TRANSCRIPTION REGULATION OF THE EXPRESSION OF THE PLASMID ENCODED TOXIN OF ENTEROAGGREGATIVE ESCHERICHIA COLI

by

Jacob Henry Duddy

A thesis submitted to
The University of Birmingham
for the degree of
MASTER OF PHILOSOPHY

February 2013

School of Biosciences
The University of Birmingham

Abstract

The pathogenic properties of Enteroaggregative *Escherichia coli* strain 042 results from the synchronised expression of virulence factors, which include the Plasmid Encoded Toxin. Pet is a member of the serine protease autotransporter of the Enterobacteriaceae family and contributes to infection by cleaving α-fodrin, disrupting the actin cytoskeleton of host cells. The expression of Pet is induced by global transcription factor CRP with further enhancement by the nucleoid associated protein Fis. This study identifies the residues of RNA polymerase, Fis and CRP required for the induction of transcription, thereby clarifying the mechanism of activation employed by the transcription factors. Fis activates transcription from the Fis binding site via a direct interaction with RNA polymerase, facilitated by protein specific determinants. This interaction is dependent on the position of the Fis binding site on the DNA and it subsequent orientation on the helical face of the DNA. CRP induces transcription from its binding site via direct interactions with RNA polymerase, facilitated by protein specific determinants.

Acknowledgements

Firstly I offer my gratitude to my supervisor, Prof Stephen Busby, whose inexhaustible patience and support were invaluable. This work would not be possible without the support of those in my life both within and outside of the lab. My thanks lies with the Henderson laboratory who offered their knowledge and time; Amanda Rossiter in particular. My thesis would not of been completed though without the constant support, financially and emotionally, of my parents. Special thanks lies with Katy Madden and her housemate whose hugs from the former and tea from the latter; the impact of which is immeasurable.

Table of Contents

Chapter 1: Introduction 1-16

1.1 ***Escherichia coli*** 2

1.2 **Pathogenic *Escherichia coli*** 3

1.3 **Gene Expression** 5

1.4 **The Transcription Cycle** 7

1.4.1 *Formation of the Initiation Complex* 7

1.4.2 *Promoter Clearance* 9

1.4.3 *Elongation* 9

1.4.4 *Termination* 10

1.4.5 *Promoter Strength and Regulation of Promoters* 11

1.4.6 *cAMP Receptor Protein* 13

1.4.7 *Factor for Inversion Stimulation* 13

1.5 **Enteroaggregative *Escherichia coli*** 14

1.5.1 *Plasmid Encoded Toxin* 14

1.6 **Aims** 16

Chapter 2: Materials and Methods 17-36

2.1 **Suppliers** 18

2.2 **Buffers, Solutions and Reagents** 18

2.2.1 Gel Electrophoresis of DNA 18

2.2.2 Extraction and Purification of Nucleic Acids 19

2.2.3 β-galactosidase Assays 19

2.2.4 Preparing Competent Cell by Rubidium Chloride Method 19

2.3 **Bacterial Growth Media** 19

2.3.1 Liquid media 19

2.3.2 Solid media 19

2.3.3 Antibiotics 20

2.4 **Bacterial Strains** 20

2.5 **Plasmids** 20

2.6 Gel Electrophoresis of DNA 20

2.6.1 Agarose Gel Electrophoresis of DNA 20

2.6.2 Polyacrylamide Gel Electrophoresis of DNA 20

2.7 Extraction and Purification of Nucleic Acids 23

2.7.1 Phenol/Chloroform Extraction of DNA 23

2.7.2 Ethanol Precipitation of DNA 23

2.7.3 Purification of DNA using QIAquick PCR Purification Kit 23

2.7.4 Electroelution of DNA from Polyacrylamide Gels 23

2.7.5 Small-Scale Preparation of Plasmid DNA ("mini-prep") 24

2.7.6 Large-Scale Preparation of Plasmid DNA ("maxi-prep") 24

2.8 Bacterial Transformations 24

2.8.1 Preparation of Competent Cells Using Rubidium Chloride Method 24

2.8.2 Transformation of Competent Cells with Plasmid DNA 25

2.9 DNA Manipulations 25

2.9.1 Polymerase Chain Reaction 25

2.9.2 Restriction Digestion of DNA 25

2.9.3 DNA Ligations 25

2.9.4 DNA Sequencing 26

2.9.5 Introduction of Point Mutations Using Megaprimer PCR 26

2.10 Cloning of Promoter Fragments into pSR and pRW50 26

2.10.1 Preparation of EcoRI-HindIII Promoter Fragments by PCR 26

2.10.2 Construction of aer1 and FnCC(-40.5) derivatives 31

2.11 β-galactosidase Assays 31

Chapter 3: Regulatory Region of the EAEC Pet Operon 37-43

3.1 Introduction 38

3.2 Spatial Constrains of Fis Activation 38

3.3 Contributions of Fis DNA Binding Sites to Expression 38

3.4 Discussion 40

Chapter 4: Emulation of Pet Promoter Architecture in a Model Promoter 44-54

4.1 Introduction 45

4.2 Emulation of PET Promoter Architecture 45

4.3 Activity of FnCC(-40.5) Promoter Series 45

4.4 Effect of α-CTD alanine substitution mutants on the expression of the FnCC(-40.5) Series 47

4.5 Effect of the Substitution of Critical Fis Residues on Expression of F_{25}CC(-40.5) 47

4.6 Discussion 50

Chapter 5: Conclusions 55-58

References 59-70

Index of Figures

Figure 1.1: Structural overview of RNAP and σ 6

Figure 2.1: pRW50 22

Figure 2.2: pSR 22

Figure 2.3: Megaprimer PCR 30

Figure 2.4: aer1 34

Figure 2.5: aer1 –fisI 34

Figure 2.6: aer1 –fisII 34

Figure 2.7: aer1 –fisI/II 34

Figure 2.8: aer1 Δ5 34

Figure 2.9: aer1 Δ10 34

Figure 2.10: CC(-41.5) 35

Figure 2.11: CC(-40.5) 35

Figure 2.12: F_{15}CC(-40.5) 35

Figure 2.13: F_{24}CC(-40.5) 35

Figure 2.14: F_{25}CC(-40.5) 35

Figure 2.15: F_{26}CC(-40.5) 35

Figure 2.16: F_{30}CC(-40.5) 35

Figure 2.17: F_{35}CC(-40.5) 36

Figure 2.18: F_{39}CC(-40.5) 36

Figure 2.19: F_{40}CC(-40.5) 36

Figure 2.20: F_{41}CC(-40.5) 36

Figure 2.21: F_{45}CC(-40.5) 36

Figure 2.22: $Fc_{66}F_{35}$CC(-40.5) 36

Figure 3.1: Point Mutations and Nested Deletions Within the aer1 Fragment 39

Figure 3.2: Deletions Within the aer1 Fragment 41

Figure 3.3: Point Mutations Within the aer1 Fragment 41

Figure 3.4: Fis Regulatory Sites in *E. coli* K12 43

Figure 4.1: The FnCC(-40.5) series 46

Figure 4.2: Implications of a Second Upstream Fis Binding Site 48

Figure 4.3: Transcriptional Activity of the FnCC(-40.5) Series 48

Figure 4.4: Impact of Alanine Substitutions Within α on FnCC(-40.5) Series 49

Figure 4.5: Impact of Amino Acid Substitutions Within Fis on F_{25}CC(-40.5) 53

Figure 5.1: Model of Transcription Regulation at the *pet* Promoter 57

Index of Tables

Table 2.1: *Escherichia coli* K12 strains used in this study 21

Table 2.2: Plasmids used in this study 21

Table 2.3: Oligonucleotides Primers Used In This Study 27

Table 2.4: PCR Cycle 29

Table 2.5: EcoRI-HindII fragments used in this study 33

List of Abbreviations

A/C/G/T/U	adenosine/cytosine/guanine/thymidine/uracil
APS	ammonium persulphate
cAMP	3',5' cyclic adenosine monophosphate
bp	Nucleotide Triphosphate base pairs
CRP	cAMP receptor protein
CTD	carboxy-terminal domain
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	2'-deoxynucleotide 5'-triphosphate (N=A,C,G or T)
dsDNA	double stranded deoxyribonuclease
EC	elongation complex
GI	gastro-intestinal
Kbp	kilobase
LB	lennox broth
mRNA	messenger ribonucleic acid
nt	nucleotide triphosphate
NTP	nucleoside 5'-triphosphate (N = A, C, G or T)
NTD	amino-terminal domain
OD	optical density
ONPG	2-nitrophenyl-β-D-galatopyranoside
PAGE	polyacrylamide gel
PAI	pathogenicity island
PCR	polymerase chain reaction
RNA	ribonucleic acid
RNAP	RNA polymerase
RP_c	RNAP-promoter closed complex
RP_o	RNAP-promoter open complex
rpm	revolutions per minute
S.D	standard deviation
TE	Tris EDTA buffer

TEMED	N,N,N’,N’-tetramethylethylenediamine
tRNA	transfer ribonucleic acid
UV	ultraviolet
v/v	volume per volume
WT	wild type
w/v	weight per volume

Amino acids are given as their designated single letter codes

Chapter 1

Introduction

1.1 *Escherichia coli*

Escherichia coli is a rod-shaped, gram-negative, facultative anaerobe of the class gammaproteobacteria. *E. coli* specifically colonises endotherms as a commensal or strict pathogen depending on the strain; commensal colonisation of the human lower gastro-intestinal (GI) tract occurring soon after birth (Bettelheim *et al.* 1974). Transmission occurs through an oral-faecal route, with sporadic transition through water and sediment, between hosts; environmental sources such as these are estimated at harbouring half of the entire *E. coli* population (Savageau 1983; Luna *et al.* 2010). Within this environmental population, there are thought to be strains present that are functioning saprophytes (Solo-Gabriele *et al.* 2000; Power *et al.* 2005). Due to the ease of cultivation and manipulation within the laboratory, the genome of *E. coli* was one of the first sequenced; cementing *E. coli* K12 as a founding pillar of modern molecular microbiology (Blattner *et al.* 1997).

In the human intestine a myriad of bacteria make up the microbiota, with the following distribution of phylum: *Firmicutes* (64%), *Bacteroidetes* (23%), *Proteobacteria* (8%), *Actinobacteria* (3%), and other minor taxonomic divisions where high diversity is observed (Eckburg *et al.* 2005). The host and its gut flora have a symbiotic relationship. This deep rooted relationship is exemplified by the dependence of the host on its intestinal microbiota, which expresses genes that are functionally essential for but absent from the host (Backhed *et al.* 2005; Gill *et al.* 2006; Kurokawa *et al.* 2007; Claus *et al.* 2011). This symbiotic relationship also provides protection against pathogenic organisms, since nutritional niches and attachment sites of the host gut are monopolised by the resident commensal community (Hudault *et al.* 2001; Asahara *et al.* 2004; Stecher *et al.* 2007; Leatham *et al.* 2009). Humans, at any one time, can be simultaneously colonised by at least several commensal *E. coli* strains. This includes a resident strain that is typically predominant and present for months, or years, as well as transient strains, that are present for only a few days or weeks (Sears *et al.* 1950; Sears and Brownlee 1952; Sears *et al.* 1956; Bettelheim *et al.* 1972; Caugant *et al.* 1981). These commensal strains are located within the mucosa of the large intestine, specifically, in the caecum and the colon. In this nutritional niche *E. coli* is the predominant facultative anaerobe, making up 0.1% of the indigenous microbiota (Freter *et al.* 1983; Poulsen *et al.* 1994; Eckburg *et al.* 2005). The predominance of *E. coli,* in the niche of the mucosa, is thought to have arisen due to the efficiency with which it utilises gluconate for respiration, via anaerobic mixed acid fermentation (Sweeney *et al.* 1996; Chang *et al.* 2004).

The interior of *E. coli* cells consist of the liquid cytoplasm, within which the majority of the macromolecular machinery of the cell resides; enabling cell growth, metabolism, and replication (Vendeville *et al.* 2011). The

cytoplasm is encased by a sheath of an inner and outer membrane, separated by an intermediary periplasmic space, containing peptidoglycan. The outer membrane is in contact with the external environment and acts as a selective barrier; facilitating the uptake of beneficial molecules whilst excluding harmful molecules. Proteins including flagella, adhesins, pores, sensors, amongst others are dotted across the outer membrane, imbued with diverse functions (van Bloois *et al.* 2011). Some strains also possess an extracellular complex polysaccharide structure, the capsule, which protects the cell from desiccation and phagocytosis whilst also promoting adherence to host cells (Whitfield 2006; Beloin *et al.* 2008).

E. coli has a circular, double-stranded DNA genome which forms a distinct structured unit within the cytoplasm, the nucleoid. The nucleoid is a complex consisting of the DNA genome, encoding the genes required for the survival of the organism; and the nucleoid associated proteins (NAPs) that maintain the structure of the nucleoid and modulate gene expression (Reyes-Lamothe *et al.* 2008; Browning *et al.* 2010). The genomes of many *E. coli* strains have now been sequenced and comparisons between these strains highlight the wide diversity within the species. K12 MG1655 contains 4288 annotated protein coding genes with anywhere between 4000 and 5500 genes in related strains (Blattner *et al.* 1997). Surprisingly, when 207 *E. coli* isolates were compared, a core set of only 2256 genes were common to all strains (Jackson *et al.* 2011). The remaining portion of each genome is comprised from a mosaic of insertions and deletions, fabricated from a heterogeneous mix of the 18000 genes of the dispensable genome (Rasko *et al.* 2008; Touchon *et al.* 2009). The phenotypic plasticity observed between strains is caused by variation in the combination of genes, from the pan-genome, in the strains dispensable genome.

1.2 Pathogenic *Escherichia coli*

The genomes of pathogenic *E. coli* differ from commensal strains in size and content, where the genomes can be up to 1 Mbp larger, due to the loss and gain of superfluous pan-genomic elements. The evolutionary move from commensal to pathogen is initiated by the acquisition of virulence factors which enhance microbial transmission, survival or colonisation, in context of the host, followed by the loss of genomic elements required for environmental survival reducing the viability of the organism outside of the host (Hacker and Kaper 2000; Ochman and Jones 2000; Hong *et al.* 2012). Following selective pressures, resulting from the exposure to the host immune system, which selects for variants that can deal with these pressures; these variants tend to have higher mutation and recombination rates, although this phenotype tend to tail away after beneficial mutations are gained (Cooper and Lenski 2000; Denamur *et al.* 2000; Wirth *et al.* 2006; Woods *et al.* 2011).

The evolution of bacterial genomes can occur endogenously, through mutations, rearrangements and gene duplication or exogenously by horizontal gene transfer (HGT) (Chopra *et al.* 2003; Matias Rodrigues and Wagner 2009). Mobile elements including plasmids, transposons and bacteriophages, are transferred between organisms by HGT, transferring DNA sequences, corresponding to single genes or several genes, to the recipient. Some of these genes may contribute, directly or indirectly, to the pathogenicity of a bacterium; although the classification of a factor being virulent is arbitrary as a factor that promotes virulence in the host may also promote fitness in the environment, virulence depends upon environmental context (Groisman and Ochman 1996; Hacker *et al.* 1997; Croxen and Finlay 2010). The loss and gain of mobile elements can influence the evolution of the bacterial genome, with the gain of genomic islands or single genes permitting the rapid acquisition of new traits; including amongst others, resistance to antimicrobials, or the ability to colonise and promote virulence within a host (Waldor *et al.* 1996; Boltner *et al.* 2002; Herold *et al.* 2004; Anantham and Hall 2012). The acquisition of these traits can bestow a pathogenic phenotype or provide access to previously occluded niches within a host, uncovering an evolutionary path geared towards the generation of pathogenic organisms. These mechanisms of gene acquisition, and subsequent selective pressures, have resulted in eight separate pathovars of *E. coli* which can colonise and cause disease within the GI tract, urinary tract or meninges; depending on the bespoke assortment of virulence factors ascribed to each of the pathovars. The six diarrhoeagenic pathovars are enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli (*EAEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (DAEC). The two extraintestinal pathovars are uropathogenic *E. coli* (UPEC) and meningitis-associated *E. coli* that cause urinary tract infections and meningitis/sepsis respectively.

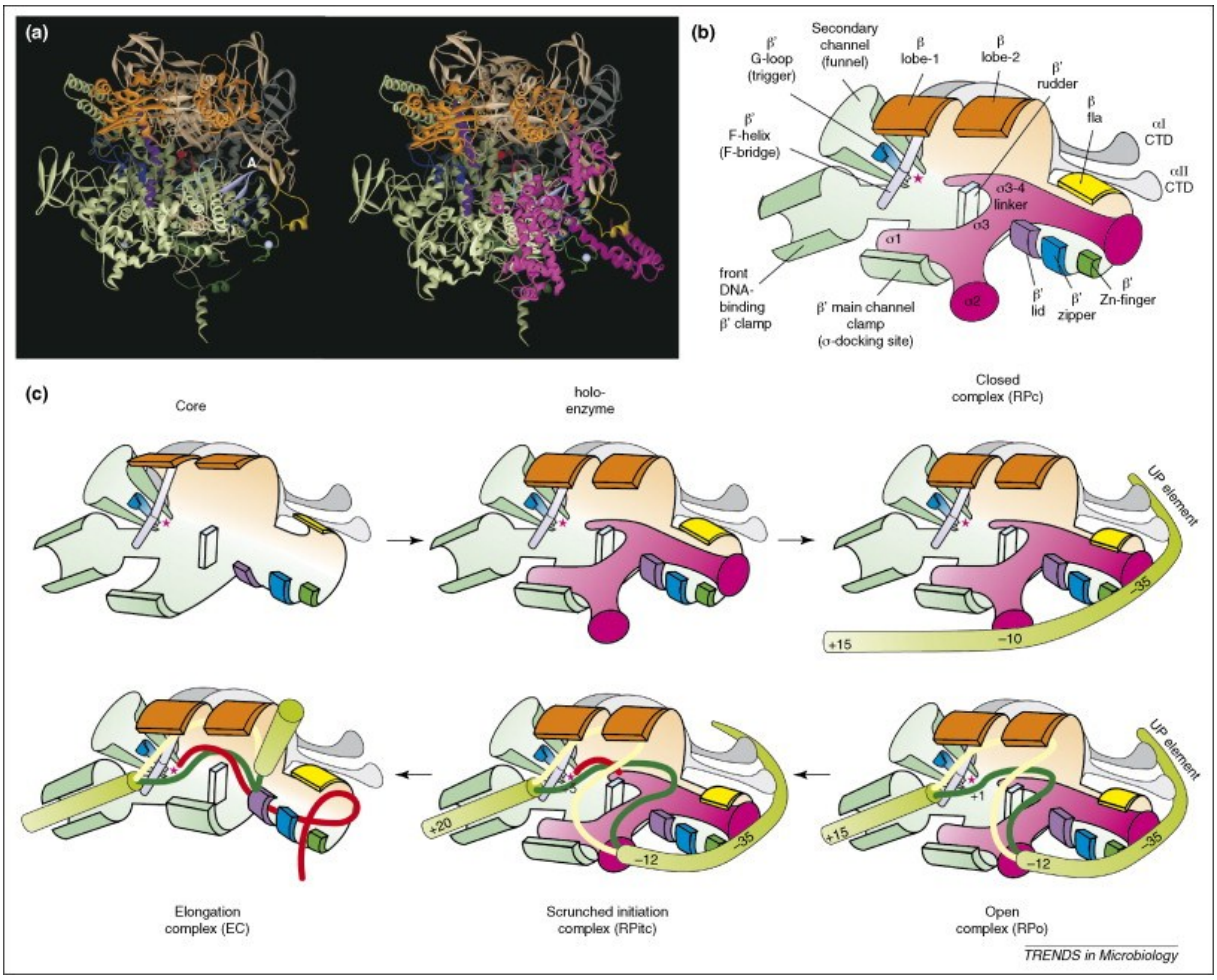
E. coli pathovars all share a common strategy in their pathogenesis and often target the same host machinery, although the individual mechanisms and virulence factors employed vary (Finlay and Falkow 1997; Croxen and Finlay 2010). This strategy starts with adherence to host cells, followed by proliferation and distribution throughout the host, aided by virulence factors that aid evasion of host defences or improve the competiveness of the pathogen to the host microbiome, occasionally causing damage to the host (Betancourt-Sanchez and Navarro-Garcia 2009; Shames and Finlay 2010). Pathogenicity is not achieved by any single virulence factor, rather it requires an ensemble; with regulatory factors synchronising the expression of virulence factors, in response to extracellular physiochemical stimuli, to be expressed in the right place at the right time. Virulence factors are regulated in response to physiochemical signals indicative of the host, inducing the cascade of

required factors to promote bacterial survival and proliferation. The regulons of the regulatory factor can overlap allowing multiple inputs to modulate the activities of specific genes, tuning their expression. Inductions of virulence genes are frequently regulated by pathogen-specific transcriptional regulators encoded on PAIs or plasmids or by regulators present in commensal *E. coli* providing a link between essential cellular processes and virulence. Taken together, a picture emerges where a constantly changing extracellular physiochemical signals are responsively countered by constantly fluctuating gene expression; the virulence factors expressed supporting the survival of the organism whilst also resulting in pathogenicity.

1.3 Gene Expression

Gene expression is the process by which a gene is firstly transcribed, where a strand of RNA whose sequence is complementary to the template strand of the gene is assembled; sequential translation, yields any resultant protein. Gene expression is regulated in response to extracellular physiochemical stimulus, adjusting the cells gene expression profile; ensuring survival. Regulation can occur at any stage of gene expression; interfering with the steps of the transcription and translation cycles or altering the half-life of the RNA transcript or the resultant protein. Transcription, is the first step in gene expression and therefore offers the most economical point for regulation to occur; the most common regulatory target being the initiation of transcription. Transcription is reliant upon RNA polymerase (RNAP), an enzyme which synthesises RNA complementary to the DNA template of the genome.

The *E. coli* core RNAP enzyme shares significant sequence and structural homology with its archaeal and eukaryotic homologs, which all resemble a 'crab claw'; yet the composition of the subunits that form this claw differ between species ($\alpha_2\beta\beta'\omega$ subunits for *E. coli* core RNAP) (Figure 1.1a) (Ebright *et al.* 1985; Ebright *et al.* 1990; Zhang and Ebright 1990; Parkinson *et al.* 1996; Passner and Steitz 1997; Chen *et al.* 2001). The pincers of the claw are formed by the β and β′ subunits, separated by a ~27 Å cleft, with the active centre at the base cleft containing the two crucial Mg^{2+} ions needed for nucleotide addition cycle during transcript elongation. Two α subunits are present, each binding either β (αI) or β' (αII), and form a dimer, holding the complex together (Ebright *et al.* 1984; Passner and Steitz 1997). Each α subunit consists of two domains connected by a flexible linker: an N-terminal domain (αNTD) responsible for dimerisation and interaction with β' and β subunits, and a C-terminal domain (αCTD) able to interact with promoter DNA and transcription factors



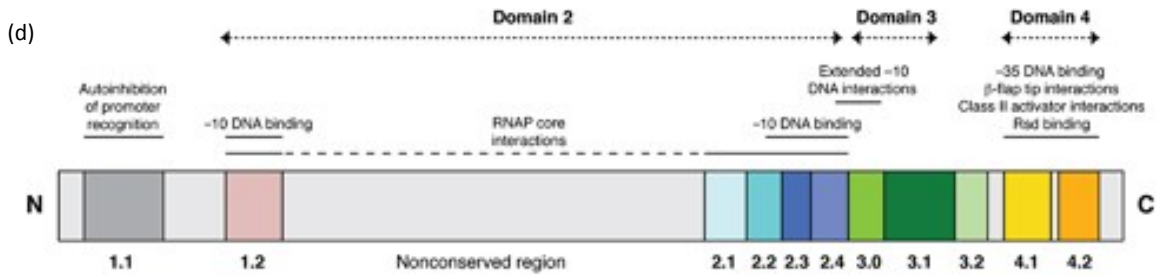


Figure 1.1: Structural overview of RNAP and σ (Borukhov and Nudler 2003; Ghosh *et al.* 2010)

(a) Crystal structure of *Thermus aquaticus* RNAP core (left) and holoenzyme (right) (b) Cartoonised representation of RNAP holoenzyme (c) The major transitional steps formed on the progression from the core RNAP to EC (d) σ^{70} family domain map, with major domain regions label with names and their interaction partners. *(RNAP subunit colours: αI, light grey; αII, grey; β, light brown; β', light green; ω, dark green; σ, magenta)*

(Caddick *et al.* 1986; Johnson *et al.* 1986a; Ishihama 1992; Browning *et al.* 2010). The small ω subunit interacts exclusively with β' near the base of the pincer, acting as a chaperone, maintaining the required conformation of β' to be incorporated into the $\alpha_2\beta$ complex during the formation of the core enzyme (Ghosh *et al.* 2001; Ghosh *et al.* 2003; Mathew and Chatterji 2006).

1.4 *The Transcription Cycle*

Transcription can be separated into 3 major steps each characterised by several structural transitions of RNAP; formation of the initiation complex and promoter clearance, transcript elongation and termination. For transcription to occur the elongation complex (EC) must be assembled and translocated along the DNA, catalysing the synthesis of the RNA transcript as it goes (Figure 1.1c) (Ebright *et al.* 1987). Initially RNAP binds the promoter DNA, just upstream of the gene, forming the closed initiation complex (RP_c), isomerisation of which gives the open initiation complex (RP_o) able to transition to the subsequent Elongation Complex (EC). Termination of transcription is facilitated by various mechanisms, all of which cause RNAP to release the RNA transcript and DNA template, ready for the next round of transcription.

1.4.1 Formation of the Initiation Complex

The core RNAP enzyme is catalytically competent for transcription, but the complex still lacks the essential ability to bind specifically and strongly to promoter DNA (Ishihama 1990). Transient binding of a σ factor to the core to form the holoenzyme, imbuing the complex with ability to bind, specificity, to promoters within the σ factors regulon, whilst inhibiting nonspecific DNA interaction (Figure 1.1a,b) (Parkinson *et al.* 1996; Ferguson *et al.* 2000; Wade *et al.* 2006). Bacteria can encode multiple σ factors, belonging to either the σ^{70} or σ^{54} family; which differ in sequence, structure and mechanism of transcription initiation (Wosten 1998; Ghosh et al. 2010). In *E. coli*, σ^{70} is the housekeeping σ factor; its regulon encompassing the genes required for normal growth (Reppas et al. 2006). σ^{70} family factors are formed of four conserved domains (σ1, σ2, σ3 and σ4) connected by flexible linkers (Figure 1.1 d) (Lonetto et al. 1992; Campbell et al. 2002; Paget and Helmann 2003). σ^{70} binds the RNAP core through several independent contacts; the σ2 region conveying the major contact with the β' (Zhang and Ebright 1990; Malhotra *et al.* 1996; Severinova *et al.* 1996; Sharp *et al.* 1999; Campbell *et al.* 2002; Opalka *et al.* 2010). When σ is unbound, the σ1.1region mimics DNA, occluding the regions essential for promoter binding (σ2 and σ4), inhibiting promoter association; upon formation of the holoenzyme, σ1.1 undergoes a conformational change becoming located within the active-site channel of RNAP, exposing σ2 and σ4, de-

repressing promoter recognition (Dombroski *et al.* 1993; Parkinson *et al.* 1996; Gopal and Chatterji 1997; Bowers and Dombroski 1999; Mekler *et al.* 2002).

Following its formation, the holoenzyme tracks along the DNA until it locates a promoter; characterised by a set of promoter sequence elements, specific for the bound σ factor, each of which conveys an additive contribution to binding, forming the resultant RP_c (Guthold *et al.* 1999; Sakata-Sogawa and Shimamoto 2004; Miroslavova and Busby 2006; Hook-Barnard and Hinton 2007). σ^{70} promoters possess the hexameric elements, −35 and −10, separated by a spacer of 17(±1) nt (consensus sequences $^{-35}$TTGACA^{-30} and $^{-12}$TATAAT^{-7} respectively), which interact with σ4.2 and σ2.4 respectively (Rosenberg and Court 1979; McClure 1985; Siegele *et al.* 1989; Lisser and Margalit 1993; Dombroski 1997). The precise sequence of the spacer normally has little impact on transcription but strict conservation of the length is essential (deHaseth and Helmann 1995; Hook-Barnard and Hinton 2009; Singh *et al.* 2011; Sztiller-Sikorska *et al.* 2011; Yuzenkova *et al.* 2011). Besides the −10 and −35 elements, other elements can be present such as extended −10 ($^{-15}$TGn^{-14}) and the GC-rich discriminator (ranging from positions +1 to +20), both of which interact with σ^{70}, and the UP element (variable region upstream of the -35 element) which interacts with the αCTD (Travers 1984; Ross *et al.* 1993; Barne *et al.* 1997; Estrem *et al.* 1998; Burr *et al.* 2000; Mitchell *et al.* 2003; Haugen *et al.* 2008). Promoters generally contain a combination of a few, but not all, of these elements, each of which contributes to the affinity of the holoenzyme to the promoter. The following step, the isomerisation of RP_c to RP_o, is initiated by the melting of the DNA duplex at the -10 position. This and subsequent steps are driven by the favourable transition from RP_c to RP_o (Feklistov and Darst 2011; Chakraborty *et al.* 2012). Upon melting, the highly conserved -11A, of the -10 element flips, associating with σ2.3, stabilising the melted conformation of the strand and forming the upstream edge of the transcription bubble (Matlock and Heyduk 2000; Chen *et al.* 2001; Lim *et al.* 2001; Lee *et al.* 2004; Schroeder *et al.* 2008; Schroeder *et al.* 2009). The template strand is then loaded into the active site of RNAP, interacting with σ2.3 further contributing to stability of the transcription bubble (Aiyar *et al.* 1994; Chen *et al.* 2001; Murakami and Darst 2003). The dsDNA downstream of the -10 element scrunches into the active-site channel, further opening the transcription bubble from -10 to +2, whilst strain accumulates (Chen *et al.* 2010). This strain is released by the bending and kinking of downstream DNA (Rees *et al.* 1993; Rivetti *et al.* 1999). The bent dsDNA is loaded into the active-site channel by a proposed "rope-swing" mechanism, where repulsion between the dsDNA and σ1.1 stimulates segments of β subunit to transiently rearrange, widening the active-site channel sufficiently for loading the dsDNA (Murakami and Darst 2003; Chen *et al.* 2010). The width of the channel is restored, following the loading of the DNA, locking it firmly in place, displacing σ1.1 from the channel; denoting the first

step towards the dissociation of σ from the core (Mekler *et al.* 2002; Chen *et al.* 2010). The isomerisation event concludes, when the +1 transcription start site is aligned in the correct position of the active site, ready for transcription (Naryshkin *et al.* 2000; Chen *et al.* 2001; Murakami and Darst 2003; Vignon *et al.* 2003).

1.4.2 Promoter Clearance

Upon formation of the RP_o, NTP is supplied to the active site through the secondary channel, initiating the nucleotide addition cycle and initial synthesis of the RNA transcript. During this time RNAP remains stationary, bound to the promoter, and the downstream template DNA, required for transcript elongation, is scrunched into the complex, accumulating stress (Straney and Crothers 1987; Cheetham and Steitz 1999; Kapanidis *et al.* 2006; Revyakin *et al.* 2006; Tang *et al.* 2008). As the transcript length reaches 12 nt it encounters the RNA exit channel, which is blocked by σ3.2, and the releasing of the accumulated stress leads to one of two outcomes; the transcript can successful dislodge the σ3.2 loop and synthesis can continue or the synthesis of the transcript can be aborted (Cashel *et al.* 2003; Murakami and Darst 2003; Kulbachinskiy and Mustaev 2006). Normally the transcript is unsuccessful in dislodging the σ3.2 loop, with successive failure resulting in the production of many short abortive transcripts (<12 nt); promoters with the highest consensus sequences have the highest affinity association between the promoter and holoenzyme, preventing dissociation and resulting in extensive abortion (Hsu *et al.* 2003; Vo *et al.* 2003; Miroslavova and Busby 2006). After reaching a length of 12 nt and displacing the σ3.2 loop, the transcript, must reach a length of 16–17 nt and displace σ4, which partly occludes the end of the RNA exit channel(Parkinson *et al.* 1996; Geszvain *et al.* 2004; Gregory *et al.* 2005). σ remains bound to the core only through the interactions of σ2, the σ2-3 linker, and σ3 the stochastic loss of which results in the total dissociation of σ, permits translocation, signifying the formation of the EC. Occasionally σ can remaining bound to the core, translocating for significant distances, and contributing to proximal-pausing (Shimamoto *et al.* 1986; Chen and Ebright 1993; Ring *et al.* 1996; Brodolin *et al.* 2004; Kapanidis *et al.* 2005; Raffaelle *et al.* 2005).

1.4.3 Elongation

Following promoter escape, transcript elongation occurs with transient pauses punctuating transcript synthesis; occurring when there is misalignment of the additive nucleotide, downstream promoter-like elements or the EC adopts the elemental pause conformation (Ring *et al.* 1996; Adelman *et al.* 2002; Herbert *et al.* 2006; Xie 2008; Walmacq *et al.* 2009). Entry into the elemental pause conformation occurs through the isomerisation of the active site causing inhibition of nucleotide addition, where the conformation of the G (trigger)-loop is altered, which normally acts in tandem with the F (bridge)-helix, facilitating nucleotide addition through a brownian

ratchet mechanism (Komissarova and Kashlev 1997; Abbondanzieri *et al.* 2005; Bar-Nahum *et al.* 2005; Toulokhonov *et al.* 2007; Xie 2008; Landick 2009). Pausing can be prolonged by transcript backtracking, RNA hairpin formation or regulator binding (Chan and Landick 1989; Artsimovitch and Landick 2000; Artsimovitch and Landick 2002; Bai *et al.* 2004; Landick 2009; Maoileidigh *et al.* 2011). RNAP can be restored to an active conformation, from the backtracked state, by GreB or Mfd, which restore the 3' register of the transcript within the active site; entry into a paused confirmation is partially prevented by GreA (Borukhov *et al.* 1993; Selby and Sancar 1993; Marr and Roberts 2000; Toulme *et al.* 2000; Park *et al.* 2002; Laptenko *et al.* 2003; Roberts and Park 2004). GreB binds RNAP, inserting a Mg^{2+} topped protrusion through the secondary channel, swapping with and acting as a substitute for the G(trigger)-loop, cleaving the nascent RNA, generating a new 3′-terminus (Steitz and Steitz 1993; Opalka *et al.* 2003; Sosunova *et al.* 2003; Vassylyeva *et al.* 2007; Roghanian *et al.* 2011). Mfd binds the duplex DNA immediately upstream of the transcription bubble, binding RNAP and then ratchets the RNAP forward, causing the reversal of the backtracking until the 3' terminus is in proper register within the active site or in an intractable roadblock causing RNAP to dissociates from the DNA (Roberts and Park 2004). Pausing is a necessary part of elongation as it allows translation to be coupled with transcription through the interaction between the RNAP bound NusG, and the ribosomal NusE; preventing prolonged pausing and Rho-dependent termination.

1.4.4 Termination

Throughout elongation the EC enters pause phases, although the complex is rescued, resuming transcription, but distinct sequence signals causes a pause that cannot be rescued; committing the EC to the transcription termination pathway (McDowell et al. 1994; Gusarov and Nudler 1999; Peters et al. 2011). These signals are either intrinsic, dependent upon the interactions between the EC and the transcript, or dependant on the recruitment and subsequent action of the Rho factor (Epshtein et al. 2007; Epshtein et al. 2010). The intrinsic mechanism of termination requires the presence of a conserved termination sequence; GC-rich inverted repeat followed by a track of U at the 3' terminus of the transcript (Brendel *et al.* 1986; Telesnitsky and Chamberlin 1989; Martinez-Trujillo *et al.* 2010). When transcribed, the U-track induces pausing providing the time for the inverted repeat region to hybridise and form a strong hairpin in the RNA exit channel (Gusarov and Nudler 1999). Hairpin formation is facilitated by the increased thermodynamic stability over the DNA-RNA hybrid and is accommodated within the exit channel due to widening, facilitated by flexibility of parts within this region (zipper, flap, and zinc finger) inducing folding of G (trigger)-loop causing RNAP to enter a prolonged, stabilised pause (Farnham and Platt 1981; Epshtein *et al.* 2007; Toulokhonov *et al.* 2007). As the hairpin grows it is

relocates into the main channel, melting the DNA-RNA hybrid, and interacting with the G (trigger)-loop, locking the EC into a unrecoverable state, inducing the DNA-clamp to open, resulting in the release RNAP from both the DNA template and RNA transcript (Bar-Nahum *et al.* 2005; Vassylyev *et al.* 2007).

Factor dependent terminating requires Rho, which forms as a homo-hexameric ring with the transcript loaded through its central cavity and translocates in a 5′→ 3′ direction along the transcript in an ATP dependant mechanism (Roberts 1969; Geiselmann *et al.* 1992; Adelman *et al.* 2006; Skordalakes and Berger 2006). Loading relies on the presence of a *rut* site; 70-80 nt length of C-rich, unstructured RNA (Morgan *et al.* 1985; Sharp *et al.* 1986; Alifano *et al.* 1991; Kim and Patel 2001). Subsequent binding of Rho to the EC is essential for termination, but whether Rho forms this complex prior to translocation or whether translocation precedes complex formation is unclear (Mooney *et al.* 2009; Epshtein *et al.* 2010; Kalyani *et al.* 2011). In either case the rates of translocation and transcription are coupled, so that the rate of translocation must exceed transcription for termination to occur (Jin *et al.* 1992; Sashni *et al.* 2012). Termination occurs exclusively within the termination region, 60–90 nt downstream of the *rut* site, which pauses the EC (Galloway and Platt 1988; Banerjee *et al.* 2006; Ciampi 2006). The type of pause is important in Rho-dependant termination, with the ability of the paused EC to dissociation reduced, when due to irreversible backtracking or RNA hairpin formation (Kassavetis and Chamberlin 1981; Dutta *et al.* 2008). The release of the paused EC occurring only when the EC is restored to a catalytically competent state, by the forces produced by Rho through ATP-dependant translocation (Dutta *et al.* 2008).

1.4.5 Promoter Strength and Regulation of Promoters

Bacteria can encode multiple σ factors, belonging to the σ^{70} or σ^{54} family. *E. coli* possesses one σ^{54} factor and six σ^{70} family factors (σ^{70}, σ^{38}, σ^{32}, σ^{28}, σ^{24}, σ^{19}) (Gruber and Gross 2003; Paget and Helmann 2003). Each σ factor has a degenerate regulons defined by presence of factor specific promoter elements. Under normal physiological conditions σ^{70} is highly expressed stimulating the expression of its regulon of housekeeping gene, when stricken by physiochemical stress, a stress response σ factor is expressed accessing its regulon, rapidly reprogramming the profile of gene expression to respond to the stress. Promoters within a regulon of each of the σ factors display varying “strengths”, with strong promoters inducing the highest levels of gene expression. The strength of a promoter is determined by the affinity of the promoter to RNAP and its isomerisation efficiency. Promoters become stronger as the conserved promoter elements (such as UP, −35, −10 and extended -10 elements for σ^{70}) approach full consensus although this strength declines when the interactions between the RNAP and promoter

are sufficiently strong to hinder promoter escape (Miroslavova and Busby 2006). The quantities of available RNAP are limited and therefore prudent and adaptive distribution of RNAP to promoters is essential (Ishihama 2000; Grainger *et al.* 2005). Therefore, promoters are made up of different combinations of non-consensus promoter elements; essential genes being the strongest and response specific genes weakest, but whose activity can be induced in response to stimuli by transcription factors.

Transcription factors are responsible for the modulation of expression of a promoter in response to stimuli and therefore their expression and allosteric activation are regulated to avoid inappropriate cellular responses (Busby and Ebright 1997). Transcription factors affect the activity of a promoter negatively or positively by modulating the strength of the promoter; altering the affinity to RNAP or affecting the efficiency of the progression to the elongation complex. Transcription factors bind sequence specific sites within the promoter and directly or indirectly modulate promoter activity (Ishihama 1992; Kahramanoglou *et al.* 2011). The position of the transcription factor binding site is important where activators can repress from specific locations (Bendtsen *et al.* 2011). Repressors bind downstream of promoters or over promoter elements interfering with transcription and/or elongation. Activators normally bind upstream of the RNAP binding elements increasing the affinity between the promoter and RNAP or supporting isomerisation. Activators can act via many diverse mechanisms at σ^{70}-dependent promoters, including class I and class II mechanism, functioning through direct interaction between the transcription factor and RNAP holoenzyme (Browning and Busby 2004). Class II activators increase promoter affinity by binding just upstream and overlapping the -35 element and interact with the proximal α subunit and often domain 4 of σ^{70} (Busby and Ebright 1997; Dove *et al.* 2003). Class I activators bind further upstream and interact with the distal α subunit; the position of the binding site can be variable due to the flexibility of the α subunits linker domain (Ebright 1993). Regulation of a promoter is not limited to one transcription factor; the combined input of multiple factors tunes a promoters' activity in response to a range of stimuli. Regulons of transcription factors can ranges in size, anywhere from 1 to 100+ promoters, the latter regulated by transcription factors coined global regulators (Martinez-Antonio and Collado-Vides 2003). Two such global regulators are the Factor for Inversion Stimulation (Fis) and cAMP Receptor Protein (CRP), each of which is expressed sequentially through growth; CRP during lag phase and Fis early log phase (Ball *et al.* 1992; Azam and Ishihama 1999).

1.4.6 cAMP Receptor Protein

CRP is a global transcription regulator whose regulon is comprised of at least 180 genes with a diverse range of functions, many of which are involved in utilisation of secondary carbon sources (Grainger *et al.* 2005; Deutscher 2008). CRP forms a homodimer in solution, each dimer formed of two domains; the N-terminal domain (1–133) responsible for subunit dimerisation and cAMP binding and the C-terminal domain (139–209) for interactions with the palindromic 16 bp CRP DNA binding site (TGTGAnnnnnnTCACA) (Ebright *et al.* 1989; Zhou *et al.* 1993; Parkinson *et al.* 1996; Lawson *et al.* 2004). CRP is allostericly activated by binding to a co-factor, cAMP, which is produced when glucose concentrations are low. When active, CRP binds to its binding site, exerting a 90-125° bend in DNA, flanking the binding site, altering the architecture of the promoter (Parkinson *et al.* 1996; Passner and Steitz 1997; Lin and Lee 2003; Popovych *et al.* 2009). Gene regulation, mediated by changes in promoter architecture, is common; were the position of additional transcription factors binding sites can be altered to allow, or disallow, the factor to contribute to transcription regulation. The position of the CRP DNA binding site in regards to a promoter determines the effect CRP has on the expression from a promoter; cAMP-CRP able to both induce and repress gene expression (Lawson *et al.* 2004). Transcription activation requires an interaction between RNAP the activating regions of CRP; activation region 1 (AR1) located in the C-terminal domain (αCTD) and activation region 2 (AR2) the N-terminal domain (Igarashi and Ishihama 1991; Zhou *et al.* 1994; Savery *et al.* 1996). Whether CRP induces or represses transcription depends on the position of the binding site; the optimum positions for activation, in respect to the transcription start site, being centred at position -41 for class II and positions -93, -83, -72, or -62 for class I (Tebbutt *et al.* 2002; Rossiter *et al.* 2011). CRP also represses transcription through promoter element occlusion, anti-activation or by hindering promoter clearance (Polayes *et al.* 1988; Mollegaard *et al.* 1993; Lee and Busby 2012).

1.4.7 Factor for Inversion Stimulation

The nucleoid is a distinct complex consisting of the genomes and several nucleoid associated proteins (NAPs) that maintain this structure and some of which affect gene (Browning *et al.* 2010). Fis is one such NAP, and is involved in a diverse set of cellular processes, including regulating transcription, maintaining nucleoid structure, site specific DNA inversion and modulating level of negative DNA topology (Johnson *et al.* 1986b; Ussery *et al.* 2001). Fis exists as a homodimer that binds to its 15 bp binding site characterised by partial dyad symmetry and AT rich, degenerate sequence (Grainger *et al.* 2006; Shao *et al.* 2008). This binding inducing DNA bending between 40° and 90° which leads to the stabilisation of the supercoils of the DNA helix, compacting the DNA (Travers and Muskhelishvili 1998; Skoko *et al.* 2006). There are at least 894 Fis DNA binding regions scattered

throughout the genome within both promoter and intergenic regions; transcription regulation modulation occurring from a subset of these (Schneider *et al.* 2001; Grainger *et al.* 2006; Cho *et al.* 2008; Browning *et al.* 2010). Fis is able to modulate transcription by directly affecting the activity of RNA polymerase or indirectly by modulating the level of negative DNA supercoiling (Travers and Muskhelishvili 1998; Skoko *et al.* 2006; Browning *et al.* 2010). Fis is able to affect transcription through stabilising negative supercoiling which captures RNAP which is aligned with the promoter element, this alignment generating positive writhe counteracting the negative supercoiling. Fis then causing the negative supercoil to be released, the stored torsion transmitted to RNAP; contributing to the isomerisation of the open complex to the Elongation Complex, and subsequent transcript elongation.

1.5 Enteroaggregative *Escherichia coli*

EAEC is a heterogeneous pathovar, characterised by the "stacked bricked" pattern of adhesion with which it adheres to HEp-2 cells; many strains of the pathovar have the ability to causes persistent watery diarrhoea (Nataro and Kaper, 1998; Nataro *et al*. 1998; Harrington, *et al*. 2006). When EAEC is subsequently mentioned in this work, it refers to the prototypical strain 042 (Chaudhuri *et al*. 2010). The pathogenesis of EAEC involves several distinct processes; adherence to intestinal mucosa, biofilm formation and toxin expression, facilitated by the coordinated expression of virulence factors (Wakimoto, *et al*. 2004; Harrington, *et al*. 2005; Nataro 2005.) The pAA plasmid encodes many of these virulence factors ; one of which is the prototypical SPATE (serine protease autotransporter of the *Enterobacteriaceae*), Pet (Eslava, *et al*. 1998; Henderson *et al*. 1999; Chaudhuri, *et al*. 2010). Pet contributes to the pathogenesis of EAEC strain 042 by targeting α-fodrin, disrupting the actin cytoskeleton of host cells (Eslava et al. 1998; Nataro 2005; Croxen and Finlay 2010). The action of Pet causes the homeostasis of the host gut to be disturbed allowing EAEC to outcompete the commensal flora, allowing EAEC to colonise the nutritional niche of the intestinal mucosa (Kamada *et al*. 2012).

1.5.1 Plasmid Encoded Toxin

The *pet* gene is encoded on the pAA plasmid of EAEC strain 042 and expressed during early exponential phase resulting in the secretion and accumulation of extracellular Pet (Betancourt-Sanchez and Navarro-Garcia 2009; Chaudhuri *et al.* 2010). Previous work has been done to assess aspects of transcription the *pet* promoter (Rossiter *et al.* 2011). A transposon mutagenesis library of EAEC was screened for strains deficient in Pet secretion; over 150 candidates were identified, several of which had insertion in around the genes encoding the global transcription factors CRP and Fis (Rossiter *et al.* 2011). An *in vitro* transcription assay of region containing the

predicted translation start site of *pet*, was assessed to determine the location of transcription start site. The downstream of the transcription start site several potential promoter elements, including CRP and Fis DNA binding sites, were identified. *In vitro* footprinting was used to define the location of the DNA binding sites. One CRP binding site was shown to be centred -40.5 and two Fis binding sites centred -91 (fisI) and -122 (fisII), in respect to the transcription binding site.

The contribution of each these sites on the expression from the *pet* promoter was assessed by the creation of several *pet* promoter derivatives contained single or double nucleotide substitution that knocked out each proposed promoter element individually. These promoter derivatives were cloned into a gene reporter system (pRW50) that allowed gene expression to be assessed by the observed functional activity of an expressed enzyme; β-galactosidase. The *pet* promoter was confirmed to possess several promoter elements; extended -10 element, CRP DNA binding site and two Fis DNA binding sites. Both the extended -10 element and CRP DNA binding site were seen to be essential for transcription; when knocked out greatly showed drastically reduced the levels of expression. The downstream Fis DNA binding site showed significantly reduced expression, but to a lesser extent. Suggesting transcription activation is mediated by CRP and enhanced by Fis, from the downstream Fis binding site, in a synergic manner. When the upstream Fis binding site was knocked out expression was seen to rise showing this site is repressive. An *in vitro* band shift assay was performed showing that the downstream Fis binding site has a higher affinity than that of the upstream site but both may be occupied at the same time.

This work then proceeds, further defining the effects of CRP in the context of the *pet* promoter. CRP dependent induction was assumed to be directed via a class II activation mechanism, from CRP DNA binding site. An assay conducted in a Δ*crp* genetic background completed with plasmids contain CRP with either activating regions (AR1 and AR2) activity impaired; both regions impairments showed significantly reduced expression. The optimum position for CRP DNA binding site for class II transcription activation is centred at location -41.5 in respects to the transcription start site (Gaston *et al.* 1990). 1 bp was inserted between position −21 and −22 of the *pet* promoter to alter the position of the CRP binding site to the optimal position. The expression of this promoter was compared to that of the wild type in Δ*fis*, Δ*crp* and wild type genetic backgrounds. The expression from this promoter was at equal levels in both Δ*fis* background and wild type genetic backgrounds, meaning promoter is no longer activated by Fis and the contribution of CRP to transcription activation is increased.

1.6 Aims

The previous work was able to identify the important promoter elements that regulate transcription at the *pet* promoter and provided some insights into the mechanisms of regulation. Transcription regulation at this promoter requires further investigation to clarify how CRP and Fis collaborate to modulate expression. Fis is seen to regulate from two DNA binding sites in the promoter but the mechanisms employed are unclear. Furthermore, the dependence on CRP for the activation due to Fis has not been addressed. The aims of this work are to firstly elucidate how Fis induces transcription from the downstream Fis binding site. Secondly, the mechanism of repression employed by Fis bound to the upstream site will be assessed. Finally, any suspected interactions between RNAP and transcription factors will be assessed to identify specific determinants to confirm the presence of class I or II interactions.

Chapter 2

Materials and Methods

2.1 Suppliers

Unless stated otherwise, all chemicals were ordered from Sigma-Aldrich or Fischer Scientific. Oligonucleotides were supplied by Alta Biosciences, University of Birmingham. Restriction endonucleases, Calf alkaline phosphatase and Phusion® High-Fidelity DNA Polymerase were supplied by New England Biolabs. T4 DNA ligase was supplied by Invitrogen. dNTP was supplied by Bioline. All enzymes were used as according to manufacturer's instructions and with supplied buffers.

2.2 Buffers, Solutions and Reagents

Solutions for use with bacterial or DNA manipulation were autoclaved prior to use for 20 minutes at 120^{O}C and 15 psi or filter sterilised using a Acrodisc 32 mm syringe filter with 0.2 μm supor® membrane (PALL).

2.2.1 Gel Electrophoresis of DNA

General

DNA loading dye: 5X DNA loading buffer (Bioline)

WebGreen DNA Stain: 2-5 μL per 100mL 0.8-3.0% agarose solution (Web Scientific)

DNA markers: 100 bp and hyperladder 1 (Bioline)

Agarose gel electrophoresis

50x TAE buffer: 2 M Tris acetate, 100 mM Na_2EDTA. Diluted to 1x for use as running buffer

Agarose solutions: 0.8-1.2% agarose in 1X TAE, 800 Watt in a microwave until dissolved

Polyacrylamide gel electrophoresis

5x TBE buffer: 0.335M Tris borate pH8.3, 10mM Na_2EDTA (National Diagnostics). Diluted to 1x for use as running buffer

Stock acrylamide: 30% (w/v) acrylamide, 0.8% bisacrylamide stock solution (Protégél - National Diagnostics)

7.5% acrylamide: 125 mL stock acrylamide solution, 100mL 5x TBE, 20ml glycerol, 255ml distilled water

2.2.2 Extraction and Purification of Nucleic Acids

Phenol/Chloroform: phenol/chloroform/isoamyl alcohol (25/24/1 v/v), pH 8.0 (Sigma)

TE buffer: 10 mM Tris pH 8.0, 0.5 mM EDTA

2.2.3 β-galactosidase Assays

Z-buffer: 0.75 g/L KCl, 0.25 g/L $MgSO_4.7H_20$, 8.53 g/L Na_2PO_4, 4.87 g/L $NaH_2PO_4.2H_2O$, 2.70 ml/L β-mercaptoethanol (β-mercaptoethanol added after autoclaves, immediately prior to use)

ONPG: o-nitrophenyl-β-D-galactopyranoside 0.8 g/L (dissolved in Z buffer)

Sodium Carbonate: 1 M Na_2CO_3

2.2.3 Preparing Competent Cell by Rubidium Chloride Method

TFB1: 12.1 g/L RbCl, 9.9 g/L $MnCl_2.4H_2O$, 2.9 g/L CH_3CO_2K, 1.1 g/L CaCl2, 1.5% (v/v) glycerol (Filter sterilise, keep at 4 degrees)

TFB2: 10 mM MOPS, 75 mM CaCl2, 10 mM RbCl2, 15 % (v/v) glycerol, pH to 6.5 (Filter sterilise, keep at 4 degrees)

2.3 Bacterial Growth Media

2.3.1 Liquid media

Liquid media is prepared by central services, The University of Birmingham. It is prepared with distilled water and autoclaved for 20 minutes at 120^{O}C and 15 psi.

LB: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl

2.3.2 Solid media

Solid media is prepared by central services, The University of Birmingham. It is prepared with distilled water and autoclaved for 20 minutes at 120^{O}C and 15 psi. When required for use, media was heated at 100^{O}C until melted and left to cool to 40-50^{O}C before addition of antibiotics (section 2.3.3). Media is poured into petri dishes aseptically, and store at 4^{O}C.

LB: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 10 g/L agar

Lactose MacConkey: 50 g/L MacConkey lactose agar (OXOID)

2.3.3 Antibiotics

Antibiotics were prepared as stated and then filter sterilised through an Acrodisc syringe filter.

Ampicillin: 50 mg/mL in sterile distilled water (stored -20^{O}C)

Tetracycline: 100 mg/mL in methanol (stored -20^{O}C)

Kanamycin: 50 mg/mL in sterile distilled water (stored -20^{O}C)

Antibiotics were added to media after autoclaving to select for plasmid/chromosomally encoded antibiotic resistance, at final concentrations; ampicillin 50 μg/mL, tetracycline 100 μg/mL, kanamycin 50 μg/mL.

2.4 Bacterial Strains

Bacterial strains used within this study are listed in Table 2.1. All strains used are stored at -80^{O}C as glycerol stocks. Before use these were re-streaked on LB agar and grown at 37^{O}C overnight. Overnight cultures were grown from a single colony, in 5-10 mL of LB and antibiotic supplemented were appropriate, and aerated at 37^{O}C for 14-16 hrs. Monitoring the levels of growth of cultures were accessed by measuring the OD at 600 nm using a spectrophotometer (Ultraspec 2100 pro).

2.5 Plasmids

Plasmids used within this study are listed in Table 2.2. Plasmid maps are shown Figure 2.1-2.2.

2.6 Gel Electrophoresis of DNA

2.6.1 Agarose Gel Electrophoresis of DNA

Agarose gels were used to analysis DNA fragments >500 bp. Molten agarose solution (0.8-1.2%) was cooled to around 60^{O}C, at which point WebGreen, which associates with DNA and fluoresces allowing visualiation of DNA, was added and the solution poured into a gel casting plate, up to final concentration of 0.01x. DNA fragments were mixed with 5:1 DNA loading dye and loaded into gels. Samples were run in 1X TAE running buffer at 105 V for 45 minutes against a DNA ladder.

2.6.2 Polyacrylamide Gel Electrophoresis of DNA

Polyacrylamide gels were used to analyse DNA fragments smaller than 1 Kbp. 7.5% polyacrylamide solution was polymerised by the addition of 0.001 volumes of TEMED and 0.01 volumes of 10% (w/v) ammonium persulphate, and immediately pipetted into a gel casting plate. DNA fragments were mixed with 0.2 volumes of DNA loading dye and loaded into gels. Samples were run in 1X TBE running buffer at 125 V for 90 minutes

Table 2.1: *Escherichia coli* K12 strains used in this study

Strain	Description	Source
RLG221	*rec*A56 *ara*D139 Δ(*ara-leu*)7697 *lac*X74 *gal*U *gal*K *hsd*R *str*A	R.Gourse University of Wisconsin-Madison
BW25113	*lac*I^q *rrn*B_{T14} Δ*lac*Z_{WJ16} *hsd*R514 Δ*ara*BAD_{AH33} Δ*rha*BAD_{LD78}	(Baba *et al.* 2006)
BW25113 *crp*::*aph*	Δ*rha*BAD_{LD78} *crp::aph*	(Baba *et al*. 2006)
BW25113 *fis*::*aph*	Δ*rha*BAD_{LD78} *fis::aph*	(Baba *et al*. 2006)

Table 2.2: Plasmids used in this study

Strain	Description	Source
pRW50	Low copy number, broad host range, *lac* expression vector. EcoRI-HindIII flanked promoter fragments are inserted into a region upstream of *lac*ZYA, *ori*V (Tet^R) (Figure 2.1)	(Lodge *et al.* 1992)
pSR	High copy number, pBR322 derivative. EcoRI –HindIII fragments upstream of *λoop* terminator (Amp^R) (Figure 2.2)	(Kolb *et al.* 1995)
pREIIα (and derivatives)	Plasmid containing *rpo*A, expressing RNAP α subunit (derivatives containing alanine substitutions at positions 273-329)	(Blatter *et al.* 1994; Gaal *et al.* 1996; Wood *et al.* 1997)
pHTfIα (and derivatives)	Plasmid containing *rpo*A, expressing RNAP α subunit (derivatives containing alanine substitutions at positions 255-271)	(Tang *et al.* 1994; Gaal *et al.* 1996)
pKK223-3 (and derivatives)	Plasmid is an expression vector that encodes and expresses Fis from a non-native promoter (derivatives containing nucleotide substitution of *fis*, causing amino acid substitutions at positions 71 or 72)	(McLeod *et al.* 1999)

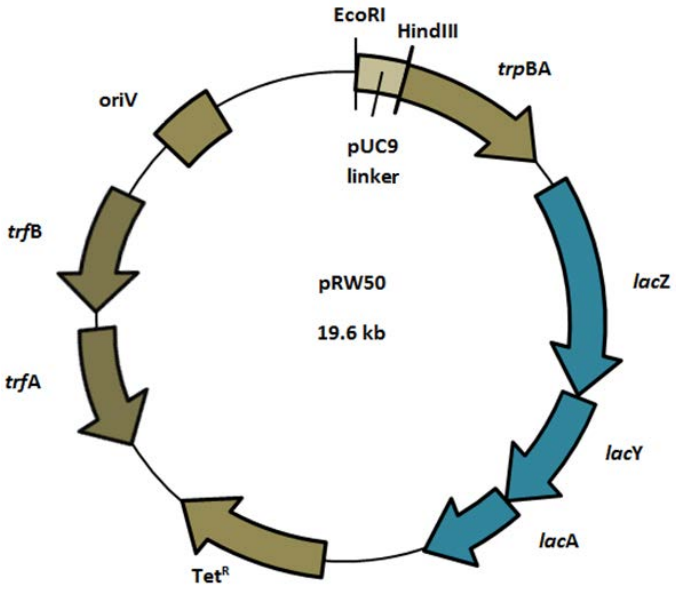


Figure 2.1: pRW50

EcoRI-HindIII fragments are cloned upstream of *trp*BA replacing pUC9 linker, placing the expression of the *lac* operon under the control of the promoter on the EcoRI-HindIII fragment. pRW50 also contains the tetracycline resistance gene (Tet^R), origin of replication, genes responsible for plasmid replication (*trf*A and *trf*B) and *trp*BA which is part of the lac operon.

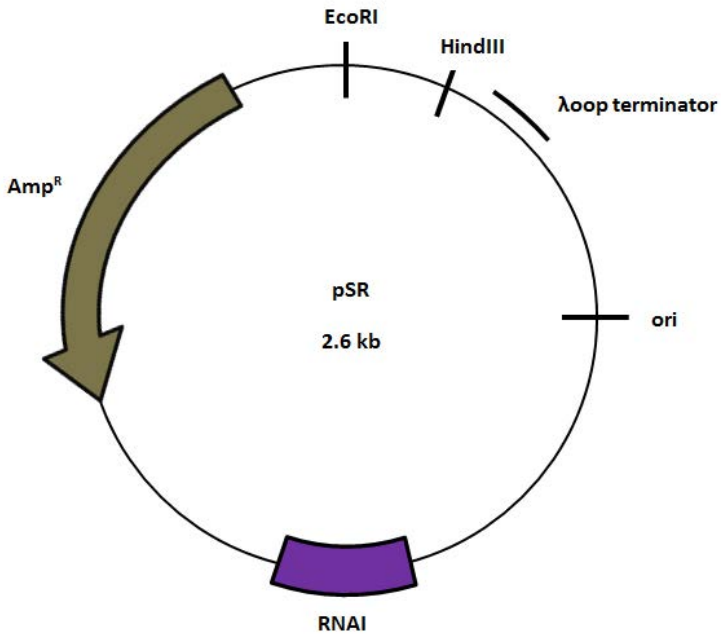


Figure 2.2: pSR

EcoRI-HindIII fragments are cloned upstream of λ*oop* terminator. pSR allow the measurement of *in vitro* transcription from the EcoRI-HindIII fragment; transcription terminating at the λ*oop* terminator, producing a transcript of a defined length. RNA1 transcribes a control transcript. pSR also contains the *bla* gene providing ampicillin resistance and the origin of replication (ori).

against a DNA ladder and stained through 5-10 minutes incubation with 0.01x WebGreen solution. Fragments were visualised used a UV transilluminator.

2.7 Extraction and Purification of Nucleic Acids

2.7.1 Phenol/Chloroform Extraction of DNA

To remove contaminating protein, DNA solutions were mixed 1:1 with Phenol/chlololform, vortexed and centrifuged for 2 minutes at 13,000 rpm. The top aqueous phase was transferred into a fresh microfuge tube and concentrated by ethanol precipitation (section 2.7.2).

2.7.2 Ethanol Precipitation of DNA

Ethanol precipitation was used to concentrate DNA and remove previous buffers. 1/10 volumes of 3 M sodium acetate pH 5.2, and 2 volumes of ice-cold 100% ethanol were added to DNA solution. The sample was then incubated overnight (14-16 hours) at -80^{O}C. After this period the sample was centrifuged at 14000 rpm and 4^{O}C for 15 minutes. The supernatant was decanted and 1 mL ice-cold 70% ethanol was added and centrifuged as before. The supernatant again was decanted and 1 mL ice-cold 100% ethanol was added and centrifuged as before. The supernatant was decanted and the DNA pellet dried under vacuum for 10 minutes. The pellet was re-suspended in sterile distilled water.

2.7.3 Purification of DNA using QIAquick PCR Purification Kit

Purification of DNA fragment post PCR or restriction digest was performed using the QIAquick PCR Purification Kit (QIAGEN) following the manufacturer's instructions. The DNA is applied to a silica membrane to which the fragment, >100 bp, bind. The membrane is washed with two rounds of PE buffers, which are passed across the membrane by centrifugation, removing primers, nucleotides, enzymes, salts and other impurities from DNA samples. The DNA fragments were eluted by with 50 µL of low-salt elution buffer EB, which dissociates the fragments from the membrane (QIAGEN).

2.7.4 Electroelution of DNA from Polyacrylamide Gels

After Polyacrylamide Gel Electrophoresis of DNA fragments (section 2.7.1), the required DNA bands were excised from the gel and placed into 6.0 mm dialysis tubing with 200 µL 0.1X TBE and sealed with clips. The tubing was placed in an electroelution tank and run at 30 mA for 30 minutes. The DNA solution from the tubing was transferred to a microfuge tube. 200 µL sterilised distilled water was used to wash the interior of the dialysis

tubing before being combined with the DNA solution. The DNA solution was then purified using Phenol/ Chloroform extraction (section 2.7.1) and ethanol precipitation (section 2.7.2).

2.7.5 Small-Scale Preparation of Plasmid DNA ("mini-prep")

A culture containing the plasmid to be purified was grown overnight in LB, supplemented with any required antibiotic. 10 mL of culture (low copy plasmids) or 5 mL culture (high copy plasmid) was used to extract plasmid DNA, performed using the QIAprep Spin Miniprep Kit (QIAGEN) following the manufacturer's instructions. The culture is lysed and the lysate is separated from the suspended cell debris by centrifugation. The clear lysate is then loaded on the silica membrane of the spin kit, which binds the DNA. To avoid extraneous endonuclease activity acting on the plasmid DNA 0.5 mL Buffer PB was applied across the membrane filtration assisted by centrifugation. Other Impurities such as salt, enzymes and chromosomal DNA are washed away by two rounds of PE buffer which are run across the membrane. DNA was eluted in a small volume of elution buffer, which releases the DNA by altering the ionic conditions within the silica membrane. DNA was eluted with 50 µL elution buffer EB (QIAGEN).

2.7.6 Large-Scale Preparation of Plasmid DNA ("maxi-prep")

A culture containing the plasmid to be purified was grown overnight in LB, supplemented with any required antibiotic. 500 mL culture was used to extract plasmid DNA, performed using the QIAprep Spin Maxiprep Kit (QIAGEN) following the manufacturer's instructions. The principles of DNA purification are the same as those describes in section 2.7.5 but on a larger scale. Plasmid DNA was eluted with 2.0 mL TE (QIAGEN).

2.8 Bacterial Transformations

2.8.1 Preparation of Competent Cells Using Rubidium Chloride Method

50 mL LB was inoculated with 0.5 mL of an overnight culture of the strain to be transformed and grown at 37^{O}C under aeration until OD_{650} 0.4-0.6 (mid-exponential phase). The culture was then transferred to a sterile 50 mL centrifuge tube incubated in ice for 5-10 minutes, and then centrifuged at 4,000 rpm and 4^{O}C for 15 minutes. The supernatant is discarded and the pellet is resuspended in TFB1 buffer and incubated in ice for 90 minutes. The solution was then centrifuged at 4,000 rpm and 4^{O}C for 15 minutes, the supernatant discarded and the pellet resuspended in 1 mL of TFB2 buffer, this solution was transferred to microfuge tubes as 200 µL aliquots and stored at -80^{O}C.

2.8.2 Transformation of Competent Cells with Plasmid DNA

50 µL of competent cells were mixed with 5 µL of plasmid DNA and incubated in ice for 30 minutes. The cells were then shocked at 42^{O}C for 45 seconds. 500 mL of LB was added to the cells, which were allowed to recover at 37^{O}C with aeration for 60 minutes. The cells were centrifuged at 4,000 rpm for 15 minutes, and the pellets were resuspended in minimal supernatant (around 80 µL) and plated on to agar supplemented with required antibiotics, and grown overnight at 37^{O}C.

2.9 DNA Manipulations

2.9.1 Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was done using Phusion® High-Fidelity DNA Polymerase (New England Biolabs) and supplied HF buffer. The 4 µL of template (plasmid/PCR product) was added directly from the 50 µL elute of QIAprep Spin Miniprep Kit (QIAGEN) or QIAquick PCR Purification Kit (QIAGEN). dNTP (Bioline) was added, with a final concentration of 1 mM (0.25 mM of each). Oligonucleotides primers (Alta Biosciences) (Table 2.3) were added at a final concentration of 0.75 mM each. The typical PCR cycle used is shown in Table 2.4.

2.9.2 Restriction Digestion of DNA

38 µL DNA solution from the 50 µL elute of QIAprep Spin Miniprep Kit (QIAGEN) or QIAquick PCR Purification Kit (QIAGEN) was double digested with 3 µL of each enzyme and 10 µL of appropriate buffer, determined using New England Biolabs online Tm calculator (http://www.neb.com/nebecomm/DoubleDigest Calculator.asp?). This solution was incubated at 37^{O}C for 3 hours. Plasmid DNA that was to be used as a vector in cloning was then treated with 3 µL calf alkaline phosphatase (New England Biolabs) and incubated at 37^{O}C for 3 hours to remove terminal 5' phosphate groups. The digest was purified by Phenol/Chloroform extraction (section 2.7.1) or by using QIAquick PCR Purification Kit (QIAGEN) (section 2.7.3).

2.9.3 DNA Ligations

10 µL of restriction digested insertion DNA, 5 µL restriction digested, dephosphorylated vector DNA, 1 µL T4 DNA ligase and 2 µL T4 DNA ligase buffer were mixed and incubated at room temperature for 3 hours. This ligate mixture was added to 50 µL RLG221 competent cells as described (section 2.8.2). When transforming into pRW50 vector, the cells were plated on MacConkey agar, allowing red/white screening of overnight transformants for candidates. Red colonies were selected and grown overnight to extract plasmids using QIAprep

Spin Miniprep Kit (QIAGEN) (section 2.7.5). Plasmids were screened by PCR, using primers flanking the insertion sites on the plasmid. Positive candidates were sequenced using the appropriate primer (section 2.9.4).

2.9.4 DNA Sequencing

Plasmid-to-profile sequencing was performed by the Function Genomics and Proteomics Laboratory, University of Birmingham. 5 µL of plasmid was mixed with 3 µL 1 mM sequencing primer up to 10 µL in sterile distilled water. The sequencing primers are listed in Table 2.3.

2.9.5 Introduction of Point Mutations Using Megaprimer PCR

In the first round of PCR the template EcoRI-HindIII fragment was present in pSR vector. The point mutation is inserted through a mutagenic primer and flanking primer, Psr_R (Table 2.3). The fragment was purified by QIAquick PCR purification kit was eluted in 30 µL of EB buffer. The second round of PCR was performed with the previous template but this time with 10 µL of the purified megaprimer and the alternate flanking primer D5431. This produced the full length EcoRI-HindIII fragment with the required mutation. This fragment was purified by QIAquick PCR purification kit, double digested by EcoRI and HindIII, and cloned into pRW50 and/or pSR (section 2.10). Mutagenic primers are listed in Table 2.3 and resultant EcoRI-HindIII fragment listed in Table 2.5.

2.10 Cloning of Promoter Fragments into pSR and pRW50

2.10.1 Preparation of EcoRI-HindIII Promoter Fragments by PCR

All the promoter fragments used in this study are listed in Table 2.5 and their sequences shown Figures 2.4 -2.22. *aer1* and derivatives where created from pSR *aer1* by megaprimer PCR (Section 2.9.5), using primers D5431 and Psr_R and the relevant mutagenic primer (Table 2.3). FnCC(-40.5) and derivatives (F_{45}CC(-40.5), F_{40}CC(-40.5), F_{35}CC(-40.5), F_{30}CC(-40.5), F_{25}CC(-40.5), F_{15}CC(-40.5)) were created using megaprimer PCR from pSR F_{35}CC(-40.5) using primers D5431 and Psr_R and the relevant mutagenic primer (Table 2.3). F_{24}CC(-40.5) and F_{26}CC(-40.5) were created from pSR F_{25}pCC(-40.5) D5431 and Psr_R and the relevant mutagenic primer (Table 2.3). $Fc_{66}F_{25}$CC(-40.5) was created from pRW50 F_{35}CC(-40.5) using PCR and the primers F2CCcons and D10527.

Table 2.3: Oligonucleotides Primers Used In This Study

	Name	Sequence (5'→3')	Use
Sequencing	D10520	CCCTGCGGTGCCCCTCAAG	Anneals upstream of EcoRI in pRW50. Sequencing and amplification of insert.
	D10527	GCAGGTCGTTGAACTGAGCCT GAAATTCAGG	Anneals downstream of HindIII in pRW50. Sequencing and amplification of insert.
	D5431	ACCTGACGTCTAAGAAACC	Anneals upstream of EcoRI in pSR. Sequencing and amplification of insert.
	Psr_R	GACGACGACATGGCTCGATTG	Anneals downstream of HindIII in pSR. Sequencing and amplification of insert.
Mutagenic	F_{15}CC(-40.5)	GGAGAGCTCATGTGATGTACA TCACATGG	Deletes 20 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{24}CC(-40.5)	GGAGAGCTCATCAGGTAAAT G	Deletes 11 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{25}CC(-40.5)	GGAGAGCTCGATCAGGTAAA TG	Deletes 10 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{26}CC(-40.5)	GGAGAGCTCCGATCAGGTAA ATG	Deletes 11 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{30}CC(-40.5)	GGAGAGCTCCCGGGGATCAG GTA	Deletes 5 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{35}CC(-40.5)	GGAGAATTCAAAAATGATTA ATTTATGATCAATCATTGCTC GGTACCCGGGG	Inserts fisI binding site of aer1 upstream of CC(-40.5) centred at position -91 in regards to transcription start site

	F_{39}CC(-40.5)	GGAGAGCTCCATGGGTACCCGGGGATC	Inserts 4 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{40}CC(-40.5)	GGAGAGCTCCATGGGGTACCCGGGGATC	Inserts 5 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{41}CC(-40.5)	GGAGAGCTCCATGGAGGTACCCGGGGATC	Inserts 6 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F_{45}CC(-40.5)	GGAGAGCTCCATGGATATCGGTACCCGGGGATC	Inserts 10 nt between fisI site and CRP site of F_{35}pCC(-40.5)
	F2CCcons	GGAGAATTCATAATGGATCAAATTTTGAGCAAAAAAAAACAAAAATGATTAATTTATGATC	Inserts 37 nt of *aer1* sequence upstream of F35pCC(-40.5), includes fisII site which is altered to the consensus Fis binding site sequence
	PetFISIIC	GGAACCTGATAATGCTCCTAAATAGGAGCAAAAAAAAC	2 nt substitution in fisII site of *aer1* which is unable to recruit Fis
	aer1_5	CAATCAGCATTGGTTGACATCGTATCA	5 nt deletion between -10 element and crp site, in *aer1*
	aer1_10	CAATCAGCATTGGTTCGTATCATTAAC	10 nt deletion between -10 element and crp site, in *aer1*
	Pet_R	GGGAAGCTTGTATTTTATTCATATATTCTCTCAACTCATTTATTG	Anneals downstream of *pet* promoter and inserts a HindIII site

Table 2.4: PCR Cycle

Cycle step	Cycles	Temp	Time
Initial Denaturation	1	98°C	30 seconds
Denaturation		98°C	5–10 seconds
Annealing	30	45–72°C	10–30 seconds
Extension		72°C	15–30 seconds per kb
Final extension	1	72°C	5–10 minutes
Hold	1	4°C	∞

The annealing temperature required for the PCR reaction was determined using New England Biolabs online Tm calculator (http://www.neb.com/nebecomm/tech_reference/TmCalc/ Default.asp?). The extension time was for the expected PCR product was determined by the manufacturer's instructions (15–30 seconds per kb).

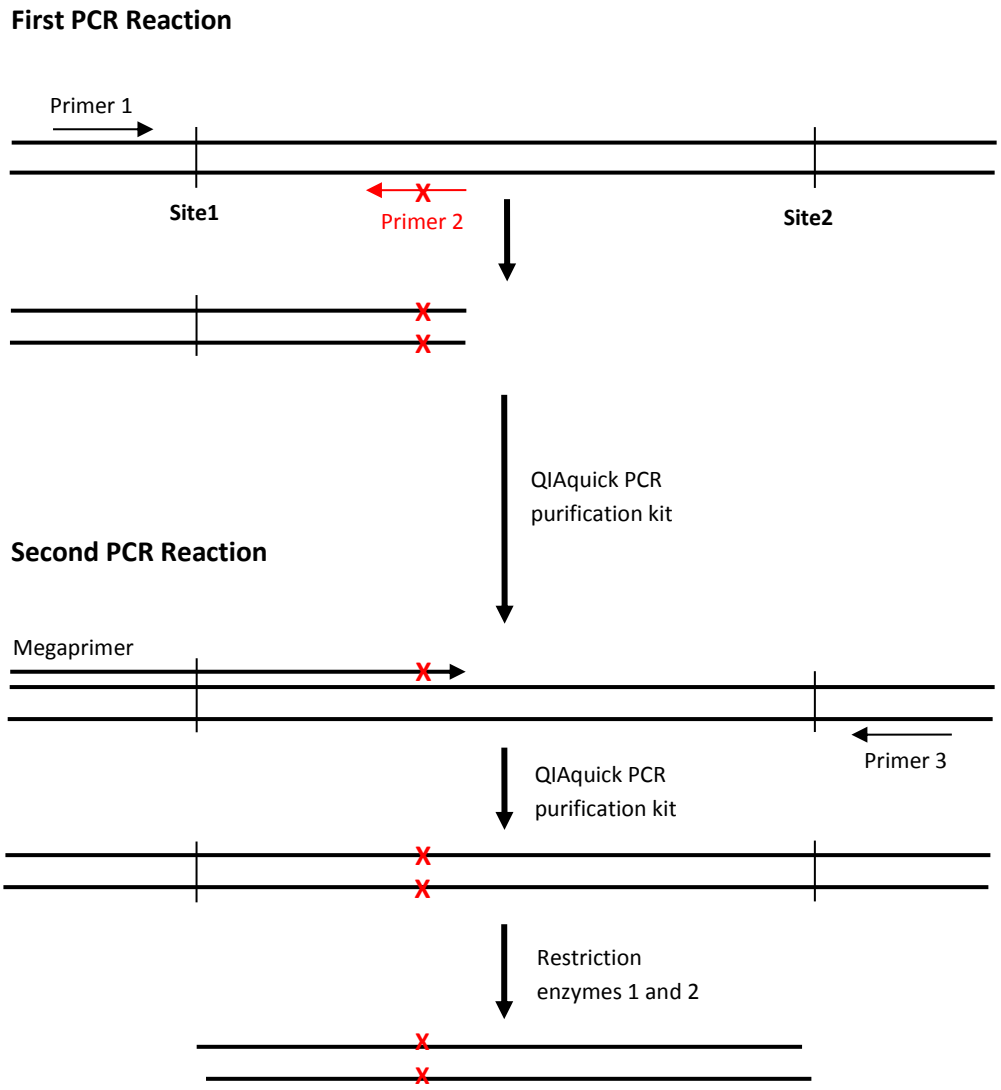


Figure 2.3: Megaprimer PCR

The megaprimer is generated in the First PCR reaction using a primer flanking the fragment (primer 1) and a mutagenic primer (primer 2). The PCR product is used in the second reaction to generate the full mutagenic fragment using the megaprimer and a primer that flanks the opposite side of the fragment (primer 3). This is then digested with restrictive enzymes and ligated in to an appropriate vector.

2.10.2 Construction of aer1 and FnCC(-40.5) derivatives

Derivatives of *aer1* and FnCC(-40.5) were prepared from pSR *aer1* and pSR FnCC(-40.5). pSR was extracted by QIAspin miniprep kit, double digested with EcoRI and HindIII and treated with calf alkaline phosphatase, and purified by phenol/chloroform purification followed by ethanol. EcoRI-HindIII fragments (section 2.10.1) (Table 2.5) were inserted into pSR by ligation for 3 hours at room temperature, and this ligate was transformed in rubidium chloride treated RLG221 competent cells. These cells were selected for by overnight growth on an Ampicillin supplemented LB agar plate. Candidates were verified by sequencing using primers D5431 and/or Psr_R. Successful candidates were grown overnight and plasmids extracted by QIAspin miniprep kit, double digested with EcoRI and HindIII and treated with calf alkaline phosphatase, and purified by phenol/Chloroform purification followed by ethanol. pRW50 was extracted by QIAspin miniprep kit, double digested with EcoRI and HindIII and treated with calf alkaline phosphatase, and purified by phenol/Chloroform purification followed by ethanol precipitation. EcoRI-HindIII fragments were inserted into pRW50 by ligation for 3 hours at room temperature, and this ligate was transformed in Rubidium Chloride treated RLG221 competent cells. These cells were selected for by overnight growth on a tetracycline supplemented MacConkey agar plate. Candidates were selected by red coloured colonies and verified by sequencing using primers D10520 and/or D10527.

2.12 β-galactosidase Assays

β-galactosidase assays were used to measure the activities of a plasmid encoded promoter:*lacZ* fusions in several genetic backgrounds (Miller, 1972). 5 mL of LB was supplemented with appropriate antibiotics and inoculated with a freshly transformed colony, containing pRW50 with the promoter of interest, and incubated overnight at 37°C. The following day, 5 mL LB was inoculated with the overnight culture at a 1:100 dilution for measurements of β-galactosidase activity during aerobic growth. Once cells had reached the required OD_{650}=0.5, they were immediately placed on ice and the exact OD_{600} reading was recorded. To lyse cells, 2 mL samples were added to 30 µL toluene and 30 µL 0.1% sodium deoxycholate and incubated for 30 min at 37°C, with aeration. For measurements of β-galactosidase activity, 100 µLof cell lysate was added to 2 ml Z buffer (0.75 g/L KCl, 0.25 g/L $MgSO_4.7H_2O$, 8.53 g/L Na_2HPO_4, 4.87 g/L $NaH_2PO_4.2H_20$); β-mercaptoethanol was added to Z buffer prior to use, at 0.27 mL per 100 mL Z buffer. The assay was initiated by the addition of 0.5 ml o-nitrophenyl-β-D-galactopyranoside (ONPG) solution, dissolved in Z buffer (3.92 g/L). Once the solution had changed from a clear colour to a straw yellow colour or 1 h had elapsed, 1 mL 1 M sodium carbonate was added to stop the reaction and the reaction time recorded. The absorbance at 420 nm of the resultant solution was measured and the calculation of β-galactosidase activity was calculated as follows:

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

Where:

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

3.6 = final assay volume (mL)

$^{1000}/_{4}$ = factor for conversion of OD_{420} into nmol o-nitrophenyl (ONP), based on 1nmol mL^{-1} ONP having an OD_{420} of 0.0045

t = incubation time (min)

v = volume of lysate added (in mL)

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

Table 2.5: EcoRI-HindII fragments used in this study

Name	Description	Source
aer1	-275 to +89 in respects to the transcription start site of the Pet promoter (EAEC strain 042)	(Rossiter *et al.* 2011)*
aer1-fisI	Derivative of *aer1* with G at position -84 and C at position -98	(Rossiter *et al.* 2011)*
aer1-fisII	Derivative of *aer1* with G at position -119 and C at position -126	(Rossiter *et al.* 2011)*
aer1-fisI/II	Derivative of *aer1* with G at positions -84 and -119 and C at positions -98 and -126	This study
aer1 Δ5	*aer1* with 5 bp deleted between -10 element and crp site	This study
aer1 Δ10	*aer1* with 10 bp deleted between -10 element and crp site	This study
CC(-41.5)	*E. coli melR* promoter derivative with consensus DNA site for CRP centred at –41.5	(Gaston *et al*. 1990)*
CC(-40.5)	CC(-41.5) with 1 bp deleted between -10 element and crp site	(West *et al*. 1993)*
F_{15}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -71	This study
F_{24}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -80	This study
F_{25}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -81	(Rossiter *et al.* 2011)*
F_{26}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -82	This study
F_{30}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -86	(Rossiter *et al.* 2011)*
F_{35}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -91	(Rossiter *et al.* 2011)*
F_{39}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -95	This study
F_{40}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -96	(Rossiter *et al.* 2011)*
F_{41}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -97	This study
F_{45}CC(-40.5)	Derivative of CC(-40.5) with fisI of *aer1* centred at position -101	(Rossiter et al. 2011)*

* Promoters were kindly provided by Henderson and Busby labs

Figure 2.4: aer1

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCGTAAATATGAGCAAAAAAAA
ACAAAAATGATTAATTTATGATCAATCAGCATTGGTTTACATGACATCGTATCATTAACGAGAGCATTGTCACAC
ATTAACAATATAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCC
TTTGTTCATTCAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.5: aer1 -fisI

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCGTAAATATGAGCAAAAAAAA
ACAAAAATCATTAATTTATGATGAATCAGCATTGGTTTACATGACATCGTATCATTAACGAGAGCATTGTCACAC
ATTAACAATATAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCC
TTTGTTCATTCAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.6: aer1 -fisII

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCCTAAATAGGAGCAAAAAAAA
ACAAAAATGATTAATTTATGATCAATCAGCATTGGTTTACATGACATCGTATCATTAACGAGAGCATTGTCACAC
ATTAACAATATAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCC
TTTGTTCATTCAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.7: aer1 –fisI/II

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCCTAAATAGGAGCAAAAAAAA
ACAAAAATCATTAATTTATGATGAATCAGCATTGGTTTACATGACATCGTATCATTAACGAGAGCATTGTCACAC
ATTAACAATATAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCC
TTTGTTCATTCAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.8: aer1 Δ5

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCGTAAATATGAGCAAAAAAAA
ACAAAAATGATTAATTTATGATCAATCAGCATTGGTTTACATCGTATCATTAACGAGAGCATTGTCACACATTAA
CAATATAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCCTTTGT
TCATTCAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.9: aer1 Δ10

GAATTCTTCCATGTGACATCTCAGATGGAAGCTCAGATGGAACCTGATAATGCTCGTAAATATGAGCAAAAAAAA

ACAAAAATGATTAATTTATGATCAATCAGCATTGGTTCGTATCATTAACGAGAGCATTGTCACACATTAACAATA

TAGAACTGTTACTTTTTACGGGATATTAGTAACAAACCATTACTAATGGTTTTAATTCTTAATCCTTTGTTCATT

CAATAAATGAGTTGAGAGAATATATGAATAAAATATACAAGCTT

Figure 2.10: CC(-41.5)

GAATTCGAGCTCGGTACCCGGGGATCAGGTAAAGTGATGTACATCACATGGATCCCCCTCACTCCTGCCATAATT

CTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.11: CC(-40.5)

GAATTCGAGCTCGGTACCCGGGGATCAGGTAAAGTGATGTACATCACATGGATCCCCTCACTCCTGCCATAATTC

TGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.12: F_{15}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCATGTGATGTACATCACATGGATCCCCTCACTCCT

GCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.13: F_{24}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCATCAGGTAAATGTGATGTACATCACATGGATCCC

CTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.14: F_{25}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCGATCAGGTAAATGTGATGTACATCACATGGATCC

CCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.15: F_{26}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCGATCAGGTAAATGTGATGTACATCACATGGATC

CCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.16: F_{30}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCCGGGGATCAGGTAAATGTGATGTACATCACATG

GATCCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Figure 2.17: F_{35}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCGGTACCCGGGGATCAGGTAAATGTGATGTACATC
ACATGGATCCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCT
T

Figure 2.18: F_{39}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCATGGGTACCCGGGGATCAGGTAAATGTGATGTA
CATCACATGGATCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAA
GCTT

Figure 2.19: F_{40}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCATGGGGTACCCGGGGATCAGGTAAATGTGATGT
ACATCACATGGATCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAA
AGCTT

Figure 2.20: F_{41}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCATGGAGGTACCCGGGGATCAGGTAAATGTGATG
TACATCACATGGATCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACAGATA
AAGCTT

Figure 2.21: F_{45}CC(-40.5)

GAATTCAAAAATGATTAATTTATGATCAATCATTCGAGCTCCATGGATATCGGTACCCGGGGATCAGGTAAATGT
GATGTACATCACATGGATCCCCTCACTCCTGCCATAATTCTGATATTCCAGGAAAGAGAGCCATCCATGAATACA
GATAAAGCTT

Figure 2.22: $Fc_{66}F_{35}$CC(-40.5)

GAACCTGATAATGCTCGTAAATATGAGCAAAAAAAAACAAAAATGATTAATTTATGATCAATCATTCGAGCTCC
ATGGATATCGGTACCCGGGGATCAGGTAAATGTGATGTACATCACATGGATCCCCTCACTCCTGCCATAATTCT
GATATTCCAGGAAAGAGAGCCATCCATGAATACAGATAAAGCTT

Chapter 3

Regulatory Region of the Pet Operon

3.1 Introduction

The *pet* gene is present on the pAA plasmid of EAEC strain 042 and encodes a serine protease autotransporter, Pet. As determine by prior work by Rossiter *et a.l* 2011, the transcription activation of *pet* is CRP dependant; where CRP binds and interacts with RNAP from its site centred at -40.5, with respect to transcription start site. Two Fis binding sites reside upstream of the CRP binding site; CRP dependant activation is enhanced by the presence of the downstream Fis DNA binding site (fisI; centred -91) but the presence of the upstream Fis DNA binding site (fisII; centred -122) appears to be repressive. Fis is shown to bind the fisI and fisII sites *in vitro,* but the mechanism by which enhancement and repression occurs is still undefined. The following work has the aim of firstly, confirming the face of helix dependence of the fisI site within the *pet* promoter and secondly, examining the mechanism of repression occurring due to the presence of the fisII site. The following experiments utilised *aer1* and related derivatives, which contained the regulatory region of the *pet* gene, to further elucidate the actions of Fis at the *pet* promoter.

3.2 Spatial Constrains of Fis Activation

The Fis site is known to activate transcription when bound to the fisI binding site, through an undetermined mechanism (Rossiter et al. 2011). A number of derivatives of the *pet* promoter were created to assess the spatial constraints on the synergistic activation by Fis and CRP, providing insights regarding the mechanisms involved. These constructs where constructed from *aer1* (Sequence -275 to +89; in respects to the transcription start site of the *pet* promoter of EAEC strain 042). Either 5 nt (*aer1 Δ5*) or 10 nt (*aer1 Δ10*) were deleted from within the region between the fisI and crp sites altering the orientation of the fisI site on the DNA helix; to either the opposite side of the helix (*aer1 Δ5*) or the original side of the helix, albeit closer to the crp site (*aer1 Δ10*). When the activities of these promoters were assessed the level of expression from the *aer1* and *aer1 Δ10* were the same, whilst the levels of *aer1 Δ5* were significantly reduced (Figure 3.2). These results show the enhancement of the fisI is face-of-helix dependant, with activation only occurring when the fisI site was position on the same face of the DNA helix as the crp site.

3.3 Contributions of Fis DNA Binding Sites to Expression

When Fis binds the *pet* promoter it either activates transcription, bound to the fisI binding site, or represses transcription bound to the fisII. To assess the contribution of each of the Fis binding sites upon the expression from the *pet* promoter, several derivatives containing nucleotide substitutions were created (Rossiter *et al.* 2011).

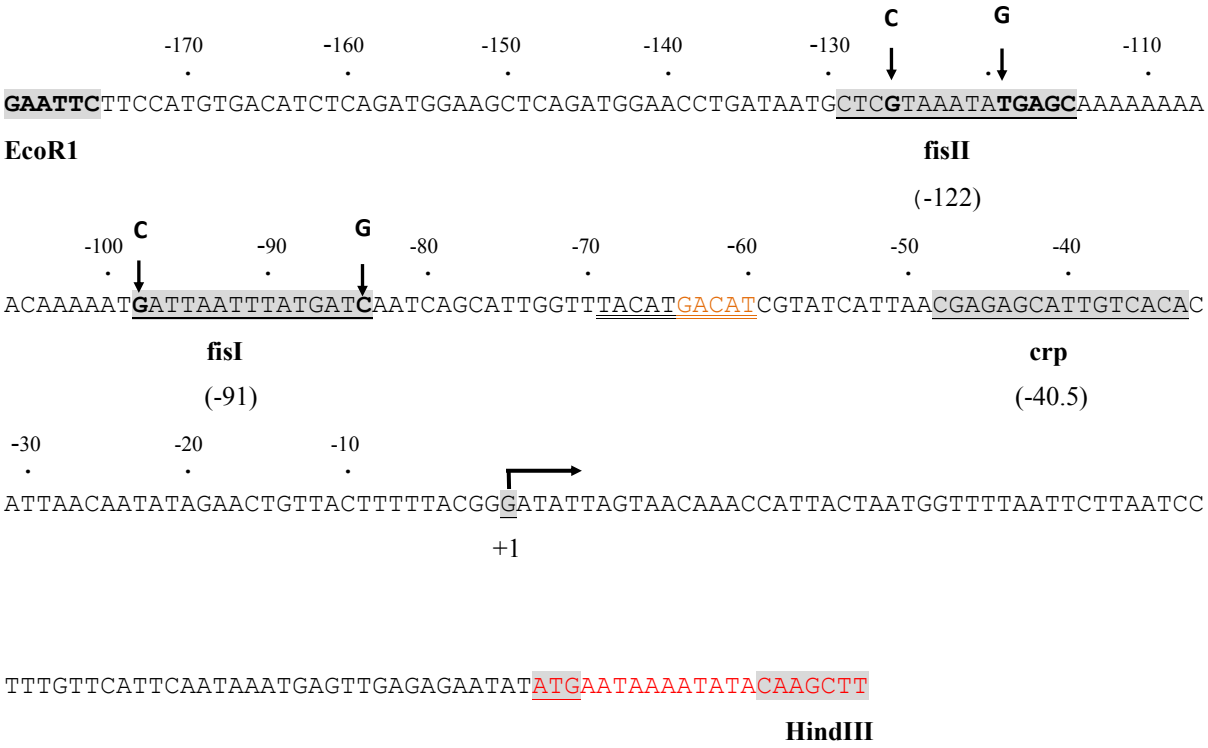


Figure 3.1: Point Mutations and Deletions Within the aer1 Fragment

The sequence of the *aer1,* which contains the *pet* regulatory region and downstream sequence, including the translational start site AUG (underlined). The transcription factor binding regions are underlined and labelled with the sites name and positions in regards to the transcription start side (+1). The vertical arrows identify the bases which were substituted for the **BOLD** base; 2 for each site which inactive the site by reducing the homology of the site to the consensus transcription factor binding sequence. The double underlined text The substitution of nucleotides within promoter elements can reduce or increase the homology of a promoter element, compared to the highest affinity sequence referred to as the consensus sequence, reducing or increasing identifies the nested deletions; the orange text signifying the 5 nt deletion (*aer1 Δ5*) and the entire underlined sequence for the 10 nt deletion (*aer1 Δ10*).

Derivatives of *aer1* were created were the Fis binding sites were altered by substituting two nucleotide bases, critical for factor binding, reducing the ability of the sites to recruit and bind Fis (shown in Figure 3.1). These derivatives had their fisI site disrupted (*aer1 -fisI*), fisII site disrupted (*aer1 –fisII*) or both fisI and fisII sites disrupted (Rossiter *et al.* 2011). First the activity observed when the fisII was disrupted showed a slightly raised activity when compared to aer1 (Figure 3.3). When the fisI site is disrupted the levels of expression are greatly diminished, although when the fisII is also disrupted this reduced activity is not seen to change (Figure 3.3). Suggesting the repressive action of the fisII site is dependant the presence of a functional on fisI site.

3.4 Discussion

Fis plays an important role in the positive and negative regulation of transcription at the *pet* promoter. The results obtained during this and the past study demonstrate Fis can only enhancement activation, when the Fis binding site is present on the on the same face of the DNA helix as the CRP binding site (Rossiter *et al*. 2011). The face of helix dependence on transcriptional activity of Fis suggests a direct interaction with RNAP may be present; activating transcription via a class I mechanism. Fis cannot enhance expression alone and requires the presence of CRP to function, due to the distance of the fisI site from the RNAP binding region. A set of Fis regulated promoters were assessed in regards to the locations of the DNA binding sites and the mechanism of regulation (Figure 3.4). Promoters that are only regulated Fis, in a class I manner, do so from sites positioned between a specific (Orange box Figure 3.4); -58 and -76 in regards to the transcription start site. Furthermore, promoters with activatory Fis binding sites positioned close to -91, as in the *pet* promoter, generally have a second, downstream transcription factor binding site. This second site has seen to be a Fis binding sites, most commonly positioned around -70, and a CRP binding site; one promoter possessing a downstream CRP binding site at -63.5 (Keseler *et al.* 2011). No promoters possessing the *pet* promoter architecture had been studied prior to Rossiter *et al.* 2011. This suggests the mechanism of transcription activation by Fis, bound around position -91, requires the action of a downstream factor bound on the same face of the DNA helix. This requirement is presumably due to the DNA bending provided by the downstream factor (Fis or CRP) which facilitates activation by the upstream Fis. Favourable variation of DNA topology offers another possibility for fisI mediated enhancement where DNA bending that accompanies Fis binding confers a beneficial conformational change of the local downstream (Muskhelishvili and Travers 2003). Whether or not a direct interaction between the RNAP and upstream Fis in the *pet* promoter occurs is unclear, and therefore requires further investigation to determine the precise mechanism of activation. Mutational analysis of the promoter failed to show Fis repressing

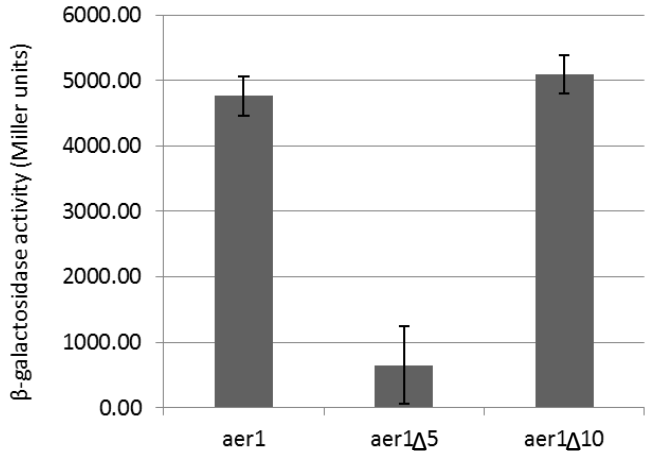


Figure 3.2: Nested Deletions Within the aer1 Fragment

A graph displaying the relative activities of constructs of *aer1*, containing nested deletions, when cloned into pRW50 and transformed into E. *coli* BW25113; following a β-galactosidase assay. Where *aer1 Δ5* has contains a 5 nt deletion from the spacer between the CRP and fisI site and *aer1 Δ10* contains a 10 nt deletion from the same region. Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by error bars.

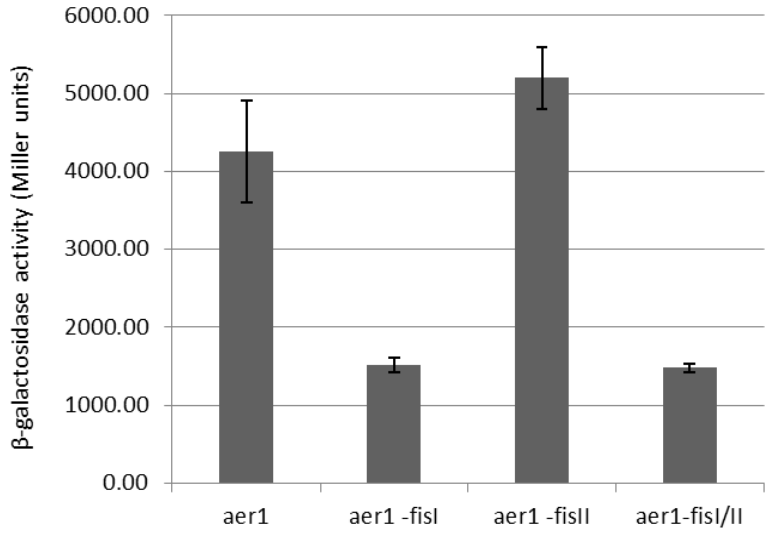


Figure 3.3: Point Mutations Within the aer1 Fragment

A graph displaying the relative activities of aer1constructs, containing point mutation, when cloned into pRW50 and transformed into E. *coli* BW25113; following a β-galactosidase assay. Where the fisI site of *aer1* has been disrupted (*aer1 –fisI*), the fisII site been disrupted (*aer1 –fisII*) and where both fisI and fisII sites have been disrupted (*aer1 –fisI/II*) (Rossiter *et al.* 2011). Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by error bars.

expression from the fisII site; when both the fisI and fisII were deleteriously mutated. The repressive actions of Fis at the fisII site must therefore dependent on the presence of a functional fisI site. Several mechanisms whereby fisII represses transcription, requiring a functional fisI site, are possible. (**1**) A direct interaction between Fis bound at fisII and RNAP, facilitated by the bending of the DNA by the downstream bound CRP and Fis, could modulate transcription by inhibiting promoter escape (Lee and Busby 2012). (**2**) The fisII site competes with fisI site for the recruitment of Fis, reducing the levels of expression by reducing the occupation of fisI (Lickwar *et al.* 2012). (**3**) Fis may modulate transcription by negatively altering the topology of the promoter, although which factors within the promoter this might affect is unclear (Schneider *et al.* 1999; Grainger *et al.* 2008). Further investigation is required to determinate which mechanism of repression occurs.

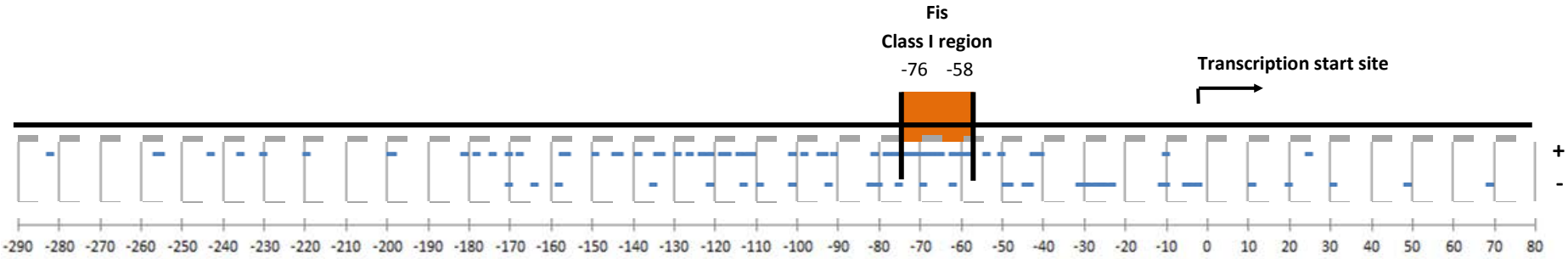


Figure 3.4: Fis Regulatory Sites in *E. coli* K12

This diagram shows the positions Fis DNA binding sites in regards to the transcription start site, from a database of sequenced and annotated promoters (ecocyc). Fis binding sites were plotted based on the position of the binding site and if it activates (+) or represses (-) gene expression (Keseler *et al.* 2011). The orange box shows the region, -58 to-76, where Fis is seen to promote expression solely in a class I manner. Activating Fis site upstream of this region (-58 to -76) are seen to contain a second downstream transcription binding site.

Chapter 4

Emulation of the *pet* Promoter Architecture in a Model Promoter

4.1 Introduction

There are many questions regarding the transcription regulation of *pet* that still remain. The contribution of promoter architecture upon transcription regulation was investigated; the architecture of the *pet* promoter was emulated in a well-studied promoter (Rossiter *et al*. 2011). The well-studied promoter, CC(-40.5), derived from MelR, whose CRP binding site is centred 40.5 bp upstream of the transcription start site. The fisI site was cloned 35 bp upstream of the CRP binding site; resulting in the promoter F_{35}CC(-40.5) (Rossiter *et al*. 2011). This model system provided a base upon which positions of the transcription factor binding sites could be altered, allowing the effect of the position of these binding sites on gene expression to be assessed. This work presented in this chapter evaluates the mechanisms utilised by Fis in the transcription regulation of *pet*-like promoters, in terms of architecture, and probing the proposed intimate interactions between the α subunit of RNAP and Fis.

4.2 Emulation of PET Promoter Architecture

The mechanistic action of the fisII site of the *pet* promoter is not well defined; questions regarding how transcription is repressed still remain. To address this matter, a consensus Fis binding site (Fc) was cloned 66 bp upstream of the FisI site of F_{35}CC(-40.5); producing $Fc_{66}F_{35}$CC(-40.5) (Figure 4.1). The addition of an upstream Fis binding site, stronger than that of the fisII of *pet*, should have the effect of causing repression equal to if not greater than that observed in the *pet* promoter. The levels of expression from the F_{35}CC(-40.5) and $Fc_{66}F_{35}$CC(-40.5) show no significant change, implying the addition of the upstream Fis binding site has no noticeable effect on expression (figure 4.2).

4.3 Activity of FnCC(-40.5) Promoter Series

Fis is able to activate transcription, directly and indirectly, via variety of mechanisms. Fis appears to activated transcription, bound at fisI, via a class I activation mechanism (Travers and Muskhelishvili, 1998). To probe the mechanism by which Fis collaborates with CRP to induce transcription, a series of promoter constructs were created (FnCC(-40.5) where the spacer region between the Fis and CRP binding sites was varied in length (n = number of bases between the centres of the Fis and CRP DNA binding site) (Figure 4.1) (Rossiter *et al* 2011). The expressions of these promoters were assessed under wild type and Δ*fis* genetic backgrounds. Several peaks of expression were observed (F_{25}CC(-40.5), F_{35}CC(-40.5) and F_{45}CC(-40.5)); the level of expression decreased as "n" increased (figure 4.3). Those promoters with n+/-1 of the F_{25}CC(-40.5) peak (F_{24}CC(-40.5) and F_{26}CC(-40.5)) were Fis dependant, but had levels of expression lower than the peak. The promoters that generated these three peaks of expression have fisI sites on the same face of the DNA helix as the crp site; the

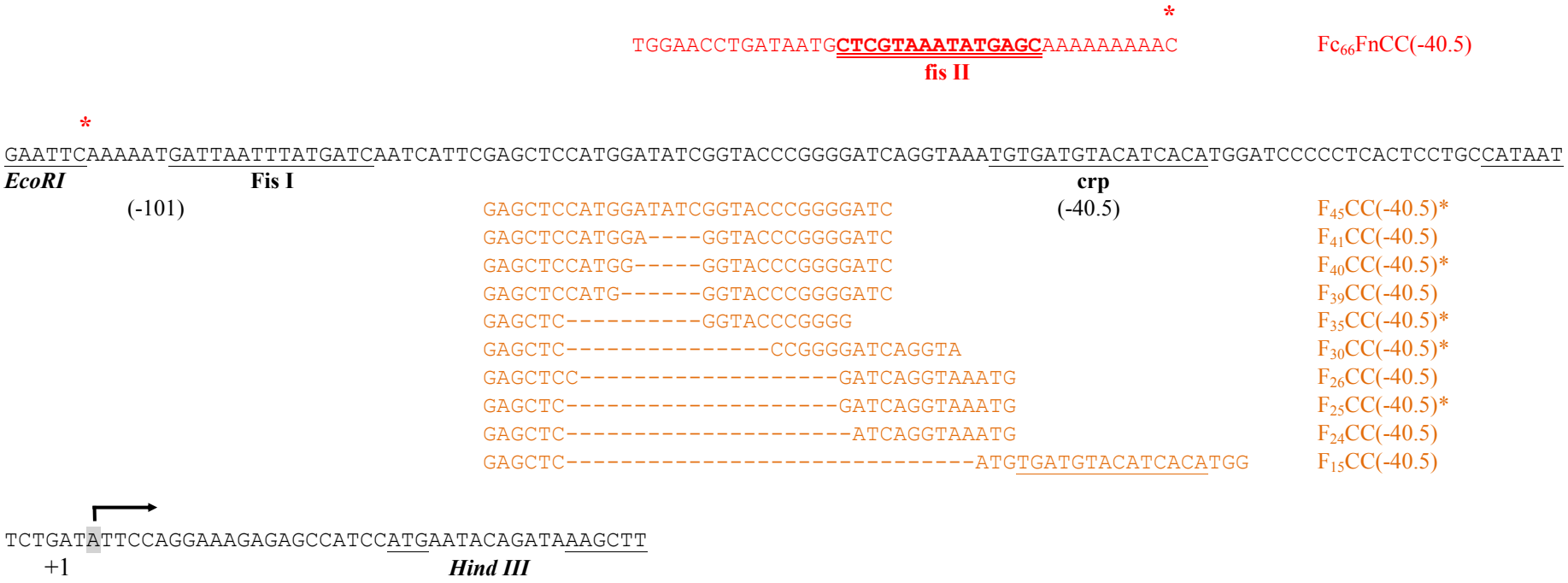


Figure 4.1: The FnCC(-40.5) series

The sequence of the FnCC(-40.5) series with transcription factor binding regions are underlined and labelled with the sites name and positions in regards to the transcription start side (+1). The series differs by length of sequence located between the CRP site (-40.5) and the fisI site; where n signifies the number of bases between the centres of these two site. The sequence for F_{45}CC(-40.5) is shown in black with the orange dashes (-) flanked by complementary sequences denoting the bases removed in the construction of each of the FnCC(-40.5) series. The construction of $Fc_{66}F_{35}$CC(-40.5) differs from the series by the addition of the red sequence from the red asterisk (*).

* Rossiter *et al.* 2011

peaks are separated by 10 bp, or one completed turn of the DNA helix. This suggests that Fis is able to induce transcription when correctly orientation on the same side of the DNA helix as the CRP DNA binding site. As fisI is moved away further from the 3 peak positions (n=25, 35 & 45) the expression is reduced until Fis does not induce expression, with levels reduced to that to that of CC-40.5. These results, when put together, lend credence to the proposed class I interaction between Fis and RNAP, but direct evidence of this interaction is still lacking.

4.4 Effect of α-CTD alanine substitution mutants on the expression of the FnCC(-40.5) Series

Transcription factors are able to modulate gene expression, particularly transcription initiation, through a number of mechanisms. Class I and class II activators are able to induce transcription by binding to specific sites within the promoter region and recruit RNAP to the promoter; through direct interactions with the carboxyl terminal domain of the RNAP α subunit (α-CTD). The residues of the α subunit involved in this interaction depends on the transcription factor involved and the position of the transcription factor binding site in the promoter. By substituting the interacting residues, for alanine, electrostatic interactions between the transcription factor and α-CTD can be disrupted, reducing the transcriptions factor ability to recruit RNAP and induce transcription. By measuring the reduction in gene expression it is possible to measure the impact an individual amino acid on the overall interactions. CRP interacts with positions 271 and 273 of the α subunit when acting as a class II activator and Fis interacts with positions 271, 285 and 287 of the α subunit as class I activator (Savery *et al.* 1998; Aiyar *et al.* 2002). These 4 residues (271, 273, 285 and 287) and the control residue (261) were assessed across 4 promoters (CC(-40.5), F_{25}pCC(-40.5), F_{30}pCC(-40.5) and F_{35}pCC(-40.5)). The alanine substitution of residues 285 and 287 affected all of the promoters, consistent with the activation of all of the promoters by CRP (Figure 4.4). The alanine substitution of residue 271 reduces the expression of F_{30}CC(-40.5) and F_{35}CC(-40.5), and the alanine substitution of residue 273 affects the Fis induced promoters (F_{25}(CC-40.5) and F_{35}(CC-40.5)). These two results are consistent with the notion that promoter F_{25}(CC-40.5) and F_{35}(CC-40.5) are induced by a class I Fis mechanism. The substitution of residue 261 reduced the expression of the CC(-40.5) promoter, which was unexpected as this reduction was not previously seen in CC(-41.5) but may be due to the architectural differences between these promoters activator (Savery *et al.* 1998; Aiyar *et al.* 2002). These results require further proof to determine if an interaction between Fis and the α-CTD occurs.

4.5 Effect of the Substitution of Critical Fis Residues on Expression of F_{25}CC(-40.5)

Fis is able to act as a transcriptional activator by interacting with the α-CTD through the critical residues 71 and 72, of Fis; substitution of these residues with other amino acids is able to modulate the nature of this interaction

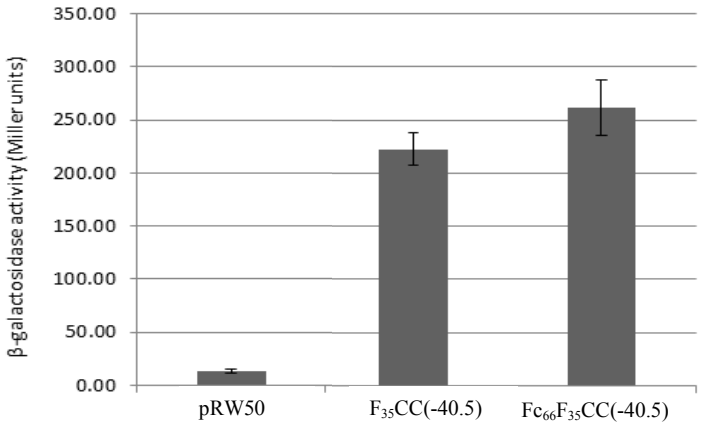


Figure 4.2: Implications of a Second Upstream Fis Binding Site

A graph displaying the relative activities of $F_{35}CC(40.5)$ and $Fc_{66}F_{35}CC(40.5)$ constructs when cloned into pRW50 and transformed into E. *coli* BW25113; following a β-galactosidase assay. F**c**n signifies the presence a Fis binding site, of consensus sequence, and F**p**n the fisI site from the *pet* promoter; where n indicates the how many nucleotides the centre of the Fis binding site is upstream of the centre of the downstream CRP binding site, centred at -40.5, in respects to the transcription start site. Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by error bars.

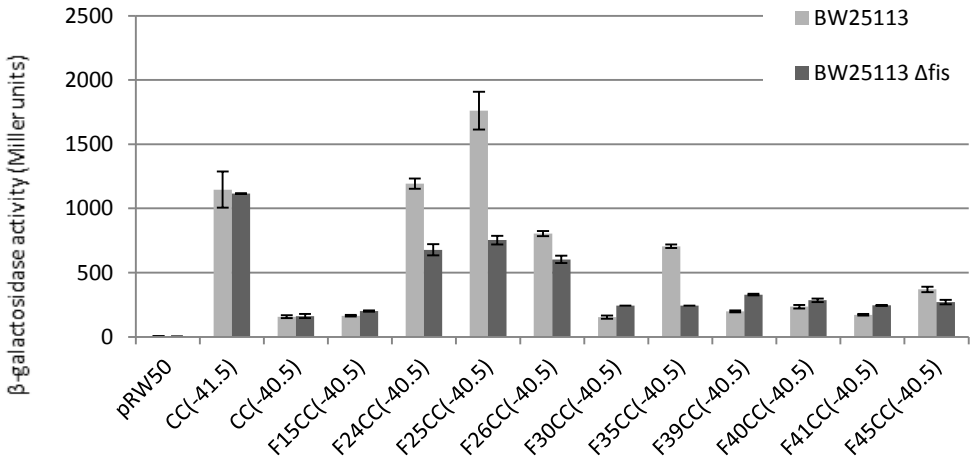


Figure 4.3: Transcriptional Activity of the FnCC(-40.5) Series

A graph displaying the relative activities of FnCC(40.5) series constructs when cloned into pRW50 and transformed into E. *coli* BW25113 or E. *coli* BW25113 Δ*fis*; following a β-galactosidase assay. Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by error bars. The activities were also normalised by determining the average difference in β-galactosidase activity between BW25113 and BW25113 Δ*fis*, for CC(-40.5) and CC(-41.5), and applying the normalisation factor to the β-galactosidase activities of subsequent constructs.

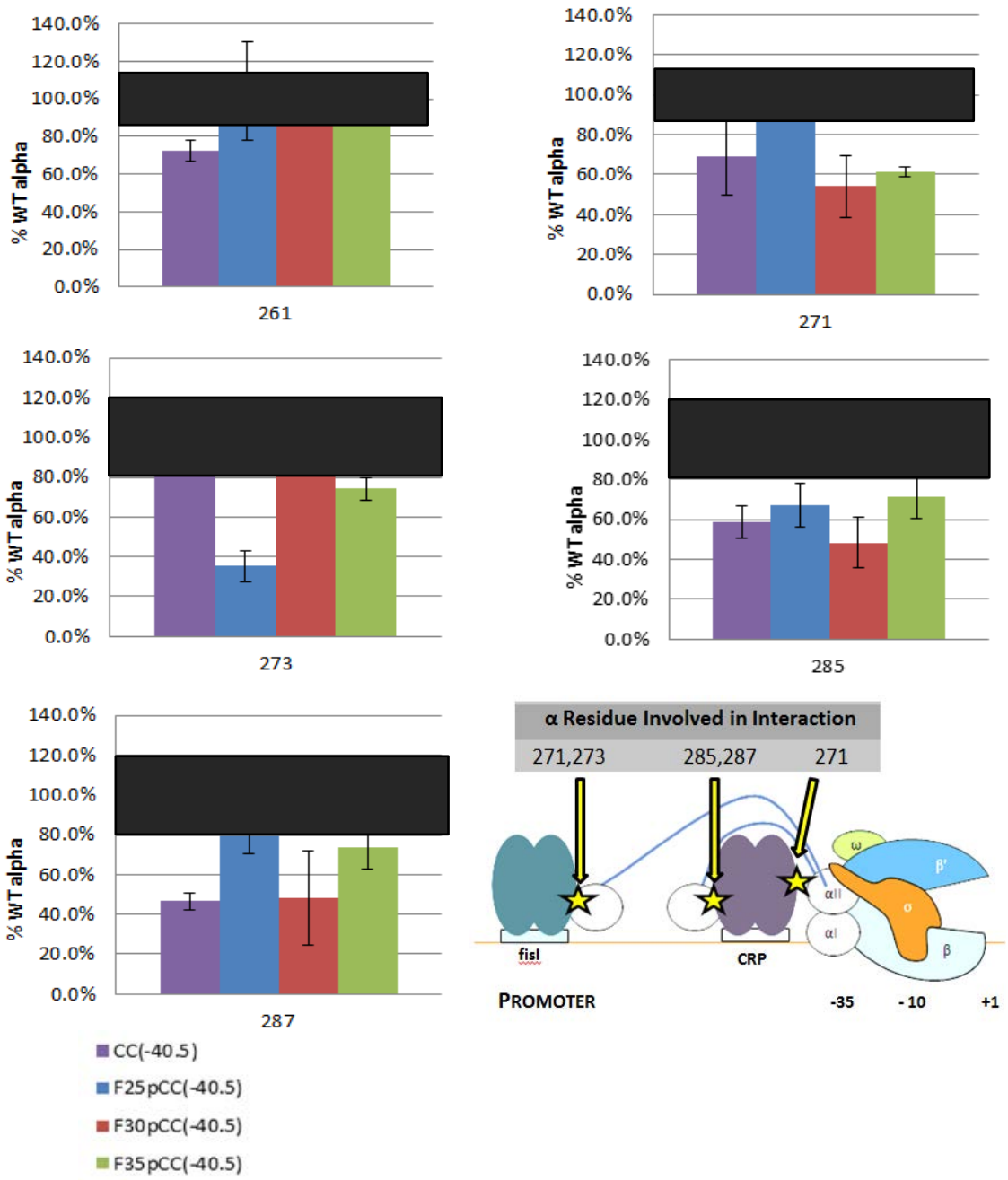


Figure 4.4: Impact of Alanine Substitutions Within α on FnCC(-40.5) Series

β-galactosidase assay displaying the activities of FnCC(40.5) constructs carried by pRW50 within E. *coli* BW25113 containing either the pHTfIα (PHT) or pREIIα (PRE) plasmid(Blatter *et al.* 1994; Gaal *et al.* 1996; Wood *et al.* 1997; Tang *et al.* 1994; Gaal *et al.* 1996)); containing either WT α or α containing an alainine substitution at positions 261 or 271 (PHT), or 273, 285 or 287(PRE). Results are shown as a percentage of WT α β-galactosidase activity when in plasmid, PRE or PHT; WT α is shown as 100% and the greatest standard deviation of WT α is displayed as the grey shaded box. Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by bars. The bottom right cartoon represents the pet promoter with RNAP, CRP and Fis bound; residues of α that know determinates in the class I Fis interaction and class II CRP are indicated.

(Bokal *et al.* 1997; Cheng *et al.* 2000; Aiyar *et al.* 2002). Residue 71 is important as it contributes to an electrostatic interaction with the α-CTD, whilst residue 72 maintains the environment of the α-CTD contacting region, with the disruption of either through deleterious substitutions, able to negatively affect transcription (Pan *et al.* 1996). The contributions of these residues was tested for the promoter F_{25}CC(-40.5) in a Δ*fis* genetic background complemented with wild type Fis or Fis containing an amino acid substitution (R71Y, E71K or G72A). When residue 71 is substituted for Y or K the expression is reduced by 50% when compared to wild type (WT+ Fis. When residue 72 is substituted for A expression is reduced to 25% of wild type. These results suggest residue 71 is one of a number of determinants for the interaction and when disrupted expression is reduced and when residue 72 is substituted for A the entire determinant region is disrupted reducing expression further.

4.6 Discussion

The emulation of *pet* promoter architecture in the model promoter allowed the effect of the *pet* promoter architecture to be assessed independently. These results tend to agree the proposed regulatory mechanisms that occur at the *pet* promoter, although there are a few inconsistencies that remained to be rectified. The first disparity between the promoter systems is the inability of an inserted fisII site to reduce the expression from the F_{35}CC(-40.5) promoter (Figure 4.2). Previously the repressive action of the fisII site was taken as being due to several possibilities: a direct interaction with RNAP, sequestration of Fis from fisI site or a conformational change to promoter, with these mechanisms dependant on a function fisI site. From these results any direct interaction and sequestration of Fis are unlikely and can be ruled out as the architecture of the *pet* promoter was fully reconstructed, indeed with a higher affinity Fis binding site. A conformation model appears to be the most likely candidate, where the change in promoter conformation due to Fis binding at the fisII site in appears to only affect the pet promoter negatively presumably due to subtle sequence difference between the promoters.

Fis is able to act as an activator in these promoters where the fisI site is postioned on the same side of the DNA helix as the CRP site; presumably this allows the two α subunits to interact with the two transcription factors (Figure 4.3). As fisI is placed at positions on the DNA helix other than that of peak positions (n=25, 35 & 45) the ability of Fis to induce expression was seen to the gradually diminishes until the induction by Fis is lost (n= 30, 39, 40 & 41). The ability of promoters, where fisI is located adjacent to the peak positions (n= 24 & 26), to induce expression, albeit at lowers level, is accommodated by the flexibility of the α subunits linker region. The decrease in expression presumably occurs because of the higher free energy costs required for the complex to adopt suboptimal conformations. The optimal positions for the fisI site occur every 10 bp (1 revolution of the

DNA helix), the expression from the optimal positions increasing as the spacer between the fisI and crp sites decrease, lending credence to the possibility of a class I interaction. This trend is broken with the F_{15}CC(-40.5) promoter, where Fis is no longer able to induce expression, likely as in this case the Fis binding site is located too close to the CRP binding site, causing a steric clash, preventing binding and subsequent activation. These results together hint at a direct interaction between Fis and α-CTD, but does not rule out the possibility of Fis acting as a conformational activator.

Class I and II activation requires a direct interaction, between the transcription factor and specific residues of the α-CTD. These results show that the expressions from those promoters, which are activated by Fis, expressions are reduced when residue 273 is substituted with alanine. Whilst all of the promoters are negatively affected by the alanine substitutions of the CRP specific residues 271, 285 and 287, except F_{25}CC(-40.5), the only promoter not affected by residue 271 (Figure 4.4). When F_{25}CC(-40.5) is compared to F_{35}CC(-40.5), residue 273 has a much greater effect on F_{25}CC(-40.5) along with the inactivity of residue 271. This suggests the change the change in spacer length is having an effect on the interaction with Fis and the α-CTD channelling the interaction through fewer residues. This could explain why F_{25}CC(-40.5) is more highly expressed, as reducing the strength of an unduly strong promoter can aid promoter escape and ultimately increase expression (Miroslavova and Busby 2006). A discrepancy in this experiment, is the reduced expression from the CC(-40.5) promoter due to the substitution of residue 261. CC(-40.5) has not been alanine screened before and the reduction in expression, due to the alanine substation of residue 261, might be an intinstic charaterisic of the CC(-40.5) promoter. These results require further evidence to provide improve the confidence in the existence of a direct interaction between Fis and α-CTD.

Fis is able to interact with the α-CTD through determinants including residues 71 – 73. Substitution of residue 71 can affect the direct electrostatic interaction with the α-CTD and substitution of residue 72 can disrupt the surrounding residues, abolishing the interaction with the α-CTD (McLeod et al 1999). Firstly the results show the when the pKK233-3 plasmid does not express Fis the levels of expression from F_{25}CC(-40.5) are lower than that that when pKK233-3 expresses G72A Fis (Figure 4.5). Fis is responsible for repressing the expression of CRP in a manner not dependent on interactions with the α subunit. Therefore all of the pKK233-3 plasmids that express any of the Fis mutants are able to control the levels of CRP, whilst when pKK233-3 is does not express Fis repression does not occur, leading to higher levels of CRP and consequently higher levels of expression from

the F_{25}CC(-40.5) promoter. When the residue 71 is substituted for Y/K the electrostatic interaction with the α-CTD is disrupted leading to reduced levels of expression. When residue 72 is substituted expression

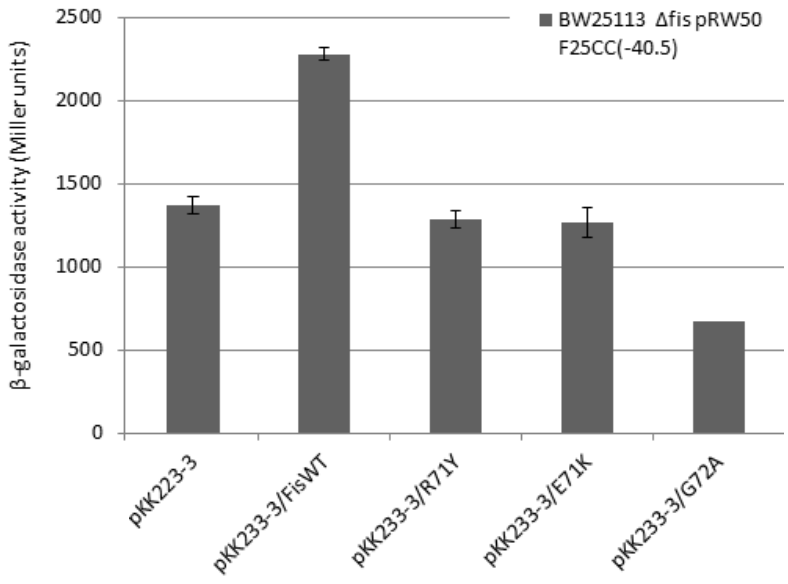


Figure 4.5: Impact of Amino Acid Substitutions Within Fis on F_{25}CC(-40.5)

β-galactosidase assay displaying the activities of F_{25}pCC(40.5) constructs carried by pRW50 within E. *coli* BW25113 Δ*fis* containing the either the PKK233-3 plasmid or the PKK233-3 plasmid containing WT Fis or Fis containing an amino acid substitution at position 71 (Y/K) or position 72 (A). Cell were grown anaerobically at 37°C in LB broth until early log phase (OD_{650}=0.4) and results are the average of three independent assays; standard deviation shown by error bars.

becomes even lower, this is consist with the idea that altering residue 72 disrupts the surrounding α-CTD interacting residues abolishing the interactions. These results collaborate with the α-CTD substitution and the spacer variation experiments, and leads to the conclusion that Fis activates transcription in a class I mechanism in the FnCC(-40.5) series.

Chapter 5

Conclusions

The pathogenic properties of EAEC strain 042 result from the combination of several distinct processes, adherence to intestinal mucosa, biofilm formation and toxin expression, facilitated by the coordinated expression of many virulence factors including Pet, itself targeting α-fodrin, disrupting the actin cytoskeleton of host cells (Eslava et al. 1998; Nataro 2005; Croxen and Finlay 2010). The expression of Pet is governed by the global transcription factors CRP and Fis, CRP activating transcription from a non-canonical bindings site (-40.5), with further enhancement provided by Fis bound at the fisI site (-91) (Rossiter et al. 2011). When EAEC shifts from a nutrient limited to a nutrient rich environment it soon enters the log phase of growth where levels of cAMP fall resulting in the gradual repression of both CRP (Hanamura and Aiba 1991; Nasser et al. 2001). However Pet expression occurs in a temporal window at the initiation of nutrient upshift where a transient peak of Fis expression occurs; expression of Pet tailing off as the levels of active CRP are depleted (Ball et al. 1992). This mechanism of co-activation allows the expression of Pet to be restricted to early log phase when EAEC encounters a nutrient rich environment to compete with the host microbiome and initiate infection (Betancourt-Sanchez and Navarro-Garcia 2009).

Previous work was able to show transcription regulation of *pet* expression is dependent on CRP and Fis. The worked assessed the mechanisms employed to affect transcription to regulate transcription. CRP activates transcription through a class II mechanism from its non-canonical binding site; this interaction dependant on the residues 185, 187 and 173 of the α-CTD and AR1 and AR2 of CRP (Figure 4.4) (Fig. S3 Rossiter *et al.* 2011). Fis was known to play an important role in the regulation of the *pet* promoter, activating from the fisI site and repressing from the fisII site. Fis has been shown to function as an activator from the fisI site when CRP is bound to the promoter, facilitated by residues 173 and 171 of the α-CTD and residues 71 and 73 of Fis (Figure 4.4; Figure 4.5). The dependence on these residues of Fis and the α-CTD alongside the helical phasing dependence (Figure 3.3; Figure 4.3) strongly suggests a Fis dependant class I interaction, from the fisI site, induces transcription. The reason for the dependence on the presence of a bound CRP for the functioning of Fis has yet to be fully established, although it seems likely the DNA bending provided by CRP compensates for the limited length of the α subunit linker region and the larger spacer between fisI and the -10 element, allowing the interaction to occur. The second Fis binding site, fisII, was shown to be repressive in the *pet* promoter; despite further work the underlying mechanism remains elusive (Figure 3.2) (Rossiter *et al*. 2011). The compilation of the results presented delivers the most complete model yet for the transcription regulation of the *pet* promoter (Figure 5.1). In summary, this study has been able to provide some deeper understanding of the principal interactions that regulate the expression of the pet promoter, with the focus residing on the fisI site.

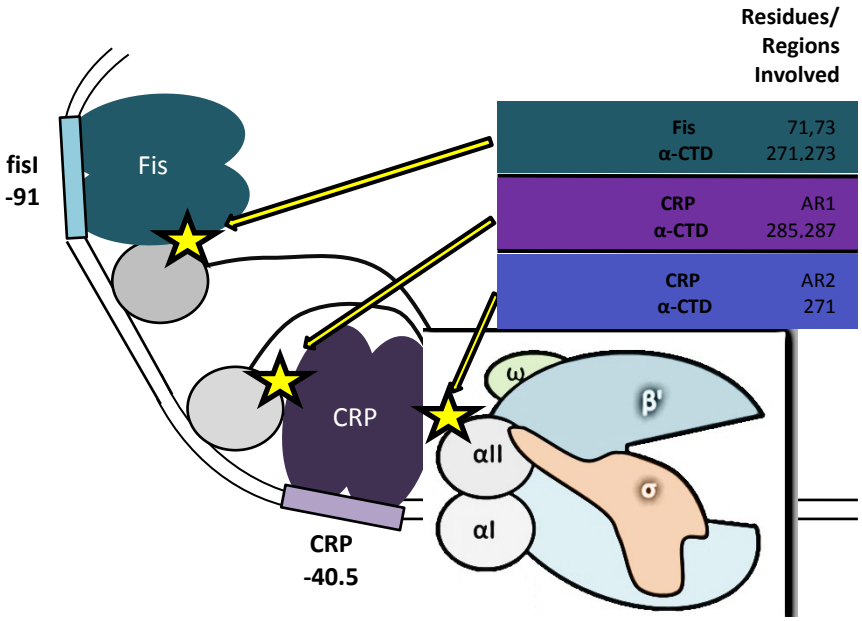


Figure 5.1: Model of Transcription Regulation at the *pet* Promoter

Model displaying the proposed interactions implicated in transcription regulation at the *pet* promoter. The yellow stars indicate the location of direct interactions with the residues of the transcription factor and α-CTD confirmed to be involved in induction of transcription. CRP binds to the non-canonical DNA binding site, causing the DNA to bend; presumably the DNA bending facilitates the interaction between Fis, bound at fisI, and RNAP.

Nevertheless, there are many questions regarding of the pet system that remain unanswered; the effect of DNA bending on the interaction between Fis and RNAP, the specific residues responsible for the interaction between α-CTD and each transcription factor, and the repressive nature of the fisII site. Further understanding the *pet* promoter will underpin the understanding of other promoters of virulence factors of related architecture (Rossiter *et al.* 2011).

References

Abbondanzieri, E. A., W. J. Greenleaf, J. W. Shaevitz, R. Landick and S. M. Block (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature*. **438**, 460-5.

Adelman, J. L., Y. J. Jeong, J. C. Liao, G. Patel, D. E. Kim, G. Oster and S. S. Patel (2006) Mechanochemistry of transcription termination factor Rho. *Molecular cell*. **22**, 611-21.

Adelman, K., A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts and M. D. Wang (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proceedings of the National Academy of Sciences of the United States of America*. **99**, 13538-43.

Aiyar, S. E., Y. L. Juang, J. D. Helmann and P. L. deHaseth (1994) Mutations in sigma factor that affect the temperature dependence of transcription from a promoter, but not from a mismatch bubble in double-stranded DNA. *Biochemistry*. **33**, 11501-6.

Aiyar, S. E., S. M. McLeod, W. Ross, C. A. Hirvonen, M. S. Thomas, R. C. Johnson and R. L. Gourse (2002) Architecture of Fis-activated transcription complexes at the *Escherichia coli rrnB P1* and rrnE P1 promoters. *Journal of molecular biology*. **316**, 501-16.

Alifano, P., F. Rivellini, D. Limauro, C. B. Bruni and M. S. Carlomagno (1991) A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell*. **64**, 553-63.

Anantham, S. and R. M. Hall (2012) pCERC1, a Small, Globally Disseminated Plasmid Carrying the dfrA14 Cassette in the *strA* Gene of the *sul2-strA-strB* Gene Cluster. *Microbial drug resistance*.

Artsimovitch, I. and R. Landick (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proceedings of the National Academy of Sciences of the United States of America*. **97**, 7090-5.

Artsimovitch, I. and R. Landick (2002) The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell*. **109**, 193-203.

Asahara, T., K. Shimizu, K. Nomoto, T. Hamabata, A. Ozawa and Y. Takeda (2004) Probiotic bifidobacteria protect mice from lethal infection with Shiga toxin-producing *Escherichia coli* O157:H7. *Infection and immunity*. **72**, 2240-7.

Azam, T. A. and A. Ishihama (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *The Journal of biological chemistry*. **274**, 33105-13.

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

Backhed, F., R. E. Ley, J. L. Sonnenburg, D. A. Peterson and J. I. Gordon (2005) Host-bacterial mutualism in the human intestine. *Science*. **307**, 1915-20.

Bai, L., A. Shundrovsky and M. D. Wang (2004) Sequence-dependent kinetic model for transcription elongation by RNA polymerase. *Journal of molecular biology*. **344**, 335-49.

Ball, C. A., R. Osuna, K. C. Ferguson and R. C. Johnson (1992) Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *Journal of bacteriology*. **174**, 8043-56.

Banerjee, S., J. Chalissery, I. Bandey and R. Sen (2006) Rho-dependent transcription termination: more questions than answers. *Journal of microbiology*. **44**, 11-22.

Bar-Nahum, G., V. Epshtein, A. E. Ruckenstein, R. Rafikov, A. Mustaev and E. Nudler (2005) A ratchet mechanism of transcription elongation and its control. *Cell*. **120**, 183-93.

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

Beloin, C., A. Roux and J. M. Ghigo (2008) *Escherichia coli* biofilms. *Current topics in microbiology and immunology*. **322**, 249-89.

Bendtsen, K. M., J. Erdossy, Z. Csiszovszki, S. L. Svenningsen, K. Sneppen, S. Krishna and S. Semsey (2011) Direct and indirect effects in the regulation of overlapping promoters. *Nucleic acids research*. **39**, 6879-85.

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

Bettelheim, K. A., A. Breadon, M. C. Faiers, S. M. O'Farrell and R. A. Shooter (1974) The origin of O serotypes of *Escherichia coli* in babies after normal delivery. *The Journal of hygiene*. **72**, 67-70.

Bettelheim, K. A., M. Faiers and R. A. Shooter (1972) Serotypes of *Escherichia coli* in normal stools. *Lancet*. **2**, 1223-4.

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

Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*. **277**, 1453-62.

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

Boltner, D., C. MacMahon, J. T. Pembroke, P. Strike and A. M. Osborn (2002) R391: a conjugative integrating mosaic comprised of phage, plasmid, and transposon elements. *Journal of bacteriology*. **184**, 5158-69.

Borukhov, S. and E. Nudler (2003) RNA polymerase holoenzyme: structure, function and biological implications. *Current opinion in microbiology*. **6**, 93-100.

Borukhov, S., V. Sagitov and A. Goldfarb (1993) Transcript cleavage factors from *E. coli. Cell*. **72**, 459-66.

Bowers, C. W. and A. J. Dombroski (1999) A mutation in region 1.1 of sigma70 affects promoter DNA binding by *Escherichia coli* RNA polymerase holoenzyme. *The EMBO journal*. **18**, 709-16.

Brendel, V., G. H. Hamm and E. N. Trifonov (1986) Terminators of transcription with RNA polymerase from *Escherichia coli*: what they look like and how to find them. *Journal of biomolecular structure & dynamics*. **3**, 705-23.

Brodolin, K., N. Zenkin, A. Mustaev, D. Mamaeva and H. Heumann (2004) The sigma 70 subunit of RNA polymerase induces *lacUV5* promoter-proximal pausing of transcription. *Nature structural & molecular biology*. **11**, 551-7.

Browning, D. F. and S. J. Busby (2004) The regulation of bacterial transcription initiation. *Nature reviews. Microbiology*. **2**, 57-65.

Browning, D. F., D. C. Grainger and S. J. Busby (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Current opinion in microbiology*. **13**, 773-80.

Burr, T., J. Mitchell, A. Kolb, S. Minchin and S. Busby (2000) DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic acids research*. **28**, 1864-70.

Busby, S. and R. H. Ebright (1997) Transcription activation at class II CAP-dependent promoters. *Molecular microbiology*. **23**, 853-9.

Caddick, M. X., H. N. Arst, Jr., L. H. Taylor, R. I. Johnson and A. G. Brownlee (1986) Cloning of the regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. *The EMBO journal*. **5**, 1087-90.

Campbell, E. A., O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz and S. A. Darst (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular cell*. **9**, 527-39.

Cashel, M., L. M. Hsu and V. J. Hernandez (2003) Changes in conserved region 3 of *Escherichia coli* sigma 70 reduce abortive transcription and enhance promoter escape. *The Journal of biological chemistry*. **278**, 5539-47.

Caugant, D. A., B. R. Levin and R. K. Selander (1981) Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics*. **98**, 467-90.

Chakraborty, A., D. Wang, Y. W. Ebright, Y. Korlann, E. Kortkhonjia, T. Kim, S. Chowdhury, S. Wigneshweraraj, H. Irschik, R. Jansen, B. T. Nixon, J. Knight, S. Weiss and R. H. Ebright (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science*. **337**, 591-5.

Chan, C. L. and R. Landick (1989) The Salmonella typhimurium his operon leader region contains an RNA hairpin-dependent transcription pause site. Mechanistic implications of the effect on pausing of altered RNA hairpins. *The Journal of biological chemistry*. **264**, 20796-804.

Chang, D. E., D. J. Smalley, D. L. Tucker, M. P. Leatham, W. E. Norris, S. J. Stevenson, A. B. Anderson, J. E. Grissom, D. C. Laux, P. S. Cohen and T. Conway (2004) Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America*. **101**, 7427-32.

Chaudhuri, R. R., M. Sebaihia, J. L. Hobman, M. A. Webber, D. L. Leyton, M. D. Goldberg, A. F. Cunningham, A. Scott-Tucker, P. R. Ferguson, C. M. Thomas, G. Frankel, C. M. Tang, E. G. Dudley, I. S. Roberts, D. A. Rasko, M. J. Pallen, J. Parkhill, J. P. Nataro, N. R. Thomson and I. R. Henderson (2010) Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PloS one*. **5**, e8801.

Cheetham, G. M. and T. A. Steitz (1999) Structure of a transcribing T7 RNA polymerase initiation complex. *Science*. **286**, 2305-9.

Chen, J., S. A. Darst and D. Thirumalai (2010) Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proceedings of the National Academy of Sciences of the United States of America*. **107**, 12523-8.

Chen, S., J. Vojtechovsky, G. N. Parkinson, R. H. Ebright and H. M. Berman (2001) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: DNA binding specificity based on energetics of DNA kinking. *Journal of molecular biology*. **314**, 63-74.

Chen, Y. and R. H. Ebright (1993) Phenyl-azide-mediated photocrosslinking analysis of Cro-DNA interaction. *Journal of molecular biology*. **230**, 453-60.

Cheng, Y. S., W. Z. Yang, R. C. Johnson and H. S. Yuan (2000) Structural analysis of the transcriptional activation region on Fis: crystal structures of six Fis mutants with different activation properties. *Journal of molecular biology*. **302**, 1139-51.

Cho, B. K., E. M. Knight, C. L. Barrett and B. O. Palsson (2008) Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome research*. **18**, 900-10.

Chopra, I., A. J. O'Neill and K. Miller (2003) The role of mutators in the emergence of antibiotic-resistant bacteria. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*. **6**, 137-45.

Ciampi, M. S. (2006) Rho-dependent terminators and transcription termination. *Microbiology*. **152**, 2515-28.

Claus, S. P., S. L. Ellero, B. Berger, L. Krause, A. Bruttin, J. Molina, A. Paris, E. J. Want, I. de Waziers, O. Cloarec, S. E. Richards, Y. Wang, M. E. Dumas, A. Ross, S. Rezzi, S. Kochhar, P. Van Bladeren, J. C. Lindon, E. Holmes and J. K. Nicholson (2011) Colonization-induced host-gut microbial metabolic interaction. *mBio*. **2**, e00271-10.

Cooper, V. S. and R. E. Lenski (2000) The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature*. **407**, 736-9.

Croxen, M. A. and B. B. Finlay (2010) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature reviews. Microbiology*. **8**, 26-38.

deHaseth, P. L. and J. D. Helmann (1995) Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Molecular microbiology*. **16**, 817-24.

Denamur, E., G. Lecointre, P. Darlu, O. Tenaillon, C. Acquaviva, C. Sayada, I. Sunjevaric, R. Rothstein, J. Elion, F. Taddei, M. Radman and I. Matic (2000) Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell*. **103**, 711-21.

Deutscher, J. (2008) The mechanisms of carbon catabolite repression in bacteria. *Current opinion in microbiology*. **11**, 87-93.

Dombroski, A. J. (1997) Recognition of the -10 promoter sequence by a partial polypeptide of sigma70 *in vitro*. *The Journal of biological chemistry*. **272**, 3487-94.

Dombroski, A. J., W. A. Walter and C. A. Gross (1993) Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes & development*. **7**, 2446-55.

Dove, S. L., S. A. Darst and A. Hochschild (2003) Region 4 of sigma as a target for transcription regulation. *Molecular microbiology*. **48**, 863-74.

Dutta, D., J. Chalissery and R. Sen (2008) Transcription termination factor rho prefers catalytically active elongation complexes for releasing RNA. *The Journal of biological chemistry*. **283**, 20243-51.

Ebright, R. H. (1993) Transcription activation at Class I CAP-dependent promoters. *Molecular microbiology*. **8**, 797-802.

Ebright, R. H., P. Cossart, B. Gicquel-Sanzey and J. Beckwith (1984) Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli*. *Nature*. **311**, 232-5.

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

Ebright, R. H., Y. W. Ebright, P. S. Pendergrast and A. Gunasekera (1990) Conversion of a helix-turn-helix motif sequence-specific DNA binding protein into a site-specific DNA cleavage agent. *Proceedings of the National Academy of Sciences of the United States of America*. **87**, 2882-6.

Ebright, R. H., A. Kolb, H. Buc, T. A. Kunkel, J. S. Krakow and J. Beckwith (1987) Role of glutamic acid-181 in DNA-sequence recognition by the catabolite gene activator protein (CAP) of *Escherichia coli*: altered DNA-sequence-recognition properties of [Val181]CAP and [Leu181]CAP. *Proceedings of the National Academy of Sciences of the United States of America*. **84**, 6083-7.

Ebright, R. H., S. F. Le Grice, J. P. Miller and J. S. Krakow (1985) Analogs of cyclic AMP that elicit the biochemically defined conformational change in catabolite gene activator protein (CAP) but do not stimulate binding to DNA. *Journal of molecular biology*. **182**, 91-107.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman (2005) Diversity of the human intestinal microbial flora. *Science*. **308**, 1635-8.

Epshtein, V., C. J. Cardinale, A. E. Ruckenstein, S. Borukhov and E. Nudler (2007) An allosteric path to transcription termination. *Molecular cell*. **28**, 991-1001.

Epshtein, V., D. Dutta, J. Wade and E. Nudler (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature*. **463**, 245-9.

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

Estrem, S. T., T. Gaal, W. Ross and R. L. Gourse (1998) Identification of an UP element consensus sequence for bacterial promoters. *Proceedings of the National Academy of Sciences of the United States of America*. **95**, 9761-6.

Farnham, P. J. and T. Platt (1981) Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription *in vitro*. *Nucleic acids research*. **9**, 563-77.
Feklistov, A. and S. A. Darst (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*. **147**, 1257-69.
Ferguson, A. L., A. D. Hughes, U. Tufail, C. G. Baumann, D. J. Scott and J. G. Hoggett (2000) Interaction of sigma 70 with *Escherichia coli* RNA polymerase core enzyme studied by surface plasmon resonance. *FEBS letters*. **481**, 281-4.
Finkel, S. E. and R. C. Johnson (1992) The Fis protein: it's not just for DNA inversion anymore. *Molecular microbiology*. **6**, 3257-65.
Finlay, B. B. and S. Falkow (1997) Common themes in microbial pathogenicity revisited. *Microbiology and molecular biology reviews : MMBR*. **61**, 136-69.
Freter, R., H. Brickner, J. Fekete, M. M. Vickerman and K. E. Carey (1983) Survival and implantation of *Escherichia coli* in the intestinal tract. *Infection and immunity*. **39**, 686-703.
Gaal, T., W. Ross, E. E. Blatter, H. Tang, X. Jia, V. V. Krishnan, N. Assa-Munt, R. H. Ebright and R. L. Gourse (1996) DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes & development*. **10**, 16-26.
Galloway, J. L. and T. Platt (1988) Signals sufficient for rho-dependent transcription termination at trp t' span a region centered 60 base pairs upstream of the earliest 3' end point. *The Journal of biological chemistry*. **263**, 1761-7.
Geiselmann, J., T. D. Yager, S. C. Gill, P. Calmettes and P. H. von Hippel (1992) Physical properties of the *Escherichia coli* transcription termination factor rho. 1. Association states and geometry of the rho hexamer. *Biochemistry*. **31**, 111-21.
Geszvain, K., T. M. Gruber, R. A. Mooney, C. A. Gross and R. Landick (2004) A hydrophobic patch on the flap-tip helix of *E.coli* RNA polymerase mediates sigma(70) region 4 function. *Journal of molecular biology*. **343**, 569-87.
Ghosh, P., A. Ishihama and D. Chatterji (2001) *Escherichia coli* RNA polymerase subunit omega and its N-terminal domain bind full-length beta' to facilitate incorporation into the alpha2beta subassembly. *European journal of biochemistry / FEBS*. **268**, 4621-7.
Ghosh, P., C. Ramakrishnan and D. Chatterji (2003) Inter-subunit recognition and manifestation of segmental mobility in Escherichia coli RNA polymerase: a case study with omega-beta' interaction. *Biophysical chemistry*. **103**, 223-37.
Ghosh, T., D. Bose and X. Zhang (2010) Mechanisms for activating bacterial RNA polymerase. *FEMS microbiology reviews*. **34**, 611-27.
Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett and K. E. Nelson (2006) Metagenomic analysis of the human distal gut microbiome. *Science*. **312**, 1355-9.
Gopal, V. and D. Chatterji (1997) Mutations in the 1.1 subdomain of *Escherichia coli* sigma factor sigma70 and disruption of its overall structure. *European journal of biochemistry / FEBS*. **244**, 613-8.
Grainger, D. C., M. D. Goldberg, D. J. Lee and S. J. Busby (2008) Selective repression by Fis and H-NS at the *Escherichia coli dps* promoter. *Molecular microbiology*. **68**, 1366-77.
Grainger, D. C., D. Hurd, M. D. Goldberg and S. J. Busby (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic acids research*. **34**, 4642-52.
Grainger, D. C., D. Hurd, M. Harrison, J. Holdstock and S. J. Busby (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 17693-8.
Gregory, B. D., B. E. Nickels, S. A. Darst and A. Hochschild (2005) An altered-specificity DNA-binding mutant of *Escherichia coli* sigma70 facilitates the analysis of sigma70 function *in vivo*. *Molecular microbiology*. **56**, 1208-19.
Groisman, E. A. and H. Ochman (1996) Pathogenicity islands: bacterial evolution in quantum leaps. *Cell*. **87**, 791-4.
Gruber, T. M. and C. A. Gross (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology*. **57**, 441-66.
Gusarov, I. and E. Nudler (1999) The mechanism of intrinsic transcription termination. *Molecular cell*. **3**, 495-504.
Guthold, M., X. Zhu, C. Rivetti, G. Yang, N. H. Thomson, S. Kasas, H. G. Hansma, B. Smith, P. K. Hansma and C. Bustamante (1999) Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase. *Biophysical journal*. **77**, 2284-94.
Hacker, J., G. Blum-Oehler, I. Muhldorfer and H. Tschape (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Molecular microbiology*. **23**, 1089-97.
Hacker, J. and J. B. Kaper (2000) Pathogenicity islands and the evolution of microbes. *Annual review of microbiology*. **54**, 641-79.

Hanamura, A. and H. Aiba (1991) Molecular mechanism of negative autoregulation of *Escherichia coli crp* gene. *Nucleic acids research*. **19**, 4413-9.
Harrington S. M., M. C. Strauman, C. M. Abe, and J. P. Nataro. (2005) Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative *Escherichia coli*. *Cellular Microbiology*. **7**, 1565–1578.
Harrington, S. M., Dudley, E. G. and Nataro, J. P. (2006), Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiology Letters.* **254**, 12–18.
Haugen, S. P., W. Ross, M. Manrique and R. L. Gourse (2008) Fine structure of the promoter-sigma region 1.2 interaction. *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 3292-7.
Henderson, I.R., Hicks, S., Navarro-Garcia, F., Elias, W.P., Philips, A.D., and Nataro, J.P. (1999) Involvement of the enteroaggregative *Escherichia coli* plasmid-encoded toxin in causing human intestinal damage. *Infection and Immunity.* **67**, 5338–5344.
Herbert, K. M., A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick and S. M. Block (2006) Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell*. **125**, 1083-94.
Herold, S., H. Karch and H. Schmidt (2004) Shiga toxin-encoding bacteriophages--genomes in motion. *International journal of medical microbiology : IJMM*. **294**, 115-21.
Hong, S. H., X. Wang, H. F. O'Connor, M. J. Benedik and T. K. Wood (2012) Bacterial persistence increases as environmental fitness decreases. *Microbial biotechnology*.
Hook-Barnard, I. G. and D. M. Hinton (2007) Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene regulation and systems biology*. **1**, 275-93.
Hook-Barnard, I. G. and D. M. Hinton (2009) The promoter spacer influences transcription initiation via sigma70 region 1.1 of *Escherichia coli* RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 737-42.
Hsu, L. M., N. V. Vo, C. M. Kane and M. J. Chamberlin (2003) *In vitro* studies of transcript initiation by *Escherichia coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry*. **42**, 3777-86.
Hudault, S., J. Guignot and A. L. Servin (2001) *Escherichia coli* strains colonising the gastrointestinal tract protect germfree mice against Salmonella typhimurium infection. *Gut*. **49**, 47-55.
Igarashi, K. and A. Ishihama (1991) Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell*. **65**, 1015-22.
Ishihama, A. (1990) Molecular assembly and functional modulation of *Escherichia coli* RNA polymerase. *Advances in biophysics*. **26**, 19-31.
Ishihama, A. (1992) Role of the RNA polymerase alpha subunit in transcription activation. *Molecular microbiology*. **6**, 3283-8.
Ishihama, A. (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annual review of microbiology*. **54**, 499-518.
Jackson, S. A., I. R. Patel, T. Barnaba, J. E. LeClerc and T. A. Cebula (2011) Investigating the global genomic diversity of *Escherichia coli* using a multi-genome DNA microarray platform with novel gene prediction strategies. *BMC genomics*. **12**, 349.
Jin, D. J., R. R. Burgess, J. P. Richardson and C. A. Gross (1992) Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proceedings of the National Academy of Sciences of the United States of America*. **89**, 1453-7.
Johnson, J. P., R. A. Gatti, T. S. Sears and R. L. White (1986a) Inverted duplication of JH associated with chromosome 14 translocation and T-cell leukemia in ataxia-telangiectasia. *American journal of human genetics*. **39**, 787-96.
Johnson, R. C., M. F. Bruist and M. I. Simon (1986b) Host protein requirements for *in vitro* site-specific DNA inversion. *Cell*. **46**, 531-9.
Kahramanoglou, C., A. S. Seshasayee, A. I. Prieto, D. Ibberson, S. Schmidt, J. Zimmermann, V. Benes, G. M. Fraser and N. M. Luscombe (2011) Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic acids research*. **39**, 2073-91.
Kalyani, B. S., G. Muteeb, M. Z. Qayyum and R. Sen (2011) Interaction with the nascent RNA is a prerequisite for the recruitment of Rho to the transcription elongation complex *in vitro*. *Journal of molecular biology*. **413**, 548-60.
Kamada N, Kim Y, Sham H.P, Vallance B.A, Puente J.L, Martens E.C and Núñez G (2012) Regulated Virulence Controls the Ability of a Pathogen to Compete with the Gut Microbiota. *Science*. **336**, 1325-1329.
Kapanidis, A. N., E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss and R. H. Ebright (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*. **314**, 1144-7.
Kapanidis, A. N., E. Margeat, T. A. Laurence, S. Doose, S. O. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R. H. Ebright and S. Weiss (2005) Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis. *Molecular cell*. **20**, 347-56.

Kassavetis, G. A. and M. J. Chamberlin (1981) Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro. *The Journal of biological chemistry*. **256**, 2777-86.
Keseler, I. M., J. Collado-Vides, A. Santos-Zavaleta, M. Peralta-Gil, S. Gama-Castro, L. Muniz-Rascado, C. Bonavides-Martinez, S. Paley, M. Krummenacker, T. Altman, P. Kaipa, A. Spaulding, J. Pacheco, M. Latendresse, C. Fulcher, M. Sarker, A. G. Shearer, A. Mackie, I. Paulsen, R. P. Gunsalus and P. D. Karp (2011) EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic acids research*. **39**, D583-90.
Kim, D. E. and S. S. Patel (2001) The kinetic pathway of RNA binding to the *Escherichia coli* transcription termination factor Rho. *The Journal of biological chemistry*. **276**, 13902-10.
Kolb, A., D. Kotlarz, S. Kusano and A. Ishihama (1995) Selectivity of the *Escherichia coli* RNA polymerase E sigma 38 for overlapping promoters and ability to support CRP activation. *Nucleic acids research*. **23**, 819-26.
Komissarova, N. and M. Kashlev (1997) RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. *The Journal of biological chemistry*. **272**, 15329-38.
Kulbachinskiy, A. and A. Mustaev (2006) Region 3.2 of the sigma subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *The Journal of biological chemistry*. **281**, 18273-6.
Kurokawa, K., T. Itoh, T. Kuwahara, K. Oshima, H. Toh, A. Toyoda, H. Takami, H. Morita, V. K. Sharma, T. P. Srivastava, T. D. Taylor, H. Noguchi, H. Mori, Y. Ogura, D. S. Ehrlich, K. Itoh, T. Takagi, Y. Sakaki, T. Hayashi and M. Hattori (2007) Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA research : an international journal for rapid publication of reports on genes and genomes*. **14**, 169-81.
Lan, R. and P. R. Reeves (2002) *Escherichia coli* in disguise: molecular origins of Shigella. *Microbes and infection / Institut Pasteur*. **4**, 1125-32.
Landick, R. (2009) Transcriptional pausing without backtracking. *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 8797-8.
Laptenko, O., J. Lee, I. Lomakin and S. Borukhov (2003) Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *The EMBO journal*. **22**, 6322-34.
Lawson, C. L., D. Swigon, K. S. Murakami, S. A. Darst, H. M. Berman and R. H. Ebright (2004) Catabolite activator protein: DNA binding and transcription activation. *Current opinion in structural biology*. **14**, 10-20.
Leatham, M. P., S. Banerjee, S. M. Autieri, R. Mercado-Lubo, T. Conway and P. S. Cohen (2009) Precolonized human commensal *Escherichia coli* strains serve as a barrier to *E. coli* O157:H7 growth in the streptomycin-treated mouse intestine. *Infection and immunity*. **77**, 2876-86.
Lee, D. J. and S. J. Busby (2012) Repression by Cyclic AMP Receptor Protein at a Distance. *mBio*. **3**.
Lee, H. J., H. M. Lim and S. Adhya (2004) An unsubstituted C2 hydrogen of adenine is critical and sufficient at the -11 position of a promoter to signal base pair deformation. *The Journal of biological chemistry*. **279**, 16899-902.
Lickwar, C. R., F. Mueller, S. E. Hanlon, J. G. McNally and J. D. Lieb (2012) Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature*. **484**, 251-5.
Lim, H. M., H. J. Lee, S. Roy and S. Adhya (2001) A "master" in base unpairing during isomerization of a promoter upon RNA polymerase binding. *Proceedings of the National Academy of Sciences of the United States of America*. **98**, 14849-52.
Lin, S. H. and J. C. Lee (2003) Determinants of DNA bending in the DNA-cyclic AMP receptor protein complexes in *Escherichia coli*. *Biochemistry*. **42**, 4809-18.
Lisser, S. and H. Margalit (1993) Compilation of *E. coli* mRNA promoter sequences. *Nucleic acids research*. **21**, 1507-16.
Lodge, J., J. Fear, S. Busby, P. Gunasekaran and N. R. Kamini (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS microbiology letters*. **74**, 271-6.
Lonetto, M., M. Gribskov and C. A. Gross (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *Journal of bacteriology*. **174**, 3843-9.
Luna, G. M., C. Vignaroli, C. Rinaldi, A. Pusceddu, L. Nicoletti, M. Gabellini, R. Danovaro and F. Biavasco (2010) Extraintestinal *Escherichia coli* carrying virulence genes in coastal marine sediments. *Applied and environmental microbiology*. **76**, 5659-68.
Malhotra, A., E. Severinova and S. A. Darst (1996) Crystal structure of a sigma 70 subunit fragment from *E. coli* RNA polymerase. *Cell*. **87**, 127-36.
Maoileidigh, D. O., V. R. Tadigotla, E. Nudler and A. E. Ruckenstein (2011) A unified model of transcription elongation: what have we learned from single-molecule experiments? *Biophysical journal*. **100**, 1157-66.
Marr, M. T. and J. W. Roberts (2000) Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Molecular cell*. **6**, 1275-85.

Martinez-Antonio, A. and J. Collado-Vides (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Current opinion in microbiology*. **6**, 482-9.
Martinez-Trujillo, M., A. Sanchez-Trujillo, V. Ceja, F. Avila-Moreno, R. M. Bermudez-Cruz, D. Court and C. Montanez (2010) Sequences required for transcription termination at the intrinsic lambdatI terminator. *Canadian journal of microbiology*. **56**, 168-77.
Mathew, R. and D. Chatterji (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends in microbiology*. **14**, 450-5.
Matias Rodrigues, J. F. and A. Wagner (2009) Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS computational biology*. **5**, e1000613.
Matlock, D. L. and T. Heyduk (2000) Sequence determinants for the recognition of the fork junction DNA containing the -10 region of promoter DNA by *E. coli* RNA polymerase. *Biochemistry*. **39**, 12274-83.
McClure, W. R. (1985) Mechanism and control of transcription initiation in prokaryotes. *Annual review of biochemistry*. **54**, 171-204.
McDowell, J. C., J. W. Roberts, D. J. Jin and C. Gross (1994) Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. *Science*. **266**, 822-5.
McLeod, S. M., J. Xu, S. E. Cramton, T. Gaal, R. L. Gourse and R. C. Johnson (1999) Localization of amino acids required for Fis to function as a class II transcriptional activator at the RpoS-dependent *proP P2* promoter. *Journal of molecular biology*. **294**, 333-46.
Mekler, V., E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A. N. Kapanidis, W. Niu, Y. W. Ebright, R. Levy and R. H. Ebright (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*. **108**, 599-614.
Miroslavova, N. S. and S. J. Busby (2006) Investigations of the modular structure of bacterial promoters. *Biochemical Society symposium*, 1-10.
Mitchell, J. E., D. Zheng, S. J. Busby and S. D. Minchin (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic acids research*. **31**, 4689-95.
Mollegaard, N. E., P. B. Rasmussen, P. Valentin-Hansen and P. E. Nielsen (1993) Characterization of promoter recognition complexes formed by CRP and CytR for repression and by CRP and RNA polymerase for activation of transcription on the *Escherichia coli deoP2* promoter. *The Journal of biological chemistry*. **268**, 17471-7.
Mooney, R. A., S. E. Davis, J. M. Peters, J. L. Rowland, A. Z. Ansari and R. Landick (2009) Regulator trafficking on bacterial transcription units *in vivo*. *Molecular cell*. **33**, 97-108.
Morgan, W. D., D. G. Bear, B. L. Litchman and P. H. von Hippel (1985) RNA sequence and secondary structure requirements for rho-dependent transcription termination. *Nucleic acids research*. **13**, 3739-54.
Murakami, K. S. and S. A. Darst (2003) Bacterial RNA polymerases: the wholo story. *Current opinion in structural biology*. **13**, 31-9.
Muskhelishvili, G. and A. Travers (2003) Transcription factor as a topological homeostat. *Frontiers in bioscience : a journal and virtual library*. **8**, d279-85.
Naryshkin, N., A. Revyakin, Y. Kim, V. Mekler and R. H. Ebright (2000) Structural organization of the RNA polymerase-promoter open complex. *Cell*. **101**, 601-11.
Nasser, W., R. Schneider, A. Travers and G. Muskhelishvili (2001) CRP modulates fis transcription by alternate formation of activating and repressing nucleoprotein complexes. *The Journal of biological chemistry*. **276**, 17878-86.
Nataro J. P. and J. B. Kaper. (1998) Diarrheagenic *Escherichia coli. Clinical Microbiology Reviews*, **11**, 142cal .
Nataro J. P., T. Steiner, and R. L. Guerrant (1998) Enteroaggregative*Escherichia coli. Emerging Infectious Diseases*. **4,** 251ging.
Nataro, J. P. (2005) Enteroaggregative *Escherichia coli* pathogenesis. *Current opinion in gastroenterology*. **21**, 4-8.
Ochman, H. and I. B. Jones (2000) Evolutionary dynamics of full genome content in *Escherichia coli. The EMBO journal*. **19**, 6637-43.
Opalka, N., J. Brown, W. J. Lane, K. A. Twist, R. Landick, F. J. Asturias and S. A. Darst (2010) Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS biology*. **8**.
Opalka, N., M. Chlenov, P. Chacon, W. J. Rice, W. Wriggers and S. A. Darst (2003) Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell*. **114**, 335-45.
Paget, M. S. and J. D. Helmann (2003) The sigma70 family of sigma factors. *Genome biology*. **4**, 203.
Pan, C. Q., S. E. Finkel, S. E. Cramton, J. A. Feng, D. S. Sigman and R. C. Johnson (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *Journal of molecular biology*. **264**, 675-95.
Park, J. S., M. T. Marr and J. W. Roberts (2002) *E. coli* Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell*. **109**, 757-67.
Parkinson, G., C. Wilson, A. Gunasekera, Y. W. Ebright, R. H. Ebright and H. M. Berman (1996) Structure of the CAP-DNA complex at 2.5 angstroms resolution: a complete picture of the protein-DNA interface. *Journal of molecular biology*. **260**, 395-408.

Passner, J. M. and T. A. Steitz (1997) The structure of a CAP-DNA complex having two cAMP molecules bound to each monomer. *Proceedings of the National Academy of Sciences of the United States of America*. **94**, 2843-7.
Peters, J. M., A. D. Vangeloff and R. Landick (2011) Bacterial transcription terminators: the RNA 3'-end chronicles. *Journal of molecular biology*. **412**, 793-813.
Polayes, D. A., P. W. Rice, M. M. Garner and J. E. Dahlberg (1988) Cyclic AMP-cyclic AMP receptor protein as a repressor of transcription of the *spf* gene of *Escherichia coli*. *Journal of bacteriology*. **170**, 3110-4.
Popovych, N., S. R. Tzeng, M. Tonelli, R. H. Ebright and C. G. Kalodimos (2009) Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 6927-32.
Poulsen, L. K., F. Lan, C. S. Kristensen, P. Hobolth, S. Molin and K. A. Krogfelt (1994) Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. *Infection and immunity*. **62**, 5191-4.
Power, M. L., J. Littlefield-Wyer, D. M. Gordon, D. A. Veal and M. B. Slade (2005) Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. *Environmental microbiology*. **7**, 631-40.
Pupo, G. M., R. Lan and P. R. Reeves (2000) Multiple independent origins of Shigella clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proceedings of the National Academy of Sciences of the United States of America*. **97**, 10567-72.
Raffaelle, M., E. I. Kanin, J. Vogt, R. R. Burgess and A. Z. Ansari (2005) Holoenzyme switching and stochastic release of sigma factors from RNA polymerase *in vivo*. *Molecular cell*. **20**, 357-66.
Rasko, D. A., M. J. Rosovitz, G. S. Myers, E. F. Mongodin, W. F. Fricke, P. Gajer, J. Crabtree, M. Sebaihia, N. R. Thomson, R. Chaudhuri, I. R. Henderson, V. Sperandio and J. Ravel (2008) The pangenome structure of Escherichia coli: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of bacteriology*. **190**, 6881-93.
Rees, W. A., R. W. Keller, J. P. Vesenka, G. Yang and C. Bustamante (1993) Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science*. **260**, 1646-9.
Reppas, N. B., J. T. Wade, G. M. Church and K. Struhl (2006) The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Molecular cell*. **24**, 747-57.
Revyakin, A., C. Liu, R. H. Ebright and T. R. Strick (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*. **314**, 1139-43.
Reyes-Lamothe, R., X. Wang and D. Sherratt (2008) *Escherichia coli* and its chromosome. *Trends in microbiology*. **16**, 238-45.
Ring, B. Z., W. S. Yarnell and J. W. Roberts (1996) Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*. **86**, 485-93.
Rivetti, C., M. Guthold and C. Bustamante (1999) Wrapping of DNA around the *E.coli* RNA polymerase open promoter complex. *The EMBO journal*. **18**, 4464-75.
Roberts, J. and J. S. Park (2004) Mfd, the bacterial transcription repair coupling factor: translocation, repair and termination. *Current opinion in microbiology*. **7**, 120-5.
Roberts, J. W. (1969) Termination factor for RNA synthesis. *Nature*. **224**, 1168-74.
Roghanian, M., Y. Yuzenkova and N. Zenkin (2011) Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. *Nucleic acids research*. **39**, 4352-9.
Rosenberg, M. and D. Court (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annual review of genetics*. **13**, 319-53.
Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov and R. L. Gourse (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*. **262**, 1407-13.
Rossiter, A. E., D. F. Browning, D. L. Leyton, M. D. Johnson, R. E. Godfrey, C. A. Wardius, M. Desvaux, A. F. Cunningham, F. Ruiz-Perez, J. P. Nataro, S. J. Busby and I. R. Henderson (2011) Transcription of the plasmid-encoded toxin gene from enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Molecular microbiology*. **81**, 179-91.
Sakata-Sogawa, K. and N. Shimamoto (2004) RNA polymerase can track a DNA groove during promoter search. *Proceedings of the National Academy of Sciences of the United States of America*. **101**, 14731-5.
Sashni, R., S. Mishra, B. S. Kalayani and R. Sen (2012) Suppression of *in vivo* Rho-dependent transcription termination defects: evidence for kinetically controlled steps. *Microbiology*.
Savageau, M. A. (1983) Escherichia-Coli Habitats, Cell-Types, and Molecular Mechanisms of Gene-Control. *Am Nat*. **122**, 732-744.
Savery, N., V. Rhodius and S. Busby (1996) Protein-protein interactions during transcription activation: the case of the *Escherichia coli* cyclic AMP receptor protein. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. **351**, 543-50.

Savery, N. J., G. S. Lloyd, M. Kainz, T. Gaal, W. Ross, R. H. Ebright, R. L. Gourse and S. J. Busby (1998) Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *The EMBO journal*. **17**, 3439-47.

Schneider, R., R. Lurz, G. Luder, C. Tolksdorf, A. Travers and G. Muskhelishvili (2001) An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic acids research*. **29**, 5107-14.

Schneider, R., A. Travers, T. Kutateladze and G. Muskhelishvili (1999) A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Molecular microbiology*. **34**, 953-64.

Schroeder, L. A., T. J. Gries, R. M. Saecker, M. T. Record, Jr., M. E. Harris and P. L. DeHaseth (2009) Evidence for a tyrosine-adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of *Escherichia coli* RNA polymerase to promoter DNA. *Journal of molecular biology*. **385**, 339-49.

Schroeder, L. A., M. E. Karpen and P. L. deHaseth (2008) Threonine 429 of *Escherichia coli* sigma 70 is a key participant in promoter DNA melting by RNA polymerase. *Journal of molecular biology*. **376**, 153-65.

Sears, H. J. and I. Brownlee (1952) Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of bacteriology*. **63**, 47-57.

Sears, H. J., I. Brownlee and J. K. Uchiyama (1950) Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of bacteriology*. **59**, 293-301.

Sears, H. J., H. Janes, R. Saloum, I. Brownlee and L. F. Lamoreaux (1956) Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *Journal of bacteriology*. **71**, 370-2.

Selby, C. P. and A. Sancar (1993) Molecular mechanism of transcription-repair coupling. *Science*. **260**, 53-8.

Severinova, E., K. Severinov, D. Fenyo, M. Marr, E. N. Brody, J. W. Roberts, B. T. Chait and S. A. Darst (1996) Domain organization of the *Escherichia col*i RNA polymerase sigma 70 subunit. *Journal of molecular biology*. **263**, 637-47.

Shames, S. R. and B. B. Finlay (2010) Breaking the stereotype: virulence factor-mediated protection of host cells in bacterial pathogenesis. *PLoS pathogens*. **6**, e1001057.

Shao, Y., L. S. Feldman-Cohen and R. Osuna (2008) Functional characterization of the *Escherichia coli* Fis-DNA binding sequence. *Journal of molecular biology*. **376**, 771-85.

Sharp, J. A., S. K. Guterman and T. Platt (1986) The rho-115 mutation in transcription termination factor rho affects its primary polynucleotide binding site. *The Journal of biological chemistry*. **261**, 2524-8.

Sharp, M. M., C. L. Chan, C. Z. Lu, M. T. Marr, S. Nechaev, E. W. Merritt, K. Severinov, J. W. Roberts and C. A. Gross (1999) The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized. *Genes & development*. **13**, 3015-26.

Shimamoto, N., T. Kamigochi and H. Utiyama (1986) Release of the sigma subunit of *Escherichia coli* DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. *The Journal of biological chemistry*. **261**, 11859-65.

Siegele, D. A., J. C. Hu, W. A. Walter and C. A. Gross (1989) Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *Journal of molecular biology*. **206**, 591-603.

Singh, S. S., A. Typas, R. Hengge and D. C. Grainger (2011) *Escherichia coli* sigma senses sequence and conformation of the promoter spacer region. *Nucleic acids research*. **39**, 5109-18.

Skoko, D., D. Yoo, H. Bai, B. Schnurr, J. Yan, S. M. McLeod, J. F. Marko and R. C. Johnson (2006) Mechanism of chromosome compaction and looping by the *Escherichia coli* nucleoid protein Fis. *Journal of molecular biology*. **364**, 777-98.

Skordalakes, E. and J. M. Berger (2006) Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor. *Cell*. **127**, 553-64.

Solo-Gabriele, H. M., M. A. Wolfert, T. R. Desmarais and C. J. Palmer (2000) Sources of *Escherichia coli* in a coastal subtropical environment. *Applied and environmental microbiology*. **66**, 230-7.

Sosunova, E., V. Sosunov, M. Kozlov, V. Nikiforov, A. Goldfarb and A. Mustaev (2003) Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 15469-74.

Stecher, B., R. Robbiani, A. W. Walker, A. M. Westendorf, M. Barthel, M. Kremer, S. Chaffron, A. J. Macpherson, J. Buer, J. Parkhill, G. Dougan, C. von Mering and W. D. Hardt (2007) *Salmonella enterica serovar typhimurium* exploits inflammation to compete with the intestinal microbiota. *PLoS biology*. **5**, 2177-89.

Steitz, T. A. and J. A. Steitz (1993) A general two-metal-ion mechanism for catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America*. **90**, 6498-502.

Straney, D. C. and D. M. Crothers (1987) A stressed intermediate in the formation of stably initiated RNA chains at the *Escherichia coli* lac UV5 promoter. *Journal of molecular biology*. **193**, 267-78.

Sweeney, N. J., D. C. Laux and P. S. Cohen (1996) *Escherichia coli* F-18 and *E. coli* K-12 eda mutants do not colonize the streptomycin-treated mouse large intestine. *Infection and immunity*. **64**, 3504-11.

Sztiller-Sikorska, M., E. Heyduk and T. Heyduk (2011) Promoter spacer DNA plays an active role in integrating the functional consequences of RNA polymerase contacts with -10 and -35 promoter elements. *Biophysical chemistry*. **159**, 73-81.

Tang, G. Q., R. Roy, T. Ha and S. S. Patel (2008) Transcription initiation in a single-subunit RNA polymerase proceeds through DNA scrunching and rotation of the N-terminal subdomains. *Molecular cell*. **30**, 567-77.

Tang, H., K. Severinov, A. Goldfarb, D. Fenyo, B. Chait and R. H. Ebright (1994) Location, structure, and function of the target of a transcriptional activator protein. *Genes & development*. **8**, 3058-67.

Tebbutt, J., V. A. Rhodius, C. L. Webster and S. J. Busby (2002) Architectural requirements for optimal activation by tandem CRP molecules at a class I CRP-dependent promoter. *FEMS microbiology letters*. **210**, 55-60.

Telesnitsky, A. and M. J. Chamberlin (1989) Terminator-distal sequences determine the *in vitro* efficiency of the early terminators of bacteriophages T3 and T7. *Biochemistry*. **28**, 5210-8.

Touchon, M., C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguenec, M. Lescat, S. Mangenot, V. Martinez-Jehanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Medigue, E. P. Rocha and E. Denamur (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS genetics*. **5**, e1000344.

Toulme, F., C. Mosrin-Huaman, J. Sparkowski, A. Das, M. Leng and A. R. Rahmouni (2000) GreA and GreB proteins revive backtracked RNA polymerase *in vivo* by promoting transcript trimming. *The EMBO journal*. **19**, 6853-9.

Toulokhonov, I., J. Zhang, M. Palangat and R. Landick (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Molecular cell*. **27**, 406-19.

Travers, A. and G. Muskhelishvili (1998) DNA microloops and microdomains: a general mechanism for transcription activation by torsional transmission. *Journal of molecular biology*. **279**, 1027-43.

Travers, A. A. (1984) Conserved features of coordinately regulated *E. coli* promoters. *Nucleic acids research*. **12**, 2605-18.

Ussery, D., T. S. Larsen, K. T. Wilkes, C. Friis, P. Worning, A. Krogh and S. Brunak (2001) Genome organisation and chromatin structure in *Escherichia coli. Biochimie*. **83**, 201-12.

van Bloois, E., R. T. Winter, H. Kolmar and M. W. Fraaije (2011) Decorating microbes: surface display of proteins on *Escherichia coli. Trends in biotechnology*. **29**, 79-86.

Vassylyev, D. G., M. N. Vassylyeva, A. Perederina, T. H. Tahirov and I. Artsimovitch (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*. **448**, 157-62.

Vassylyeva, M. N., V. Svetlov, A. D. Dearborn, S. Klyuyev, I. Artsimovitch and D. G. Vassylyev (2007) The carboxy-terminal coiled-coil of the RNA polymerase beta'-subunit is the main binding site for Gre factors. *EMBO reports*. **8**, 1038-43.

Vendeville, A., D. Lariviere and E. Fourmentin (2011) An inventory of the bacterial macromolecular components and their spatial organization. *FEMS microbiology reviews*. **35**, 395-414.

Vignon, G., R. Kohler, E. Larquet, S. Giroux, M. C. Prevost, P. Roux and A. P. Pugsley (2003) Type IV-like pili formed by the type II secreton: specificity, composition, bundling, polar localization, and surface presentation of peptides. *Journal of bacteriology*. **185**, 3416-28.

Vo, N. V., L. M. Hsu, C. M. Kane and M. J. Chamberlin (2003) In vitro studies of transcript initiation by *Escherichia coli* RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. *Biochemistry*. **42**, 3798-811.

Wade, J. T., D. Castro Roa, D. C. Grainger, D. Hurd, S. J. Busby, K. Struhl and E. Nudler (2006) Extensive functional overlap between sigma factors in *Escherichia coli. Nature structural & molecular biology*. **13**, 806-14.

Wakimoto N., J. Nishi, J. Sheikh, Nataro J.P, Sarantuya J, Iwashitam M, Manago K, Tokufa K, Yoshinaga M and Kawonoy Y (2004) Quantitative biofilm assay using a microtiter plate to screen for enteroaggregative *Escherichia coli. American Journal of Tropical Medicine and Hygiene*. **71**, 687–690.

Waldor, M. K., H. Tschape and J. J. Mekalanos (1996) A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. *Journal of bacteriology*. **178**, 4157-65.

Walmacq, C., M. L. Kireeva, J. Irvin, Y. Nedialkov, L. Lubkowska, F. Malagon, J. N. Strathern and M. Kashlev (2009) Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. *The Journal of biological chemistry*. **284**, 19601-12.

Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli. Annual review of biochemistry*. **75**, 39-68.

Wirth, T., D. Falush, R. Lan, F. Colles, P. Mensa, L. H. Wieler, H. Karch, P. R. Reeves, M. C. Maiden, H. Ochman and M. Achtman (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Molecular microbiology*. **60**, 1136-51.

Wood, L. F., N. Y. Tszine and G. E. Christie (1997) Activation of P2 late transcription by P2 Ogr protein requires a discrete contact site on the C terminus of the alpha subunit of *Escherichia coli* RNA polymerase. *Journal of molecular biology*. **274**, 1-7.
Woods, R. J., J. E. Barrick, T. F. Cooper, U. Shrestha, M. R. Kauth and R. E. Lenski (2011) Second-order selection for evolvability in a large *Escherichia coli* population. *Science*. **331**, 1433-6.
Wosten, M. M. (1998) Eubacterial sigma-factors. *FEMS microbiology reviews*. **22**, 127-50.
Xie, P. (2008) A dynamic model for transcription elongation and sequence-dependent short pauses by RNA polymerase. *Bio Systems*. **93**, 199-210.
Yang, J., H. Nie, L. Chen, X. Zhang, F. Yang, X. Xu, Y. Zhu, J. Yu and Q. Jin (2007) Revisiting the molecular evolutionary history of *Shigella spp*. *Journal of molecular evolution*. **64**, 71-9.
Yuzenkova, Y., V. R. Tadigotla, K. Severinov and N. Zenkin (2011) A new basal promoter element recognized by RNA polymerase core enzyme. *The EMBO journal*. **30**, 3766-75.
Zhang, X. P. and R. H. Ebright (1990) Substitution of 2 base pairs (1 base pair per DNA half-site) within the *Escherichia coli lac* promoter DNA site for catabolite gene activator protein places the *lac* promoter in the FNR regulon. *The Journal of biological chemistry*. **265**, 12400-3.
Zhou, Y., S. Busby and R. H. Ebright (1993) Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell*. **73**, 375-9.
Zhou, Y., P. S. Pendergrast, A. Bell, R. Williams, S. Busby and R. H. Ebright (1994) The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *The EMBO journal*. **13**, 4549-57.