complete extension of the all amplified fragments. The PCR products carrying randomly mutagenized promoter fragments were purified either by a QIAquick PCR purification kit (Qiagen), by extraction from 1.3% agarose gel or by electroelution from 7.5% polyacrylamide gels (see sections 2.5.3-2.5.5), depending on the size of the promoter fragment. PCR product was then digested and ligated into a cloning vector. Cloned plasmid was transformed into M182 with the wild type plasmid being used to screen the library. Transformants were plated on MacConkey lactose plates, supplemented with suitable

antibiotics. Phenotype change was monitored either up or down mutants that were then restriked onto MacConkey lactose plates to ensure the phenotype and to get single colonies. Then a single colony was sat for overnight culture to miniprep the plasmid. The plasmid was then retransformed to check for the phenotype and digested to check for the fragment size. Finally, the plasmid was sequenced using an appropriate primer.

2.9.4 Site-directed mutagenesis

Site-directed mutagenesis was performed to introduce specific mutation in a DNA sequence of interest, cloned in plasmid pRW224 or pRW225 using a downstream or upstream oligonucleotide primer containing the desired base change, together with an upstream or a downstream primer. The amplified products were then digested with appropriate restriction enzymes and cloned into plasmid pRW224 and pRW225. A number of oligonucleotide primers used for site-directed mutagenesis are listed in table 2.4.

2.9.5 Restriction digestion of DNA

Restriction digestion was used to prepare both a DNA fragment and a cloning vector for ligation. 50 µl DNA (purified PCR product or plasmid miniprep) was digested using 2 µl of each restriction enzyme (all from New England BioLabs® Inc.) in a final volume of 60 µl of the appropriate buffer, as determined using the New England BioLabs® Inc double digest finder. The procedure was performed on ice. Then the reaction was incubated for three hours in a 37°C heat block. For the vector, a further treatment with alkaline phosphatase was performed to remove terminal 5' phosphate groups and stop the vector from religating. 3 µl

alkaline phosphatase added to the reaction and the mixture was incubated for one hour at 37°C, then the vector was phenol/chloroform treated and gel extracted. For a DNA fragment digest, it was run on either agarose gel or 7.5% PAGE and then electroeluted (see sections 2.5.3-2.5.5).

2.9.6 Ligation

Once the cloning vector and the insert were purified after restriction digestion, they were ligated. The ligation was performed in a 20 µl reaction mix containing 50 ng of vector, 50ng insert, 2 µl ligation buffer, and 1 µl T4 ligase enzyme. The reaction was left at room temperature for ten minutes then on ice for ten minutes. For enhanced efficiency of the ligation, the mixture was left to ligate at 16°C overnight. Then 10 µl of the reaction was used for transforming M812 competent cells (see section 3.6). Transformants were plated onto nutrient agar supplied with appropriate antibiotic. Candidates were screened for successful cloning by first re-striking the colonies, then miniprepping the plasmid, which was then digested to make sure it produced the desired fragment size. Colony PCR was also performed on the plasmid to check for the presence of a desired fragment. The plasmid was then sequenced using sequencing primers to ensure correct fragment sequence.

2.9.7 Sequencing

To sequence a DNA fragment cloned into a plasmid, plasmid-to-profile sequencing was carried out by the Functional Genomics and Proteomics Laboratory, University of

Birmingham, UK. For sequencing of plasmid templates, 6.8 µl of plasmid miniprep was mixed with 3.2 µl of 1 µM sequencing primer. For sequencing of PCR products, 3 µl of purified PCR product was mixed with 3 µl of 1 µM sequencing primer and 4 µl of sterile distilled water. Primers used for sequencing inserts in plasmids are listed in Table 2.4.

2.10 Cloning of promoter fragments and/or target genes

Promoter fragments used in this study, genes, and regulator sequences are shown in Fig. 2.8, 2.10, 2.11, and 2.12. The fragments were amplified from *E. coli* O157: H7 Sakai chromosome by PCR. In the case of promoter mutations, a pre-cloned promoter into a plasmid was used as a template. The PCR primers were constructed with restriction sites at each end of both the forward and the reverse primers. The restriction sites were the same on the cloning vectors used to clone a fragment.

2.10.1 Cloning into pRW224

Promoter fragments for cloning into the *lac*Z fusion vector pRW224 were digested with EcoRI and HindIII enzymes and gel extracted from 1.3% agarose gel and then purified. The fragments were then cloned into pRW224/U9 after preparing it as mentioned in section 2.9.6. and 2.9.7. The sequence of candidates was confirmed (see section 2.9.8. LEE10-568, LEE20-203), and GS promoter derivatives were cloned into pRW224/U9.

2.10.2 Construction of pRW225 derivatives

Vector pRW225 was used to clone promoter fragments as translation phusion to *lacZ*. The vector was digested with *Eco*RI and *Hind*III enzymes as mentioned in section 2.9.6. The promoter fragments to be cloned were further prepared by digesting them with EcoRI and HindIII enzymes after amplifying them. Fragments were then ligated into pRW225 as mentioned in section 2.6.7. Successful candidates were then sequence confirmed as mentioned in section 2.9.8. LEE30-275, LEE30-275 derivatives, LEE150, and LEE150 derivatives were cloned into pRW225 in this study.

2.10.3 Construction of pRW400 derivatives

For constructing the pRW400 vector, pRW224/U9 was digested with BamHI and NheI enzymes. The vector was then purified and extracted from 0.8% agarose gel. This digestion removed lac genes in pRW224 ~ 6kb resulting in a 10kb vector. Green fluorescent protein was amplified from vector pJB (Bryant *et al.*, 2014), and it was digested with BamHI and NheI enzymes. The fragment was then purified and gel extracted from 1.3% agarose gel. The fragment was then cloned into pRW224. The resultant plasmid was renamed pRW400. Promoters LEE10-568, LEE20-203, and LEE20-203 99T were cloned into pRW400. Successful candidates where then sequenced as mentioned in section 2.9.8.

2.10.4 Construction of pSR derivatives

LEE150-1 fragment was constructed in the pSR vector. The primers used to construct *LEE*150-1 are mentioned in Table. 2.4. The vector was digested with EcoRI and HindIII enzymes, purified and gel extracted. *LEE*150-1 was made as two inserts. The first insert was digested with *Eco*RI and *Nco*I enzymes, and the second insert was digested with *Nco*I and *Hind*III enzymes. The fragments were electroeluted after extracting the gel slices from 7.5% Polyacrylamide gel. The fragments were then ligated into the pSR vector using three-way ligation. Successful candidates were confirmed by DNA sequencing. The insert *LEE*150-1 was then excised by digesting with *Eco*RI and *Hind*III enzymes and cloned into the pRW255 cloning vector as mentioned in section 2.11.2.

2.11 β-galactosidase assays

The activity of *lacZ* under the control of a promoter was measured by β-galactosidase assays. The assays were carried out using various bacterial backgrounds and under different growth conditions.

2.11.1 β-galactosidase assays during exponential growth phase

One overnight culture per strain each strain carrying a plasmid-encoded promoter fused to *lacZ* to be assayed. This was performed by inoculating a single colony in 5 ml LB with antibiotics. Then the flasks were left to grow overnight for 16-18 hours at 37°C in a shaker. Then on the next day, each overnight night culture was sub-cultured into a new flask. This was performed by, pipetting 100 µl of overnight culture and sub-culturing into 5 ml of LB

with antibiotics. Cells were left to grow at 37°C in a shaker to the mid logarithmic phase where OD_{650} = 0.3-0.6. Once the optimum optical density was reached, two drops each of toluene and 1% sodium deoxycholate were added to lyse the cells. They were then vortexed for 15 seconds and placed in the shaker at 37°C without bungs for 20 minutes. In the meantime, test tubes were placed in a 37°C water bath and 2.5 ml Z-buffer supplied with both ONPG and beat mercaptoethanol. When cells were lysed 100 µl of each cell's lysate was added to 2.5 ml Z-buffer and assayed for β-galactosidase activity. The reaction was performed at 37°C until a yellow colour developed, at which point the reaction was stopped by adding 1 ml 1 M sodium carbonate and the OD_{420} was measured. The time was recorded at which the lysate of the cells was assayed, as was the time taken by each reaction to develop the yellow colour, and the time difference was used in the formula:

$$\text{β-galactosidase activity} = \frac{1000 \text{ X } 2.5 \text{ X } 3.6 \text{ X } OD_{420nm}}{OD_{650nm} \text{ x } 4.5 \text{ x } t \text{ x } v} \text{ nmol/minute/mg bacterial mass}$$

Where,

2.5 = factor for conversion of OD_{650} into bacterial mass, based on OD_{650} of 1 being equivalent to 0.4 mg/ml bacteria (dry weight).

3.6 = final assay volume (ml)

100/4.5 = 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 (minutes)

v = volume of lysate added (ml)

The assay was performed in triplicate and repeated twice. The mean β-galactosidase activity and standard deviation were calculated each time. As a control, in all assays cells harbouring empty plasmid were used to blank the measurements.

2.11.2 β-galactosidase assays during stationary growth phase

For this condition, the assay was performed from an overnight culture after diluting the cells in a 1:10 mix with a suitable media. The OD_{650} was measured and the assay was carried out as mentioned in section 2.12.1. The assay was performed in triplicate and the experiment was repeated twice at least.

2.12 Attachment experiment of *E.coli* to eukaryotes

In order better to understand the relationship between the pathogen and the host, attachment experiments were used in part of this study. The eukaryote cell lines used were Caco-2 cells, which are intestinal, epithelial cells, mimicking host cell lines. In addition, HeLa cell lines were also used.

2.12.1 Cell splitting and preparation

In day one, Caco-2 cells were split into new 10 cm plates at a ratio of 1:10 as follows: the confluence of the cells was checked under the microscope (10 x lens or 40 x lens); when cells were near confluent, they were split by first adding 1 ml trypsin – the plate was washed in an upward and backward motion; the trypsin was then sucked away and another 1 ml trypsin was added to the plate before the plate was placed back in an oven at 37°C for five minutes. Cells coming off the plate were identified by seeing clouds moving and by checking under the microscope. Then 9 ml DMEM media containing foetal calf serum, penicillin, and streptomycin to prevent bacterial contamination was added to the plate. The plate was rinsed with media and the contents were removed into a 50 ml Falcon tube. Then 2 ml was dispensed into each plate. The plates were then left at 37°C in an oven and checked for a period of three to four days for confluence.

2.12.2 Bacterial cell inoculation

On day three, strains to be assayed were inoculated into LB media with antibiotics, so a fresh single colony was inoculated into 5 ml LB with antibiotics and left to grow overnight at 37°C in a shaker. The strains assayed in these experiments were EHEC Sakai strains.

2.12.3 Preparation of bacterial culture for adhesion experiment

On day four, from the overnight cultures, 100 µl was sub-cultured into new 5 ml LB media with antibiotics and left to grow to the mid logarithmic phase where the OD_{650}= 0.3-0.6. The optical density was noted for each strain. Then the amount of the culture needed for the

adhesion experiment was calculated as follows: OD_{650} of the strains/30 × 6 ml DMEM media to give a multiplicity of infection of 100. From each strain, 1 ml culture was collected into 1.5 ml microfuge tubes and lysed by toluene and 1% sodium deoxycholate and kept aside. At the same time, the cells were plated into LB agar in different dilutions to count the number of colony-forming units. The remaining cultures were used further for the attachment experiment. First, the media in cell line dishes prepared according to 2.12.1 were removed and the plates were washed twice with 5 ml PBS. Then the cultures were added to each plate (10 cm tissue culture plate), and the plates were incubated at 37°C for three hours. After that, the media was removed and the plates were washed with 5 ml PBS three times. In the final wash, 1 ml of the cells was collected into 1.5-microfuge tubes and lysed with toluene and 1% sodium deoxycholate and kept aside. At the same time, cells were diluted and plated into LB agar to count colony-forming units. Then the cells in the plates were lysed with 1 ml PBS and 1% triton-100. A lysis buffer was used to release the bacterial cells attached to the Caco-2 cells by lysing the Caco-2 cells and releasing the bacterial cells. Again, cells were collected into 1.5 ml microfuge tubes and lysed with toluene and 1% sodium deoxycholate. At the same times, cells were diluted and plated on LB agar to count colony-forming units. The samples collected in the microfuge tubes were then assayed by β-galactosidase (see section 2.11). The number of colonies was counted and used to calculate the β-galactosidase activity as follows: the OD of each culture was back calculated number of colonies at OD_{650} was counted, then number of colonies at OD_{650}=1 was calculated. Later OD_{650} of 1 was used as a constant to back calculate the OD_{650} of each step. Thus the optical density of each culture was calculated and entered into the β-galactosidase formula.

2.12.4 Attachment to HeLa cells

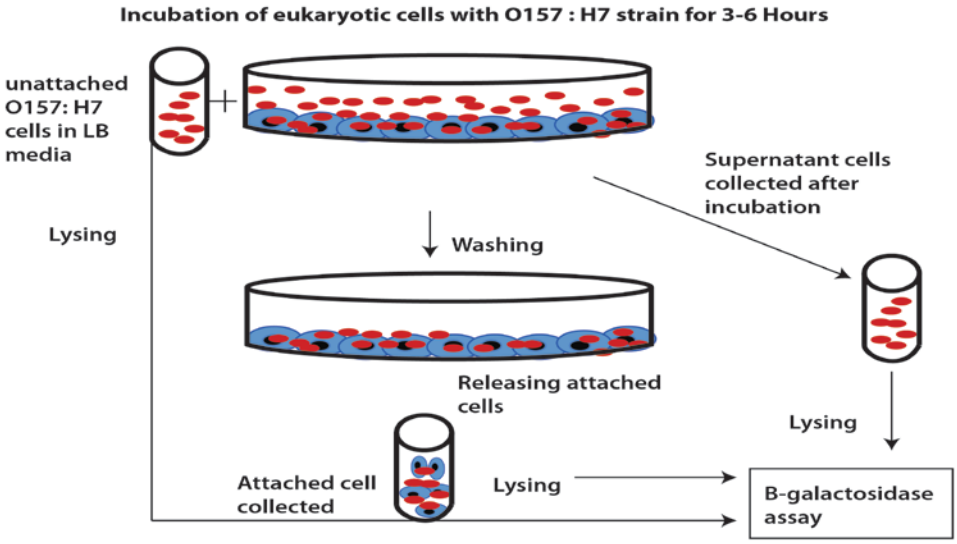


Fig. 2.13: Schematic representation of attachment procedure.
Sakai strains were grown to exponential growth phase then diluted in DMEM and incubated with Caco-2 cells for 3-6 hours. Cells were then collected for assay as described in 2.12.3.

Attachment to HeLa cells was performed on a 1 ml plate covered with HeLa cell lines. The bacterial culture was prepared as described in sections 2.12.1, 2.12.2 and 2.12.3,except that the amount of the media used to dilute the bacterial culture was 1 ml DMEM instead of 6 ml DMEM. Also, β-galactosidase assay was not done for this assay. The attachment to HeLa cells was done in a 6-well plate, instead of a 10 cm tissue culture plate. The plate was incubated for three hours at 37°C in an oven. The cells were then lysed with PBS and 0.1% Triton-100 and plated into LB agar to form colony-forming units.

2.13 Immunostaining for fluorescence microscopy

Fluorescence microscopy was another method used to visualize the impact of bacterial attachment to host cells. For this method, a promoter fragment was cloned into a *gfp*-expressing vector, pRW400, then the vector was co-transformed with pACYC184/GrlA vector into Sakai strains. These were then fixed onto HeLa cells coating cover slips and stained for microscopy.

2.13.1 Setting up cells on cover slips

Cells were grown near to the point of confluence (see section 2.12.1). Sterile coverslips were then used as follows: a paper towel was sprayed with ethanol in a hood, then individual cover slips were sprayed with ethanol on both sides of the towel. A second paper towel was sprayed with ethanol and placed over the cover slips to blot the cover slips to dry. Then the top towel was removed to allow the cover slips to air dry in the hood. Once cover slips were sterile and

dry, forceps sprayed with ethanol were used to place the cover slips into a six-well plate. HeLa cells were treated with trypsin and diluted in DMEM 1×10^5 then 2 ml of the cell suspension was pipetted into each well. The plate was placed in the 37°C oven overnight to allow the cells to grow near confluence.

2.13.2 Bacterial culture preparation

Strains to be tested were inoculated into 5 ml LB with antibiotics and left to grow over night at 37°C in a shaker. On the following day, 100 µl of each culture was sub-cultured into fresh 5 ml LB with antibiotics and left to grow to mid logarithmic phase at 37°C in a shaker. Then bacterial culture was diluted in 1 ml DMEM media after measuring the $OD_{650.}$

2.13.3 Infection of HeLa cells with bacterial cells

The cell lines in the six-well plate were checked for confluence then the media was removed and washed twice with 1 ml PBS. After that, 1 ml from each strain was pipetted into a well. The plate was then placed in an oven at 37°C for three hours.

2.13.4 Staining protocol

After incubation for three hours, the media was removed, and each well was washed with 1 ml PBP by tilting the plate. The solution was added slowly away from the cover slip. Then

PBS was aspirated and the cells were fixed with 1 ml 3.2% formaldehyde in PBS for 15 minutes at room temperature. Otherwise, fixation was left overnight in the fridge. Formaldehyde was then removed and the cells were made permeable ready for staining by adding 1 ml PBS+0.1 Triton X-100 to each well for five minutes. The solution was then aspirated and washed once with 1 ml PBS. 200 µl dye solution was added directly to the cover slips, left for ten minutes to stain, then washed three times with 1 ml PBS. The last wash was left on the cover slips to prevent drying out. Slides were labelled and 10 µl of the mounting solution was added to each slide. A bent needle was then used to flip up the cover slips. Tweezers were used to grasp the cover slips and place them onto the drop of the mounting media with cells facing down. The cover slips were then sealed with nail polish and the slides were kept in a dark slide box.

2.14 Coating of glass microscope slides

To test for a single molecule trigger of GrlA activation, glass slides were coated with single molecules. Cover slips were coated with either poly-L-lysine, which is a non-specific positively charged molecule that adheres to the cover slips, or collagen, which is a specific host molecule that may trigger the activation of GrlA. Cover slips were also coated with another specific animal protein, fibronectin.

2.14.1 Coating of glass microscope slides with poly-L-lysine and collagen

Firstly, cover slips to be coated were rinsed with ethanol and left to dry. Then they were laid out in petri dishes. Each cover slip was covered with 200 µl poly-L-lysine solutions. The cover slips were incubated at room temperature for one hour. Then when collagen was tested, firstly, poly-L-lysine was removed from the cover slips and 200 µl collagen was placed on each cover slip. Collagen was diluted in PBS from a stock solution of 100 µg/ml to make 10 µg/cm^2. Cover slips were incubated at room temperature for one hour. Once the cover slips coated with poly-L-lysine only, and those coated with collagen were ready, bacterial culture was added which was prepared first as mentioned in sections 2.12.2 and 2.12.3. Then the plates were incubated at 37°C in an oven for three hours. Then the staining protocol was performed as described in section 2.13.4.

2.14.2 Coating of glass microscope slides with fibronectin

A similar coating strategy to that in 2.14.1 was used with the fibronectin. However, the stock solution of fibronectin was diluted 1:5 in 1 x PBS. The working solution was 10 µl/cm^2.

Chapter Three: Bacterial immobilization induces *LEE* gene expression.

3.1 Introduction

In the *LEE* operons, two *LEE*-encoded regulators, GrlR and GrlA, are especially important for the regulation of *LEE* gene expression. This is because firstly, they are encoded within the *LEE* and expressed from a bicistronic operon between the *LEE1* and the *LEE2* operons. Second, they are major players in regulating expression of the *LEE* operons (Pallen *et al.*, 2005). GrlR binds to GrlA and stops it from binding to the LEE1 promoter (Deng *et al.*, 2004), but the mechanism by which GrlR-dependent repression of GrlA is relieved is unknown. Once it occurs, GrlA with Ler forms a positive regulatory loop and the formation of the T3SS is enhanced (Barba *et al.*, 2005). GrlA binds to the LEE1 promoter (Islam *et al.*, 2011) that results in enhanced expression of the *LEE1* operon. In particular, expression of Ler, which is believed to be the master regulator of the *LEE* operons, because Ler regulates the expression from all five *LEE* operons, including its own *LEE1* operon, is increased. Ler positively regulates expression from the *LEE* operons by counteracting H-NS, but negatively regulates its own operon, the *LEE1* operon. Expression of Ler is in turn regulated by a number of factors, including global regulators such as IHF, FIS, H-NS, and quorum sensing, and specific regulators, such as PerC homologues (which activate), and Ler itself (which represses).

The mechanism by which Ler regulates the expression of the different *LEE* operons is believed to be by alleviating the silencing effect of H-NS (Stoebel *et al.*, 2008). Ler relieves this repression under suitable conditions (Bustamante *et al.*, 2001). The condition under which GrlA is triggered, and the expression of the whole system is initiated, is unclear. It could be a factor that releases GrlA from GrlR, which would then allow GrlA to activate

expression from the LEE1 promoter. However, this factor has not been identified, and there is no proof that the factor acts via GrlR. In fact, in this work, I show that attachment plays a key role, and, surprisingly, its action is signalled via free GrlA.

EHEC and EPEC utilize a number of attachment factors that facilitate their initial and intimate attachment to host cells. First, there are the initial stages of attachment of EHEC and EPEC prior to the intimate intimin-mediated attachment (Nicholls *et al.*, 2000). The initial pre-intimate attachment stage was found to be caused by the BFP (bundle forming pili), which divide EPEC into two strains, typical EPEC containing BFP, and atypical EPEC, lacking BFP (Giron *et al.*, 1991), which is encoded by *bfp* on the EAF (EPEC adherence factor) plasmid (Baldini *et al.*, 1983). It is also believed that BFP are involved in causing EPEC to make micro-colonies, and that they aid in defying the host's immune responses. Additionally, BFP play a role in the localized adherence feature of EPEC (Cleary *et al.*, 2004). In EHEC, the long polar fimbriae are found to be important for adhesion. Hence, a deletion in the *lpf1* and *lpf2* genes stopped the adherence of EHEC to Hep-2 cells (Fitzhenry *et al.*, 2006). Additionally, EHEC attach to the Peyer's patch FAE and form A/E lesions (Phillips *et al.*, 2000). It was concluded that mutations in *lpf* resulted in EHEC adhesion to different places than the usual site of infection, i.e. the large intestine (Fitzhenry *et al.*, 2006). EHEC and EPEC use intimin, which is an outer membrane protein expressed from the *eae* gene in the *LEE5* operon (Luo *et al.*, 2000), to adhere intimately to the host cells. The role of intimin in tight attachment to host cells has been studied, and it was shown that EHEC with a mutation in the gene encoding intimin adhered diffusely and failed to form micro-colonies and pedestals, as in the wild type strain on Hep-2 cells (Cookson and Woodward, 2002).

It is clear that bacterial attachment is important for the formation of micro-colonies and pedestals, leading to the questions of whether attachment could have an influence on *LEE* gene expression, by affecting expression of the *LEE* operons either positively or negatively, and whether attachment was crucial for the expression from the *LEE* operons. To discover whether attachments to host cells influences expression from the LEE, Caco-2 cell systems were introduced to mimic the interaction between the bacterial cells and the host cells. In addition, some previous data (Islam *et al.,* 2011) were revisited and compared, but in a different host strain.

3.2 Activation of expression from the LEE1 promoter by GrlA in planktonic culture

Expression from the *LEE* operons is believed to be tightly regulated and it involves a number of factors and regulators that are either *LEE*-encoded or non-*LEE*-encoded. Among the regulators, Ler is the most important as it up-regulates the expression from the *LEE* operons but it down-regulates its own expression, i.e. Ler down regulates expression from the LEE1 operon. *LEE1* operon regulation is studied here by looking at the regulation of its promoter, LEE1.

In a previous study, the impact of GrlA's regulatory role on the LEE1 promoter was studied. It was found that GrlA played a positive role in the regulation of each promoter fragment of the LEE1 promoter (Islam *et al.,* 2011). In addition, it was found that the predicted DNA target site for GrlA-binding at P1 was located in the spacer region between the promoter -10 and -35 elements, and GrlA compensates for the non-optimal P1 spacer, permitting the RNAP

enzyme to bind so that transcription can be initiated. However, the host strain that was used in this study was *E. coli* K-12 strain. It was therefore decided that it should be retested in an EHEC-strain host to imitate the actual scenario of infection in the gut. Hence, Sakai wild type (*Δsxt* 1 and 2) EHEC strain was tested in this study. In addition to testing the impact of GrlA on the LEE1 promoter, the effects of the GrlR regulator alone, and GrlR together with GrlA (GrlR+A) were tested as well.

The starting LEE10-568 fragment carries the complete LEE1 promoter, containing both P1 and P2 promoters, is 572 base pairs in size (Fig 3.1), and was selected to study regulation by GrlA. To study the effects of GrlA and GrlR at the LEE1 promoter, measurements were made in a "wild type" Sakai strain, and in *Δgrl*A and *ΔgrlR* mutant derivative strains. Each strain was transformed with pRW224/LEE10-568 and the β-galactosidase assay was performed. Fig 3.2 illustrates the level of expression from the LEE10-568 promoter in the different strains. The results imply that, the promoter is quite active in the absence of GrlA. The presences of GrlA causes a ~ 40 % increase in the promoter's activity compared to *Sakai (ΔgrlA)* data. So, GrlA is responsible for the increase in activity and not wholly responsible for activity. The level of activation was restored to a similar level to the wild type, when *grlR* was deleted. Shown in the figure as well proter fragments LEE20-203 and LEE20-203 99T used later in the chapter.

LEE10-568 (549 bp)

EcoRI -568

GAATTCTGCACCCGTTCCAGGTTAGTGCTGGCTGTAGCTTATGTCCGGGAAACAGCTAATAGATATATATACTCGTCATACTTCAAGTTGCATGTGC

TGCGACTGCGTTCGCTTACCCCAATCACTTACTTATGTAAGCTCCTGGGGATTCACTCGCTTGCCGCCTTCCTGTAACTCGAATTAAGTAGAGTATA

GTGAAACGGTTCAGCTTGGTTTTTATTCTGTTTTATTTGTTTATGCAATGAGATCTATCTTATAAAGAGAAACGCTTAACTAAATGGAAATGCAATT

-100 -90 -80 P1 UP elements -35 element 18 bp

ATTAAAGTCGTTTGTTAACGAGATGATTTTCTTCTATATCATTGATTTTAAATGGATTTTAAAAATATATGATTTTTTTGTTGACATTTAATGATAA

-10 element

TGTATTTTACACATTAGAAAAAAGAGAATAATAACATTTTAA**GGTGG**TTGTTTGATGAAATAGATGTGTCCTAATTTGATAGATAAACGTTATCTCA

P2

CATAATTTATATCATTTGATTAATTGTTGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGC**AAGCTT**

HindIII

LEE20-203 (46 bp)

GAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

-35 element -10 element

LEE20-203 99T(46 bp)

GAATTCTTGACATTTAATGATAATGTATTTTA**T**ACATTAGAAAAAAGAGAAT**AAGCTT**

Fig. 3.1: DNA sequences of LEE10-568, LEE20-203 and LEE20-203 99T promoters.
The figure shows the DNA sequences of LEE10-568, LEE20-203, and LEE20-203 99T promoter fragments of *LEE*1 regulatory region. EcoRI and HindIII sites are in bold, -10 and -35 elements are underlined for P1. -10 and -35 elements for P2 are in grey shades. Predicted GrlA binding site is heighted in yellow.

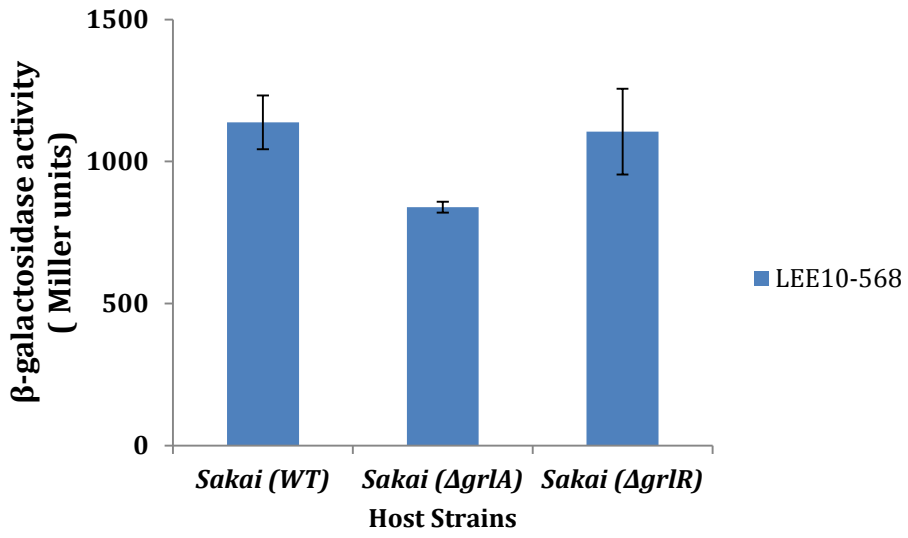


Fig. 3.2: Activation of expression from the LEE10-568 promoter fragment in different Sakai background strains.

The bar chart shows the measured β-galactosidase expression in Sakai wild type, Sakai *ΔgrlA* and Sakai *ΔgrlR* strains, each containing pRW224 carrying a LEE10-568::*lac* fusion. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

Table 3.1: Levels of expression from the LEE10-568 promoter fragment in different Sakai background strains.

	β-galactosidase activity (Miller units ± SD)	
	pRW224/LEE10-568	**pRW224/U9**
Sakai (WT)	1138	184 ± 91
Sakai (*ΔgrlA*)	839	218 ± 8
Sakai (*ΔgrlR*)	1105	171 ± 17

[2]

[2] The table shows the measured β-galactosidase activity in Sakai wild type, Sakai (*ΔgrlA*), and Sakai (*ΔgrlR*) strains each containing empty pRW224/LEE10-568 or pRW224/U9. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~ 0.5 at 650 nm.

Thus, in planktonic culture in the Sakai host strain, the level of expression from the LEE10-568 promoter was modest. This level was not enhanced when the negative regulator gene *grlR* was deleted, despite the presence of chromosomal GrlA. To control the experiment in Fig 3.2, the Sakai strains were also transformed with a derivative of pRW224 (pRW224/U9; Fig 2.1) containing a linker stuffer sequence rather than a promoter fragment, and the β-galactosidase assays were repeated. The data (Table 3.1) show that expression levels were low, and it can therefore be concluded that measured activities resulting from the LEE10-568 fragment were significant.

Because disruption of *grlA* and *grlR* in the Sakai strain had little impact on the level of expression from the LEE1 promoter, *grlA*, *grlR* and *grlA + grlR* were introduced on a multi copy number plasmid pACYC184. To do this, the different genes were cloned into the pACYC184 vector (Fig. 2.3). Then, the Sakai strain was transformed with pRW224/LEE10-568 with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. The transformants were plated onto MacConkey agar containing tetracycline and chloramphenicol antibiotics. They were left in the incubator at 37 °C overnight. On the following day, three separate colonies were picked from each plate and inoculated into 5 ml LB media in 25 ml flasks containing tetracycline and chloramphenicol. The flasks were then left on a shaker at 37 °C overnight. Finally, on the following day, β-galactosidase assays were performed, as described in materials and methods (section 2.11). The expression from the LEE10-568 promoter was then assayed with the regulators in (Fig. 3.3). It appears that the promoter was activated ~ 1.5-fold by the presence of the overexpressed free GrlA in the planktonic culture. The level of activation was reduced when GrlR or GrlR+A were tested. This data showed that even when the GrlA was over-expressed, the level of activation of the expression from the LEE10-568 promoter was modest. This

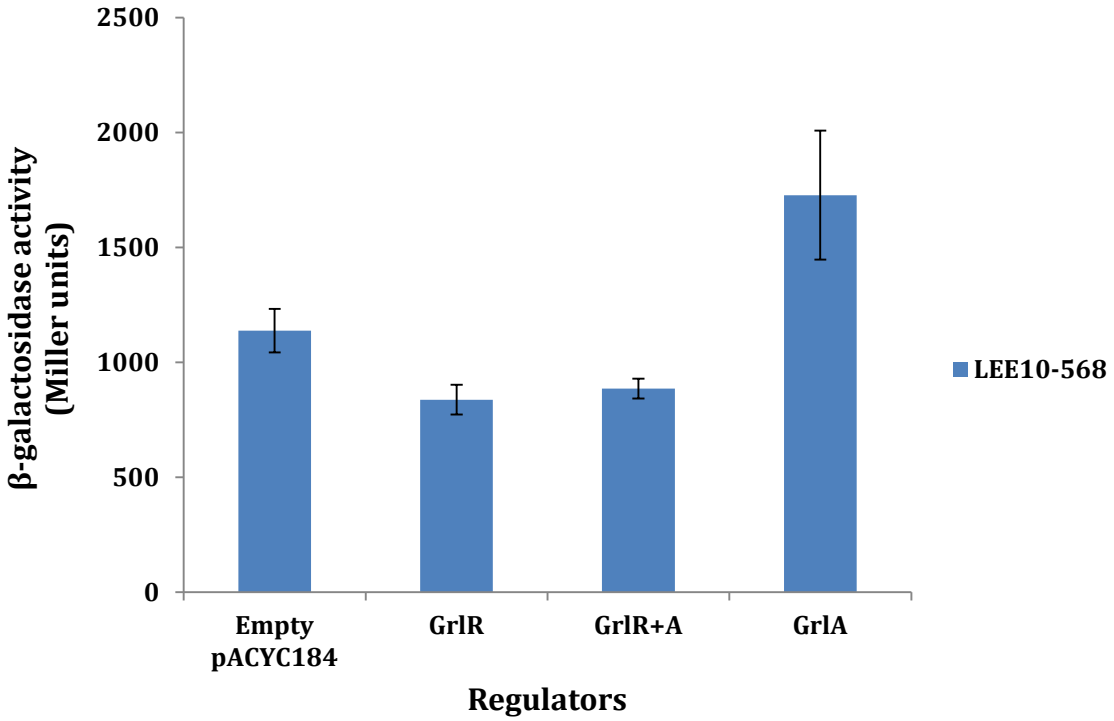


Fig. 3.3: Activation of expression from the LEE10-568 promoter fragment by GrlA.
This bar chart shows the measured β-galactosidase expression in Sakai wild type strains, each containing pRW224 carrying a LEE10-568::*lac* fusion together with empty pACYC184, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

raises the question of whether the promoter is activated by GrlA either directly, by free GrlA, or indirectly, by releasing the bound GrlA from the GrlR+A complex, or by a different mechanism.

Additionally, to control the experiment illustrated in Fig. 3.3, the Sakai strain was transformed with pRW224/U9 with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA plasmid constructs and the β-galactosidase assay was performed. The level of expression was also low, as shown in Table 3.2, compared to the level of expression with the LEE10-568 promoter fragment.

3.3 Activation of expression from the LEE1 promoter by GrlA after attachment to Caco-2 host cells

The next stage was to introduce a system in which the bacterial cultures came into contact with eukaryotic host cells. To do this Caco-2 cells were chosen as an experimental system. Caco-2 cells originate from the human epithelial cells of the gut lining, originally coming from a colon adenocarcinoma. These cells form a monolayer, after spontaneous differentiation in culture, after which they express several morphological and functional features of the mature monocyte (Sambuy *et al.,* 2005).

To begin, in order to investigate the activation of the promoter of the *LEE1* transcription unit, Sakai strains were transformed with pRW224/LEE10-568, and either pACYC184/ΔHN, or

Table 3.2: Effect of GrlA and GrlR on activity of the LEE1 promoter in EHEC.

	β-galactosidase activity (Miller units ± SD)	
	pRW224/LEE10-568	**pRW224/U9**
pACYC184/ΔHN	1138	184 ± 91
pACYC184/GrlR	837	235 ± 41
pACYC184/GrlR+A	886	196 ± 37
pACYC184/GrlA	1727	198 ± 32

[3]

[3] The table shows the measured β-galactosidase activity in Sakai strains each containing empty pRW224/U9 with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~ 0.5 at 650 nm.

pACYC184/GrlA. Then, overnight cultures from each strain were sub-cultured into new LB media with chloramphenicol and tetracycline antibiotics, and grown to the mid exponential phase OD_{650} ~ 0.5. 1 ml from each culture was kept aside and lysed for the β-galactosidase assays and labelled as "undiluted culture". The rest of each culture was then diluted into DMEM media and incubated with Caco-2 cells in tissue culture plates for three hours at 37°C**.** After this incubation, Caco-2 cells infected with Sakai strains then 1 ml from each strain was collected and labelled "supernatant". These free cells were lysed and assayed for β-galactosidase expression, and these measurements reflect promoter activity in the free-swimming strains that had not attached to Caco-2 cells. Cells were then washed with 1 ml of PBS.

To release the attached Sakai bacteria from the Caco-2 cells, 0.1% Triton-X100 was added so that epithelial cells ruptured, and the attached Sakai cells were released. Released cells, labelled "attached", were collected in a tube and lysed for β-galactosidase assays. Fig 3.4 illustrates data showing how GrlA activates the LEE10-568 promoter in the attached cells. In addition, the figure shows data for the "undiluted" and "supernatant" cells, in order to compare expression with non-attached cells.

Alongside β-galactosidase activity measurements, cells from the "undiluted" (planktonic) sample, the "supernatant", and the "attached" fractions were plated onto LB agar (after dilution) for colony counting. The figure shows β-galactosidase activities, normalized to cell number. This was achieved by including the colony count numbers to work out the initial OD 650 of the supernatant and the equivalent value for the attached stages. It can be inferred that,

without GrlA, the basal level of LEE10-568 activity was enhanced ~ 1.6-fold in the attachment stage compared to the undiluted stage.

However, with GrlA, activation of expression was enhanced ~ 20-fold in the attachment stage compared to the undiluted stage. The results in Fig. 3.4 demonstrate clear and significant contact-dependent, GrlA-dependent activation at the *LEE1* regulatory region. Next, whether this activation is down regulated by GrlR was questioned. In order to address this, the pACYC184-derivative-carrying GrlR+A was used. Thus, Sakai strains were transformed with pRW224/LEE10-568 and either pACYC184/ΔHN, or pACYC184/GrlR+A+. Fig 3.5 shows the measured normalized β-galactosidase assay data. Overexpression of the GrlR+A complex resulted in a slight reduction in promoter activity in the undiluted stage (planktonic culture). However, the level of expression from LEE10-568 is enhanced by GrlR+A in the attachment stage compared to the undiluted stage by a factor of ~ 17. Presumably, this is accounted for under these circumstances by the excess of GrlA in comparison to GrlR. In the attachment stage, host factor(s) must be acting on free GrlA and possibly destabilizing the GrlR+A complex, thereby freeing some GrlA molecules that could then bind to the promoter elements and recruit RNAP to initiate transcription. Taken together, the results in Fig. 3.4 and Fig. 3.5 argue that attachment of Sakai strain cells to the Caco-2 host cells triggers direct activation of GrlA (maybe by toggling it from an inactive into an active state.

3.4 Overexpression of GrlR results in reduced promoter activity

Next, the effect of overexpression of GrlR alone on the basal level of expression from the LEE1 promoter was tested. To do this, the pACYC184-derivative-carrying GrlR was used.

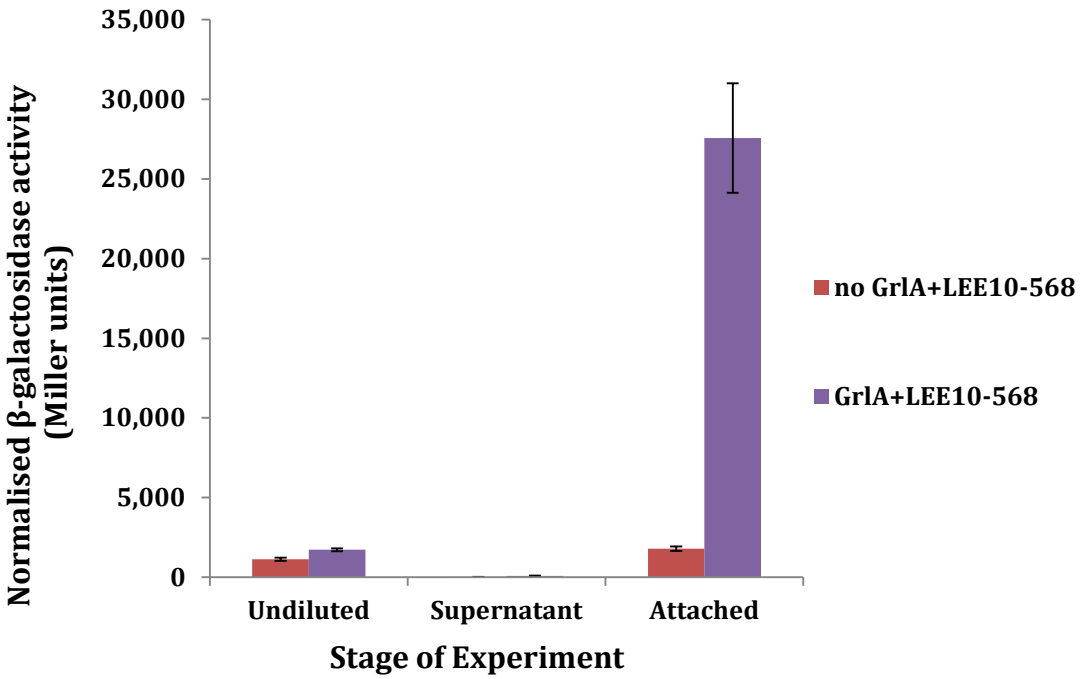


Fig. 3.4: Activation of expression from the LEE10-568 promoter fragment by GrlA after attachment to Caco-2 cells.

The bar chart shows the measured normalized β-galactosidase expression in Sakai wild type cells each containing pRW224 carrying LEE10-568::*lac* fusion, or empty pRW224 together with pACYC184/GrlA, or empty pACYC184. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact (attached stage). Error bars represent the standard errors of three independent repeats. No GrlA: not cloned on pACYC184 vector, but is expressed from the chromosome.

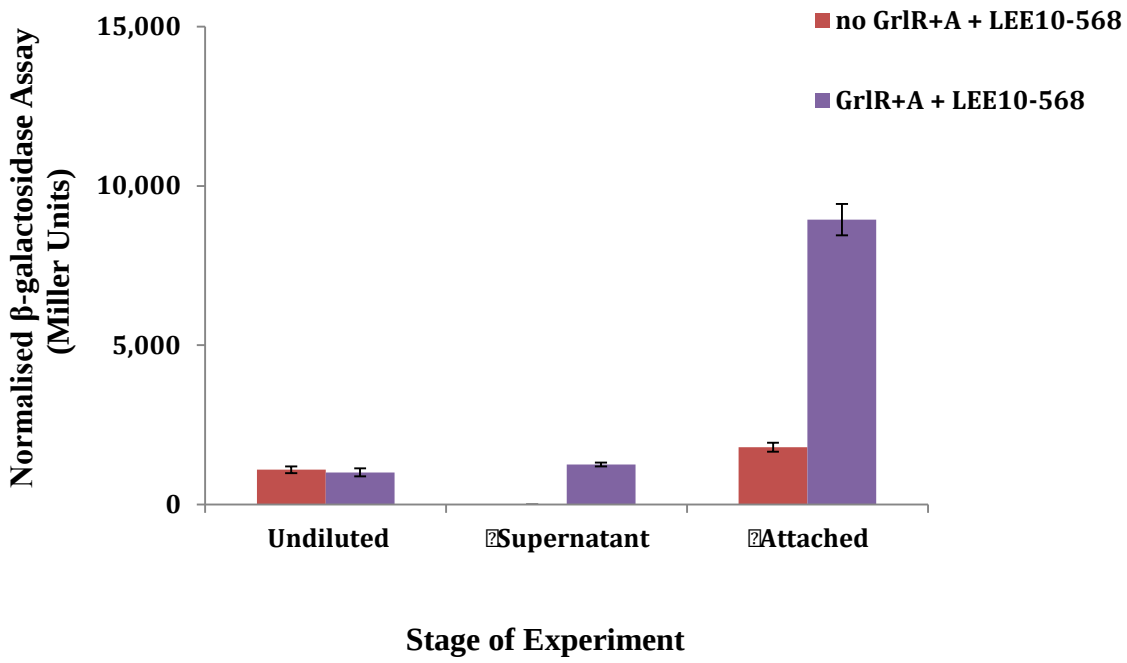


Fig. 3.5: Activation of expression from the LEE10-568 promoter fragment by GrlR+A after attachment to Caco-2 cells.

The bar chart shows the measured normalized β-galactosidase expression in Sakai wild type strains each containing pRW224, carrying an LEE10-568::*lac* fusion or empty pRW224 together with pACYC184/GrlR+A or empty pACYC184. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact (attached stage). Error bars represent the standard errors of three independent repeats. No GrlR+A: not cloned on pACYC184 vector, but is expressed from the chromosome.

Sakai strains were transformed with pRW224/LEE10-568 with pACYC184/ΔHN, or pACYC184/GrlR. Fig. 3.6 illustrates normalized β-galactosidase levels measured in this experiment. It can be seen that the range of activities in these assays was lower than the range in Fig. 3.4, where excess GrlA was present. Overexpression of GrlR in the planktonic culture decreased the level of expression from the LEE10-568 promoter. This likely means that, when GrlR was not present, the background chromosome-encoded GrlA was activating expression from the LEE10-568 promoter.

However, when GrlR was overexpressed, it suppressed GrlA activity, resulting in a reduced promoter activity, and it can be said to have reduced it to the basal level. Moreover, in the attachment stage, expression from the LEE10-568 promoter was almost completely suppressed when GrlR was overexpressed. Again it is thought that GrlR suppresses activation of the promoter by GrlA, but in this case something else could have happened that caused this very low promoter activity of about ~ 15 times less in the attachment stage compared to the undiluted stage.

3.5 Control experiments and overview

To control the experiments in Fig. 3.4, Fig. 3.5, and Fig. 3.6, Sakai strains were transformed with pRW224/U9, together with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A +A, or pACYC184/GrlA. The collected data are shown together in Table 3.3. The measured levels of expression were low compared to the expression level when LEE10-568 promoter

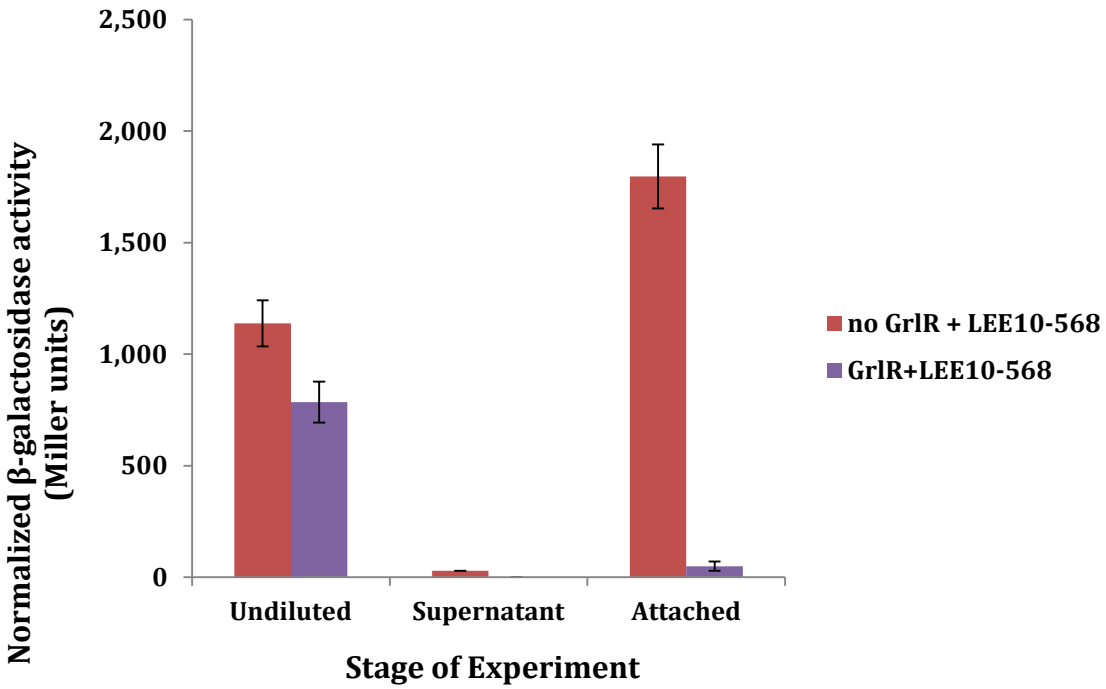


Fig. 3.6: GrlR-dependent effects on expression from the LEE10-568 promoter fragment after attachment to Caco-2 cells.

The bar chart shows the measured normalized β-galactosidase expression in Sakai wild type strains each containing pRW224, carrying an LEE10-568::*lac* fusion or empty pRW224 with pACYC184/GrlR, or empty pACYC184. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact (attached stage). Error bars represent the standard errors of three independent repeats. No GrlR: not cloned on pACYC184 vector, but is expressed form the chromosome.

was tested (data recapped in Table 3.4). Note that all the data is illustrated together on the same scale in Fig. 3.7. From the data, it looks like host factor(s) trigger activation of GrlA, possibly in two ways. In the first, the trigger may bind directly to the free, unbound GrlA, which will then bind to the promoter elements to recruit RNAP to initiate transcription at the LEE1 promoter. In the other, the trigger may bind to bound GrlA in the GrlR+A complex, thus making GrlR release bound GrlA, and this, again, may result in activation of transcription initiation. Finally, GrlA function may be triggered via an alternative route, which might involve other factors before GrlA. Fig. 3.7 shows combined data concerning expression from the LEE10-568 regulatory region, comparing all four situations: pACYC184/ΔHN, GrlR, GrlR+A, and GrlA. It is clear that expression from LEE10-568 was enhanced in the attachment stage in all cases, except in the GrlR case, where the level of expression was reduced. This expression was significantly enhanced when GrlA was present in the attachment stage. GrlR+A also enhanced expression in the attachment stage, but less than GrlA. In the supernatant, which represents an internal control, the level of expression was low compared to the other two stages. This stage was included in the assay to confirm that the enhanced level of expression was dependent on the Sakai cells being attached to the Caco-2 cells. Where there were some free-swimming, unattached Sakai strains, the level of expression was even lower than in the undiluted stage.

To exclude the possibility that the activation of expression of LEE10-568 seen in the attachment stage was artificially caused by the presence of Triton X-100, β-galactosidase assays were performed with Sakai cells transformed with the plasmid constructs mentioned in Section 3.2. Subcultures from each strain were grown to the exponential phase growth, then 1 ml of PBS with 0.1% Triton X-100 was added to each culture, and the assay was performed.

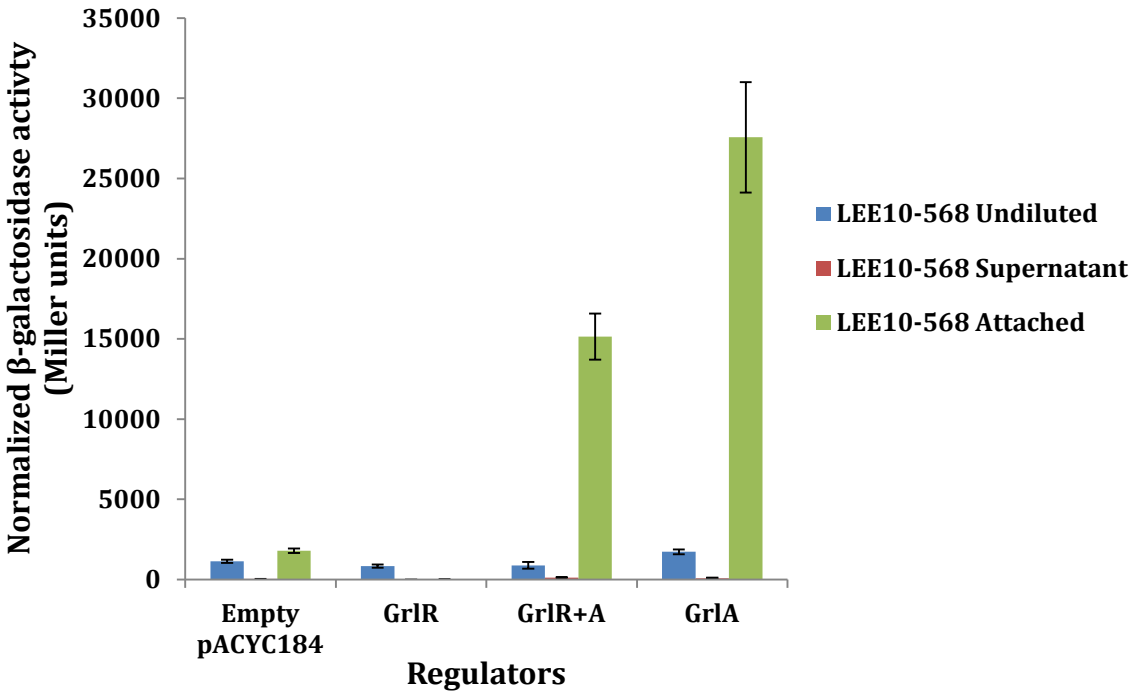


Fig. 3.7: Activation of expression from the LEE10-568 promoter fragment by GrlR, GrlR+A, and GrlA after attachment to Caco-2 cells.

This bar chart shows merged measured normalized β-galactosidase expression in Sakai wild type strains, each containing pRW224 carrying an LEE10-568::*lac* fusion or empty pRW224 together with pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact. Error bars represent the standard errors of three independent repeats.

Table 3.3: Basal level of expression after attachment to Caco-2 cells with GrlR, GrlR+A, and GrlA constructs.

Normalized β-galactosidase activities (Miller units ± SD)			
Plasmids	**Undiluted**	**Supernatant**	**Attached**
PACYC184/ΔHN + pRW224/ U9	104 ± 15	0 ± 0	2 ± 3
pACYC184/GrlR + pRW224/ U9	93 ± 19	0 ± 0	0 ± 0
pACYC184/GrlR+A + pRW224/U9	106 ± 18	8 ± 2	35 ± 8
pACYC184/GrlA + pRW224/U9	90 ± 3	0 ± 0	0 ± 0

[4]

[4] The table shows the measured normalised β-galactosidase activity in Sakai wild type strains each containing the empty pRW224/U9 with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~ 0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact (attached stage). Error bars represent the standard errors of three independents repeats.

Table 3.4. Activation of expression from the LEE10-568 promoter fragment after attachment to Caco-2 cells with GrlR, GrlR+A, and GrlA constructs.

Normalized β-galactosidase activities (Miller units ± SD)			
Plasmids	**Undiluted**	**Supernatant**	**Attached**
PACYC184/ΔHN + pRW224/LEE10-568	1138 ± 15	30 ± 0	1796 ± 204
pACYC184/GrlR + pRW224/LEE10-568	785.3 ± 131	0 ± 0	50 ± 36
pACYC184/GrlR+A + pRW224/LEE10-568	886 ± 217	125.6 ± 10	15134.62 ± 1435
pACYC184/GrlA + pRW224/LEE10-568	1727 ± 150	85 ± 43	27567.57 ± 4861

[5]

[5] This table shows the measured normalised β-galactosidase activity in Sakai wild type strains, each containing the empty pRW224/U9 with pACYC184/ΔHN, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~ 0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours contact (attached stage). Error bars represent the standard errors of three independent repeats.

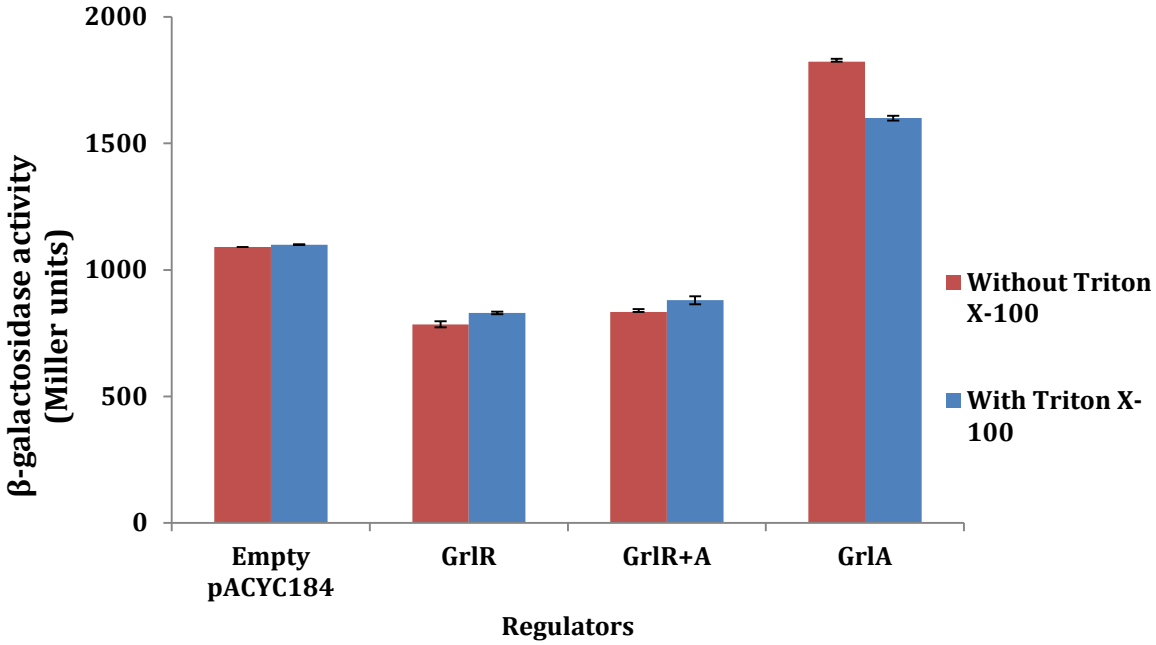


Fig. 3.8: Activation of expression from the LEE10-568 promoter fragment by GrlA with and without Triton X-100.

This bar chart shows the measured β-galactosidase expression with Triton X-100 in Sakai wild type strains each containing pRW224, carrying an LEE10-568::*lac* fusion together with empty pACYC184, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

Data in Fig. 3.8 show that the presence of the Triton in the assay reaction did not influence the promoter activity, and the assays gave similar data to those illustrated in Fig. 3.3. Therefore, the enhanced promoter activity seen was due to attachment to the host cells not to the presence of chemicals in the reaction.

3.6 Expression from the LEE1 promoter in single cells

It was clear that maximum activation of expression from the LEE1 promoter by GrlA was achieved when host cells were included. To look at the expression from individual cells, the LEE10-586 promoter fragment was fused to *gfp* to make LEE10-568::*gfp* fusion. To do this, the *lac* genes in pRW224 downstream of the cloning sites were removed and replaced with a *gfp* gene to make pRW400 (Fig. 2.5).

Expression was visualized by the production of green fluorescence under the fluorescent microscope. Sakai strains were transformed with pRW400/LEE10-568 with pACYC184/ΔHN or with pACYC184/GrlA. Overnight cultures were incubated with HeLa cells for three hours then stained for microscopy after washing non-attached cells with PBS. Note that, for these experiments, HeLa cells were found to be better than Caco-2 cells.

The stains used were Hoechst, which stained the nucleic acids in both the HeLa cells and in the bacterial cells, and a preparation of FITC-labelled derivative of phalloidin, a fungal toxin that binds to actin filaments only. The FITC-phalloidin was used to

visualize the impact of Sakai infection on host cell cytoskeleton rearrangement and pedestal formation.

Fig. 3.9 represents images taken from various channels on a microscope to visualise different features by highlighting different fluorescent markers.

Fig. 3.9Ai illustrates the Hoechst fluorescence staining of nucleic acids of both Sakai and HeLa cells, shown in Ai and Bi. In Ai, the numbers of Sakai cells are higher than in Bi, this is because in Ai, GrlA was present. So, when GrlA was supplied on a multi-copy plasmid, the number of bacteria attached to the surface of HeLa cells was more. In the next picture, Aii and Bii, shown is the GFP expression, and as can be seen GFP was expressed when GrlA was present in Aii. However, in the bottom panel in Bii, GFP was not expressed.

Next the effect of the Sakai strain on the host cell actin filaments (Fig 3.9Aiii and Biii) was examined. This channel transmits fluorescence to show phalloidin-bound actin elements in the HeLa cells. In the top photo (+GrlA), the Sakai strain managed to cause rearrangement of the actin filaments and produce pedestals to a greater extent than is visible in the bottom photo (-GrlA). This shows that when GrlA is present, more pedestals are formed, which is logical because there are already more attached bacterial cells, so eventually more pedestals would be caused. To check the system, Hoechst stain was used to visualise both HeLa cells and Sakai cells (Fig. 3. 9Ai, Bi). Here, the nucleic acids of both HeLa and Sakai cells show as blue, confirming real

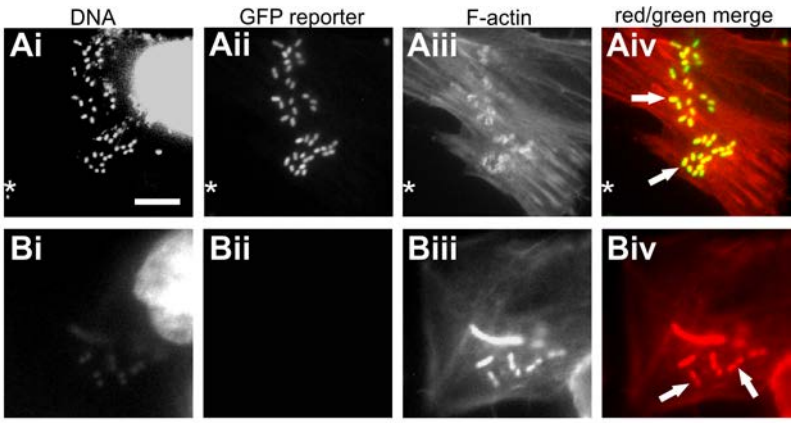


Fig. 3.9: Expression from the LEE10-568 promoter fragment in single cells.

This figure shows the GFP expression in Sakai wild type strains, each containing pRW400 carrying a LEE10-568::*gfp* fusion together with pACYC184/GrlA, or pACYC184/ΔHN. Fluorescence microscopy was performed after fixing a Sakai culture to HeLa cells and staining with phalloidin (red), which stained the actin filaments and showed pedestal formation, and Hoechst (blue) that stained the nucleic acids in both Sakai cells and HeLa cells. The ultraviolet channel (UV) transmitted at UV wavelength to show the blue fluorescence of the nucleic acid (Ai, and Bi). The blue channel transmitted at blue wavelength to show green fluorescence when GFP was expressed (Aii, and Bii). The green channel transmitted a green wavelength to show the red fluorescence of the actin filaments and pedestals (Aiii and Biii). Green+red, a merge of the Green and the Blue channels, showed whether glowing cells were causing pedestal formation (Aiv and Biv). The top panel shows GFP expression from LEE10-568 with GrlA, and the bottom panel shows GFP expression from LEE10-568 without GrlA.

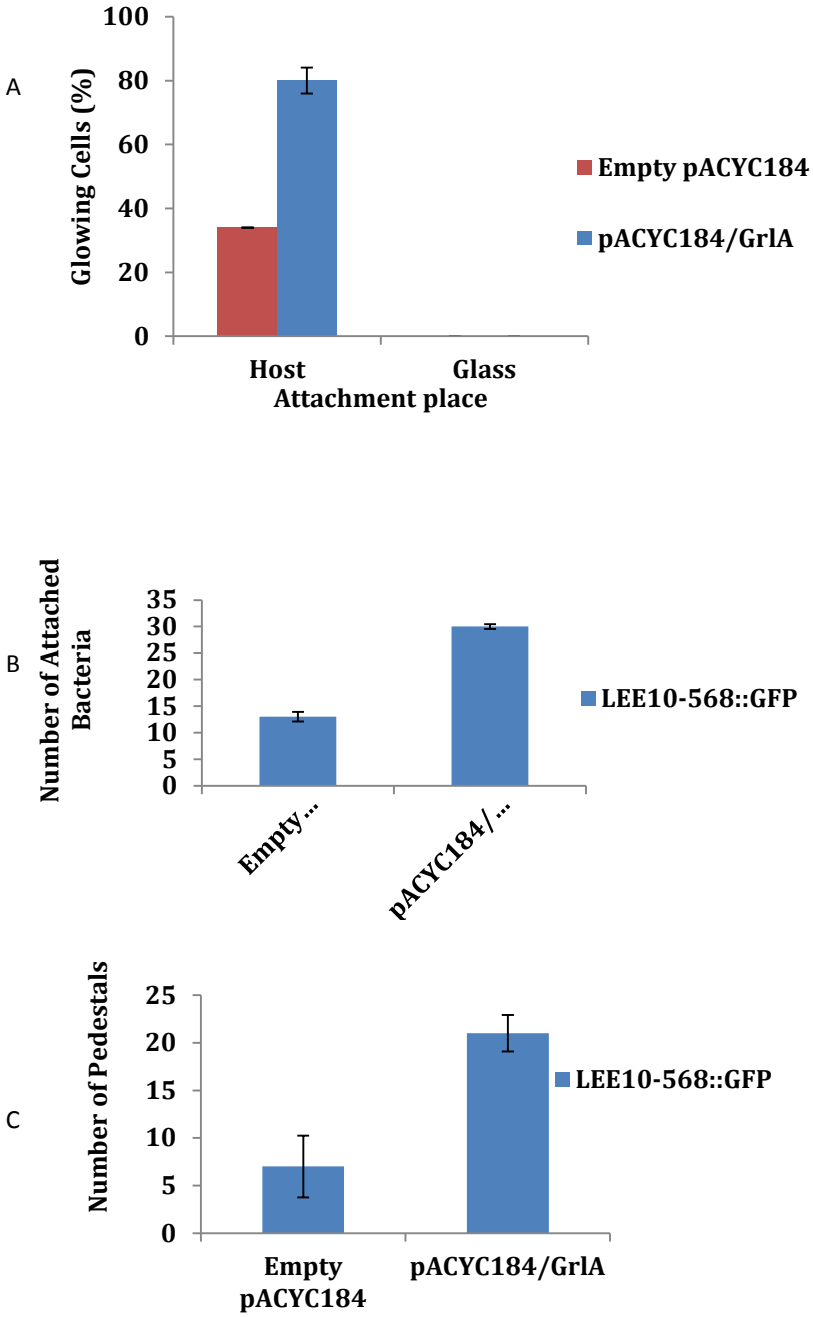


Fig. 3.10: Quantitative analysis of attached Sakai to HeLa cells.

These bar charts show the statistical analysis of the results shown in Fig. 3.9. In A, the number of cells attached to both host cells and the glass was counted and the percentage of glowing cells is indicated. In B, the number of bacteria attached to the host cells, with and without GrlA, was counted and is indicated. In C, the number of pedestals per cell was counted and is indicated. Error bars represent the standard errors of three independent repeats.

interactions between the pathogen and the host cells. It was seen that GrlA caused more attachment to HeLa cells, as well as more pedestal formation.

The next experiment, illustrated in Fig. 3.9 Aii, Bii, visualises GrlA-dependent expression from the LEE10-568 promoter fragment by detecting green fluorescence. It is clear that GrlA had a positive effect on the expression from the LEE10-568 promoter fragment as the GFP was expressed (upper panel), but it was not expressed when GrlA was absent (lower panel). Note that the exposure time for both pictures with GrlA or without GrlA was the same, 400 ms, so the images could be compared. The results are consistent with the ensemble β-galactosidase assays (Table 3.4). Also, consistent with the β-galactosidase assays, expression from the LEE10-568 promoter fragment was enhanced after attachment even when GrlA was not present (Fig. 3.9Bii, lower panel).

The fluorescence microscopy experiment permitted to ask whether GFP was expressed in bound or free bacteria (or both). To address this, an overlay was performed of the images in Fig. 3.9Aii (GFP channel) and Fig. 3.9Aiii (F-actin). The overlaid image (Fig. 3.9Aiv) shows that Sakai cells attached to HeLa cells are green, but Sakai cells attached to the glass slide were not fluorescent, even when GrlA was present (upper overlay). This was not seen when GrlA was not present in the bottom overlay picture. Hence, only Sakai cells attached to HeLa cells expressed GFP. This tells us that the system is turned on by attachment to specific substrates but not by random weak attachment. Moreover, from Fig. 3.9Aiv, pedestals were seen by the Sakai strain expressing GFP, the images of phalloidin location (Fig. 3.9Aiv) and GFP

expression (Fig. 3.9Aii) were overlaid. The resulting image (Fig. 3.9iv) shows that only Sakai cells in which GFP is overexpressed are able to cause the rearrangement of the actin filaments and produce pedestals.

In addition to looking at the impact of GrlA on attachment, pedestals, and GFP expression, quantitative analysis was also performed at each stage. Data in Fig. 3.10A show that the percentage of glowing cells, when bacterial cells were attached to HeLa cells, was 80% when GrIA was present, compared to 34% glowing cells when GrlA was absent, but 0% of the bacterial cells glowed when they were attached to the glass slide when GrlA was either present or absent. Numbers of bacteria attached to HeLa cells were counted. The number of attached bacteria was ~ 2.3-fold higher when GrlA was present (Fig. 3.10B). In addition to counting the number of attached bacteria, the number of pedestals was counted as well. When GrlA was present, there was a ~ 3-fold increase in pedestal formation.

3.7 Expression from the constitutive promoter in single cells

In addition to measuring the expression of GFP controlled by the LEE10-568 fragment, the LEE20-203 99T promoter fragment, which contains a point mutation in the -10 element, was assayed. Note that 99T is a single base change that results in the LEE1 promoter being independent of GrlA (Islam *et al*., 2011). Fig. 3.11 illustrates the experiment, in which empty pRW400 was used as a negative control, and, to be assured that observed effects were due to the Sakai strain, uninfected HeLa cells were

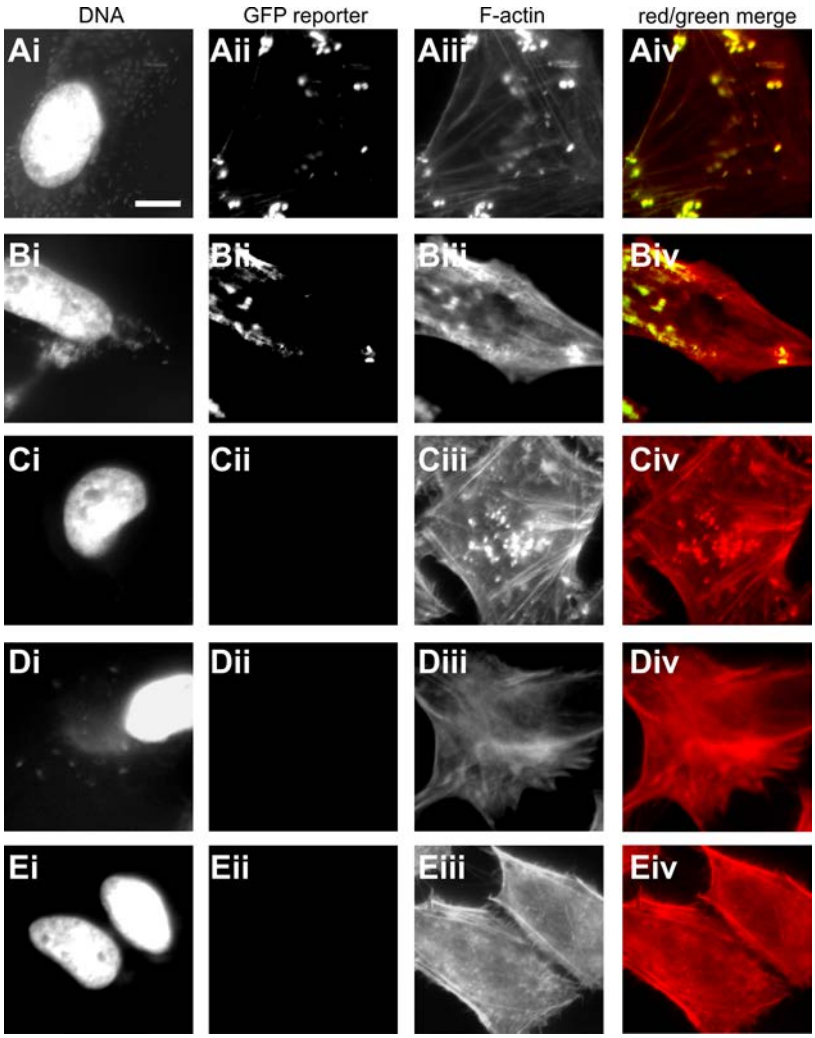


Fig. 3.11: Expression of LEE20-203 99T promoter fragment in single cells.
This figure shows the GFP expression in Sakai wild type strains, each containing pRW400 carrying an LEE20-203 99T::*gfp* fusion, and pRW400 together with pACYC184/GrlA or pACYC184/ΔHN, and an uninfected cells**.** Fluorescence microscopy was done after fixing Sakai-strain cultures to HeLa cells and staining them with phalloidin (red), which stained the actin filaments and showed pedestal formation, and Hoechst (blue) that stained the nucleic acids in both Sakai cells and HeLa cells. The ultraviolet channel (UV) transmitted at UV wavelength to show the blue fluorescence of the nucleic acid (Ai, Bi, Ci, Di, and Ei). The blue channel transmitted a blue wavelength to show green fluorescence when GFP was expressed (Aii, Bii, Cii, Dii, and Eii). The Green channel transmitted a green wavelength to show the red fluorescence of the actin filaments and pedestals formation (Aiii, Biii, Ciii, Diii, and Eiii). Green+Red, a merge of the green and red fluorescence, showed whether glowing cells caused pedestal formation (Aiv, Biv, Civ, Div, and Eiv). (Scale bar, 10 μm)

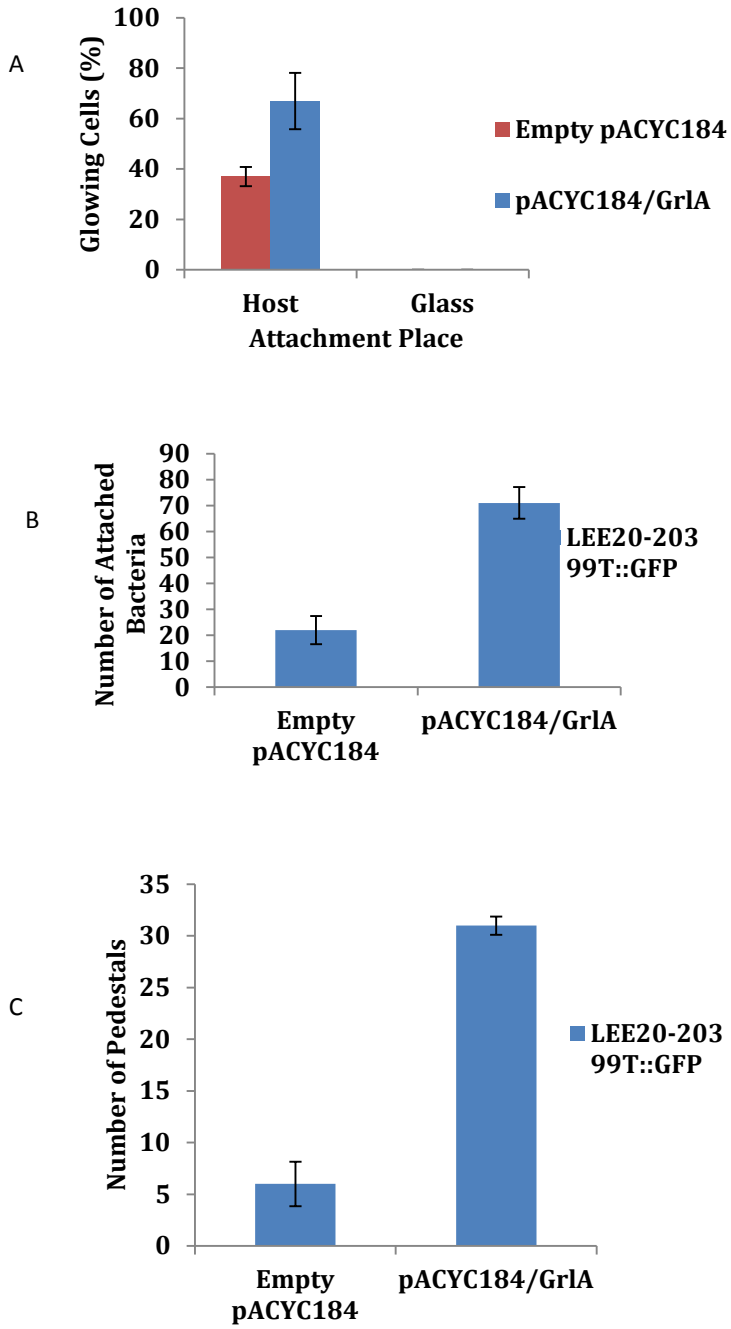


Fig. 3.12: Quantitative analysis of attached Sakai to HeLa cells.
These bar charts show statistical analysis of the data presented in Fig. 3.11.A. In A, the number of cells attached to both host cells and the glass was counted and the percentage of glowing cells is indicated. In B, the number of bacteria attached to the host cells, with and without GrlA, was counted and is indicated. In C, the number of pedestals per cell was counted and is indicated. Error bars represent the standard errors of three independent repeats.

also included. It is clear that the Sakai strain with the mutant promoter was able to attach to HeLa cells whether GrlA was present or absent (Fig. 3.11Aii and Bii), and the attached Sakai strain produced pedestals (Fig. 3.11Aiii and Biii). These characteristics were caused by true bacterial-host interaction, as Fig. 3.11Ai and Bi indicate, showing the nucleic acids of both Sakai and HeLa cells. Expression from the LEE20-203 99T promoter fragment was seen as GFP expression (Fig 3.11Aii) and an overlay of Fig. 3.11Aii and Fig. 3.11Aiii showed that bacteria attached to HeLa cells appeared green (Fig. 3.11Aiv). Additionally, the cells expressing GFP (Fig. 3.11Aiii) also showed pedestal formation (Fig. 3.11Aiv). Note that when GrlA was absent, the pattern of green fluorescence was similar to when GrlA was present, but the intensity of fluorescence was lower.

In the control Sakai strains containing 'empty' pRW400, GrlA-induced pedestals still formed, likely because overexpression of GrlA induces LEE expression. Finally, no pathogen-associated features were seen in HeLa cells when they were not infected with the Sakai strain.

Quantitative analysis was also performed on the cells carrying pRW400 with the constitutive promoter (Fig. 3.12). The percentage of glowing cells when GrlA was present was ~ 1.8-fold higher than when GrlA was not present (Fig. 3.12A). This showed that, even though the LEE20-203 99T promoter exhibited constitutive activity, its expression was higher when GrlA was present. Not only that, but the number of attached bacteria was higher when GrlA was present. The number of attached bacteria when GrlA was present was ~ 3.2 times higher than the number of

attached bacteria when GrlA was absent (Fig. 3.12B). In addition, the number of pedestals formed per cell, when GrlA was present, was ~ 5 times higher than when GrlA was absent (Fig. 3.12C). In the negative control, pedestals were formed when GrlA was overexpressed, in contrast to when GrlA was absent, when fewer pedestals were formed, showing that the pedestals were GrlA-dependent.

3.8 Activation of expression from the LEE1 promoter increases attachment to HeLa cells

GrlA-dependent activation at the LEE10-568 promoter was ~ 19 times higher when bacteria were attached to host cells. The higher expression was demonstrated by both increased β-galactosidase and GFP expression. In conditions where LEE10-568 promoter activity was enhanced, it was believed that the number of bacteria attached to host cells would increase. To test this, the Sakai strain was transformed with LEE10-568/pRW224 and, either pACYC184/ΔHN, or pACYC184 encoding GrlR, GrlR+A, or GrlA. Overnight cultures of each strain were then incubated with HeLa cells. After three hours, unattached bacteria were washed away, the Hela cells were lysed, and diluted lysates were plated on LB agar for colony counting to quantify bacterial attachment.

As can be seen in Fig. 3.13, the number of bacteria attached to HeLa cells increased ~ 8.5-fold when GrlA was present compared to the colony count of attached bacterial

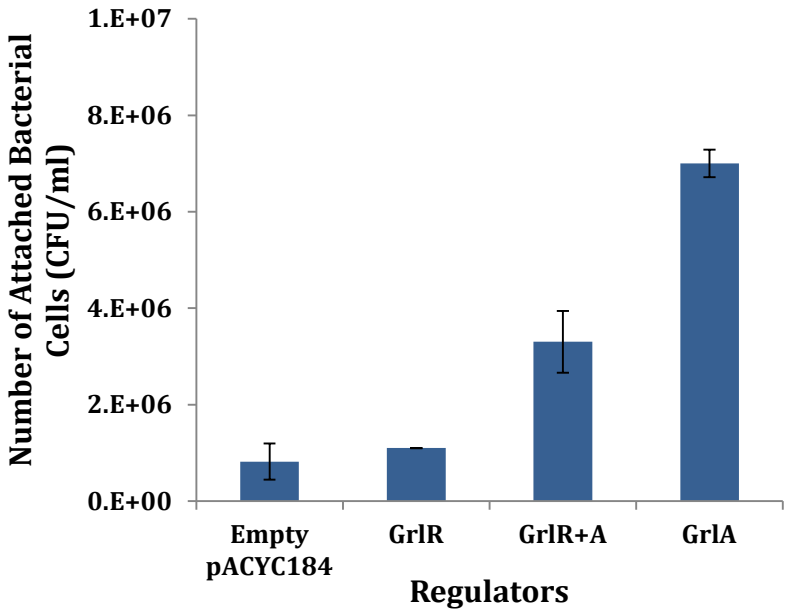


Fig. 3.13: Number of bacteria attached to HeLa cells.
This bar chart represents the counted colony-forming units of the Sakai wild type strains, each containing pRW224 carrying LEE10-568::*lac* fusion together with either empty pACYC184, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Sakai cells were in contact with HeLa cells at 37 °C for three hours. Cells were then lysed and plated into LB agar and incubated at 37 °C overnight. Error bars represent the standard errors of three independent repeats.

cells with the pACYC184/ΔHN. The number of attached colonies when GrlR was present was similar to the colony count of attached bacterial cells with the empty pACY184. The number of attached colonies when GrlR+A was present was ~ 4 times higher than the count of attached bacterial cells with the pACYC184/ΔHN. In short, when bacterial cells attach to host cells, free GrlA becomes active. Active GrlA activates the expression from the LEE1 promoter, which has a positive impact on expression from the other LEE operons, resulting in enhanced attachment to host cells (Fig. 3.13).

3.9 Activation of expression from the LEE20-203 promoter fragment by GrlA in planktonic culture and after attachment to Caco-2 cells

The LEE20-203 promoter fragment is the shortest fully functional derivative fragment from the LEE1 promoter (Fig. 3.1) and it is believed that it contains the GrlA binding site (Islam *et al.*, 2011).

Here, the experiments that had been performed with the longer LEE10-568 fragment were repeated with the short LEE20-203 fragment. The results, illustrated in Fig. 3.14, show that, in planktonic culture, expression from the LEE20-203 promoter when GrlA was present was ~ 4.4 times higher than the expression from the LEE20-203 promoter with pACYC184/ΔHN. Expression from the LEE20-203 promoter when

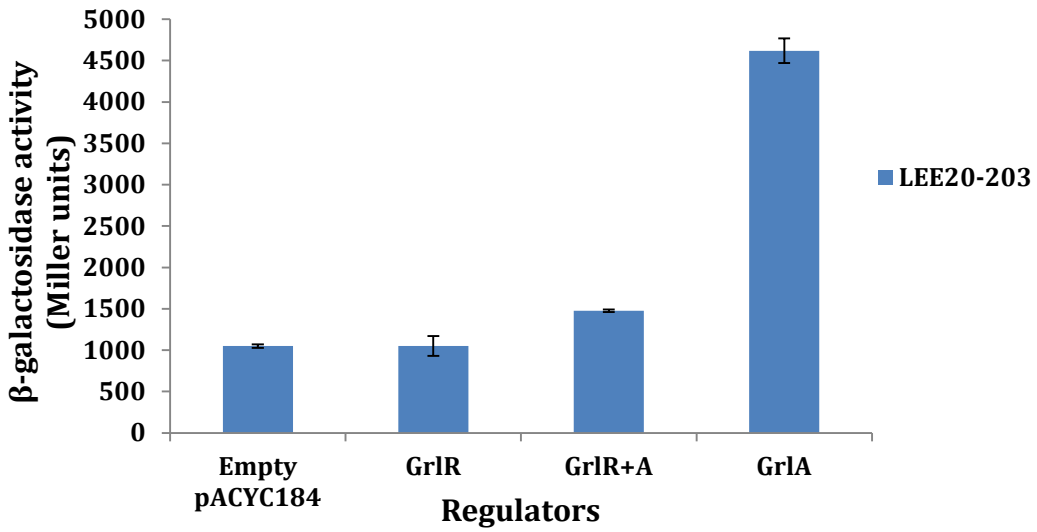


Fig. 3.14: Activation of expression from the LEE20-203 promoter by GrlA.
This bar chart shows the measured β-galactosidase expression in Sakai wild type strains, each containing pRW224 carrying a LEE20-203::*lac* fusion together with empty pACYC184, pACYC184/GrlR, pACYC184/GrlR+A, or pACYC184/GrlA. Measurements were made after growing the cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

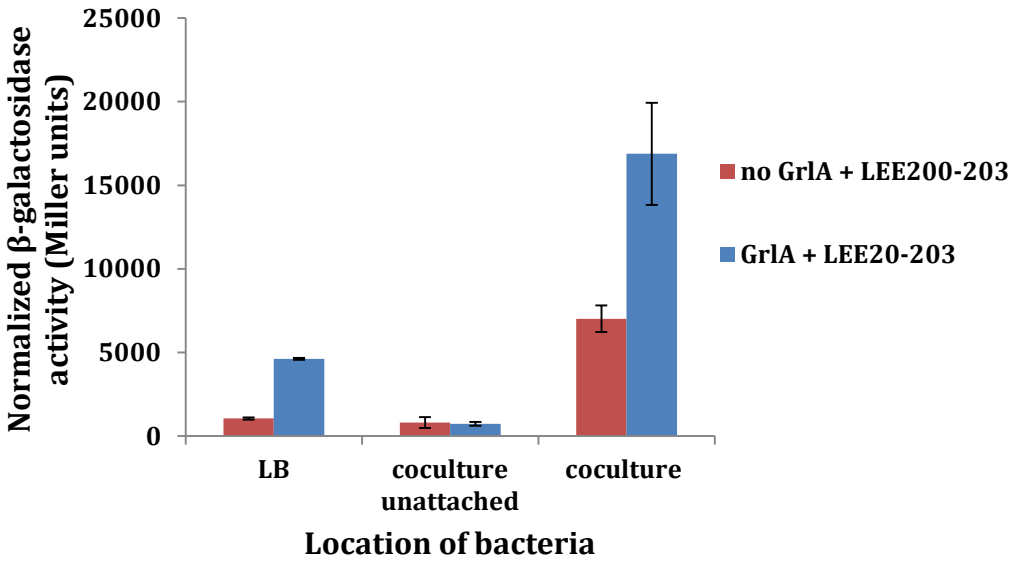


Fig. 3.15: Activation of expression from the LEE20-203 promoter by GrlA after attachment to Caco-2 cells.

This bar chart represents the measured normalized β-galactosidase expression in Sakai wild type strains, each containing pRW224 carrying an LEE20-203::*lac* fusion or empty pRW224 together with pACYC184/GrlR+A (blue bars) or empty pACYC184 (burgundy bars). Measurements were made after growing cells in LB medium at 37 °C to an optical density of ~0.5 at 650 nm. The assay was done in three different stages. At the undiluted stage, cells were lysed and assayed in planktonic culture. At the supernatant stage, cells that did not attach to the Caco-2 were lysed after three hours of contact with the Caco-2 cells, followed by those that attached to Caco-2 cells after three hours of contact (attached stage). Error bars represent the standard errors of three independent repeats.

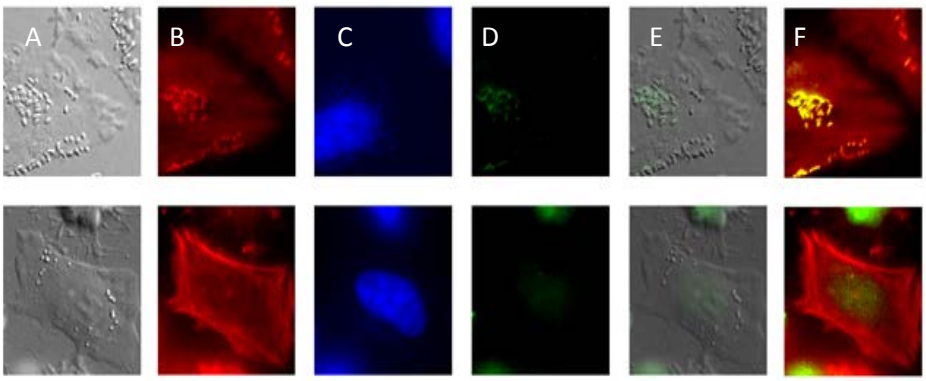
GrlR was present was similar to the expression from the LEE20-203 promoter with the pACYC184/ΔHN. When GrlR+A was present there was a ~ 1.4-fold increase in the level of expression from LEE20-203 compared to the level of expression from LEE20-203 with the pACYC184/ΔHN. The level of expression from the LEE20-203 was lower than LEE10-568.

Next, Sakai strains were transformed with pRW224/LEE20-203, with either GrlA or pACYC184/ΔHN, then overnight cultures from each strain were incubated with Caco-2 cells for three hours. Cells were lysed and assayed for β-galactosidase activity. Fig. 3.15 shows that the level of expression from the LEE20-203 promoter fragment with pACYC184/ΔHN was increased ~ 6.6 times in the attachment stage compared to the undiluted stage. Moreover, the level of expression from the LEE20-203 promoter when GrlA was present in the attachment stage was ~ 3.7 times higher than the level of expression from the LEE20-203 when GrlA was present in the undiluted stage. To sum up, the maximal activation of expression from the LEE20-203 promoter by GrlA was achieved when cells were introduced to host cells. However, by comparing the increase of the activation of expression from LEE10-568 with GrlA to the increase of the activation of expression from LEE20-203 with GrlA, it is clear that GrlA was activating LEE10-568 better than LEE20-203 after attachment.

3.10 Expression from the LEE20-203 promoter fragment in single cells

It was seen previously that attachment to HeLa cells triggered expression from the LEE10-568 promoter when GrlA was present. LEE10-568 represents the full length of the LEE1 promoter (Fig. 3.1). It was thought that the shortest promoter fragment

LEE20-203::*gfp* + pACYC184/GrlA



LEE20-203::*gfp* + pACYC184/ΔHN

Fig. 3.16: Expression of LEE20-203 promoter fragment in single cell.
This figure shows the GFP expression in Sakai wild type strains, each containing pRW400 carrying a LEE20-203::*gfp* fusion, together with either pACYC184/GrlA or pACYC184/ΔHN. Fluorescence microscopy was done after fixing Sakai strain cultures to HeLa cells and staining with phalloidin (red), which stained the actin filaments and showed the pedestal formation, and Hoechst (blue) that stained the nucleic acids in both Sakai strains and HeLa cells. A differential interference contrast (DIC) channel transmitted lights to show the structure's shape (A). The green channel transmitted at green wavelength to show the red fluorescence of the actin filaments and the pedestal formation (B). The ultraviolet channel (UV) transmitted at UV wavelength to show the blue fluorescence of the nucleic acid (C). The blue channel transmitted at blue wavelength to show green fluorescence when GFP was expressed (D). DIC+Blue, a merge of the DIC and blue channels showed which cells were glowing (E), and Green+Blue, a merging of the green and the blue channels showed whether the glowing cells were causing pedestal formation (F). (Scale bar, 10 μm)

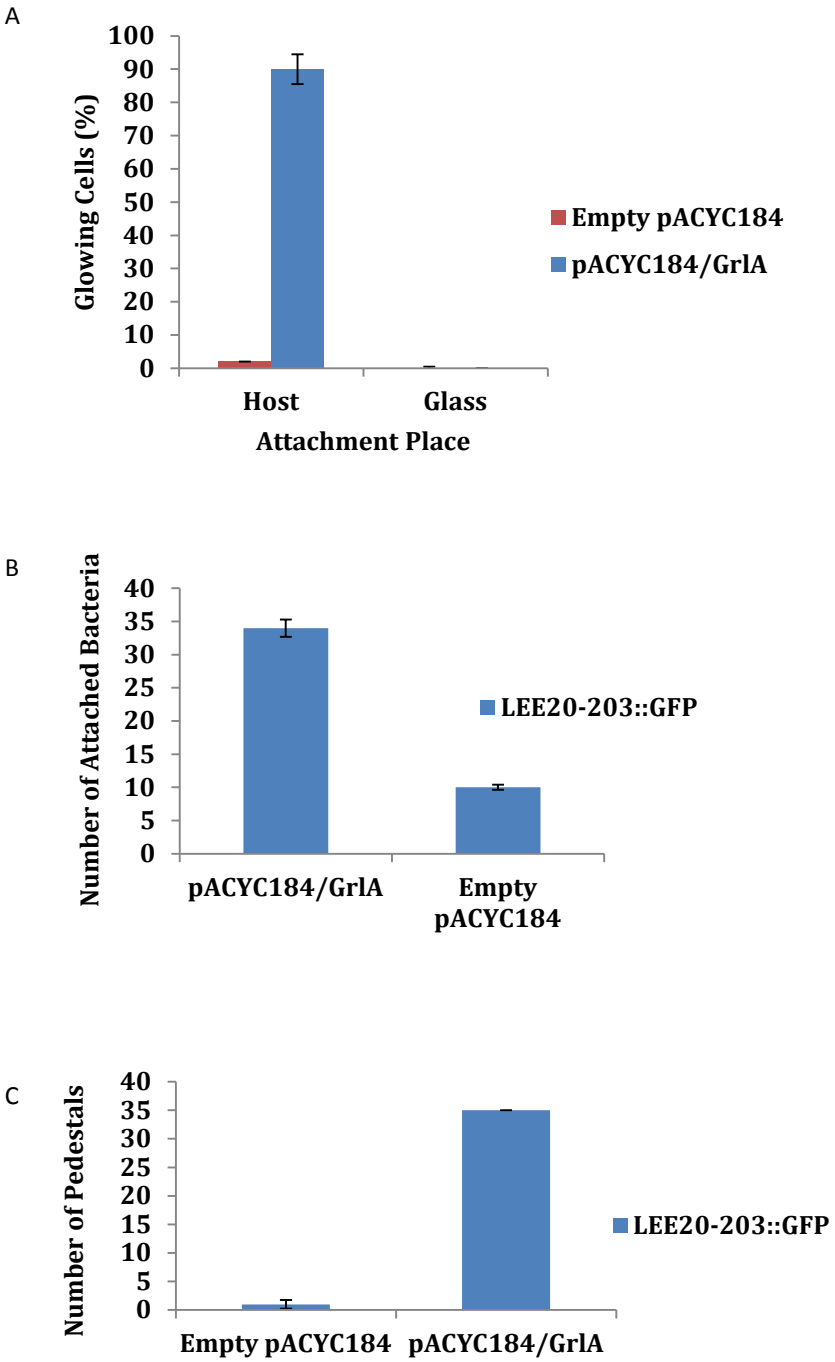


Fig. 3.17: Quantitative analysis of attached Sakai to HeLa cells.
This bar chart shows the statistical analysis of the data presented in fig 3.16. In A, the number of attached Sakai-strain cells to both host cells and the glass was counted and the percentage of glowing cells is indicated. In B, the number of bacteria attached to the host cells, with and without GrlA, was counted and is indicated. In C, the number of pedestals per cell was counted and is indicated. Error bars represent the standard errors of three independent repeats.

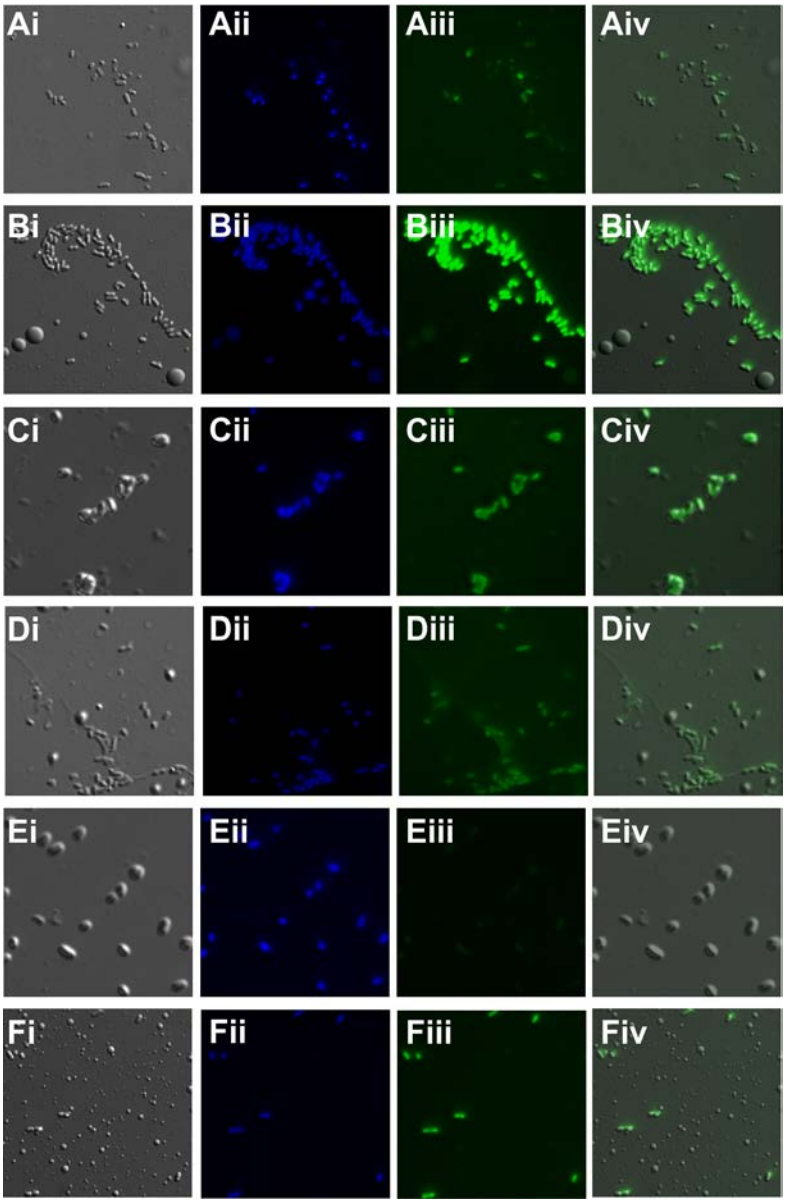
should be tested to confirm that the binding site of the triggered GrlA was the same site as that predicted (Islam *et al.*, 2011). To do this, Sakai strains were transformed with pRW400 carrying a LEE20-203::*gfp* fusion with pACYC184/GrlA or empty pACYC814 to look at single-cell expression. The overnight culture from each strain was incubated with HeLa cells for three hours. Cells were then fixed and stained for fluorescence microscopy.

It can be inferred from Fig. 3.16 that attachment to HeLa cells triggered the expression from the LEE20-203 promoter and that GFP was expressed when GrlA was present, unlike the bottom picture where GFP was not expressed with the pACYC184/ΔHN. The percentage of glowing cells when GrlA was present was ~ 45 times greater than the percentage of glowing cells with pACYC184/ΔHN (Fig. 3.17A-C). In addition, the number of bacteria attached to HeLa cells when GrlA was present was ~ 3.4 fold higher than the number of bacteria attached to the HeLa cells with pACYC184/ΔHN. The number of pedestals formed from bacterial cells when GrlA was present was higher than the number of pedestals formed from the bacterial cells with pACYC184/ΔHN, by ~ 35 times. However, the intensity with which Sakai cells with LEE20-203::*gfp* fusion glowed was lower than the intensity of glowing with LEE10-568::*gfp*.

3.11 Expression from the LEE10-568 promoter fragment in response to pure substrates

In the single-cell experiment mentioned earlier, Sakai strains were introduced to host cells. Thus, the strains were introduced to various host factors that could trigger the activation of expression from the LEE10-568 promoter. The next task was, therefore, to test a single cell substrate to see whether the expression from the LEE10-568 promoter could be triggered similarly to the attachment to HeLa cells or not.

To do this, three single substrates were used: poly-L-lysine, collagen, and fibronectin. Poly-L-lysine was used as a non-specific, positively charged molecule that bacterial cells attach to, whereas collagen and fibronectin were used as specific triggers from the host and they are components of the host cell system. All three substrates were fixed on glass slides separately then overnight cultures of the Sakai strains were transformed with pRW400 carrying a LEE10-568::*gfp* fusion with pACYC184/GrlA or with the pACYC184/ΔHN. These were incubated with each substrate for three hours. After that, cells were fixed and stained for fluorescence microscopy. Fig. 3.18A shows in the case of poly-L-lysine, that the level of expression from the LEE10-568 promoter was higher when GrlA was present (Biii) than the level of expression from the LEE10-568 promoter with the pACYC184/ΔHN. Despite that the fact that there were some glowing cells with the pACYC184/ΔHN (Aiii) the number and intensity of glowing cells when GrlA was present was ~ 16 times higher (Fig 3.18B). When collagen substrate was used (Fig3. 18A Ciii and Diii), the number of attached bacteria was lower


Ai
Aii
Aiii
Aiv
Bi
Bii
Biii
Biv
Ci
Cii
Ciii
Civ
Di
Dii
Diii
Div
Ei
Eii
Eiii
Eiv
Fi
Fii
Fiii
Fiv

G

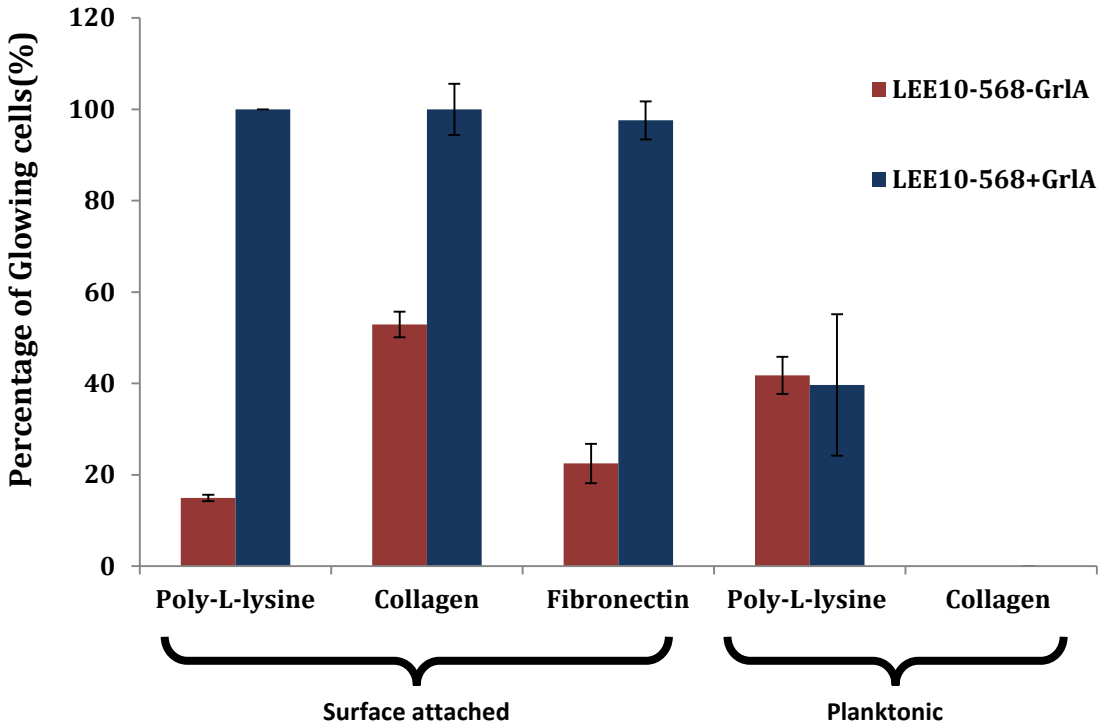


Fig. 3.18: Expression of LEE10-568 promoter in single cell upon attachment to single substrates.

A-F. This figure represents the GFP expression in Sakai wild type strain, each containing pRW400 carrying a LEE10-568::*lac* fusion promoter fragment together with either pACYC184/GrlA or empty pACYC184. Fluorescence microscopy was done after fixing Sakai strain cultures to poly-L-lysine (A and B), collagen (C and D), or fibronectin (E and F) and staining with Hoechst (blue), which stained the nucleic acids in Sakai cells. A differential interference contrast (DIC) channel transmitted lights to show the structure's shape (Fig. Ai, Bi, Ci, Di, Ei, and Fi). The ultraviolet channel (UV) transmitted at UV wavelength to show the blue fluorescence of the nucleic acid (Aii, Bii, Cii, Dii, Eii, and Fii). The blue channel transmitted at blue wavelength to show green fluorescence when GFP was expressed (Aiii, Biii, Ciii, Diii, Eiii, and Fiii). DIC+Blue, a merger of the DIC and the blue channel showed which cells were glowing (Aiv, Biv, Cvi, Div, Eiv, Fiv). (Scale bar, 10 μm)

G. Bar chart showing the quantitative analysis of the data presented in Fig. 3.18A. The percentages of glowing Sakai strains when the GFP was expressed after attachment to three single-cell substrates (poly-L-lysine, collagen, and fibronectin) are shown. The blue bar shows the percentage of glowing cells with GrlA, and the burgundy bar shows the percentage of the glowing cells with without GrlA. Error bars represent the standard errors of three independent repeats. Surface attached indicates that the substrates were first fixed on glass slides, and then cultures were added. "Planktonic" labelling indicates that both the substrates and the Sakai were incubated together in liquid culture then were fixed and stained.

than when poly-L-lysine was used. However, collagen was also used to trigger the GrlA function and enhance the expression from the LEE10-568 promoter with a ~ 2.8-fold more glowing cells than with the pACYC184/ΔHN (Fig. 3.18.B). In addition, fibronectin substrate was able to trigger activation of expression from the LEE10-568 promoter when GrlA was present (Fig. 3.18A Fiii and Eiii). The percentage of glowing cells when fibronectin was used in the presence of GrlA was ~ 4.5 times higher than when fibronectin was used with pACYC184/ΔHN (Fig. 3. 18B).

To see whether the triggered activation of expression was due to sensing a molecule in the liquid medium or it can be triggered after fixation, a control was included (Fig. 3.18B). To do this, ~ 2 ml of overnight culture of each strain was incubated in a 1.5 ml microfuge tube with poly-L-lysine or collagen then left at 37 °C for three hours. Strains were then fixed and stained for fluorescence microscopy. The number of attached bacteria was lower than the number of attached bacteria with poly-L-lysine. However, collagen was also able to trigger the GrlA. It was inferred that expression from the LEE10-568 promoter could be triggered by sensing the molecule in solution when poly-L-lysine was used, regardless of the presence or absence of GrlA. However, expression from the LEE10-568 promoter could not be triggered by collagen in a solution in either the presence or absence of GrlA.

It was clear that a non-specific substrate could trigger expression from the LEE10-568 promoter, unlike a specific substrate, with which attachment was required to

trigger the function. To conclude, both a non-specific molecule like, in this case, poly-L-lysine, and the specific molecules, collagen and fibronectin, can trigger GrlA.

However, the non-specific molecule can trigger the function in solution, but the specific substrate can trigger expression only after attachment.

3.12 Discussion

Regulation of expression from the *LEE* operons is believed to be a complicated process that is tightly regulated by a number of factors (Kaper *at al.*, 2004). In particular, GrlA is one of the major regulators of the *LEE* operon as it positively regulates expression from the *LEE1* operon, which then positively regulates the rest of the *LEE* operons. In this study, I have focused on analysing how, unbound GrlA is regulated in the pathogenic host, so I have used EHEC Sakai strains throughout as a host strain and I have also introduced a host cell system to gain a better understanding of how host cell factors may affect gene expression either positively or negatively. Attachment to a host cell is one of the main steps of bacterial pathogenesis (Dean-Nystrom *et al.*, 1998). In the case of EHEC, a number of attachment determinants are employed to aid in intimate attachment to host cells, mainly intimin and Tir receptors.

The issue to address here was whether attachment to host cells regulates gene expression in EHEC, as it is known that two molecules of GrlR bind to one molecule of GrlA, and that the binding affinity of the GrlR molecule to the GrlA molecule is higher than the binding affinity of the GrlA molecule to the DNA elements

(Padavannil *et al.,* 2013). In addition, in the absence of GrlR, GrlA still only modestly activated the expression from the LEE1 promoter, as shown in this study. Not only that, but in planktonic culture free, unbound GrlA could activate expression from the LEE1 promoter by ~ 1.5 times in the EHEC host strain, thus, raising the question of whether attachment to the host cell can trigger the free GrlA. In addition, the questions arise of whether attachment to the host cell affects bound or free GrlA or both, and whether the LEE1 promoter is up regulated by attachment to the host cells.

It has been shown that attachment to host cells can trigger the free, unbound GrlA (Fig. 3.19), which then activates expression from the LEE1 promoter. In addition, attachment also affects bound GrlA, but the impact on free GrlA was higher than the impact on bound GrlA, as shown previously (Padavannil *et al.,* 2013). Attachment to host cells could, therefore, trigger the bound molecule. Thus, GrlR releases some GrlA that is able to activate expression from the LEE1 promoter and even greater impact was achieved on the free GrlA resulting in a maximum activation of the LEE1 promoter. In addition, I have shown that LEE1 promoter activity was significantly enhanced after attachment to host cells. Moreover, LEE1 promoter activated by free GrlA resulted in enhanced attachment to the host cells, as shown by the increase in colonies attached to host cells after enhanced expression from the LEE1 promoter by free GrlA.

Single cell expression method has also been introduced to visualize the impact of cellular attachment on expression. It was clear that attaching to host cells resulted in triggered GrlA that was then able to bind and activate expression from the LEE1

promoter. This was demonstrated by the expression of the GFP and by visualizing the cytoskeletal rearrangement and pedestal formation features in HeLa cells and the higher number of bacterial cells attached to HeLa cells.

To determine whether active GrlA was picking up a site in the promoter that was not the predicted binding site of GrlA, the shortest promoter fragment was tested. The result showed that triggered GrlA could also activate expression from the shortest fragment but less than the LEE10-568 promoter fragment. This can be explained by looking back at the sequence of the shortest construct which is a 46-base-pair fragment containing only the -35 and -10 elements with the 18-base-pair spacer region between them. It lacks the AT-rich region upstream from -35, which is required for optimal binding of GrlA to the promoter. To conclude, triggered GrlA was picking the predicted binding site in the promoter, which is the spacer region between -35 and -10 elements.

The next question to address was whether a single-molecule substrate can trigger GrlA and result in expression from the LEE1 promoter. The data showed that GrlA could be triggered by both a non-specific substrate, which was in this case Poly-L-Lysine, and also by specific substrates such as collagen and fibronectin. The next question was whether this activation could be triggered when the cells were attached or only when the bacterial cells actually sensed the triggering molecules. The data showed that the non-specific trigger, Poly-L-Lysine, could trigger GrlA when it is fixed on a glass slide and in a liquid suspension. However, collagen can only trigger

GrlA on a fixed slide. This data shows how important attachment is in triggering the GrlA function and thereby up-regulating gene expression.

To sum up, gene expression from the LEE operons is a complicated and carefully regulated mechanism. This system can be triggered when the bacterial cells attach to the host cells and results in higher gene expression and, in turn, higher attachment to host cells in a positive feedback loop (Fig 3.19). As already shown, this system can only be triggered upon attachment to host cells. Otherwise, in planktonic culture, the amount of activation is modest. The number of bacterial cells was proportionally related to the level of expression. Thus, more attached bacteria result in a higher level of gene expression. Performing the FAS test determined the actual level of gene expression in the single-cell experiment. In addition, the quick response of the triggered GrlA and the high intensity of GFP production can also be monitored by performing QPCR in the future.

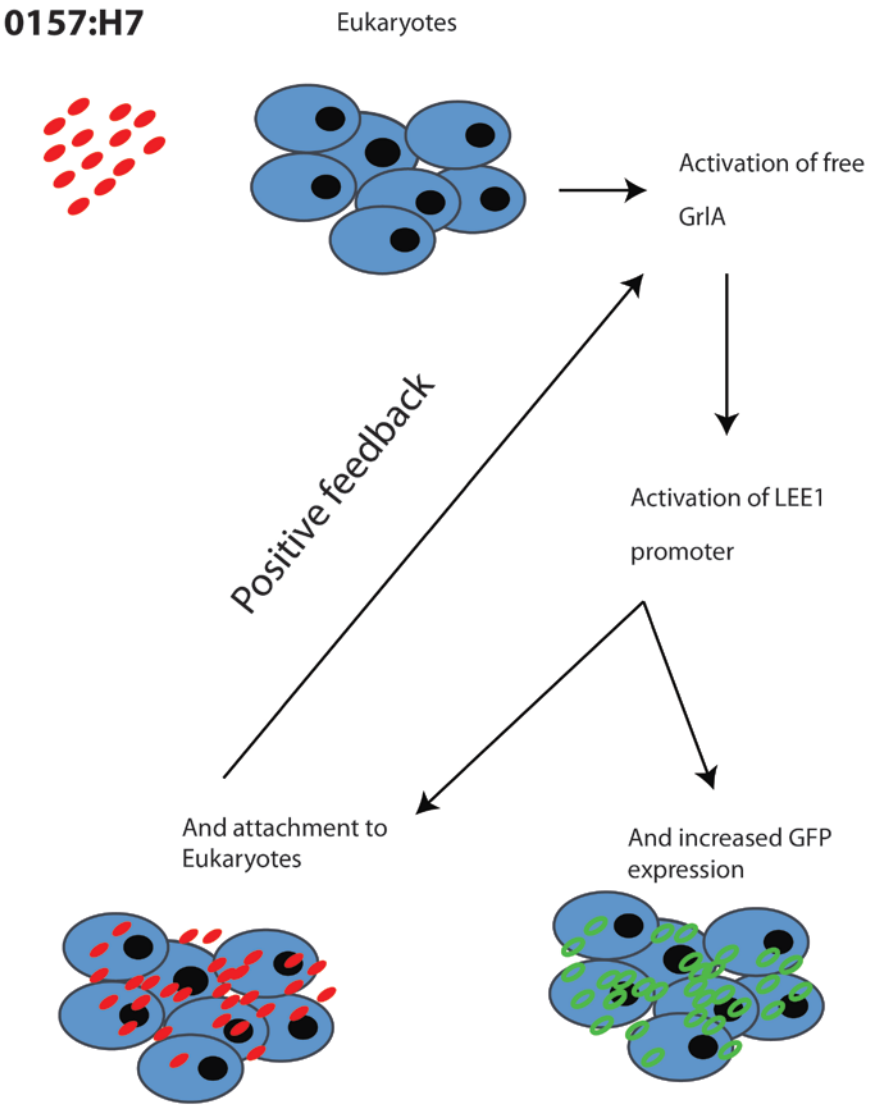


Fig. 3.19: Model for activation of LEE1 promoter by GrlA after attaching to host cells. This figure shows the proposed scenario by which the LEE1 promoter is activated by free GrlA. Once O157:H7, represented by the Sakai strain in this study, attach to host cells, free unbound GrlA becomes activated. Active GrlA results in the up-regulation of expression from the *LEE1* operon, as proposed previously by binding to the spacer region in the LEE1 promoter between -10 and -35 elements. Activation of the *LEE1* regulatory region resulted in the activation of the rest of the *LEE* operons and increased expression of the Type III secretion system, which then increased the attachment of the bacterial cells to the host. Attachment to the host cells also resulted in increased GFP expression.

Chapter Four:

Regulation at the LEE1 leader sequence and promoter

4.1 Regulation at the *ler* leader sequence and role of the minigene

In many bacterial mRNAs, the first translation initiation signal is some distance from the 5' end and does not involve the first AUG methionine codon (Kozak, 1983; O'Donnell *et al.,* 2001). The Shine-Dalgarno sequence is considered to be the primary determinant that assures translation initiation, and it is a short sequence composed of a 4-6 base-pair sequence that is complementary to a sequence at the 3' end of the 16S ribosomal RNA (Shine and Dalgarno, 1974; Steitz and Jakes, 1975). A compilation of the 5' end of *Escherichia coli* messenger RNAs has displayed that most messages have 40-80 un-translated bases; however, there are some with much longer un-translated sequences (Shultzaberger *et al.,* 2007). These sequences are called leader sequences, and they can fold into various secondary structures. Moreover, these leader sequences are believed to be involved in regulatory functions, such as posttranscriptional regulation, but most leader sequences are un-translated (Gardner, 1979; Gavini and Pulakat, 1991).

The translation start site, AUG, of Ler, which is the first ORF (open reading frame) in the *LEE1* operon, is ~170 base downstream from the 5' end of the *LEE1* mRNA (Fig. 4.1). This unusually long leader sequence contains mostly adenosine and uridine. It also contains a potential AUG at +38 relative to the transcription start site, which causes expression from a small open reading frame (minigene). The mingene of *ler* is a short sequence containing two codons, and a Shine-Dalgarno sequence, which is GGTGG, seven bases away from the AUG at +38. Ler AUG is located at +170. It also contains other AUGs at +47 and +152, which are non functional.

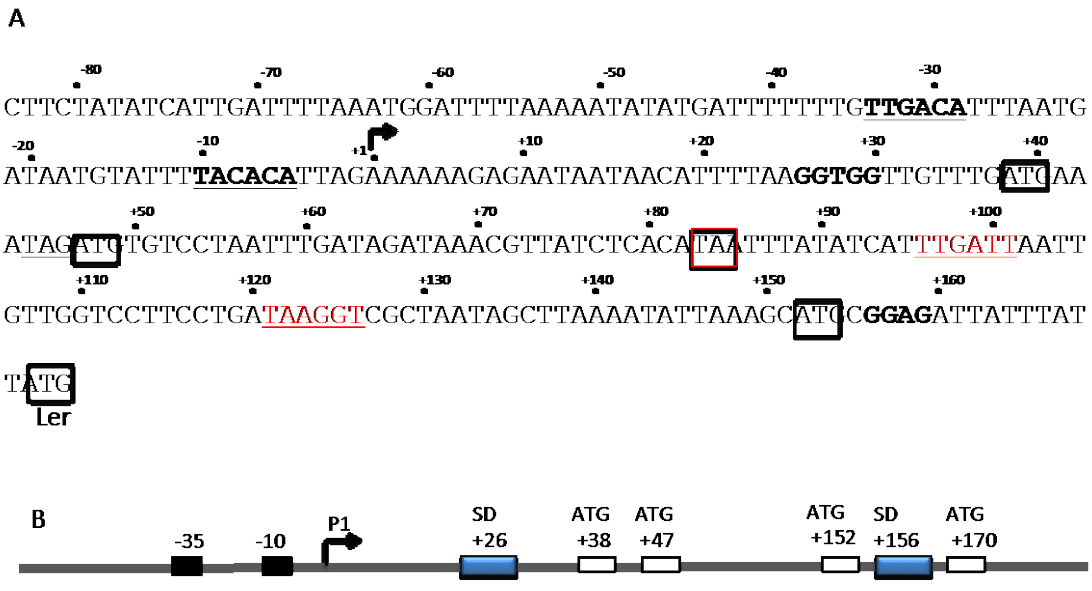


Fig. 4.1: Base Sequence and a schematic representation of the *LEE* regulatory region.

A. Base sequence of the *LEE* regulatory region. The predicted -10 and -35 promoter elements are shown underlined. The transcription start-site (TSS) is indicated by a bent arrow. The Shine-Dalgarno sequences are in bold black, and the potential translation start-sites (AUGs), are boxed. Stop codons are either underlined or red boxed at +44 and +83 respectively.

B. A schematic representation of the *LEE* regulatory region showing the promoter elements in black boxes, the TSS with a bent arrow, the position of the Shine-Dalgarno sequences in blue boxes, and the potential AUG positions relative to the transcription start-site in white boxes.

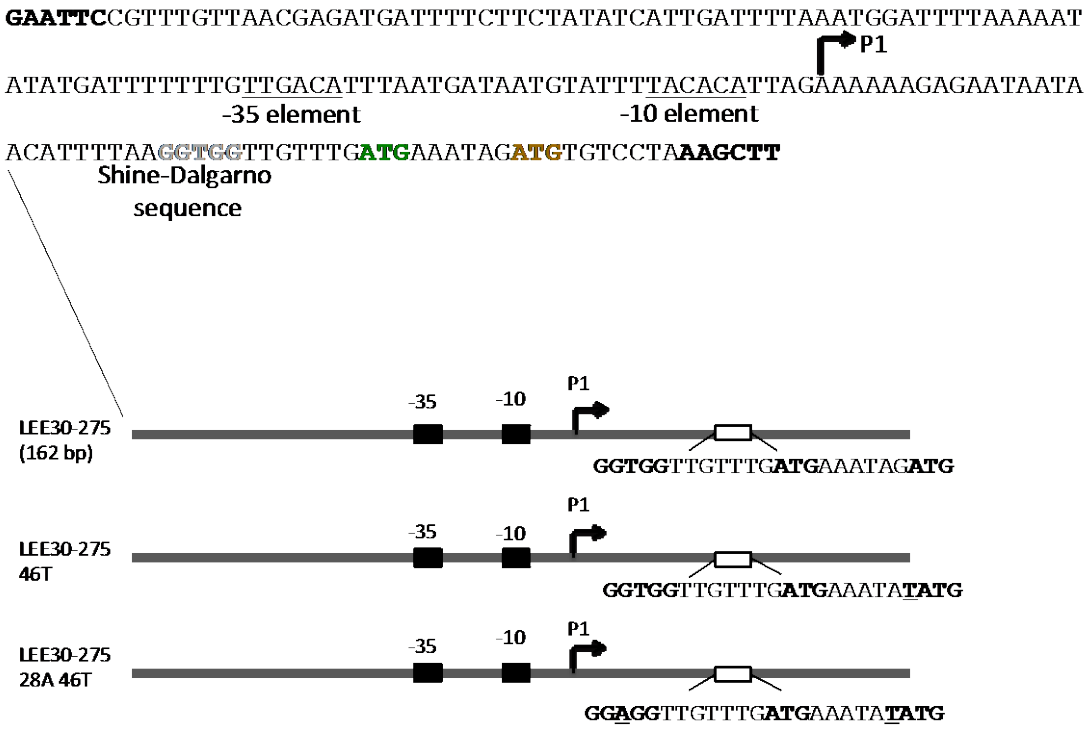
It was found that expression from this short open reading frame had influenced Ler expression, and interfering with the minigenes alters Ler expression and the ability of EHEC to interact with host cells (Islam *et al.,* 2012).

Here, the role of the leader sequence in regulating Ler expression and the impact of altering the ribosomal binding sites on expression were investigated. This was to see if altering the SD-sequences would have an effect on the expression from the minigene described by Islam *et al*. Later, a similar mutation in LEE150, the longer fragment, was introduced to see whether or not it affects RNA folding and expression of Ler. Then the effect of modifying the SD-sequence adjacent to the Ler initiation codon to the Ler affects its expression was looked at by introducing point mutations in the sequence. Finally, the role of the leader sequence was investigated by deleting it and assaying the effect on *ler* expression. Answering these questions led to understanding how *ler* expression is regulated by cis-acting RNA sequences.

4.1.1 Modification of the Shine-Dalgarno sequence affects the expression from the minigene

Altering the sequence upstream of the minigene of Ler affected its expression positively (LEE30-275 46T). This was done by cloning a short fragment that runs from -275 to -114 relative to the translation start-site of Ler, LEE30-275 (Fig. 4.2A) into pRW225 (Fig. 2.4), as a *lacZ* translation fusion vector. The construct was then

A



B

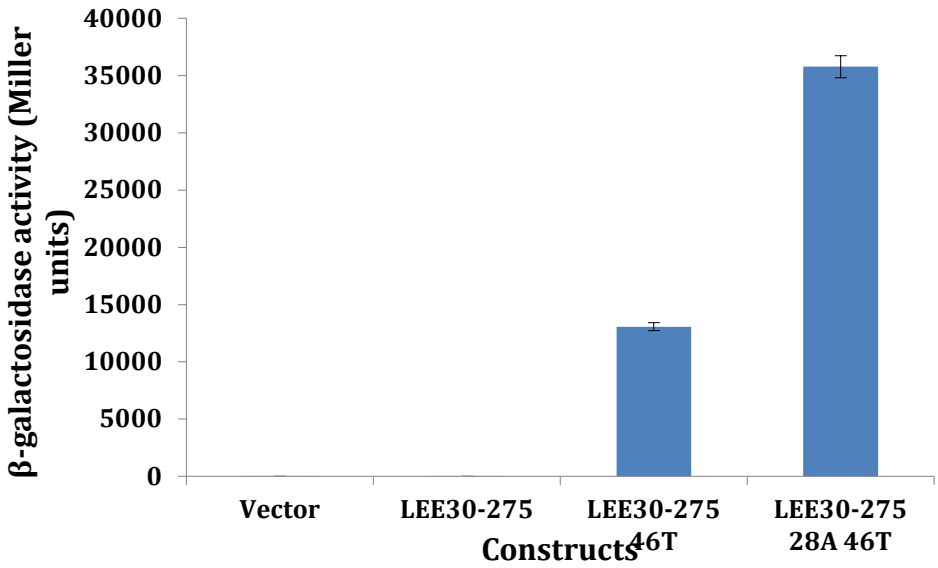


Fig. 4.2: Activation of expression from the LEE30-275 and derivatives.

A. A schematic representation of the LEE30-275 fragment, starting from -106 to +65, relative to the *ler* translation start-site. The Shine-Dalgarno sequence is shaded grey, the potential ATG green and brown.

B. Measured β-galactosidase expression in M182 strains, each containing pRW225 carrying a LEE30-275::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

transformed into *E. coli* K-12 (M182 strain) and the β-galactosidase assay was carried out to measure the level of minigene expression (Fig. 4.2B). It is clear that the level of expression was low. To alter this situation, a point mutation was introduced in the stop codon prior to the second start codon (ATG at position +47 relative to the transcription start-site). This was a change from TAG at +46 to TAT, which then extend the length of the minigene product, such that translation stops at +83. This newly constructed fragment was labelled LEE30-275 (46T) and, as shown in Fig. 4.2B, the level of expression was increased in comparison to the LEE30-275 (WT) fragment.

The role of the Shine-Dalgarno sequence of the minigene at position +26 and was then considered, and it was questioned whether altering it could have an impact. In this case the SD-sequence was changed to the consensus sequence so, instead of GGTGG, was constructed to GGAGG. This mutation was combined with the earlier one, the 46T, to make the LEE30-275 (28A, 46T) fragment, which was then fused to the *lacZ* expression vector as translation fusion. The construct was introduced into *E. coli* K-12 (M182 strain), and the β-galactosidase assay was performed. It is clear that changing the Shine-Dalgarno sequence had a positive effect on the expression from the minigene (Fig. 4.2B), which might indicate that the original ribosomal binding site was not optimal.

Fig. 4.2 shows that making the SD-sequence closer to the consensus had a positive impact on the expression of the mini ORFs as the level of β-galactosidase was

significantly higher than both the LEE30-275 wild type and LEE30-275 (46T). In addition to the β-galactosidase assay, SDS-PAGE was performed for all constructs to see the enhanced level of protein expression, which is clearly visible in Fig. 4.3.

Lanes B and C represent the protein profile of cells containing pRW225/LEE30-275 fused to *lacZ*, where two different loading volumes were used, 5 and 10µl respectively. Lanes D and E represent the protein profile of cells containing pRW225/LEE30-275 (46T). In this construct a point mutation was introduced in the stop codon at position 46. The protein gel shows, a new band, which is believed to be LacZ. Indicating altering the codon caused message read though seen in protein expression. Again, the samples were loaded using two different volumes. Lanes F and G represent the protein profile of cells containing pRW225/LEE30-275 (28A, 46T), where a second mutation was introduced at the SD-sequence beside mutation at stop codon. In this profile, the protein expression of the LacZ was enhanced compared to LEE30-275 (46T). This gel showed with data from B-gal that altering SD-sequence, influenced expression from the minigene positively. Lane A was a mixture of protein marker.

4.1.2 Effect of modifying the Shine-Dalgarno sequence on *ler* expression

The next question was whether altering the minigene's expression would have an effect on the expression of Ler. To answer this question, the full-length promoter

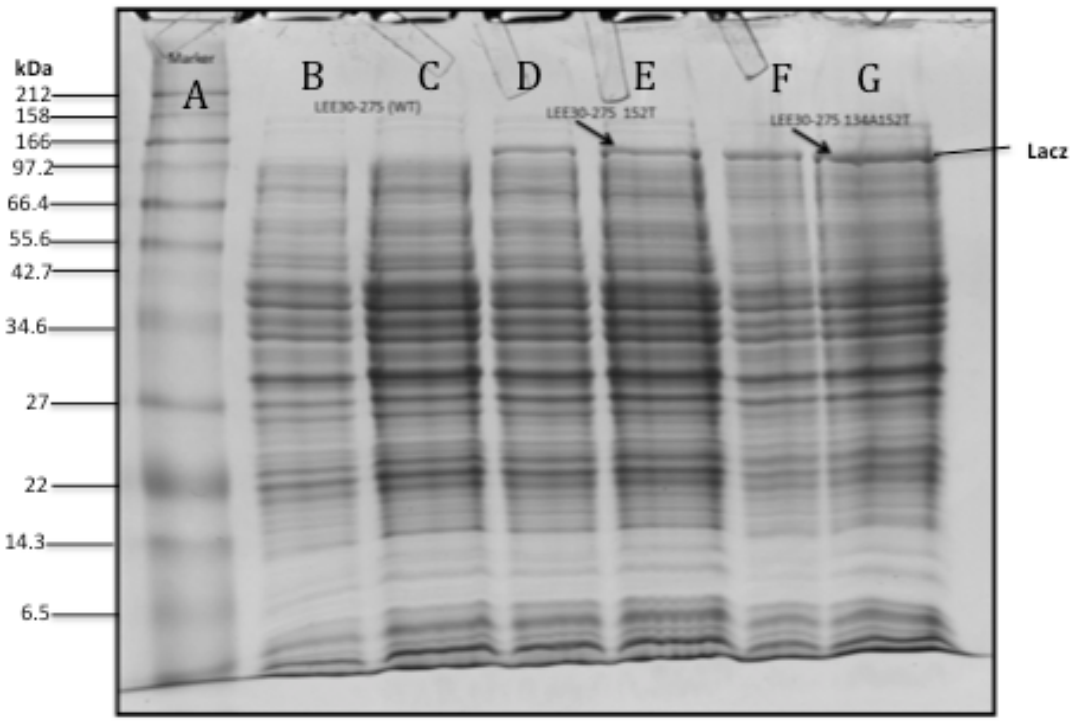


Fig. 4.3: SDS-PAGE profile of proteins expressed from the minigene in LEE30-275 and derivatives.

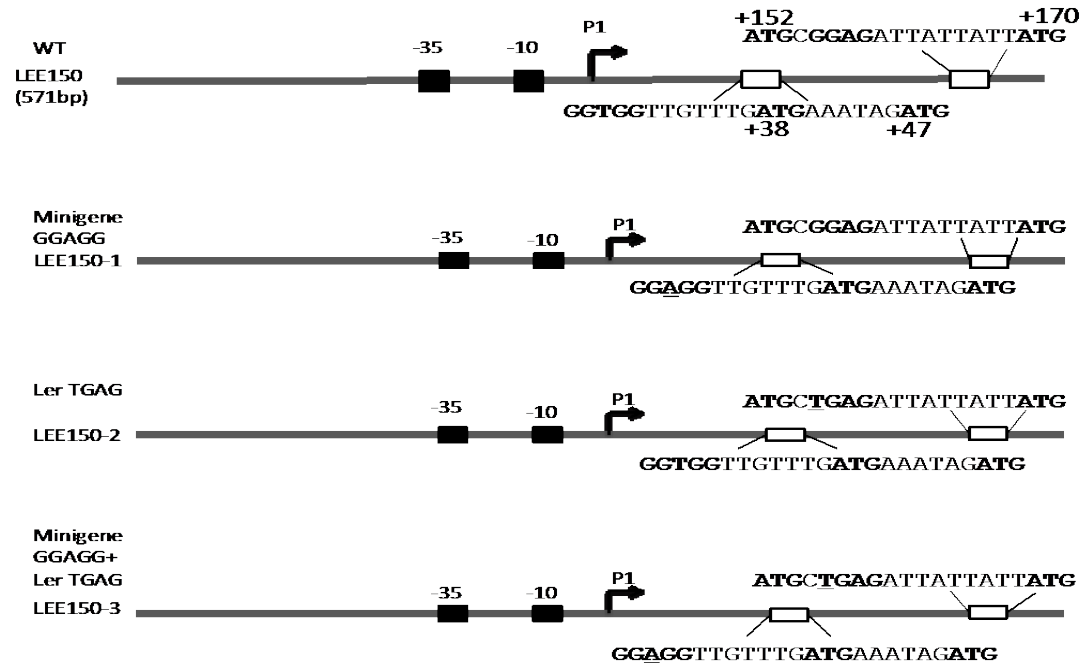
This figure shows the protein profile of the *E. coli* cells containing three constructs: LEE30-275 lanes B and C, LEE30-275 (46T) lanes D and E and LEE30-275 (28A 46T) lanes F and G. Each sample was run in duplicate in the first lane (5µl) and in the second lane (10µl). Lane A is mixed protein marker was run alongside the samples. The arrows point to protein encoded by the minigene *lacZ* fusion.

containing the leader sequence LEE150 was used. It was clear that introducing a base mutation in the Shine-Dalgarno sequence fused to the minigene increased the expression significantly. The same point mutation was then introduced into the longer fragment, LEE150, to see if *ler* expression could be affected.

To do this, the LEE150 fragment, which starts from -568 to +169 relative to the translation start-site of *ler*, and which carries the LEE1 promoter and the leader sequence, was fused to a pRW225 cloning vector, where the *lacZ* was fused to the AUG of *ler* (Fig. 4.4A).

Thus, any effect on the *lacZ* expression reflects the *ler* expression. The construct was transformed into *E. coli* K-12 (M182) strain and the β-galactosidase assay was carried out. Fig. 4.4B shows the basal level of *ler* expression in the LEE150 fragment in Miller Units. Then the Shine-Dalgarno sequence at +26 was altered from GGTGG to GGAGG, and the newly constructed fragment was labelled LEE150-1. Four primers were used to construct mutation in the SD-sequence. This is creating two PCR fragments that then were used in three way ligation method was followed in this experiment The level of Ler expression is shown in Fig. 4.4B and here it can be seen that altering the ribosomal-binding site of the minigene had no effect on the level of expression of Ler, hence the β-galactosidase activity was similar to LEE150.

A



B

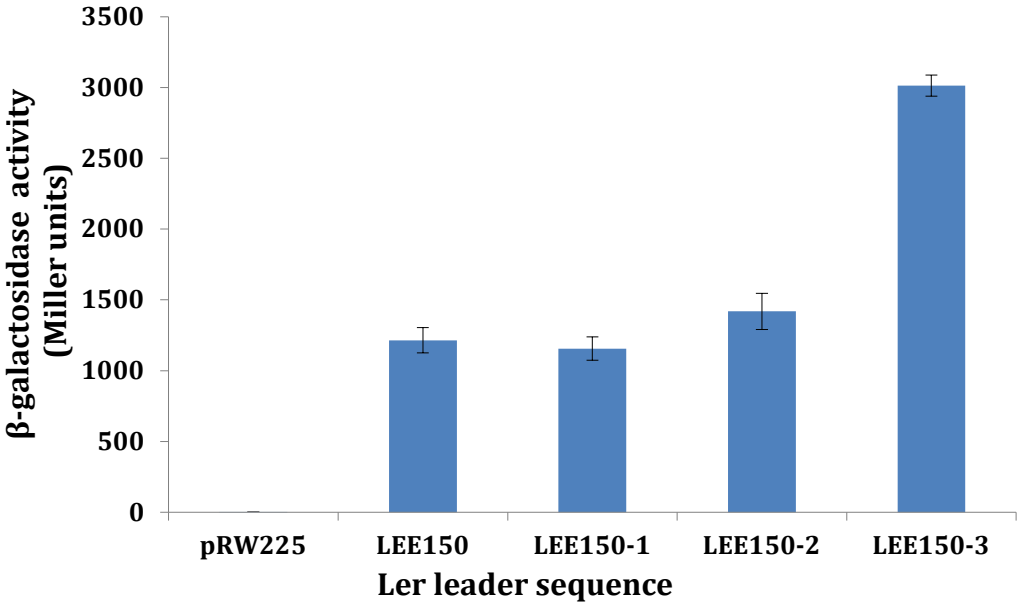


Fig. 4.4: Activation of expression from the LEE150 and derivatives.
A. A schematic representation of the LEE150 fragment and derivatives.

B. Measured β-galactosidase expression in M182 strains, each containing pRW225 carrying LEE150::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

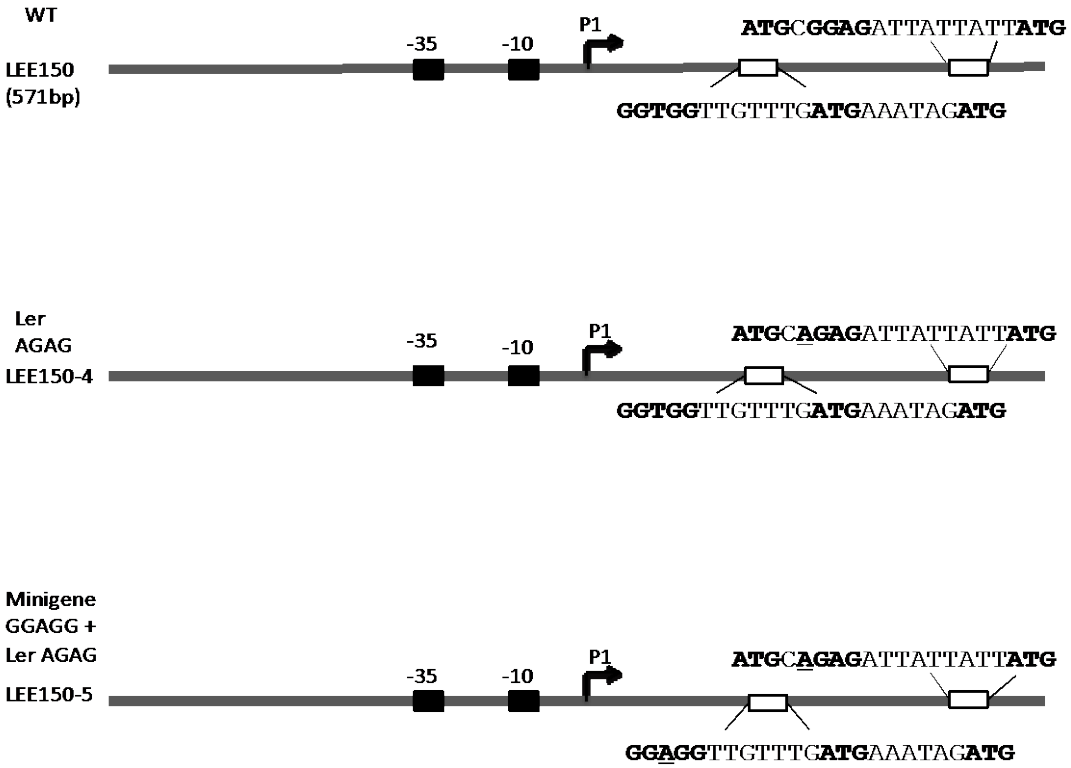
In order to establish what can affect *ler* expression, either positively or negatively, the down-stream Shine-Dalgarno sequence at +156 fused to Ler was mutated as well. To do this a downstream primer carrying a point mutation in the Shine-Dalgarno sequence at +156 was designed, and the fragment, LEE150-2, was constructed.

Modifying the ribosomal binding site from GGAG into TGAG, which was thought to change it into the non-consensus sequence, did not result in enhanced or reduced Ler expression (Fig. 4.4B).

Then, both mutations mentioned earlier in fragment LEE150-1 and LEE150-2 were combined into a one fragment and assay the effect, speculating that the enhanced ribosomal binding of the minigenes and the mutation in the Shine-Dalgarno fused to *ler* might enhance *ler* expression. The newly made fragment was LEE150-3 and mutations in both Shine-Dalgarno sequences, upstream and downstream, had positive effects on Ler expression (Fig. 4.4B). The level of Ler expression was ~ 3 times higher than the basal level of expression from the LEE150.

Then a different point mutation was introduced in the Shine-Dalgarno sequence fused to *ler* to bring it near to the consensus sequence, namely, GGAG into AGAG. The fragment was then labelled LEE150-4 and assayed. The point mutation showed enhanced Ler expression compared to the LEE150 level of expression (Fig. 4.5B). To see if this enhanced level of expression would be higher when combined with the mutated Shine-Dalgarno sequence upstream, a new fragment was constructed in

A



B

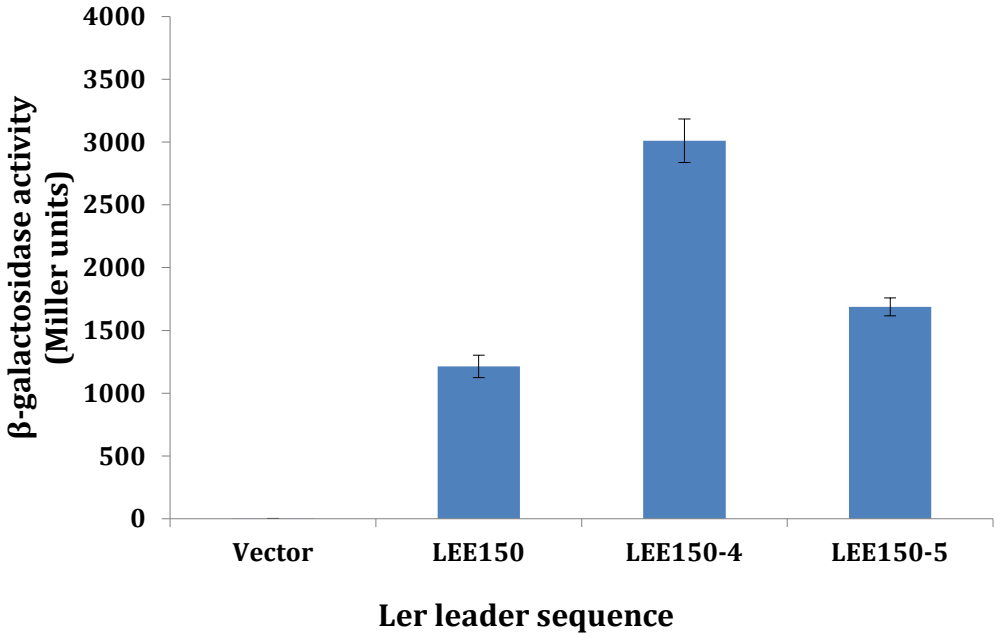


Fig. 4.5: Activation of expression from the LEE150 and derivatives.
A. A schematic representation of LEE150 fragment and derivatives.

B. Measured β-galactosidase expression in M182 strains, each containing pRW225 carrying LEE150::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

which the upstream Shine-Dalgarno sequence was altered from GGTGG into GGAGG and the downstream Shine-Dalgarno sequence was altered from GGAG to AGAG to make LEE150-5. The construct was then transformed into *E. coli* K-12 (M182 strain) and assayed. In this case, the level of *ler* expression was reduced to the basal level of *ler* expression. This suggests that when the Shine-Dalgarno sequence for *ler* mRNA translation, was changed to become nearer to the consensus, the level of expression could not be increased with a different mutation (Fig. 4.5B).

The data showed that the expression from the minigene was enhanced when a point mutation was introduced in both the Shine-Dalgarno sequence and in the stop codon. This mutation had no effect on Ler expression. Ler expression was positively enhanced when a specific point mutation was introduced in the Shine-Dalgarno sequence fused to Ler, which is GGAG instead of AGAG. This is because when the mutation was changed from A into T, there was no effect on the Ler expression, unless a second mutation was introduced in the upstream Shine-Dalgarno sequence.

4.1.3 Effect of deleting the leader sequence on Ler expression

It was seen earlier that modifying the ribosomal binding site sequence could enhance expression of *ler*. Later, it was investigated whether the leader sequence itself has an impact on the level of expression. As mentioned above, LEE1 is unique for having a long leader sequence of ~ 170 bases. This unusual long leader sequence might have a regulatory function on *ler* expression that yet to be identified. To test this hypothesis,

the leader sequence was deleted starting downstream from the -10 elements, from -163 down to -19 relative to *ler* transcription start site. The starting fragment was LEE150 (WT) and was used as a template for the PCR (Fig. 4.6A). To construct this fragment, four oligonucleotides were designed. The first two oligos were used to amplify the promoter fragment down to the -10 elements. The second two oligos were used to amplify the section from the *ler* SD-sequence and AUG down to the *lacZ*. Then the two resultant fragments were cloned into a pRW225 cloning vector (Fig. 4.6A). Next, a β-galactosidase assay was carried out after growing cells at 37°C as shown in Fig. 4.6B. The level of *ler* expression was significantly higher when the leader sequence was deleted in the LEE150 *Δl* than in the LEE150. The result shown in Fig. 4.6B suggests that the leader sequence function might be a binding site for negative regulators, a repressor binding to the RNA, a structure in the RNA, or a degradation effect. Next, the same constructs were tested in various temperature conditions to see if expression is dependent on temperature variations (Fig. 4.6A and B). It was thought that the RNA forms a stable complex at lower temperatures, which blocks translation, but when the leader sequence is deleted activity markedly increased. At a higher temperature, the level of expression did not change in either LEE150 or in LEE150 *Δl*, but in this case and at 37° C, the level of expression was higher in the LEE150 *Δl* than in LEE150. When the temperature was lowered to 30° C, the level of expression from the LEE150 increased and the level of expression from the LEE150 *Δl* was higher, but still less than at 37° C. However, the levels of expression at 20° C were low for both the LEE150 and the LEE150 *Δl* (Fig. 4.7A).

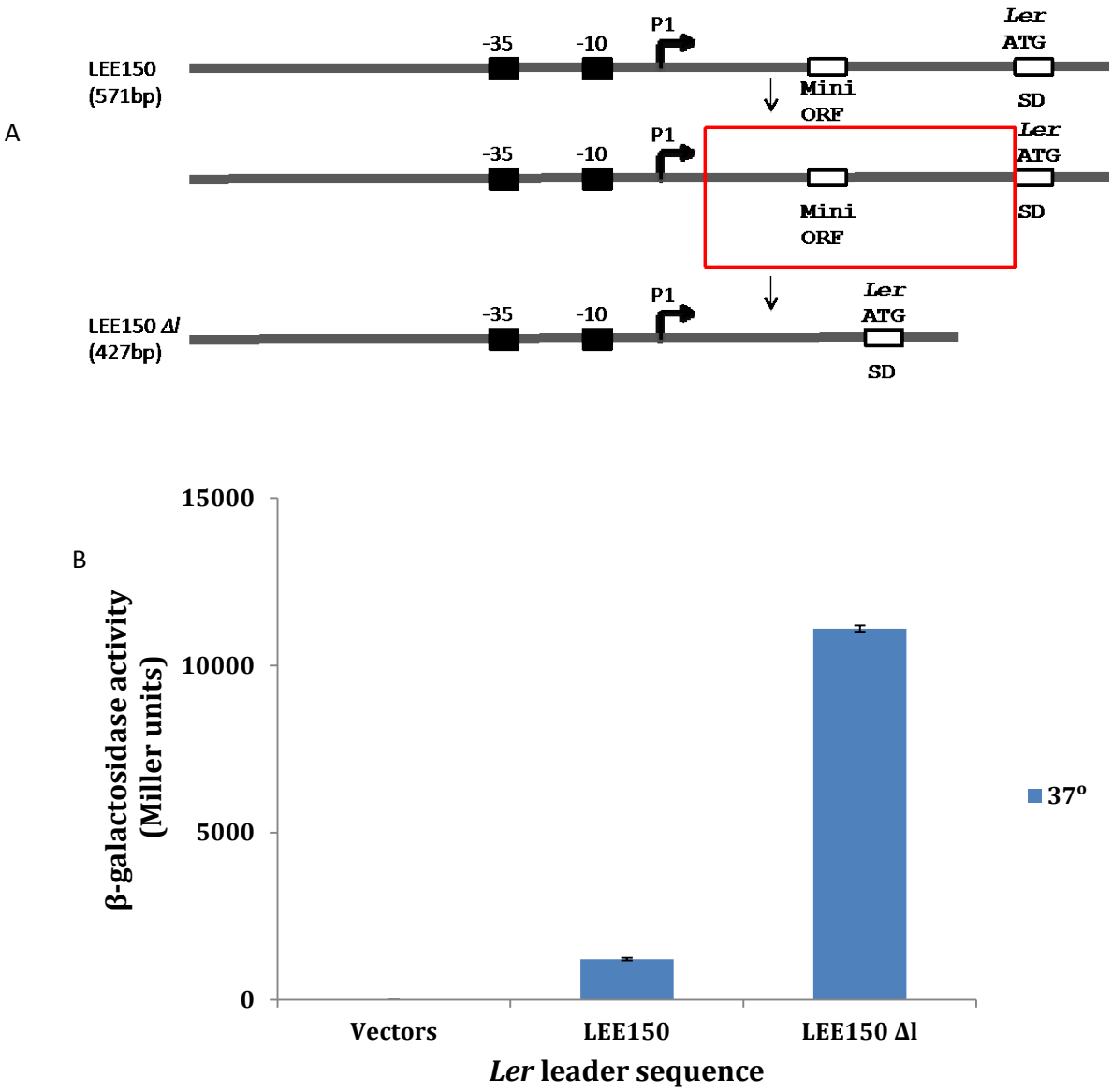
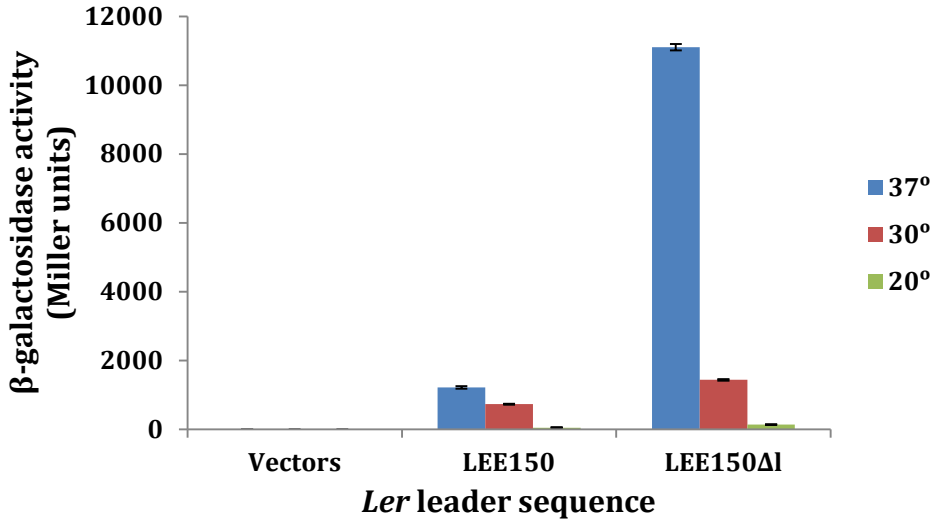


Fig. 4.6: Activation of expression from the LEE150 and LEE150 *Δl*.

A. A schematic representation of the LEE150 fragment and derivative.

B. Measured β-galactosidase expression in M182 strains, each containing pRW225 carrying LEE150::*lac* and LEE150 *Δl* fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats

A



B

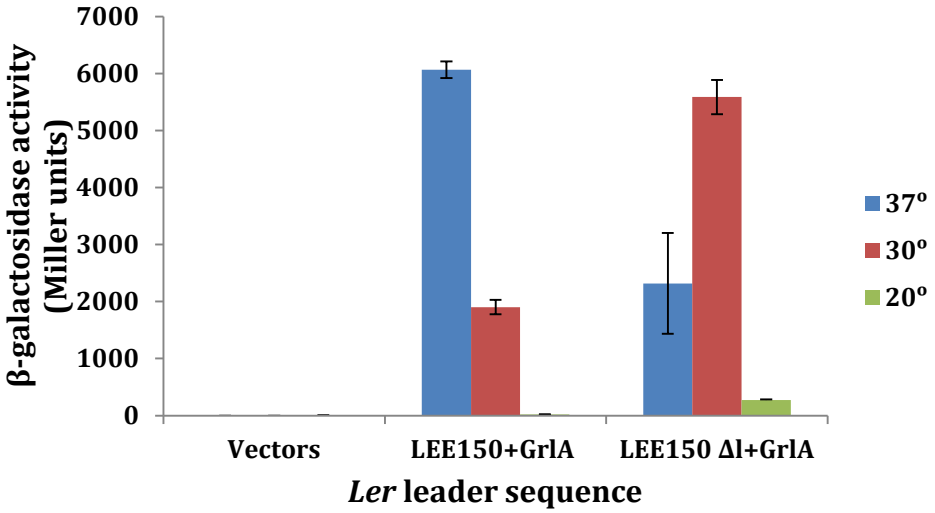


Fig. 4.7: Activation of expression from the LEE150 and LEE150 *Δl* in different growth temperatures.

Measured β-galactosidase expression in M182 strains, each containing pRW225 carrying LEE150::*lac* and LEE150 *Δl* fusions. Measurements were made after growing cells in LB medium at 37° C, 30° C, and 20° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

A. Shows the basal level of expression from each construct.

B. Shows the level of expression with GrlA.

Then, the impact of GrlA on the level of expression from each construct and at different temperatures was investigated (Fig. 4.7B). In LEE150, the level of expression was enhanced at 37° C and 30° C, but the level was very low at 20° C. However, in LEE150 *Δl,* the level of expression was reduced at 37° C, but enhanced at both 30° C and 20° C. The maximal levels of expression from LEE150 and LEE150 *Δl* are plotted in Fig. 4.8A and B.

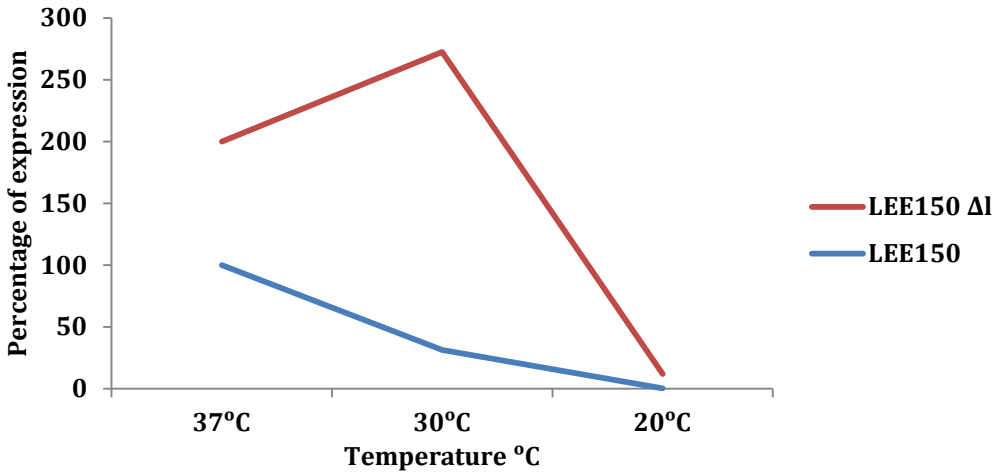
To conclude, the leader sequence might function as a binding site for negative regulators, such as HN-S, or other silencing proteins, or there could be some structures binding to the RNA, or other degradation factors associated with it. This could be investigated further by testing the system in strains with negative regulators deleted.

4.1.4 Conclusion

It is evident that the *LEE1* promoter contains a long leader sequence, which has three AUG start codons along with a *ler* AUG start codon; it is believed that expression from these mini ORFs has regulatory functions.

In a previous study (Islam *et al*., 2012), it was shown that altering these mini ORFs caused altered *ler* expression, as well as EHEC interaction with host cells. In the current study, it was evident that enhancing the Shine-Dalgarno sequence had a

A



B

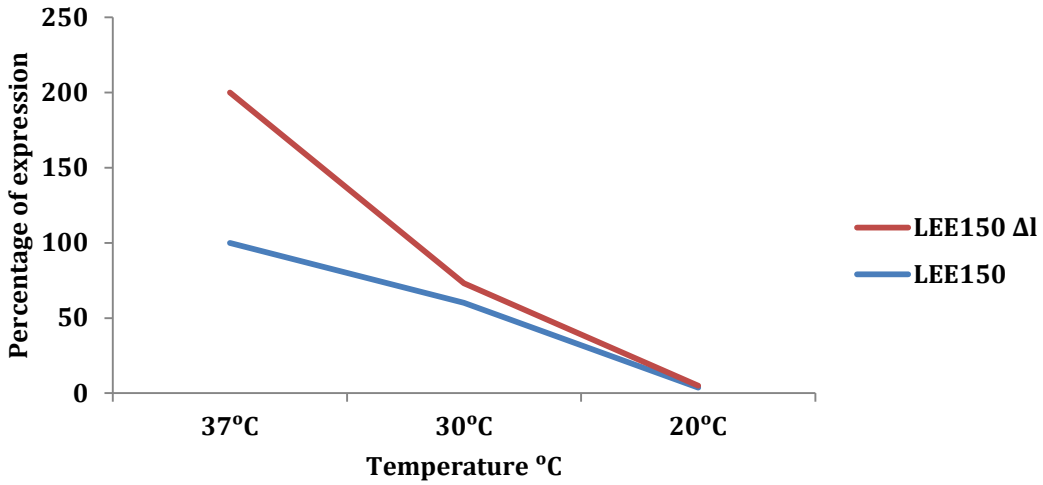


Fig. 4.8: Maximal level of expression of the LEE150 and derivative.
This linear chart shows the maximal level of expression from LEE150 and derivative, where the level of expression from LEE150 was set to 100%, at 37° C, and the rest were related to it.

A. Shows the level of expression without GrlA.

B. Shows the level of expression with GrlA.

positive impact on the expression of the mini ORF. It is believed that these mini ORFs work as a device that sets levels of *ler* gene expression and these mini ORFs are conserved in all pathogenic *E. coli* containing the *LEE* pathogenicity island, as well as in *Citrobacter rodentium*.

The level of *ler* gene expression is also regulated by the Shine-Dalgarno sequence as point mutations in LEE150 SD-sequence, altered Ler expression. In addition, the impact of deleting the leader sequence on overall bacterial growth in different growing temperatures has revealed that it might contain a binding site for a negative regulator(s).

4.2 Regulation at hybrid promoters

In previous work, it was found that the *LEE1* promoter of EHEC is defective because the spacer between the promoter -10 and -35 region is not optimal, and that the GrlA, LEE-encoded regulator, compensates for this defect by binding to the spacer, thereby activating the *LEE 1* promoter by an unusual mechanism (Islam *et al.*, 2011). However, there are other transcription factors, which have major and global roles in the regulation of various transcription units, such as cAMP receptor protein (CRP). CRP is a well-characterised transcription factor known to activate transcription at target promoters by binding to upstream sequences of the promoter as a Class I transcription factor. Alternatively, the transcription can be initiated when the CRP site overlaps with part of the RNAP subunits at the -35 elements, αCTD**,** as Class II simple activation. Finally, to activate the transcription initiation simply, a transcription factor binds to the spacer region between the -10 and -35 elements of the promoter, and bends the region, making it suitable for the RNAP to bind and for the transcription to be initiated known as conformational change transcription factor. This class of activation is referred to as conformational change and GrlA is an example of such a transcription factor (Browning and Busby, 2004).

Here, in addition to looking at how the leader sequence of *ler* was regulated, how the *LEE1* promoter was regulated was a matter of interest. In this work, whether CRP could activate synergistically with GrlA at the *LEE1* operon promoter was investigated. It was found that GrlA and CRP do not activate synergistically at a

hybrid promoter, but, rather, activate independently. In addition, the hybrid LEE1 promoter was found to be activated mainly by CRP, not GrlA. When CRP was not present, however, GrlA was able to gain its regulatory role and activate the LEE1 promoter.

4.2.1 CRP is the major activator at a hybrid promoter

The activation of transcription at promoters can be by three simple activating transcription factors which are classified as Class I, Class II, and conformational change transcription factors (Browning and Busby, 2004). It is also evident that transcription initiation can be activated by Class I transcription factors, or by both Class I and Class II transcription factors. However, there was no documented case of transcription initiation activation by Class I and conformational change transcription factors. To address this, a hybrid promoter at which both GrlA, as a conformational change transcription factor, and CRP, as a Class I transcription factor, bind was constructed to find if both transcription factors can bind and transcription can be initiated.

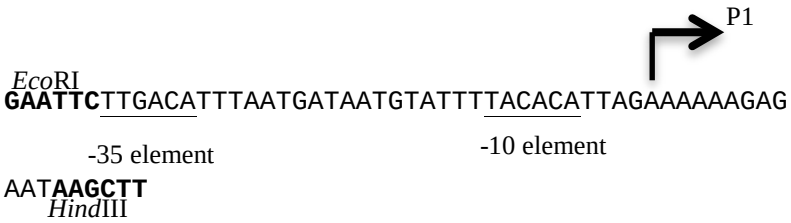
To do this, the shortest fragment of the LEE1 promoter, LEE20-203, was picked, which runs from -203 to -158 upstream to the transcription start site, and is believed to contain the potential binding site of GrlA, (Islam *et al.*, 2011). A CRP binding site

was then cloned upstream of the -35 element of LEE20-203 promoter fragment of the LEE1 promoter at position -61.5 (Fig. 4.9A). This was constructed by designing a primer that contained the CRP binding site upstream to -35 element of the promoter. After that, the new fragment (GS100) was cloned into pRW224, a *lacZ* expression vector, as transcription fusion. Then, GS100/pRW224 was transformed into M182 *E. coli* strains with GrlA (positive regulator of the LEE1 promoter), which was cloned into a pACYC184 vector, or with an empty vector pACYC184/ΔHN. Then, β-galactosidase assays were carried out to measure the level of expression from the *lacZ* fused to pRW224.

As shown in Fig. 4.9B1, inserting a CRP site upstream of -35 elements of the LEE20-203 DNA fragment of the LEE1 promoter enhanced the promoter strength approximately fourfold, compared to the parent fragment's (LEE20-203) activity. However, the GS100 promoter appeared to have higher activity when GrlA was available compared to the LEE20-203 DNA fragment. Fold of activation was still lower than LEE20-203 promoter's, which showed more than three fold, enhanced promoter activity with GrlA. This suggests that the CRP could be the major activator of the GS100 promoter and GrlA was only able to add some to the promoter activity.

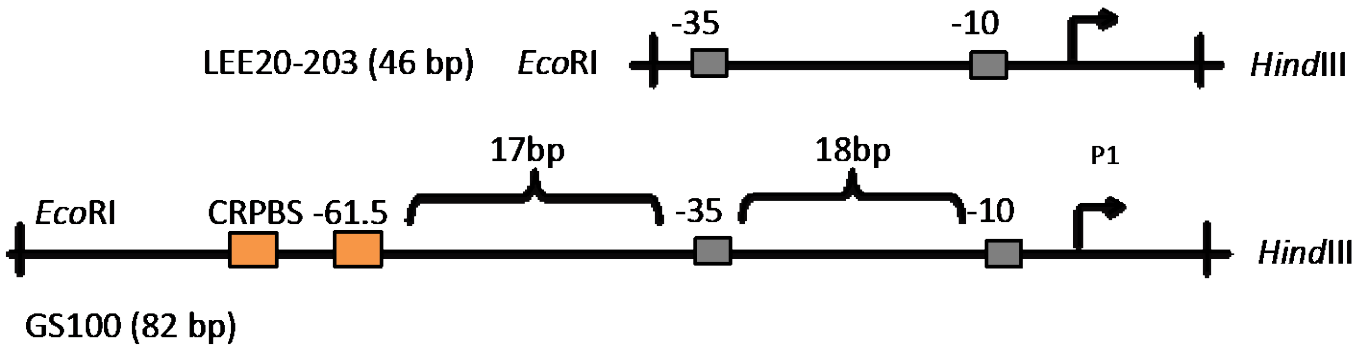
A

LEE20-203 (46 bp)

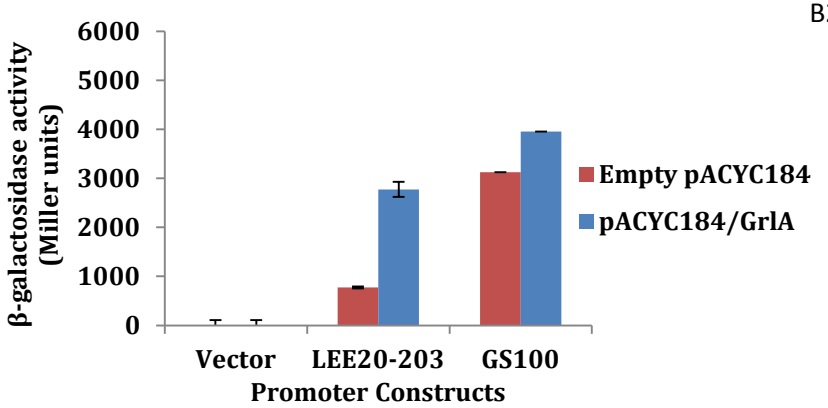


GS100 (82 bp)

GAATTCAAATGTGATGTACATCACATGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAATAAGCTT



B1



B2

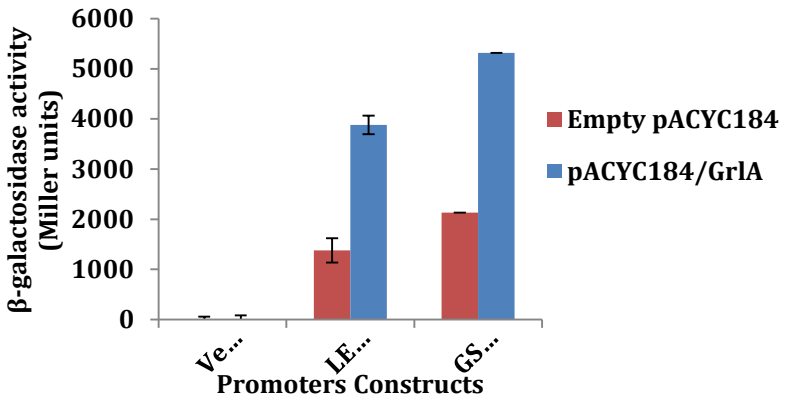


Fig. 4.9: Activation of expression from the GS100 promoter.
A. Sequence and a schematic representation of the LEE20-203 and GS100 fragments.

B1. Measured β-galactosidase expression in M182 strains, each containing pRW224 carrying a LEE20-203::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

B2. Measured β-galactosidase expression in M182 *Δcrp* strains, each containing pRW224 carrying a LEE20-203::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

Next, the aforementioned LEE20-203 and GS100 plasmid constructs were transformed into M182 *E. coli* Δ*crp*. Additionally, pACYC184/GrlA and pACYC184/ΔHN were transformed into the same strains. The activity of the promoter was measured by β-galactosidase assay. GrlA was able to bind to the GS100 promoter and activate it, more than when it was tested in *E. coli crp+*. The promoter activity was more than twice that of the promoter activity without GrlA (Fig. 4.9B2).

4.2.2 Regulation of expression from the GS 100 promoter and derivatives in *E. coli crp+* and Δ*crp* strains

It is clear that CRP and GrlA functions are not synergistic. Disruption of the -10 element of GS100 was carried out to test its effect on the promoter by introducing a mutation in the second base. It was then labelled GS100(-9C). In a separate promoter construct, the CRP binding site was disrupted while not altering the core promoter elements, and the construct was labelled GS101 (Fig. 4.10A). The constructs were then cloned into a pRW224 *lacZ* expression vector, and transformed into M182 *E. coli +crp* and Δ*crp* strains. Additionally, pACYC184/GrlA and pACYC184/ΔHN were transformed into the same strains. The promoters' activities in trans were measured by β-galactosidase assay, as a product of *lacZ*.

In Fig. 4.10B1, the level of expression from GS100 and derivatives is shown .A point mutation in the -10 element of GS100-(9C) promoter almost killed promoter activity in the +*crp* strains. Furthermore, a mutation in the CRP binding site caused reduction

A

GS100 (82 bp)

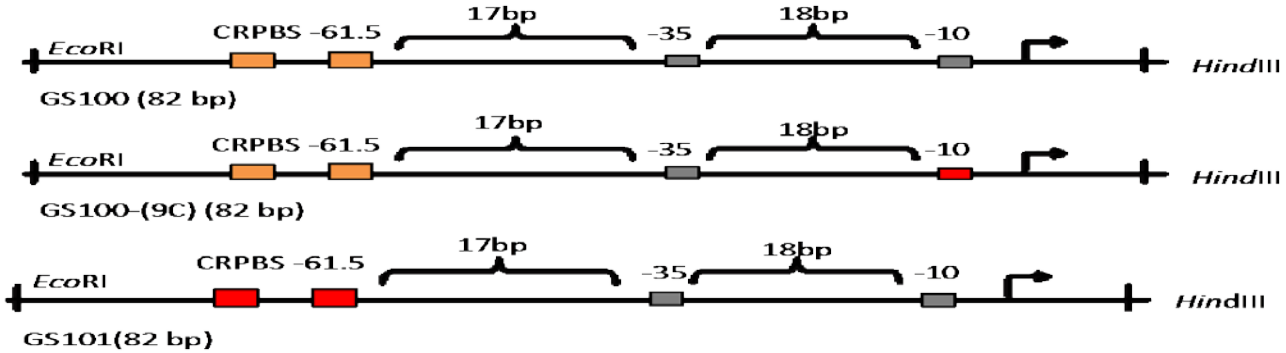
GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS100 (-9C) (82 bp)

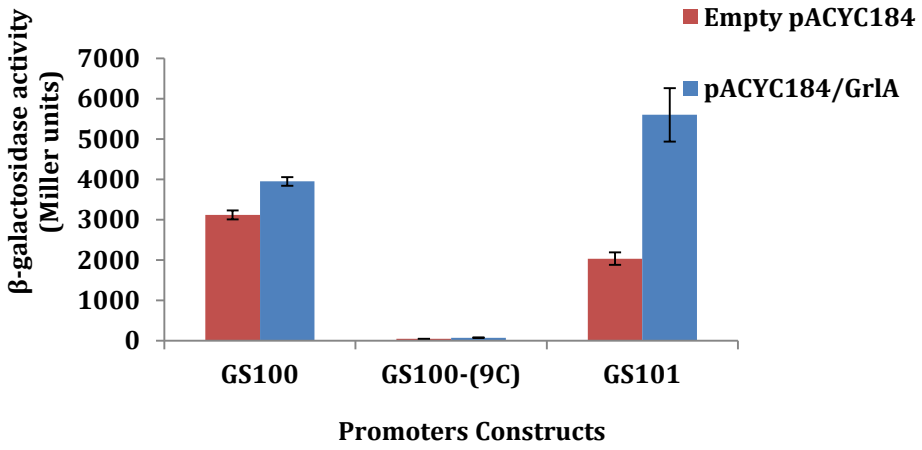
GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTT**C**CACATTAGAAAAAAGAGAAT**AAGCTT**

GS101 (82 bp)

GAATTCAAA**TGTCA**TGTACA**TGACA**TGGATCCAGATCAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**



B1



B2

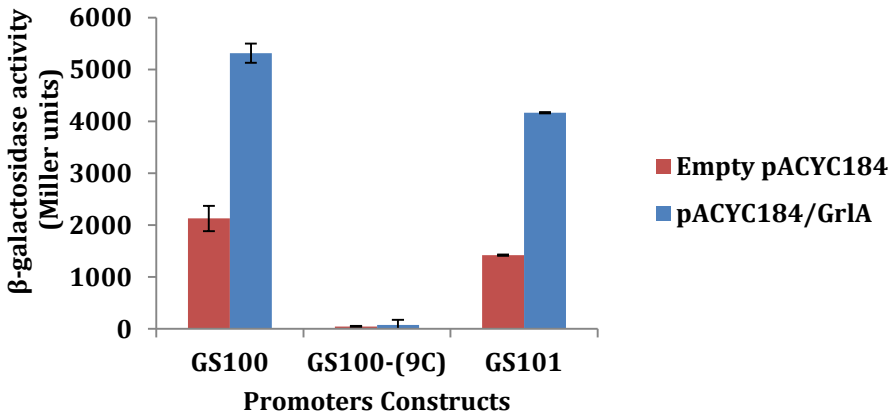


Fig. 4.10: Activation of expression from the GS100 promoter and derivatives.
A. Sequence and a schematic representation of the GS 100 and derivatives.

B1. Measured β-galactosidase expression in M182 strains, each containing pRW224 carrying a GS100::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

B2. Measured β-galactosidase expression in M182*Δcrp* strains, each containing pRW224 carrying a GS100::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeat

in the promoter activity, which was then enhanced in the presence of GrlA. This suggests the importance of the second base in the -10 element for the promoter function. Even when the CRP site was disrupted, GrlA was able to carry out the activation of the transcription initiation in this case. In Fig. 4.10B2, the constructs were transformed into M182 *E. coli* Δ*crp* strains and assayed. As shown, the level of expression from the GS100(-9C) and GS101 decreased similarly to Fig. 4.10B2.

4.2.3 Misplacing the CRP binding site results in GrlA-dependent promoter

In the GS100 promoter fragment, the spacer region between the -35 element and the CRP binding site is 17 bp. Therefore, the spacer was reduced to see if misplacing the CRP binding site affects the level of expression (Fig. 4.11A).

To do this, GS102 promoter was constructed, where the CRP binding site was cloned upstream of the -35 element of the LEE1 promoter at position -55.5. Then the promoter fragment was cloned into a pRW224 *lacZ* expression vector. GS102/pRW224 was transformed into M182 *E. coli*. In addition, GrlA was cloned into pACYC184 then it was transformed into M182 *E. coli* as well as an empty vector pACYC184/HN. β-galactosidase assay was then carried out.

The level of expression from the GS102 promoter was higher than the LEE20-203, but lower than the GS100 (Fig. 4.11B1). This shows that reducing the spacer region between the -35 element and the CRP binding site reduces the promoter activity. However, when GrlA was present, the promoter was activated 2.3-fold. This shows

A

LEE20-203 (46 bp)

P1

*Eco*RI

GAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAG

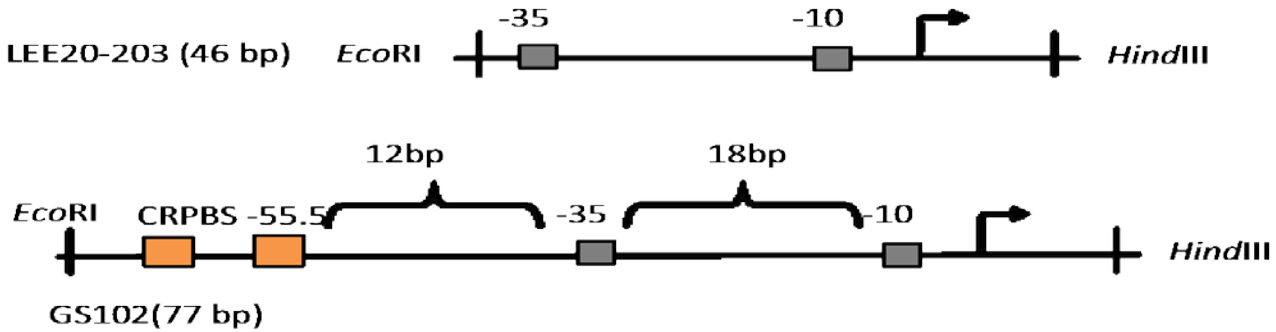
-35 element -10 element

AAT**AAGCTT**

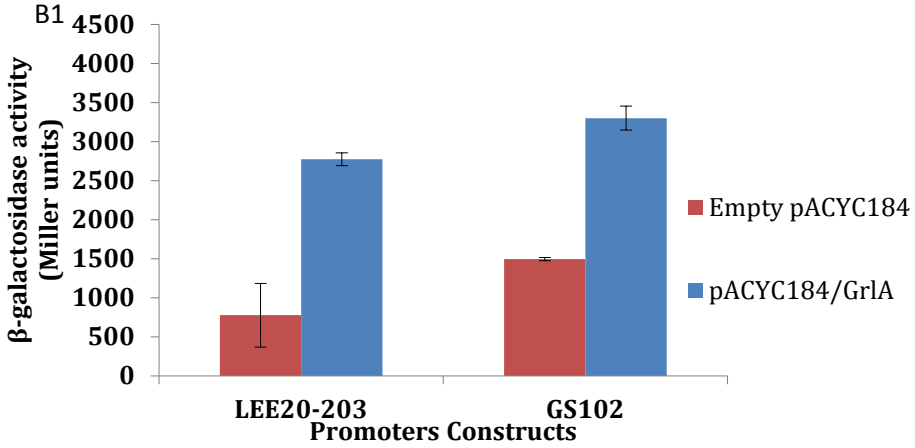
*Hind*III

GS102 (77 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCAGATCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**



B1



B2

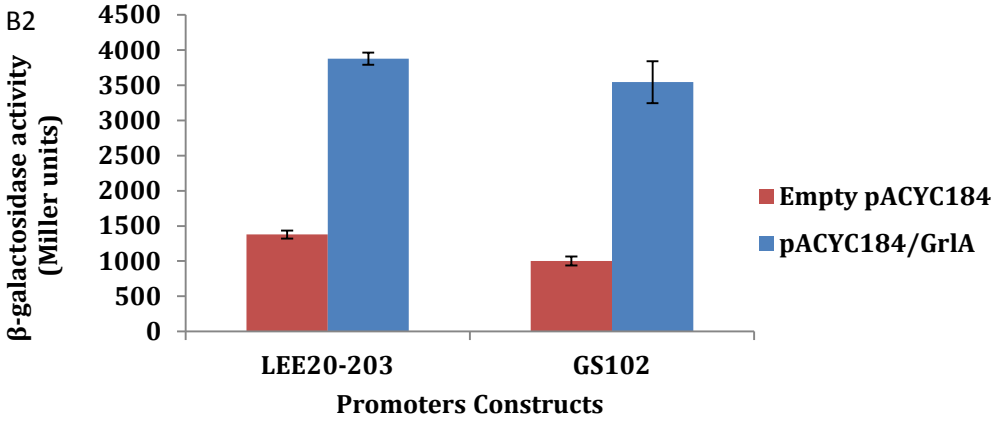


Fig. 4.11: Activation of expression from the GS102 promoter.
A. Sequence and a schematic representation of the LEE20-203 and GS102 fragments.

B1. Measured β-galactosidase expression in M182 strains, each containing pRW224 carrying a GS102::*lac* fusion. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

B2. Measured β-galactosidase expression in M182*Δcrp* strains, each containing pRW224 carrying a GS102::*lac* fusion. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

that in this case the function of CRP and GrlA could be synergistic at this promoter (Fig. 4.11B1). When the constructs were tested, in *E. coli* Δ*crp,* the level of expression from the GS 102 was reduced due to absence of CRP, but when GrlA was supplied, the level of expression increased 3.5 times (Fig. 4.11B2).

4.2.4 Misplacing the CRP in another position results in slightly GrlA-dependent promoter

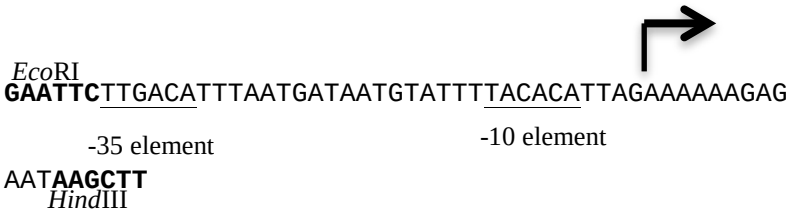
A new construct was created in which a CRP binding site was cloned upstream of the -35 element of the LEE1 promoter at position -51.5 (Fig. 4.12A). It was then cloned into pRW224 and transformed into M182 *E. coli*, as well as pACYC184/GrlA and pACYC184/ΔHN.

Then, a β-galactosidase assay was carried out. In the *E. coli +crp*, GrlA was able to bind and add less activity to the promoter compared to LEE20-203. However, in *E. coli* Δ*crp* (Fig. 4.12B1, and Fig. 4.12B2) GrlA was able to cause more fold of activation than in *E. coli +crp* strain. The reason could be that the two activators do not activate synergistically. Hence, when CRP was absent, GrlA fold of activation was higher.

A point mutation was then made in the -10 element of the GS103 promoter and tested in the two background strains (Fig. 4.13A). In the *+crp* strain, a mutation in the -10 element of GS103 was not able to reduce the promoter activity: the promoter was still active regardless of the presence of GrlA (Fig. 4.13B1). However, when *E. coli* Δ*crp* strains were tested, the GS103 promoter was not active (Fig. 4.13B2).

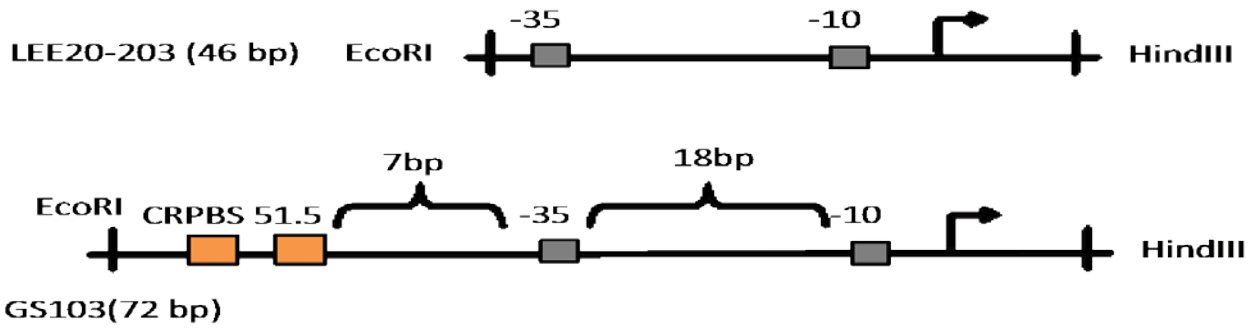
LEE20-203 (46bp)

A

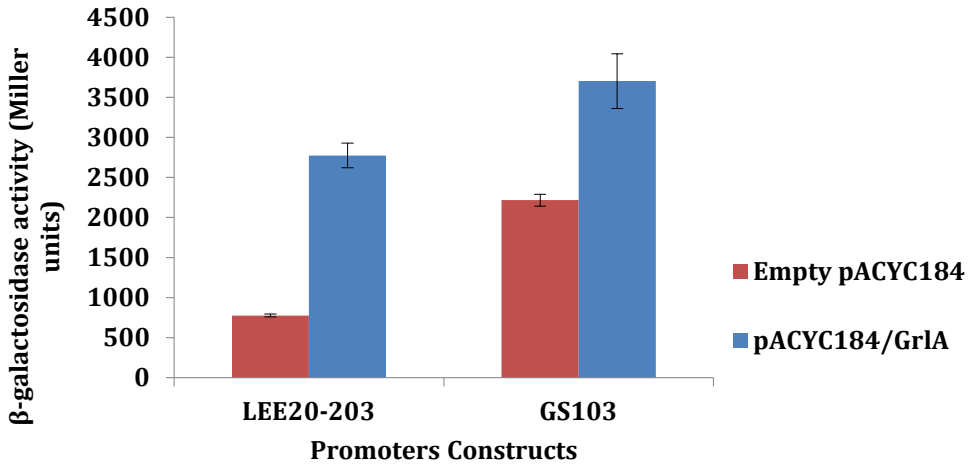


GS103 (72 bp)

GAATTCAAATGTGATGTACATCACATGGATCCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAATAAGCTT



B1



B2

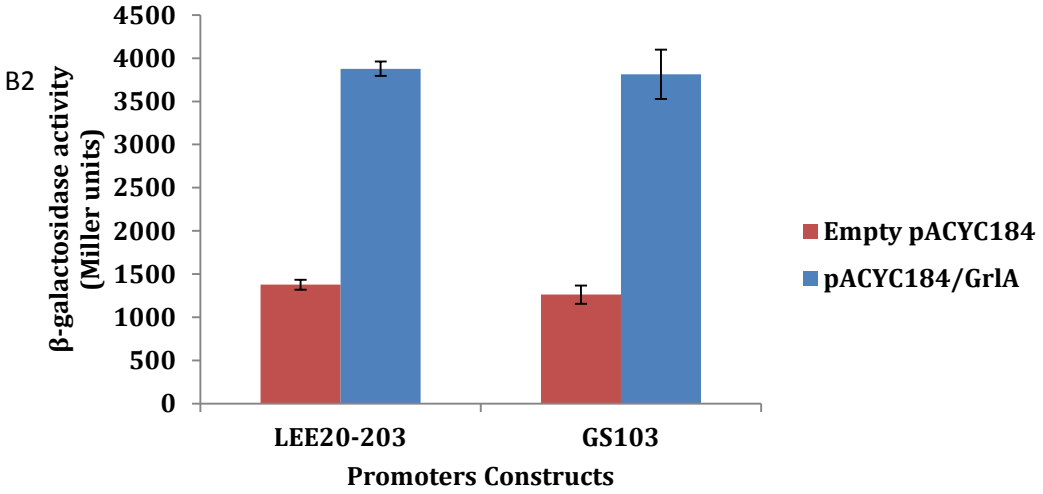


Fig. 4.12: Activation of expression from the GS103 promoter.
A. Sequence and a schematic representation of the LEE20-203 and GS103 fragments.

B1. Measured β-galactosidase expression in M182 strains, each containing pRW224 carrying a GS103::*lac* fusion. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

B2. Measured β-galactosidase expression in M182*Δcrp* strains each containing pRW224 carrying a GS103::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

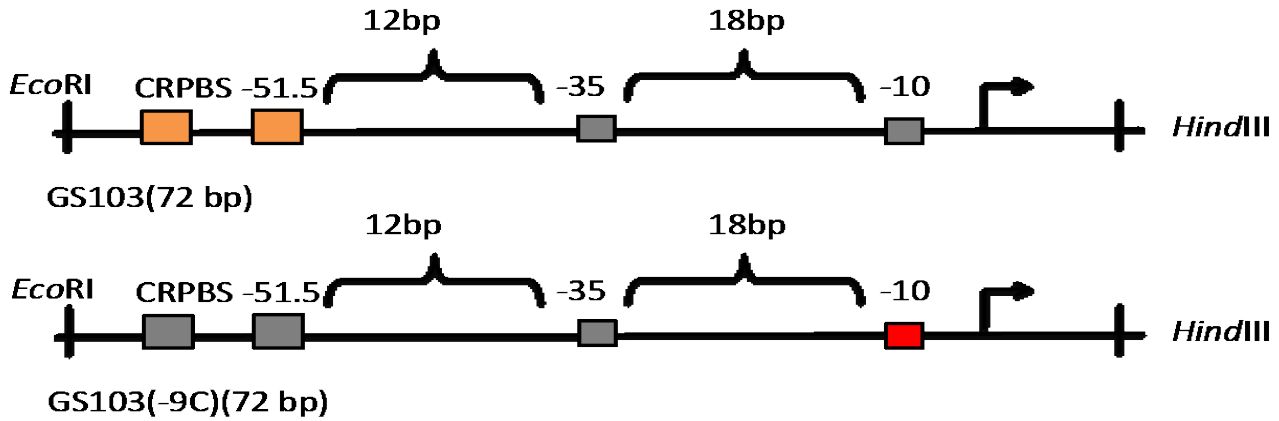
A

GS103 (72 bp)

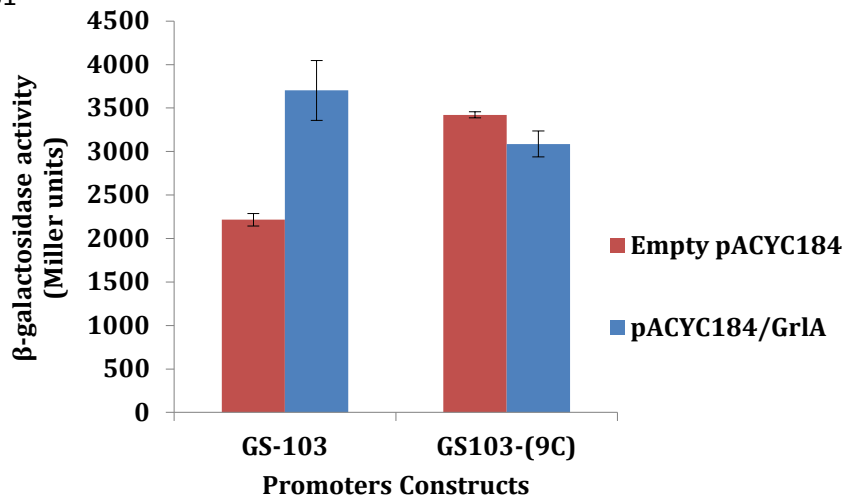
GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAGAGAAT**AAGCTT**

GS103 (-9C) (72 bp)

GAATTCAAA**TGTGA**TGTACA**TCACA**TGGATCCTTGACATTTAATGATAATGTATTTT**C**CACATTAGAAAAAAGAGAAT**AAGCTT**



B1



B2

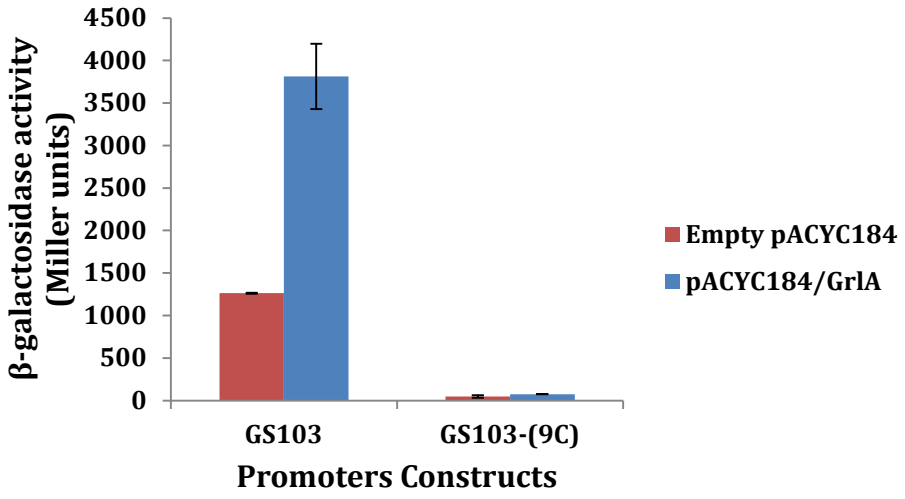


Fig. 4.13: Activation of expression from the GS103 and GS103-(9C) promoter derivatives.
A. Sequence and a schematic representation of the GS103 and derivative fragments.

B1. Measured β-galactosidase expression in M182 strains, each containing pRW224 carrying a GS103::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~ 0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

B2. Measured β-galactosidase expression in M182*Δcrp* strains, each containing pRW224 carrying a GS103::*lac* and derivative fusions. Measurements were made after growing cells in LB medium at 37° C to an optical density of ~0.5 at 650 nm. Error bars represent the standard errors of three independent repeats.

4.2.5 Conclusion

The results of this study imply that LEE20-203, which is a shortest DNA fragment of the *LEE1* promoter, is up-regulated by GrlA. This result supports previous findings on the positive impact of GrlA on *LEE1* operon in a manner independent from GrlR, since the background strain used in this study was M182 (*E. coli* K-12), which means the absence of virulence factors including GrlA (Huang and Syu, 2008; Barba *et al.*, 2005).

In contrast to the previous result, when a CRP site was cloned upstream of the LEE20-203 -35 element, at different positions, a different effect was seen. When the CRP site was located at poistion -61.5 upstream of the promoter (GS100), promoter activity was significantly enhanced compared to the LEE20-203 promoter (Fig. 4.9).

By contrast, when GrlA was present, it could add some activity to the GS100 promoter activity as it did when it was present for the LEE20-203 promoter. It is believed that the enhanced promoter activity for GS100 was derived mainly from CRP binding upstream because, and GrlA added some activity to the promoter. It is thought that when CRP was located at position -61.5, GrlA might bind and add little activity but the predominant activator in this case could be CRP. To confirm this result, the same constructs were transformed into *E. coli* Δ*crp* strains and β-galactosidase assays were carried out. As the results of the GS100 promoter showed (Fig. 4.9), when CRP was not present, GrlA could bind and, significantly, activate the

promoter. Thus, this result gives a very similar pattern to the LEE20-203 promoter when GrlA activates it.

A point mutation was made in the -10 element of the GS100 promoter making the GS100-(9C) promoter. As the result indicated (Fig. 4.10), the promoter activity was lost in spite of the absence of GrlA. In addition, a point mutation was made in the CRP binding site, creating the GS101 promoter. As stated, the promoter activity was negatively affected; however, the promoter activity of GS101 was still greater than that of LEE20-203, which might mean that the mutation in the CRP binding site did prevented CRP from binding entirely. This meant that there was still some binding but it was not as efficient as, for example, in the case of the GS100 promoter. The same constructs were also transformed into *E. coli* Δ*crp* strains and assayed, giving similar results as in *E. coli crp*+ strains (Fig. 4.10).

A second case was misplacing the CRP site to position -55.5, as indicated in Fig. 4.11, to make the GS102 promoter. This promoter produced a similar type of activity to the LEE20-203 it was activated significantly in the presence of GrlA. This is thought to mean that when the CRP binding site was misplaced, it could add little activity to the promoter compared to GS100. So, in this case, GrlA was able to bind and activate the promoter and synergised with CRP. Clearly, the activity of GS102 is derived from GrlA binding. When the same construct was transformed into *E. coli* Δ*crp* strains and assayed (Fig. 4.9), it gave a similar result as when CRP was present. To sum up, misplacing the CRP binding site caused CRP to be unable to bind and the predominant activator in this case was GrlA.

Finally, when CRP was positioned at position -51.5, to make the GS103 promoter, and assayed (Fig. 4.12), at first it gave a similar result to GS100: CRP might be the

major activator of the promoter. The promoter activity in Δ*crp* is 1250 Mu, but in *crp*+ activity ~ 2200 Mu. When GS100 was assayed in *E. coli* Δ*crp* strains, GrlA was able to bind and activate the promoter. A similar case was found for GS103 in *E. coli* Δ*crp* strains (Fig. 4.12). However, it was believed that the CRP was picking an ectopic promoter rather than the -10 element of the GS103 promoter. It could be due to that CRP was not able to activate at -51.5. To confirm this, a point mutation in the -10 element of GS103 was made to create GS103(-9C), and it was assayed in *E. coli* *crp*+ and Δ*crp* strains (Fig. 4.13). As expected, CRP was picking an ectopic promoter because, when the -10 element of P1 was not active, the promoter was still active in the presence of CRP; however, the promoter was dead when assayed in the *E.coli* Δ*crp* strain.

This experiment was designed to understand how a combination of a Class I and an activator that binds to the spacer region between -10 and -35 (GrlA) activator would act at a given promoter, in this case, the LEE1 promoter. It is known that CRP is a conventional transcription factor that binds upstream of the -35 element of a promoter and causes RNAP to bind and initiate transcription of a gene. GrlA, however, is thought to be acting in a more complicated manner. It acts as the MerR family of transcription factors, which bind in the spacer region between -10 and -35 elements, causing a twist in the promoter. The spacer between -10 and -35 is not optimal for the case of LEE1, having 18 base pairs instead of the optimal 17 base pairs, causing the promoter to be fairly weak. However, with GrlA binding, RNAP is recruited and transcription initiation begins (Islam *et al.*, 2010).

In this experiment, it was evident that CRP was the predominant activator and GrlA could not bind and activate the promoter unless CRP was absent, i.e. in Δ*crp* strains. Therefore, in this experiment, CRP and GrlA do not activate the promoter synergistically; instead, they do so independently. It was also evident that, when CRP was misplaced, it could not function and, in this case, GrlA was able to bind and activate the promoter. Finally, it was also found that CRP can pick an ectopic promoter, which could be identified by using a primer extension assay, other than the GS103 promoter; because the LEE1 promoter is especially AT rich

4.3 Summary

This study has shown how the leader sequence is regulated and how regulation takes place at a hybrid promoter. The regulation of *LEE* was demonstrated to complicated and involves a number of factors, either LEE or non-LEE encoded regulators. Previously, in Chapter Three, the role of attachment in triggering the expression of the *LEE* operons was shown. So attachment, minigenes sequence, ribosomal binding sites sequences, the leader sequence, and CRP influences the expression from the *LEE* operons. There are other factors that could be covered in order to understand better how this multi-machinery system works, such as the role of flagella, the role of PerC homologues, the role of quorum sensing, etc.

Chapter five:
Final summary

5.1 Adaptation and plasticity in the bacterial world

It was believed that RNA played a dominant role in the first living cells, as it served both to carry genetic information and to supply the enzymatic activities to make new RNA, and to reproduce itself (Eigen *et al.,* 1981). Early replicative machineries appeared to have been led by RNA activity, until DNA appeared, having evolved by reverse transcribing RNA. The crucial feature of DNA was the replacement of ribose by deoxyribose and uracil by thymine. Since then, DNA has served as the carrier for the genetic material and encodes all the RNA functions, holding the transcribed message for translation. Yet, RNA still possesses some of its original activities (Guerrier *at al.*, 1983; Gilbert, 1986; Zaug and Cech, 1986).

There are a number of approaches have been used to study the origin of life on earth each with different pros and cons. To gain a better understanding of how living creatures developed, comparative phylogenic studies have been used. This is accompanied by integrating insights from palaeontology and developmental biology, and then the information provided by molecular and cellular biology (Cavalier-Smith, 2006). Comparative polygenetic studies are useful in that they provide detailed and accurate accounts of the evolutionary relationships across the entire tree of life, and they also point to ancestral states for most of the organisms on the tree. However, phylogenic studies required the information provided from paleontological studies in order to give time scale. Otherwise, information gained without these integrative studies would lack direct knowledge of totally vanished groups, and recently molecular and cellular biological approaches have proved to be important as they are mostly in agreement with the fossil record.

The first living cells were believed to be formed of double membrane layers and are called Negibacteria and are believed to have appeared 3.5 giga-years ago (Fig. 5.1). The cells containing two lipid bilayers were called Eubacteria. Subsequent membrane changes and incorporation of other new features led to the evolution of new cells. Thus, the lipopolysaccharide layer and formation of glycobacteria developed 2.8 giga-years ago. Then the outer membrane became thicker to enable the living organism to colonize the soil. Then, 0.9 giga-years ago, the outer membrane was lost and the cells evolved to contain one membrane, and were then called unimembrana. Unimembrana evolved from the thickened murine-wall eubacteria into two environments. One, which persisted in high temperatures, a hyperthermophiliv arachebacteria, utilized the methane. The second was phototrophic, eukaryote bacteria. Both the eukaryotes and the arachebacteria are classified as neomura unimebrana (Cavalier-Smith, 2006).

It is believed that the great differences in the gene ranges of pathogenic bacteria are caused by the highly dynamic role of gene evolution (Koonin and Wolf, 2008). Genome evolution can be the result of a number of things, including increase in the genome size, or expansion. Interestingly, the expansion of genomes is not linear, hence, the size of genome and content varies among species within the same genus and furthermore, it is different between strains of the same species (Kaas *et al.*, 2012; Thompson *et al.*, 2013). So, the reasons for gene expansion could be due to HGT (horizontal gene transfer), gene duplication, de novo emergence of genes, or gene contraction by gene loss (Koonin *et al.*, 2001; Snel *et al.*, 2002; Mirkin *et al*., 2003). A number of methods were used to prove genome expansion, including, comparative

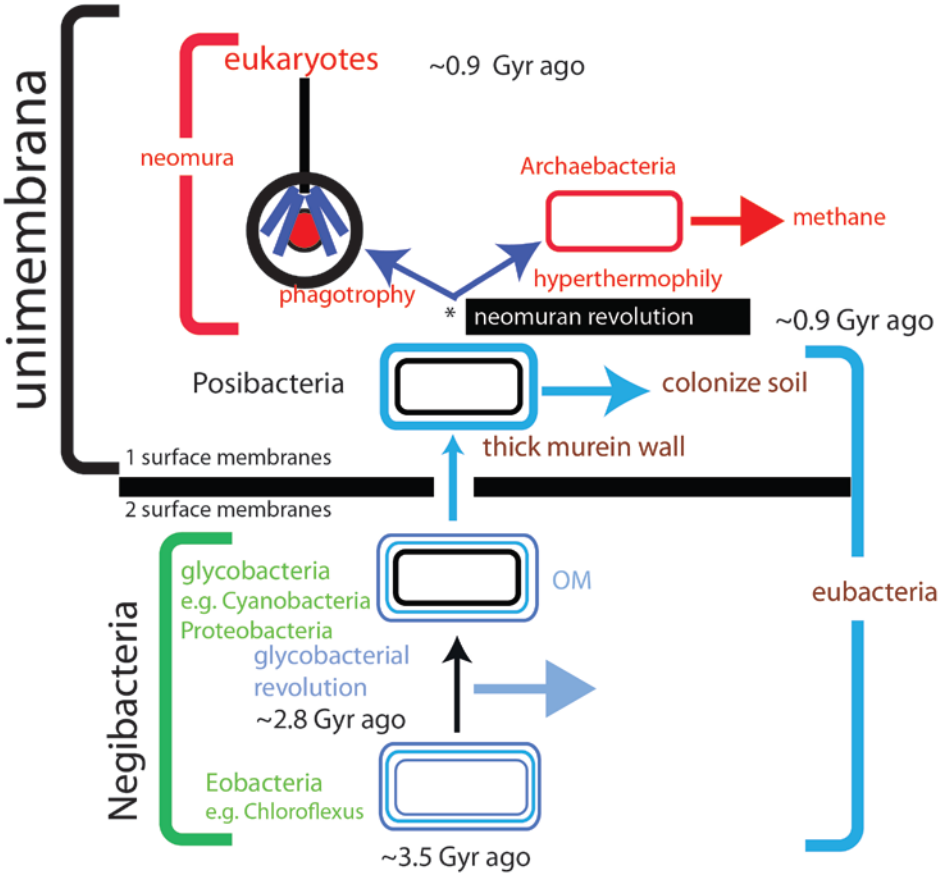


Fig. 5.1: Evolution pattern of prokaryotic cells. The figure represents time-scale revolutionary steps in the development of primitive cells from basic double-layered prokaryotes, into more complex single-layered eukaryotes (figure was redrawn from Cavalier-Smith, 2006).

genomic observations, and explicit evolutionary reconstruction by using maximum parsimony (Snel *et al.,* 2002). Among reasons for genome expansion gene loss is believed to be most prominent. It is streamlining, more uniform, and considered to be the default evolutionary option of organisms (Snel *et al*., 2002; Richards *et al*., 2014).

The fact that the genome size of bacteria is larger than the size of the bacteria itself makes the compaction of the chromosome very important (Browning *et al.,* 2010). NAPs are mainly characterized in *E. coli*. Early studies revealed the measured binding affinities of the NAPs to DNA, and identified their preferred targets within the DNA. The levels of NAPs in different growth phases were also quantified (Azam and Ishihama, 1999; Azam *et al.,* 1999). The knowledge of NAPs and their role was gained by the chromatin immunoprecipitation method and by whole genome sequencing (Grainger and Busby, 2008). Both methods help in showing the distribution of NAPs across the chromosome. So the NAPs can compact the genome by bending, bridging, wrapping and clustering after binding to DNA (Luisterburg *et al.,* 2008). Here it can be seen that the bacterial chromosome is highly compacted and the ways of compaction are highly regulated. Therefore, the expression from genes in this highly folded and compacted chromosome is dependent on the dosage and the positions of the genes to be transcribed (Bryant *et al.,* 2014).

Hence, bacteria developed a natural competence for accepting foreign DNA, which can be taken by conjugation, transduction, or transformation. Foreign DNA can be up taken from either the environment or from other strains of species (Seitz and Blokesch, 2013). Natural competence is also a highly regulated mechanism, and it

results in evolution of the organism (Chen *et al.*, 2005; Claverys *et al.*, 2006). It also helps the organism to adapt to new living environments. The most recent example of natural competence is the German outbreak (Seitz and Blokesch, 2013). Studies revealed that the outbreak was caused by an enteroaggregative *E. coli* strain, O104:H4, that acquired Shiga-like toxins that are normally gained by EHEC (Bielaszewska *et al.*, 2011).

So here, horizontal gene transfer reflects both genome expansion and the natural competence of bacteria. Horizontally acquired elements are up taken by the organism through transformation, hence, mobile genetic elements, such as plasmids, are not required for the DNA to be transformed. HGT is counted as a major factor in the evolution of prokaryotes and a key source of revolution and adaptation to new atmospheres and lifestyles (Seitz *et al.*, 2013). Horizontally acquired genes can be shaped by both the habitat of living organisms and its niches (Smillie *et al*., 2011; Polz *et al*., 2013). HGT occur between species that share similar living space (Boussau *et al.*, 2008), the frequency of this process is high, and such species as have higher rates of sharing genes are known as being connected by highways of gene sharing (Beiko *et al.*, 2005). The horizontally acquired genes or foreign elements are characterized by a high content of AT, more than the ancestral genome (Ochman *et al*., 2000). Consequently, they serve as favourite sites for silencing effects mediated by the nucleoid-associated proteins, such as H-NS (Lucchini *et al.*, 2006). Hence, several regulators work by relieving the silencing effect of H-NS, in order to facilitate gene expression in response to environmental signals (Stoebel et al., 2008).

5.2 Transcriptional regulation

The exact time when transcription regulation was evolved is as yet unknown and the bacterial cells are believed to have had no regulation strategy at first. The bacterial cells were empty cells “chassis” without regulatory networking, and the current modern bacteria are derived form continuous evolution and adaptation to different niches (Pallen, 2009). However, with adaptation to different niches, bacterial cells evolved transcription factors to regulate the gene expression. The evolution of the transcription regulation network is believed to have happened rapidly. Additionally, their evolution is independent of their target genes (Madan Babu *at al.,* 2006). This might be a cause of the natural selection that leads to loss or gain of transcription factors and their target genes independently. Despite this, some organisms have evolved orthologous transcription factors, which have DNA-binding properties. Specific transcription factors have evolved that have DNA-binding properties but these are not orthologous. These transcription factors were gained as the size of the genome increased, hence the need for more transcription factors (Aravind *et al.,* 2005). So these different transcription factors regulate genes that respond differently to environmental signals. However, organisms that have similar life pattern, tends to have homologous regulatory interactions (Babu *et al.,* 2006).

Transcription factors are derived from NAPs; hence they bind to elements on the DNA regulatory regions and also have allosteric metabolic interaction. In a simple system, where the bacteria interact only with the surrounding environment, i.e., before the presence of eukaryotes, transcription factors responded to ion signals such as phosphate or oxygen. The environmental signalling was transducted to the TFs

through metabolites or covalent modification. The results of signal transduction from stimuli are either expression or repression of genes (Pollack et al., 2002). According to Martinez-Antonio, there are seven regulators in *E. coli* identified as global transcription factor regulators: CRP, IHF, FNR, Fis, ArcA, Lrp, and Hns. These regulators regulate expression in response to cAMP, regulation of energy production, phosphate level, regulating nutritional state, and modelling of the DNA topology to regulate transcription (Busby and Kolb, 1996; Landgraft *et al.*, 1996; Hatfield and Benham, 2002). Here, global regulators work to respond to signals from the environment and as bacteria were superior to mammals, their systems were simpler and worked to enable them live in the environment. However, with the arrival of mammals new interactions were established between bacteria and mammals. This may result in simple symbiosis or in pathogenesis. Gain or loss of genes might determine the pathogenicity or an organism(s). Either way, regulation of gene expression is more complicated due to this relationship with eukaryotes. Hence, specific regulators developed that regulate specific genes, and which are mainly gained through horizontal gene transfer. However, global transcription factors still regulate gene expression from the pathogenicity genes that are usually present on islands.

5.3 Global and bespoke regulators

Then comes the question of the role of the housekeeping regulators in regulating the acquired pathogenicity island. It is known that the *LEE* PI is regulated by a number of housekeeping regulators, such as, IHF, FIS, H-NS, quorum sensing, GadXE, etc. (Kaper *et al*., 2004). These regulators act either positively or negatively to regulate

the expression from the *LEE* PI, by either directly interacting with the DNA elements, or by interacting with *LEE* cis regulators. For example, through GrlA, these regulators might transmit the signals through to GrlA. Despite the presence of housekeeping regulators, *LEE* PI still encodes for cis regulators, indicating their beneficial advantages for natural selection and adaptation to new environments. Frankly, it seems logic that a newly acquired PI encodes for its own regulators.

EPEC and EHEC are among the most important gastrointestinal pathogens causing bloody diarrhoea and other complications associated with acquiring virulence factors. The main acquired pathogenic feature in these organisms is the *LEE* PI. LEE Pathogenicity Island was horizontally acquired and incorporated within the bacterial chromosome. This island is characterized by high AT content, making it a target for H-NS silencing. Under suitable conditions H-NS silencing effect is relieved and the expression is initiated. As well as acquiring pathogenic genes on the *LEE, LEE* encodes for its own regulator. So pathogenic *E. coli* acquired both the genes and their regulators. Acquiring regulators is believed to have had a selective advantage for the pathogens as they aid in overcoming hostile aspects of the environment in the gut, leading to selection. Moreover, horizontally acquired regulators to adapt to cell physiology and environment, hence, these regulators regulate genes outside the pathogenicity island (Brown *et al.*, 2014). For example, GrlA, a horizontally acquired *LEE* regulator, regulates the flagger and haemolysin genes in EHEC. So, here lies the importance of having acquired regulators are well as genes on the pathogenicity island.

5.4 Promoter organization and activation

There are two mechanisms to activate the transcription initiation. One can be on the level of the promoter, and the other mechanism is on the level of the RNAP (Lee *et al.,* 2012). Activation of transcription initiation on the level of promoter, or promoter-centric activation that is used for adaptation can be achieved by recruiting proteins, which strengthen the promoter to attract the RNAP. This is done by supplying other functional proteins to the promoter, or it can be achieved by relieving the repressor effect on the promoter.

On the level of the RNAP, or RNAP-centric activation, that is used for life and death decisions, a factor(s) can bind to the RNAP to alter its promoter selection. At in the promoter-centric model, the activation of transcription initiation is achieved by the binding of transcription factor(s) to different locations on the promoter elements to help in recruiting the RNAP holoenzyme to the promoter and initiating transcription. Alternatively, the transcription initiation can be repressed by the binding of a repressor at a different location to the promoter elements to block the binding of transcription factors or to block the entry of the RNAP to the promoters' elements. In the RNA-centric model the promoter preference could be changed for the RNAP by simply alternating the sigma factors, thus, redirecting the RNAP promoters in response to environmental signalling. Additionally, RNAP can be modified in response to different metabolites. Hence, the transcription factors bind to the RNAP instead of to the promoter elements, and result in RNAP pre-recruitment (Lee *et al.*, 2012).

Transcription regulation can be done on either promoter-centric or RNAP-centric integration of promoter and the RNAP-centric model is also involved in regulation. In here, to activate the transcription initiation, interaction happens between TFs and σ^{70} domains of RNAP. An example of such interaction is seen at the bacteriophage λ P_{RM} promoter. Activation at this promoter is derived by binding of the cI dimer protein next to -35 elements of the promoter and by interaction of λ cI residues E34 D38 with residues R588 and R596 of domain 4 on σ^{70} subunit of RNAP (Li *et al.*, 1994). Other examples include the AraC family, FNR in *E. coli*, and bacteriophage Mu Mor protein. In these regulators TFs overlap with the -35 elements and interact with domain 4 on the σ^{70} subunit of RNAP.

Regulation of transcription initiation often requires integration of transcription factors. The regulation is controlled by the involvement of two transcription-factor activators at a promoter, an activator and a repressor, or the interaction of two repressors. For the interaction of two transcription-factor activators there could be interaction of class I with class I TFs, or class I with class II TFs. Transcription factors bind to both the RNA and promoter elements in order to initiate transcription. For example, transcription factors bind to a proximal site within the promoter overlapping with -35 as class II or upstream -35 elements as class I TFs, but the transcription is not initiated until second TFs make contact with αCTD as class II TFs. Then, RNAP could initiate transcription; this way is known as independent contact. Pet promoter is activated by this mechanism, as it requires both Fis and CRP to regulate transcription initiation (Rossiter *et al.*, 2011). Other integration of TFs example is at a given promoter two class I TFs bind co-operatively to activate

transcription. This mechanism is employed by MelR and CRP TFs at the *melAB* promoter in *E. coli* or with ToxR and TcpP in *V. cholerae* (Ebright and Busby, 1995; Wade *at al.,* 2001). This mechanism of synergy could be explained as the bacteria evolving to ensure their survival and mixing TFs at a promoter could be a way of doing it (Lee *et al.,* 2012). On other promoters, regulation is achieved by the interaction of activator and repressor TFs. Examples exist of such promoters where one TF can activate the transcription initiation, but due to the presence of repressor(s) upstream of the promoter, the TF could not activate. To activate the transcription initiation, another TF is needed to relieve the repression caused by repressor(s), and then the initial TF could initiate the transcription. This means of integrative regulation is known as anti-repression. *Nir* promoter is activated by FNR, however it is repressed by IHF and Fis and to relieve this repression NarL and NarP work to displace IHF, hence the promoter is activated (Browning *et al.,* 2004). Repositioning of TFs is another example of integrative regulation where TF binds upstream from its target and, for it to initiate transcription, another TF binds upstream and causes repositioning of the initial TF to regulate transcription initiation. CRP works to reposition MalT TF at *malE* promoter in *E. coli* (Brown *et al.,* 2003). In another case, a TF could bind to distance elements upstream from the promoter, and for it to contact the RNAP, another TF that causes bending of the DNA is required, so the initial TF is in contact with the RNAP holoenzyme by contact with a sigma subunit. At the *narG* promoter in *E. coli*, IHF binds upstream promoter elements to bend the DNA; hence TF comes into contact with an alternative sigma subunit and initiates transcription (Schroder *et al.,* 1993).

In this work, shown that in a semi-synthetic system transcription factors were employed and I provided evidence to understand how they co-operatively regulate the expression from a promoter. It is known that transcription activation can be caused by two transcription factors binding, either class I with class I or class I with class II transcription factors. However, there were no reports on how a conformational change transcription factor would activate with class I or class II transcription factors. For example, in *pet* in enteroaggregative *E. coli*, on *sat* on UPEC and on *sigA* promoters in *S. sonnei*. The requirements for both CRP and Fis TFs were studied on semi-synthetic promoters. CRP and Fis are global transcription factors and on a number of promoters, they are known to co-regulate, I showed that on a semi-synthetic promoter there is no co-regulation of promoter expression by a class1 transcription factor, which is CRP and a conformational change TF, which is GrlA.

5.5 What is special about Pathogenicity Islands?

Despite the fact that pathogenicity islands play a major role in bacterial pathogeneses (Fig. 5.2), and their evolution though horizontal gene transfer, which confers selective advantage to the organism, they are only needed for a short time. In my case, EHEC or EPEC use the Type III secretion system to inject toxins or effector molecules into the host cells only when required, namely, when the bacteria reach their destination. But in environments such as soil this system loses its usefulness, and it can be lost, as it has been gained.

The *LEE* pathogenicity island is silenced by H-NS and only during appropriate moments is this repression relieved. Ler, which is a global regulator for the *LEE* and H-NS homologue, is responsible for relieving this repression by counteracting H-NS. Ler, with GrlA, forms a positive feedback loop. The more Ler is expressed, the more GrlA is expressed, hence more activation of the expression from *ler*. Ler regulate *LEE1* operon negatively as it was found to form a loop of the DNA elements that then traps RNAP open complex process (Bhat *et al.*, 2014). GrlA activation of expression is due to response to signals. I showed that GrlA significantly activates expression from the *LEE* promoter after attachment to host cells. This indicates that EHEC or EPEC or a microbe in general knows when to switch on gene expression. And this is in agreement with the careful and tight regulation of the *LEE* pathogenicity island. I also showed a positive feedback loop between GrlA activation and the impact on attachment. I showed that with attachment GrlA activates expression that then results in more attachment to host cells. However, the mechanism of signal transduction to GrlA is as yet unknown. It could be some surface protein on the host cells or some chemicals that are present in the gut. Further studies into signal transduction will lead to a better understanding of how GrlA is triggered. Although pet promoter in Enteroaggregative *E. coli* is regulated by two housekeeping transcription factors, CRP and FNR, and they show co-regulation of pathogenicity, in my work, the regulation of the *LEE* pathogenicity island was mainly by two cistronic bespoke regulators, Ler and GrlA. Both Ler and GrlA turn on the entire *LEE* operon regulation. In a relative study

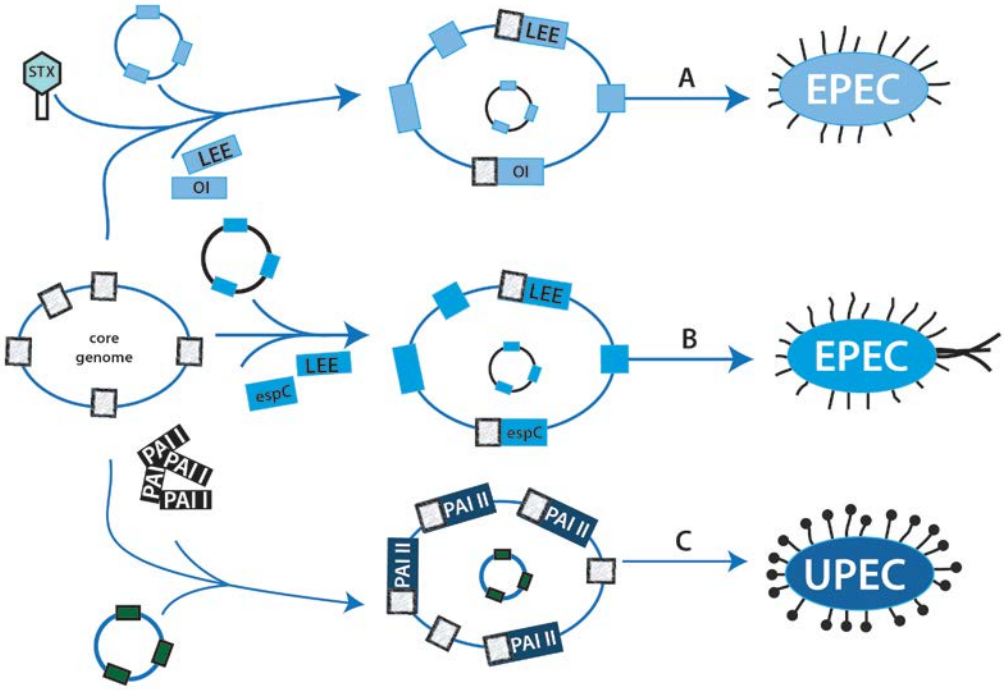


Fig. 5.2: Acquisition of Pathogenicity Island.
The figure represents how the pathogenic *E. coli*, including EHEC, EPEC, and UPEC acquire pathogenicity islands. As shown, horizontal gene transfer plays a major role in delivering the mobile genetic elements, Shiga-toxins, foreign plasmids, *LEE*, etc. to cells and as a result, transforming them into pathogenic organisms. Acquired genes are then incorporated into the core genome of the organism and use their own regulators as well as housekeeping regulators to regulate their gene expression (figure was redrwan from Schmidt and Hensel, 2004).

by Siryaporn *et al* 2014, where *P. aeruginosa*, which is one of the main opportunistic pathogens, the switch of pathogenesis, found to be induced after attachment to host cells. And this was demonstrated by production of GFP in contrast to non-adherent Strains. This increasingly, proved argument that attachment is crucial to switch on pathogenic features of microbes, in this case *E. coli*.

References

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181:** 6361-6370.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181:** 6361-6370.

Almiron, M., Link, A.J., Furlong, D., and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev* **6:** 2646-2654.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M., and Iyer, L.M. (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29:** 231-262.

Arthur, T.M., and Burgess, R.R. (1998) Localization of a sigma70 binding site on the N terminus of the Escherichia coli RNA polymerase beta' subunit. *J Biol Chem* **273:** 31381-31387.

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:** 33105-33113.

Baldini, M.M., Kaper, J.B., Levine, M.M., Candy, D.C., and Moon, H.W. (1983) Plasmid-mediated adhesion in enteropathogenic Escherichia coli. *J Pediatr Gastroenterol Nutr* **2:** 534-538.

Barba, J., Bustamante, V.H., Flores-Valdez, M.A., Deng, W., Finlay, B.B., and Puente, J.L. (2005) A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. *J Bacteriol* **187:** 7918-7930.

Bates, S.R., and Quake, S.R. (2014) Mapping of protein-protein interactions of E. coli RNA polymerase with microfluidic mechanical trapping. *PLoS One* **9:** e91542.

Beiko, R.G., Harlow, T.J., and Ragan, M.A. (2005) Highways of gene sharing in prokaryotes. *Proc Natl Acad Sci U S A* **102:** 14332-14337.

Berg, H.C. (2003) *E. coli in Motion*. United States: Springer-Verlag.
Bergholz, T.M., Wick, L.M., Qi, W., Riordan, J.T., Ouellette, L.M., and Whittam, T.S. (2007) Global transcriptional response of Escherichia coli O157:H7 to growth transitions in glucose minimal medium. *BMC Microbiol* **7:** 97.

Bhat, A., Shin, M., Jeong, J., Kim, H., Lim, H., Rhee, J.H., *et al.* (2014) DNA looping-dependent autorepression of LEE1 P1 promoters by Ler in enteropathogenic Escherichia coli (EPEC). *Proceedings of the National Academy of Sciences* **111:** E2586-E2595.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., *et al.* (2011) Characterisation of the< i> Escherichia coli</i> strain associated with an

outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study **11:** 671-676.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J., and Denamur, E. (1998) Phylogenetic analysis of Escherichia coli strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J Infect Dis* **177:** 642-650.

Bingle, L.E., Constantinidou, C., Shaw, R.K., Islam, M.S., Patel, M., Snyder, L.A., *et al.* (2014) Microarray analysis of the ler regulon in enteropathogenic and enterohaemorrhagic escherichia coli strains **9:** e80160.

Blanco, J.E., Blanco, M., Mora, A., and Blanco, J. (1997) Production of toxins (enterotoxins, verotoxins, and necrotoxins) and colicins by Escherichia coli strains isolated from septicemic and healthy chickens: relationship with in vivo pathogenicity. *J Clin Microbiol* **35:** 2953-2957.

Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschape, H., and Hacker, J. (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an Escherichia coli wild-type pathogen. *Infect Immun* **62:** 606-614.

Boussau, B., Gueguen, L., and Gouy, M. (2008) Accounting for horizontal gene transfers explains conflicting hypotheses regarding the position of aquificales in the phylogeny of Bacteria. *BMC Evol Biol* **8:** 272-2148-8-272.

Branchu, P., Matrat, S., Vareille, M., Garrivier, A., Durand, A., Crépin, S., *et al.* (2014) NsrR, GadE, and GadX Interplay in Repressing Expression of the Escherichia coli O157: H7 LEE Pathogenicity Island in Response to Nitric Oxide **10:** e1003874.

Brown, N.L., Stoyanov, J.V., Kidd, S.P., and Hobman, J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27:** 145-163.

Brown, N.F., Rogers, L.D., Sanderson, K.L., Gouw, J.W., Hartland, E.L., and Foster, L.J. (2014) A horizontally acquired transcription factor coordinates Salmonella adaptations to host microenvironments. *MBio* **5:** e01727-14.

Brown, N.L., Stoyanov, J.V., Kidd, S.P., and Hobman, J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27:** 145-163.

Brown, N.L., Stoyanov, J.V., Kidd, S.P., and Hobman, J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27:** 145-163.

Browning, D.F., Cole, J.A., and Busby, S.J. (2004) Transcription activation by remodelling of a nucleoprotein assembly: the role of NarL at the FNR- dependent Escherichia coli nir promoter. *Mol Microbiol* **53:** 203-215.

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:** 773-780.

Buck, M., Engl, C., Joly, N., Jovanovic, G., Jovanovic, M., Lawton, E., *et al.* (2015) In vitro and in vivo methodologies for studying the Sigma 54-dependent transcription. *Methods Mol Biol* **1276:** 53-79.

Burgess, R.R., and Anthony, L. (2001) How sigma docks to RNA polymerase and what sigma does. *Curr Opin Microbiol* **4:** 126-131.
Busby, S., and Kolb, A. (1996) The CAP modulon. In *Regulation of gene expression in Escherichia coli* The CAP modulon: Springer, pp. 255-279.

Busby, S., Kotlarz, D., and Buc, H. (1983) Deletion mutagenesis of the Escherichia coli galactose operon promoter region. *J Mol Biol* **167:** 259-274.

Bustamante, V.H., Santana, F.J., Calva, E., and Puente, J.L. (2001) Transcriptional regulation of type III secretion genes in enteropathogenic Escherichia coli: Ler antagonizes H-NS-dependent repression. *Mol Microbiol* **39:** 664-678.

Bustamante, V.H., Villalba, M.I., García-Angulo, V.A., Vázquez, A., Martínez, L.C., Jiménez, R., and Puente, J.L. (2011) PerC and GrlA independently regulate Ler expression in enteropathogenic Escherichia coli. *Mol Microbiol* **82:** 398-415.

Buttner, M.J., and Lewis, C.G. (1992) Construction and characterization of Streptomyces coelicolor A3(2) mutants that are multiply deficient in the nonessential hrd-encoded RNA polymerase sigma factors. *J Bacteriol* **174:** 5165-5167.

Casadaban, M.J., and Cohen, S.N. (1980) Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J Mol Biol* **138:** 179-207.

Cavalier-Smith, T. (2006) Cell evolution and Earth history: stasis and revolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361:** 969-1006.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., *et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* **337:** 591-595.

Chang, A.C., and Cohen, S.N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134:** 1141-1156.

Chen, C., Stevens, B., Kaur, J., Smilansky, Z., Cooperman, B.S., and Goldman, Y.E. (2011) Allosteric vs. spontaneous exit-site (E-site) tRNA dissociation early in protein synthesis. *Proceedings of the National Academy of Sciences* **108:** 16980-16985.

Chen, I., Christie, P.J., and Dubnau, D. (2005) The ins and outs of DNA transfer in bacteria. *Science* **310:** 1456-1460.

Cho, B.K., Knight, E.M., Barrett, C.L., and Palsson, B.O. (2008) Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. *Genome Res* **18:** 900-910.

Choy, H., and Adhya, S. (1996) Negative control **1:** 1287-1299.
Churchman, L.S., and Weissman, J.S. Native elongating transcript sequencing (NET-seq)14.4. 1-14.4. 17.

Claret, L., and Rouviere-Yaniv, J. (1997) Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival. *J Mol Biol* **273:** 93-104.

Claverys, J., Prudhomme, M., and Martin, B. (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. *Microbiology* **60:** 451.

Cleary, J., Lai, L.C., Shaw, R.K., Straatman-Iwanowska, A., Donnenberg, M.S., Frankel, G., and Knutton, S. (2004) Enteropathogenic Escherichia coli (EPEC) adhesion to intestinal epithelial cells: role of bundle-forming pili (BFP), EspA filaments and intimin. *Microbiology* **150:** 527-538.

Collinson, S.K., Emody, L., Muller, K.H., Trust, T.J., and Kay, W.W. (1991) Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. *J Bacteriol* **173:** 4773-4781.

Collinson, S.K., Emody, L., Trust, T.J., and Kay, W.W. (1992) Thin aggregative fimbriae from diarrheagenic Escherichia coli. *J Bacteriol* **174:** 4490-4495.

Cookson, A.L., and Woodward, M.J. (2003) The role of intimin in the adherence of enterohaemorrhagic Escherichia coli (EHEC) O157 : H7 to HEp-2 tissue culture cells and to bovine gut explant tissues **292:** 547-553.

Cordeiro, T.N., Schmidt, H., Madrid, C., Juárez, A., Bernadó, P., Griesinger, C., *et al.* (2011) Indirect DNA readout by an H-NS related protein: structure of the DNA complex of the C-terminal domain of Ler.

Cui, Y., Wang, Q., Stormo, G.D., and Calvo, J.M. (1995) A consensus sequence for binding of Lrp to DNA. *J Bacteriol* **177:** 4872-4880.

Dalziel, R.G., Fitzgerald, J.R., Mims, C.A., Nash, A.A., and Stephen, J. (2000) *Mims' pathogenesis of infectious disease*: Academic Press.
Dame, R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* **56:** 858-870.

Dame, R.T., Noom, M.C., and Wuite, G.J. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444:** 387-390.

Dean-Nystrom, E.A., Bosworth, B.T., Moon, H.W., and O'Brien, A.D. (1998) Escherichia coli O157:H7 requires intimin for enteropathogenicity in calves. *Infect Immun* **66:** 4560-4563.

Deng, S., Stein, R.A., and Higgins, N.P. (2005) Organization of supercoil domains and their reorganization by transcription. *Mol Microbiol* **57:** 1511-1521.

Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vazquez, A., *et al.* (2004) Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc Natl Acad Sci U S A* **101:** 3597-3602.

Diepold, A., and Wagner, S. (2014) Assembly of the bacterial type III secretion machinery. *FEMS Microbiol Rev* **38:** 802-822.

Dillon, S.C., and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8:** 185-195.

Donnenberg, M.S., Tzipori, S., McKee, M., O'brien, A., Alroy, J., and Kaper, J. (1993) The role of the eae gene of enterohemorrhagic Escherichia coli in intimate attachment in vitro and in a porcine model. *J Clin Invest* **92:** 1418.

Dorman, C.J. (2009) Nucleoid-associated proteins and bacterial physiology. *Adv Appl Microbiol* **67:** 47-64.

Dorman, C.J., and Deighan, P. (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet Dev* **13:** 179-184.

Dove, S.L., Darst, S.A., and Hochschild, A. (2003) Region 4 of sigma as a target for transcription regulation. *Mol Microbiol* **48:** 863-874.

Dramsi, S., Magnet, S., Davison, S., and Arthur, M. (2008) Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol Rev* **32:** 307-320.

Drlica, K., and Rouviere-Yaniv, J. (1987) Histonelike proteins of bacteria. *Microbiol Rev* **51:** 301-319.

Dworkin, J., and Losick, R. (2002) Does RNA polymerase help drive chromosome segregation in bacteria?. *Proceedings of the National Academy of Sciences* **99:** 14089-14094.

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

Ebright, R.H. (1993) Transcription activation at Class I CAP-dependent promoters. *Mol Microbiol* **8:** 797-802.
Eigen, M., Gardiner, W., Schuster, P., and Winkler-Oswatitsch, R. (1981) The origin of genetic information. *Sci Am* **244:** 88-92, 96, et passim.

Elliott, S.J., Hutcheson, S.W., Dubois, M.S., Mellies, J.L., Wainwright, L.A., Batchelor, M., *et al.* (1999) Identification of CesT, a chaperone for the type III secretion of Tir in enteropathogenic Escherichia coli. *Mol Microbiol* **33:** 1176-1189.

Elliott, S.J., Wainwright, L.A., McDaniel, T.K., Jarvis, K.G., Deng, Y.K., Lai, L.C.,

et al. (1998) The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic Escherichia coli E2348/69. *Mol Microbiol* **28:** 1-4.

Evans, D.J.,Jr, and Evans, D.G. (1983) Classification of pathogenic Escherichia coli according to serotype and the production of virulence factors, with special reference to colonization-factor antigens. *Rev Infect Dis* **5 Suppl 4:** S692-701.

Feklistov, A., and Darst, S.A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **147:** 1257-1269.

Fitzhenry, R., Dahan, S., Torres, A.G., Chong, Y., Heuschkel, R., Murch, S.H., *et al.* (2006) Long polar fimbriae and tissue tropism in Escherichia coli O157:H7. *Microb Infect* **8:** 1741-1749.

Frank, J., and Gonzalez, R.L.,Jr (2010) Structure and dynamics of a processive Brownian motor: the translating ribosome. *Annu Rev Biochem* **79:** 381-412.

Galán, J.E., and Collmer, A. (1999) Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284:** 1322-1328.

Gardner, J.F. (1979) Regulation of the threonine operon: tandem threonine and isoleucine codons in the control region and translational control of transcription termination. *Proceedings of the National Academy of Sciences* **76:** 1706-1710.

Gavini, N., and Pulakat, L. (1991) Role of translation of the pheA leader peptide coding region in attenuation regulation of the Escherichia coli pheA gene. *J Bacteriol* **173:** 4904-4907.

Gerber, D., Maerkl, S.J., and Quake, S.R. (2009) An in vitro microfluidic approach to generating protein-interaction networks. *Nat Methods* **6:** 71-74.
Ghosh, P., Ishihama, A., and Chatterji, D. (2001) Escherichia coli RNA polymerase subunit ω and its N- terminal domain bind full- length β′ to facilitate incorporation into the α2β subassembly **268:** 4621-4627.

Ghosh, P. (2004) Process of protein transport by the type III secretion system. *Microbiol Mol Biol Rev* **68:** 771-795.
Gilbert, W. (1986) Origin of life: The RNA world. *Nature* **319:** .
Gilbert, W., and Muller-Hill, B. (1966) Isolation of the lac repressor. *Proc Natl Acad Sci U S A* **56:** 1891-1898.

Giron, J.A., Ho, A.S., and Schoolnik, G.K. (1991) An inducible bundle-forming pilus of enteropathogenic Escherichia coli. *Science* **254:** 710-713.

Glauert, A.M., and Thornley, M.J. (1969) The topography of the bacterial cell wall. *Annu Rev Microbiol* **23:** 159-198.

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

Grainger, D., and Busby, S.W. (2008) Methods for studying global patterns of DNA

binding by bacterial transcription factors and RNA polymerase. *Biochem Soc Trans* **36:** 754-757.

Grainger, D.C., Aiba, H., Hurd, D., Browning, D.F., and Busby, S.J.W. (2007) Transcription factor distribution in Escherichia coli: studies with FNR protein. *Nucleic Acids Res* **35:** 269-278.

Grainger, D.C., Hurd, D., Harrison, M., Holdstock, J., and Busby, S.J.W. (2005) Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome. *Proc Natl Acad Sci U S A* **102:** 17693-17698.

Gualerzi, C.O., and Pon, C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry (N Y)* **29:** 5881-5889.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849-857.

Gunnelius, L., Hakkila, K., Kurkela, J., Wada, H., Tyystjarvi, E., and Tyystjarvi, T. (2014) The omega subunit of the RNA polymerase core directs transcription efficiency in cyanobacteria. *Nucleic Acids Res* **42:** 4606-4614.

Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R., and Goebel, W. (1990) Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal Escherichia coli isolates. *Microb Pathog* **8:** 213-225.

Hacker, J., and Blum-Oehler, G. (2007) In appreciation of Theodor Escherich **5:** 902-902.

Hartzog, G.A., and Kaplan, C.D. (2011) Competing for the clamp: promoting RNA polymerase processivity and managing the transition from initiation to elongation. *Mol Cell* **43:** 161-163.

Hatfield, G.W., and Benham, C.J. (2002) DNA topology-mediated control of global gene expression in Escherichia coli. *Annu Rev Genet* **36:** 175-203.

Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C., and Gourse, R.L. (2006) rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase. *Cell* **125:** 1069-1082.

Hedegaard, J., Steffensen, S.A., Norskov-Lauritsen, N., Mortensen, K.K., and Sperling-Petersen, H.U. (1999) Identification of Enterobacteriaceae by partial sequencing of the gene encoding translation initiation factor 2. *Int J Syst Bacteriol* **49 Pt 4:** 1531-1538.

Henkin, T.M. (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* **22:** 3383-3390.

Herring, C.D., Raffaelle, M., Allen, T.E., Kanin, E.I., Landick, R., Ansari, A.Z., and Palsson, B.O. (2005) Immobilization of Escherichia coli RNA polymerase and location of binding sites by use of chromatin immunoprecipitation and microarrays. *J Bacteriol* **187:** 6166-6174.

Herring, C.D., Glasner, J.D., and Blattner, F.R. (2003) Gene replacement without selection: regulated suppression of amber mutations in Escherichia coli. *Gene* **311:** 153-163.

Hicks, S., Frankel, G., Kaper, J.B., Dougan, G., and Phillips, A.D. (1998) Role of intimin and bundle-forming pili in enteropathogenic Escherichia coli adhesion to pediatric intestinal tissue in vitro. *Infect Immun* **66:** 1570-1578.

Honda, N., Iyoda, S., Yamamoto, S., Terajima, J., and Watanabe, H. (2009) LrhA positively controls the expression of the locus of enterocyte effacement genes in enterohemorrhagic Escherichia coli by differential regulation of their master regulators PchA and PchB. *Mol Microbiol* **74:** 1393-1341.

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

Ishihama, A. (1992) Role of the RNA polymerase alpha subunit in transcription activation. *Mol Microbiol* **6:** 3283-3288.

Ishihama, A. (2000) Functional modulation of Escherichia coli RNA polymerase. *Annu Rev Microbiol* **54:** 499-518.

Ishihama, A., and Ito, K. (1972) Subunits of RNA polymerase in function and structure. II. Reconstitution of Escherichia coli RNA polymerase from isolated subunits. *J Mol Biol* **72:** 111-123.

Islam, M.S., Shaw, R.K., Frankel, G., Pallen, M.J., and Busby, S.J. (2012) Translation of a minigene in the 5' leader sequence of the enterohaemorrhagic Escherichia coli LEE1 transcription unit affects expression of the neighbouring downstream gene. *Biochem J* **441:** 247-253.

Islam, M.S., Bingle, L.E.H., Pallen, M.J., and Busby, S.J.W. (2011) Organization of the LEE1 operon regulatory region of enterohaemorrhagic Escherichia coli O157:H7 and activation by GrlA. *Mol Microbiol* **79:** 468-483.

Ito, K., Iwakura, Y., and Ishihama, A. (1975) Biosynthesis of RNA polymerase in Escherichia coli. III. Identification of intermediates in the assembly of RNA polymerase. *J Mol Biol* **96:** 257-271.

Iyoda, S., Honda, N., Saitoh, T., Shimuta, K., Terajima, J., Watanabe, H., and Ohnishi, M. (2011) Coordinate control of the locus of enterocyte effacement and enterohemolysin genes by multiple common virulence regulators in enterohemorrhagic Escherichia coli. *Infect Immun* **79:** 4628-4637.

Iyoda, S., Koizumi, N., Satou, H., Lu, Y., Saitoh, T., Ohnishi, M., and Watanabe, H.

(2006) The GrlR-GrlA regulatory system coordinately controls the expression of flagellar and LEE-encoded type III protein secretion systems in enterohemorrhagic Escherichia coli. *J Bacteriol* **188:** 5682-5692.

Iyoda, S., and Watanabe, H. (2004) Positive effects of multiple pch genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic Escherichia coli O157 : H7 to HEp-2 cells. *Microbiology* **150:** 2357-2571.

Izoré, T., Job, V., and Dessen, A. (2011) Biogenesis, Regulation, and Targeting of the Type III Secretion System. *Structure* **19:** 603-612.

Jacob, F., Brenner, S., and Cuzin, F. (1963) On the Regulation of DNA Replication in Bacteria. *Cold Spring Harbor Symposia on Quantitative Biology* **28:** 329-348.

Jeon, Y.H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., and Kyogoku, Y. (1995) Solution structure of the activator contact domain of the RNA polymerase alpha subunit. *Science* **270:** 1495-1497.

Jeong, J.H., Kim, H.J., Kim, K.H., Shin, M., Hong, Y., Rhee, J.H., *et al.* (2012) An unusual feature associated with LEE1 P1 promoters in enteropathogenic Escherichia coli (EPEC). *Mol Microbiol* **83:** 612-622.

Jerse, A.E., Yu, J., Tall, B.D., and Kaper, J.B. (1990) A genetic locus of enteropathogenic Escherichia coli necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc Natl Acad Sci U S A* **87:** 7839-7843.

Jimenez, R., Cruz-Migoni, S.B., Huerta-Saquero, A., Bustamante, V.H., and Puente, J.L. (2010) Molecular characterization of GrlA, a specific positive regulator of ler expression in enteropathogenic Escherichia coli. *J Bacteriol* **192:** 4627-4642.

Jobichen, C., Li, M., Yerushalmi, G., Tan, Y.W., Mok, Y.K., Rosenshine, I., *et al.* (2007) Structure of GrlR and the implication of its EDED motif in mediating the regulation of type III secretion system in EHEC. *PLoS Pathog* **3:** e69.

Johnson, J.R., and Stamm, W.E. (1989) Urinary tract infections in women: diagnosis and treatment. *Ann Intern Med* **111:** 906-917.

Kaas, R.S., Friis, C., Ussery, D.W., and Aarestrup, F.M. (2012) Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse Escherichia coli genomes. *BMC Genomics* **13:** 577-2164-13-577.

Kaguni, J.M. (2006) DnaA: controlling the initiation of bacterial DNA replication and more. *Annu Rev Microbiol* **60:** 351-375.

Kailasan Vanaja, S., Bergholz, T.M., and Whittam, T.S. (2009) Characterization of the Escherichia coli O157:H7 Sakai GadE regulon. *J Bacteriol* **191:** 1868-1877.

Kamio, Y., and Nikaido, H. (1976) Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase C and cyanogen bromide

activated dextran in the external medium. *Biochemistry (N Y)* **15:** 2561-2570.

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:** 1144-1147.

Kapanidis, A.N., Margeat, E., Laurence, T.A., Doose, S., Ho, S.O., Mukhopadhyay, J., *et al.* (2005) Retention of Transcription Initiation Factor σ70 in Transcription Elongation: Single-Molecule Analysis. *Mol Cell* **20:** 347-356.

Kaper, J.B., and O'Brien, A.D. (1997) *Escherichia Coli 0157:H7 and Other Shiga Toxin-producing E. Coli Strains* : American Society for Microbiology.

Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic escherichia coli **2:** 123-140.

Kavenoff, R., and Bowen, B. (1976) Electron microscopy of membrane-free folded chromosomes from Escherichia coli. *Chromosoma* **59:** 89-101.

Kellenberger, Edouard.,Ryter, Antoinette., (1958) *Cell wall and cytoplasmic membrane of Escherichia coli.* New York: Rockefeller inst. press.

Kolb, A., Kotlarz, D., Kusano, S., and Ishihama, A. (1995) Selectivity of the Escherichia coli RNA polymerase Es38 for overlapping promoters and ability to support CRP activation. *Nucleic Acids Research* **23:** 819-826.

Koo, B., Rhodius, V.A., Campbell, E.A., and Gross, C.A. (2009) Mutational analysis of Escherichia coli σ28 and its target promoters reveals recognition of a composite−10 region, comprised of an 'extended−10'motif and a core−10 element. *Mol Microbiol* **72:** 830-843.

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification 1 **55:** 709-742.

Koonin, E.V., and Wolf, Y.I. (2008) Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res* **36:** 6688-6719.

Kornberg, A. (1988) Special Issue: Enzymology of DNA Replication DNA replication **951:** 235-239.

Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol Rev* **47:** 1-45.
Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* **234:** 187-208.

Krawiec, S., and Riley, M. (1990) Organization of the bacterial chromosome. *Microbiol Rev* **54:** 502-539.

Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S.L., Rogulja, D., Nickels, B.E., *et al.* (2002) A role for interaction of the RNA polymerase flap domain with the

sigma subunit in promoter recognition. *Science* **295:** 855-857.

Landgraf, J.R., Wu, J., and Calvo, J.M. (1996) Effects of nutrition and growth rate on Lrp levels in Escherichia coli. *J Bacteriol* **178:** 6930-6936.

Laursen, B.S., Sorensen, H.P., Mortensen, K.K., and Sperling-Petersen, H.U. (2005) Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* **69:** 101-123.

LeClerc, J.E., Li, B., Payne, W.L., and Cebula, T.A. (1996) High mutation frequencies among Escherichia coli and Salmonella pathogens. *Science* **274:** 1208-1211.

Lee, D.J., Minchin, S.D., and Busby, S.J. (2012) Activating transcription in bacteria. *Annu Rev Microbiol* **66:** 125-152.

Lee, D., Bingle, L., Heurlier, K., Pallen, M., Penn, C., Busby, S., and Hobman, J. (2009) Gene doctoring: a method for recombineering in laboratory and pathogenic Escherichia coli strains **9:** 252.

Lemon, K.P., and Grossman, A.D. (2001) The extrusion-capture model for chromosome partitioning in bacteria. *Genes Dev* **15:** 2031-2041.

Lewis, D.E., and Adhya, S. (2015) Molecular Mechanisms of Transcription Initiation at gal Promoters and their Multi-Level Regulation by GalR, CRP and DNA Loop. *Biomolecules* **5:** 2782-2807.

Li, M., Moyle, H., and Susskind, M.M. (1994) Target of the transcriptional activation function of phage lambda cI protein. *Science* **263:** 75-77.

Lim, C.J., Lee, S.Y., Kenney, L.J., and Yan, J. (2012) Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing **2:** .

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* **95:** 271-276.

Lodge, J., Williams, R., Bell, A., Chan, B., and Busby, S. (1990) Comparison of promoter activities in Escherichia coli and Pseudomonas aeruginosa: use of a new broad-host-range promoter-probe plasmid. *FEMS Microbiol Lett* **67:** 221-226.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* **174:** 3843-3849.

Lonetto, M.A., Brown, K.L., Rudd, K.E., and Buttner, M.J. (1994) Analysis of the Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *Proc Natl Acad Sci U S A* **91:** 7573-7577.

Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C.

(2006) H-NS mediates the silencing of laterally acquired genes in bacteria **2:** e81.

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.

Luo, Y., Frey, E.A., Pfuetzner, R.A., Creagh, A.L., Knoechel, D.G., Haynes, C.A., *et al.* (2000) Crystal structure of enteropathogenic Escherichia coli intimin-receptor complex. *Nature* **405:** 1073-1077.

Madan Babu, M., Teichmann, S.A., and Aravind, L. (2006) Evolutionary Dynamics of Prokaryotic Transcriptional Regulatory Networks. *J Mol Biol* **358:** 614-633.
Mallick, E.M., Garber, J.J., Vanguri, V.K., Balasubramanian, S., Blood, T., Clark, S., *et al.* (2014) The ability of an attaching and effacing pathogen to trigger localized actin assembly contributes to virulence by promoting mucosal attachment. *Cell Microbiol* **16:** 1405-1424.

Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R.R. (2003) Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *Cell* **113:** 577-586.

Martínez-Antonio, A., and Collado-Vides, J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* **6:** 482-489.

Mathew, R., and Chatterji, D. (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol* **14:** 450-455.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A* **92:** 1664-1668.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., *et al.* (1999) Food-related illness and death in the United States. *Emerg Infect Dis* **5:** 607-625.

Meier, M., Sit, R.V., and Quake, S.R. (2013) Proteome-wide protein interaction measurements of bacterial proteins of unknown function. *Proceedings of the National Academy of Sciences* **110:** 477-482.

Mellies, J.L., Barron, A.M., and Carmona, A.M. (2007) Enteropathogenic and enterohemorrhagic Escherichia coli virulence gene regulation. *Infect Immun* **75:** 4199-4210.

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

Milon, P., Maracci, C., Filonava, L., Gualerzi, C.O., and Rodnina, M.V. (2012) Real-time assembly landscape of bacterial 30S translation initiation complex. *Nat Struct Mol Biol* **19:** 609-615.

Mirkin, B.G., Fenner, T.I., Galperin, M.Y., and Koonin, E.V. (2003) Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evol Biol* **3:** 2.
Mitchell, J.E., Oshima, T., Piper, S.E., Webster, C.L., Westblade, L.F., Karimova, G., *et al.* (2007) The Escherichia coli regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70. *J Bacteriol* **189:** 3489-3495.

Mitchell, J.E., Zheng, D., Busby, S.J., and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in Escherichia coli. *Nucleic Acids Res* **31:** 4689-4695.

Mitchell, P. (1961)
Approaches to the analysis of specific membrane transport. . *biological structure and function*581-603.

Mobley, H.L., Koch, A.L., Doyle, R.J., and Streips, U.N. (1984) Insertion and fate of the cell wall in Bacillus subtilis. *J Bacteriol* **158:** 169-179.

Moore, P.B. (2012) How should we think about the ribosome?. *Annu Rev Biophys* **41:** 1-19.

Moxon, E., and Kroll, J. (1990) The role of bacterial polysaccharide capsules as virulence factors. In *Bacterial capsules* The role of bacterial polysaccharide capsules as virulence factors: Springer, pp. 65-85.

Müller-Hill, B. (1996) The lac operon: a short history of a genetic paradigm Walter de Gruyter.

Muller-Hill, B. (1998) The function of auxiliary operators. *Mol Microbiol* **29:** 13-18.

Mullineaux, C.W., Nenninger, A., Ray, N., and Robinson, C. (2006) Diffusion of Green Fluorescent Protein in Three Cell Environments in Escherichia Coli. *Journal of Bacteriology* **188:** 3442-3448.

Murakami, K.S. (2015) Structural Biology of Bacterial RNA Polymerase **5:** 848-864.

Murakami, K.S. (2013) X-ray crystal structure of Escherichia coli RNA polymerase sigma70 holoenzyme. *J Biol Chem* **288:** 9126-9134.

Murakami, K.S., and Darst, S.A. (2003) Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol* **13:** 31-39.

Murakami, K.S., Masuda, S., and Darst, S.A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science* **296:** 1280-1284.

Nataro, J.P., and Kaper, J.B. (1998) Diarrheagenic Escherichia coli. *Clinical Microbiology Reviews* **11:** 142-201.

Neuhaus, F.C., and Baddiley, J. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67:** 686-723.

Nicholls, L., Grant, T.H., and Robins-Browne, R.M. (2000) Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic Escherichia coli to epithelial cells. *Mol Microbiol* **35:** 275-288.

Nielsen, H.J., Ottesen, J.R., Youngren, B., Austin, S.J., and Hansen, F.G. (2006) The Escherichia coli chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol Microbiol* **62:** 331-338.

Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67:** 593-656.

Niki, H., Jaffe, A., Imamura, R., Ogura, T., and Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO J* **10:** 183-193.

Noom, M.C., Navarre, W.W., Oshima, T., Wuite, G.J., and Dame, R.T. (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* **17:** R913-4.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299-304.

O'Donnell, S.M., and Janssen, G.R. (2001) The initiation codon affects ribosome binding and translational efficiency in Escherichia coli of cI mRNA with or without the 5' untranslated leader. *J Bacteriol* **183:** 1277-1283.

Olsen, A., Jonsson, A., and Normark, S. (1989) Fibronectin binding mediated by a novel class of surface organelles on Escherichia coli. *Nature* **338:** 652-655.

Ouwehand, A.C., Salminen, S., and Isolauri, E. (2002) Probiotics: an overview of beneficial effects. *Antonie Van Leeuwenhoek* **82:** 279-289.

Owen-Hughes, T.A., Pavitt, G.D., Santos, D.S., Sidebotham, J.M., Hulton, C.S., Hinton, J.C., and Higgins, C.F. (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71:** 255-265.

Padavannil, A., Jobichen, C., Mills, E., Velazquez-Campoy, A., Li, M., Leung, K.Y., *et al.* (2013) Structure of GrlR-GrlA complex that prevents GrlA activation of virulence genes. *Nat Commun* **4:** 2546.

Paget, M.S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution **5:** 1245-1265.

Pallen, M. (2009) *The rough guide to evolution*: Dorling Kindersley Ltd.

Pallen, M.J., Beatson, S.A., and Bailey, C.M. (2005) Bioinformatics analysis of the

locus for enterocyte effacement provides novel insights into type-III secretion. *BMC Microbiol* **5:** 9.

Paradis, S., Boissinot, M., Paquette, N., Belanger, S.D., Martel, E.A., Boudreau, D.K., *et al.* (2005) Phylogeny of the Enterobacteriaceae based on genes encoding elongation factor Tu and F-ATPase beta-subunit. *Int J Syst Evol Microbiol* **55:** 2013-2025.

Parkhill, J., and Brown, N.L. (1990) Site-specific insertion and deletion mutants in the mer promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. *Nucleic Acids Res* **18:** 5157-5162.

Perdue, S.A., and Roberts, J.W. (2011) σ 70-dependent transcription pausing in Escherichia coli. *J Mol Biol* **412:** 782-792.

Perna, N.T., Mayhew, G.F., Posfai, G., Elliott, S., Donnenberg, M.S., Kaper, J.B., and Blattner, F.R. (1998) Molecular evolution of a pathogenicity island from enterohemorrhagic Escherichia coli O157:H7. *Infect Immun* **66:** 3810-3817.

Phillips, A.D., Navabpour, S., Hicks, S., Dougan, G., Wallis, T., and Frankel, G. (2000) Enterohaemorrhagic Escherichia coli O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. *Gut* **47:** 377-381.

Pollack, J.R., and Iyer, V.R. (2002) Characterizing the physical genome. *Nat Genet* **32:** 515-521.

Polz, M.F., Alm, E.J., and Hanage, W.P. (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure **29:** 170-175.

Porter, M.E., Mitchell, P., Free, A., Smith, D.G., and Gally, D.L. (2005) The LEE1 promoters from both enteropathogenic and enterohemorrhagic Escherichia coli can be activated by PerC-like proteins from either organism. *J Bacteriol* **187:** 458-472.

Postow, L., Hardy, C.D., Arsuaga, J., and Cozzarelli, N.R. (2004) Topological domain structure of the Escherichia coli chromosome. *Genes Dev* **18:** 1766-1779.

Poxton, I. (2010) Bacterial communities in the gut and the consequences of upsetting them. *SGM* **37:** 232-235.

Provence, D.L., and Curtiss, R.,3rd (1992) Role of crl in avian pathogenic Escherichia coli: a knockout mutation of crl does not affect hemagglutination activity, fibronectin binding, or Curli production. *Infect Immun* **60:** 4460-4467.

Raetz, C.R. (1990) Biochemistry of endotoxins. *Annu Rev Biochem* **59:** 129-170.

Raetz, C.R., and Dowhan, W. (1990) Biosynthesis and function of phospholipids in Escherichia coli. *J Biol Chem* **265:** 1235-1238.

Raffaelle, M., Kanin, E.I., Vogt, J., Burgess, R.R., and Ansari, A.Z. (2005) Holoenzyme switching and stochastic release of sigma factors from RNA polymerase

in vivo. *Mol Cell* **20:** 357-366.

Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K., and Whittam, T.S. (2000) Parallel evolution of virulence in pathogenic Escherichia coli. *Nature* **406:** 64-67.

Reshes, G., Vanounou, S., Fishov, I., and Feingold, M. (2008) Cell Shape Dynamics in Escherichia coli. *Biophys J* **94:** 251-264.

Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314:** 1139-1143.

Richards, V.P., Palmer, S.R., Pavinski Bitar, P.D., Qin, X., Weinstock, G.M., Highlander, S.K., *et al.* (2014) Phylogenomics and the dynamic genome evolution of the genus Streptococcus. *Genome Biol Evol* **6:** 741-753.

Robinson, A., and van Oijen, A.M. (2013) Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies. *Nat Rev Microbiol* **11:** 303-315.
Rose, R.E. (1988) The nucleotide sequence of pACYC184. *Nucleic Acids Res* **16:** 355.

Rossiter, A.E., Browning, D.F., Leyton, D.L., Johnson, M.D., Godfrey, R.E., Wardius, C.A., *et al.* (2011) 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:** 179-191.

Russell, R.M., Sharp, F.C., Rasko, D.A., and Sperandio, V. (2007) QseA and GrlR/GrlA regulation of the locus of enterocyte effacement genes in enterohemorrhagic Escherichia coli. *J Bacteriol* **189:** 5387-5392.

Russo, T.A., and Johnson, J.R. (2006) Extraintestinal isolates of Escherichia coli: identification and prospects for vaccine development. *Expert Rev Vaccines* **5:** 45-54.

Saavedra, J.M. (2001) Clinical applications of probiotic agents. *Am J Clin Nutr* **73:** 1147S-1151S.

Saecker, R.M., Record, M.T.,Jr, and Dehaseth, P.L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J Mol Biol* **412:** 754-771.

Saitoh, T., Iyoda, S., Yamamoto, S., Lu, Y., Shimuta, K., Ohnishi, M., *et al.* (2008) Transcription of the ehx enterohemolysin gene is positively regulated by GrlA, a global regulator encoded within the locus of enterocyte effacement in enterohemorrhagic Escherichia coli. *J Bacteriol* **190:** 4822-4830.

Sala, C., Grainger, D.C., and Cole, S.T. (2009) Dissecting regulatory networks in host-pathogen interaction using chIP-on-chip technology **5:** 430-437.

Saldana, Z., Xicohtencatl-Cortes, J., Avelino, F., Phillips, A.D., Kaper, J.B., Puente, J.L., and Giron, J.A. (2009) Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli. *Environ Microbiol* **11:** 992-1006.

Sambuy, Y., De Angelis, I., Ranaldi, G., Scarino, M.L., Stammati, A., and Zucco, F. (2005) The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell Biol Toxicol* **21:** 1-26.

Sauer, F.G., Mulvey, M.A., Schilling, J.D., Martinez, J.J., and Hultgren, S.J. (2000) Bacterial pili: molecular mechanisms of pathogenesis. *Curr Opin Microbiol* **3:** 65-72.

Savageau, M.A. (1983) Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. *The American Naturalist* **122:** 732-744.

Schilling, J.D., Mulvey, M.A., Vincent, C.D., Lorenz, R.G., and Hultgren, S.J. (2001) Bacterial invasion augments epithelial cytokine responses to Escherichia coli through a lipopolysaccharide-dependent mechanism. *J Immunol* **166:** 1148-1155.

Schleifer, K.H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36:** 407-477.

Schmidt, H., and Hensel, M. (2004) Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* **17:** 14-56.

Schnaitman, C.A. (1970) Protein composition of the cell wall and cytoplasmic membrane of Escherichia coli. *J Bacteriol* **104:** 890-901.

Schneider, D.A., and Gourse, R.L. (2004) Relationship between growth rate and ATP concentration in Escherichia coli: a bioassay for available cellular ATP. *J Biol Chem* **279:** 8262-8268.

Schroder, I., Darie, S., and Gunsalus, R.P. (1993) Activation of the Escherichia coli nitrate reductase (narGHJI) operon by NarL and Fnr requires integration host factor. *J Biol Chem* **268:** 771-774.

Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008) A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* **15:** 1091-1103.

Seitz, P., and Blokesch, M. (2013) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram- negative bacteria. *FEMS Microbiol Rev* **37:** 336-363.

Seitz, P., and Blokesch, M. (2013) DNA-uptake machinery of naturally competent Vibrio cholerae. *Proc Natl Acad Sci U S A* **110:** 17987-17992.
Sekine, S., Tagami, S., and Yokoyama, S. (2012) Structural basis of transcription by bacterial and eukaryotic RNA polymerases. *Curr Opin Struct Biol* **22:** 110-118.

Sharma, V.K., and Zuerner, R.L. (2004) Role of hha and ler in transcriptional regulation of the esp operon of enterohemorrhagic Escherichia coli O157:H7. *J Bacteriol* **186:** 7290-7301.

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:** 5392-5399.

Shin, M., Lagda, A.C., Lee, J.W., Bhat, A., Rhee, J.H., Kim, J., *et al.* (2012) Gene silencing by H-NS from distal DNA site. *Mol Microbiol* **86:** 707-719.

Shine, J., and Dalgarno, L. (1974) The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A* **71:** 1342-1346.

Shulman, S.T., Friedmann, H.C., and Sims, R.H. (2007) Theodor Escherich: the first pediatric infectious diseases physician?. *Clin Infect Dis* **45:** 1025-1029.

Shultzaberger, R.K., Chen, Z., Lewis, K.A., and Schneider, T.D. (2007) Anatomy of Escherichia coli σ70 promoters. *Nucleic Acids Research* **35:** 771-788.

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

Silverman, M., and Simon, M. (1974) Flagellar rotation and the mechanism of bacterial motility. *Nature* **249:** 73-74.

Simonetti, A., Marzi, S., Jenner, L., Myasnikov, A., Romby, P., Yusupova, G., *et al.* (2009) A structural view of translation initiation in bacteria. *Cell Mol Life Sci* **66:** 423-436.

Siryaporn, A., Kuchma, S.L., O'Toole, G.A., and Gitai, Z. (2014) Surface attachment induces Pseudomonas aeruginosa virulence. *Proceedings of the National Academy of Sciences* **111:** 16860-16865.

Smillie, C.S., Smith, M.B., Friedman, J., Cordero, O.X., David, L.A., and Alm, E.J. (2011) Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480:** 241-244.

Smith, J.L., Fratamico, P.M., and Gunther, N.W. (2007) Extraintestinal pathogenic Escherichia coli. *Foodborne Pathog Dis* **4:** 134-163.

Snel, B., Bork, P., and Huynen, M.A. (2002) Genomes in flux: the evolution of archaeal and proteobacterial gene content. *Genome Res* **12:** 17-25.

Søgaard-Andersen, L., Martinussen, J., Møllegaard, N.E., Douthwaite, S.R., and Valentin-Hansen, P. (1990) The CytR repressor antagonizes cyclic AMP-cyclic AMP receptor protein activation of the deoCp2 promoter of Escherichia coli K-12. *Journal of Bacteriology* **172:** 5706-5713.

Sogaard-Andersen, L., Martinussen, J., Mollegaard, N.E., Douthwaite, S.R., and Valentin-Hansen, P. (1990) The CytR repressor antagonizes cyclic AMP-cyclic AMP receptor protein activation of the deoCp2 promoter of Escherichia coli K-12. *J Bacteriol* **172:** 5706-5713.

Sperandio, V., Li, C.C., and Kaper, J.B. (2002) Quorum-sensing Escherichia coli regulator A: a regulator of the LysR family involved in the regulation of the locus of enterocyte effacement pathogenicity island in enterohemorrhagic E. coli. *Infect Immun* **70:** 3085-3093.

Sperandio, V., Mellies, J.L., Nguyen, W., Shin, S., and Kaper, J.B. (1999) Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic Escherichia coli. *Proc Natl Acad Sci U S A* **96:** 15196-15201.

Sperandio, V., Kaper, J.B., Bortolini, M.R., Cruz Neves, B., Keller, R., and Trabulsi, L.R. (1998) Characterization of the locus of enterocyte effacement (LEE) in different enteropathogenic Escherichia coli (EPEC) and Shiga-toxin producing Escherichia coli (STEC) serotypes. *FEMS Microbiol Lett* **164:** 133-139.

Sperandio, V., Mellies, J.L., Delahay, R.M., Frankel, G., Crawford, J.A., Nguyen, W., and Kaper, J.B. (2000) Activation of enteropathogenic Escherichia coli (EPEC) LEE2 and LEE3 operons by Ler. *Mol Microbiol* **38:** 781-793.

Sperandio, V., Torres, A.G., Jarvis, B., Nataro, J.P., and Kaper, J.B. (2003) Bacteria–host communication: The language of hormones. *Proceedings of the National Academy of Sciences* **100:** 8951-8956.

Steitz, J.A., and Jakes, K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. *Proceedings of the National Academy of Sciences* **72:** 4734-4738.

Stoebel, D.M., Free, A., and Dorman, C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154:** 2533-2545.

Stoebel, D.M., Free, A., and Dorman, C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154:** 2533-2545.

Sueoka, N., and Quinn, W.G. (1968) Membrane attachment of the chromosome replication origin in Bacillus subtilis. *Cold Spring Harb Symp Quant Biol* **33:** 695-705.

Swinger, K.K., and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol* **14:** 28-35.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., *et al.* (2010) Crystal structure of bacterial RNA polymerase bound with a transcription

inhibitor protein. *Nature* **468:** 978-982.

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

Tatsuno, I., Nagano, K., Taguchi, K., Rong, L., Mori, H., and Sasakawa, C. (2003) Increased adherence to Caco-2 cells caused by disruption of the yhiE and yhiF genes in enterohemorrhagic Escherichia coli O157:H7. *Infect Immun* **71:** 2598-2606.

Thompson, C.C., Silva, G.G., Vieira, N.M., Edwards, R., Vicente, A.C.P., and Thompson, F.L. (2013) Genomic taxonomy of the genus prochlorococcus. *Microb Ecol* **66:** 752-762.

Toro, E., and Shapiro, L. (2010) Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol* **2:** a000349.

Trabulsi, L.R., Keller, R., and Gomes, T.A.T. (2002) 10.321/eid0805. Typical and Atypical Enteropathogenic Escherichia coli **8:** 508.

Tree, J.J., Wolfson, E.B., Wang, D., Roe, A.J., and Gally, D.L. (2009) Controlling injection: regulation of type III secretion in enterohaemorrhagic Escherichia coli. *Trends Microbiol* **17:** 361-370.

Tsai, A., Petrov, A., Marshall, R.A., Korlach, J., Uemura, S., and Puglisi, J.D. (2012) Heterogeneous pathways and timing of factor departure during translation initiation. *Nature* **487:** 390-393.

Turner, S.M., Scott‐Tucker, A., Cooper, L.M., and Henderson, I.R. (2006) Weapons of mass destruction: virulence factors of the global killer enterotoxigenic Escherichia coli. *FEMS Microbiol Lett* **263:** 10-20.

Umanski, T., Rosenshine, I., and Friedberg, D. (2002) Thermoregulated expression of virulence genes in enteropathogenic Escherichia coli. *Microbiology* **148:** 2735-2744.

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:** 6325-6330.

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:** 712-719.

Vitreschak, A.G., Lyubetskaya, E.V., Shirshin, M.A., Gelfand, M.S., and Lyubetsky, V.A. (2006) Attenuation regulation of amino acid biosynthetic operons in proteobacteria: comparative genomics analysis. *FEMS Microbiol Lett* **234:** 357-370.

Vollmer, W. (2008) Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiol Rev* **32:** 287-306.

Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman,

S.R., Zhang, Y., *et al.* (2014) Transcription. Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* **344:** 1285-1289.

Wade, J.T., Belyaeva, T.A., Hyde, E.I., and Busby, S.J. (2001) A simple mechanism for co- dependence on two activators at an Escherichia coli promoter. *EMBO J* **20:** 7160-7167.

Wagner, R. (1984) *Transcription Regulation in Prokaryotes*. Oxford, UK: Oxford University Press.

Walters, M., and Sperandio, V. (2006) Autoinducer 3 and epinephrine signaling in the kinetics of locus of enterocyte effacement gene expression in enterohemorrhagic Escherichia coli. *Infect Immun* **74:** 5445-5455.

Wang, X., Liu, X., Possoz, C., and Sherratt, D.J. (2006) The two Escherichia coli chromosome arms locate to separate cell halves. *Genes Dev* **20:** 1727-1731.

White, M.A., Eykelenboom, J.K., Lopez-Vernaza, M.A., Wilson, E., and Leach, D.R. (2008) Non-random segregation of sister chromosomes in Escherichia coli. *Nature* **455:** 1248-1250.

Wong, A.R., Pearson, J.S., Bright, M.D., Munera, D., Robinson, K.S., Lee, S.F., *et al.* (2011) Enteropathogenic and enterohaemorrhagic Escherichia coli: even more subversive elements. *Mol Microbiol* **80:** 1420-1438.

Wülfing, C., and Plückthun, A. (1994) Protein folding in the periplasm of Escherichia coli. *Mol Microbiol* **12:** 685-692.

Yamamoto, T., and Yokota, T. (1983) Plasmids of enterotoxigenic Escherichia coli H10407: evidence for two heat-stable enterotoxin genes and a conjugal transfer system. *J Bacteriol* **153:** 1352-1360.

Young, B.A., Gruber, T.M., and Gross, C.A. (2004) Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science* **303:** 1382-1384.

Yu, J., and Oster, G. (2012) A small post-translocation energy bias aids nucleotide selection in T7 RNA polymerase transcription. *Biophys J* **102:** 532-541.

Zaug, A.J., and Cech, T.R. (1986) The intervening sequence RNA of Tetrahymena is an enzyme. *Science* **231:** 470-475.

Zhang, N., and Buck, M. (2015) A Perspective on the Enhancer Dependent Bacterial RNA Polymerase **5:** 1012-1019.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012) Structural basis of transcription initiation. *Science* **338:** 1076-1080.

Zhu, C., Agin, T.S., Elliott, S.J., Johnson, L.A., Thate, T.E., Kaper, J.B., and Boedeker, E.C. (2001) Complete nucleotide sequence and analysis of the locus of

enterocyte Effacement from rabbit diarrheagenic Escherichia coli RDEC-1. *Infect Immun* **69:** 2107-2115.

Index

Host attachment and fluid shear are integrated into a mechanical signal regulating virulence in *Escherichia coli* O157:H7

Ghadah Alsharif[a], Sadia Ahmad[a], Md Shahidul Islam[a,b], Riddhi Shah[a], Stephen J. Busby[a], and Anne Marie Krachler[a,1]

[a]Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, B15 2TT Birmingham, United Kingdom; and [b]Department of Biotechnology, Bangladesh Agricultural University, Mymensingh 2202, Bangladesh

Edited by Frederick M. Ausubel, Harvard Medical School, Massachusetts General Hospital, Boston, MA, and approved March 25, 2015 (received for review December 2, 2014)

Enterohemorrhagic *Escherichia coli* (EHEC) is a foodborne pathogen causing hemorrhagic colitis and hemolytic uremic syndrome. EHEC colonizes the intestinal tract through a range of virulence factors encoded by the locus of enterocyte effacement (LEE), as well as Shiga toxin. Although the factors involved in colonization and disease are well characterized, how EHEC regulates its expression in response to a host encounter is not well understood. Here, we report that EHEC perceives attachment to host cells as a mechanical cue that leads to expression of LEE-encoded virulence genes. This signal is transduced via the LEE-encoded global regulator of LEE-encoded regulator (Ler) and global regulator of Ler and is further enhanced by levels of shear force similar to peristaltic forces in the intestinal tract. Our data suggest that, in addition to a range of chemical environmental signals, EHEC is capable of sensing and responding to mechanical cues to adapt to its host's physiology.

enterohemorrhagic *Escherichia coli* | locus of enterocyte effacement | attaching/effacing pathogens | gastrointestinal infection | mechanosensing

MICROBIOLOGY

Author contributions: G.A., S.A., M.S.I., R.S., S.J.B., and A.M.K. designed research; G.A., S.A., M.S.I., R.S., and A.M.K. performed research; G.A., S.A., M.S.I., and R.S. contributed new reagents/analytic tools; G.A., S.A., M.S.I., R.S., S.J.B., and A.M.K. analyzed data; and G.A., S.J.B., and A.M.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: a.krachler@bham.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1422986112/-/DCSupplemental.