Bacterial Chromosome Organisation and Transcription

by

Laura Sellars

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

The University of Birmingham

April 2014

Abstract

The bacterial chromosome has to be condensed to fit inside the cell, forming a compact structure called the nucleoid, which is confined to a particular region of the cell without constriction by a membrane. Originally, the nucleoid was thought to be packed into the cell in a disordered way, unlike the highly organised chromatin of eukaryotic cells. More recently, the bacterial nucleoid has been shown to be far more structured than previously thought, with DNA present in topologically distinct loops, which are then arranged into macrodomains. Some of the proteins involved in structuring the *E. coli* chromosome are also known to have important roles in regulating transcription, and at least one transcription factor is known to cause distant DNA sites to cluster upon binding. These factors lead to the idea that chromosome structure could be affected by local gene expression.

To investigate the possible link between chromosome structure and gene expression, Fluorescent Reporter/Operator Systems (FROS) were used to study the positions of different inducible promoters, with and without induction. The FROS method was adapted to use a smaller insert, therefore causing less disruption to the chromosome structure. The transcription factor MalI was also developed as a novel FROS reporter. Of the five promoters studied, only *araFGH* showed any movement upon induction, moving away from the cell pole. In cells at the point of division, induction of the *araFGH* promoter caused segregation of sister chromatids adjacent to *araFGH*. These results suggest that induction of promoters can cause a change in local chromosome structure, although this is not seen at all promoters. The diffusion of fluorescently tagged transcription factors, used as reporters in FROS, can also be studied using super resolution microscopy methods.

Acknowledgements

Over the 4 years it has taken to complete this thesis I have had help and advice from many people for which I am very grateful. Firstly I would like to thank my supervisor, Prof Steve Busby for giving me this opportunity and for his support and guidance over the past few years. I would also like to thank Dr Dave Lee, Rita Godfrey and Dr Doug Browning for sharing their experience and wisdom with me and helping me with all my practical problems. Thank you to Dr Jack Bryant for introducing me to the lab and spending a lot of his time helping me and discussing our projects. Thanks also to Dr Maritoñi Sánchez Romero for her patience and assistance in training me on the microscope, without which this project would never have got started. A massive thank you to Dr Eugenio Sanchez-Moran and his lab for allowing me to use their microscope and helping me with my problems, as well as making me feel very welcome. For the PALM experiments in Chapter 5, I thank Dr Achillefs Kapanidis and Federico Garza de Leon for carrying out the experiments. I look forward to continuing this collaboration. I would like to thank Stephen Bevan for constructing the strains for studying MntR controlled promoters used in Chapter 4, and also for being an excellent student and causing me very little trouble.

I thank all members of the Busby and Cole labs, past and present, for their support and friendship. It has made the past 4 years a lot more pleasant than it otherwise would have been. I would like to thank my family, especially my parents and Lindsey for all of their encouragement through the whole of my life. I hope I've made you proud. Finally, I thank Andy, for getting me through this especially stressful year. Your turn next, then we will have a rest!

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Abbreviations

A	Adenine
AFM	Atomic force microscopy
APS	Ammonium persulfate
bp	base pairs
BSA	Bovine serum albumin
C	Cytosine
ChIP-Chip	Chromatin immunoprecipitation on microarray
CRP	cAMP
CTD	carboxy-terminal domain
dif	deletion induced filamentation
DNA	deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate (N = A, C, G or T)
dps	DNA-binding protein from starved cells
EMCCD	Electron multiplying charge coupled device
Fis	factor for inversion stimulation
FROS	Fluorescent Repressor/Operator System
G	Guanine
GFP	Green Fluorescent Protein
H-NS	Histone-like nucleoid structuring protein
HR	Homology Region
IHF	Integration host factor
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	kilobase
LB	Lennox broth
Mb	Megabase

MSD	Mean Square Displacement
NAP	Nucleoid associated protein
NSL	Non-structured Left Macrodomain
NSR	Non-structured Right Macrodomain
OD	Optical density
NTD	amino-terminal domain
NPG	2-nitrophenyl- β-D-galactopyranoside
PA	Photo-activatable
PALM	Photoactivated localisation microscopy
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
RNA	ribonucleic acid
RNAP	RNA polymerase
rpm	revolutions per minute
T	Thymine
TIRF	Total internal reflection fluorescence
UV	Ultraviolet
w/v	weight per volume

1. Introduction

1.1 Escherichia coli

Bacteria are prokaryotic organisms, usually between 1 and 20 μm in size, that are found in diverse environments almost anywhere on earth. Many bacteria have adapted to live inside other organisms, including humans, where they far outnumber human cells. Some bacteria enter other organisms as pathogens, but many are commensal and actually benefit the host. *Escherichia coli* is a member of the *Enterobacteriaceae* family of enteric bacteria. This is a large family of rod shaped Gram-negative bacteria, many of which are found in the gastrointestinal tracts of mammals. Although many are harmless, there are several pathogenic members of this family including *E. coli*, *Salmonella* and *Shigella* (Murray et al., 2009). Although *E. coli* can cause disease in humans, the majority of its strains are non-pathogenic and have in fact proved beneficial to humans as model organisms.

E. coli cells are about 2 μm in length and 0.5 μm in diameter and have a volume of 0.6-0.7 μm^3 (Murray *et al.*, 2009). The cell is surrounded by a cell envelope consisting of an inner membrane, periplasmic space containing peptidoglycan and outer membrane, as with all *Enterobacteriaceae*. The outer membrane largely consists of lipopolysaccharide (LPS), which is heat stable, on the outer leaflet and phospholipid on the inner leaflet. The inner membrane consists of a phospholipid bilayer. Both the inner and outer membrane have a variety of proteins inserted into them, with the proteins found in the outer membrane being of particular interest as they can be involved in virulence and drug resistance (Knowles *et al.*, 2009). Sometimes a capsule is also present on the surface of *E. coli* formed of polysaccharides. This helps to protect the cell from phagocytosis and prevents antibodies from binding to the outer membrane and is therefore a virulence factor (Whitfield, 2006). Many *E. coli* are motile and possess flagella, which also helps in colonisation for pathogenesis (Murray *et al.*, 2009).

Virulent strains of *E. coli* can affect the host in several ways with some causing gastroenteritis and others causing extraintestinal infections. There are five major groups of *E. coli* that cause gastroenteritis: enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroaggregative *E. coli* (EAEC), enterohemorrhagic *E. coli* (EHEC) and enteroinvasive *E. coli* (EIEC) (Croxen and Finlay, 2009, Kaper *et al.*, 2004). These groups vary between colonising the small and large intestine, symptoms and also in the way the infection spreads. This can be from contaminated water, uncooked food or, in some cases when the infectious dose is low, person-to-person contact. Most of these infections lead to diarrhoea and in some cases vomiting and fever (Croxen and Finlay, 2009). Of the extraintestinal infections caused by *E. coli,* urinary tract infections (UTI) are seen when uropathogenic *E. coli* (UPEC) from the colon infect the urethra and bladder and, in some cases, the kidneys (Russo and Johnson, 2003). Neonatal meningitis *E. coli* (NMEC), along with group B streptococci, is one of the leading causes of neonatal meningitis (Croxen and Finlay, 2009). Finally, infections in the urinary tract and intestines can lead to septicaemia in rarer cases. Most of these strains have become pathogenic due to acquisition of virulence factors often on plasmids, pathogenicity islands or bacteriophage DNA (Murray *et al.*, 2009).

Non-pathogenic *E. coli* strains are often used as model organisms for research, as results are often applicable to other bacteria. Lab cultivated strains often lose their virulence factors meaning they can be studied without any risk of infecting researchers. Strain K-12 is an example of this, and is one of the most widely used model organisms. Some of the characteristics of *E. coli* that make it desirable as a model organism are its fast growth and the comparative ease with which it can be cultured in laboratory conditions. The genome of *E. coli* K-12 strain MG1655 was sequenced and annotated in 1997 and revealed the genome to

contain 4,639,221 bp (Blattner *et al.*, 1997). The availability of sequence data has allowed for the development of many gene manipulation techniques which, again, made *E. coli* K-12 a favourable choice for molecular microbiology experiments.

Specific aspects of *E. coli* pathogenesis are often studied with a view to combating infection. However, more general aspects of how bacteria function are studied in *E. coli* K-12, for example promoter regulation, membrane structure, stress responses and chromosome structure. The findings can then be applied to other bacteria, particularly other *Enterobacteriaceae*, many of which are pathogens themselves.

1.2 Chromosome Structure in *Escherichia coli*

The *E. coli* K-12 MG1655 circular chromosome is around 4.6 Mb and contains 4489 genes. There is an origin of replication, *oriC*, located around position 3923 kb. Directly opposite this, at 1588 kb, is the *dif* site (deletion-induced filamentation) (Keseler *et al.*, 2011). This site, and the surrounding area of the chromosome, are responsible for the termination of replication and the separation of daughter chromosomes. The chromosome translates to a physical length of 1.6 mm in Watson-Crick structure, but a typical *E. coli* cell is just 2 μm in length, meaning the chromosome would need to be compacted by a factor of at least 1000 just to fit inside the cell (Holmes and Cozzarelli, 2000). The DNA in this tightly packed structure still needs to be accessible to proteins and other molecules involved in the processes of transcription, replication and DNA repair. As more research is being carried out into the structure of the bacterial chromosome, it is becoming clearer that a structure that can accommodate all of these needs must be very complex.

1.2.1 The Nucleoid

Unlike its eukaryotic counterpart a bacterial chromosome is not divided from the cytoplasm by a membrane, instead forming a compact structure called the nucleoid. Despite not being membrane bound, the nucleoid is confined to a particular area of the cell which has been suggested to be as low as 15% of the cellular volume (Reyes-Lamothe *et al.*, 2008b). However, while the bacterial chromosome has traditionally been represented as being unstructured and disorganised, evidence was found to suggest this was not the case over 30 years ago. Kavenoff and Bowen, who visualised the *E. coli* nucleoid using electron microscopy, saw that the nucleoid actually consisted of loops of supercoiled DNA arranged around a central region (Kavenoff and Bowen, 1976) (see figure 1.1). Further work showed that a single loop could be relaxed with a nick of the DNA while the rest remained supercoiled (Worcel and Burgi, 1972). This suggests that the loops of supercoiled DNA are independent of each other, and there is some kind of barrier to supercoiling, meaning the loops are topologically distinct. These loops are now known as domains. Estimates of the size and number of domains in the *E. coli* chromosome has varied greatly over the years but recent estimates believe there are around 500 domains with an average length of approximately 10 kb although there is a large range of sizes (Postow *et al.*, 2004). These domains are further organised into larger macrodomains (see section 1.2.2). Fluorescent microscopy and atomic force microscopy (AFM) have shown that the compaction of the nucleoid varies throughout cell cycle, with the nucleoid appearing more loosely packed in log phase than in stationary phase (Kim *et al.*, 2004). It has also been shown that the protein composition of the nucleoid changes depending on growth phase (Ali Azam *et al.*, 1999), indicating that the nucleoid is a dynamic structure, which can be modified to suit the needs of the cell. Presumably this allows

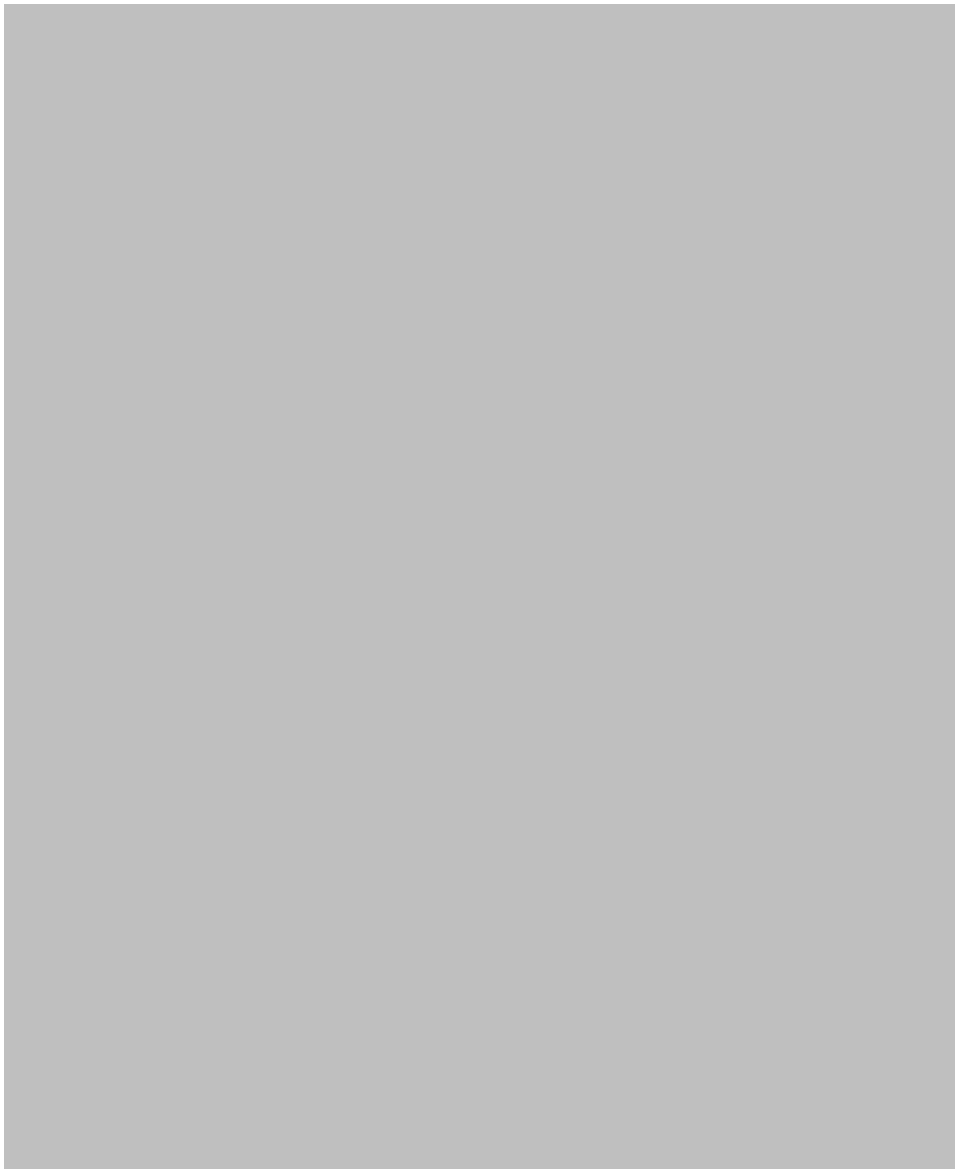


Figure 1.1: The *E. coli* nucleoid visualised by electron microscope

Membrane-free folded chromosomes from *E. coli* were visualised by electron microscopy, revealing loops of DNA spreading out from a core. (Kavenoff and Bowen 1976). Scale bar represents 1 μm.

bacteria to have structured chromosomes that are still accessible for processes like DNA replication and transcription.

By mass, only about 3% of the nucleoid is DNA (Reyes-Lamothe *et al.*, 2008b) suggesting there are other factors bound to DNA which may have a structural role. There exists a class of 12 proteins in *E. coli* with functions involving architecture of the chromosome, known as nucleoid associated proteins or NAPs. The functions of NAPs are diverse, with some responsible for DNA compaction and some for maintaining chromosome structure, and many have other functions, for example, transcription regulation. Some NAPs are found in abundant concentrations in *E. coli* and are well studied while others are found at lower concentrations and are less well understood (Azam and Ishihama, 1999). There are also NAPs whose concentrations fluctuate depending on the phase of the cell cycle. Although bacteria do not have histone proteins, several of the NAPs have been classed as "histone-like", despite having no structural similarity to eukaryotic histones (Salerno *et al.*, 2009). Many NAPs have an important role in chromosome compaction and are reviewed in section 1.2.3. Dps, a NAP known to have another function protecting cells from oxidative damage (Almiron *et al.*, 1992), is very abundant during stationary phase, making it the main protein component of the nucleoid, in non-growing cells. Dps causes the nucleoid to become a much more compact structure (Kim *et al.*, 2004). It has also been shown that in stress situations, Dps in *E. coli* can bind to DNA and form a crystalline structure, protecting the DNA from damage (Wolf *et al.*, 1999). Another group of proteins called structural maintenance of chromosome proteins (SMC), homologous to SMC proteins found in eukaryotes, play a role in chromosome structure in bacteria. This is the MukBEF complex in *E. coli* (Soppa, 2001). Although many proteins are involved in maintaining chromosome structure in bacteria, no single protein

seems be essential or responsible for the whole process, suggesting that the bacterial nucleoid is a combined effect of many proteins.

It has been observed that, in bacteria, ribosomes are excluded from the nucleoid (Hobot *et al.*, 1985, Li and Jensen, 2009). It has long been thought that bacterial transcription and translation are linked (Miller *et al.*, 1970), unlike eukaryotic cells where the nuclear membrane separates the two processes in space and time. If transcription and translation are spatially and temporally linked, this suggests that RNA polymerase and the gene being transcribed need to be in the same place as the ribosome, at the edge of the nucleoid. Fluorescently tagged RNAP has been shown to be predominantly in the nucleoid, sometimes forming foci in rapidly growing cells (Lewis *et al.*, 2000, Cabrera and Jin, 2003). However, much of the RNA synthesis in bacteria is making stable RNA that will not be translated, for example rRNA and tRNA. 60% of transcription in a cell is at the rRNA genes (Schneider *et al.*, 2003). The absence of ribosomes from the nucleoid suggests that chromosome structure may need to be rearranged for transcription and translation to take place.

1.2.2 Macrodomains

As well as the chromosome being organised into the domains described above, the domains themselves are then further arranged into larger structures known as macrodomains. Valens *et al.* 2004 first described this organisation and identified six macrodomains, defined by the ability of sites across the chromosome to interact with each other. They found that sites within the same macrodomain could readily interact with each other, but interacted very little with sites in other macrodomains. Two large macrodomains, of around 1 Mb, were identified around the replication origin and replication terminus named Ori and Ter respectively. Either

side of the Ter macrodomains are two regions of 0.7 Mb called the Right and Left domains (see figure 1.2). The final two macrodomains are found flanking the Ori domain and are less structured than the others. These are called Non-structured Left (NSL) and Non-structured Right (NSR) (Valens *et al.*, 2004). Loci in these macrodomains were able to interact with loci in other domains across long distances. To visualise the positions of the macrodomains in the cell, *parS* sites at various locations of interest on the chromosome, and a ParB::GFP fusion protein was used to follow these locations using time-lapse fluorescent microscopy.These experiments revealed "home" positions, where sites within the individual macrodomains occupy certain areas in the cell. This area is two to four times smaller for the structured macrodomains than for the two non-structured domains. This study also revealed some information about the dynamics of the macrodomains. Loci in the non-structured regions were shown to be highly mobile whereas loci in the structured domains were also mobile but restrained within a specific region of the cell. Sites in the non-structured macrodomains are more likely to be able to interact with sites in other macrodomains, presumably because they are less restrained (Espéli *et al.*, 2008). Little is known about the factors responsible for structuring most of the macrodomains, but it is known that a protein called MatP has a role in structuring the Ter macrodomain. MatP binds to a 13 bp motif called *matS*. There are 23 *matS* sites in the Ter macrodomain, compared to just four other sites found on the rest of the *E. coli* chromosome, with 2 each in the Left and Right macrodomains. A MatP::GFP fusion showed that MatP accumulates in the cell as a focus that colocalises with the Ter domain, and MatP inactivation resulted in a less compaction of DNA. (Mercier *et al.*, 2008). MatP binds to *matS* as a dimer but can interact by a coiled-coil domain to tetramerise, linking distant *matS* sites (Dupaigne *et al.*, 2012). All of these results suggested that MatP could be the main factor in organising the Ter macrodomain.

Figure 1.2: Positions of the macrodomains on the *E. coli* chromosome

The positions of the 6 macrodomains of the *E. coli* chromosome are shown as well as the *oriC* and *dif* as defined by Kesler *et al.,* 2011. The macrodomain containing *oriC* is called Ori which is flanked by two non-structured domains, Non-Structured Left (NSL) and Non-Structured Right (NSR). The macrodomain containing the *dif* is called Ter and is flanked by the Left and Right macrodomains (Valens *et al.,* 2004).

1.2.3 Compaction of DNA

As previously described, the *E. coli* chromosome needs to be compacted by 1000 to 10,000 fold to fit inside the cell. This compaction is comparable to that of a chromosome in a eukaryotic cell, but in *E. coli* is achieved without histone proteins (Sherratt, 2003). The mechanisms of compaction also need to allow the chromosome to remain functional, and have been the subject of many studies.

One of the factors important in compaction of DNA is supercoiling. In the Watson-Crick structure of DNA, there is a helical turn every 10.6 bp. If turns are introduced to or removed from the DNA it will become over- or under-wound, or supercoiled. The DNA of the nucleoid is anchored by various structural proteins so the supercoils cannot diffuse and the DNA twists to accommodate the extra turns. These twists mean that supercoiled DNA is much more compact form of DNA than Watson-Crick structure, and so supercoiling is a useful tool for packing DNA into the nucleoid. It is thought that approximately 4% of the helical turns in the *E. coli* chromosome are removed, compared to the relaxed structure of DNA (Trun and Marko, 1998). In *E. coli*, topoisomerases cause DNA to be underwound resulting in the formation of supercoils (Thanbichler and Shapiro, 2006). These enzymes function by cutting DNA, rotating it and religating (Liu *et al.*, 1980). Cell processes, such as DNA replication, that involve unwinding of the DNA helix cause the superhelicity of the local area to change. These processes often involve large protein complexes which, instead of following the helical twist of the DNA, cause the DNA to rotate. This results in a region of positive supercoiling ahead of the complex and negative supercoiling behind. Topoisomerases I and II (Topo I and gyrase) are responsible for maintaining the level of supercoiling in the *E. coli* chromosome with Topoisomerase IV (Topo IV) also playing a role, with a balance of the three needed for normal supercoiling (Zechiedrich *et al.*, 2000). Gyrase is essential in *E. coli* (Levine *et al.*,

1998) and is the only topoisomerase able to insert negative supercoils. Other topoisomerases help to maintain the overall underwound state by removing positive supercoils. Cells lacking Topo I can grow but only in the presence of gyrase mutations (Dinardo *et al.*, 1982). Topo I and Topo IV both work by removing positive supercoils, with Topo I having a much faster rate but only on very negatively supercoiled DNA. Topo IV works slower but is able to completely relax DNA (Zechiedrich *et al.*, 2000). Supercoiling also alters the twist of the DNA to create a more open complex required for many processes including replication, transcription and DNA repair (Peter *et al.*, 2004). It has also been suggested that these processes, in turn, may have a role in compaction of DNA as they involve many proteins binding to the DNA (Bendich, 2001). Although supercoiling has an essential role in the compaction of DNA, it alone cannot compact the DNA to the extent required, therefore, there must be other factors involved (Reyes-Lamothe *et al.*, 2008b).

Another important factor in the compaction of the nucleoid are the NAPs. Several of these have functions that involve bending or bridging of the DNA which could be used to structure the nucleoid. One well studied NAP is Fis (factor for inversion stimulation), a homodimeric protein which causes bending of DNA between 50º and 90º when it binds (Pan *et al.*, 1996). Its concentration in the cell fluctuates, peaking at 60,000 during exponential growth and reducing to less than 100 molecules per cell in stationary phase (Ali Azam *et al.*, 1999). Fis is known to preferentially bind AT rich sequences and a 15 bp consensus sequence for a binding, GNTCAAATTTTGANC, was identified (Finkel and Johnson, 1992). However, studies using this consensus sequence to predict Fis binding sites generate varying results. More recently, an improved consensus sequence has been identified using Fis binding sites that had been confirmed using ChIP-chip (Cho *et al.*, 2008). Electron microscopy and AFM

have shown that Fis is able to cause DNA branching, condensing and organising the nucleoid (Schneider *et al.*, 2001). Another NAP known to compact DNA by bending is IHF (Integration host factor), a heterodimeric protein that wraps DNA around itself, introducing a bend of around 160º (Rice *et al.*, 1996). IHF is most abundant in early stationary phase, where it is thought to number around 55,000 monomers per cell (Ali Azam *et al.*, 1999). A consensus sequence has been identified for IHF, but the surrounding sequence is also important (Friedman, 1988). Another way that NAPs can contribute to chromosome condensation is by bridging DNA. One protein that works this way is H-NS (Histone-like nucleoid structuring protein), which preferentially binds curved DNA, but also has a consensus sequence (Lang *et al.*, 2007). The C-terminal domain of H-NS is the DNA binding domain and the N-terminus is responsible for dimerisation, allowing the possibility of higher-order structures of H-NS (Esposito *et al.*, 2002). As with other NAPs, AFM has been used to show the effect of binding on DNA. It was observed that H-NS could bring distant regions of DNA close together by bridging, and therefore condensing the DNA (Dame *et al.*, 2000). A *hns* mutation results in cells with a reduced number of origins. Evidence suggests that this reduction is an indirect effect of H-NS, indicating that it may have an important role in chromosome replication and possibly structure (Atlung and Hansen, 2002). However, it has been shown that deletion of some of the NAPs does not result in a less compact nucleoid, suggesting that their role may not be as important as first thought (Zimmerman, 2006)

DNA is also compacted by an effect called macromolecular crowding. This effect was first noticed in 1971 when it was observed that DNA collapsed into a compact structure when placed in a solution with high concentrations of polymers and salts (Lerman, 1971). The concentration of total protein and RNA in the *E. coli* cytoplasm is 340 mg/ml, which means it

is packed with proteins and macromolecules (Zimmerman and Trach, 1991). This high concentration could apply force on the nucleoid and contribute to compaction. It may also make it more energetically favourable for other factors to bind the DNA which may in turn stabilise the structure of the nucleoid (Murphy and Zimmerman, 1995). Although molecular crowding, supercoiling and NAPs are all important factors involved in the compaction of the chromosome none of them alone are sufficient to achieve the level of compaction required. Instead, it is likely that the combination of these factors, along with others, are responsible for chromosome compaction. Throughout the cell cycle they compact the nucleoid in a way that also allows it to be dynamic and fit with other processes, such as DNA replication, repair and transcription.

1.3 DNA replication and chromosome segregation in *Escherichia coli*

1.3.1 DNA replication

E. coli has a circular chromosome, which is replicated bidirectionally from a single origin of replication, *oriC*. The two halves of the chromosome are referred to as replichores, with one either side of *oriC* (Wang *et al*., 2006). Each replichore is replicated by one replisome, a complex of replication machinery. Once the whole chromosome has been replicated, the cell will divide to make two daughter cells. As with many bacteria, multiple cycles of DNA replication can be underway in the same cell, due to the fact that replication of an entire chromosome takes longer than the time for a cell to grow and divide. If growing conditions are good, replication will be initiated again, before the previous cycle has finished, at both copies of the origin of replication. This results in there being 4 copies of the region surrounding the origin of replication, and when the cell divides to make two daughter cells, chromosome replication is already well underway in both.

The system for initiating replication in *E. coli* is not entirely understood, but it involves the replication initiation protein, DnaA, being activated by ATP. DnaA-ATP binds *oriC* in a stepwise manner which opens the DNA helix. This allows DnaC, the helicase loader, to load DnaB, the replicative helicase, leading to the formation of a complex of proteins, called the replisome. One replisome, consisting of around 13 proteins, associates with one strand of DNA, making a replication fork, which continues around the circular chromosome until it reaches the terminus region, approximately opposite *oriC* on a circular chromosome map. Unlike some other bacteria, *E. coli* has one polymerase which replicates both the leading and lagging strand (Dervyn *et al.*, 2001).

Replication initiation is regulated in several ways to ensure that replication is not initiated prematurely and to coordinate DNA replication with cell division. The concentration of DnaA-ATP is kept at a low level by HdaA, which hydrolyses the ATP to make DnaA-ADP. HdaA is bound to the β sliding clamp in the replisome (Reyes-Lamothe *et al.*, 2012, Katayama *et al.*, 2010). Replication can be initiated where there are low levels of replication, where the sliding clamp is far away, allowing DnaA-ATP to accumulate. This is combined with the action of little understood mechanisms that generate DnaA-ATP. The build up of DnaA-ATP causes replication to be initiated at all *oriC* sequences in the cell. DnaA-ATP is not hydrolysed during initiation of replication but a second mechanism prevents multiple initiation events. SeqA sequesters hemimethylated GATC sites to prevent binding by Dam methylase or DnaA. It is known that there is a cluster of GATC sites at *oriC* in *E. coli*, which will be hemimethylated immediately after replication (Henaut *et al.*, 1996). Fluorescently tagged SeqA has been shown to form a focus, presumably due to association with newly replicated DNA (Brendler *et al.*, 2000).

Every helical turn in the chromosome needs to be removed during replication. In the 4.6 Mb *E. coli* chromosome this is 4.6 x 10^5 links, and any links not removed will result in interlinked chromosomes called catenanes (Ip *et al.*, 2003). DNA gyrase, a member of the type II toposiomerase family, is responsible for removing these supercoils to maintain the overall balance of supercoiling in the nucleoid. The movement of the replisome along the DNA helix and unwinding of the helix for replication causes an accumulation of positive supercoiling ahead of the replication fork which cannot diffuse as the chromosome is a closed circular structure. This can cause rotation of DNA at the replication fork leading to the two newly replicated chromosomes being intertwined, known as precatenanes (Bermejo *et al.*, 2008). These structures have to be removed to allow the sister chromatids to be segregated. If they are not removed before the replication process is complete, the new chromosomes will be catenated. Topoisomerase IV (Topo IV) is responsible for both the removal of precatenanes and the separation of catenated chromosomes. Topo IV has two subunits, ParC, which interacts with the replisome, and ParE, which associates with FtsK, a component of the septum. Topoisomerase III (Topo III) is also capable of decatenating chromosomes (Nurse *et al.*, 2003).

Dividing cells were observed to have replicated sister chromatids in opposite halves of the cell, where they had, presumably, been segregated. It was originally thought that both replisomes were anchored at mid-cell, between the sister chromatids, and DNA to be replicated was pulled through the replicative machinery and pushed out of the other side. This idea, known as “replication factories”, was disproved by Reyes-Lamonthe *et al*. by following fluorescently tagged replication proteins. This revealed that although the replisomes assembled at *oriC,* usually around mid-cell, the two replication forks separated around 5

minutes after initation and moved independently along the DNA in opposite halves of the cell. Towards termination, the replication forks were then seen to move back towards mid-cell (Reyes-Lamothe *et al.*, 2008a). Fluorescent Reporter/Operator System (FROS, see section 3.1.1) experiments also showed that the newly replicated sister chromatids were localised to opposite cell halves, so one copy of the chromosome was in each future daughter cell.

Termination of replication occurs in the terminus region of the chromosome. As previously mentioned, high levels of positive supercoiling are found ahead of the replication fork, which are usually removed by the activity of topoisomerases. However, as the two replication forks near each other and the terminus region, the accessibility of the DNA by gyrase is restricted and the positive supercoiling accumulates. This accumulation of supercoiling can cause the replication forks to stall (Sherratt, 2003). Progress of replication forks is stopped by a DNA binding protein, Tus, which blocks the replicative helicase, DnaB. Tus binds to 10 *Ter* sites on the *E. coli* chromosome, including two primary sites, *TerA* and *TerB*. *TerA* and *TerB* are 350 kb apart, and create what is known as a "replication fork trap" (Hill *et al.*, 1988a). The sites are 23 bp long and are not palindromic. The *tus* gene is immediately downstream of the *TerB* site, which overlaps the *tus* ribosome binding site and -10 region. Tus has been shown to self-regulate by binding to *TerB* and repressing *tus* expression (Natarajan *et al.*, 1991). The 10 *Ter* sites are found in two sets of 5, with all 5 sites in the same orientation. The replication fork is able to pass the sites only in one direction, so they can enter the replication fork trap but cannot progress beyond the other side (Hill *et al.*, 1988b). When a replication fork approaches a *Ter* site from the non-permissive direction, Tus interacts with DnaB, the replicative helicase, and forms a stalled complex (Mulcair *et al.*, 2006, Mulugu *et al.*, 2001). When both replication forks are stalled at *Ter* sites, DnaB separates a few nucleotides of

duplex DNA, dissociating Tus and allowing the replication forks to continue and replicate the final section of the chromosome (Mulcair *et al.*, 2006).

Studies on DNA replication observed that, in cells with a single chromosome, the left and right replichores are located in opposite halves of the cell, with *oriC* in the centre. Replication begins at *oriC*, in the centre of the cell, and then the replication fork moves outwards towards opposite cell poles. As replication progresses, replicated DNA is segregated to opposite cell halves and the replication forks return to the centre of the cell, where the remaining unreplicated DNA is located. When the chromosome has been replicated, each half of the cell has one copy of each replichore, and in the majority of cells, the arrangement along the length of the cell is left, *oriC*, right, left, *oriC*, right. This asymmetric organisation suggests that the segregation of sister chromatids is not random and may influence chromosome structure further down the line (Wang *et al.*, 2006, Nielsen *et al.*, 2006) (see figure 1.3).

1.3.2 Chromosome Segregation

E. coli replicates by dividing at mid-cell to give two daughter cells, each with a copy of the chromosome. For this to happen, the replicated chromosome first needs to be segregated. Unlike eukaryotic cells, where chromosomes are completely replicated before being separated by mitotic spindles, the segregation of bacterial chromosomes occurs alongside DNA replication. Some bacteria, such as *Bacillus subtilis*, encode a ParA/B system, similar to that used to segregate plasmids, to segregate chromosomes. This system is not present in eubacteria, including *E. coli*, suggesting other mechanisms must be present (Autret and Errington, 2003). It seems that many different processes contribute to the segregation of newly replicated chromosomes prior to cell division but the whole picture is not clear yet. It is

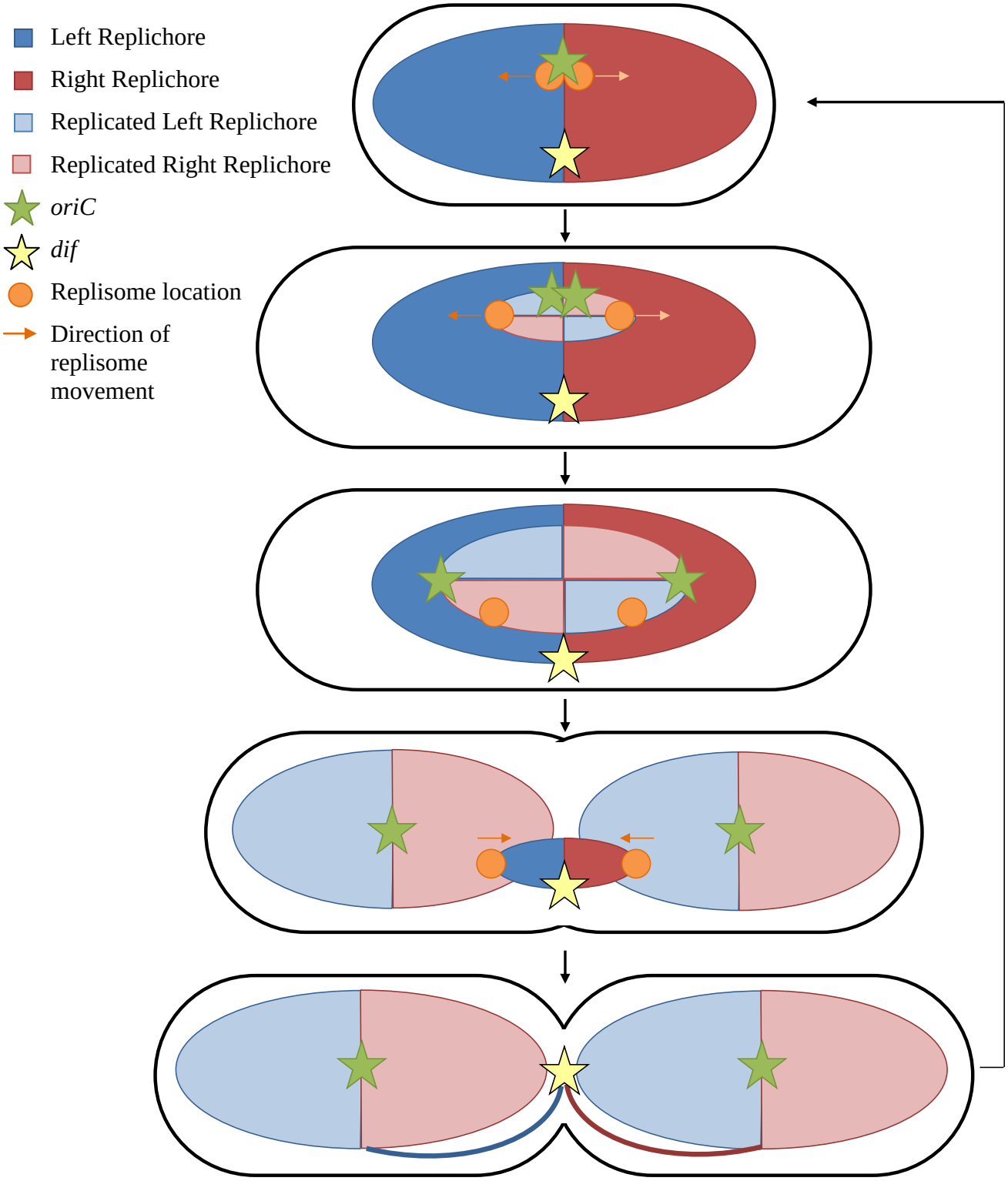


Figure 1.3: Diagram of the segregation of the left and right replichores during chromosome replication

A model of a bacterial cell throughout the cell cycle showing the segregation of the left and right replichores into future daughter cells. Also shown are the approximate positions of *oriC*, *dif* and the replisome.Arrows show the approximate direction of movement of replisomes, away from the centre of the cell into opposite cell halves and then back towards the cell centre again. Figure adapted from Nielsen *et al.*, 2006 and Reyes-Lamonthe *et al.*, 2008b.

known that early in the replication process, the sister chromatids begin to segregate to opposite halves of the cell, and the rest of the chromosome follows at various rates after it is replicated. Newly replicated Ori macrodomains localise to positions of one quarter and three quarters relative to cell length. It has been suggested that the reason for this positioning is that after cell division, the one quarter and three quarter positions will become mid-cell in each of the new daughter cells. If *oriC* is at this position it is in the correct place for replication to initiate and, later in the cell cycle, the septum to form (Sherratt, 2003). The rate at which regions of the chromosome segregate seems to be related to their macrodomain. In general, the non-structured macrodomains segregate quicker than structured macrodomains (Espéli *et al.*, 2008). After DNA is replicated there may be a cohesion step, of varying length depending on chromosomal position, before segregation.

The Ori macrodomain is first to be replicated but it is thought the region immediately adjacent to *oriC* remains colocalised for around 14 minutes (Bates and Kleckner, 2005), although estimates of periods of chromosome cohesion vary between studies. The Ori macrodomain is flanked by two non-structured macrodomains. Once the sister chromatid segregation begins in the Ori macrodomain, the two non-structured domains follow in quick succession. The Left and Right macrodomains have a window of chromosome cohesion before separating in the same order as replication. The Ter macrodomain has the longest period of chromosomal cohesion, with many estimates being over 60 minutes (Espéli *et al.*, 2008, Bates and Kleckner, 2005). It has been suggested that the MatP, responsible for organisation of the Ter macrodomain, could also be involved in the delay in segregation. MatP has been shown to interact with ZapB, a cell division factor associated with the FtsZ ring, which anchors the Ter macrodomain at mid-cell. Espéli *et al.* demonstrated that MatP associated to ZapB could hold

together distance sites in the Ter macrodomain, even when these are not physically linked, in a linear chromosome. They go on to suggest that segregation of the Ter macrodomain could be caused by ZapB dissociating from the FtsZ ring (Espéli *et al.*, 2012). Two sites have also been identified that flank the Ter domain and are thought to delay segregation. The 12 bp sites are called *tidL* and *tidR*, located just outside the Ter, 50 kb into the Left macrodomain and 130 kb into the right macrodomain respectively. They are bound by a protein called YfbV and, when bound, are thought to insulate the rest of the chromosome from the effects of proteins such as MatP, which restrict the segregation of Ter (Thiel *et al.*, 2012).

After the newly replicated DNA has been segregated to its new "home" position in the cell, the structural features that were present in the chromosome before replication need to be restored. The structural maintenance of chromosomes (SMC) family of proteins are involved in condensing the sister chromatids after segregation (Sherratt *et al.*, 2001). One member of this family, the MukBEF complex, has an important role in condensing sister chromatids, with MukB being required for normal chromosome partitioning (Niki *et al.*, 1991). In the absence of MukB, sister chromatids do not separate into opposite halves of the cell, suggesting that MukB may be responsible for this arrangement in wild type cells (Danilova *et al.*, 2007). GFP tagged MukB was shown to form at least one focus within the nucleoid, dependent on the presence of MukE and MukF, mostly at the one quarter and three quarter positions of the cell (Ohsumi *et al.*, 2001). These foci were later shown to colocalise with *oriC* in the majority of cells (Danilova *et al.*, 2007).

As well as segregating new chromosomes, it is important that the cell divides at mid-cell to give daughter cells of approximately equal size, each containing a full chromosome. An

important factor in this process is FtsZ, a tubulin-like protein that forms a ring shaped septum (Löwe and Amos, 1998). Formation of the FtsZ ring is then followed by recruitment of other cell division machinery, so correct positioning and timing of FtsZ ring formation is essential. If this is assembled too soon it can lead to the cell dividing when there is still unsegregated DNA at mid-cell and the chromosome is cut. Formation of the FtsZ ring near a cell pole results in small, nucleoid-free cells. There are two processes known to regulate FtsZ ring formation in *E. coli*, the Min system and nucleoid occlusion. The Min system, consisting of MinC, MinD and MinE, prevents the FtsZ ring from forming too close to cell poles. MinD oscillates from one cell pole to the other, changing pole approximately every 40 seconds (Huang *et al.*, 2003), indicating the position of mid-cell (Raskin and de Boer, 1999). MinD recruits MinC, which also oscillates. MinC has been shown to interact with FtsZ and prevent polymerisation, therefore preventing the formation of the FtsZ ring (Hu *et al.*, 1999). MinC is found at the highest concentration close to cell poles, so prevents the FtsZ ring from forming at this undesirable location. MinE forms a ring at the edge of the area occupied by MinD, which will be at around mid-cell. MinE causes ATP bound by MinD to be hydrolysed, which leads to MinD:ADP returning to the cytoplasm (Huang *et al.*, 2003). The MinE ring inhibits MinC function and allows FtsZ filaments to form the FtsZ ring. The MinE ring is independent of FtsZ, but ensures that its formation is in the correct place (Raskin and de Boer, 1999).

A second process called nucleoid occlusion works with the Min system by preventing the formation of the divisome in the immediate surroundings of the nucleoid. Fluorescent tagging has shown that arcs of FtsZ begin to form between segregated nucleoids, not over the nucleoid, indicating the presence of a nucleoid occlusion mechanism in *E. coli* (Zaritsky and Woldringh, 2003). A DNA binding protein called SlmA, which binds to specific DNA

sequences, has been shown to colocalise with the *E. coli* nucleoid, and was required for nucleoid occlusion (Bernhardt and De Boer, 2005). To play this role, SlmA binds, and prevents FtsZ polymerisation, an effect that is amplified when SlmA is bound to DNA (Cho *et al.*, 2011). The 50 SlmA DNA sites on the *E. coli* chromosome are not distributed evenly across the macrodomains, with a lower frequency in the Left and Right macrodomains, and they are completely absent from the Ter macrodomain (Dame *et al.*, 2011). It is thought that free SlmA would not be able to inhibit FtsZ polymerisation at the concentrations present in the cell, so only DNA-bound SlmA will have an effect (Cho *et al.*, 2011). This suggests that SlmA prevents FtsZ ring formation over the Ori and non-structured macrodomains but allows it to form around the Ter domain towards the end of chromosome replication, which would locate it at mid-cell (Männik *et al.*, 2012). A *slmA* mutation is not lethal in cells where the Min system is still functional, but a double mutant is lethal, due to incorrect positioning of the FtsZ ring (Bernhardt and De Boer, 2005).

The formation of the FtsZ ring leads to the recruitment of other cell division proteins, including FtsK, which has an important role in coordinating DNA segregation and cell division. FtsK is a DNA translocase which is capable of travelling at up to 5kb per second *in vitro* (Pease et al., 2005). *In vivo* it is attached to the septum, so is immobilised, and hence it is DNA that moves. The N terminus of FtsK is essential for cell viability and is involved in septum formation and closure (Draper *et al.*, 1998). The C terminus of FtsK is the DNA translocase and has 3 domains, α, β and γ (Sivanathan *et al.*, 2006). The DNA translocase activity of FtsK has several important roles in the facilitating the segregation of newly replicated chromosomes. Firstly Topo IV interacts with FtsK to decatenate chromosomes, probably with an interaction between ParC of Topo IV and domain 3 of FtsK (Espéli *et al.*,

2003). Before the cell divides, FtsK moves any DNA in the cell division site to one of the future daughter cells (Thanbichler and Shapiro, 2006).

The DNA translocase activity of FtsK also helps solve another issue that affects the segregation of sister chromatids, the presence of chromosome dimers. These are caused by homologous recombination between the sister chromatids. The period of sister chromatid cohesion, explained previously, may be to allow homologous recombination to correct any errors made during replication. Chromosome dimers, observed in about 14% of cells, need to be resolved before the final chromosome segregation and cell division (Pérals *et al.*, 2001). In *E. coli*, tyrosine recombinases XerCD resolve chromosome dimers at a specific site called *dif*, located approximately opposite *oriC* on a circular map of the chromosome. Resolution of chromosome dimers occurs after the end of chromosome replication, and relies on the presence of the FtsK C terminal to align *dif* sites on the sister chromatids (Pérals *et al.*, 2001). The FtsK C terminal activates the XerCD recombinase (Aussel *et al.*, 2002), before the DNA translocase activity is used to move the DNA to the correct position for chromosome dimer resolution. FtsK is a bidirectional motor, but direction is determined by 8 bp DNA sites, FtsK orientating polar sequences (KOPS). The γ subunit of FtsK recognises KOPS and allows translocation in one direction only. These are arranged mostly on the leading strand of DNA and act to direct FtsK towards *dif* (Sivanathan et al., 2006).

One candidate for the driving force of chromosome segregation is transcription. This idea is enforced by the observations that nucleoids appear to be more dispersed in cells grown in minimal media compared to those grown in rich medium. In rich media, much transcription is of ribosomal RNA, whereas in minimal media, when growth is slower, transcription is spread

across a range of genes, suggesting RNA polymerase has some effect on nucleoid structure (Reyes-Lamothe *et al.*, 2012). Inhibiting both transcription and translation, using rifampicin and chloramphenicol respectively, caused the nucleoid to appear more condensed (Zusman *et al.*, 1973), but the addition of rifampicin did not affect the rate of DNA segregation (Reyes-Lamothe *et al.*, 2008b). RNA polymerase is one of the strongest molecular motors in the cell, able to move DNA with a force of 30 pN, making it a good candidate for moving DNA in chromosome segregation (Wang *et al.*, 1998). One popular theory about chromosome segregation is that the transcription of membrane proteins could provide the driving force. It has been proposed that when membrane proteins are expressed a section of DNA containing the gene moves to the membrane and the gene is transcribed and translated, and the protein is inserted into the membrane sequentially, a process called transertion. This attachment to the membrane would move the DNA away from the replication machinery and towards the two daughter cells (Woldringh, 2002). However, since inhibition of transcription and translation did not disrupt chromosome segregation, it is likely that transertion is not the sole or main driving force of chromosome segregation (Reyes-Lamothe *et al.*, 2008b).

1.4 Regulation of transcription in *Escherichia coli*

1.4.1 Bacterial RNA polymerase

Unlike eukaryotes, most bacteria have one RNA polymerase responsible for all RNA synthesis (Ebright, 2000). This multi-subunit enzyme is found in two forms, the core enzyme and the holoenzyme. The core enzyme is able to catalyse RNA synthesis but cannot recognise promoter sequences and initiate transcription. The core enzyme consists of two α subunits, β and β' subunits and an ω subunit. RNAP has been crystallised, including recently the *E. coli* enzyme (Zuo *et al.*, 2013), and the structure has been described as a "crab-claw structure",

with the end of the β and β' subunits forming one arm of the claw each. Between the large β and β' subunits is a channel, containing the active site and one essential Mg^{2+} ion (Darst *et al.*, 1989, Zhang *et al.*, 1999). This active site has very high structural homology to the yeast RNAP II (Ebright, 2000). The α subunit dimer associates with the β and β' subunits, triggering assembly of the core enzyme. The N-terminal domains of the α subunits are responsible for dimerisation and construction of the multi-subunit enzyme (Hayward *et al.*, 1991). The function of the ω subunit is not clear, but it is thought it may act as a chaperone in the folding of the β' subunit (Mukherjee *et al.*, 1999, Ghosh *et al.*, 2001).

Although the core enzyme is catalytically active and can bind to DNA non-specifically, it cannot recognise the promoter sequences that signal the point for RNAP to bind and begin transcription. This function is added to the core enzyme by interaction with a σ factor to make the holoenzyme. The σ factor not only recognises the promoter sequence but by binding to it, ensures that the holoenzyme is correctly positioned to initiate transcription. Once the holoenzyme is positioned at the promoter sequence, the σ factor facilitates the unwinding of DNA at the transcription start site to allow transcription to advance (Murakami *et al.*, 2002). The σ subunit from *Thermus aquaticus* has been crystallised revealing three domains σ_2, σ_3 and σ_4 joined by flexible linkers (Campbell *et al.*, 2002). There are different families of σ factors which allow RNAP to recognise different sets of promoters. Under normal conditions, most promoters in *E. coli* are recognised by the "housekeeping" σ factor σ^{70}, a member of the σ^{70} family of σ factors. In response to stresses experienced by the cell, other σ factors will start to accumulate and compete with σ^{70} for binding to RNAP (Maeda *et al.*, 2000). These σ factors will recognise different promoter elements to σ^{70} and direct RNAP to transcribe genes, whose products will help the cell to deal with the stress.

Most bacteria have multiple σ factors each recognising different sets of promoters, with *E. coli* containing 7 σ factors (Maeda *et al.*, 2000). For example, in *E. coli*, σ^{38} is responsible for transcribing genes required in the stationary phase of growth (Hengge-Aronis, 1993). There are around 2000 molecules of RNAP per cell, of which around 700 are free for binding for transcription (Ishihama, 2000). However there are around 1200 molecules of the various σ factors so there is competition for binding to RNAP. In most conditions σ^{70} is found at the highest concentration and it has 16 fold higher affinity for RNA polymerase than σ^{38} (Maeda *et al.*, 2000). σ^{70} factors have 4 domains, σ_1, σ_2, σ_3 and σ_4. σ_1 has important roles in preventing free σ factor from recognising promoters and also in open complex formation. The other domains are highly conserved and have roles in recognising promoter elements or interacting with RNAP (Murakami and Darst, 2003).

1.4.2 Promoter recognition

The regulation of gene expression is essential to ensure that gene products are available when they are needed, but energy is not wasted making unnecessary proteins. The expression of genes is controlled by promoters and their interaction with the holoenzyme. There are around 4000 genes in *E. coli* competing for a limited supply of RNAP (Ishihama, 2000). The recognition of promoter sequence is an important factor in defining how highly a gene is expressed. There are four elements of a promoter sequence that can be recognised by the RNAP holoenzyme. The two main elements of the promoter are the -10 and -35 hexamers, with consensus sequences of 5'-TATAAT-3' and 5'-TTGACA-3' respectively. As well as the sequences of these elements the spacing between them is also important. Promoters with good matches to these two consensus sequences can function, but if the -10 and -35 elements are further from the consensus sequence there are other promoter elements that can strengthen the promoter. The extended -10 sequence is a TGn motif found immediately upstream of the -10

sequence and UP element is with an AT rich region of around 20 bp located upstream of the -35 sequence (Browning and Busby, 2004). Promoters are often referred to as "strong" or "weak", with the strongest sequences giving highest gene expression and often having promoter elements close to the consensus sequence for a given σ factor. The sequence of the promoter outside of the -10 and -35 elements can also have some effect, by stabilising the formation of the open complex and changing binding affinity.

As previously mentioned, the σ factor is essential for recognition of these promoter elements by the holoenzyme. The -10 sequence is recognised by region 2.4 of σ_2 and the -35 element by region 4.2 of σ_4 (Murakami *et al.*, 2002) (see figure 1.4). The extended -10 sequence is recognised by regions 2.4 and 3.0 of the σ factor (Sanderson *et al.*, 2003). Finally, the α subunit C-terminal domains of the holoenzyme sometimes interact with the UP element (Ross *et al.*, 1993). The similarity of these elements to consensus sequence helps to suggest the strength of the promoter although there are other factors involved. Part of the role of the promoter elements is to ensure RNAP is positioned at the correct place to initiate transcription. Once the RNAP holoenzyme is bound in the correct position for initiation of transcription it forms the open complex of DNA where the two strands have started to unwind. Once the DNA is melted, the -10 element can be recognised by σ and RNAP interacts with downstream DNA up to position +2 (Zhang *et al.*, 2012). Aromatic amino acid residues in region 2.3 of σ contact the base at position -12 and stabilise the open complex. The single stranded DNA is inserted into the channel of RNAP containing the active site (Murakami *et al.*, 2002). This process may be repeated several times due to abortive initiation, when a transcript of about 12 nt has been made, RNAP can escape from the promoter. Once RNAP has entered the elongation phase of transcription, σ is no longer needed, is released and

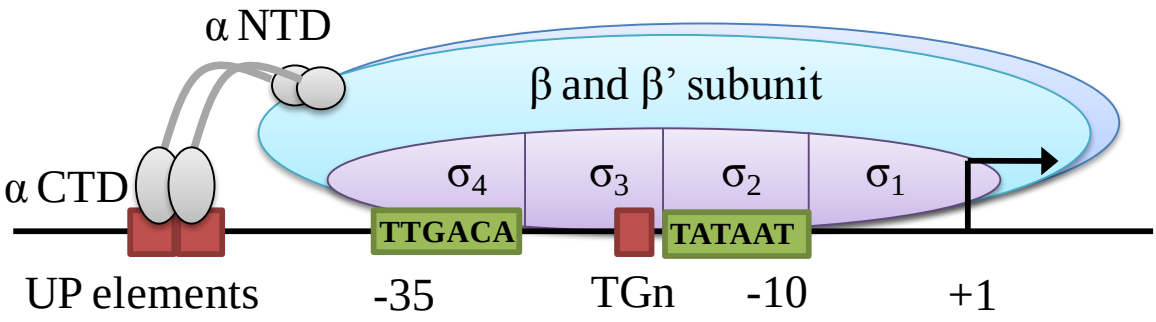


Figure 1.4: Structure of RNA Polymerase and its interactions with promoter elements

The RNA polymerase subunits α, β, β', and σ are shown with the four domains of σ labelled. Also shown are the transcription start site and promoter elements with consensus sequences for the -10 and -35 elements

Adapted from Browning and Busby, 2004.

the core enzyme completes the transcript (Browning and Busby, 2004, Murakami and Darst, 2003) .

1.4.3 Transcription factors

The activity of many promoters is regulated by transcription factors which can act as activators or repressors. They also link the activity of the promoter to signals relevant to the gene being regulated. There are 300 genes for predicted transcription factors in *E. coli*, equating to 8% of the genes (Pérez-Rueda and Collado-Vides, 2000). Most of these transcription factors have consensus sequences they recognise and bind to, directing them to a specific promoter or set of promoters. This binding can affect transcription by activating or repressing. Activators make the promoter more efficient by increasing its affinity for RNAP.

There are three common methods by which a single transcription factor can activate transcription. In the first, Class I activation, the transcription factor binds to a site located upstream of the -35 element and contacts the CTD of the α subunit of the holoenzyme (see figure 1.5a) (Ebright, 1993). As the linker between the α CTD and NTD is flexible, the exact location of this binding site can vary. Class II activation involves a binding site that overlaps the -35 element. The transcription factor then binds to this site and contacts domain 4 of the σ factor (see figure 1.5b) (Dove *et al.*, 2003). In both of these cases, binding of the transcription factor helps to recruit RNAP to the promoter. In a third mechanism the transcription factor binds between the -10 and -35 elements and aligns the promoter elements for binding by the holoenzyme (see figure 1.5c) (Browning and Busby, 2004). This mechanism is particularly seen in promoters where the spacing between the -10 and -35 elements is not optimal.

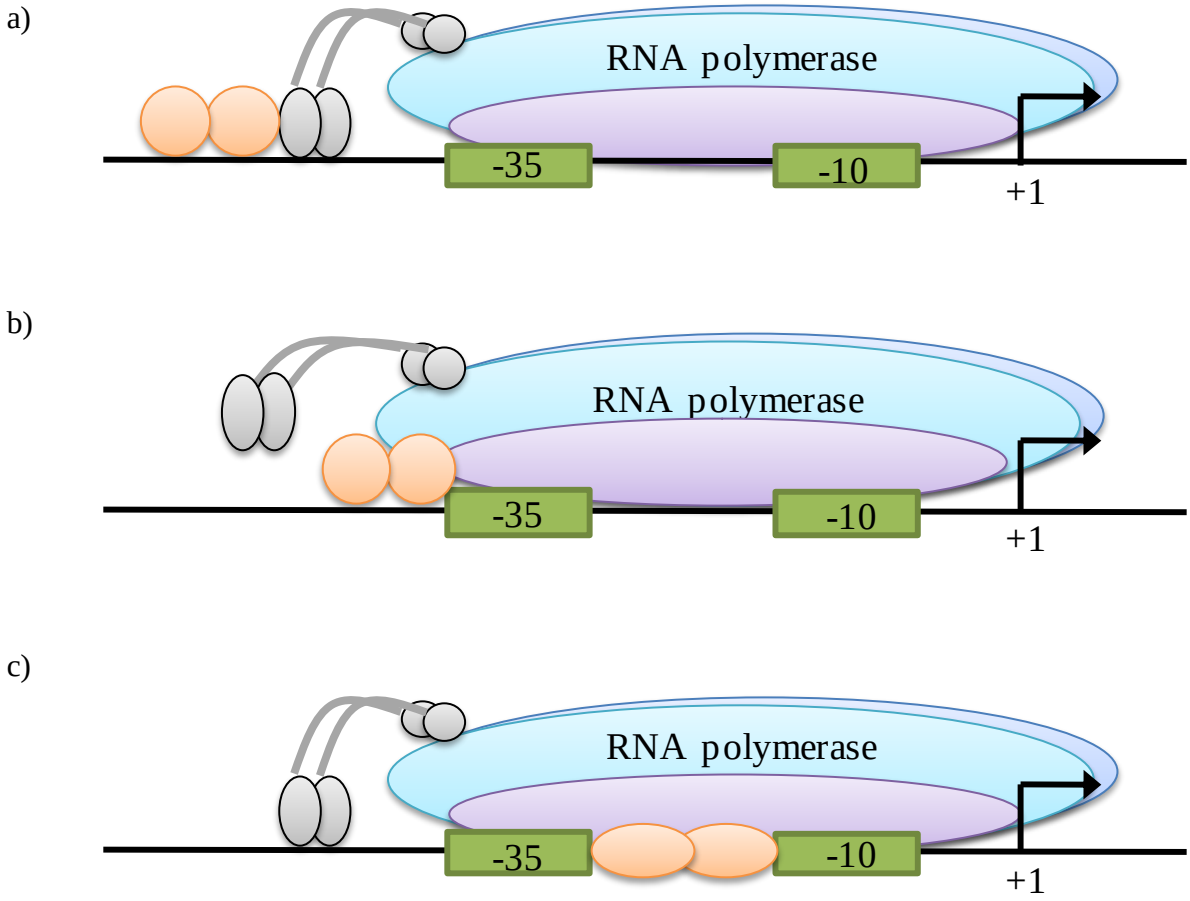


Figure 1.5: Activation at simple promoters

a) Class I activation – The activator binds to an upstream site and contacts the α CTD of RNA polymerase recruiting the polymerase to the promoter.

b) Class II activation – The activator binds to a target that is adjacent to the promoter -35 element and interacts with domain 4 of σ^{70}

c) Activation by conformational change – The activator binds at or near to the promoter elements and realigns the -10 and -35 elements so that RNA polymerase can bind to the promoter.

Adapted from Browning and Busby (2004).

There are also three common mechanisms for repression of promoters by transcription factors. One mechanism involves steric hindrance where the transcription factor binds over or near to the promoter elements preventing the holoenzyme from binding (see figure 1.6a) (Hudson and Fried, 1990). Another method involves two sets of binding sites present in the promoter which can both be bound by repressor proteins. The two sets of repressor proteins can then interact causing the DNA to loop and again preventing the holoenzyme from binding the promoter (see figure 1.6b) (Semsey *et al.*, 2002). Lastly, the repressor can bind to an activator and prevent it from performing its function, therefore preventing the promoter from being activated (see figure 1.6c) (Shin *et al.*, 2001, Browning and Busby, 2004).

The number of promoters controlled by any transcription factor differs from case to case, with 20% of transcription factors regulating two or less promoters. In contrast, 51% of genes in *E. coli* are directly regulated by at least one of a group of seven transcription factors (Martínez-Antonio and Collado-Vides, 2003). All seven of these had previously been identified as controlling a particularly large number of genes, due to regulating other transcription factors and thereby increasing the number of genes they indirectly control. These factors, called global transcription regulators, includes ArcA, FNR, FIS, CRP, IHF, LRP and H-NS (Martínez-Antonio and Collado-Vides, 2003). CRP was identified as the most important of these global regulators as it regulated 197 promoters, 22 of which were for other transcription factors. CRP senses the levels of cAMP in the cell (Busby and Ebright, 1999) thereby linking gene expression to the energetic status of the cell.

Almost half of all *E. coli* promoters are controlled by more than one transcription factor allowing for multiple signals to influence gene expression. There are 47 transcription factors

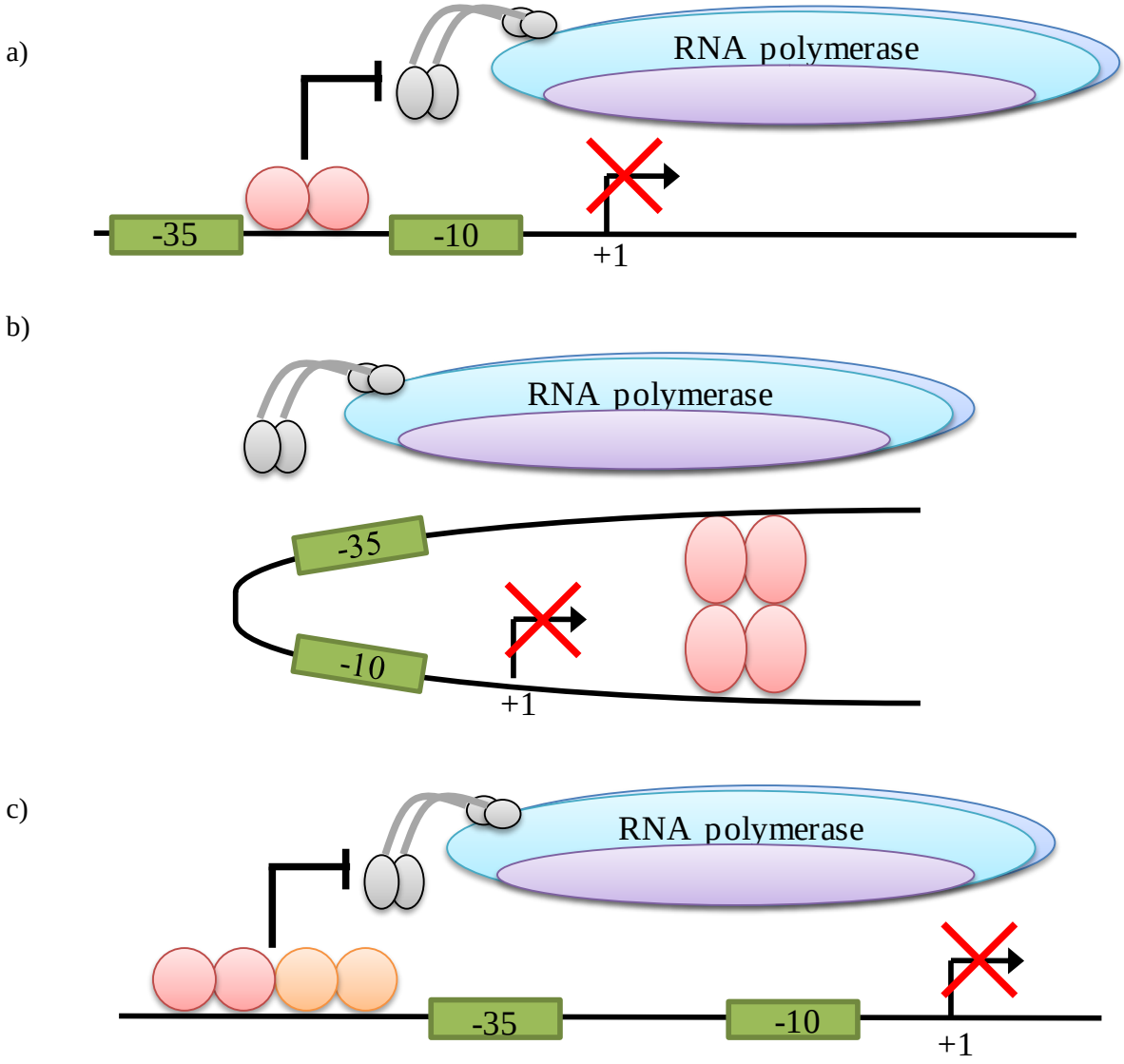


Figure 1.6: Mechanisms of repression

a) Repression by steric hindrance – The repressor binding site overlaps core promoter elements and blocks recognition of the promoter by RNA polymerase

b) Repression by looping – Repressors bind to distal sites and interact by looping, repressing the promoter.

c) Repression by modulation of an activator – The repressor binds to an activator and prevents the activator from functioning by blocking promoter recognition by RNA polymerase.

Adapted from Browning and Busby (2004).

known to regulate genes also regulated by CRP. In many cases a global regulator co-regulates with a more specific transcription factor which "fine-tunes" the response reacting to a precise signal (Martínez-Antonio and Collado-Vides, 2003). For example, the *araBAD* operon, encoding genes responsible for the metabolism of arabinose, is co-regulated by CRP and AraC. CRP senses high cellular cAMP levels, signalling low glucose levels, and AraC is activated by binding to arabinose. In combination, these two transcription factors ensure that *araBAD* is only expressed in the presence of arabinose but the absence of glucose, the exact conditions when the cell would want to be metabolising arabinose (Miyada *et al.,* 1984) (See section 4.1.3 for more detail). The combination of global and local transcription factors make a regulatory network allowing the *E. coli* cell to respond to small changes in the environment and link the regulation of genes with a related function.

1.4.4 Link between nucleoid structure and transcription regulation

As described previously, NAPs have an important role in compacting DNA and holding the nucleoid structure in place. Many of them also have another important role in the cell in the regulation of transcription. Out of the seven "global regulators" mentioned in section 1.4.3, three are also NAPs with clear links to modifying chromosome structure. These are Fis, IHF and H-NS (Martínez-Antonio and Collado-Vides, 2003). Fis alone is thought to cause at least a slight change in transcription in around 21% of currently annotated *E. coli* genes, either activating or repressing, directly or indirectly (Cho *et al.,* 2008). Chromatin immunoprecipitation analysis also showed that targets bound by Fis and H-NS are also associated with RNA polymerase. All three of these NAPs preferentially bind to AT rich sequences which suggests a possible link to promoter sequences (Grainger *et al.,* 2006). The theory of the process of transertion suggests that transcription of membrane proteins may

involve rearrangement of the chromosome to position the gene at the membrane, further evidence of a link between gene expression and chromosome structure (Woldringh *et al.*, 1995).

It is possible that the transcription of any protein encoding gene may involve some chromosome remodelling, due to the coupling of transcription and translation and the absence of ribosomes from the nucleoid (Miller *et al.*, 1970, Hobot *et al.*, 1985). It is also possible that instead of the processes of chromosome folding and transcription occurring side by side in the cell they actually work together with some factors taking on two roles simultaneously and helping towards both of these processes.

1.5 Aims of the project

The overall aim of this work is to attempt to establish if there is a link between chromosome structure and the regulation of transcription in bacteria. Both areas have been studied previously and there are indications that they could influence each other but there is little known about this possible link. The binding of the GalR repressor has been shown to have an effect on chromosome structure, as well as its previously known function of repressing transcription (Qian *et al.*, 2012). In particular, the ability of GalR to bring together binding sites at distant locations on the chromosome suggests at least some transcription factors may be able to influence chromosome structure (see section 4.1.1 for more detail). Epifluorescence microscopy was exploited to observe the structure of the chromosome and behaviour of transcription factors in the cell. Epifluorescent microscopy was chosen over confocal microscopy for these experiements as the level of fluorescence is relatively low and fluorescent foci are prone to photobleaching. In epifluorescent microscopy, light is collected

from the whole sample, in contrast to confocal microscopy where the whole cell is illuminated but light is only collected from one focal plane. Exposure times are also longer for confocal imaging.

A position on the chromosome can be tagged using a DNA binding protein fused to a fluorescent protein targeted to the site of interest, a technique known as fluorescent reporter operator systems (FROS, see section 3.1.1). The initial aim of this work was to develop a new FROS method to give the most reliable results with the least disruption to the chromosome, Also, a second FROS system was developed using the transcription factor MalI as a FROS reporter, as this has been well studied and is known to have some similarities to LacI. This provides an alternative to using TetR as a FROS reporter in combination with LacI (Chapter 3). This method was used to study inducible promoters in the presence and absence of their inducer. Sites adjacent to promoters were tagged with a fluorescent protein and followed by microscopy to establish whether the position of the promoter within the cell changes when transcription is taking place. Promoters controlled by the AraC and MntR transcription factors were chosen for study as they are well studied and promoters can be easily induced. The same method was used to investigate whether promoters regulated by the same transcription factor colocalise in the cell, despite being found at distant locations on the circular chromosome, which would indicate a possible role in chromosome organisation for transcription factors. Again, members of the AraC and MntR regulons were studied in the presence and absence of inducers. The effect of inducing transcription on the chromosome structure in cells that were at the point of division was also studied (Chapter 4).

Finally, photoactivated localisation microscopy (PALM) was used to investigate the diffusion of LacI on a single molecule level. This gives information about motility of individual molecules under different conditions. The method had previously been used to study DNA repair by visualising individual DNA polymerase molecules, and preliminary experiments in Chapter 5 show its suitability for studying transcription factors.

In combination, these results aim to uncover a potential link between transcription regulation and chromosome structure in bacteria. Although FROS has been used previously, it has not been used to study the effects of transcription. This study aims to achieve this and adapt the FROS method to cause less disruption to chromosome structure.

2. Materials and Methods

2.1 Buffers, Solutions and Reagents

All solutions to be used for bacterial growth or DNA manipulations were autoclaved at 120°C and 15 psi for 20 minutes. All chemicals were purchased from Sigma-Aldrich or Fisher Scientific unless stated otherwise. All solutions were made up with distilled and deionised water (resistivity 18.2 MΩcm). All solutions with concentrations expressed as a percentage are w/v.

2.2 Growth Media

All liquid media was autoclaved at 120°C and 15 psi for 20 minutes, and stored at room temperature. Solid media was also autoclaved 120°C and 15 psi for 20 minutes.

2.2.1 Solid Media

Nutrient Agar: 23 g nutrient agar (Difco) in 1 L ddH2O

MacConkey Agar: 40 g MacConkey agar base (Difco) in 1 L ddH2O

Solid media was made up in distilled and deionised water (ddH_2O) as above, then autoclaved at 120°C and 15 psi for 20 minutes and cooled before any supplements were added. Molten agar was then poured into sterile petri dishes under sterile conditions and stored at 4°C. An appropriate carbon source was added to MacConkey agar before autoclaving to give the required final concentration.

2.2.2 Liquid Media

Lennox broth*:* 20 g Tryptone, 10 g Yeast Extract, 10 g NaCl in 1 L ddH_2O

M9 Minimal Media 10 x Salts*:* 60 g Na_2HPO_4, 30 g KH_2PO_4,5 g NaCl,10 g NH_4Cl in 1 L ddH_2O

M9 minimal media: 10 ml 10 X M9 Salts, 90 ml Sterile ddH_2O, 200 µl 1 M $MgSO_4$, 100 µl 0.1 M $CaCl_2$, 500 µl 20% w/v Casamino acids, 1.5 ml 20% w/v Carbon source.

All components were autoclaved separately before use and fresh media was made up on day of use.. Antibiotics and other supplements were added at time of use.

M9 minimal media for PALM: 10 ml 10 x M9 salts, 90 ml sterile ddH_2O, 200 µl 1 M $MgSO_4$, 100 µl 0.1 M $CaCl_2$, MEM amino acids + proline, MEM vitamins, 0.2% glycerol

2.2.3 Antibiotics and other supplements

Ampicillin: Solution of 40 mg/ml in sterile ddH_2O, stored at -20°C. Used at a final concentration of 100-200 µg/ml.

Chloramphenicol: Solution of 35 mg/ml in ethanol, stored at -20°C. Used at a final concentration of 35 µg/ml.

Kanamycin: Solution of 50 mg/ml in sterile ddH_2O, stored at -20°C. Used at a final concentration of 50 µg/ml.

Tetracycline: Solution of 10 mg/ml in ethanol, stored at -20ºC. Used at a final concentration of 35 µg/ml.

Rifampicin: Solution of 12.5 mg/ml in ethanol, stored at -20ºC. Used at a final concentration of 50 µg/ml.

IPTG: Solution of 100 mM in sterile ddH_2O, stored at -20°C. Diluted as required.

X-Gal: Solution of 50 mg/ml in dimethyl sulfoxide and stored at -20°C. Used at a final concentration of 40 µg/ml.

Arabinose: Solution of 20% w/v in sterile ddH_2O. Used at a final concentration of 0.3%

Fructose: Solution of 20% w/v in sterile ddH_2O. Used at a final concentration of 0.3%

Glucose: Solution of 20% w/v in sterile ddH_2O. Used at a final concentration of 0.3%

All stock solutions were made up as above and sterilised using a 0.2 μm syringe filter. To supplement solid media, agar was cooled to approximately 50ºC before the supplement was added to give the final concentration required and mixed well by shaking.

2.3 Bacterial Strains

Strains were stored as glycerol stocks at -80°C. Before use the strain was restreaked onto nutrient agar plates and incubated overnight at 37°C. Overnight cultures were grown by inoculating a single colony in 5 ml LB or M9 minimal media, with antibiotics if required, and incubating overnight at 37°C, or lower temperature if necessary, with orbital shaking at 250 rpm in non-baffled flasks, unless specified otherwise. To monitor growth of bacterial cultures, optical density at 650 nm was measured using a Helios Gamma spectrophotometer. Strains used in this study are listed in Table 2.1.

2.4 Plasmids

All plasmids used in this study are listed in table 2.2. Plasmids maps are shown in figures 2.1 to 2.13

Table 2.1 *E.coli* K-12 strains used in this study (continued on next page)

Strain	Description	Origin
RLG221	recA56 araD139 (ara-leu)7697 lacX74 galU galK hsdR strA	R. Gourse
MG1655	F- λ- ilvG rfb-50 rph-1	(Cherepanov and Wackernagel, 1995)
C2925	Ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) Tet^S endA1 rspL136 (Str^R) dam13::Tn9 (Cam^R) xylA-5 mtl-1 thi-1 mcrB1 hsdR2	New England Biolabs
ΔaraC 83	*araC* derivative of BW25113	(Baba *et al.*, 2006)
KH000	*lacI* derivative of MG1655	Hollands (unpublished)
KH001	Derivative of KH000 in which the *lacZ* promoter has been removed	Hollands (unpublished)
DL01	Derivative of MG1655 in which LacI has been tagged with GFP on the chromosome and the *lacZ* promoter has been removed	(Lee *et al.*, 2009)
DL02	Derivative of DL01 from which the *kan* cassette has been removed	This study
LR04	Derivative of KH001 in which 22 copies of the LacI DNA sites inserted adjacent to *araBAD*	This study
MSR02	Derivative of LR04 in which LacI has been tagged with GFP on the chromosome with the *lacZ* promoter removed	M. Sánchez-Romero (Unpublished)
LR06	Derivative of MSR02 from which the *kan* cassette has been removed	This study
LR15	Derivative of DL02 in which 6 copies of the LacI DNA sites inserted adjacent to *araBAD*	This study
LR16	Derivative of MG1655 in which 20 copies of the MalI DNA site have been inserted adjacent to *araBAD*	This study
LR17	Derivative of LR16 from which the *kan* cassette has been removed	This study
LR18	Derivative of LR15 from which the *kan* cassette has been removed	This study

Table 2.1 *E. coli* K-12 strains used in this study (continued on next page)

Strain	Description	Origin
LR19	Derivative of LR06 in which 20 copies of the MalI DNA site have been inserted adjacent to *araJ*	This study
LR20	Derivative of MG1655 in which 20 copies of the MalI DNA site have been inserted adjacent to *araF*	This study
LR22	Derivative of DL02 in which 22 LacI DNA sites have been inserted adjacent to *araJ*	This study
LR29	Derivative of MG1655 in which LacI has been tagged with PAmCherry on the chromosome and the *lacZ* promoter has been removed	This study
LR30	Derivative of LR18 in which LacI has been tagged with PAmCherry on the chromosome and the *lacZ* promoter has been removed	This study
LR31	Derivative of LR19 from which the *kan* cassette has been removed	This study
LR35	Derivative of LR29 from which the *kan* cassette has been removed	This study
LR37	Derivative of LR30 from which the *kan* cassette has been removed	This study
LR38	Derivative of LR20 from which the *kan* cassette has been removed	This study
LR39	Derivative of LR22 from which the *kan* cassette has been removed	This study
LR40	Derivative of LR06 in which 20 MalI DNA sites have been inserted adjacent to *mntH*	This study
LR42	Derivative of LR40 from which the *kan* cassette has been removed	This study
LR44	Δ*araC* derivative of LR38	This study
LR45	Derivative of SXB4 in which 20 MalI DNA sites have been inserted adjacent to *mntH*	This study
LR46	Derivative of LR06 in which 20 MalI DNA sites have been inserted adjacent to *araFGH*	This study
LR47	Derivative of LR45 from which the *kan* cassette has been deleted	This study
LR48	Derivative of LR46 from which the *kan* cassette has been deleted	This study

Table 2.1 *E. coli* K-12 strains used in this study (continued)

Strain	Description	Origin
SXB1	Derivative of MG1655 in which 20 copies of the MalI DNA site have been inserted adjacent to *mntH*	S. Bevan (unpublished)
SXB2	Derivative of MG1655 LacI::GFP in which 22 copies of the LacI DNA site have been inserted adjacent to *dps*	S. Bevan (unpublished)
SXB3	Derivative of SXB1 from which the *kan* cassette has been removed	S. Bevan (unpublished)
SXB4	Derivative of SXB2 from which the *kan* cassette has been removed	S. Bevan (unpublished)

Table 2.2 Plasmids used in this study (continued on next page)

Plasmid	Description	Origin
pDOC-C	Donor plasmid for gene goctoring with *sacB* gene for counter selection (Amp^R *sacB*) (figure 2.1)	(Lee et al., 2009)
pDOC-G	Donor plasmid for gene doctoring used for making GFP fusions on Chromosome (Kan^R Amp^R *sacB*)	(Lee et al., 2009)
pACBSR	Mutagenesis plasmid containing λ red proteins and SceI under control of an arabinose inducible promoter for use in gene doctoring (Cm^R) (figure 2.2)	Scarab Genomics
pPM301	pUC19 derivative containing 22 LacI DNA site repeats (Amp^R)	P.McGlynn (unpublished)
pPM461	pUC19 derivative containing 6 LacI DNA site repeats (Amp^R)	P.McGlynn (unpublished)
pJB10	Donor plasmid for gene doctoring used for inserting *plac GFP* onto the chromosome at *ara* locus (Kan^R Amp^R) (figure 2.3)	J.Bryant (unpublished)
pJB32	Donor plasmid for gene doctoring used for inserting *plac GFP* onto the chromosome at *rcsB* locus. Contains *sacB* counter selection (Kan^R Amp^R *sacB*) (figure 2.4)	J.Bryant (unpublished)
pCP20	Temperature sensitive plasmid encoding FLP recombinase (Amp^R Cm^R)	(Cherepanov and Wackernagel, 1995)
pUC19	Coloning vector (Amp^R) (figure 2.5)	
pUCMal20	pUC19 containing 20 repeats of the MalI DNA site (Amp^R)	This Study
pET20b	Protein expression vector (Amp^R)	Novagen
pET22b	Protein expression vector containing constitutively expressed LacI (Amp^R)	Novagen
pACYC184	Cloning vector (Tet^R Cm^R) (figure 2.6)	(Chang and Cohen, 1978)
pACYCΔHN	Cloning vector with *tetR* deletion (Cm^R)	C. Webster (unpublished)
pmCherry-N1	Plasmid carrying mCherry suitable for N terminal fusions	ClonTec
pJW15Δ100	Plasmid for fusing genes to the *melR* promoter (Amp^R) (figure 2.7)	(Kahramanoglou *et al.*, 2006)

Table 2.2 Plasmids used in this study (continued on next page)

Plasmid	Description	Origin
pLR1	Derivative of pJB10. Donor plasmids for inserting 6 LacI DNA sites adjacent to *araBAD* (Amp^R)	This study
pLR2	Derivative of pJB10. Donor plasmids for inserting 22 LacI DNA sites adjacent to *araBAD* (Amp^R)	This study
pLR5	Derivative of pLR1. Donor plasmid for inserting 6 LacI DNA sites adjacent to *araBAD* with *kan* cassette (Kan^R Amp^R)	This study
pLR6	Derivative of pLR2. Donor plasmid for inserting 22 LacI DNA sites adjacent to *araBAD* with *kan* cassette (Kan^R Amp^R)	This study
pLR7	Derivative of pLR5 with pDOC-C backbone. Donor plasmid for inserting 6 LacI DNA sites adjacent to *araBAD* (Kan^R Amp^R *sacB*)	This study
pLR8	Derivative of pLR6 with pDOC-C backbone. Donor plasmid for inserting 22 LacI DNA sites adjacent to *araBAD*(Kan^R Amp^R *sacB*)(figure 2.8)	This study
pLR11	Derivative of pJB32. Donor plasmid containing 20 copies of the MalI DNA site (Kan^R Amp^R *sacB*)(figure 2.9)	This study
pLR15	Derivative of pJB32. Donor plasmid containing *ftnB* homology region upstream of the *kan* cassette (Kan^R Amp^R *sacB*)	This study
pLR16	Derivative of pLR15. Donor plasmid containing *yecJ* homology region downstream of *placZGFP* (Kan^R Amp^R *sacB*)	This study
pLR17	Derivative of pLR16. Donor plasmid for inserting 20 copies of the MalI DNA site adjacent to *araFGH* (Kan^R Amp^R *sacB*) (figure 2.9)	This study
pLR19	Derivative of pLR8. Donor plasmid for inserting 20 copies of the MalI DNA site adjacent to *araBAD* (Kan^R Amp^R *sacB*) (figure 2.9)	This study
pLR20	Derivative of pLR17. Donor plasmid for inserting 22 LacI DNA sites adjacent to *araFGH* (Kan^R Amp^R *sacB*) (figure 2.8)	This study
pLR22	Derivative of pJB32 Donor plasmid containing *araJ* homology region upstream of the *kan* cassette (Kan^R Amp^R *sacB*)	This study

Table 2.2 Plasmids used in this study (continued on next page)

Plasmid	Description	Origin
pLR23	Derivative of pLR22. Donor plasmid containing *mak* homology region downstream of *placZGFP* (Kan^R Amp^R *sacB*)	This study
pLR24	Derivative of pLR23. Donor plasmid for inserting 20 copies of the MalI DNA site adjacent to *araJ* (Kan^R Amp^R *sacB*) (figure 2.9)	This study
pLR25	Derivative of pLR24. Donor plasmid for inserting 22 copies of the LacI DNA site adjacent to *araJ* (Kan^R Amp^R *sacB*) (figure 2.8)	This study
pSB1	Derivative of pJB32. Donor plasmid containing *ypeC* homology region upstream of the *kan* cassette (Kan^R Amp^R *sacB*)	S. Bevan (unpublished)
pSB2	Derivative of pJB32. Donor plasmid containing *dps* homology region downstream of *placZGFP* (Kan^R Amp^R *sacB*)	S. Bevan (unpublished)
pSB3	Derivative of pSB1. Donor plasmid containing *mntH* homology region downstream of *placZGFP* (Kan^R Amp^R *sacB*)	S. Bevan (unpublished)
pSB4	Derivative of pSB2. Donor plasmid containing *rhtA* homology region upstream of the *kan* cassette (Kan^R Amp^R *sacB*)	S. Bevan (unpublished)
pSB5	Derivative of pSB3. Donor plasmid for inserting 20 copies of the MalI DNA site adjacent to *mntH* (Kan^R Amp^R *sacB*) (figure 2.9)	S. Bevan (unpublished)
pSB6	Derivative of pSB4. Donor plasmid for inserting 22 copies of the LacI DNA site adjacent to *dps* (Kan^R Amp^R *sacB*) (figure 2.8)	S. Bevan (unpublished)
pLER101	Derivative of pACYC184 containing LacI fused to GFP under the control of the *lacI* promoter (Cm^R) (figure 2.10)	This study
pLER104	Derivative of pLER101 containing MalI fused to mCherry under the control of the *malI* promoter (Cm^R) (figure 2.11)	This study
pLER105	Derivative of pJW15 containing MalI::mCherry fused to *melR* promoter (Cm^R) (figure 2.12)	This study

Table 2.2 Plasmids used in this study (continued)

Plasmid	Description	Origin
pLER107	Derivative of pLER101 containing LacI fused to FLAG tag under the control of the *lacI* promoter (Cm^R)	This study
pLER108	Derivative of pLER105 containing MalI fused to mCherry under the control of the *melR* promoter (Cm^R) (figure 2.13)	This study
pRW901	Derivative of pRW50 containing the *lacZ* gene without a promoter for use as a negative control in β-galactosidase assays (Tet^R)	(Butala *et al.*, 2009)
pBAD	Derivative of pRW901 containing the *lacZ* gene under the control of the *araBAD* promoter (Tet^R)	M. Sánchez-Romero
pBAD-HisB-PAmCherry	Plasmid encoding PAmCherry	Clontech
pDOC-PAM-lac	Derivative of pDOC-G (Lee *et al.*, 2009) containing PAmCherry and homology regions to tag AraC with PAmCherry (Kan^R Amp^R *sacB*) (figure 2.14)	This study

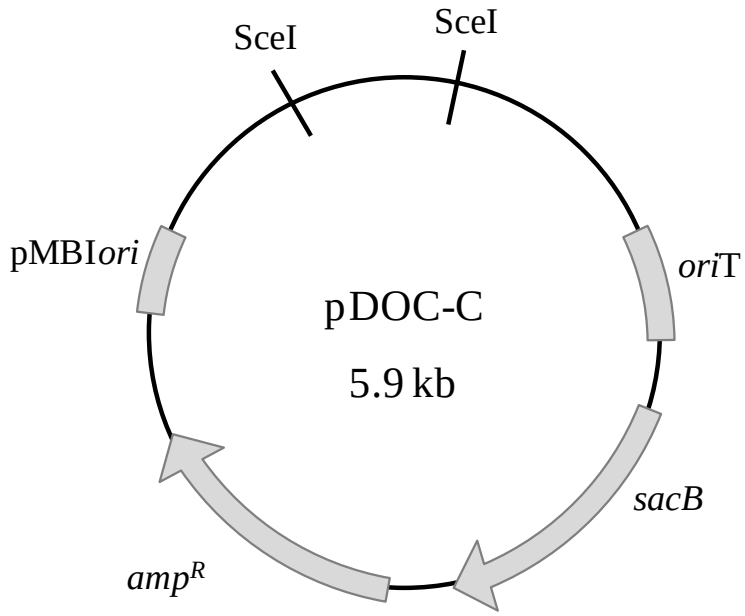


Figure 2.1: Plasmid map of pDOC-C

pDOC-C is used for making gene doctoring donor plasmids. SceI sites are shown flanking a cloning region. Also shown are the ampicillin resistance gene (*amp*^{R}), the origins of replication (*ori*T and pMBI*ori*) and the sucrose sensitivity gene (*sacB*).

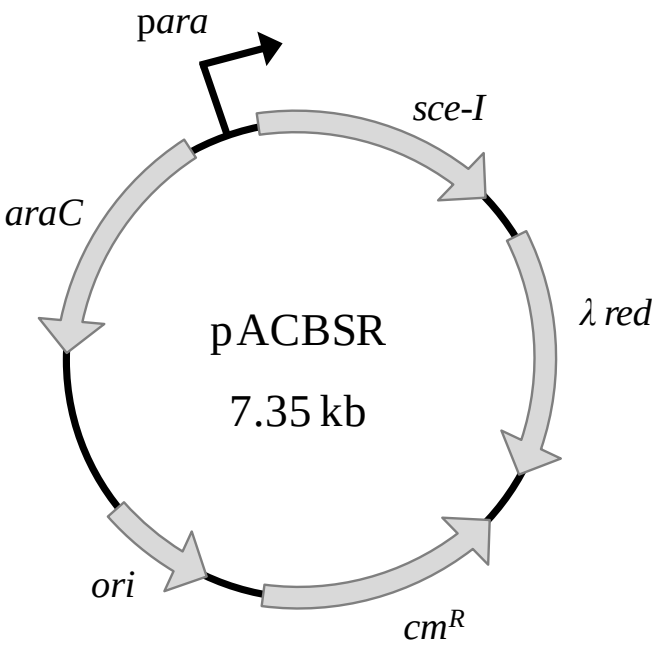


Figure 2.2: Plasmid map of pACBSR

pACBSR is a mutagenesis plasmid used in gene doctoring. It encodes the λ Red recombination proteins and the SceI yeast meganuclease, both under the control of the *araBAD* promoter (p*ara*). The *araC* gene is included to provide the activator for the *araBAD* promoter. Also shown are the chloramphenicol resistance gene (cm^R) and origin of replication (*ori*).

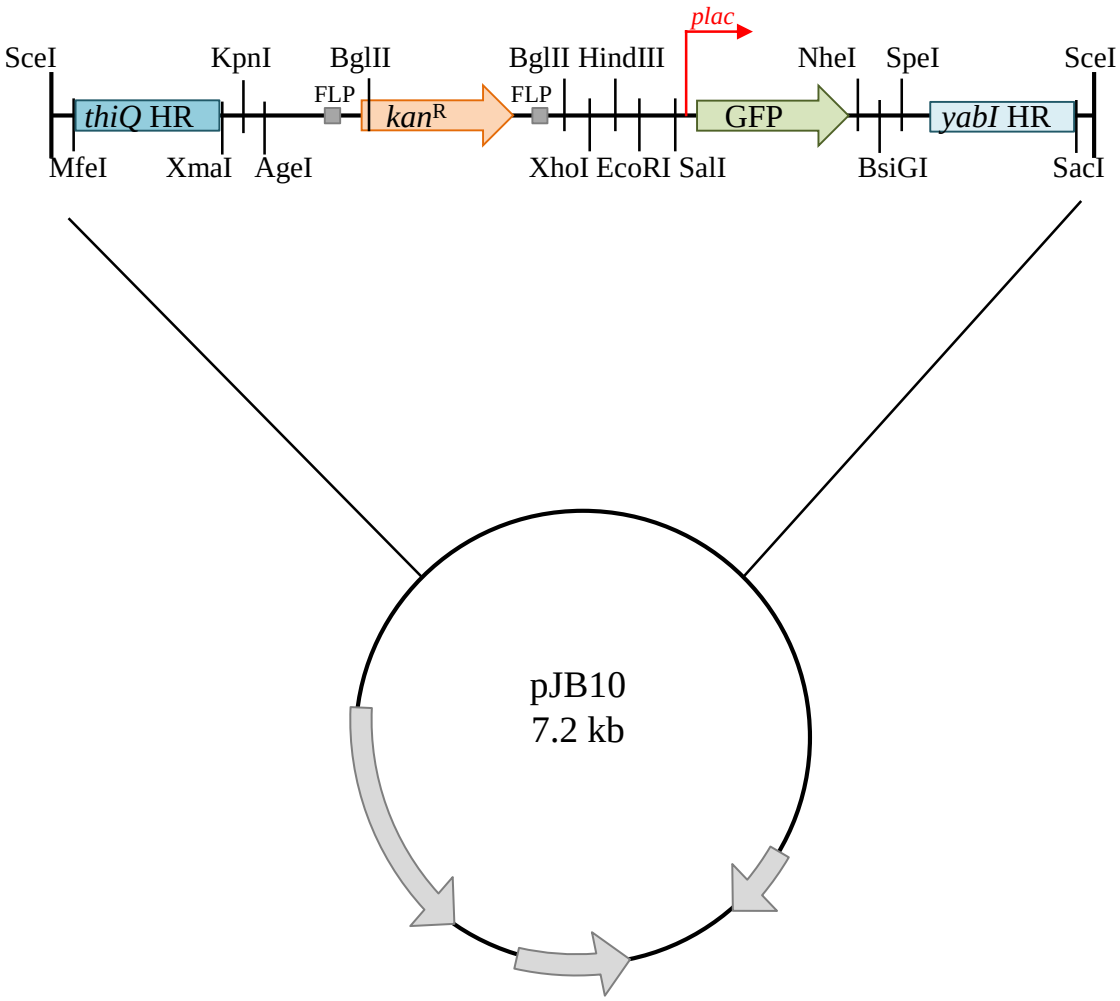


Figure 2.3: Plasmid map of pJB10

pJB10 is a gene gorging plasmid for inserting p*lacZ gfp* adjacent to the *araBAD* promoter. It carries 500 bp of homology to the 3' end of *thiQ* flanked by MfeI and XmaI restriction sites. Also carried is 500 bp of homology to the 3' end of *yabI* flanked by NheI and SacI restriction sites. Between the two homology regions are a kanamycin resistance cassette (kan^R) and the *lacZ* promoter fused to *gfp*. Also shown are the ampicillin resistance gene (*bla*), the origin of replication (*ori*) and the plasmid replication gene (*rop*).

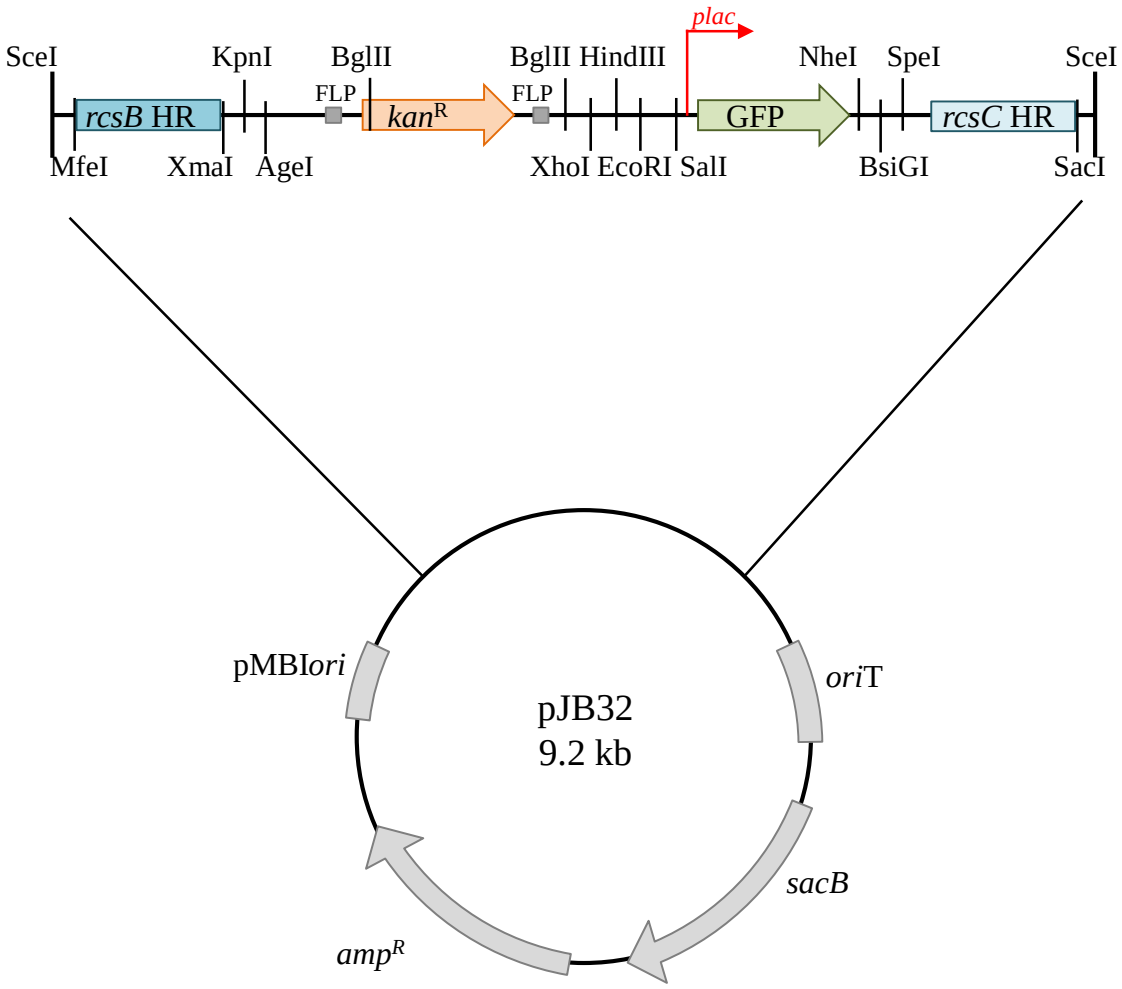


Figure 2.4: Plasmid map of pJB32

pJB32 is a gene doctoring plasmid for inserting p*lacZ gfp* adjacent to the *araBAD* promoter. It carries 500 bp of homology to the 3' end of *thiQ* flanked by MfeI and XmaI restriction sites. Also carried is 500 bp of homology to the 3' end of *yabI* flanked by NheI and SacI restriction sites. Between the two homology regions are a kanamycin resistance cassette (kan^R) and the *lacZ* promoter fused to *gfp*. Also shown are the ampicillin resistance gene (amp^R), the origins of replication (*ori*T and pMBI*ori*) and the sucrose sensitivity gene (*sacB*).

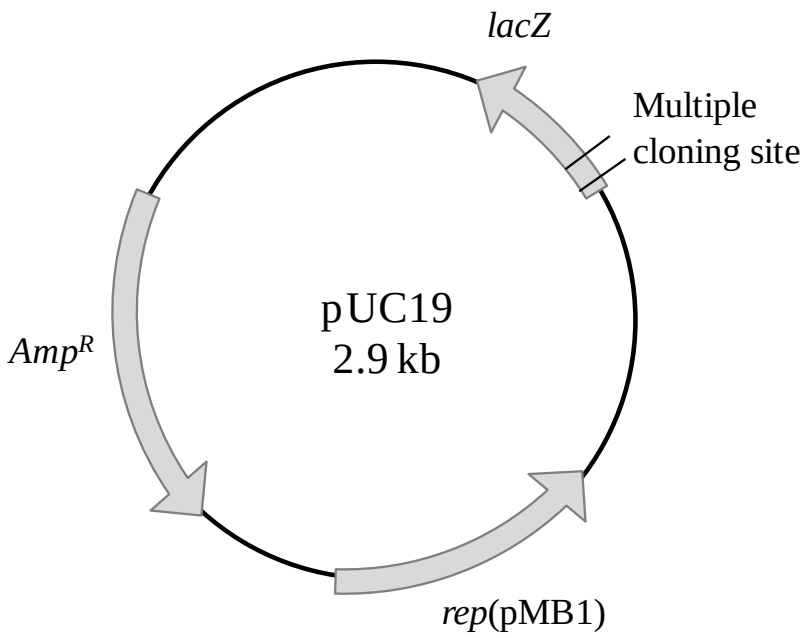


Figure 2.5: Plasmid map of pUC19

pUC19 is a cloning vector encoding *lacZ* with a multiple cloning site within the gene. Successful inserts into this multiple cloning site can be selected for by blue-white selection on X-gal plates. Also shown are the ampicillin resistance gene (Amp^R) and origin of replication (*rep*(pMB1)).

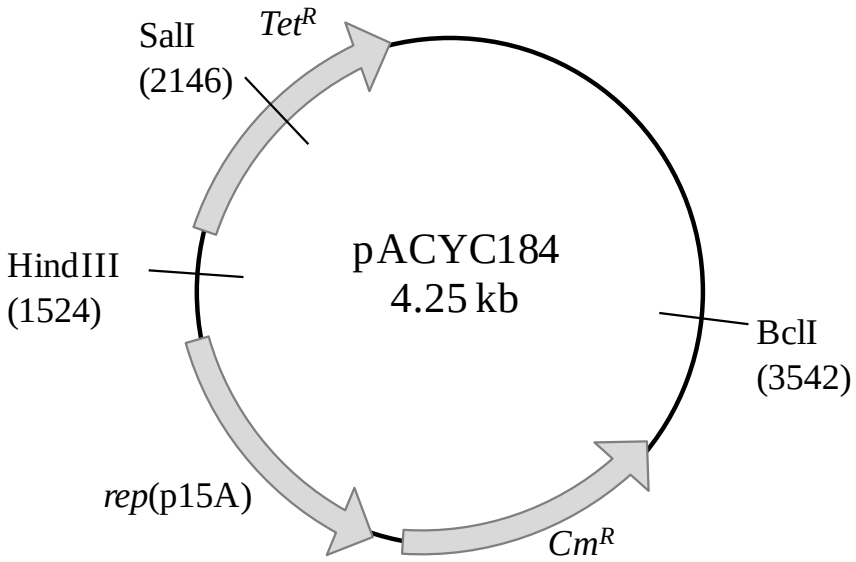


Figure 2.6: Plasmid map of pACYC184

pACYC184 is a small, medium copy number *E. coli* plasmid. It contains the replicon *rep* for replication of the plasmid, a tetracycline resistance gene (tet^R) and a chloramphenicol resistance gene (cm^R). Also shown are the positions or restriction sites for HindIII, BclI and SacI.

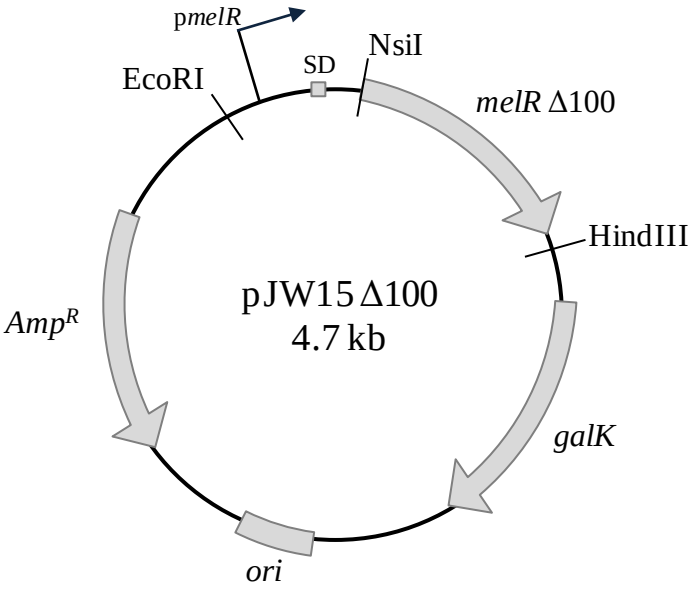


Figure 2.7: Plasmid map of pJW15 Δ100

pJW15 Δ100 is a multicopy plasmid carrying the *melR* promoter. NsiI and HindIII restriction sites flank a truncated version of the *melR* gene. Genes of interest can be cloned into the plasmid to place the gene under the control of the *melR* promoter. The ATG of the gene is removed and the NsiI site (ATGCAT) provides the translation start codon. Also shown are the ampicillin resistance gene (Amp^R), the *galK* gene, the origin of replication (*ori*) and Shine-Dalgarno sequence (SD).

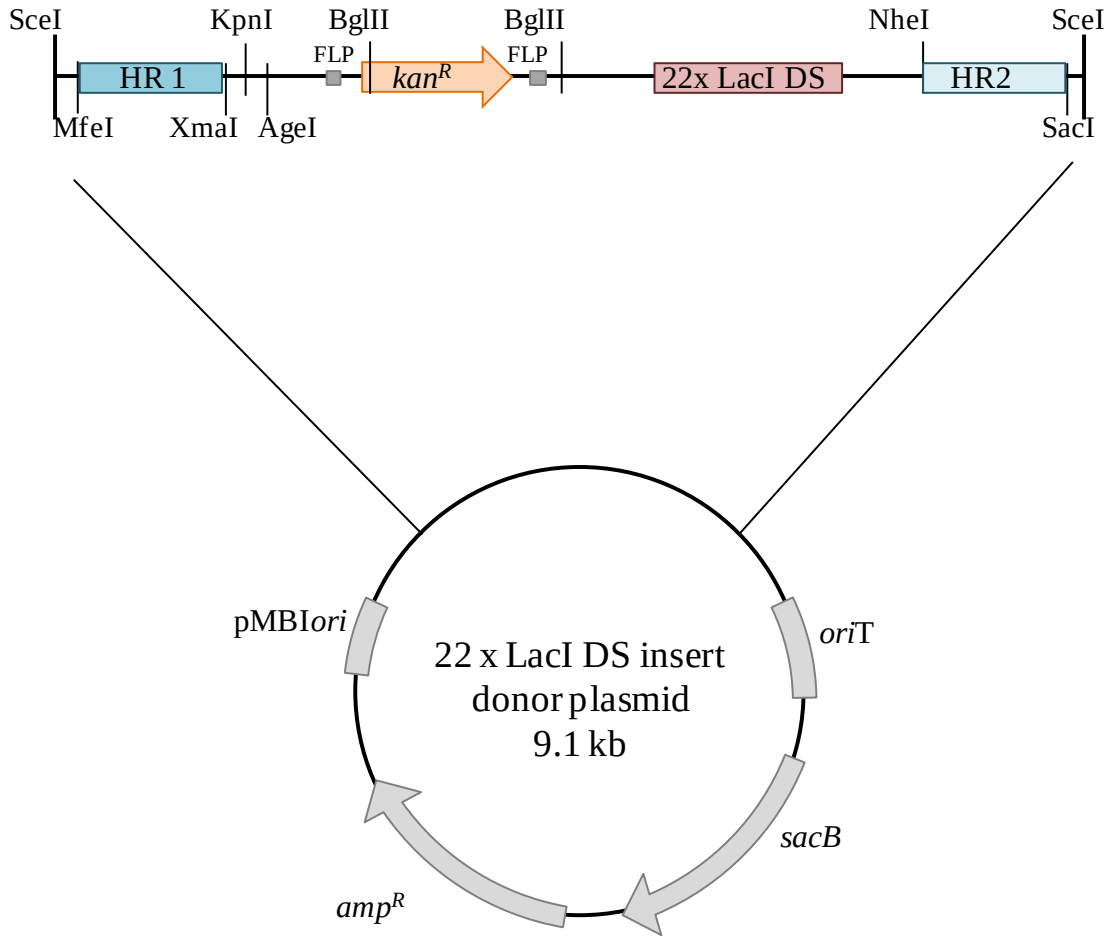


Figure 2.8: Plasmid map of donor plasmids used for inserting 22 copies of the LacI DNA site onto the chromosome

Gene doctoring donor plasmids for inserting 22 LacI DNA site onto the chromosome have the same basic structure shown. They carry 500 bp of homology to gene upstream of the insertion site flanked by MfeI and XmaI restriction sites. Also carried is 500 bp of homology to the gene downstream of the insertion site flanked by NheI and SacI restriction sites. Between the two homology regions are a kanamycin resistance cassette (*kan*^{R}) and 22 copies of the LacI DNA site (LacI DS). Also shown are the ampicillin resistance gene (*amp*^{R}), the origins of replication (*ori*T and pMBI*ori*) and the sucrose sensitivity gene (*sacB*).

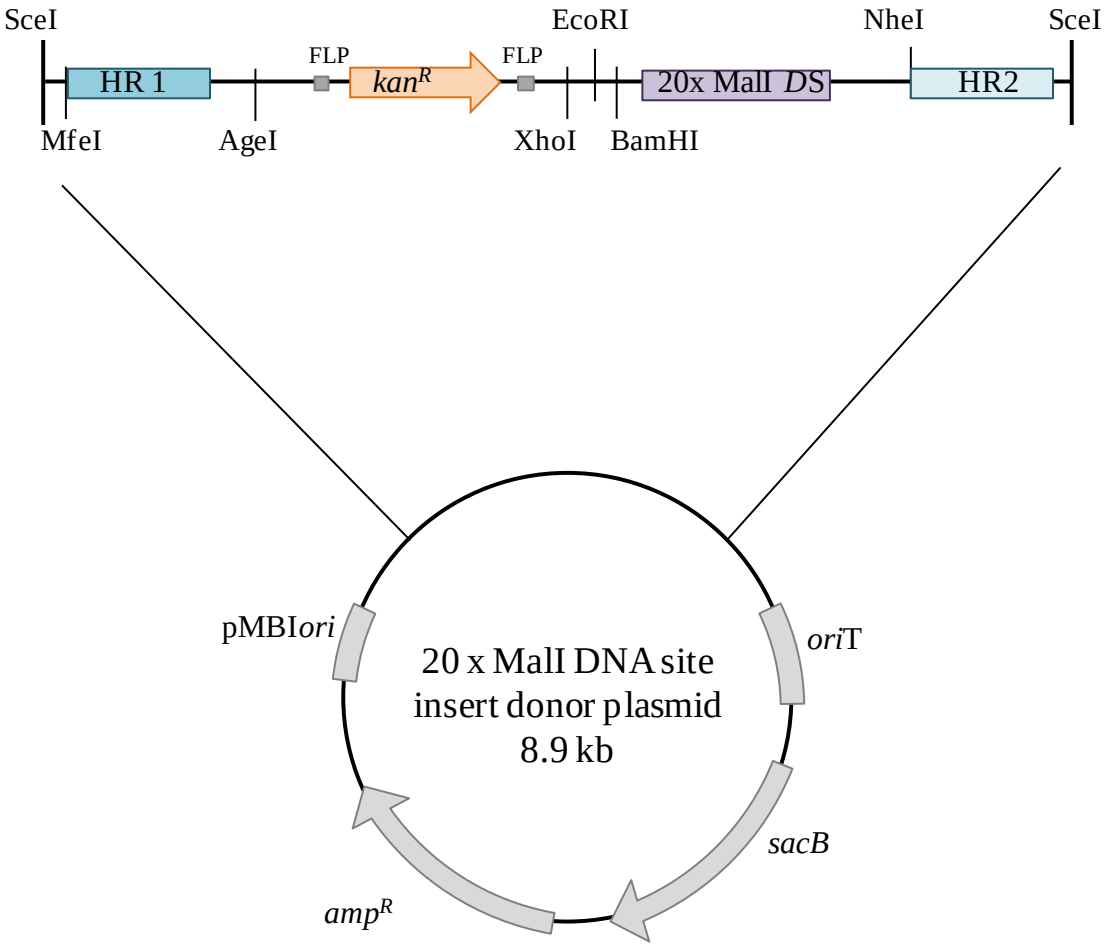


Figure 2.9: Plasmid map of donor plasmid used inserting 20 copies of the MalI DNA site onto the chromsome

Gene doctoring donor plasmids for inserting 20 MalI DNA sites onto the chromosome have the same basic structure shown. They carry 500 bp of homology to gene upstream of the insertion site flanked by MfeI and XmaI restriction sites. Also carried is 500 bp of homology to the gene downstream of the insertion site flanked by NheI and SacI restriction sites. Between the two homology regions are a kanamycin resistance cassette (*kan*^{R}) and 20 copies of the MalI DNA site (MalI DS). Also shown are the ampicillin resistance gene (*amp*^{R}), the origins of replication (*ori*T and pMBI*ori*) and the sucrose sensitivity gene (*sacB*).

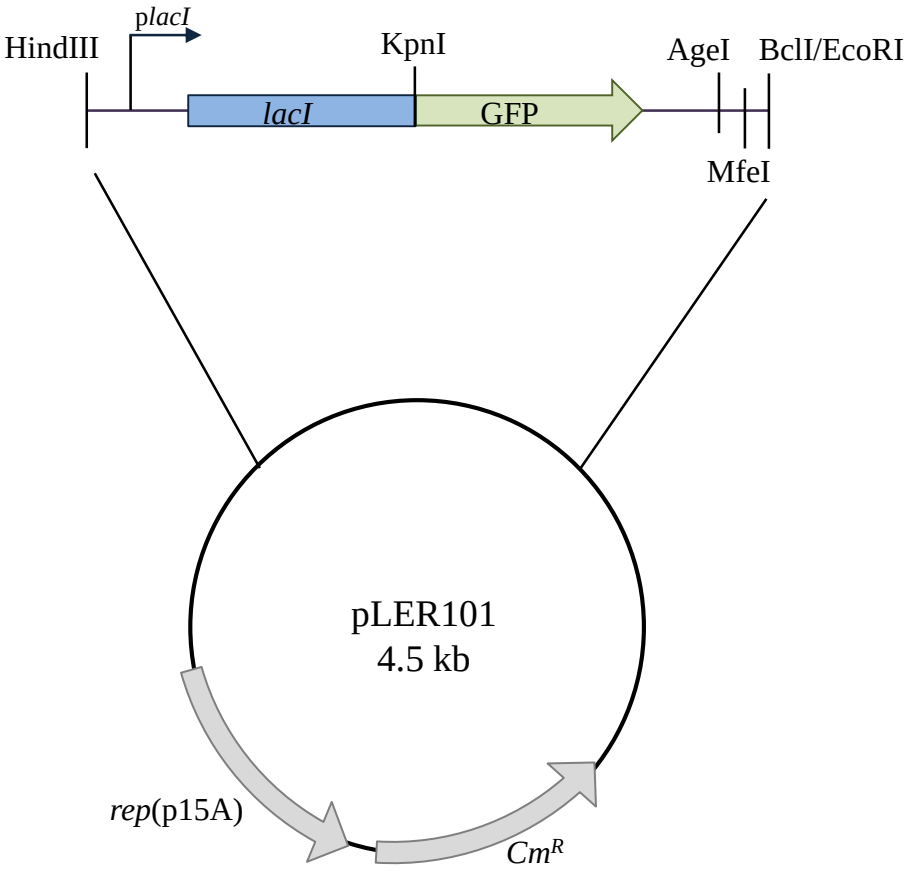


Figure 2.10: Plasmid map of pLER101

pLER101 is a pACYC184 derivative encoding LacI::GFP under the control of the *lacI* promoter (p*lacI*). Also shown are the origin of replication (*rep*(p15A)) and chloramphenicol resistance gene (Cm^R).

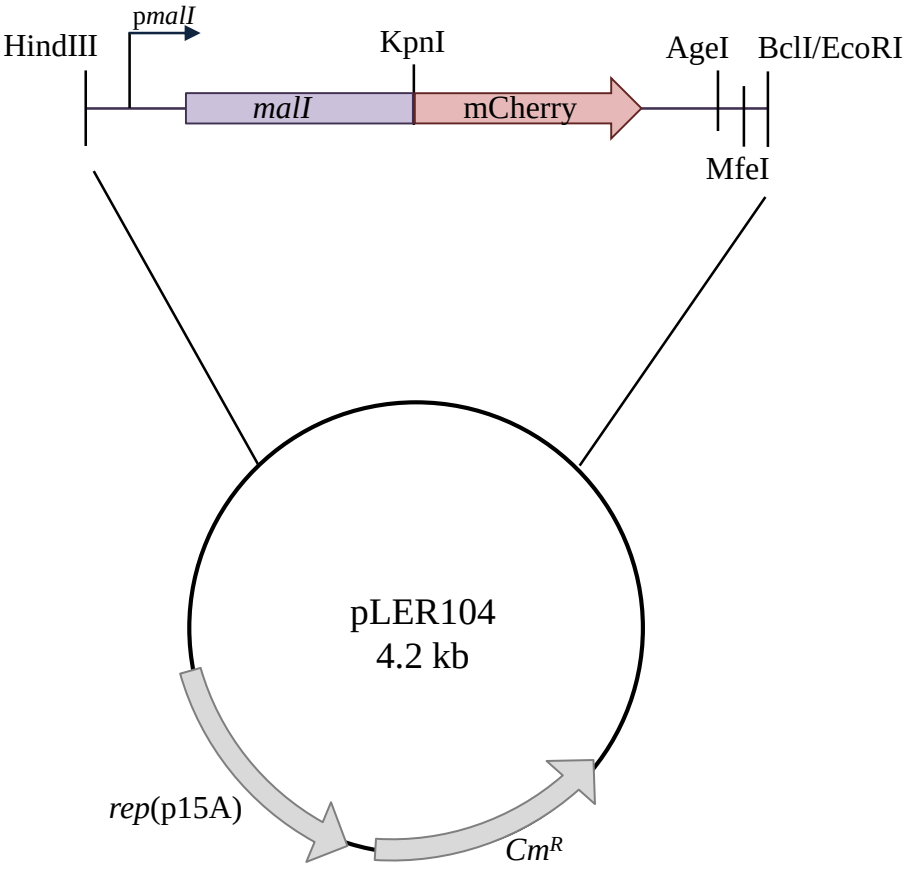


Figure 2.11: Plasmid map of pLER104

pLER104 encodes MalI::mCherry under the control of the *malI* promoter (p*malI*). Also shown are the origin of replication (*rep*(p15A)), chloramphenicol resistance gene (Cm^R) and MalI binding sites.

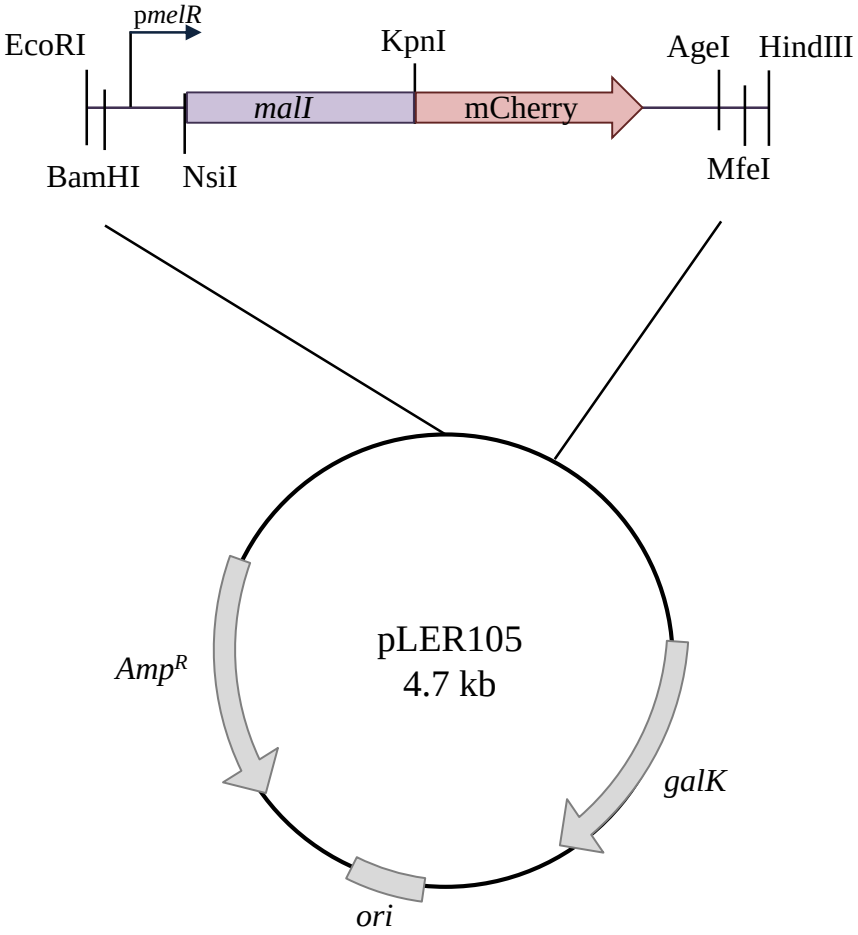


Figure 2.12: Plasmid map of pLER105

pLER105 encodes MalI::mCherry under the control of the *melR* promoter. Also shown are the origin of replication (*ori*), ampicillin resistance gene (Amp^R) and *galK*.

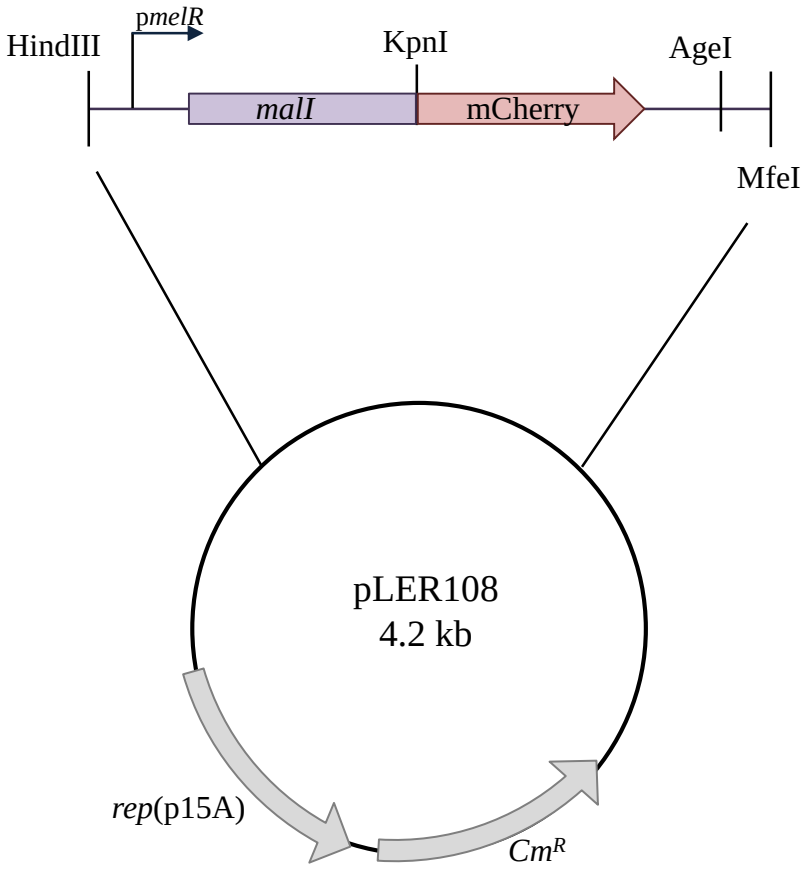


Figure 2.13: Plasmid map of pLER108

pLER108 encodes MalI::mCherry under the control of the *melR* promoter. Also shown are the origin of replication (*rep*(p15A)) and chloramphenicol resistance gene (Cm^R).

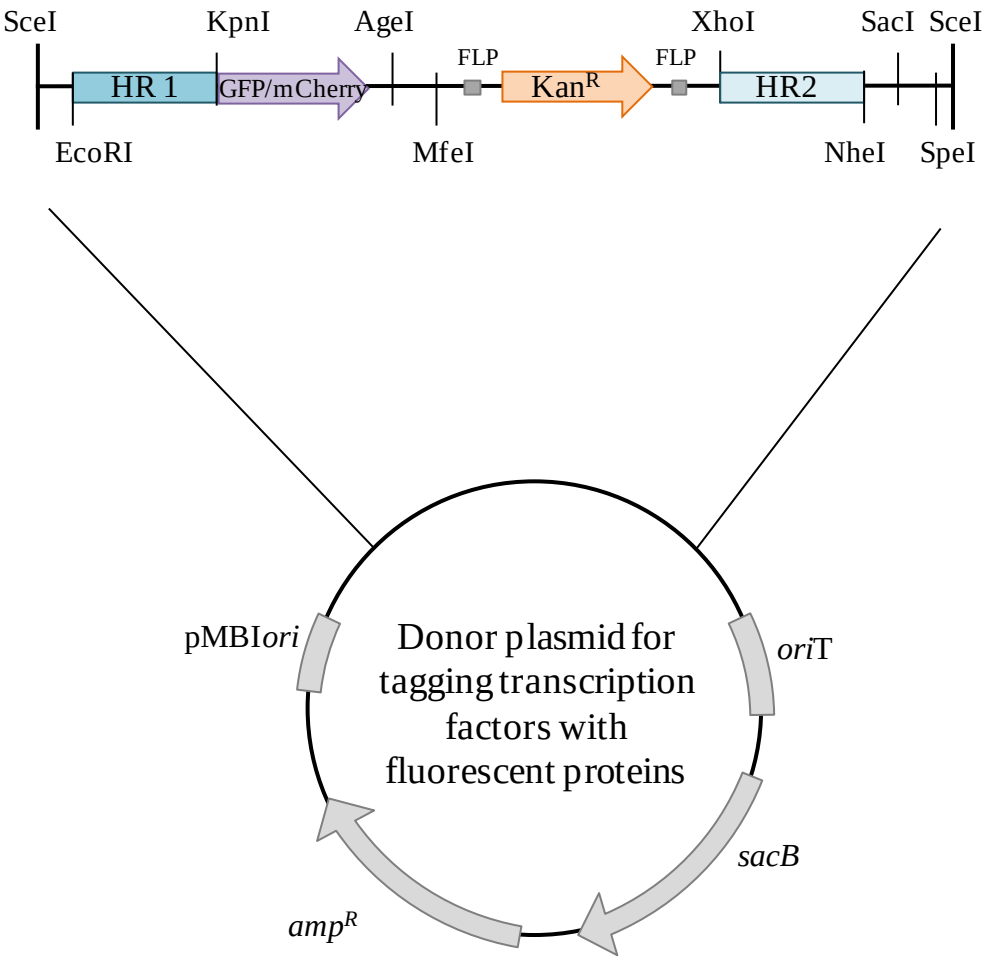


Figure 2.14: Plasmid map of donor plasmids used for tagging transcription factors with fluorescent proteins on the chromosome

Gene doctoring donor plasmids for tagging transcription factors with fluorescent proteins on the chromosome have the same basic structure shown. They carry a region of homology to gene upstream of the insertion site flanked by EcoRI and KpnI restriction sites. Also carried is a region of homology to the gene downstream of the insertion site flanked by XhoI and NheI restriction sites. Between the two homology regions are a kanamycin resistance cassette (*kan*^{R}) and a fluorescent protein gene, flanked by KpnI and AgeI restriction sites. Also shown are the ampicillin resistance gene (*amp*^{R}), the origins of replication (*ori*T and pMBI*ori*) and the sucrose sensitivity gene (*sacB*).

2.5 Gel electrophoresis of DNA

2.5.1 Agarose gel electrophoresis

DNA loading dye: 0.025 % Bromophenol Blue, 0.025 % Xylene Cyanol F, 20% glycerol, 10 mM TRIS, 1 mM EDTA

DNA Markers: 100 bp and 1 kb DNA ladders (NEB)

Ethidium Bromide*:* 10 mg/ml (Biorad)

SYBR Safe: (Invitrogen)

40 x TAE Buffer: 2 M Tris acetate, 100 mM Na_2EDTA (National Diagnositics). Diluted to 1 x for use as running buffer.

0.8% Agarose solution: 0.8g agarose dissolved in 100 ml 1x TAE buffer.

Agarose gels were used to analyse and purify DNA fragments of more than 200 bp in length. 0.8% agarose solution was heated in a microwave to dissolve agarose before bring cooled to around 50°C and poured into a gel casting tray with a comb. DNA samples were mixed with loading dye in a ratio of 5:1 and loaded into the wells. The gel was run at 100 V in TAE buffer for 30-45 minutes. Gels were stained in a solution of 0.5mg/ml ethidium bromide solution or SYBR Safe, diluted approximately 100 fold from concentrate, for 30 minutes and then visualised using a UV transilluminator. For gel extractions, gels were stained with SYBR Safe and visualised using a blue light box and orange filter to avoid UV damage to DNA to be used in downstream processes.

2.5.2 Polyacrylamide gel electrophoresis

5 x TBE Buffer: 0.44 M Tris borate pH8.3, 10 mM Na_2EDTA (National Diagnositics). Diluted to 1 x for use as running buffer.

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

7.5% acrylamide solution: 125 ml acrylamide stock solution, 100ml 5x TBE, 20 ml glycerol, made up to 500 ml with ddH_2O

Polyacrylamide gels were used for analysis of small fragments of DNA or to differentiate between two bands when the size difference was very small. For a small gel containing 9 wells 10 ml working solution was used. This was polymerised by adding 100 µl 10% APS and 15 µl TEMED (N,N,N',N'-tetramethylethane-1,2-diamine). DNA samples were mixed with DNA loading dye and loaded into the wells. Gels were run in 1 x TBE buffer at 30-40 mA constant current for 30 minutes to 1 hour, stained in ethidium bromide solution and visualised using a UV transilluminator.

2.6 Extraction and purification of nucleic acids

2.6.1 Purification of DNA using QIAquick PCR purification kit

Purification of PCR products or DNA from a restriction digest was carried out using a QIAquick PCR purification kit (QIAgen) or ISOLATE PCR and Gel kit (Bioline) following manufacturer's instructions. Purified DNA was eluted in 50 µl elution buffer provided in the kit or 50 µl ddH_2O.

2.6.2 Extraction of DNA from agarose gels

DNA samples were run on a 0.8% agarose gel as described previously. The gels were then stained in SYBRSafe for 30 minutes and a blue light box was used to visualise the DNA in combination with an orange filter. The required bands were then excised and DNA fragments were purified using the QIAquick gel extraction kit (QIAgen) or ISOLATE PCR and Gel kit (Bioline) following the manufacturer's instructions. Samples were eluted into 50 µl elution buffer provided in the kit or 50 µl ddH_2O.

2.6.3 Preparation of plasmid DNA

A fresh colony of a strain carrying the required plasmid was used to inoculate an overnight culture in LB medium supplemented with the appropriate antibiotic. Plasmid DNA was extracted from the culture using a QIAprep Spin Miniprep Kit (QIAgen) or ISOLATE Plasmid Mini Kit (Bioline) following manufacturer's instructions. Plasmid DNA was eluted into 50 µl elution buffer provided in the kit or 50 µl ddH_2O.

2.7 Transformations

2.7.1 Preparation of competent cells using the calcium chloride method

Calcium chloride: 0.1 M

The strain to be made competent was streaked out to provide fresh, single colonies. One colony was used to inoculate an overnight culture in LB media. The following day, 1 ml of overnight culture was used to inoculate 50 ml LB media and the culture was incubated at 37ºC until OD_{650} reached around 0.3-0.5, mid-exponential phase. The culture was decanted into a sterile centrifuge tube and incubated on ice for 10 minutes, before the cells were harvested by centrifugation for 10 minutes at 3400 x *g* and 4ºC. The supernatant was removed

and the pellet resuspended in 25 ml ice cold 0.1 M $CaCl_2$. The cells were incubated on ice for 20 minutes before being harvested by centrifugation for 10 minutes at 3400 x *g* and 4ºC. The pellet was resuspended in 3.3 ml ice cold 0.1 M $CaCl_2$ with 15% glycerol and incubated on ice for at least 2 hours or overnight prior to use. Cells were stored as aliquots at -80ºC.

2.7.2 Transformation

Approximately 100 ng DNA of plasmid were added to 100 µl competent cells and incubated on ice for 30 minutes. Cells were then heat shocked at 42°C for 90 seconds and incubated on ice for a further 5 minutes. 500 µl LB was added to the cells and they were incubated at 37°C for 1 hour with shaking. Cells were harvested by centrifugation at 18,000 x *g* and resuspended in approximately 100 µl LB, before being plated out onto nutrient agar plates, supplemented with the appropriate antibiotic. Plates were incubated overnight at 37°C or an appropriate temperature.

2.8 DNA Manipulations

All primers used in this study are shown in table 2.3

2.8.1 Standard PCR

Phusion polymerase (NEB) was used for most standard PCR reactions using the buffer supplied. Template DNA was either approximately 1 ng genomic DNA or approximately 1 ng plasmid DNA from a miniprep, diluted 1 in 100 in sterile ddH_2O. dNTPs (Bioline) were present at a final concentration of 1 µM each. Primers (Alta Biosciences) were used at a final concentration of 0.5 µM each. PCR reactions were made up to 50 µl in sterile ddH_2O. 0.5 µl of enzyme was used (2 units). A typical PCR cycle used is shown in figure 2.15

Table 2.3 DNA oligonucleotides used in this study (continued on next page)

Name	Sequence (5'→3')	Use
Oligonucleotides used for gene doctoring donor plasmids - General		
D58793	GGATGTGCTGCAAGG	Sequencing primer to check homology regions inserted between NheI and SacI. Binds downstream of SceI site
D58794	TATGCTTCCGGCTCG	Sequencing primer to check homology regions inserted between MfeI and XmaI. Binds downstream of SceI site
D76207	CCTTTAGCAGCCCTTGCGCC	Sequencing primer to check homology regions inserted between MfeI and XmaI. Binds upstream of Kan^R
D69988	TCAAGATCTGATCAAGAGACAGGATGAGG	Primer for amplifying Kan^R cassette. Binds upstream of Kan^R cassette over BglII site
D68556	TTTACGTCGCCGTCCAG	Primer for amplifying Kan^R from pJB10. Binds in GFP downstream of *lacZ* promoter
Oligonucleotides used for gene doctoring donor plasmids – Insert adjacent to araBAD		
D69231	GCCGCAATTGCCGGGATTGAAACTGAACG	Upstream primer for amplification of the *thiQ* homology region. Carries MfeI site.
D69232	GCCGCCCGGGCGACGCTTGCCGCGTCTTATC	Downstream primer for amplification of the *thiQ* homology region. Carries XmaI site.
D69234	GCCGGCTAGCCATCAGGCAACCCCGCAC	Upstream primer for amplification of the *yabI* homology region. Carries NheI site.
D69233	GCCGGAGCTCCTGAACATGCGTTGCATCAAC	Upstream primer for amplification of the *yabI* homology region. Carries SacI site.
D69747	GTCGCACAGAACATCGG	Upstream primer for checking inserts adjacent to *araBAD*
D69748	TCGCTGGTCATTTCTGAAG	Downstream primer for checking inserts adjacent to *araBAD*
Oligonucleotides used for gene doctoring donor plasmids – Insert adjacent to araFGH		
D74949	CGGCAATTGCTCTCAAATGAACCGCGA	Upstream primer for amplification of the *ftnB* homology region. Carries MfeI site.
D74950	TAACCCGGGCAGATGAGGCAGCGG	Downstream primer for amplification of the *ftnB* homology region. Carries XmaI site.
D74951	CAAGCTAGCTGTTTGAAGCAGCGG	Upstream primer for amplification of the *ypeC* homology region. Carries NheI site.

Table 2.3 DNA oligonucleotides used in this study (continued on next page)

Name	Sequence (5'→3')	Use
D74952	GTCGAGCTCTGTCATATTATAAGCGC	Upstream primer for amplification of the *ypeC* homology region. Carries SacI site.
D75296	AGGTATGGCAACCGCTGG	Upstream primer for checking inserts adjacent to *araFGH*
D75297	TGCTGCGACAATGGCCG	Downstream primer for checking inserts adjacent to *araFGH*
Oligonucleotides used for gene doctoring donor plasmids – Insert adjacent to araJ		
D75738	GCTCAATTGTGCGCGGGATTATTTGCC	Upstream primer for amplification of the *araJ* homology region. Carries MfeI site.
D75739	GCTCCCGGGATCATGCCTGATGCGACG	Downstream primer for amplification of the *araJ* homology region. Carries XmaI site.
D75740	GCTGCTAGCGCGCCAATTGCCTACGTT	Upstream primer for amplification of the *mak* homology region. Carries NheI site.
D75741	GCTGAGCTCATCGGCACGGGATGCG	Downstream primer for amplification of the *mak* homology region. Carries SacI site.
D76827	TTCACCACTGCGCATTGCAGC	Upstream primer for checking inserts adjacent to *araJ*
D76828	TTCAGAAGCAGTAGATGGCGCG	Downstream primer for checking inserts adjacent to *araJ*
Oligonucleotides used for gene doctoring donor plasmids – Insert adjacent to dps		
D75742	GCTCAATTGTGTGGTTCCTGCTACCG	Upstream primer for amplification of the *rhtA* homology region. Carries MfeI site.
D75743	GCTCCCGGGGAGAAATTCTGCATGGTTATGC	Downstream primer for amplification of the *rhtA* homology region. Carries XmaI site.
D75744	GCTGCTAGCGCTACTTTTCCTCTACACCG	Upstream primer for amplification of the *dps* homology region. Carries NheI site.
D75745	GCTGAGCTCCCCCAGAGCTACACCG	Downstream primer for amplification of the *dps* homology region. Carries SacI site.
D76491	TTTCGTCTGGGTTGTGCTGGC	Upstream primer for checking inserts adjacent to *dps*
D76492	CGTTGTGGATGTCCAGCG	Downstream primer for checking inserts adjacent to *dps*

Table 2.3 DNA oligonucleotides used in this study (continued on next page)

Name	Sequence (5'→3')	Use
Oligonucleotides used for gene doctoring donor plasmids – Insert adjacent to mntH		
D75746	GCTCAATTGATGAGGCTTATCTGACGC	Upstream primer for amplification of the *mntH* homology region. Carries MfeI site.
D75747	GCTCCCGGGGCCAATGGAGCACAATGC	Downstream primer for amplification of the *mntH* homology region. Carries XmaI site.
D75748	GCTGCTAGCGGACGCGTTTAATGGCG	Upstream primer for amplification of the *ypeC* homology region. Carries NheI site.
D75749	GCTGAGCTCGTGCTGGTGGTAACACG	Downstream primer for amplification of the *ypeC* homology region. Carries SacI site.
D76493	GCTACAGCTGCGGCGGC	Upstream primer for checking inserts adjacent to *mntH*
D76494	GCGGCAATAACCGTTTCTTGCG	Downstream primer for checking inserts adjacent to *mntH*
Oligonucleotides used for gene doctoring donor plasmids – Tagging lacI		
D59400	GCTGAATTCCAAACAGGATTTTCGCCTGC	Upstream primer for amplification of the *lacI* homology region for tagging *lacI*. Carries EcoRI site.
D59401	GCTGGTACCCTGCCCGCTTTCCAGTC	Downstream primer for amplification of the *lacI* homology region for tagging *lacI*. Carries KpnI site.
D59402	GCTCTCGAGGTTGATGAAAGCTGGCTAC	Upstream primer for amplification of the *lacZ* homology region for tagging *lacI*. Carries XhoI site.
D59403	GTAGCTAGCTCCGCCACATATCCTG	Downstream primer for amplification of the *lacZ* homology region for tagging *lacI*. Carries NheI site.
D61347	CGTTGGTGCGGATATCTCGG	Upstream primter for checking inserts at the 3' end of *lacI*
D73820	GCGGGATCCATGGTGAGCAAGGGCGAG	Upstream primer for amplifying PAmCherry. Carries a KpnI site
D75378	GTCACCGGTGTTTTATCAGACCGCTTC	Downstream primer for amplifying PAmCherry. Carries a AgeI site

Table 2.3 DNA oligonucleotides used in this study (continued on next page)

Name	Sequence (5'→3')	Use
Oligonucleotides used for checking araC *deletion*		
D76717	CTGGCATGAATGGCTTAACT GGCCG	Upstream primer for checking inserts at the 3' end of *araC*
D76718	CGGCCACAGCAGGCAGCCG	Downstream primer for checking inserts at the 3' end of *araC*
Oligonucleotides used for cloning repressor protein::fluorescent protein tags into plasmids		
D63048	CTCTTCAAATGTAGCACCTG AAG	Sequencing primer which binds upstream of the HindIII site in pACYC184 derived plasmids
D69688	GCCAAGCTTTGGCACGGGAA CC	Upstream primer for amplifying *lacI::gfp* from the chromosome of strain MG1655 LacI::GFP. Carries a HindIII site
D69689	GCCGGATCCTCTTGAAGTTC CT	Downstream primer for amplifying *lacI::gfp* from the chromosome of strain MG1655 LacI::GFP. Carries a BamHI site
D71000	CTGGGTACCATGGTGAGCAA GGG	Upstream primer for amplifying *mCherry* from plasmid pmCherry-N1. Carries a KpnI site
D71001	CTGCAATTGCTAGAGTCGCG GCC	Downstream primer for amplifying *mCherry* from plasmid pmCherry-N1. Carries a MfeI site
D63433	CGATAAGCTTCAAAACGTTT TATCAAATTTTAGTG	Upstream primer for amplifying *malI* from pACYCMalI. Carries a HindIII site
D71192	TTAGGTACCTTTCGCTGCAA TGAGCC	Downstream primer for amplifying *malI* from pACYCMalI. Carries a KpnI site
D72022	CATGATGCATGCTACCGCCA AAAAAATAACC	Upstream primer for amplifying *malI::mCherry* from pLER104 without the *malI* promoter. Carries a NsiI site
D71850	CATAAAGCTTCAATTGCTAG AGTCGCGG	Downstream primer for amplifying *malI::mCherry* from pLER104 without the *malI* promoter. Carries a NsiI site
D77566	GTGAAGCTTCCGGGGATCCC GGGAAGA	Upstream primer for amplifying *malI::mCherry* under the control of the *melR* promoter from pLER105. Carries a HindIII site

Table 2.3 DNA oligonucleotides used in this study (continued on next page)

Name	Sequence (5'→3')	Use
D77567	GTGCAATTGCTAGAGTCGCG GCCGCTA	Downstream primer for amplifying *malI::mCherry* under the control of the *melR* promoter from pLER105. Carries a MfeI site
Oligonucleotides used for generating a multiple MalI DNA site array		
D71689	CGAGTCGACACGT**GATAAA ACGTTTTATC**AGGACTCTA GAGGATCCCCGGG	Primer for amplifying pUC19 introducing a MalI DNA site (bold) and SalI site (boxed). Carries a XbaI site (underlined)
D71690	CGACTCGAGCATG**GATAAA ACGTTTTATC***GCTAGC*TGCA AGCTTGGCGTAATCATGGTC	Primer for amplifying pUC19 introducing a MalI DNA site (bold), XhoI site (boxed) and NheI site (italics). Carries a HindIII site (underlined)
D71587	GGTGTGAAATACCGCACAG ATGC	Check primer used to check the size of the insert in the multiple cloning site of pUC19. Anneals upstream of *lacZα* in pUC19
D71588	GCGAGTCAGTGAGCGAGGA AG	Check primer used to check the size of the insert in the multiple cloning site of pUC19. Anneals downstream of *lacZα* in pUC19
M13 Forward	GTAAAACGACGGCCAGT	Sequencing primer used to check inserts in the multiple cloning in pUC19. Anneals upstream of XbaI in pUC19

Temperature	Time		Purpose
98ºC	1m30 s		"hot start"
98ºC	10 s	35 cycles	Melting
T_A	20 s		Annealing
72ºC	X s		Extension
72ºC	5 mins		Final extension

Figure 2.15: Standard PCR cycle

Table shows a typical PCR cycle for Phusion High-Fidelity DNA Polymerase. An annealing temperature (T_A) of 3ºC higher than the melting temperature of the primer with the lower melting temperature. Extension time (X s) was calculated depending on the length of the product being made, allowing 15 s per kb.

2.8.2 Colony PCR

Colony PCRs were used to amplify DNA from the chromosomes and to screen for chromosomal insertions or deletions. In cases where the PCR product was to be used in downstream processes such as cloning, Phusion polymerase (NEB) was used as described above, by smearing a small amount of a colony on the inside of the PCR tube before adding the rest of the reaction mix. The initial denaturation time was extended to 10 minutes to lyse the cells. Biomix red (Bioline) was used for colony PCRs for checking chromosomal inserts. A colony to be checked was picked and resuspended in 40 µl sterile ddH_2O. 23 µl of this was added to 25 µl Biomix red (Bioline) and 1 µl each primer at a concentration of 10 µM to give a 50 µl reaction. A typical cycle is shown in figure 2.16.

2.8.3 Restriction Digestion of DNA

Purified DNA from PCR purification or plasmid miniprep (between 5 and 100 ng) was digested in a reaction with a total volume of 60 µl, also containing 2 µl enzyme (New England Biolabs, used as per manufacturers instructions), 6 µl of the appropriate 10x buffer as recommended by NEB and, if necessary, 1 µl BSA (final concentration, 100 µg/ml). Where possible, restriction digests involving two enzymes were done as a double digests using the compatible buffer but, if this was not possible, then two separate digests were carried out with a purification step in between. Restriction digests were incubated at 37°C in a heating block for 3 hours. Where capping of the vector was required to prevent re-ligaion of plasmid DNA, alkaline phosphatase was added to digested vector which removed the terminal 5’ phosphate. 3 µl Calf Intestinal Phosphatase (CIP, NEB) and 6 µl buffer 3 were added to purified digested DNA with a final volume of 60 µl and incubated for 1 hour at 37°C

Temperature	Time		Purpose
95ºC	10 mins		"hot start"
95ºC	10 s	30 cycles	Melting
T_A	30 s		Annealing
71ºC	X s		Extension
71ºC	10 mins		Final extension

Figure 2.16: Colony PCR cycle

Table shows a typical PCR cycle for Taq DNA Polymerase in Biomix Red. An annealing temperature (T_A) of the melting temperature of the primer with the lower melting temperature was used. Extension time (X s) was calculated depending on the length of the product being made, allowing 15-30 s per kb.

2.8.4 DNA Ligation

The concentration of vector and insert DNA were estimated using gel electrophoresis and compared to a band of a similar size on a ladder with known concentrations of DNA in each band (Hyperladder I, Bioline). NEB recommends 50 ng of vector are used in each reaction, in a ratio of 3/1 insert/vector, with a total DNA concentration of between 1 and 10 μg/ml. Usingthese references, the optimum volumes of vector and insert DNA were calculated and added to 2 μl NEB T4 ligase buffer and 1 μl T4 ligase (400 units), made up to 20 μl with ddH_2O. A control was also included containing vector, ligase and buffer, but no insert, to ensure there was no contamination from uncut vector. The reactions were incubated on ice for 30 minutes and then at room temperature for at least 2 hours or at 16ºC overnight. 5 μl of each ligation were then transformed into 50 μl competent cells, as described previously, and transformants were selected for on nutrient agar plates supplemented with the appropriate antibiotics.

2.8.5 DNA Sequencing

All sequencing was carried out at the Functional Genomics and Proteomics Laboratory, University of Birmingham. For sequencing reactions, primers were diluted to 1 μM. For sequencing of plasmids, 6.8 μl plasmid and 3.2 μl primer were used. For sequencing of a PCR product, 5 μl purified PCR product, 3 μl primer and 2 μl sterile ddH_2O were used.

2.9 Gene Doctoring

2.9.1 Using homologous recombination to make chromosomal modifications

Several techniques have been developed using homologous recombination to make chromosomal modifications in *E. coli*. Most of these use the λ-Red genes to catalyse the

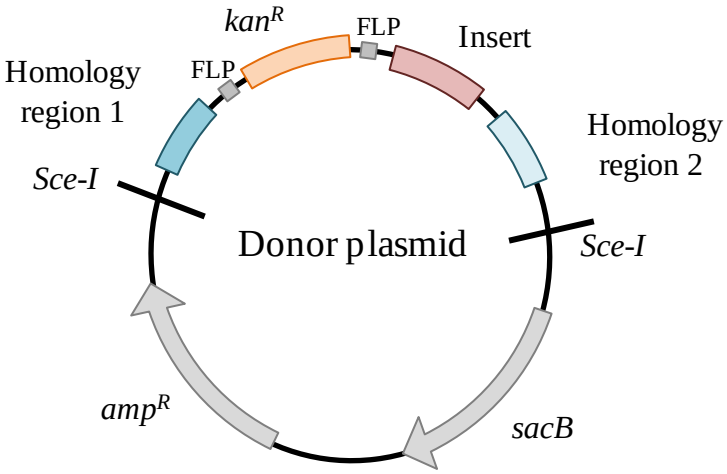
recombination and differ in the method of introducing DNA into the cell. The λ-Red proteins are used to integrate a sequence of DNA from a linear fragment onto the chromosome, at a position with homology to the regions flanking the insert sequence on the linear fragment. The specificity of this system for linear DNA should prevent unwanted recombination between similar sections of the *E. coli* chromosome or plasmids. The λ-Red system is made up of three proteins. One of these proteins, *gam*, protects single stranded DNA from degradation by the host RecBCD complex. Another, *exo*, generates single stranded overhangs on the linear DNA and the third, *bet*, catalyses recombination between homologous regions on the linear fragment and the same sequence on the chromosome, causing any sequence between these regions on the chromosome to be replaced by the sequence from the linear fragment.

The well used method developed by Datsenko and Wanner (Datsenko and Wanner, 2000) introduces the linear fragment by electroporation of a PCR product. This often includes short homology regions to a target gene of 25-40 nucleotides flanking an antibiotic resistance cassette, which results in the replacement of the target gene by the cassette. Although this has been successfully used to make deletions, the relatively low efficiency of electroporation limits this method. To eliminate the need for electroporation and increase the number of linear DNA fragments in the cells, a second technique was developed called “gene gorging”(Herring *et al.*, 2003). In this method the linear fragment is generated *in vivo* from a donor plasmid containing the homology regions, corresponding to the sequence either side of the target location on the chromosome, flanked by recognition sites for the yeast meganuclease *SceI*. The λ-Red genes are supplied on a second plasmid along with the gene for *SceI*, both under the control of an inducible promoter. When the promoter is induced *SceI* is expressed,

cleaving the donor plasmid and generating multiple copies of the linear fragment of DNA containing the homology regions, increasing the likelihood of a successful recombination event. This method was then further developed to add counter-selection to reduce the screening needed by allowing only true recombinants to grow in a method called "Gene Doctoring" (Lee *et al.*, 2009).

The donor plasmid for the gene doctoring method has two homology regions, usually about 500 bp in length each. These flank a kanamycin resistance cassette and any insert that is wanted to be incorporated onto the chromosome in the recombination process. This entire region is itself flanked by *SceI* recognition sites to form the linear fragment required. The plasmid also contains an ampicillin resistance gene and *sacB* gene, which confers sucrose sensitivity (figure 2.17). Any strain containing this plasmid will have resistance to kanamycin and ampicillin but sensitivity to sucrose. However, a strain that has successfully gone through the recombination process and incorporated the insert onto its chromosome will no longer be resistant to ampicillin or sensitive to sucrose, but will have retained the kanamycin cassette. Selecting for strains that can grow on both kanamycin and sucrose should result in only true recombinants surviving. Gene gorging and gene doctoring also require a second plasmid called the mutagenesis plasmid, pACBSR. This encodes both the *SceI* meganuclease and the λ-Red proteins, both under the control of an arabinose inducible promoter. When this promoter is induced, *SceI* is expressed to cleave the donor plasmid and the λ-Red proteins are expressed which catalyse homologous recombination of the newly formed linear fragment with the chromosome.

a)



b)

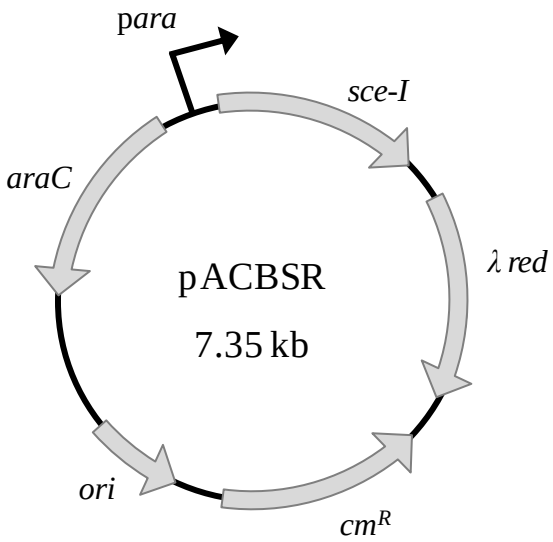


Figure 2.17: Plasmids used in gene doctoring

a) The DNA sequence to be recombined into the chromosome (insert) and antibiotic resistance cassette (*kan*^{R}) are flanked by Homology regions and *SceI* sites in a high copy number vector. Also contains *sacB* gene for counter selection
b) The mutagenesis plasmid pACBSR, encodes the *sceI* gene and λ-Red genes under the control of an arabinose inducible promoter. Also contains a chloramphenicol resistance marker.

2.9.2 Gene doctoring protocol

The strain to be gene doctored was co-transformed with pACBSR and the donor plasmid to make the insert required and grown overnight on nutrient agar plates supplemented with kanamycin and chloramphenicol. Colonies were tested for maintenance of *sacB* on the donor plasmid by patching onto plates containing kanamycin and chloramphenicol, and plates containing kanamycin and 5% sucrose. A single sucrose sensitive colony was used to inoculate 0.5 ml LB supplemented with kanamycin and chloramphenicol and incubated at 37ºC with shaking for 2-4 hours until the culture was turbid. Cells were harvested by centrifugation and washed with 0.1 X LB three times to remove any residual antibiotics. 0.3% Arabinose was then added 0.5 ml 0.1 X LB and used to resuspend the cells to induce the expression of the λ Red genes and SceI meganuclease. The culture was incubated at 37ºC with shaking for a further 2-3 hours. Cells were harvested by centrifugation and resuspended in approximately 200 µl 0.1 X LB. This volume was then evenly split between four nutrient agar plates supplemented with kanamycin and 5% sucrose. 100 µl of 1000 fold dilution of the culture, taken prior to induction and after incubation with arabinose, were plated out onto nutrient agar plates and nutrient agar supplemented with ampicillin, to compare numbers of cells containing the uncut donor plasmid pre- and post-induction. All plates were incubated at 30ºC or room temperature until colonies were visible.

2.9.3 Checking gene doctoring candidates

Plates with the pre- and post-induction dilutions should show a significant reduction in the number of colonies on the ampicillin plates when compared to the nutrient agar plates to indicate that the process has been successful and the donor plasmid has been cleaved. Colonies grown on the plates containing kanamycin and sucrose were patched onto ampicillin

and kanamycin with sucrose to check for ampicillin sensitivity. Ampicillin sensitive candidates were then checked by colony PCR, as previously described, using primers annealing on the chromosome outside of the homology regions. PCR products showing the presence of the insert were confirmed by DNA sequencing and candidates were checked for chloramphenicol sensitivity to confirm that pACBSR had been lost from the cells.

2.9.4 Removal of the kanamycin cassette from the chromosomal insert

Candidates confirmed to have the chromosomal insert and have lost the pACBSR plasmid were made competent. Cells were transformed with pCP20 and incubated at 30ºC on nutrient plates supplemented with chloramphenicol or ampicillin. Colonies were streaked onto non-selective nutrient agar plates and incubated at 37ºC to induce expression of the FLP recombinase. Colonies from the restreak were then patched onto nutrient agar plates, unsupplemented and supplemented with kanamycin and chloramphenicol. Candidates sensitive to both kanamycin and chloramphenicol were checked by colony PCR using the same primers used to check for the insert after gene doctoring. If the kanamycin cassette was successfully removed the PCR product is approximately 1.3 kb smaller than previously.

2.10 P1 Transduction

Calcium chloride: 1 M solution

Glucose: 20% solution

P1 salts solution: 10 mM $CaCl_2$, 5 mM $MgSO_4$

Sodium Citrate: 1 M solution

Chloroform

P1 transduction was used to transfer a mutation with a selective marker from one strain, the donor strain, to another, the recipient strain. Overnight cultures were grown in LB, with selection if necessary, at 37ºC. The recipient strain was diluted 50 fold into 5 ml LB and grown at 37ºC. The donor strain was diluted 100 fold the following morning into 5 ml LB supplemented with 25 µl $CaCl_2$ and 50 µl 20% glucose. This was grown for 45 minutes at 37ºC before 100 µl of a freshly prepared P1 stock was added to the culture. After approximately 3 hours further growth, the culture had become clearer with some debris floating. 100 µl chloroform was added to the culture to ensure all cells are lysed and the culture was shaken for 10 minutes. The whole culture was centrifuged at 4ºC for 10 minutes to remove any cell debris and the supernatant containing liquid phage was collected.

Recipient strain cells were prepared by harvesting cells from 1.5 ml of the culture that had been grown for approximately 4 hours by centrifugation. Cells were resuspended in 750 µl P1 salts solution. 100 µl aliquots of cells were added to glass bijou bottles with varying volumes of liquid phage, usually 1 µl, 10 µl and 100 µl. These were incubated at room temperature for 30 minutes before 1 ml LB and 200 µl 1 M sodium citrate were added to each bottle. The cells were then incubated at 37ºC with shaking for 1 hour. Cells were harvested by centrifugation and resuspended in 100 µl LB, which was then plated out onto a nutrient agar plate supplemented with 5 mM sodium citrate and an appropriate concentration of the antibiotic the donor strain carries resistance to (Thomason and Costantino, 2007). Colonies were checked for the mutation by colony PCR as described in section 2.9.3.

2.11 Microscopy

PBS Buffer: 140 mM NaCl, 2 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, to 1 L in sterile ddH_2O. All components were autoclaved before use.

poly-L-lysine: 10 mg/ml (Sigma)

Hoechst 33258: 5 µg/ml Hoechst 33258 in PBS containing 40% glycerol (v/v)

FM 4-64: 1 µg/ml FM 4-64 (molecular probes) in PBS containing 40% glycerol (v/v)

2.11.1 Growth conditions

A single colony from a freshly restreaked plate was used to inoculate a 5 ml culture of M9 minimal media, supplemented with 0.2 % casamino acids and 0.3% fructose. Fructose was chosen as a carbon source as many of the experiments involve testing the effects of arabinose and induction by fructose does not involve CRP, unlike arabinose, so the pathways of induction by the two sugars are unconnected. When necessary the cultures were also supplemented with antibiotics at half of the usual working concentration to maintain any plasmids but not reduce growth rate. 5 ml cultures were inoculated with a single fresh colony and grown, with shaking, for 24 hours at 23ºC. Previous studies stated that LacI::GFP was seen to aggregate in cells grown at 37ºC so lower temperatures were preferred (Gordon *et al.*, 1997). Cultures were then diluted 1:50 into fresh media and grown with shaking for a further 5-6 hours until OD_{650} reached approximately 0.1. For supplementation with sugars, 0.3% of the required sugar was added to the cultures for 1 minute before slides were prepared. For supplementation with manganese, 0.1 µM $MnCl_2$ was added to the solution for the final 10 minutes of growth (Waters *et al.*, 2011).

2.11.2 Preparation of microscope slides

Once cells reached the required OD_{650}, cells were harvested by centrifugation from 1.5 ml of culture. Cells were then washed 3 times with PBS. Microscope slides were prepared by spreading 5 µl poly-L-lysine onto the centre of a slide and allowing it to dry. Cells were harvested by centrifugation and all supernatant was removed using a pipette. The pellet was resuspended in 5 µl Hoechst 33258 and 5 µl FM 4-64. For cells containing mCherry, the pellet was resuspended in 10 µl Hoechst 33258. Cells were incubated at room temperature for 5 minutes before 5 µl was pipetted into the centre of the slide and a cover slip applied.

2.11.3 Imaging and analysis

Slides were imaged using a Nikon Eclipse 90i microscope, Nikon Intensilight C-HGFI lamp, Hamamatsu ORCA ER camera (1344x1024 pixels, pixel size 6.45 µm) and Nikon Plan Apo VC 100x Oil immersion lens (Numerical Aperture 1.4), with a final optical magnification of 100x. A DAPI filter set was used for visualising the Hoechst 33258 stained nucleoid, FITC filter set for GFP, TRITC filter set for the FM 4-64 stained cell membrane and TxRed filter set for mCherry (see figures 2.18 for properties of filters and 2.19 for conditions used for imaging each fluorophore). Cells were also imaged using brightfield. Microscopy was carried out at room temperature, within 30 minutes of slides being prepared.

Microscope images were analysed using Image J. For experiments looking at the position of foci in cells, the length of cells were measured using the measuring function on Image J giving an arbitrary measurement. The same tool was used to measure the distance from the pole of the cell to a focus giving a measurement in the same arbitrary units. For further analysis the position of the focus was presented relative to cell length where 1.0 is the full cell

Filter Set	Excitation Filter	Dichroic Mirror	Barrier Filter
FITC	465-495 nm	505	515-555 nm
TRITC	540-580 nm	595	600-660 nm
Tx Red	540-580 nm	595	600-660 nm
DAPI	340-380 nm	400	435-485 nm

Figure 2.18: Filter sets used for epifluorescent microscopy

Excitation and emission wavelengths for filter sets used in epifluorescence microscopy.

Fluorophore	Filter set	Exposure time	Binning
GFP	FITC	1 s	2 x 2
FM 4-64	TRITC	1 s	No binning
mCherry	Tx Red	0.5 s	2 x 2
Hoechst 33258	DAPI	90 ms	No binning

Figure 2.19: Conditions used to image different fluorophores

Filter set, exposure times and binning for imaging GFP, FM 4-64, mCherry and Hoechst.

length and 0.5 is exactly equal distance between the two poles. This allows the results to be comparable between cells of varying lengths. For cells with a single focus, the distance from the focus to cell pole was measured to the nearest pole (figure 2.20a). For cells with 2 foci, the focus closest to a cell pole was designated "1st of 2 foci" and the distance was measured from focus to the closest cell pole. The other focus was designated "2nd of 2 foci" and its position was measured relative to the same cell pole (figure 2.20b). Data for cells with 2 foci were represented as a scatter plot, with the relative position of the "1st" focus on the x axis and relative position of the "2nd" focus on the y axis. Figure 2.20c shows approximate examples of cells represented by points falling in different areas of the graph. At least 100 cells were measured in each biological repeat, for each condition. To analyse colocalisation, the position of each focus was measured using NIS elements software (Nikon), which gave a measurement in μm. ANOVA or T-test were used in Microsoft Excel to determine whether results were significant.

2.12 Photoactivated Localisation Microscopy

All lab work and analysis carried out by Federico Garza de Leon under the supervision of Achillefs Kapanidis as described by Uphoff *et al.* 2013

2.12.1 Growth conditions and preparation of microscope slides

A single colony from a freshly restreaked plate was used to inoculate a culture in LB which was grown for 4-5 hours at 37ºC. This culture was then diluted 1:10,000 into M9 medium, supplemented with MEM amino acids and L- proline, MEM vitamins and 0.2% glycerol, and grown overnight at 37ºC. Cultures were then diluted to OD_{600} of approximately 0.025 in M9 medium and grown at 37ºC for 2 hours, at which point OD_{600} was approximately 0.1. Once

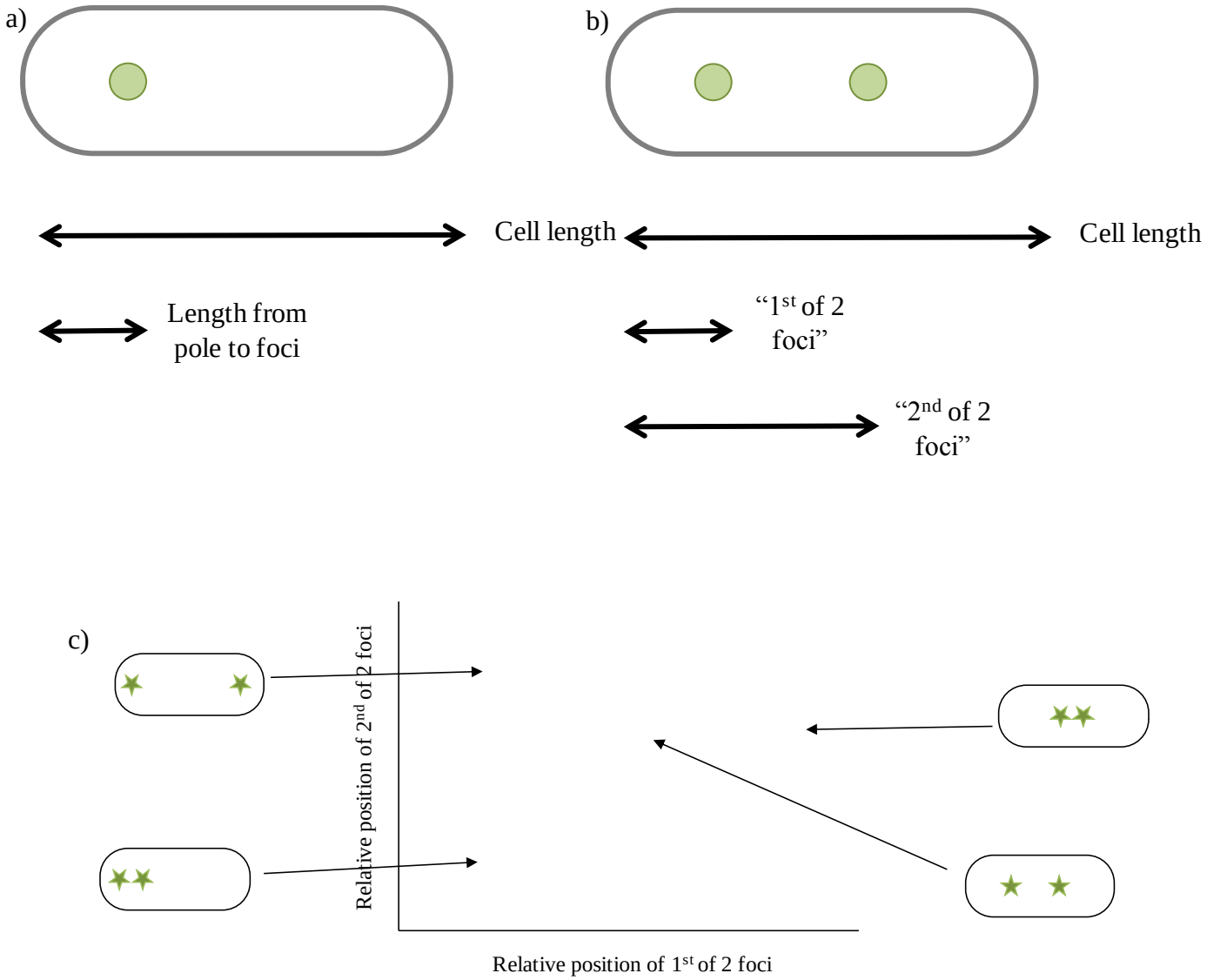


Figure 2.20: Analysis of the relative positions of foci in cells with one or two foci

a) Schematic of analysis of cells with a single focus to measure relative position of focus.

b) Schematic of analysis of cells with a two foci to measure relative position of foci.

c) Schematic of scatter graph of relative position of 1st of 2 foci plotted against relative position of 2nd of 2 foci. Arrows show the position on the graph where the points for example cells with different arrangements of foci would be plotted.

cells reached the required OD_{600}, cells were harvested by centrifugation and immobilised on agarose pads (1% agarose, molecular grade, Biorad, prepared in M9 minimal medium) between two glass coverslips (Uphoff *et al.*, 2013).

If cells were to be induced, they were immobilised in 15 µl wells with 1% polyethylenimine (Uphoff *et al.*, 2013) and images of uninduced cells were taken. Cells were induced by washing 3 times with M9 minimal media supplemented with 1 mM IPTG or 0.2% arabinose.

2.12.2 Imaging and analysis

A custom-built near TIRF (total internal reflection fluorescence) single-molecule fluorescence microscope was used to image a LacI::PAmCherry with 405 nm photoactivation laser (SLIM-561 200 mW, Oxxius) and 561 nm excitation laser (MLL-III-405 100mW, CNI). The emission was filtered by a dichroic mirror (ZT405/473/561rpc, Chroma). A 100x oil immersion objective lens (NA 1.4 Olympus) was used and PALM movies were recorded at 15.26 ms/frame for 7,500 or 10,000 frames using an EMCCD camera (Andor, iXon 897, 512x512 pixels, 114.5 nm PER PIXEL) (Uphoff *et al.*, 2013). The 405 nm laser was adjusted from 0 to 10 µW to control photoactivation of PAmCherry. The 561 nm laser was used at 3.5 mW. Microscopy was carried out at room temperature.

A custom-written MATLAB software was used to analyse images. For tracking experiments, the mean squared displacement (MSD) was calculated to analyse the diffusion of molecules. MSD is a measurement of the movement of a particle in time, taking into account the distance the particle moves in multiple time lags. The apparent diffusion coefficient (D*) was calculated to classify the diffusive behaviour of a single molecule in a single cell. As the

bacterial cell is a confined space, diffusion is non-Brownian as molecules cannot diffuse completely freely. As diffusion is not linear, a true diffusion coefficient cannot be calculated for this system. a method was developed to calculate an apparent diffusion coefficient for 2D images using the following equation (Uphoff *et al.*, 2013):

$$D^* = \frac{MSD}{4x\,\Delta t} - \frac{\sigma_{loc}{}^2}{\Delta t}$$

Where D* = Apparent diffusion coefficient ($\mu m^2/s$)

MSD = Mean square displacement

Δt = Time lag

$\sigma_{loc}{}^2$ = localisation error

One time lag is 15 ms in this case. Localisation error has been shown to introduce a positive offset to an MSD curve and therefore is subtracted in this investigation (Michalet, 2010).

An experiment with fixed cells fixed in 2.5% paraformaldehyde (v/v) gave a definition of the apparent diffusion coefficient for an immobile LacI. Any LacI molecules with a D* equal to or less than this threshold value were classified as immobile, whereas any molecules with a D* over this value were classified as mobile. For LacI::PAmCherry, this threshold was 0.1 $\mu m^2/s$.

2.13 β-galactosidase assays

Z-buffer : 75 g KCl, 0.25 g $MgSO_4.7H_2O$, 8.53 g Na_2HPO_4, 4.87 g $NaH_2PO_4.2H_2O$, 2.70 ml β-mercaptoethanol, in 1 L ddH_2O. 270 µl β-mercaptoethanol (Sigma 14.3 M 99% pure liquid) was added per 100 ml after autoclaving, on day of use.

ONPG: o-nitrophenol-β-D-galactopyranoside. 80 mg added to 100 ml Z-Buffer

Sodium carbonate: 1 M solution

The β-galactosidase enzyme is the product of the *lacZ* gene. The β-galactosidase assay (Miller, 1972) measures the activity of this enzyme in the cell, which corresponds to the activity of the *lacZ* promoter or another promoter fused to the *lacZ* gene, either on the chromosome or plasmid encoded. For each strain or plasmid to be tested, single colonies were used to inoculate 5 ml cultures with 3 biological repeats. Cultures were grown overnight with shaking at 37ºC. 100 µl of the overnight culture was used to inoculate 5 ml of the same media for each strain and condition, and the cultures were incubated at 37ºC until they reached exponential phase (OD_{650} 0.3-0.6) and the OD_{650} recorded. Each culture was lysed with 2 drops of toluene and 1% sodium deoxycholate and shaken for 20 minutes at 37ºC with the bung removed to allow the toluene to evaporate. The β-galactosidase activity was measured by adding 100 µl of the lysate to 2.5 ml of Z buffer prepared with β-mercaptomethanol and 2-Nitrophenyl β-D-galactopyranoside (ONPG) and incubated at 37ºC until a yellow colour developed. The reaction was stopped with the addition of 1 ml of 1 M sodium carbonate and the time taken for the colour change was recorded. The OD_{420} was measured. β-galactosidase activity (Miller units) was measured using the formula:

$$\beta\text{-gal activity} = \frac{1000 \times 2.5 \times 3.6 \times OD_{420nm}}{OD_{650nm} \times 4.5 \times t \times v} \text{ nmol min}^{-1}\text{mg bacterial mass}^{-1}$$

Where:

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

3.6 = final assay volume (ml)

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

t = incubation time (min)

v = volume of lysate added (in ml)

Each experiment was repeated three times and the average β-galactosidase activity and standard deviation calculated.

2.14 Constructing repressor-protein::fluorescent protein fusions

2.14.1 Construction of a plasmid encoding LacI::GFP

The LacI::GFP fusion had been made on the *E. coli* MG1655 chromosome previously, in a strain called DL01 (Lee et al., 2009). To allow for more flexibility moving the fusion between strains and to increase the copy number of LacI::GFP in the cell, the fusion was moved onto a plasmid. The cloning vector pACYC184 was used to carry the fusion. Primer D69688 was designed to bind upstream of the *lacI::gfp* fusion on the chromosome of DL01 and introduce a HindIII site. As there is little information available about the position of the *lacI* promoter and its control, the primer was designed to anneal 300 bp upstream of the predicted position of the *lacI* promoter to ensure all elements of the promoter were included. Primer D69689

was designed to bind downstream of *lacI::gfp* and the FLP site on the chromosome and introduce a BamHI site. The resulting PCR product was digested with HindIII and BamHI and the 2.2 kb fragment was gel extracted. The vector was prepared by digesting pACYC184 with BclI and HindIII, as BclI produces complementary cohesive ends to BamHI. As BclI is *dam* methylase sensitive, pACYC184 was extracted from *dam*- cells. Cutting pACYC184 with BclI and HindIII results in a fragment of 2.2 kb containing the origin of replication and chloramphenicol resistance gene and a second fragment of 2 kb containing the tetracycline resistance gene. As these fragments are very similar in size they were further digested with SalI which will only cut the tet^R containing fragment, allowing for the required fragment to be gel extracted without any risk of contamination. This vector was then treated with CIP before being ligated with the *lacI::gfp* PCR product. The resulting plasmid was pLER101 encoding LacI::GFP under the control of the *lacI* promoter. The insert was confirmed by sequencing using primers D63048 and D68556. The amino acid linker sequence between LacI and GFP is a glycine and threonine. Glycine is often used in this position to break secondary structure and ensure that the individual proteins in the fusion are folded properly.

2.14.2 Construction of a plasmid encoding MalI::mCherry

pLER101 was used as the starting point for making a plasmid encoding MalI::mCherry. Initially GFP was exchanged for mCherry. Primers were designed to amplify mCherry from pmCherry-N1 with D71000 annealing upstream and introducing a KpnI site and D71001 annealing downstream and introducing a MfeI site. The PCR product generated was digested with KpnI and MfeI and purified by gel extraction. pLER101 was digested with KpnI and MfeI, gel extracted and CIP treated before being ligated with the mCherry PCR. The resulting plasmid was pLER102 encoding LacI::mCherry under the control of the *lacI* promoter. The insert was confirmed by sequencing using primer D69689.

To exchange *lacI* and the *lacI* promoter for *malI* and its promoter, primers were designed to amplify *malI* from the plasmid pACYCMalI. Primer D63433 anneals upstream of *malI* and its promoter as defined by Lloyd *et al.* (Lloyd et al., 2010) and introduces a HindIII site. D71192 anneals at the end of the *malI* gene introducing a KpnI site and omitting the final codon of *malI*, so transcription is not terminated and runs through to the mCherry gene. This PCR product is digested with HindIII and KpnI and purified by gel extraction. pLER102 was digested with HindIII and KpnI, treated with CIP and gel extracted before being ligated with the *malI* insert. The resulting plasmid was pLER104 encoding MalI::mCherry under the control of the *malI* promoter. The insert was confirmed using the primer D63048. However, this plasmid contains two MalI DNA sites in the *malI* promoter which is not ideal for the experiments planned.

pJW15 Δ100 is a plasmid that can be used to fuse promoter-less genes to the *melR* promoter. Genes are cloned into the plasmid by removing the ATG of the gene and replacing it with an NsiI site. The ATG of the NsiI restriction site (ATGCAT) acts as a new start codon. Primer D72022 was designed to anneal after the ATG of *malI::mCherry* in pLER104 and introduce a NsiI site. Primer D71850 was designed to anneal downstream of *malI::mCherry* in pLER104 and introduce a HindIII site. The PCR product was digested with NsiI and HindIII and purified by gel extraction. Vector was prepared from pJW15 Δ100 digested with NsiI and HindIII, gel extracted and treated with CIP. Vector and insert were then ligated and the resulting plasmid was pLER105 encoding MalI::mCherry under the control of the *melR* promoter. The insert was confirmed by sequencing using primer D71192.

To transfer the fusion into a lower copy number plasmid, which would give a better concentration of MalI::mCherry in the cell, the fusion was transferred back into pACYC184. Primer D77566 was designed to anneal upstream of the *melR* promoter fused to *malI::mCherry* in pLER105, and introduce a HindIII site. Primer D77567 was designed to anneal downstream of *malI::mCherry* and introduce a MfeI site. The PCR product was digested with HindIII and MfeI and purified by gel extraction to make the insert. pLER102 was digested with HindIII and MfeI, gel extracted and treated with CIP before being ligated with the *malI::mCherry* insert. The resulting plasmid was pLER108 encoding MalI::mCherry under the control of the *melR* promoter. The insert was confirmed by sequencing using primer D63048. The amino acid linker sequence between LacI and GFP is a glycine and threonine.

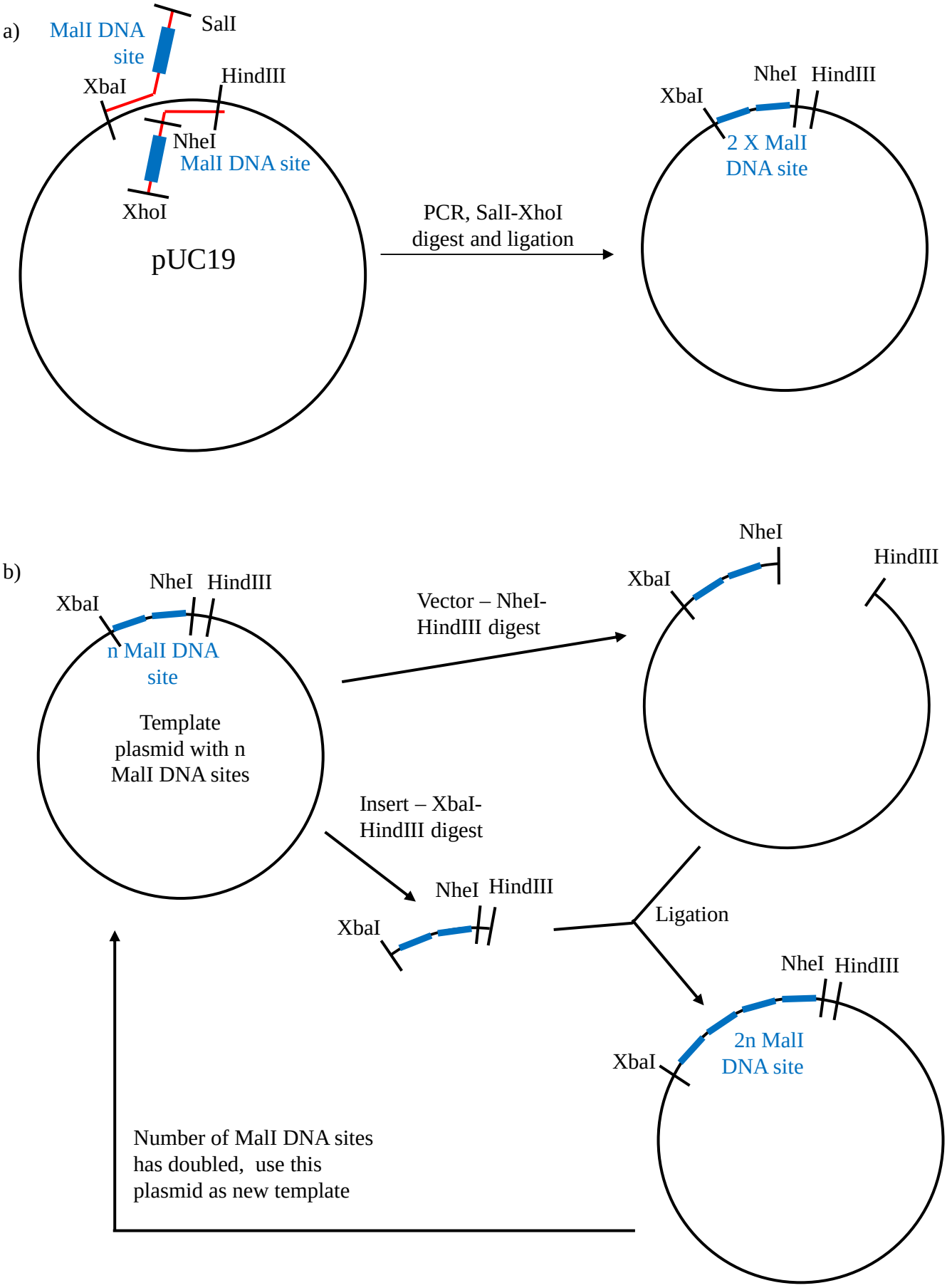
2.15 Constructing a multiple MalI DNA site array

The method for constructing the multiple MalI DNA site array was based on the method used to create LacI DNA site repeats as described by (Lau et al., 2003). An overview of the construction of the plasmid containing the multiple DNA sites is shown in figure 2.21. pUC19 was used for the construction, particularly the multiple cloning site of the plasmid. Primers were designed to anneal at the XbaI and HindIII sites in the multiple cloning site, facing in opposite directions so they would amplify the whole plasmid apart from the region between the XbaI and HindIII sites. At the 5' end of the XbaI primer a SalI site was introduced and at the 5' end of the HindIII primer an XhoI site was introduced. Between the restriction sites on each primer, a DNA site for MalI was introduced using the sequence of the DNA site in the *malX* promoter. Finally a NheI site was added to the HindIII primer between the HindIII site and MalI DNA site. The resulting spacing between MalI DNA sites was designed to be similar to that of the LacI DNA site array, which has been used previously.

Figure 2.21: Construction of a multiple MalI DNA site array

a) Primers were used to make a linear PCR product which was digested with SalI and XhoI, then ligated to make a pUC19 derived plasmid with two MalI DNA sites.

b) This plasmid was then used to make both the vector and the insert for the next cloning step, doubling the number of MalI DNA sites. This step was repeated until the required number of repeats was achieved.


a)
MalI DNA site
SalI
XbaI
HindIII
NheI
MalI DNA site
XhoI
pUC19
PCR, SalI-XhoI digest and ligation
NheI
HindIII
XbaI
2 X MalI DNA site
b)
XbaI
NheI
HindIII
n MalI DNA site
Template plasmid with n MalI DNA sites
Vector – NheI-HindIII digest
NheI
HindIII
XbaI
Insert – XbaI-HindIII digest
NheI
HindIII
XbaI
Ligation
NheI
HindIII
2n MalI DNA site
XbaI
Number of MalI DNA sites has doubled, use this plasmid as new template

These primers were used to amplify pUC19 by PCR. The linear product formed was digested with SalI and XhoI, which generate compatible cohesive ends, and ligated to create a new plasmid containing two MalI DNA sites with a hybrid SalI/XhoI site in between, a XbaI site on one side and NheI and HindIII sites on the other side. This plasmid was selected for by blue-white selection on X-gal plates.

Once this plasmid has been made it is used to produce both vector and insert for the following step. To make the vector, the plasmid is digested with NheI and HindIII and treated with CIP. To make the insert, the plasmid is digested with XbaI and HindIII. As XbaI and NheI also create compatible cohesive ends the insert can be ligated into the vector creating a plasmid containing four MalI DNA sites with a NheI/XbaI hybrid site in between. This plasmid still has an XbaI site on one side of the MalI DNA sites and NheI and HindIII on the other, but the number of MalI DNA sites has doubled. This plasmid can then be used for a further cloning step. The method of construction can be repeated to double the number of MalI DNA site with each step. Successful transformants are selected for by colony PCR using primers that anneal outside of the *lacZα* gene on pUC19. PCR reactions were separated by agarose gel electrophoresis, and if the cloning step had been successful an increase in size was seen in the PCR products. This was repeated until an array of 20 MalI DNA sites had been produced. The sequence of the multiple MalI DNA site array is shown in figure 2.22

2.16 Insertion of multiple LacI DNA sites adjacent to the *araBAD* promoter

Gene doctoring donor plasmids were derivatives of pJB10 adapted by Jack Bryant (unpublished) from plasmid pKH3 (Kerry Hollands, unpublished). pJB10 carries a *lacZ* promoter fusion to GFP with a *kan* cassette upstream all flanked by homology regions to *thiQ*

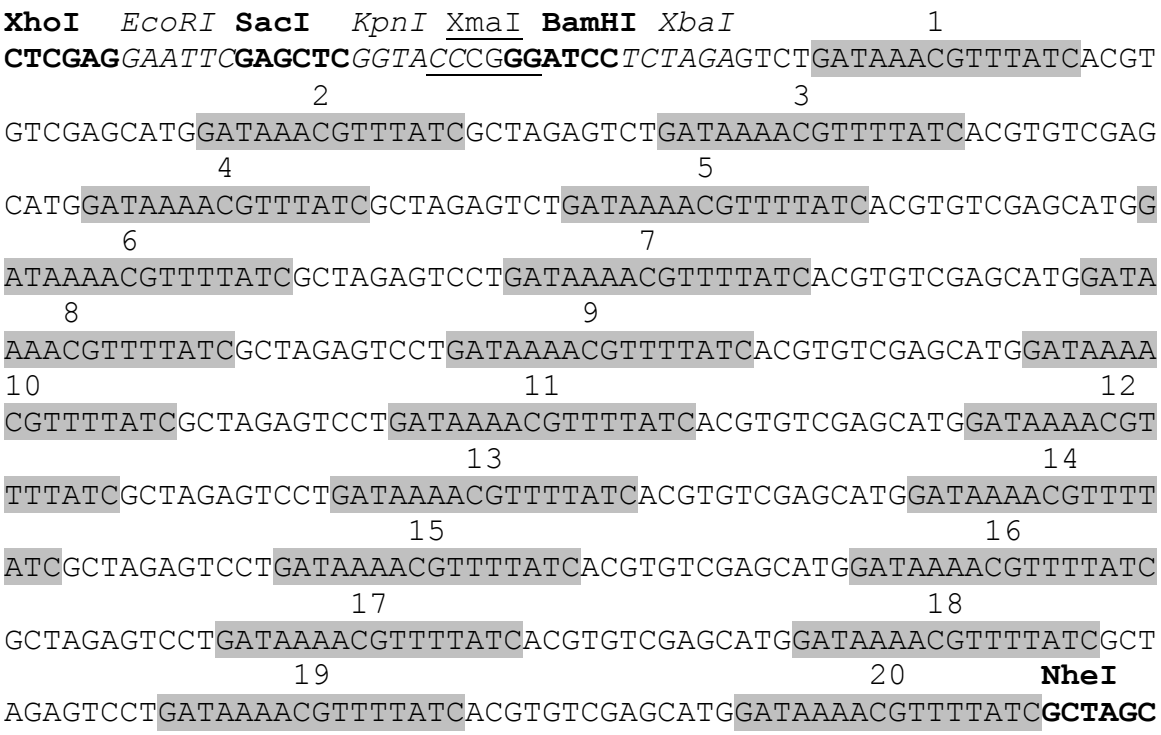


Figure 2.22: Sequence of the multiple MalI DNA site array

Sequence of the multiple MalI DNA site array, MalI DNA sites are shown in grey and numbered. Also shown are restriction sites that can be used for cloning.

and *yabI* and restriction sites for SceI (see figure 2.3). The LacI DNA site repeat sequence (sequence shown in figure 2.23) was cloned into this plasmid to create a donor plasmid for gene gorging. Sections of derivatives of this plasmid, between the SceI restriction sites, were then cloned into pDOC-C to include the presence of the *sacB* gene and allow for sucrose counter selection in the gene doctoring protocol. The resulting plasmid was used to insert the multiple LacI DNA sites at the *araBAD* locus between the *thiQ* and *yabI* genes (see figure 2.24).

2.16.1 Construction of plasmids pLR7 and pLR8

pLR1 was constructed by cloning a fragment containing six copies of the LacI DNA site from pPM461 into pJB10. pJB10 already contained the homology regions required for an insert adjacent to the *araBAD* locus (see figure 2.3) the sequence of which is shown in figure 2.25. However, it also contained a fusion of the *lacZ* promoter and GFP which needed to be replaced by multiple LacI DNA sites. Due to the repeating nature of the LacI DNA site repeats, using PCR to amplify the fragment would have a high risk of errors. There were several unique restriction sites flanking the repeat sequence so to create the insert required for cloning the fragment into the donor plasmid pPM461 was digested with BglII and NheI. The 240 bp fragment containing six copies of the LacI DNA site was purified by gel extraction from a 0.8% agarose gel, using the QIAquick gel extraction kit, and cloned into pJB10 which had also been digested with BglII and NheI to create plasmid pLR1. The insert was confirmed by sequencing using primer D69233.

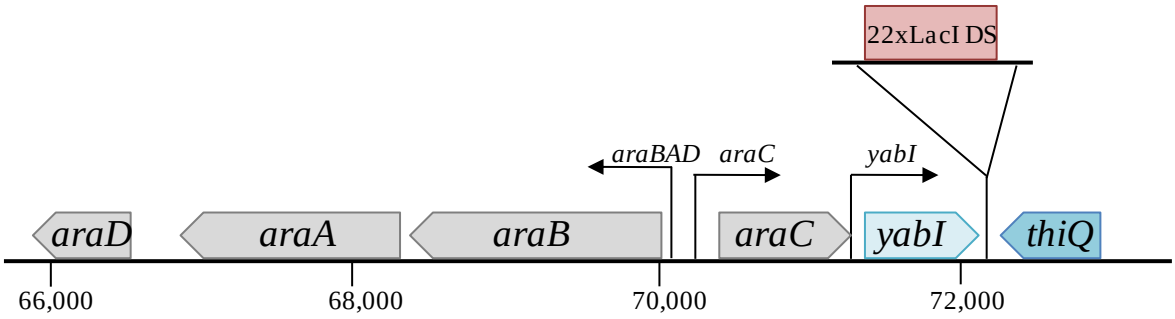
Due to the presence of two BglII site in pJB10 (see figure 2.3) the *kan* cassette was removed from the plasmid in the cloning process. The ampicillin resistance cassette was used to select

BglII *SalI* **XbaI** 1
AGATCT*GTCGAC***TCTAGA**GTCCTGCAATTGTTATCCACTCACAATTGTCTTTATTTAATT
2 3
GTTATCCGCTCACAATTACCTTGTTTCAATTGTTACCGCTCACAATTAGCCATCTGTGCT
4 5
AGAGCTCCTCCTCTAATTGTTATCCGCTCACAATTACCTCATCGCAATTGTTATCCGCTC
6
ACAATTCCAGAGGTGCAATTGTTATCCGCTCACAATTTTCCTCCGTCTGCTAGAGCTTTC
7 8
CCCCGAATTGTTATCCGCTCACGATTTTGTTTCTCCAATTGTTATCGCTCACAATTGGTT
9
CTTACAAATTGTGATCCGCTCACAATTCCTTTTCCTATGCTAGAGATATTTCCCAAATTG
10 11
TTATCCGCTCACAATTGTTTTGTCTTAATTGTTTTCTACTTCAATTGTTATCCGCTCACA
12
ATTTGCTATTTTCTGCTAGAGTCTTGTGGCCAATTGTTATCCGCTCACAATCCTTAGTAA
13 14
TCAATTGTTATCCGCTCACAATTTTCGCCACCAATTGTTATCCGCTCACAATTCGCTCG
15
AACCTGCTAGAGGTTTTGTCCAAATTGTTATCCGCTCACAATTGTCCTATGCAAATTGTT
16
ATCCGCTCACAATTTTATTTTAATTCACAATTCACCTTAACTTGCTAGAGTCCTGTCATT
17 18
AATTGTTTCCGCTCACAATTTACCTGTTGTAATTGTTATCCGCTCACAATTTCTTGCGTG
19
TAATTGTTATCCGCTCACAATTTTTTCTTCCCTGCTAGAGTATCCTACCCAATTGTTATC
20 21
CGCTCACAATTGTTTCGGTGTAATTGTTAACCGCTCACAATTATTTACTTCGAATTGTTA
22 **NheI**
TCCGCCACAATCGCTACGGTTT**GCTAGC**

Figure 2.23: Sequence of the multiple LacI DNA site array

Sequence of the multiple LacI DNA site array, LacI DNA sites are shown in grey and numbered. Also shown are restriction sites that can be used for cloning.

a)



b)

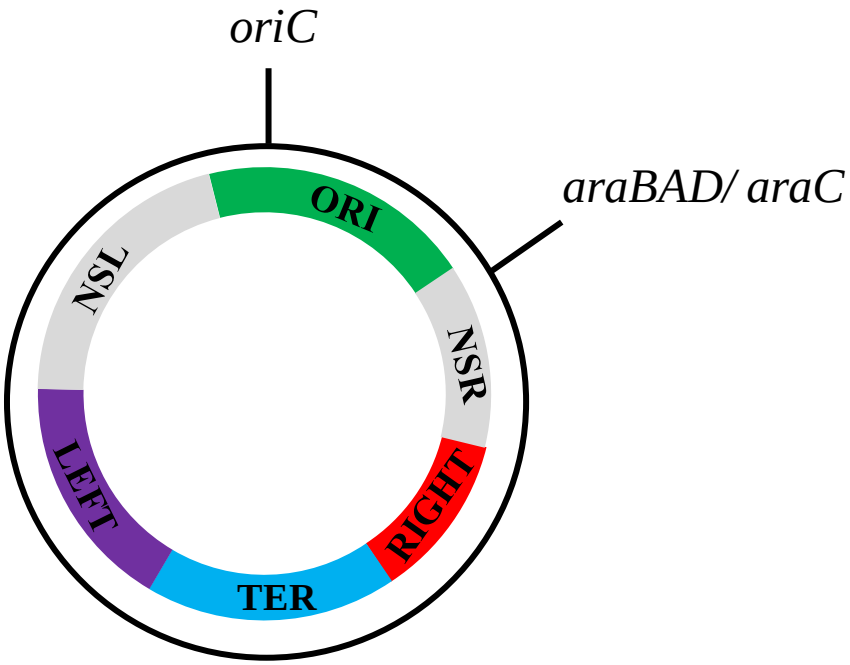


Figure 2.24: Location of multiple LacI DNA sites inserted adjacent to the *araBAD* locus

a) 22 copies of the LacI DNA site (LacI DS) were inserted onto the chromosome between convergent genes *thiQ* and *yabI*. Adjacent to these genes is the AraC controlled *araBAD* promoter which will be followed using LacI::GFP bound to the multiple operators. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Position of the *araBAD* locus relative to *oriC* and macrodomains.

Figure 2.25: DNA sequence of the chromosomal region adjacent to *araBAD* showing the homology regions and primers used to make chromosomal insertions

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *thiQ* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *yabI* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined.

araB ←

CATcgtttcactccatccaaaaaaacgggtatggagaaacagtagagagttgcgataaaaagcgtcaggtaggat

ccgctaatcttatggataaaaatgctatggcatagcaaagtgtgacgccgtgcaaataatcaatgtggacttttc

tgccgtgattatagacacttttgttacgcgtttttgtcatggctttggtcccgctttgttacagaatgcttttaa

taagcggggttaccggttgggttagcgagaagagccagtaaaagacgcagtgacggcaatgtctgatgcaatatg

araC →

gacaattggtttcttctctgaatggtgggagtatgaaaagtATGGCTGAAGCGCAAAATGATCCCCTGCTGCCGG

GATACTCGTTTAACGCCCATCTGGTGGCGGGTTTAACGCCGATTGAGGCCAACGGTTATCTCGATTTTTTTATCG

ACCGACCGCTGGGAATGAAAGGTTATATTCTCAATCTCACCATTCGCGGTCAGGGGGTGGTGAAAAATCAGGGAC

GAGAATTTGTCTGCCGACCGGGTGATATTTTGCTGTTCCCGCCAGGAGAGATTCATCACTACGGTCGTCATCCGG

AGGCTCGCGAATGGTATCACCAGTGGGTTTACTTTCGTCCGCGCGCCTACTGGCATGAATGGCTTAACTGGCCGT

CAATATTTGCCAATACGGGTTTCTTTCGCCCGGATGAAGCGCACCAGCCGCATTTCAGCGACCTGTTTGGGCAAA

TCATTAACGCCGGGCAAGGGGAAGGGCGCTATTCGGAGCTGCTGGCGATAAATCTGCTTGAGCAATTGTTACTGC

GGCGCATGGAAGCGATTAACGAGTCGCTCCATCCACCGATGGATAATCGGGTACGCGAGGCTTGTCAGTACATCA

GCGATCACCTGGCAGACAGCAATTTTGATATCGCCAGCGTCGCACAGCATGTTTGCTTGTCGCCGTCGCGTCTGT

CACATCTTTTCCGCCAGCAGTTAGGGATTAGCGTCTTAAGCTGGCGCGAGGACCAACGCATTAGTCAGGCGAAGC

TGCTTTTGAGCACTACCCGGATGCCTATCGCCACCGTCGGTCGCAATGTTGGTTTTGACGATCAACTCTATTTCT

CGCGAGTATTTAAAAAATGCACCGGGGCCAGCCCGAGCGAGTTTCGTGCCGGTTGTGAAGAAAAAGTGAATGATG

TAGCCGTCAAGTTGTCATAAttggt**aac**gaatcagacaattgacggcttgacggagtagcatagggtttgcagaa

yabI →

tccctgcttcgtccatttgacaggcacatATGCAAGCATTGCTGGAACACTTTATTACCCAATCCACCGTGTATT

CATTGATGGCGGTGGTGTTGGTGGCCTTTCTGGAGTCGCTGGCGCTGGTCGGTTTGATTCTACCCGGTACGGTGC

TGATGGCGGGGCTGGGAGCGCTGATTGGCAGCGGCGAGTTAAGTTTCTGGCACGCCTGGCTGGCAGGGATTATTG

GCTGCTTGATGGGCGACTGGATTTCTTTCTGGCTGGGTTGGCGTTTTAAAAAGCCGTTGCATCGCTGGTCATTTC D69748

TGAAGAAAAACAAAGCACTACTTGATAAAACTGAACATGCGTTGCATCAACACAGCATGTTCACCATTCTGGTCG D69233

GTCGTTTTGTTGGCCCGACGCGTCCGCTGGTGCCAATGGTGGCGGGAATGCTGGATCTGCCGGTGGCTAAATTTA

TTACGCCGAATATTATCGGCTGCCTGCTGTGGCCGCCGTTTTACTTCCTGCCAGGGATTCTGGCGGGCGCGGCGA

TCGATATTCCTGCCGGAATGCAGAGCGGTGAGTTTAAATGGTTGCTGCTGGCAACAGCGGTGTTTTTGTGGGTTG HR2

GTGGCTGGCTGTGCTGGCGGTTATGGCGCAGCGGTAAAGCGACTGACCGTTTGAGTCATTATTTGTCCCGCGGTC

GTTTGTTGTGGCTGACGCCGTTGATTTCTGCCATCGGCGTGGTGGCGCTGGTGGTGTTAATTCGCCACCCGTTGA

TGCCGGTGTATATCGATATTTTGCGTAAAGTGGTTGGGGTTTAGgagatagtcttgtgcgggttgcctgatgcga D69234

cgcttgccgcgtcttatcaggcctacaaaacgcactacccgtaggtcggataaggcgttcacgccgcatccgaca D69232

gtgcataCTAACCCGTAATCCCCAATAGTGCCGAAGCACTCGCCTTACCGCTCAACAACTCATTGGTCATACCCT

GCCAGGCGATGCGCCCGTCGGCGACTACTACCGAGCGCGTGGCGATCCGCGCCGCATCTTCCACGCTGTGCGACA

CCATCAATAGCGTCATTTTTTGCTGCTGGCAGCTCGTGCTCACCAGCGTCAACATCTCCTGACGTAACGCCGGAT HR1

CGAGCGCAGAGAACGGTTCATCGAGCAATAAAATCGGCTGTTCGCGTACCAGACAACGCGCTAACGCCACTCGCT

GTCGCTGACCGCCGGAAAGCTCGCCCGGTAACCGCGCCATTAAATTATCAATCCCCATCTGGCGGGCGATAGCGT

← *thiQ* D69231

GCATTTTCCCCTGCTGTACCGCGTTCAGTTTCAATCCCGGATTTAGCCCCAGCCCGATGTTCTGTGCGACCGTCA D69747

for successful transformants, but the *kan* cassette needed to be reinserted for when the plasmid is used in the gene doctoring process. Primers D69988 and D68556 were used to amplify the *kan* cassette from pJB10. D69988 anneals upstream of the *kan* cassette and includes one BglII site. The reverse primer, D68556, anneals to the start of the GFP gene and amplifies back over the second BglII site. The resulting product was digested with BglII and a small fragment of approximately 200 bp was lost. A larger fragment of approximately 1.1 Kb was cloned into pLR1, which had also been digested with BglII, resulting in plasmid pLR6. Only candidates containing the *kan* cassette in the correct orientation would be resistant to kanamycin as the upstream BglII site was found to be in the promoter for the *kan* cassette.This plasmid does not contain the *sacB* gene required for sucrose counter selection in gene doctoring. This is provided in pDOC-C along with amp^R flanked by SceI restriction sites. pLR6 was digested with SceI and the smaller fragment containing the two homology regions, *kan* cassette and multiple LacI DNA sites was purified by gel extraction from a 0.8% agarose gel using the QIAquick gel extraction kit. This was ligated with pDOC-C, which had also been digested with SceI and treated with CIP, to produce plasmid pLR7. Resulting colonies were checked for kanamycin and ampicillin resistance and sucrose sensitivity.

pLR8 was constructed in the same way as pLR7, making intermediates pLR2 and pLR6, to produce a donor plasmid for inserting 22 copies of the LacI DNA sites at the *araBAD* locus. This is an insert of 800 bp.

2.16.2 Construction of strains LR06 and LR18, carrying LacI DNA sites adjacent to *araBAD*

Strains LR06 and LR15 were constructed by inserting 22 and six repeats of the LacI DNA site respectively adjacent to the *araBAD* promoter. The strain modified was KH001, which has

had the *lacZ* promoter and the DNA binding domain of LacI deleted, meaning there would be no wild type LacI present in the cell and no LacI DNA sites other than the ones inserted by gene doctoring. KH001 was transformed with pACBSR and pLR8. Gene doctoring was used to insert 22 LacI DNA site repeats and the *kan* cassette onto the chromosome of KH001 at the *araBAD* locus as described in section 2.9.2. The insertion was confirmed by colony PCR using primers D69747 and D69748 and sequencing using primer D69233, resulting in strain LR04. The *kan* cassette was removed from the strain using FLP recombinase provided on the plasmid pCP20.

To visualise the LacI DNA site repeats in LR04 LacI::GFP needs to be supplied. To express LacI::GFP at a low level a second chromosomal modification was made. Using P1 transduction LacI::GFP was transferred from the MG1655 LacI::GFP strain into LR04 (Maritoñi Sánchez Romero) to make strain MSR02. The *kan* cassette was then removed from this strain to make strain LR06.

To insert six copies of the LacI DNA site onto the chromosome, DL02 was transformed with pACBSR and pLR7 and gene doctoring was used to insert six LacI DNA sites and the *kan* cassette onto the chromosome at the *araBAD* locus. The insert was confirmed by colony PCR using primers D69747 and D69748 and sequencing using primer D69233, resulting in strain LR15. The *kan* cassette was then removed from this strain to make LR18.

Sequencing confirms that the correct number of DNA sites are inserted onto the chromosome at this stage. No instability of either the LacI or MalI DNA site arrays was observed after insertion onto the chromosome. Colony PCR at a later date never showed reduction in the size

of the insert and microscopy results did not indicate that the number of DNA sites had reduced.

2.17 Insertion of Multiple MalI DNA sites adjacent to the *araBAD* promoter

2.17.1 Construction of plasmid pLR19

To insert multiple copies of the MalI DNA site onto the chromosome, the MalI DNA site array described in section 2.15 needs to be transferred into a gene doctoring donor plasmid. Due to the lack of unique restriction sites available upstream of the multiple LacI DNA sites in pLR8 (figure 2.8) the multiple MalI DNA sites could not be cloned directly into pLR8 to make a new donor plasmid. Instead pJB32 was used to make an intermediate plasmid. pUCMal20 was digested with EcoRI and NheI to make a fragment of approximately 600 bp which was purified by gel extraction. pJB32 was also digested with EcoRI and NheI, treated with CIP and the vector fragment purified by gel extraction. This vector was then ligated with the insert resulting in plasmid pLR11. In a second round of cloning, pLR11 was digested with NheI and AgeI cutting out the *kan* cassette and MalI DNA site array. This insert of 2 kb was purified by gel extraction. pLR8 was also digested with NheI and AgeI resulting in the loss of a fragment containing the *kan* cassette and multiple LacI DNA site array. The vector was treated with CIP and purified by gel extraction, before being ligated with the insert containing the *kan* cassette and multiple MalI DNA sites. The resulting plasmid was pLR19, containing 20 copies of the MalI DNA site and homology regions for insertion adjacent to the *araBAD* promoter.

2.17.2 Construction of strain LR17, carrying 20 MalI DNA sites adjacent to the *araBAD* promoter

Strain LR17 was constructed by inserting 20 copies of the MalI DNA site adjacent to the *araBAD* promoter. MG1655 was co-transformed with pACBSR and pLR19. Gene doctoring was used to insert 20 copies of the MalI DNA site and the *kan* cassette onto the chromosome adjacent to the *araBAD* promoter as described in section 2.9.2. The insertion was confirmed by colony PCR, using primers D69747 and D69748, and sequencing using primer D69233, resulting in strain LR16. The *kan* cassette was removed from the strain using FLP recombinase provided on the plasmid pCP20. The removal of the *kan* cassette was confirmed by colony PCR, using primers D69747 and D69748, and sequencing using primer D69233, resulting in strain LR17. The number of DNA sites inserted onto the chromosome was confirmed by sequencing using primer D69231. This strain was then transformed with pLER108 (figure 2.13) before the position of the insert in the cell was localised by fluorescence microscopy using FROS.

2.18 Insertion of multiple MalI DNA sites adjacent to the *araFGH* promoter

Due to the method of constructing the LacI DNA site or MalI DNA site repeats there are a high number of restriction sites present. This means donor plasmids containing the repeat sequences could not be used as a template for future donor plasmids for new target sites. Instead pJB32 was used, and three separate cloning steps carried out to make each finished donor plasmid (see figure 2.26).

2.18.1 Construction of plasmid pLR17

To reduce the chance of affecting expression of *araF* with the insertion onto the chromosome, the promoter region was avoided. Upstream of the *araF* promoter are convergent genes *ftnB*

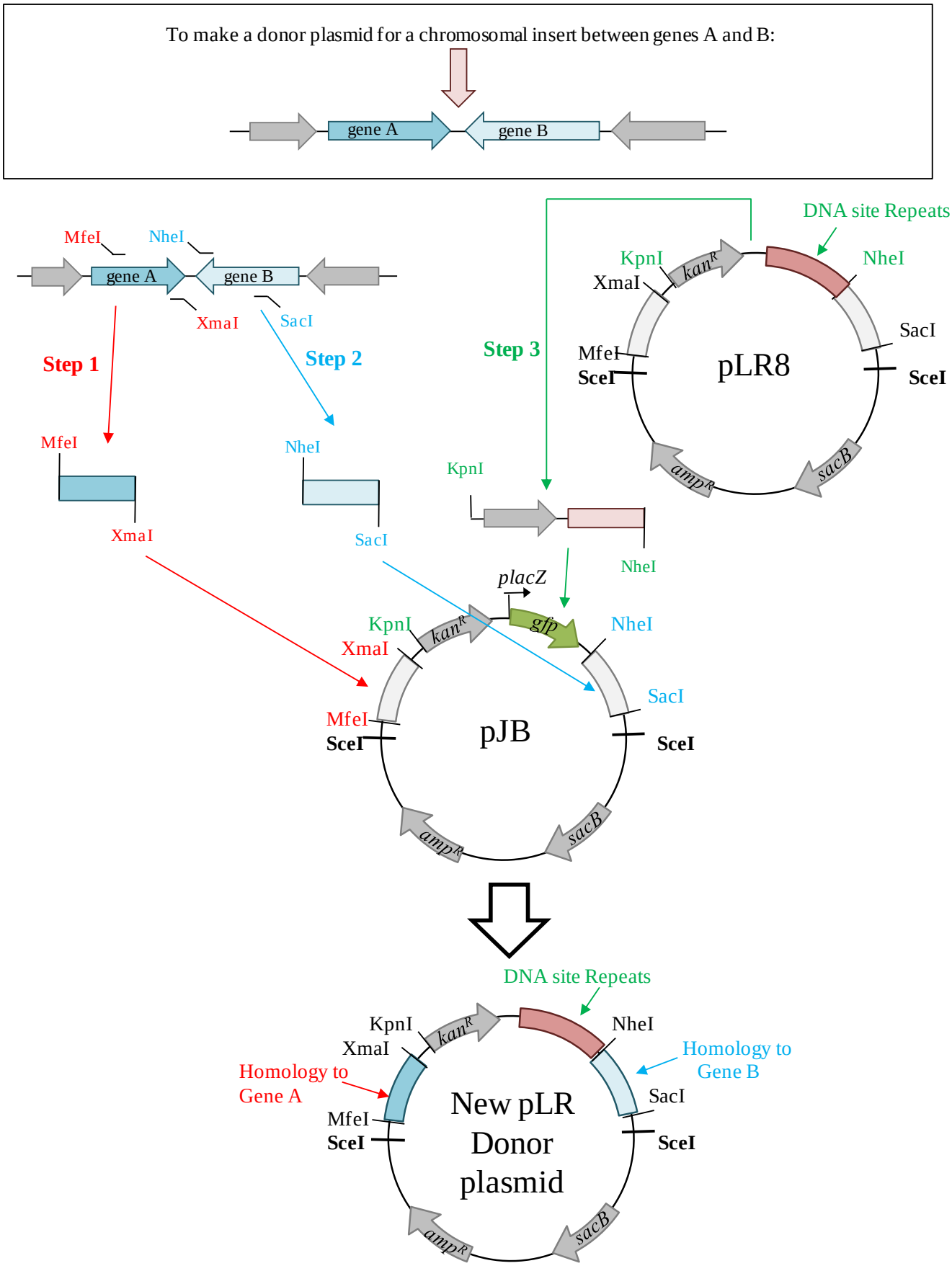
Figure 2.26: General protocol for making donor plasmids for gene doctoring

All donor plasmids for making an insert of multiple LacI or MalI DNA sites were made using the same basic protocol.

1. Homology region corresponding to gene A is amplified and cloned into plasmid pJB using restriction enzymes MfeI and XmaI.

2. Homology region corresponding to gene B is amplified and cloned into the same plasmid using restriction enzymes NheI and SacI.

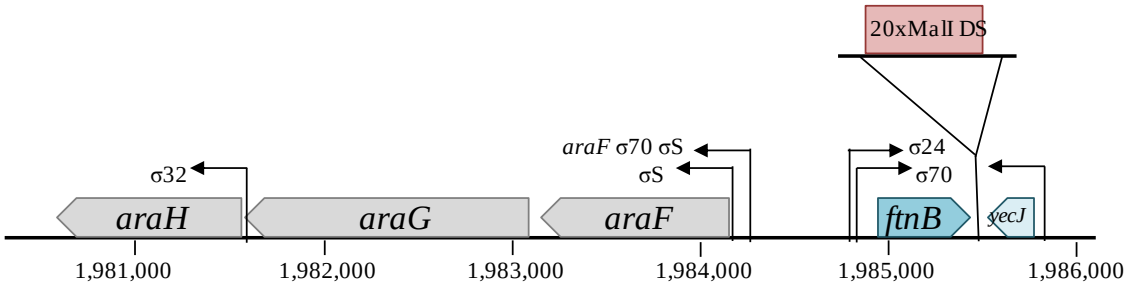
3. A plasmid already containing the multiple LacI or MalI DNA site array used to make the final insert. The plasmid was cut with KpnI and NheI to create a fragment containing the LacI DNA sites or MalI DNA sites and the *kan* cassette. This was then cloned into the plasmid containing the two new homology regions replacing *placZ::gfp*.


To make a donor plasmid for a chromosomal insert between genes A and B:
gene A
gene B
MfeI
NheI
gene A
gene B
XmaI
SacI
Step 1
Step 2
Step 3
MfeI
XmaI
NheI
SacI
KpnI
NheI
DNA site Repeats
KpnI
XmaI
kanR
NheI
SacI
MfeI
SceI
pLR8
SceI
ampR
sacB
placZ
gfp
pJB
DNA site Repeats
New pLR Donor plasmid
Homology to Gene A
Homology to Gene B

and *yecJ*, so the homology regions were designed to make the insert between them, where there would be less chance of affecting gene expression of any genes. This insert is approximately 1.3 kb from the *araF* promoter (see figure 2.27). Primers D74949 and D74950 were used to amplify the region illustrated in figure 2.27 labelled homology region 1 (HR1, figure 2.28) giving a 500 bp PCR product. D74949 introduces a MfeI restriction site upstream of the homology region and D74950 introduces a XmaI site downstream. The PCR product and pJB32 were digested with MfeI and XmaI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pLR15, which was confirmed by sequencing with D58794. The second homology region (HR2, figure 2.28) was then amplified using primers D74951 and D74952 to give a 500 bp PCR product. D74951 introduces a NheI site upstream of the homology region and D74952 introduces a SacI site downstream. The PCR product and pLR15 were digested with NheI and SacI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pLR16 which was confirmed by DNA sequencing with D58793.

pLR11 was used to make the insert for the final cloning step to avoid errors being introduced into the repeat sequence during PCR. pLR11 and pLR16 were both digested with AgeI and NheI. The 2 kb fragment containing the *kan* cassette and MalI DNA site repeats was purified by gel extraction from pLR11. From pLR16 the larger fragment of the vector was purified and treated with CIP. This was ligated with the insert from pLR11 to make plasmid pLR17, containing 20 copies of the MalI DNA site and homology regions for insertion adjacent to the *araFGH* promoter, which was confirmed by DNA sequencing with D74952.

a)



b)

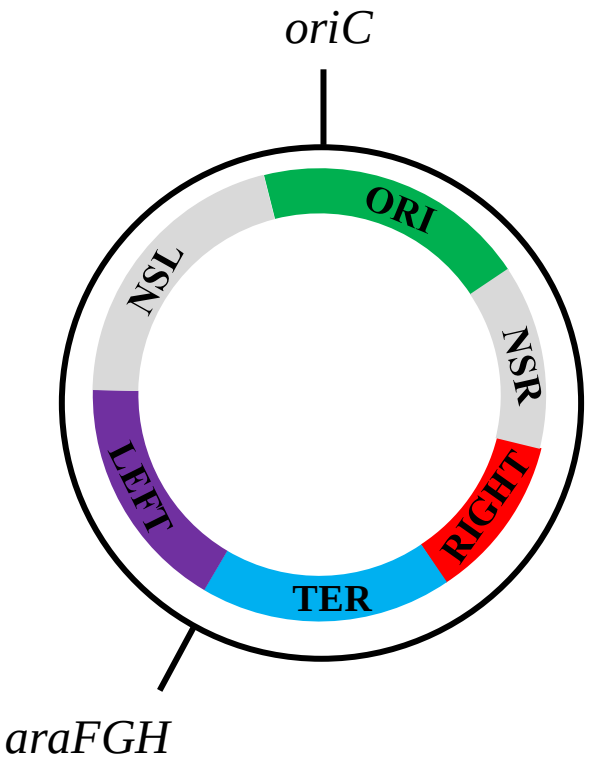


Figure 2.27: Location of multiple MalI DNA sites inserted adjacent to the *araFGH* promoter

a) 20 copies of the MalI DNA site (MalI DS) were inserted onto the chromosome between convergent genes *ftnB* and *yecJ*. Adjacent to these genes is the AraC controlled *araFGH* promoter which will be followed using MalI::mCherry bound to the multiple operators. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Position of the *araFGH* locus relative to the *oriC* and macrodomains.

Figure 2.28: DNA sequence of the chromosomal region adjacent to *araFGH* showing the homology regions and primers used to make chromosomal insertions

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *ftnB* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *yecJ* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined.

AAAACCGAGCTTCAGGTTCTCCGCCATAGCGGATTGTGACATAACGGCTGCCAGACCAATGGCTGCCAGGGCTTT

araF ←

AGTAAATTTGTGCATggttctctccagctttagtgtcgttttgtgtagggcaaaaacgaatgacattcgttaaat

taatcggaaaacaaagcattacctttttaactaaaagataagtgactgtgttgacatagttttagcgagaaattaa

ttctccataggagagcaatatcacatcgcagaattacagtgagaacgtgcataaatttagcgggaaaagacataa

gggaaagccaatttgtcagacaaattgtcgaatgcacagcagattaatccataagattagcctggaaatccttgt

tgtctttggtacccatgcgggatgtcttctttttaaccagtcaataggccgcattacctggcgttgagtttttga

aatggtgtaataaccgcaactcaaagatgtggaaaatgcacgtcattcatttcgtcattaattatcactgtgctc

attaattaacagaacacgtataatgagagccatctcgcaaaaatgaaaaaacgttttataaaatcatcacttcat

catgaattcaaattcattgattaatatcaacaagatacaaaaagcactatcattaaaattcattgcagttacatt

gatttcatcaatgaaatgtaaaaatatataaacttgatgatttaagcattttcttatacccgttcagacgttatt

cttatttcagatcatcgtcagaattgactccacgatcacatttcggaccggcagaaaggaattattctgcaaaca

ftnB →

gtaattatggtgttttgatttatcttgcacctctccacttctggatataaggatattaggtATGGCAACCGCTGG D75296

AATGCTTCTCAAACTCAACTCTCAAATGAACCGCGAGTTTTACGCATCCAATCTCTACCTTCACCTGAGTAACTG D74949

GTGTTCTGAACAGAGTCTGAACGGCACCGCCACTTTCCTTCGCGCCCAGGCACAGAGTAATGTGACCCAAATGAT

GCGCATGTTTAACTTTATGAAGAGTGTCGGCGCTACCCCCATCGTTAAAGCCATTGATGTTCCCGGTGAAAAACT

GAACTCTCTGGAAGAACTGTTCCAAAAAACGATGGAAGAATACGAGCAACGTTCTAGTACGTTGGCACAGTTAGC HR1

CGATGAAGCGAAAGAACTGAATGATGATTCAACCGTCAATTTCCTGCGCGATCTGGAAAAAGAACAGCAGCATGA

TGGTCTGTTGCTGCAAACCATTCTTGATGAAGTGCGCAGTGCGAAACTTGCGGGTATGTGCCCTGTGCAGACCGA

CCAACATGTTCTGAATGTCGTGTCACACCAGCTGCATTGAtcatcatcggcgctaatgcattgcgccgatgaagg

ttttgagaaaccgctgcctcatctgtttgaagcagcggtttttTTAATGGGATTCACCCTGTGGGGTAAACTTGA D74950

D74951

GTTCAATAAGCGCGATGGCTTTTTGGATTGCCCGCATGGTGACCGGGTCTGCGGCGGCGGGATGGTTAGTAAAGT

CGATATTCTTCAGCTGACTGGACATTTTTTCACGAACTTCAACGGGCGCGATTACATCGAGAACATCCAGAATTT

yecJ ←

GTTTGATAACCAACTGGCAAGCAACCACATCAGAAACCAATTCCTGATCGGCATTCAGCGGCTGGGACATcgtaa HR2

actcctgatagcattttgaaagccgttatagtagcgacttcacatcttcagcgatagtcacatccaccgtcatca

ggacacaaaaaaacctgccggagcaggtttttgttatcggaacatattgcctggcggtacgtctttgaacgtct

tgcaatagttattgaacatactttcaggattttgcgcagtttcatcgcggcactccgaccatttgttatacagg

tgttattgtctttgcgcttataatatgacaaccatcacaaaaatcaatctttatgtgatacaaatcacataaata D74952

ccccttttaatgttataaaaatgataatcaaaaaacagccccctatttctgacacctacagatggcaagaaatag

cgcctgccaggcgtcttttccggccattgtcgcagcactgtaacgcgtaaaatagtgctttctcttactcttctg D75297

2.18.2 Construction of strain LR38, carrying 20 MalI DNA sites adjacent to the *araFGH* promoter

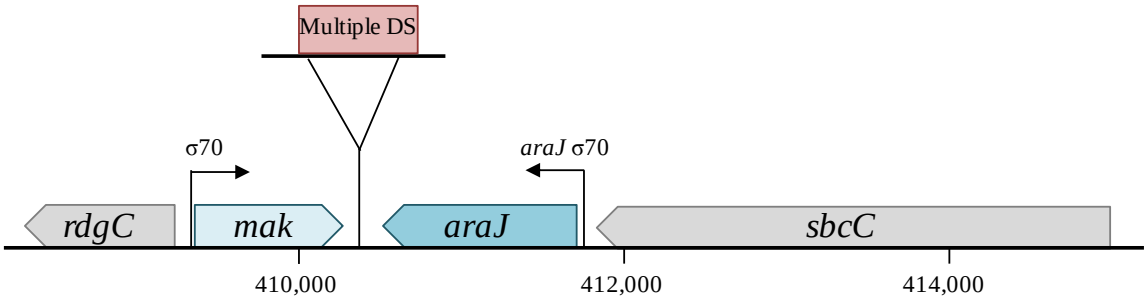
Strain MG1655 was co-transformed with pACBSR and pLR17. Gene doctoring was used to insert a *kan* cassette and 20 copies of the MalI DNA site at the *araFGH* locus. The insertion was confirmed using primers D75296 and D75297, which were designed to anneal to the chromosome outside of the homology regions, to prevent them from annealing to any remaining donor plasmid in the cell. Wild type cells give a PCR product of approximately 1 kb whereas cells with the insert give a product of approximately 3 kb. The PCR product was sequenced with D74952 to check the number of MalI DNA sites that had been inserted onto the chromosome. This was strain LR20. The *kan* cassette was removed from the strain using the FLP recombinase provided on the plasmid pCP20. The removal of the *kan* cassette was confirmed by colony PCR using primers D75296 and D75297 and sequencing with primer D74949. The resulting strain was LR38, carrying 20 MalI DNA sites adjacent to *araFGH*, which was then transformed with pLER108 to supply the cells with MalI::mCherry.

2.19 Introduction of multiple LacI DNA sites and MalI DNA sites adjacent to the *araJ* promoter

2.19.1 Construction of plasmids pLR24 and pLR25

Gene doctoring donor plasmids for making chromosomal inserts adjacent to the *araJ* promoter were made using the same method described in figure 2.26. Again, the homology regions were chosen to avoid any promoter elements by having the insertion site between convergent genes. This should minimise the disruption caused to the local chromosome structure by the insert. Downstream of *araJ* is a gene arranged in a convergent orientation called *mak*. Homology regions were designed to make the chromosomal insert at this point which is approximately 1.4 kb from the *araJ* promoter (see figure 2.29).

a)



b)

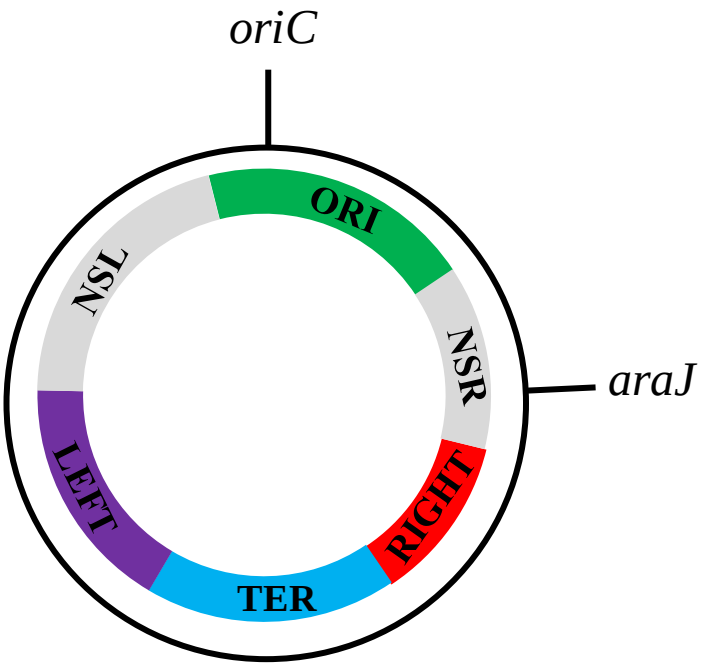


Figure 2.29: Location of multiple LacI or MalI DNA sites inserted adjacent to the *araJ* promoter

a) 22 copies of the LacI DNA site or 20 copies of the MalI DNA site (multiple DS) were inserted onto the chromosome between convergent genes *araJ* and *mak*. Adjacent to these genes is the AraC controlled *araJ* promoter which will be followed using LacI::GFP or MalI::mCherry bound to the multiple DNA sites. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Position of *araJ* relative to the *oriC* and macrodomains.

Primers D75738 and D75739 were used to amplify the region illustrated in figure 2.29 labelled homology region 1 (HR1, figure 2.30), giving a 500 bp PCR product. D75738 introduces a MfeI restriction site upstream of the homology region and D75739 introduces a XmaI site downstream. The PCR product and pJB32 were digested with MfeI and XmaI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pLR22, which was confirmed by sequencing with D58794. The second homology region (HR2, figure 2.30) was then amplified using primers D75740 and D75741 to give a 500 bp PCR product. D75740 introduces a NheI site upstream of the homology region and D75741 introduces a SacI site downstream. The PCR product and pLR22 were digested with NheI and SacI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pLR23 which was confirmed by DNA sequencing with D58793.pLR8 was used to make the multiple LacI DNA site insert for the final cloning step to avoid errors being introduced into the repeat sequence during PCR. pLR8 and pLR23 were both digested with KpnI and NheI. The 2 kb fragment containing the *kan* cassette and LacI DNA site repeats was purified by gel extraction from pLR8. From pLR23 the larger fragment of the vector was purified and treated with CIP. This was ligated with the insert from pLR8 to make plasmid pLR25 which was confirmed by DNA sequencing with D75741.

To make the donor plasmid for inserting multiple MalI DNA sites adjacent to the *araJ* promoter pLR23 was again used to make the vector but the inset came from pLR11. pLR11 and pLR23 were both digested with AgeI and NheI. The 2 kb fragment from pLR11 containing the *kan* cassette and multiple MalI DNA sites was purified by gel extraction. The larger fragment from pLR23 was also purified by gel extraction before being treated with CIP

Figure 2.30: DNA sequence of the chromosomal region adjacent to *araJ* showing the homology regions and primers used to make chromosomal insertions

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *araJ* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *mak* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined.

araJ →

ATGAAAAAAGTCATTTTATCTTTGGCTCTGGGCACGTTTGGTTTGGGGATGGCCGAATTTGGCATTATGGGCGTG

CTCACGGAGCTGGCGCATAACGTAGGAATTTCGATTCCTGCCGCCGGGCATATGATCTCGTATTATGCACTGGGG

GTGGTGGTCGGTGCGCCAATCATCGCACTCTTTTCCAGCCGCTACTCACTCAAACATATCTTGTTGTTTCTGGTG

GCGTTGTGCGTCATTGGCAACGCCATGTTCACGCTCTCTTCGTCTTACCTGATGCTCGCCATTGGTCGGCTGGTA

TCCGGCTTTCCGCATGGCGCATTTTTTGGCGTCGGAGCGATCGTGTTATCAAAAATTATCAAACCCGGAAAAGTC

ACCGCCGCCGTGGCGGGGATGGTTTCCGGGATGACAGTCGCCAATTTGCTGGGCATTCCGCTGGGAACGTATTTA

AGTCAGGAATTTAGCTGGCGTTACACCTTTTTATTGATCGCTGTTTTTAATATTGCGGTGATGGCATCGGTCTAT

TTTTGGGTGCCAGATATTCGCGACGAGGCGAAAGGAAATCTGCGCGAACAATTTCACTTTTTGCGCAGCCCGGCC

CCGTGGTTAATTTTCGCCGCCACGATGTTTGGCAACGCAGGTGTGTTTGCCTGGTTCAGCTACGTAAAGCCATAC

ATGATGTTTATTTCCGGTTTTTCGGAAACGGCGATGACCTTTATTATGATGTTAGTTGGGCTAGGGATGGTGCTG

GGAAATATGCTAAGTGGCAGGATTTCAGGACGTTATTCACCACTGCGCATTGCAGCAGTGACTGACTTTATAATT D76827

GTACTGGCACTGCTGATGCTCTTTTTCTGCGGCGGCATGAAAACAACGTCGCTTATTTTTGCTTTTATTTGTTGC D75738

GCGGGATTATTTGCCCTTTCAGCACCGCTACAAATATTGTTACTACAAAACGCCAAAGGCGGAGAGTTATTAGGT

GCCGCAGGTGGGCAAATAGCGTTTAACCTCGGTAGCGCCGTCGGCGCATATTGCGGAGGTATGATGCTGACGCTG

GGGCTGGCATATAATTACGTGGCGCTGCCTGCCGCCCTGCTTTCGTTTGCTGCGATGTCGTCGTTGCTGCTGTAT HR1

GGTCGCTATAAGCGCCAGCAAGCGGCGGATACTCCGGTGCTGGCGAAACCACTGGGGTAGgttatagtctcggtg

ctcattacttattgccggatgcggcgtgaacgccttatccgccctacgcggttctggcacattttgcaggcctga

taagacgcggcaagcgtcgcatcaggcatcggagcacttattgccggatgcggcgtgaacgccttatccggccta

cggttctggcacctttgtaggcctgataagacgcggcaagcgtcgcatcaggcatgatgcgccaattgcctacg D75836 D75740

ttttTTACTCTTGTGGCCATAACCACGCAGCGCCGCGTACGCCGCTGGAATCACCGTGCTTCGCCTTACGCACCG

GCGTTTCACATTCGCCGCCGAAGACAAATTGTTTAATCAACTGCCCAACCGTTTGATATAAACGGTCTACATTGC

TCATCCCGCCCCCCAGGACAATCACATCCGGATCGAGAATATTCACGACATGTGCCAGCGATTTTGCCAGCCGCA HR2

GCTCGTAGCGACGCAATGCCAGTTCCGCTACCGGATCGCTTTCTTCAACCAGGCGGATAATTTCACTGCCTTTCA

GCGCATGTCCGCTCAAACGACGATAATCCATCGCGAATCCCGTGCCCGAAATAAAGGTTTCAATACAACCTTGTT

TACCGCAATAACAAGGGACTTCCTCGCGATAACGCAGTTCGTCTTCGTCCATCCACGGTAGCGGATTGTGTCCCC

ACTCACCTGCCGTGCCATTGCCGCCGATATGCGCCCGCCCATTGAATGCCACGCCCGCGCCGCATCCCGTGCCGA D75741

TAATCACGGCAAATACCGTCTGCGCTCCCGCTGCCGCGCCATCTACTGCTTCTGAAACCGCCAGACAGTTAGCGT D76828

CATTTGCCAGCCGCACTTCCCGCTGCAACCTCGCGCTTAAGTCTTTATCGAATGGCTGACCGTTGAGCCAGGTTG

AATTGGCATTCTTCACCACACCGGTGTAAGGCGAAATTGAGCCAGGAATGCCCATACCTACCGTTCCGCGCTGCC

← *mak*

CCGTCGCCTGCTCCGCCATATCAACCAACGTGGCGATCGTTTCAATAGTCTGCCGGTAATCATCACGCGGCGTGG

and ligated with the insert from pLR11. The resulting plasmid was pLR24 which was confirmed by sequencing with D75741.

2.19.2 Construction of strain LR39, carrying 22 LacI DNA sites adjacent to the *araJ* promoter

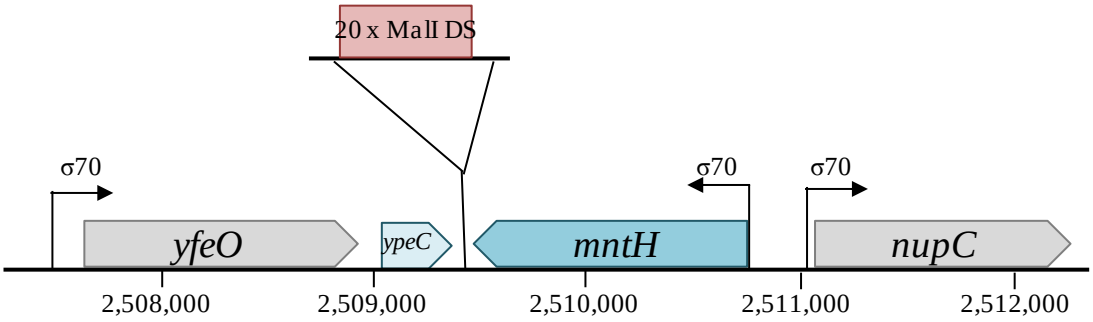
Strain MG1655 LacI::GFP was co-transformed with pACBSR and pLR25. The gene doctoring protocol described previously was used to insert a *kan* cassette and 22 copies of the LacI DNA site at the *araJ* locus. The insertion was confirmed using primers D76827 and D76828 which were designed to anneal to the chromosome outside of the homology regions to prevent them from annealing to any remaining donor plasmid in the cell. Wild type cells give a PCR product of approximately 1 kb, whereas cells with the insert give a product of approximately 3 kb. The PCR product was sequenced with D75741 to check the number of LacI DNA sites that had been inserted onto the chromosome. The resulting strain was strain LR22. The *kan* cassette was removed using the FLP recombinase provided by the plasmid pCP20. The removal of the *kan* cassette was confirmed by colony PCR using primers D76827 and D76828 and sequenced using primer D75738. The new strain was, LR39, with 22 copies of the LacI DNA site adjacent to *araJ*.

2.20 Introduction of multiple MalI DNA sites adjacent to the *mntH* promoter

2.20.1 Construction of plasmid pSB5

Gene doctoring donor plasmids for making inserts adjacent to the *mntH* promoter were made using the same basic method described in figure 2.26. To reduce the chance of any effects on expression of surrounding genes, the insertion site was between two convergent genes, *mntH* and *ypeC* (see figure 2.31). This insert is approximately 1.4 kb from the *mntH* promoter.

a)



b)

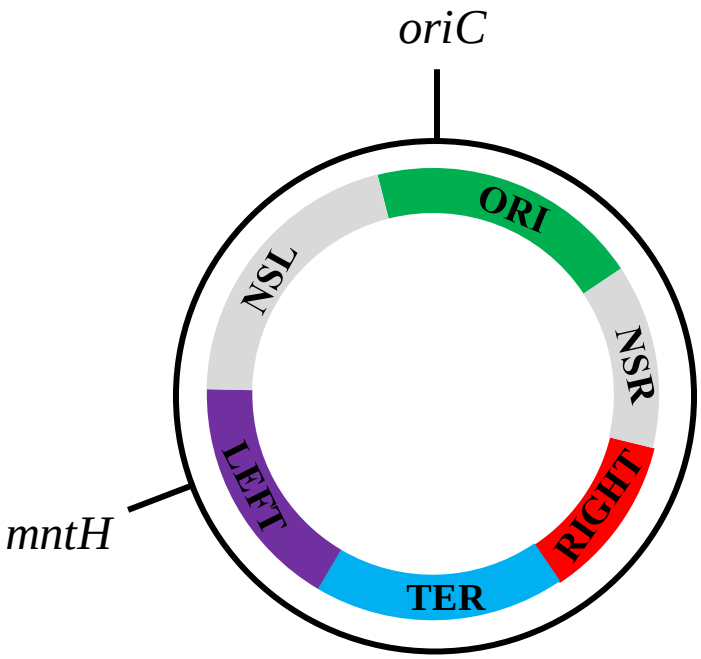


Figure 2.31: Location of multiple MalI DNA sites inserted adjacent to the *mntH* promoter

a) 20 copies of the MalI DNA site (MalI DS) were inserted onto the chromosome between convergent genes *mntH* and *ypeC*. Adjacent to these genes is the MntR regulated *mntH* promoter which will be followed using MalI::mCherry bound to the multiple DNA sites. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Position of *mntH* relative to the *oriC* and macrodomains.

Primers D75746 and D75747 were used to amplify the region illustrated in figure 2.31 labelled homology region 1 (HR1, figure 2.32), giving a 500 bp PCR product. D75746 introduces a MfeI restriction site upstream of the homology region and D75747 introduces a XmaI site downstream. The PCR product and pJB32 were digested with MfeI and XmaI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pSB1, which was confirmed by sequencing with D58794. The second homology region (HR2, figure 2.32) was then amplified using primers D75748 and D75749 to give a 500 bp PCR product. D75748 introduces a NheI site upstream of the homology region and D75749 introduces a SacI site downstream. The PCR product and pSB1 were digested with NheI and SacI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pSB3 which was confirmed by DNA sequencing with D58793.

pLR11 was used to make the multiple MalI DNA site insert for the final cloning step to avoid errors being introduced into the repeat sequence during PCR. pLR11 and pSB3 were both digested with AgeI and NheI. The 2 kb fragment from pLR11 containing the *kan* cassette and multiple MalI DNA sites was purified by gel extraction. The larger fragment from pSB3 was also purified by gel extraction before being treated with CIP and ligated with the insert from pLR11. The resulting plasmid was pSB5 which was confirmed by sequencing with D75749.

2.20.2 Construction of strain SXB3, carrying 20 MalI DNA sites adjacent to the *mntH* promoter

Strain MG1655 was co-transformed with pACBSR and pSB5. The gene doctoring protocol described previously was used to insert a *kan* cassette and 20 copies of the MalI DNA site at the *araJ* locus. The insertion was confirmed using primers D76493 and D76494, which were designed to anneal to the chromosome outside of the homology regions to prevent them from

Figure 2.32: DNA sequence of the chromosomal region adjacent to *mntH* showing the homology regions and primers used to make chromosomal insertions

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *mntH* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *ypeC* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined.

mntH →

ATGACGAACTATCGCGTTGAGAGTAGCAGCGGACGGGCGGCGCGCAAGATGAGGCTCGCATTAATGGG

ACCTGCGTTCATTGCGGCGATTGGTTATATCGATCCCGGTAACTTTGCGACCAATATTCAGGCGGGTG

CTAGCTTCGGCTATCAGCTACTGTGGGTTGTCGTTTGGGCCAACCTGATGGCGATGCTGATTCAGATC

CTCTCTGCCAAACTAGGGATTGCCACCGGTAAAAATCTGGCGGAGCAGATTCGCGATCACTATCCGCG

TCCCGTAGTGTGGTTCTATTGGGTTCAGGCAGAAATTATTGCGATGGCAACCGACCTGGCGGAATTTA

TTGGTGCGGCGATCGGTTTTAAACTCATTCTTGGTGTTTCGTTGTTGCAGGGCGCGGTGCTGACGGGG

ATCGCGACTTTCCTGATTTTAATGCTGCAACGTCGCGGGCAAAAACCGCTGGAGAAAGTGATTGGCGG

GTTACTGTTGTTTGTTGCCGCGGCTTACATTGTCGAGTTGATTTTCTCCCAGCCTAACCTGGCGCAGC

TGGGTAAAGGAATGGTGATCCCGAGTTTACCTACTTCGGAAGCGGTCTTCCTGGCAGCAGGCGTGTTA

GGGGCGACGATTATGCCGCATGTGATTTATTTGCACTCCTCGCTCACTCAGCATTTACATGGCGGTTC

GCGTCAACAACGTTATTCCGCCACCAAATGGGATGTGGCTATCGCCATGACTATTGCCGGTTTTGTCA

ATCTGGCGATGATGGCTACAGCTGCGGCGGCGTTCCACTTTTCCGGTCATACTGGTGTTGCCGATCTT D76493

GATGAGGCTTATCTGACGCTGCAACCGCTGTTAAGCCACGCTGCGGCAACGGTCTTTGGATTAAGCCT D75746

GGTTGCTGCGGGGCTGTCTTCAACGGTGGTGGGGACACTGGCGGGGCAGGTGGTGATGCAGGGCTTCA

TTCGCTTTCATATCCCGCTGTGGGTGCGTCGTACAGTCACCATGTTGCCGTCATTTATTGTCATTCTG

ATGGGATTAGATCCGACACGGATTCTGGTTATGAGTCAGGTACTGTTAAGTTTTGGTATCGCTCTGGC HR1

GCTGGTTCCACTGCTGATTTTCACCAGTGACAGCAAGTTGATGGGCGATCTGGTGAACAGCAAACGCG

TAAAACAGACAGGCTGGGTGATTGTGGTGCTGGTCGTGGCGCTGAATATCTGGTTGTTGGTGGGGACG

GCGCTGGGATTGTAGttgaatgagcgtcgcatctggcactattggcggatgcggcgtaaacgccttat

ccgccctacgcgttaaagagccggtttgtaggcctgataagacgcactagcgtcgcatcaggcattgt D75747

gctccattggcggacgcgtTTAATGGCGGTGCCCATGACCCCGGCCTTTTCCGCGATGATCGTCACGA D75748

TCGCGCCAGCCTTCACGATAGCCACGCTCATACGCTTTACGCTTATCCCAGCCACGGTGGTAGCCATT

ATCATGACGCCACCAACGGTTTTTGCGCCACTCATAATTGCGATGCCAGTAGTCACGGTCGCGCCAGT

GACCACCGTCCCAGTAATTACCGTAATGATCGCGATCGCCAATTTGTAATTTGATTGATGGCAGTAGG

GTGATTTCACCTGCGTTTGCTGCAAGCGGGGTAAATGCCATCAGGGCGGCCGCCAGAAACAGTGACCT HR2

← *ypeC*

GAACATtgttattctccttcacgctcgaagccgtcagcggcctgttaacgcaatattacggggaggta

aagccccgtatcatcgccataactcttaaatcataaaggagagcatttttgCTAAACGTTATCGTGG

GGTGGCTGTTGCTTCGGACGATTGACCATCATCATCGGCTTACCTGCTAATAACAGCCATGCCGGAAG

CATGACGATACAGAGCAGCGGTAGCAATGTGGTATTGGGTACAACGACTGCCGCCATAAAAAGACTTA

D75749

ACCAGCCATCGCGTGTTACCACCAGCACGATGCCGAGAATAGCGCAAGAAACGGTTATTGCCGCTGGT D76494

← yfeO

ACGGCGGGAACGTGCTCATGCAGCATCAACCCTAATGCCACGCCGACAAACACTGCCGGGAAGATTCG

annealing to any remaining donor plasmid in the cell. Wild type cells give a PCR product of approximately 1 kb, whereas cells with the insert give a product of approximately 3 kb. The PCR product was sequenced with D75749 to check the number of MalI DNA sites that had been inserted onto the chromosome. This was strain SXB1, which was transformed with pCP20 to remove the *kan* cassette, PCR checked and sequenced again with D75746. The new strain, SXB3, was then transformed with pLER108 to supply the cells with MalI::mCherry.

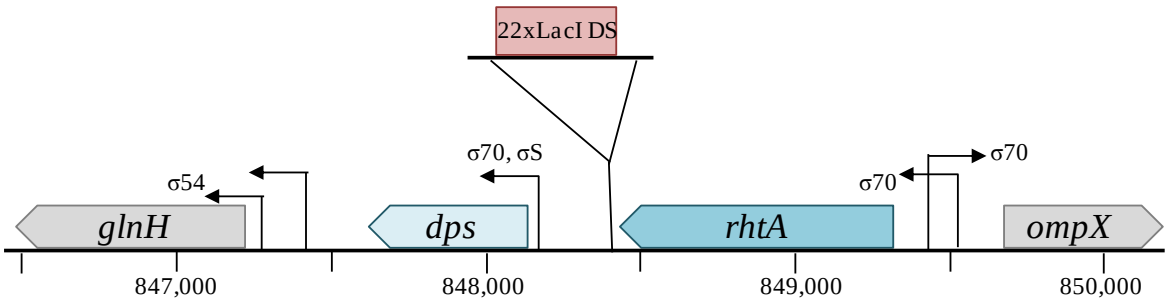
2.21 Introduction of multiple LacI DNA sites adjacent to the *dps* promoter

2.21.1 Construction of plasmid pSB6

Gene doctoring donor plasmids for making inserts adjacent to the *dps* promoter were made using the same basic method described in figure 2.26. Due to the arrangement of the genes around *dps* it was not possible to insert close to the *dps* promoter between convergent genes, which would have been the best way to ensure the insert did not change expression of any genes. Instead an insertion site was chosen upstream of the *dps* promoter between *dps* and *rhtA* and great care was taken to avoid promoter elements (see figure 2.33). This insert is approximately 250 bp from the *dps* promoter.

Primers D75744 and D75745 were used to amplify the region illustrated in figure 2.33 labelled homology region 2 (HR2, figure 2.34) giving a 500 bp PCR product. D75744 introduces a NheI restriction site upstream of the homology region and D75745 introduces a SacI site downstream. The PCR product and pJB32 were digested with NheI and SacI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pSB2, which was confirmed by sequencing with D58794. The second homology region (HR1, figure 2.34) was then amplified using primers D75742 and D75743 to give a 500 bp PCR product.

a)



b)

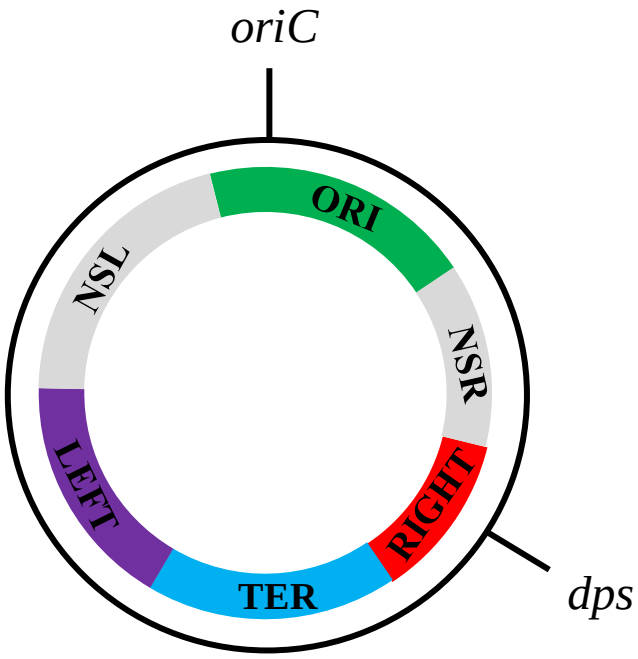


Figure 2.33: Location of multiple LacI DNA sites adjacent to the *dps* promoter

a) 22 copies of the LacI DNA sites (LacI DS) were inserted onto the chromosome between unidirectional genes *rhtA* and *dps*. Adjacent to these genes is the MntR regulated *dps* promoter which will be followed using LacI::GFP bound to the LacI DNA sites. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Position of *dps* relative to the *oriC* and macrodomains.

Figure 2.34: DNA sequence of the chromosomal region adjacent to *dps* showing the homology regions and primers used to make chromosomal insertions

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *rhtA* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *dps* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined.

rhtA →

ATGCCTGGTTCATTACGTAAAATGCCGGTCTGGTTACCAATAGTCATATTGCTCGTTGCCATGGCGTC

TATTCAGGGTGGAGCCTCGTTAGCTAAGTCACTTTTTCCTCTGGTGGGCGCACCGGGTGTCACTGCGC

TGCGTCTGGCATTAGGAACGCTGATCCTCATCGCGTTCTTTAAGCCATGGCGACTGCGCTTTGCCAAA

GAGCAACGGTTACCGCTGTTGTTTTACGGCGTTTCGCTGGGTGGGATGAATTATCTTTTTTATCTTTC

TATTCAGACAGTACCGCTGGGTATTGCGGTGGCGCTGGAGTTCACCGGACCACTGGCGGTGGCGCTGT

TCTCTTCTCGTCGCCCGGTAGATTTCGTCTGGGTTGTGCTGGCGGTTCTTGGTCTGTGGTTCCTGCTA D76491 D75742

CCGCTGGGGCAAGACGTTTCCCATGTCGATTTAACCGGCTGTGCGCTGGCACTGGGGGCCGGGGCTTG

TTGGGCTATTTACATTTTAAGTGGGCAACGCGCAGGAGCGGAACATGGCCCTGCGACGGTGGCAATTG

GTTCGTTGATTGCAGCGTTAATTTTCGTGCCAATTGGAGCGCTTCAGGCTGGTGAAGCACTCTGGCAC

TGGTCGGTTATTCCATTGGGTCTGGCTGTCGCTATTCTCTCGACCGCTCTGCCTTATTCGCTGGAAAT

GATTGCCCTCACCCGTTTGCCAACACGGACATTTGGTACGCTGATGAGCATGGAACCGGCGCTGGCTG HR1

CCGTTTCCGGGATGATTTTCCTCGGAGAAACACTGACACCCATACAGCTACTGGCGCTCGGCGCTATC

ATCGCCGCTTCAATGGGGTCTACGCTGACAGTACGCAAAGAGAGCAAAATAAAAGAATTAGACATTAA

TTAAatttacatttctgcatggttatgcataaccatgcagaatttctcgctacttttcctctacaccgtcttta D75743 D75744

tatatcgaattatgcaaaagcatatttattccgaaaattcctggcgagcagataaataagaattgttcttatcaa

tatatctaactcattgaatctttattagttttgtttttcacgcttgttaccactattagtgtgataggaacagcc

agaatagcggaacacatagccggtgctatacttaatctcgttaattactgggacataacatcaagaggatatgaa HR2

dps →

attATGAGTACCGCTAAATTAGTTAAATCAAAAGCGACCAATCTGCTTTATACCCGCAACGATGTCTCCGACAGC

GAGAAAAAAGCAACAGTAGAGTTGCTGAATCGCCAGGTTATCCAGTTTATTGATCTTTCTTTGATTACCAAACAA

GCGCACTGGAACATGCGCGGCGCTAACTTCATTGCCGTACATGAAATGCTGGATGGCTTCCGCACCGCACTGATC

GATCATCTGGATACCATGGCAGAACGTGCAGTGCAGCTGGGCGGTGTAGCTCTGGGGGACCACTCAAGTTATCAA D75745

CAGCAAAACCCCGCTGAAAAGTTACCCGCTGGACATCCACAACGTTCAGGATCACCTGAAAGAACTGGCTGACCG D76492

TTACGCAATCGTCGCTAATGACGTACGCAAAGCGATTGGCGAAGCGAAAGATGACGACACCGCAGATATCCTGAC

CGCCGCGTCTCGCGACCTGGATAAATTCCTGTGGTTTATCGAGTCTAACATCGAATAAatccatcgctgatggtg

cagaactttagtacccgataaaagcggcttcctgacaggaggccgttttgttttgcagcccacctcaacgcactt

atttagtgcatccatctgctatctccagctgattaagtaaatttttgtatccacatcatcacacaatcgttaca

taaagattgttttttcatcaggttttacgctaaataatcactgtgttgagtgcacaattttagcgcaccagattg

D75742 introduces a MfeI site upstream of the homology region and D75743 introduces a XmaI site downstream. The PCR product and pSB2 were digested with MfeI and XmaI, gel extracted and the vector treated with CIP. They were then ligated to make plasmid pSB4 which was confirmed by DNA sequencing with D58793.

pLR8 was used to make the insert containing 22 copies of the LacI DNA site for the final cloning step, to avoid errors being introduced into the repeat sequence during PCR. pLR8 and pSB4 were both digested with KpnI and NheI. The 2 kb fragment from pLR8 containing the *kan* cassette and multiple LacI DNA sites was purified by gel extraction. The larger fragment from pSB4 was also purified by gel extraction before being treated with CIP and ligated with the insert from pLR8. The resulting plasmid was pSB6 which was confirmed by sequencing with D75745.

2.21.2 Construction of strain SXB4, carrying 22 LacI DNA sites adjacent to *dps*

Strain MG1655 LacI::GFP was co-transformed with pACBSR and pSB6. The gene doctoring protocol described previously was used to insert a *kan* cassette and 22 LacI DNA sites adjacent to *dps*. The insertion was confirmed using primers D76491 and D76492, which were designed to anneal to the chromosome outside of the homology regions to prevent them from annealing to any remaining donor plasmid in the cell. Wild type cells give a PCR product of approximately 1 kb, whereas cells with the insert give a product of approximately 3 kb. The PCR product was sequenced with D75745 to check the number of LacI DNA sites that had been inserted onto the chromosome. This was strain SXB2, which was transformed with pCP20 to remove the *kan* cassette, PCR checked and sequenced again with D75742. The resulting strain was SXB4, with 22 LacI DNA sites adjacent to *dps*.

2.22 Construction of strains with two FROS inserts

To label two locations in the same cell, multiple LacI DNA sites and multiple MalI sites were inserted onto the chromosome in the same strain. Usually, P1 transduction was used to transfer multiple MalI DNA sites into a strain that already contain a multiple LacI DNA site insert and a chromosomal LacI::GFP fusion. Alternatively, MalI DNA sites were inserted into a strain carrying a LacI DNA site insert and a chromosomal LacI::GFP fusion using gene doctoring.

2.22.1 Construction of strain LR31, carrying, 22 LacI DNA sites adjacent to *araBAD* and 20 MalI DNA sites adjacent to *araJ*.

To compare the positions of the *araBAD* and *araJ* promoters both need to be tagged in the same strain using different colours. *araBAD* has already been tagged with LacI::GFP in LR06 and the *kan* cassette has been removed. LR06 was transformed with pACBSR and pLR24 and the gene doctoring was used to insert 20 copies of the MalI DNA site adjacent to *araJ*. The insert was confirmed by colony PCR using primers D76827 and D76828 which give a product of around 1 kb in wild type cells, but 3 kb if the insert is present. The PCR product was sequenced with primer D75741 to check the number of DNA sites inserted onto the chromosome. The resulting strain was LR19. The *kan* cassette was removed using the FLP recombinase provided by the plasmid pCP20. The removal of the *kan* cassette was confirmed by colony PCR using primers D76827 and D76828 and sequenced using primer D75738. The new strain, LR31, was then transformed with pLER108 to supply the cells with MalI::mCherry.

2.22.2 Construction of strain LR42, carrying 22 LacI DNA sites adjacent to *araBAD* and 20 MalI DNA sites adjacent to *mntH*

The P1 transduction method described previously (section 2.10) was used to transfer an insert from the donor strain, SXB1, to the recipient strain, LR06. SXB1 has an insert of 20 MalI DNA sites and a *kan* cassette adjacent to *mntH*. Strain LR06 carries 22 LacI DNA sites adjacent to *araBAD* and a chromosomal LacI::GFP fusion. The insert of MalI DNA sites adjacent to *mntH* was confirmed by PCR using primers D76493 and D76494, and sequencing using primer D75749. The presence of the LacI DNA sites adjacent to *araBAD* was confirmed by PCR using primers D69747 and D69748 and sequencing using primer D69233. The resulting strain was LR40. The *kan* cassette was removed from strain LR40 using the FLP recombinase, encoded on plasmid pCP20. The removal of the *kan* cassette was confirmed by PCR using primers D76493 and D76494 and sequencing using primer D75746. The resulting strain was LR42, carrying 22 LacI DNA sites adjacent to *araBAD*, 20 MalI DNA sites adjacent to *mntH* and a chromosomal LacI::GFP fusion. This strain was then transformed with plasmid pLER108, encoding MalI::mCherry.

2.22.3 Construction of strain LR47, carrying 22 LacI DNA sites adjacent to *dps* and 20 MalI DNA sites adjacent to *mntH*

The P1 transduction method described previously (section 2.10) was used to transfer an insert from the donor strain, SXB1, to the recipient strain, SXB4. SXB1 has an insert of 20 MalI DNA sites and a *kan* cassette adjacent to *mntH*. Strain SXB4 carries 22 LacI DNA sites adjacent to *dps* and a chromosomal LacI::GFP fusion. The insert of MalI DNA sites adjacent to *mntH* was confirmed by PCR using primers D76493 and D76494 and sequencing using primer D75749. The presence of the LacI DNA sites adjacent to *dps* was confirmed by PCR using primers D76491 and D76492 and sequencing using primer D75745. The resulting strain

was LR45. The *kan* cassette was removed from strain LR45 using the FLP recombinase, encoded on plasmid pCP20. The removal of the *kan* cassette was confirmed by PCR using primers D76493 and D76494 and sequencing using primer D75746. The resulting strain was LR47, carrying 22 LacI DNA sites adjacent to *dps*, 20 MalI DNA sites adjacent to *mntH* and a chromosomal LacI::GFP fusion. This strain was then transformed with plasmid pLER108, encoding MalI::mCherry.

2.22.4 Construction of strain LR48, carrying 22 LacI DNA sites adjacent to *araBAD* and 20 MalI DNA sites adjacent to *araFGH*

The P1 transduction method described previously (section 2.10) was used to transfer an insert from the donor strain, LR20, to the recipient strain, LR06. LR20 has an insert of 20 MalI DNA sites and a *kan* cassette adjacent to *araFGH*. Strain LR06 carries 22 LacI DNA sites adjacent to *araBAD* and a chromosomal LacI::GFP fusion. The insert of MalI DNA sites adjacent to *araFGH* was confirmed by PCR using primers D75296 and D75297 and sequencing using primer D74952. The presence of the LacI DNA sites adjacent to *araBAD* was confirmed by PCR using primers D69747 and D69748, and sequencing using primer D69233. The resulting strain was LR46. The *kan* cassette was removed from strain LR46 using the FLP recombinase, encoded on plasmid pCP20. The removal of the *kan* cassette was confirmed by PCR using primers D75296 and D75297, and sequencing using primer D74949. The resulting strain was LR48, carrying 22 LacI DNA sites adjacent to *araBAD*, 20 MalI DNA sites adjacent to *mntH* and a chromosomal LacI::GFP fusion. This strain was then transformed with plasmid pLER108, encoding MalI::mCherry.

2.23 AraC knockout in strain LR38

P1 transduction (section 2.10) was used to transfer a *ΔaraC* mutation into strain LR38. Strain *ΔaraC* 83, which carries a *kan* cassette for selection, was used as the donor strain. The mutation was confirmed by PCR using primers D76717 and D76718, which give a product of around 1 kb in MG1655 and no product *in ΔaraC* 83. The mutation was further confirmed by streaking candidates onto MacConkey plates supplemented with arabinose. *ΔaraC* strains have white colonies. The resulting strain was LR44, with 20 MalI DNA sites adjacent to *araFGH* and a deletion in *araC*.

2.24 Chromosomal fusion of LacI to PAmCherry

2.24.1 Construction of plasmid pDOC-PAM-lac

A new LacI fusion was needed with a fluorescent protein suitable for PALM super-resolution microscopy. A derivative of plasmid pDOC-G had already been made with homology regions for inserting a LacI::GFP fusion onto the chromosome using gene doctoring, pDOC-G-lac (Lee *et al.*, 2009). This plasmid creates a C-terminal fusion and deletes the promoter and the first 400 bp of *lacZ* (figure 2.35). Homology region 1 contains 200 bp of *lacI*, excluding the stop codon. Homology region 2 contains 200 bp of *lacZ*, from around 400 bp into the gene (figure 2.36). The area between these two homology regions will be deleted during the gene doctoring process and replaced with the fluorescent protein. As this region contains all 3 of the LacI DNA sites, there will be no DNA sites for LacI on the chromosome in the resulting strain. A photoactivatable fluorescent protein is needed for PALM, so PAmCherry was amplified from plasmid pBAD-HisB-PAmcherry using primers D73820 and D75378, which introduced a KpnI immediately upstream of PAmCherry and an AgeI site downstream. The PCR product and pDOC-G-lac plasmid were both digested with KpnI and AgeI, before being

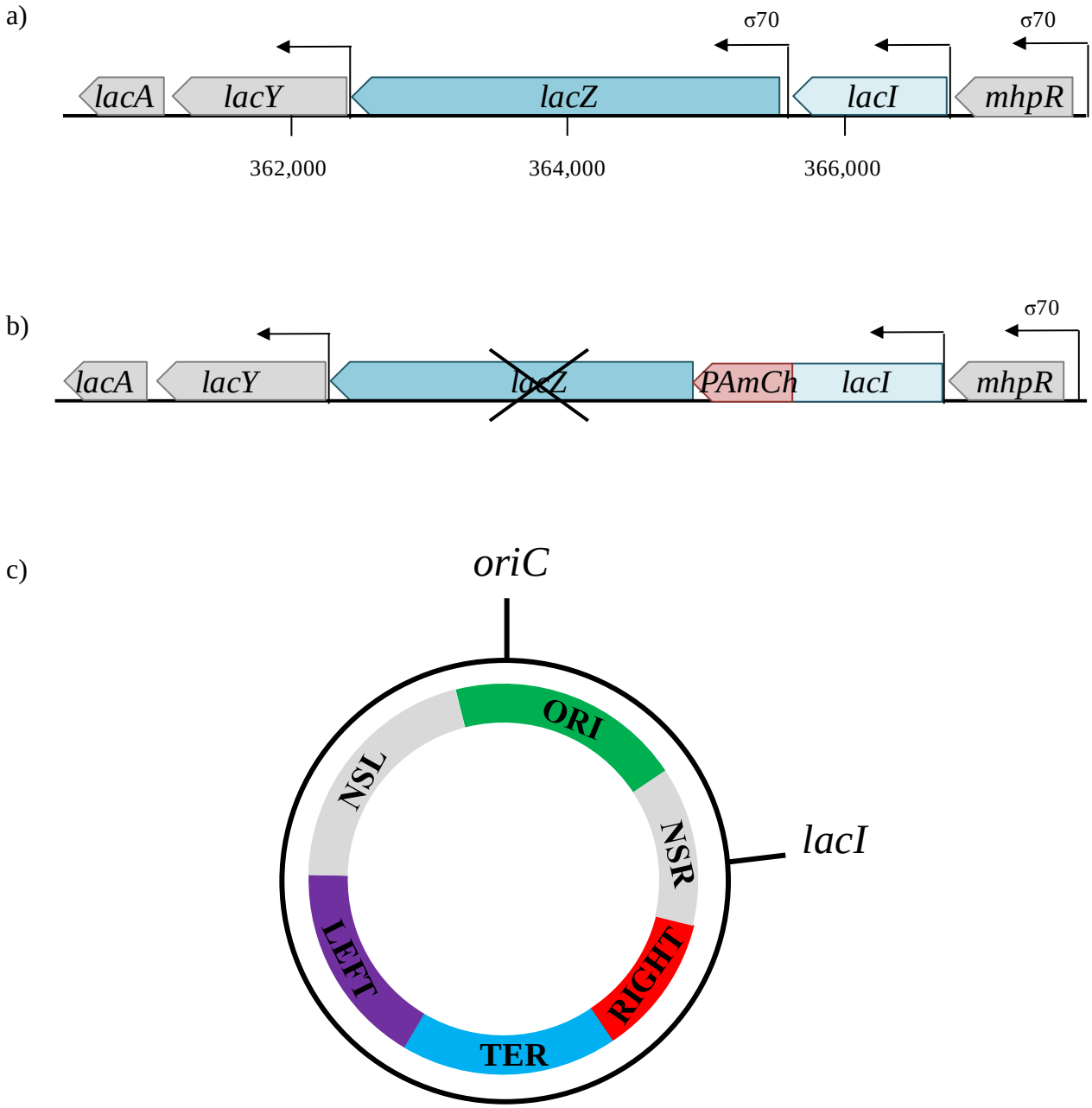


Figure 2.35: Location of the chromosomal LacI::PAmCherry fusion

a) PAmCherry was inserted onto the chromosome downstream of *lacI*. The promoter of the adjacent gene, *lacZ*, is deleted. Chromosomal position shown in bp as defined by Keseler *et al.*, 2011.

b) Chromosome structure after PAmCherry has been inserted.

c) Position of *lacI* relative to the *oriC* and macrodomains.

Figure 2.36: DNA sequence of the chromosomal region adjacent to *lacI* showing the homology regions and primers used to make chromosomal LacI::GFP fusion

Intergenic regions are shown in lowercase, sequences of genes are in capitals. The start codons of genes are marked in red letters with the direction of the gene shown by an arrow. Homology region 1 (HR1) contains part of *lacI* and is shown with dark blue shading. The sequence of the primers used to amplify homology region 1 are double underlined. Homology region 2 (HR2) contains part of *lacZ* and is shown with light blue shading. The sequence of the primers used to amplify homology region 2 are boxed. The sequence of primers used to check gene doctoring candidates for insertions are single underlined. The region between the two homology regions is lost during the gene doctoring process and, in this case, contains all 3 LacI DNA sites, shown in italics.

lacI →

GTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAAC

CAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAAC

CGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCG

CCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAA

CGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAAC

TATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTC

TCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTC

GCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGC

TGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGT

TTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCG

CTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGAT D61347

ACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGC D59400

GTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAA HR1

AGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA

CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAgcgcaacgcaatt*aatgtgagttagctcactcatt*aggcacc D59401 *O2*

ccaggctttacactttatgcttccggctcgtatgttgtgtgg*aattgtgagcggataacaatt*tcacacaggaaa *O3*

lacZ →

cagctATGACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCC

AACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT

CCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCT

GGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGC

CCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGG*TTGTT* *O2*

*ACTCGCTCACATTTAA*TGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACT D59402

CGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGACC HR2

TGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGCTGGAGTGACGGCAGTTATCTGG

AAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCA D59403

gel extracted and the vector being treated with CIP. They were then ligated to make plasmid pDOC-PAM-lac, which was confirmed by sequencing using primer D58794.

2.24.2 Construction of strains LR35 and LR37, carrying a LacI::PAmCherry chromosomal fusion

Strains MG1655 and LR18 were both co-transformed with pACBSR and pDOC-PAM-lac. The gene doctoring protocol described previously was used to insert a chromosomal fusion of LacI to PAmCherry and *kan* cassette into both strains. The insert was confirmed by PCR using primers D61347 and D59403, and sequencing using primer D59400. This resulted in strain LR29, derived from MG1655, carrying a LacI::PAmCherry fusion, and LR30, derived from LR18, carrying 6 LacI DNA sites adjacent to *araBAD* and a chromosomal LacI::PAmCherry fusion. Both strains were transformed with pCP20, encoding FLP recombinase, to remove the *kan* cassette. This was confirmed by PCR with D61347 and D59403 and sequencing using primer D59403. The resulting strains were LR35, carrying a LacI::PAmCherry fusion, and LR37, carrying 6 LacI DNA sites adjacent to *araBAD* and a LacI::PAmCherry fusion.

3. Repressor proteins as tools for tagging the *E. coli* chromosome

3.1 Introduction

3.1.1 Use of repressor proteins to visualise chromosomal loci

In order to understand more about how the *E. coli* chromosome is structured, tools are needed to follow the chromosome and study its dynamics. One commonly used technique is to use a fluorescently tagged repressor protein to tag chromosomes at a specific location. As mentioned previously, repressors are proteins that bind to a specific sequence in the promoter of a gene and regulate its transcription. One family of transcription factors whose members function in this way is the LacI/GalR family which has over 1000 members (Swint-Kruse and Matthews, 2009). Their biological function is to regulate their target genes, but these proteins are being put to a new use in visualising chromosome dynamics, where their properties of DNA recognition and binding of a consensus sequence are being exploited. An array of multiple DNA sites for a particular repressor protein is inserted at the position of interest on the chromosome. This is then visualised by tagging the repressor protein with a fluorescent protein, resulting in multiple copies of the fluorescent protein being brought together at the targeted location, forming a focus which can be followed by fluorescence microscopy. This is called a Fluorescent Repressor/Operator System (FROS).

This technique was first used in budding yeast in 1996 to study the separation of sister chromatids (Straight *et al.*, 1996). Later that year, the same technique was used in Chinese hamster ovary cells (CHO) and budding yeast where it was used to investigate the fibres formed due to large scale organisation of chromatin (Robinett *et al.*, 1996). Since then, the technique has been widely used to study chromosome dynamics in other types of cells including bacteria, *C. elegans* (Carmi *et al.*, 1998), *Drosophila* (Gasser, 2002) and *Arabidopsis thaliana* (Kato and Lam, 2001). More recently, this technique was used in *C.*

elegans to show that three genes with tissue specific promoters were localised to different areas of the nucleus in differentiated cells depending on whether they were being expressed but had no preferential position in undifferentiated embryo cells (Meister *et al.*, 2010). Many of these examples include a second chromosomal tag for comparison with the original locus. This involves using a second repressor protein tagged to a different colour fluorescent protein and inserting multiple copies of its DNA site onto the chromosome. The *lac* repressor, LacI is often used as the primary reporter, sometimes in combination with the *tet* repressor, TetR, at a reference locus (Lau *et al.*, 2003).

3.1.2 The LacI repressor

In the original FROS experiments LacI, the *E. coli lac* repressor, was used as the reporter (Straight *et al.*, 1996, Robinett *et al.*, 1996). Its usual function is to negatively regulate the *lac* operon, consisting of *lacZ, lacY* and *lacA*, all involved in the transport and metabolism of the sugar, lactose. In the absence of allolactose, the inducer, LacI binds to the *lac* promoter and represses the *lac* operon. A 21 bp LacI binding site is known as the *lac* operator and appears 3 times on the *E. coli* chromosome, at positions -82, +10 and +411, all respective to the *lac* transcription start site. Each binding site is recognised and bound by 2 LacI subunits. The site centred at position +10, known as O_1, has been shown by site directional mutagenesis to be the most important in repression and has the sequence AATTGTGAGCGGATAACAATT, which is an inverted repeat (Oehler *et al.*, 1990). Repression of the *lac* operon is thought to be achieved by two subunits of a LacI tetramer binding O_1 and two subunits binding to one of the other two sites (O_2 and O_3) simultaneously, forming a loop in the DNA, which hinders the transcription machinery and therefore represses the gene. As the N terminus is responsible for DNA binding, LacI can be fused to GFP at its C terminus without affecting its DNA binding capabilities, and has been used this way in previous studies (Elf *et al.*, 2007). This can then be

used to visualise the chromosome at a point of interest by putting multiple copies of the LacI DNA site at that point. Multiple copies of LacI then bind the chromosome at the position of interest and a focus of GFP can be seen using a microscope. Usually an array of 256 copies of the LacI DNA binding site is used with the LacI::GFP fusion being supplied by a multicopy plasmid (Lau *et al.*, 2003).

3.1.3 The MalI repressor

In 1987, a mutant was discovered that had altered expression of genes involved in maltose metabolism and transport, specifically *malK*. The corresponding mutation was mapped and the gene was called MalI, thought to be responsible for regulating genes involved in maltose metabolism and transport (Ehrmann and Boos, 1987). However, further studies showed that MalI actually regulated an operon found immediately upstream of the *malI* gene, in a divergent orientation, containing two genes called *malX* and *malY*, The physiological function of the *malX* product is unknown, but it is thought it may catalyse facilitated diffusion of an unknown substrate (Reidl and Boos, 1991). The *malY* product has the enzymatic activity of a βC-S lyase using amino acids with a βC-S linkage as a substrate to make other biologically useful molecules, for example pyruvate (Zdych *et al.*, 1995). It was shown that *malK* expression is affected by the *malY* gene product, possibly by removing an inducer. As *malY* is repressed by MalI, the effect of the *malI* mutation on *malK* expression seen by Ehrmann and Boos was in reality caused by a change in the level of the *malY* gene product, not a direct effect of MalI.

A MalI DNA site was found at the *malX* and *malI* promoters which are both co-regulated by CRP. The *malI* promoter is active in the absence of CRP, probably to ensure the repressor is

present at all times to repress *malX* and *malY* (Lloyd et al., 2008). The DNA site for MalI is 5'-GATAAAACGTTTTATC-3' located from position -24 to -9 in the *malX* promoter and +3 to +18 in the *malI* promoter (Lloyd *et al.*, 2010).

Like LacI, MalI is also a member of the LacI/GalR family of transcriptional regulators. It has 24% sequence homology to LacI with particular similarities in the N-terminal region, thought to be responsible for DNA binding (Reidl *et al.*, 1989). Although the role of MalI in the cell is unknown, it shares some properties with LacI that make it a good candidate for use as a chromosomal tag, for example DNA recognition and binding. As with LacI, multiple copies of the MalI DNA site could be used in combination with fluorescently tagged MalI protein to label a region of interest on the chromosome.

3.2 Comparison of chromosomal and plasmid-encoded LacI::GFP

All strains used for this study are derived from MG1655. Inserts for FROS experiments were introduced onto the chromosome by gene doctoring. Any insert into a chromosome risks disrupting the local chromosome structure. As FROS is often used to study chromosome structure this disruption needs to be minimised by reducing the size of the chromosomal insert as far as possible, making results collected more useful. The insert of 256 DNA sites used in most studies is around 10 kb, so a smaller insert would be preferable. This project aimed to modify FROS to be less disruptive to the chromosome. In previous studies with 256 LacI DNA sites inserted onto the chromosome, a multicopy plasmid is used to supply LacI::GFP to the cell giving a much higher concentration of LacI than is found in a wild type cell. This allows for maximum occupancy of the array of multiple LacI DNA sites by LacI::GFP. In this project, a smaller array of LacI DNA sites was used, meaning there were only 22 DNA sites

inserted at the chosen target for LacI::GFP to bind to. Due to this reduction in the number of available DNA sites, a lower cellular concentration of LacI::GFP should be needed.

Strains LR04 and LR06, both carrying 22 copies of the LacI DNA site, adjacent to the *araBAD* promoter, were used to compare plasmid encoded and chromosomal encoded LacI::GFP. Strain LR04 was transformed with plasmid pLER101 encoding LacI::GFP under the control of the wild type *lacI* promoter, now referred to as LR04-pLER101. The plasmid was derived from pACYC184 which has a copy number of approximately 15-20 and, as a result, the concentration of LacI::GFP in the cell can be expected to be 15-20 times that of LacI in wild type cells. Strain LR06 has a LacI::GFP fusion inserted onto the chromosome at the wild type location and under the control of the wild type *lacI* promoter. Both strains had the *lacZ* promoter deleted removing all other LacI DNA sites from the chromosome. When visualised by microscopy, most of the cells with chromosomal LacI::GFP contained clear foci. In contrast, in the strain containing plasmid encoded LacI::GFP, foci were not observed. Due the copy number of the plasmid increasing the level of LacI::GFP, the number of LacI::GFP molecules exceeded the number of available binding sites. A large proportion of LacI::GFP was therefore unbound in the cell contributing to a very high level of background fluorescence. The molecules of LacI::GFP that were bound to the LacI DNA site array cannot be observed as a focus over this background. Although plasmid encoded LacI::GFP was necessary for experiments using larger repeats of LacI DNA sites, a chromosomal fusion is sufficient for a smaller insert. For all subsequent experiments LacI::GFP was provided as a chromosomal fusion.

3.3 Use of 6 and 22 lac operators to tag chromosomes

To investigate the use of smaller arrays of DNA sites for FROS, arrays of 6 and 22 LacI DNA sites are inserted adjacent to the *araBAD* promoter in strains with LacI::GFP on the chromosome, strains LR18 and LR06 respectively. An array of 22 LacI DNA sites is a chromosomal insert of approximately 850 bp whereas 6 DNA sites is an insert of approximately 250 bp. These strains were checked for the presence of photostable foci. To view both strains simultaneously and compare the foci directly, strain LR06 was stained with Hoechst 33258 and LR18 was stained with FM 4-64. The cells were then washed with PBS to remove any residual stain before being mixed and transferred to a microscope slide to allow the GFP signal from both strains to be observed under the same conditions. Cells were imaged every 30 seconds to test the photostability of the foci. Strain LR06, with 22 DNA sites, had clear foci in most cells, which faded slightly over time but were still visible after being exposed to the lamp 8 times. In strain LR18, with 6 DNA sites, some cells had visible foci although these were not as clear as in LR06. The foci faded very quickly when exposed to the lamp more than once (see Figure 3.1). Although the 22 DNA site array is 500 bp larger, therefore potentially affecting the local architecture of the chromosome more than the 6 DNA site array, the appearance and photostability of the foci are greatly increased making this strain more likely to yield useful data. However, the data produced are likely to be more useful than data from a strain with a larger insert, for example the 10 kb insert widely used previously.

3.4 Activity of LacI::GFP as a transcriptional repressor

In order to function as a chromosomal tag LacI needs to retain its function as a DNA binding protein after it has been fused to GFP. The N-terminus of LacI is known to contain the DNA

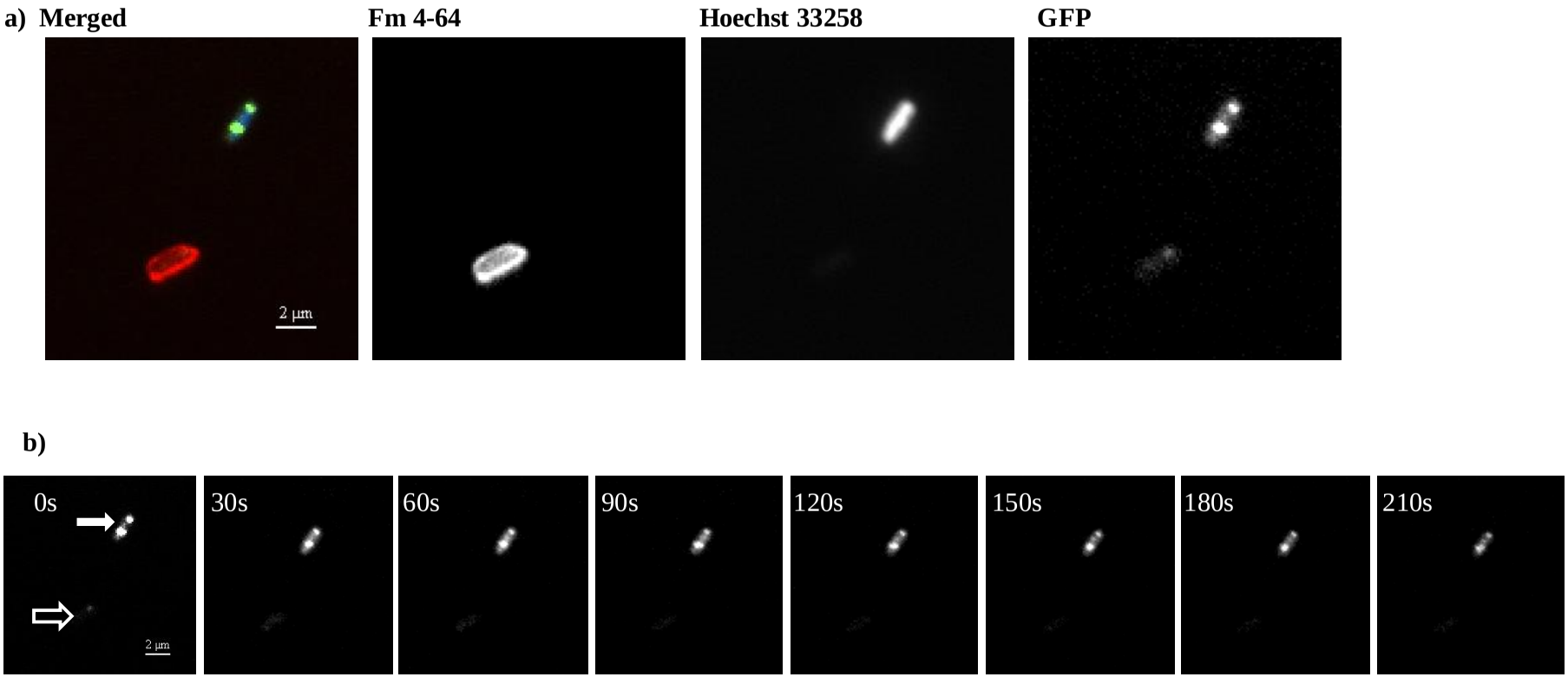


Figure 3.1: Use of 6 and 22 LacI DNA sites to tag the chromosome

a) Strain LR18, with 6 LacI DNA sites adjacent to *araBAD*, was stained with FM 4-64 to stain the membrane. Strain LR06, with 22 LacI DNA sites adjacent to *araBAD*, was stained with Hoechst 33258 to stain the nucleoid. The strains were then mixed and then transferred to the same slide to compare GFP distribution. Red shows the FM 4-64 cellular membrane stain, blue shows the Hoechst 33258 nucleoid stain and green shows the localisation of GFP. The solid arrow marks the position of a cell of strain LR06, the open arrow marks the position of a LR18 cell.

b) An image was taken every 30 seconds on the FITC channel to observe the change in GFP distribution. The solid arrow marks the position of a cell of strain LR06, The open arrow marks the position of a LR18 cell.

binding domain (Daber *et al.*, 2007) so a C-terminal GFP tag was used to reduce the possibility of the tag affecting DNA binding. To test the function of LacI::GFP as a functional repressor, it was used to complement a strain with a *lacI* deletion, KH000. In MG1655 the β-galactosidase activity is high when induced with IPTG and low when uninduced, owing to the repression of *lacZ* by LacI in the absence of IPTG. In KH000, due to the absence of *lac* repressor, the β-galactosidase activity is high both with and without induction. The level of β-galactosidase can be measured by a β-galactosidase assay (see section 2.13). A plasmid containing the LacI::GFP fusion, pLER101, was inserted into KH000 and the effect this had on the level of β-galactosidase was compared to that of a plasmid carrying non-tagged LacI, pET22b.

When KH000 is transformed with plasmid pLER101, encoding LacI::GFP, the repression of *lacZ* is restored in the absence of IPTG induction, which is shown by the low level of β-galactosidase activity (see figure 3.2). Repression was also restored when KH000 was transformed with pLER107, containing a LacI::FLAG fusion, or pET22b, which contains untagged LacI. pLER107 was included as the FLAG tag is much smaller than GFP therefore may have less affect on the properties of the protein. When the cultures were induced with IPTG, an increase in β-galactosidase activity was seen but for the strains containing plasmid encoded LacI with 3 different tags, this was only to around 50% of the level seen in MG1655 (see figure 3.2). As all 3 plasmids gave very similar results, it is likely that the lower level of β-galactosidase activity seen was not due to the presence of a tag on the LacI protein, and perhaps the increased copy number of LacI protein in the cell. This concentration of IPTG may not be high enough to allow for the increased number of LacI molecules. All plasmid-expressed LacI, with different tags, restored repression of *lacZ* in the absence of IPTG,

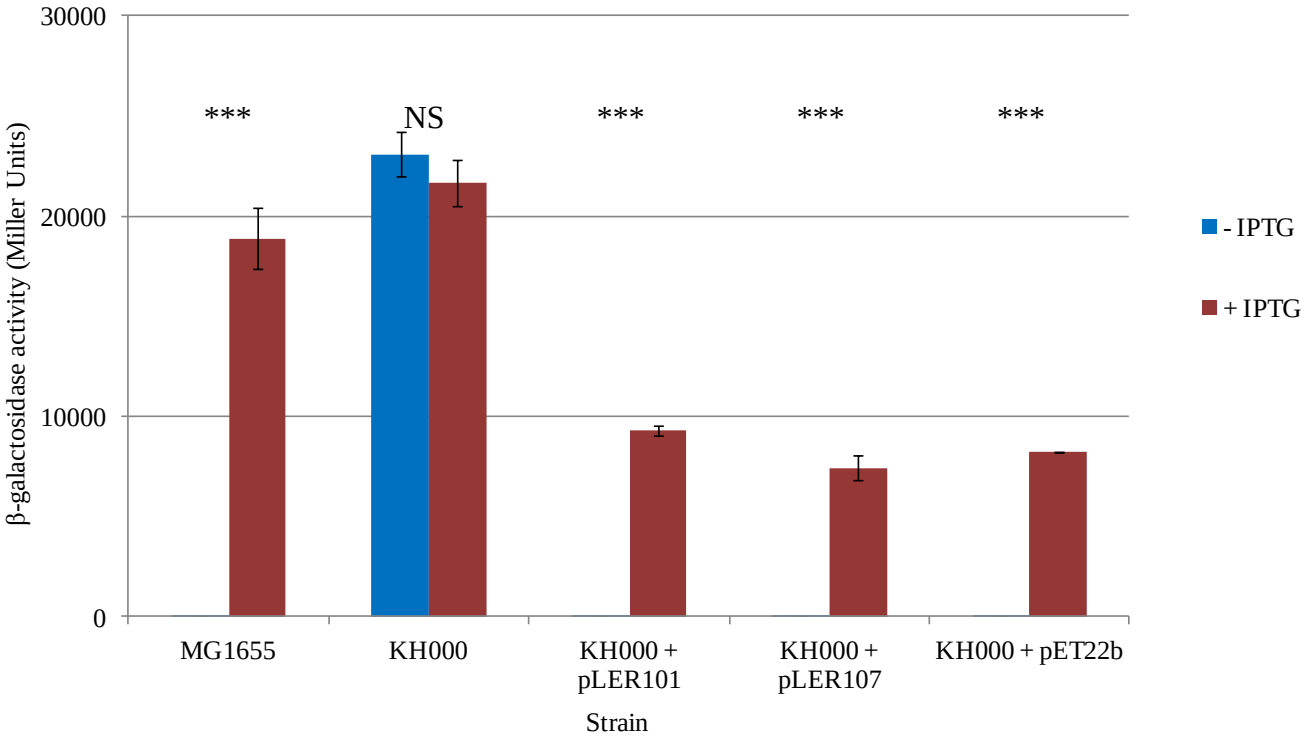


Figure 3.2: Activity of LacI::GFP as a transcriptional repressor

Strain KH000, a derivative of MG1655, has a deletion in the *lacI* gene meaning there is high β-galactosidase activity without induction by IPTG. This figure shows the repression in the absence of IPTG when LacI was added to the cell in plasmids. pLER101 contains LacI::GFP, pLER107 contains LacI::FLAG and pET22b contains wild type LacI. Strains were grown to mid-exponential phase (OD_{650} 0.3-0.5) in M9 salts medium supplemented with 0.3% fructose in the presence and absence of 0.1 mM IPTG. Data shown are representative of results from three independent experiments and error bars indicate one standard deviation from the mean. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = 9.

demonstrating that tagged LacI functions as well as non-tagged LacI as a transcriptional repressor. The GFP tag has not affected the DNA binding abilities of LacI.

3.5 Suppression of the formation of LacI::GFP foci

3.5.1 Suppression by IPTG

To confirm that the GFP foci seen in cells are due to multiple copies of LacI::GFP binding at the LacI DNA site array, an attempt was made to suppress the formation of foci by the addition of IPTG. IPTG is an inducer of *lacZ* so, in its presence, LacI should not bind its DNA sites. Strain LR06, with a chromosomal LacI::GFP fusion and 22 LacI DNA sites inserted adjacent to *araBAD*, was grown with varying concentrations of IPTG and the percentage of cells containing foci recorded. The addition of increasing concentrations of IPTG did not decrease the percentage of cells containing at least one focus ($P > 0.05$) (figure 3.3). The percentage of cells containing foci was between 87% and 96% for concentrations up to 1000 μM. This agrees with the result in section 3.4, where cells containing plasmid encoded LacI did not show full induction in the presence of IPTG, and indicates that IPTG bound LacI will still bind to its DNA site.

3.5.2 Suppression by non-tagged LacI

As an alternative method for suppressing the formation of LacI::GFP foci, non-tagged LacI was introduced to the cells in excess to compete with LacI::GFP. Plasmid pET22b was used to provide non-tagged LacI in excess and was compared to a control plasmid, pET20b, not encoding LacI. Both were transformed into strain LR06, which has a chromosomal LacI::GFP fusion and 22 LacI DNA sites inserted adjacent to *araBAD*, and is known to have foci in the

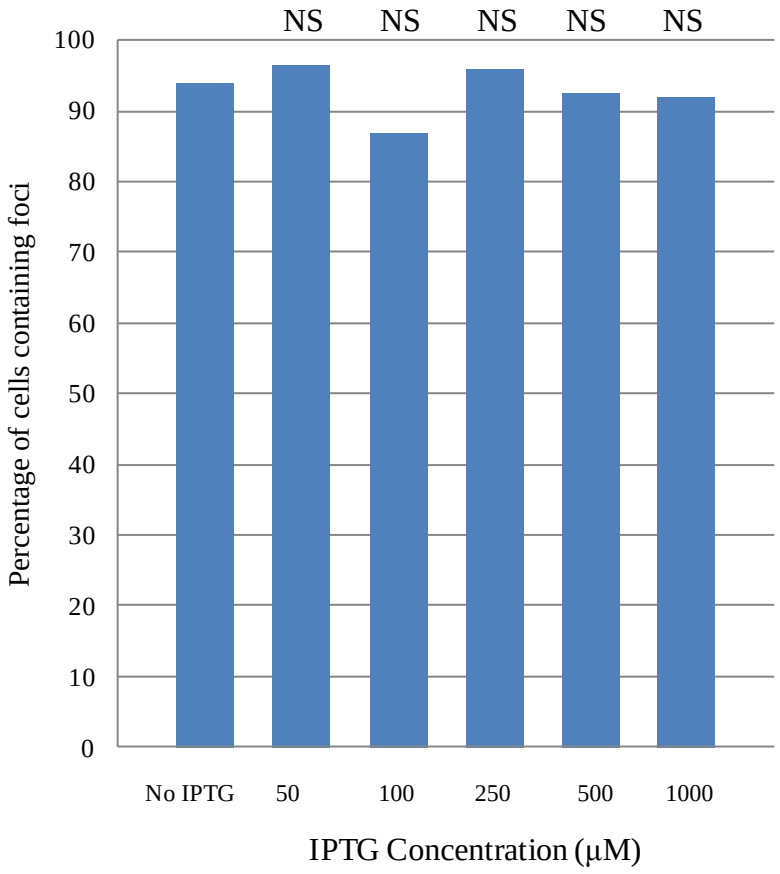


Figure 3.3: Repression of LacI::GFP foci formation with increasing concentrations of IPTG

Strain LR06, tagged with 22 LacI DNA sites adjacent to *araBAD* and carrying a chromosomal LacI::GFP fusion, was grown to an OD_{650} of approximately 0.1 in M9 salts medium supplemented with 0.3% fructose supplemented with varying concentrations of IPTG. The number of cells containing one or more foci was counted. A star above a bar represents data that is significantly different to unsupplemented cells (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different, n = > 200 cells.

majority of cells viewed. Cells with an excess of non-tagged LacI were observed to fall into 3 categories, uniform background, fuzzy foci and clearly defined foci (see figure 3.4). In the presence of non-tagged LacI, the number of cells containing a focus dropped from 93% to 15%, of which 10% were fuzzy foci and 5% were clearly defined. The majority of cells had a uniform background (see figure 3.5). This suggests that LacI::GFP was not localised to any particular area of the cell, probably because non-tagged LacI was occupying the majority of the DNA sites. Freely diffusing LacI::GFP is then observed as a uniform background. Around 5% of the cells had at least one sharp focus of GFP as typically seen in LR06, where LacI::GFP is still occupying the majority of the DNA sites. This is a considerable reduction from the 93% of cells containing foci when the control plasmid was present. The 10% of cells with fuzzy foci can be explained by the possibility of the LacI DNA site array being occupied by both LacI::GFP and non-tagged LacI. This would cause GFP molecules to accumulate but in a smaller number than usual, meaning the focus formed is not as clear and photobleaches faster. The remainder of LacI::GFP molecules are unbound in the cell contributing to the background fluorescence and making the focus less clear. This “accumulation of GFP” is similar to the focus formed when 6 LacI DNA sites were used (see figure 3.1). In conclusion, the reduction in the number of cells containing clear foci in the presence of an excess of non-tagged LacI shows that the foci were due to binding of LacI::GFP to the multiple LacI DNA sites inserted onto the chromosome.

3.6 Use of MalI as a chromosomal tag

Often LacI::GFP is used in combination with another repressor protein fused to a different colour fluorescent protein to allow two points on the chromosome to be labelled in the same cell. In most previous studies, the *tet* repressor, TetR, was used in combination with multiple

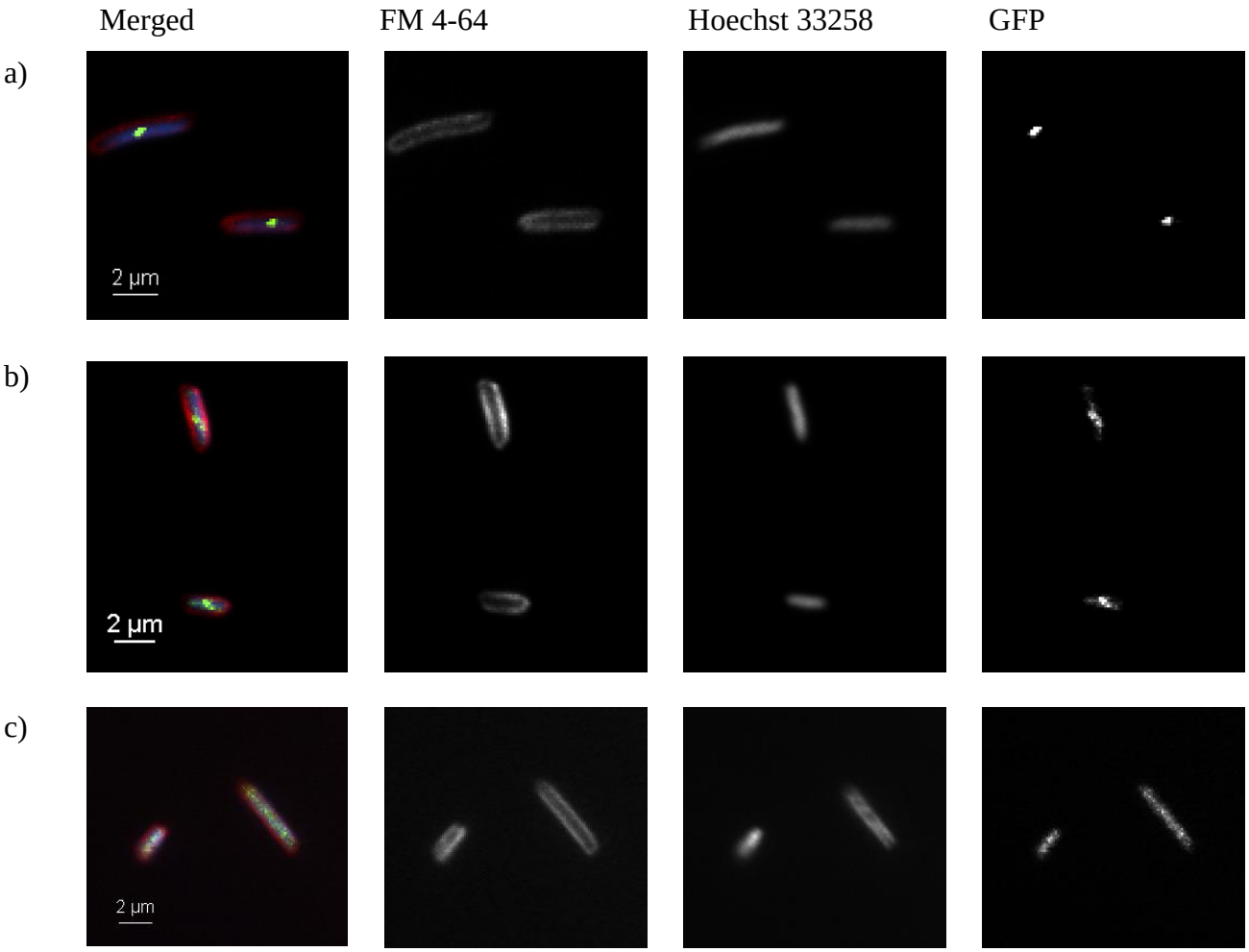


Figure 3.4: Cells observed when LacI::GFP foci formation was repressed with an excess of non-tagged LacI

LR06, carrying an insert of 22 LacI DNA site adjacent to *araBAD* and a chromosomal fusion of LacI::GFP, was transformed with either pET22b, carrying non-tagged LacI, or pET20b, empty vector. Cells were grown in M9 minimal media supplemented with 0.3% fructose, supplemented with 40 µg/ml ampicillin, at 23ºC until they reached an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid and FM 4-64 was used to stain the membrane. Cells observed fell into 3 categories:

a) LR06 with pET20b showing clear foci.

b) LR06 with pET22b showing a fuzzy focus.

c) LR06 with pET22b showing a uniform background of GFP.

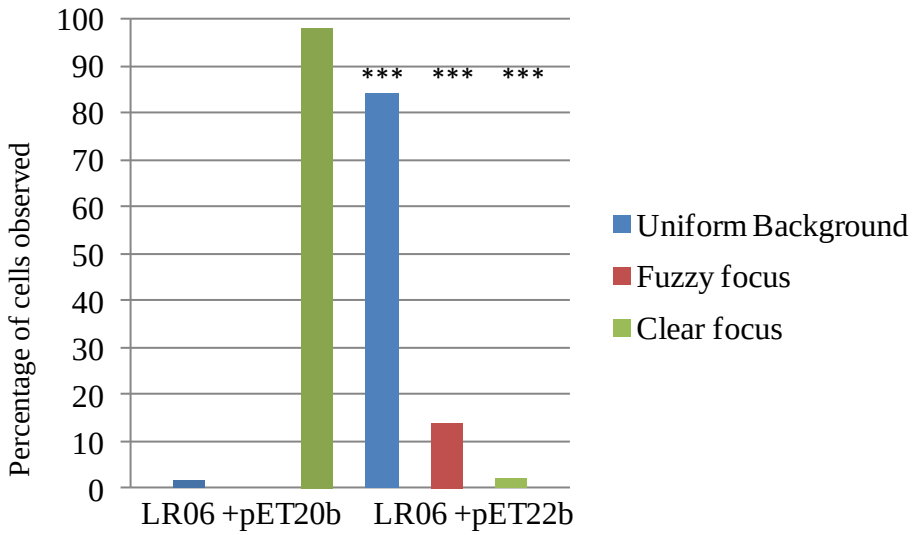


Figure 3.5: Repression of LacI::GFP foci formation with an excess of non-tagged LacI::GFP

Strain LR06, with an insert of 22 LacI DNA sites adjacent to *araBAD* and a chromosomal LacI::GFP fusion, transformed with pET22b, encoding non-tagged LacI, or pET20b, empty vector. Cells were grown to an OD_{650} of approximately 0.1 in M9 salts medium supplemented with 0.3% fructose supplemented with 40 µg/ml ampicillin and the distribution of GFP in cells was recorded. A star above a bar represents data that is significantly different to cultures with empty vector (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 500

copies of its DNA sites. As an alternative option for a repressor protein to be used as a chromosomal tag, a system was developed using MalI and multiple copies of the MalI DNA site.

A MalI::mCherry fusion was created, to be used in combination with LacI::GFP. The fusion was under the control of the *malI* promoter and carried in a multicopy plasmid, pACYC184 derivative pLER104, with a copy number of 10-15. This was introduced into strain MG1655 to check that mCherry was fluorescent and whether MalI was evenly distributed across the cell, as any localisations would make this fusion unsuitable for use in tagging the chromosome. Although a uniform background was expected, this was not seen and mCherry appeared to be localising in one pole of each cell. This strain contains the non-tagged *malI* gene on the chromosome, as well as 2 MalI DNA sites. It was thought that the presence of non-tagged MalI would not affect future experiments, as MalI::mCherry is encoded on a plasmid and therefore will be in excess over the non-tagged version. Also, as there are only 2 MalI DNA sites on the chromosome this is unlikely to form a focus of mCherry that can be seen using fluorescence microscopy. As the MalI::mCherry fusion is under the control of the *malI* promoter there is one complete and one partial MalI binding site on the plasmid. There is evidence that plasmids containing an active promoter cluster to one pole of the cell, where it is suggested transcription may be favourable (Sánchez-Romero *et al.*, 2012). Although 2 copies of the DNA site would not be enough to form a focus under normal conditions, when 15-20 copies of plasmid pLER104, each carrying 2 MalI DNA sites, are clustered at the pole of the cell, enough binding sites are in the same place to make a visible mCherry focus (see figure 3.6a).

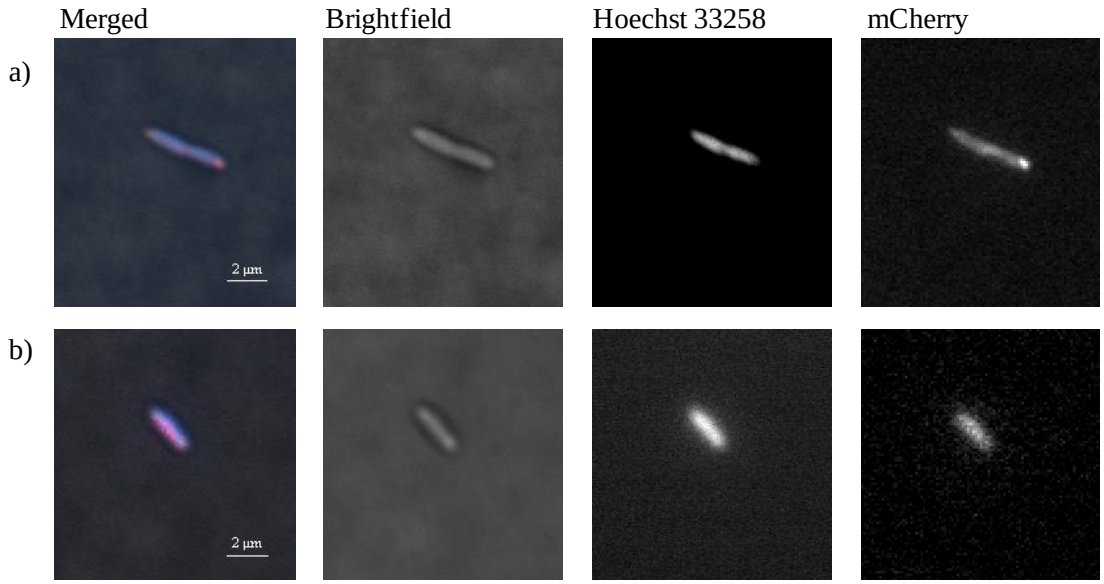


Figure 3.6: Visualisation of MalI::mCherry

a) MalI::mCherry under the control of the *malI* promoter expressed from plasmid pLER104, a pACYC184 derivative, in MG1655. Cells were grown in M9 minimal media supplemented with 0.3% fructose at 23ºC until they reached an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid.

b) MalI::mCherry under the control of the *melR* promoter expressed from plasmid pLER108, a pACYC184 derivative, in MG1655. Cells were grown in M9 minimal media supplemented with 0.3% fructose at 23ºC until they reached an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid.

To resolve this issue, the MalI::mCherry fusion was placed under the control of the *melR* transcription initiation region in plasmid pJW15Δ100 to make plasmid pLER105. The *melR* promoter and *malI::mCherry* fusion were transferred back into pACYC184 to make plasmid pLER108. Figure 3.6b shows that when this plasmid was transformed into MG1655 and visualised by microscopy MalI::mCherry was distributed evenly across the cell. The presence of 2 MalI DNA sites on the chromosome should not be a problem for future experiments and this MalI::mCherry fusion should be suitable for use in FROS experiments. Before the MalI::mCherry fusion can be used for FROS experiments, it needs to be confirmed that the tag has not affected MalI function. This was done by comparing the repression of the *malX* promoter by MalI::mCherry to that of non-tagged MalI in a β-galactosidase assay, as previously tested by Lloyd *et al.* (Lloyd *et al.*, 2010). A strain with *lacZ* deleted from the chromosome, KH001, was used for the β-galactosidase assay. It was co-transformed with plasmid pRW50malX100, containing a fragment of the *malX* promoter fused to *lacZ*, and a second plasmid either carrying MalI or empty vector. pACYCΔHN is empty vector and pACYCMalI carries non-tagged MalI. pLER104 was used to provide MalI::mCherry as it carries MalI::mCherry under the control of the *malI* promoter, giving the closest comparison to the levels of non-tagged MalI in the cell produced by pACYCMalI. On average non-tagged MalI represses the *malX* promoter by a factor of 24 and MalI::mCherry represses by a factor of 31 (see figure 3.7). These similar figures suggest that MalI::mCherry is able to behave in a similar way to non-tagged MalI. Repression of the *malX* promoter depends upon MalI binding to the DNA so the mCherry tag has not affected the DNA binding abilities of MalI.

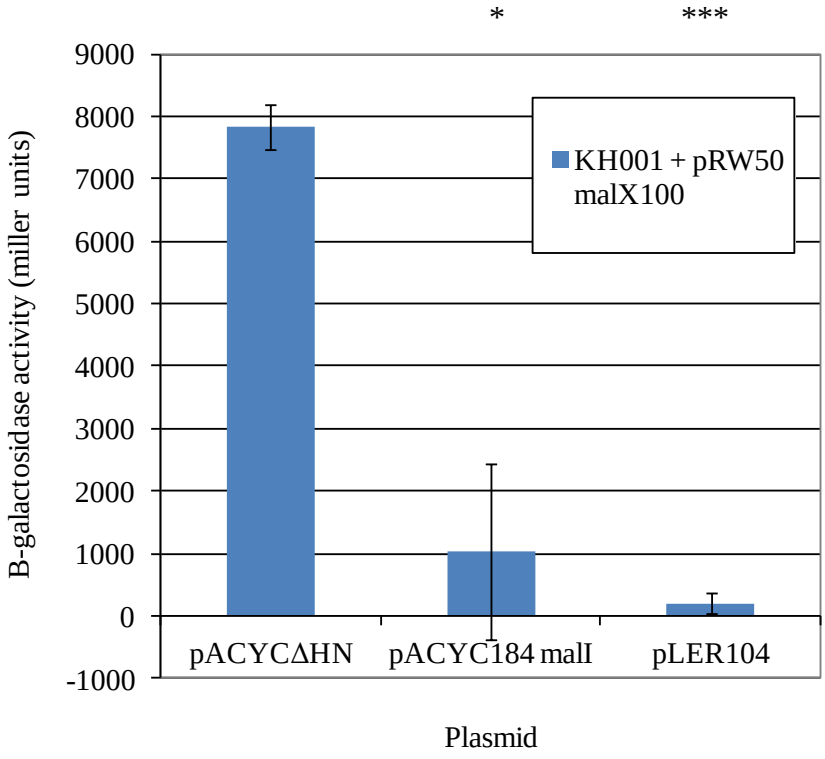


Figure 3.7: Repression of *malX* promoter with different forms of MalI

B-galactosidase activity was measured from a fusion of a fragment of the *malX* promoter to *lacZ* in the plasmid pRW50 malX100. Strain KH001 was transformed with pRW50malX100 and either pACYCΔHN, empty vector, pACYCMalI, encoding non-tagged MalI, or pLER104, encoding MalI::mCherry. Cultures were grown aerobically at 37ºC to mid-exponential phase (OD_{650} 0.3-0.5) in minimal salts medium supplemented with 0.3% fructose. pACYCMalI carries the wild type MalI and pLER104 carries MalI tagged to mCherry. pACYCΔHN was included as a negative control. Data shown are representative of three independent experiments, and error bars indicate one standard deviation from the mean. A star above a bar represents data that is significantly different to cultures with empty vector (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = 9.

3.7 Comparison of LacI and MalI as FROS reporters

To compare the repressor proteins, LacI and MalI, as FROS reporters, 2 strains were studied which both had inserts at exactly the same location, adjacent to *araBAD*. LR06 has 22 LacI DNA sites as well as chromosomal fusion of LacI::GFP. Strain LR17 has 20 MalI DNA sites and is transformed with plasmid pLER108, carrying a MalI::mCherry fusion. LR17 carrying the plasmid is referred to as LR17-pLER108. The strains were grown as previously described and cells were visualised. Clear foci of either GFP or mCherry were seen in both strains. The number of foci seen per cell was counted for both strains (see figure 3.8a). The majority of the cells from both strains had one focus per cell, with around 30% having 2 foci and a small percentage each having 3 or no foci. The LacI::GFP and MalI::mCherry reporters gave very similar results suggesting that the repressor protein::fluorescent protein combination used in FROS does not affect the result.

For the cells containing 1 or 2 foci the average position of foci was then calculated. For a single focus, the distance from the centre of the focus to the nearest cell pole was measured. If a cell contained two foci, the focus closest to a cell pole was measured first ("1st of 2 foci") and then the further away focus was measured from the same pole ("2nd of 2 foci"). These measurements were then divided by the cell length to give a position in the cell for each focus relative to cell length. Although the positions of foci varied, the averages calculated for cells with 1 or 2 foci were very similar for the strains tagged with LacI::GFP and MalI::mCherry (see figure 3.8b). Since both tags were inserted at exactly the same place on the chromosome they should give the same result unless the tag itself has some affect on the cell or the chromosome structure. LacI and MalI as FROS reporters gave results with no significant difference ($P > 0.05$) for both the number of foci per cell and the average position of foci

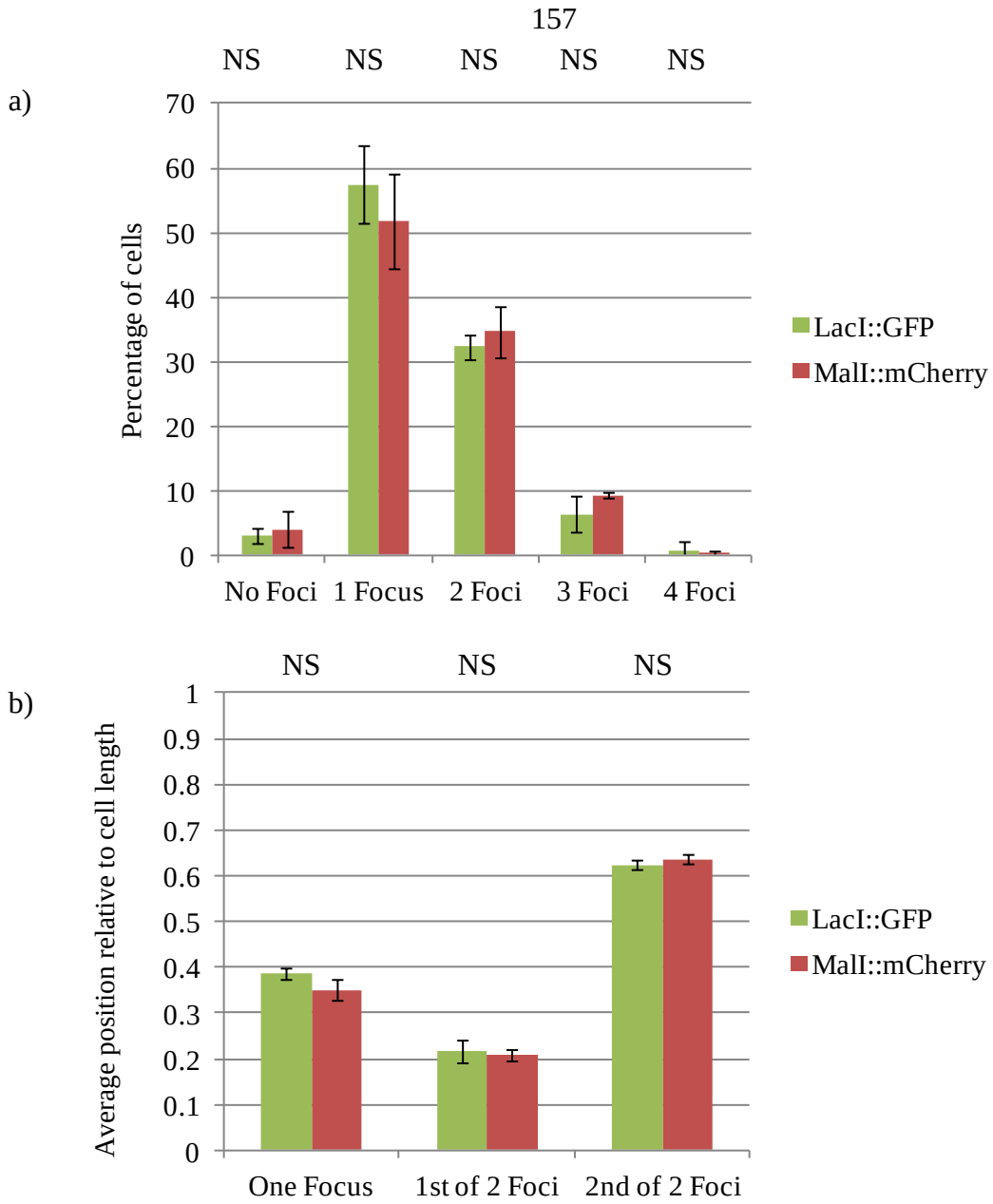


Figure 3.8: Comparison of the number of foci seen in LR06 and LR17 cells

22 copies of the LacI DNA site were inserted adjacent to the *araBAD* promoter in a strain also carrying a chromosomal fusion of LacI::GFP, strain LR06. 20 copies of the MalI DNA site were also inserted at the same location in a strain, LR17, which was then transformed with a plasmid encoding MalI::mCherry, pLER108. Cells were grown to an OD_{650} of approximately 0.1 in M9 minimal medium supplemented with 0.3% fructose, supplemented with 17.5 µg/ml chloramphenicol for LR17- pLER108.

a) A bar chart to show the percentage of cells containing different numbers of foci.

b) A bar chart to show the distance from each focus to the nearest pole of the cell was measured and represented relative to cell length. For cells with 2 foci, an average position was calculated for the focus closest to a cell pole ("1st of 2 foci") and the further away focus ("2nd of 2 foci").

Data shown are averages from 3 independent experiments and error bars show one standard deviation from the mean. A star above a bar represents data that is significantly different (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = > 400 cells.

This confirms that MalI, a novel FROS reporter, gives comparable results to LacI, and that LacI and MalI could be used in the same cell to tag different locations in future experiments.

3.8 Effect on growth rate of inserting multiple repressor DNA sites at various positions on the chromosome

3.8.1 Effect on growth rate of inserting multiple LacI DNA sites onto the chromosome

To investigate whether inserting multiple LacI DNA sites, which will be occupied by LacI::GFP, onto the chromosome, has caused a growth defect, strains with inserts at various places were compared to MG1655. All strains are derivatives of MG1655 with inserts at various places. A strain with a chromosomal LacI::GFP fusion (DL02) was included to confirm that this modification does not affect growth rate. Strains studied have this chromosomal insert and 22 LacI DNA sites inserted adjacent to the promoters of *araBAD* (LR06), *araJ* (LR39), and *dps* (SXB4). Cultures were grown at 37ºC in M9 minimal media supplemented with 0.3% fructose. OD_{650} was measured every hour for 3 biological repeats of each strain until the cultures began to reach stationary phase. All 5 strains had a very similar growth rate (see figure 3.9) suggesting that neither the LacI::GFP fusion nor the 22 LacI DNA site array have a detrimental effect on growth rate when inserted at these sites on the chromosome.

3.8.2 Effect on growth rate of inserting multiple MalI DNA sites onto the chromosome

To investigate whether the insertion of multiple MalI DNA sites onto the chromosome has affected the growth of the strains, growth rates of strains with inserts at various positions were compared to that of MG1655. Strains with 20 MalI DNA sites inserted adjacent to *araBAD* (LR17), *araFGH* (LR38), and *mntH* (SXB3) were compared. Also included was a

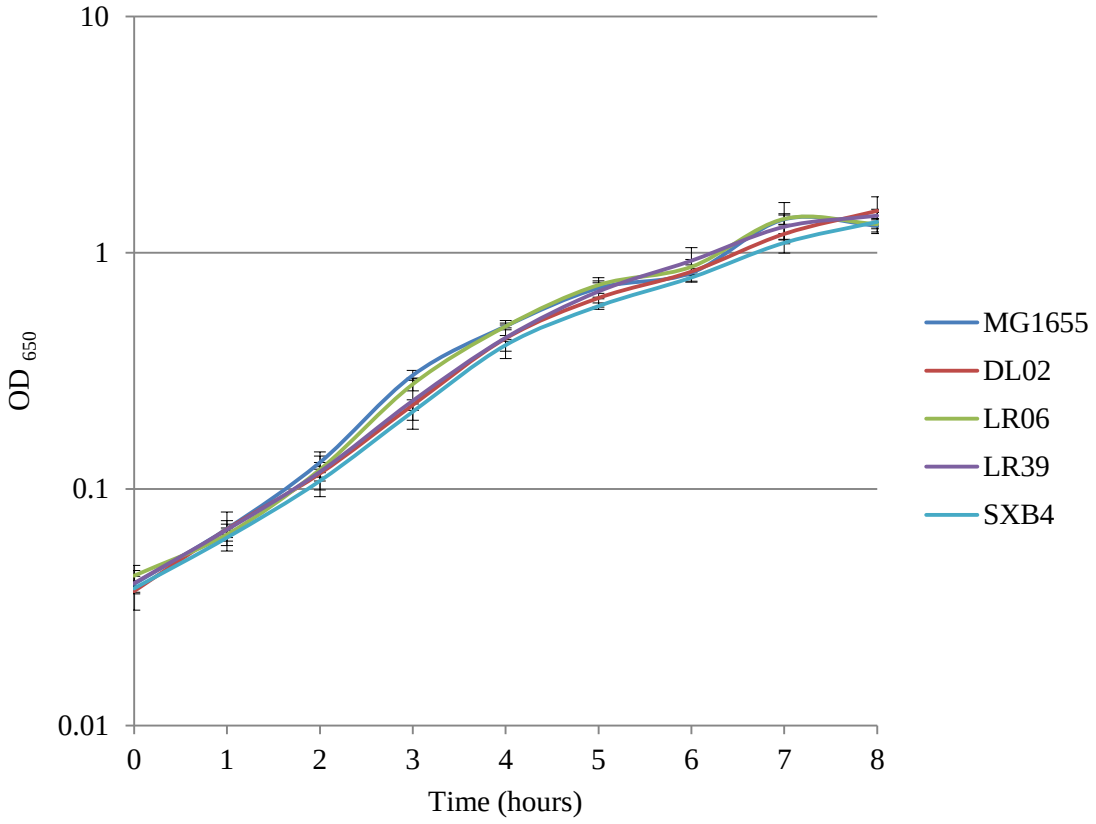


Figure 3.9: Growth rate of cells with inserts of 22 LacI DNA sites at various sites on the chromosome

The growth rate of 4 strains with different chromosomal inserts was compared to that of wild type, MG1655. The strain DL02 has a fusion of LacI::GFP on the chromosome. This fusion is then combined with inserts of 22 LacI DNA sites adjacent to *araBAD* (LR06), *araJ* (LR39) and *dps* (SXB4). Cultures were grown at 37ºC in M9 minimal medium supplemented with 0.3% fructose. OD_{650} was measured every hour for 8 hours.

strain combining both FROS reporters, carrying 20 MalI DNA sites adjacent to *araJ*, 22 LacI DNA sites adjacent to *araBAD* and a chromosomal LacI::GFP fusion (LR31). All 5 strains were transformed with the plasmid pLER108, carrying a MalI::mCherry fusion. Strains transformed with the plasmid are referred to as the strain name-pLER108. Cultures were grown at 37ºC in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol. OD_{650} was measured every hour for 3 biological repeats of each strain until the cultures began to reach stationary phase. Strains LR17-pLER108, LR39-pLER108 and LR31-pLER108, tagged with MalI at *araBAD*, *araFGH* and *araJ* respectively, showed a slightly slower growth rate when compared to MG1655-pLER108 but the difference was small and the cultures still appeared to be growing normally (see figure 3.10). Strain SXB3-pLER108, tagged at *mntH*, grew at a very similar rate to MG1655-pLER108. Overall, the reduction in growth rate in strains containing a chromosomal insert of multiple MalI DNA sites was small. This also shows that a strain with inserts of LacI and MalI DNA sites at separate locations, a chromosomal LacI::GFP fusion and a plasmid carrying MalI::mCherry, LR31-pLER108, grows at a normal rate allowing for the possibility of using the LacI and MalI FROS in the same cell to tag 2 different locations.

3.9 Discussion

3.9.1 Use of LacI as a FROS Reporter

The use of Fluorescent Repressor/Operator Systems (FROS) for studying chromosome structure in eukaryotic and bacterial cells in several different model organisms is well established. Previous studies have used an array of 256 LacI DNA sites to tag a location of interest on chromosomes and a plasmid encoded LacI::GFP fusion to visualise it. As this project aims to study the effect of promoter induction on chromosome structure, disruption to

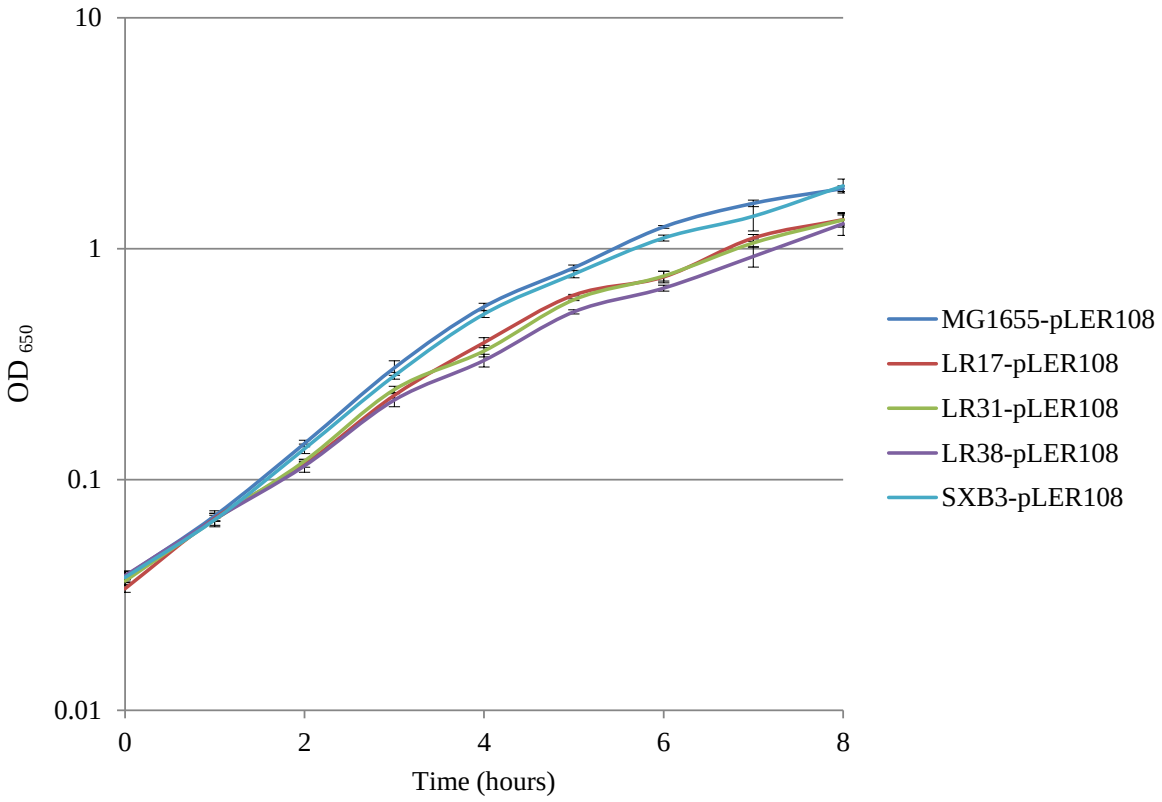


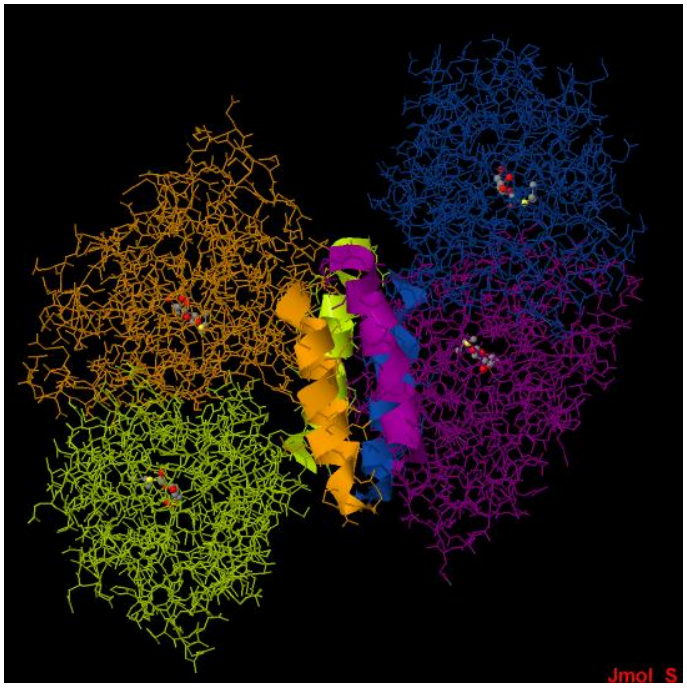
Figure 3.10: Growth rate of cells with inserts of 20 MalI DNA sites at various sites on the chromosome

The growth rate of 4 strains with different chromosomal inserts was compared to that of wild type, MG1655. Inserts of 20 MalI DNA sites have been made adjacent to *araBAD* (LR17), *araFGH* (LR38) and *mntH* (SXB3). Strain LR31 has a chromosomal fusion of LacI::GFP as well as inserts of LacI and MalI DNA sites adjacent to *araBAD* and *araJ* respectively. All strains were transformed with plasmid pLER108, carrying MalI::mCherry. Cultures were grown at 37ºC in M9 minimal medium supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol. OD_{650} was measured every hour for 8 hours.

the chromosome needs to be as small as possible. An array of 256 LacI DNA sites is an insert of 10 kb, which is 10 times the average length of an *E. coli* open reading frame, 951 bp (Blattner *et al.*, 1997). In addition to the DNA being inserted into the chromosome is the fact that it will be highly occupied by protein, LacI::GFP, which could further disrupt the chromosome. In this project, a smaller insert of 22 LacI DNA sites, around 1 kb, was used to tag positions of interest. The reduction in the number of DNA sites meant that less LacI::GFP molecules were needed to saturate the DNA sites. Plasmid encoded LacI::GFP is present at a much higher copy number due to the multiple copy number of the plasmid. Once the 22 LacI DNA sites are bound, any remaining LacI::GFP is freely diffusing, adding to background fluorescence. To reduce the copy number of LacI::GFP, the fusion was inserted onto the chromosome, under the control of the *lacI* promoter, where it would presumably be expressed at a similar level to non-tagged LacI. There are thought to be about 20 monomers of LacI per copy of the gene in a "normal" *E. coli* cell (Gilbert and Müller-Hill, 1966), which would bind to 10 DNA sites as dimers. It is not known whether it is possible for all 22 LacI DNA sites in the array to be bound simultaneously, due to steric hindrance, but it is likely that more of the 10 LacI::GFP dimers will be bound to the array rather than freely diffusing. The reduction in the percentage of unbound LacI::GFP, and therefore the reduction in background fluorescence, makes the foci much clearer. Although a smaller chromosomal insert is preferable, an attempt to use 6 LacI DNA sites for FROS gave less clear foci that photobleached very quickly (figure 3.1). As a compromise, the array of 22 LacI DNA sites was selected for further study. However, this reduction in the number of DNA sites compared to the usual 256 meant that time lapse experiments were not possible due to increased photobleaching.

Several experiments showed that LacI::GFP did not have the expected response to IPTG. When cells were induced with IPTG to suppress foci formation, no reduction in the number of cells with foci was seen (figure 3.3). Expression of *lacZ* is induced by the presence of IPTG, so LacI should not bind to its DNA site. This was also seen in the β-galactosidase assay comparing tagged and non-tagged LacI. In this case, plasmids encoding LacI::GFP, LacI::FLAG and non-tagged LacI returned repression to a *lacI*- strain in the absence of IPTG (figure 3.2). However, in IPTG induced cultures the β-galactosidase activity was around 50% lower than the level seen in wild type cultures for all 3 plasmids. This suggests that, even in the presence of IPTG, LacI::GFP does not fully release its DNA site. In this case, LacI::GFP was encoded on a multicopy plasmid, which will have increased its cellular concentration. Also, cells containing the plasmid encoding non-tagged LacI also gave a lower level of expression upon induction, suggesting that the discrepancy was not due to the GFP tag. Kinetic studies using equilibrium dialysis have shown that LacI releases its DNA site upon induction by IPTG, as the affinity of LacI for its DNA site is reduced 1000 fold (Barkley *et al.*, 1975). This is thought to be because ITPG binding disrupts the intersubunit interactions which are required for binding to the operator (Falcon and Matthews, 1999). The IPTG binding site in LacI is in the centre of the molecule, while the GFP tag is at the C-terminus (Lewis *et al.*, 1996) (see figure 3.11). Residues 340 to 357 of LacI, which are at the extreme C-terminus, form an α helix which associates with the α helix of another LacI monomer for dimerisation and then tetramerisation. In the LacI::GFP fusion, residues 359 to 361 have been removed and replaced with GFP (Lee *et al.*, 2009). It is possible that the α helix could have been disrupted by the addition of GFP meaning LacI can no longer form dimers or tetramers. The presence of four GFP molecules at the site of tetramerisation is likely to hinder the formation of tetramers, but it is thought LacI::GFP can still dimerise, as it is still functional as a transcriptional repressor. The residues responsible for IPTG binding have been identified

a)



b)

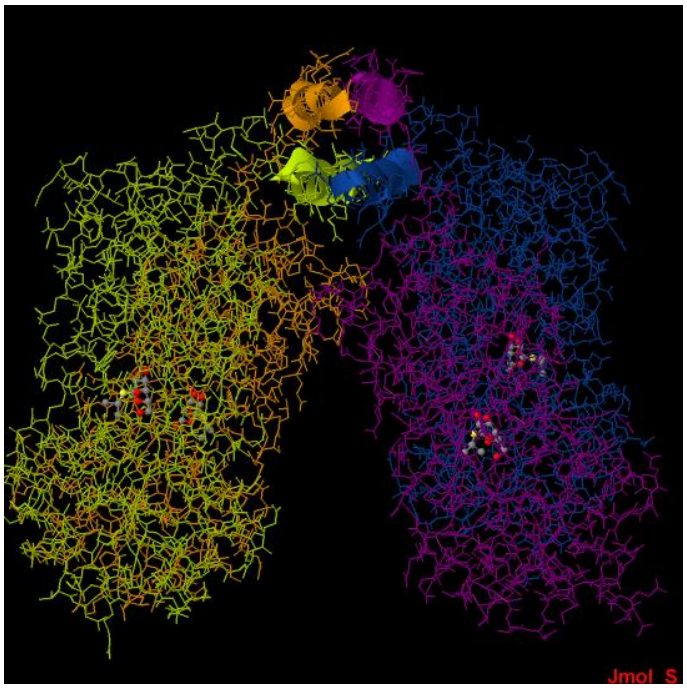


Figure 3.11: Structure of the LacI tetramer bound to IPTG

Each monomer of LacI is shown in a different colour. Residues 1-339 are shown in wireframe and residues 340-361 are shown as an α helix which associate to form the tetramer. IPTG is shown in its binding pocket as a stick and ball model. The DNA helix would be positioned at the bottom of figure b and the back of figure a.Figures a and b show the same structure with 90º rotation. Figure adapted from Lewis *et al.*, 1996.

and none of the residues in the C-terminal α helix are involved (Daber *et al.*, 2007), so there is no obvious reason why the addition of a GFP tag to LacI should change its response to IPTG. The tetramerisation of LacI::GFP could be tested using *in vitro* kinetics to measure DNA loop formation (Finzi and Gelles, 1995). An excess of non-tagged LacI could be used to suppress the formation of LacI::GFP foci. This reduced the percentage of cells containing at least one focus from 93% to 15%, showing that the foci are formed by LacI::GFP binding to the multiple DNA site array. Of the 15% of cells containing foci, only one third had a clear focus. The remainder of cells had foci that appeared fuzzy and fainter. It is thought that fuzzy foci are seen when the 22 LacI DNA site array is partially bound by LacI::GFP and partially bound by non-tagged LacI. This results in foci containing less molecules of GFP than usual.

3.9.2 MalI as a FROS reporter

LacI is often used as a FROS reporter in combination with a second repressor, fused to a different fluorescent protein, to allow two points in the same chromosome to be visualised at the same time. In the past this has been done using the *tet* repressor, TetR, and 240 copies of the TetR DNA site (Lau *et al.*, 2003). As an alternative option, the transcription repressor MalI was used as a FROS reporter using a smaller insert of 20 MalI DNA sites. MalI was tagged to mCherry and supplied to the cell encoded on a multicopy plasmid. The *malI* promoter could not be used to express MalI::mCherry because MalI autoregulates and so binds to its own promoter. Plasmids with an active promoter have been shown to cluster at one pole of the cell (Sánchez-Romero *et al.*, 2012), and since each plasmid is bound by one copy of MalI::mCherry, this results in an mCherry focus at one pole of most cells. This problem was avoided by putting the MalI::mCherry fusion under the control of the *melR* promoter in plasmid pJW15Δ100. Under the control of *melR* MalI::mCherry no longer formed foci at cell poles, but the cellular concentration was too high. This gave too much background

fluorescence and foci could not be seen in strains with the multiple MalI DNA site array inserted. To reduce the level of MalI::mCherry in the cell, the fusion was transferred into a lower copy number plasmid, which would reduce the copy number of the gene. pJW15Δ100 is a derivative of pBR322 which has a copy number of about 20. When the fusion under the control of the *melR* promoter was inserted into pACYC184, with a copy number of 10-15, foci were seen in strains with the MalI DNA site array. MalI::mCherry was shown to be able to repress the *malX* promoter, showing that the DNA binding abilities of MalI are unaffected by the mCherry tag.

As MalI had not previously been used as a FROS reporter, LacI and MalI DNA sites were used to label the *araBAD* promoter on the *E. coli* chromosome, and the results compared. When the number of foci per cell was counted there was no significant difference between the two FROS reporters. The average position of foci in cells with one or two foci was also calculated, and there was no significant difference between the two reporters. This indicates that the reporter used for FROS experiments does not change the result, and therefore LacI and MalI can be used in the same experiments. This enables two locations to be studied in the same cell, one tagged with LacI::GFP and one with MalI::mCherry, and the positions compared.

3.9.3 Asymmetry of two foci

The position of foci representing the position of the *araBAD* promoter were analysed in cells with 2 foci. It was found that the positions of these foci were not symmetrical, averaging at 0.2 and 0.6 compared to cell length. Logically, it may be expected the replicated *araBAD* locus would segregate equally, with the two halves of the cell as mirror images, but evidence

from studies of chromosome structure has shown that this is not the case (Nielsen *et al.*, 2006, Wang *et al.*, 2006). As the chromosome is replicated, the two newly replicated chromosomes are segregated into opposite halves of the cell. *oriC* is positioned approximately at the one quarter and three quarter positions and the left and right replichores are arranged along the cell length, left, right, left, right (section 1.3) (Wang *et al.*, 2006, Reyes-Lamothe *et al.*, 2008b). As a result, the region surrounding the two copies of *oriC* is an equal distance from mid-cell into each half of the cell, but most other locations are not. *araBAD* is located in the Non-structured Right macrodomain, in the right replichore. If there are two foci in the cell the region of the chromosome containing *araBAD* must have been replicated, and the sister chromatids are known to segregate rapidly in the NSR macrodomain. One copy of *araBAD* will be located close to the cell pole, but the other will be nearer to mid-cell. Figure 3.12 shows how the average positions of the *araBAD* promoter, measured using LacI and MalI FROS, fit with the expected arrangement of DNA in the cell.

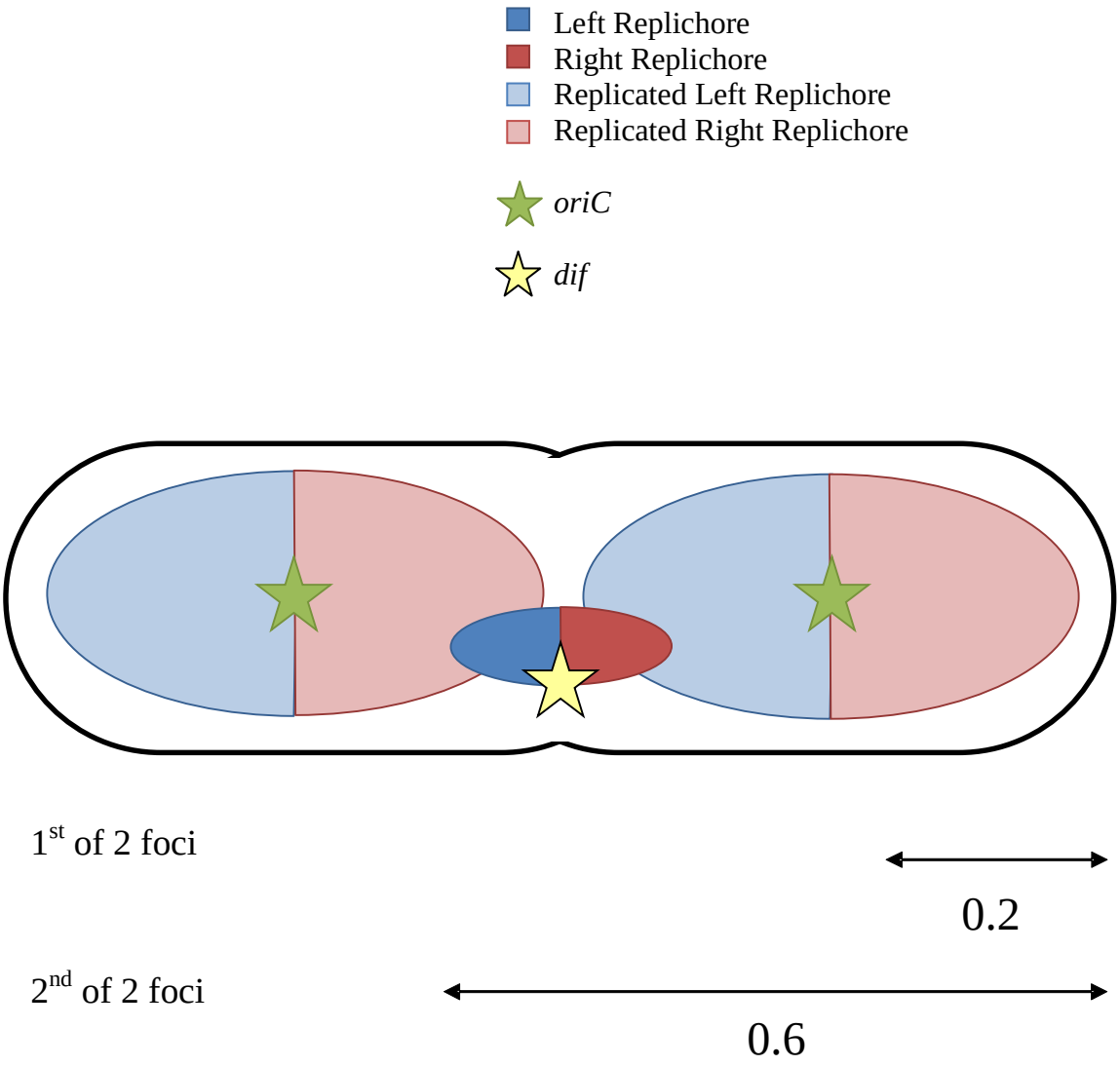


Figure 3.12: A model of chromosome segregation in a bacterial cell compared to the average position of the *araBAD* promoter

A model of a bacterial cell, with the majority of the chromosome replicated, showing the segregation of the left and right replichores into the future daughter cells. Arrows represent the average relative distance from cell pole to focus representing the *araBAD* promoter in cells with 2 foci. The focus closest to a cell pole was called "1st of 2 foci" and distance was measured from this pole to the centre of the focus. The other focus was called "2nd of 2 foci" and distance was measured from the same pole to the centre of the focus. All distances are represented relative to cell length. *araBAD* is located in the right replichore. Figure adapted from Nielsen *et al*. 2006.

4. Studies on *E. coli* promoters regulated by AraC and MntR

4.1 Introduction

4.1.1 The role of transcription factors in chromosome structure

Proteins categorised as NAPs clearly have a role in transcription regulation (Browning *et al.*, 2010). Similarly, it is possible that proteins categorised as transcription factors have an additional role in facilitating chromosome compaction and folding. Transcription factors work by binding to specific DNA sequences, usually in promoter regions of genes. Many transcription factors exist as multimers, with the opportunity to bind multiple DNA sites, often leading to looping of DNA. Whether these sites are close or distant on the chromosome, this could contribute to maintaining chromosome structure. Of the seven "global regulators", that participate in the regulation of 51% of genes in *E. coli* between them, three are also nucleoid associated proteins known to modify and maintain chromosome structure.

The link between chromosome structure and gene expression was recently investigated by Qian *et al.*, studying the transcription factor GalR. GalR regulates 5 promoters found at different positions around the *E. coli* chromosome, with only two adjacent to each other. GalR was fused to a rapidly maturing YFP to study its distribution *in vivo*. The majority of cells analysed in stationary phase contained between 1 and 3 fluorescent foci, suggesting the distribution of GalR in the cell is not uniform. These foci were not seen in strains carrying a GalR mutant unable to multimerise, indicating that the foci may be formed by the colocalisation of GalR DNA sites when dimers at distant DNA sites tetramerise, supporting the idea that transcription factors fold the chromosome. Chromosome conformation capture (3C) was then used to study the physical position of GalR-regulated promoters in the cell compared to other locations. This involves cross-linking all the DNA and protein in the cell together before the DNA is digested and re-ligated. As the DNA is still held in the position it

is found in the nucleoid, this will result in ends of DNA ligating with other fragments of DNA it is close to in space and interacting with, and not necessarily adjacent to on the chromosome. The cross-links are then reversed and PCR carried out using primers complementary to known locations, both locations of interest and control locations. If a product is seen, then the DNA at the locations of the two primers must have been close at the point of cross-linking. The frequency of these two sites colocalising can be quantified by qPCR. Qian *et al*. showed that the *galP*, *galR* and *mgl/gals* loci frequently interacted with each other, as well as 6 of the control locations, not known to be GalR targets. When GalR was deleted, interactions between all members of the GalR regulon were removed as well as most between GalR-regulated genes and the 6 new locations. The absence of GalR also prevented these 6 new sites from interacting with each other. It is suggested by the authors that there are unidentified GalR sites at each of these locations, and GalR DNA binding and multimerisation brings all of these sites together. Finally, the effect of GalR binding on mini DNA circles was studied using atomic force microscopy (AFM). The DNA circles contained 5 GalR DNA sites and in the presence of GalR were organised into loops of DNA. This was not seen with the mutant GalR, unable to multimerise, again suggesting that GalR brings together distant DNA sites by tetramerisation. All of these results indicate that GalR could be involved in chromosome structure as well as its known function as a transcription factor (Qian *et al.*, 2012). A FROS experiment could have been used to further confirm the colocalisation of GalR-regulated promoters by tagging two promoters with different FROS reporters.

This work aims to investigate promoters controlled by other transcription factors in *E. coli*. If promoters are colocalising upon induction the position of the promoter within the cell would change upon induction. Using fluorescent reporter/operator system (FROS), the positions of

two promoters regulated by the transcription factor MntR were investigated with and without induction by Mn^{2+}, and the positions of 3 promoters regulated by AraC were investigated with and without induction by arabinose.

4.1.2 The MntR regulon

Manganese is an important trace nutrient for all cells. Although manganese and other metal ions are essential for growth, they are also toxic at higher concentrations, so the levels within the cell need to be carefully controlled. Bacterial cells regulate the level of metal ions by controlling their entry into the cell across the cell membrane, by use of a specific transporter (Jakubovics and Jenkinson, 2001).

The presence of a manganese transporter was first detected in 1970, when it was identified as an active transport system that preferentially transported manganese over magnesium or calcium (Bhattacharyya, 1970). The gene encoding this transporter was confirmed as *yfeP*, which was renamed *mntH*. It was identified as being able to transport manganese, as well as iron, in a proton dependent manner, and having high similarity to the NRAMP (natural resistance-associated macrophage protein) proteins found in mammalian cells (Makui *et al.*, 2000). NRAMP proteins are found in mammalian cells where they are known to be a divalent transition metal transporter which, when present in a macrophage, have the ability to control growth of several bacterial pathogens. Genomic studies have revealed that many bacteria also possess proteins from the NRAMP family, with around 35% sequence homology to the mammalian proteins. It is thought that these bacterial NRAMPs may have a role in pathogenesis (Kehres *et al.*, 2000). *mntH* may have a role in pathogenesis as during infection, the host and pathogen compete for metal ions (Makui *et al.*, 2000). Cells with *mntH* deleted

show no growth defects under normal growth conditions. However, if the cells are exposed to peroxide stress the *mntH* gene product is essential for survival. H_2O_2 in the cell can react with free Fe^{2+} to make hydroxyl radicals which causes damage to biomolecules. The import of manganese does not directly remove the peroxide, but protects the cell by activating metalloenzymes, including superoxide dismutases, which protect the cell against reactive oxygen species (Anjem *et al.*, 2009). *E. coli* has two superoxide dismutases, one iron binding and one manganese binding. The iron bound enzyme is expressed constitutively and protects from oxidative damage by metabolic products. The manganese bound superoxide dismutase is only expressed in the presence of external oxidative stress when iron in the cell is in danger of becoming oxidised and the iron superoxide dismutase is unable to function (Geslin *et al.*, 2001).

In 2001 a transcription factor named MntR was identified that binds to the *mntH* promoter. The *mntH* promoter is repressed in the presence of Mn^{2+} and Fe^{2+}, but in a *mntR* mutant repression was reduced. A Mn^{2+}-dependent MntR DNA site was identified in the *mntH* promoter by DNase I footprinting. Also in the *mntH* promoter is a DNA site for Fur, the ferric uptake regulator, which binds to its DNA sites in the presence of Fe^{2+}. MntR was identified as being a member of the DtxR family (Patzer and Hantke, 2001). The crystal structure of the *E. coli* MntR protein was determined, and the characteristic helix-turn-helix motif was present in the N-terminal domain for DNA binding, and a dimerisation domain in the central region of the protein. There was a non-homologous region in the N-terminal domain, but all of the residues involved in DNA binding were highly conserved (Tanaka *et al.*, 2009).

Recently two studies have defined the MntR regulon and have revealed a small number of genes regulated by MntR. Chromatin immunoprecipitation was used to identify MntR binding targets across the *E. coli* chromosome, revealing four DNA sites. As expected, one of these was adjacent to *mntH* but DNA sites were also found adjacent to *mntR* itself, *yebN* and *dps* (Yamamoto *et al*., 2011). *yebN* has now been renamed *mntP* and encodes a putative efflux pump (Waters *et al.*, 2011). *dps* is a stationary phase nucleoid associated protein that protects the cell from hydrogen peroxide damage (Almiron *et al.*, 1992) and its expression is directly repressed by MntR bound to Mn^{2+}. *mntH*, *yebN* and *dps* promoters are all also regulated by either Fur or OxyR, suggesting there is a link between levels of Mn^{2+} and Fe^{2+} in the cell and oxidative stress (Yamamoto *et al.*, 2011). A second study later the same year compared whole-genome expression in wild type and Δ*mntR* cells to confirm the MntR regulon (Waters *et al.*, 2011). This transcriptomics method also identified the *mntH* and *yebN* (*mntP*) promoters as DNA sites, with MntR upregulating *mntP* in the presence of manganese. The binding at the *dps* promoter was not detected in this study, probably because the experiment was carried out in exponentially growing cells. A 42 amino acid protein previously known as *rybA* was also identified and renamed *mntS*. This was previously thought to encode a sRNA (small RNA) but has been found to also encode a small protein required for full repression on *mntH*. It is now thought that MntS may function as a manganese chaperone and transport it to the cellular locations where it is required (Waters *et al.*, 2011). *mntS* is found immediately upstream of *mntR* in a divergent orientation so the DNA site identified by chromatin immunoprecipitation in the *mntR* promoter may actually affect the *mntS* promoter.

The members of the MntR regulon are found at various locations across the chromosome (see figure 4.1). *mntR* and *mntS* are adjacent to each other with *dps* 5 kb away, and all are found in

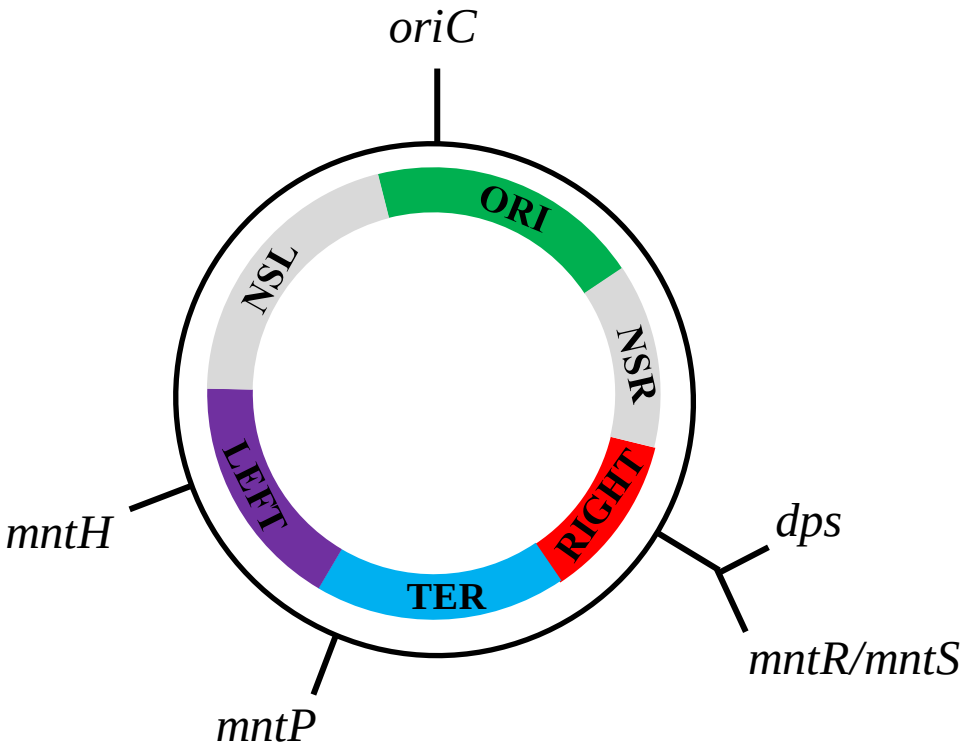


Figure 4.1: Schematic map of the *E. coli* chromosome showing locations of the MntR controlled genes

The *E. coli* chromosome is represented by a circle with *oriC* marked and macrodomains as described by Valens *et al*. shown by coloured segments**.** MntR controls 4 promoters across the *E. coli* chromosome. *mntR* itself and *mntS* are located in the right macrodomain positioned at 852 kb. *dps* is also found in the right macrodomain at position 847 kb. *mntP* (formally *yebN*) is in the Ter macrodomain positioned at 1903 kb and *mntH* is in the left macrodomain positioned at 2510 kb.

the right macrodomain. *mntP* is further away in the Ter macrodomain and *mntH* in the Left macrodomain. The aim of this work is to investigate whether MntR binding has an effect on the cellular position of the promoters it controls, and has a similar effect on the *E. coli* chromosome as GalR.

4.1.3 The AraC regulon

Arabinose is a 5-carbon sugar that can be used by *E. coli* as an energy source. The *E. coli* genome includes several genes known to be involved in the uptake and metabolism of arabinose, and a transcription factor, AraC, which controls the expression of these genes. AraC co-regulates with CRP to control the promoters so they are induced in the presence of arabinose, but only in the absence of glucose so the cell will use the preferential carbon source when it is available. AraC binds as a dimeric protein to pairs of half sites. These can either be in direct repeat or inverted repeat organisation (Carra and Schleif, 1993). AraC is a member of the large AraC/XylS family of transcriptional regulators, now numbering more than 100 transcription factors controlling genes with diverse functions. Almost all are positive regulators with a very small number, including AraC itself, acting as a repressor or functioning as both an activator and repressor (Gallegos *et al.*, 1997). Family members are characterised by a 100 residue stretch of amino acids that folds to form a domain with two DNA-binding helix-turn-helix domains. Most members of the AraC/XylS family are insoluble when over-expressed, making structural experiments difficult.

The best studied genes controlled by AraC are the *araBAD* operon, encoding ribulokinase, L-arabinose isomerise and L-ribulose 5-phosphate 4-epimerase enzymes involved in the degradation of arabinose to make D-xylulose-5-phosphate, which can then be used by the cell

(Englesberg, 1961). The *araBAD* promoter is induced in the presence of arabinose, as well as being actively repressed in its absence, and upon induction with arabinose, the activity of the *araBAD* promoter increases 300 fold above basal levels (Schleif *et al.*, 1973). Upstream of the *araBAD* promoter are two half sites for AraC called I_1 and I_2 which are each bound by a monomer of AraC in the presence of arabinose (see figure 4.2). The I_2 site partially overlaps the -35 element meaning one AraC subunit can contact RNAP to activate transcription. In the absence of arabinose, AraC remains bound to the I_1 DNA site but also binds another half site 200 bp upstream called O_2, causing looping of the DNA and repression of transcription. A previous study suggested that the presence of arabinose limits the extensibility of AraC, preventing it from binding to half sites separated by more than 21 bp. The result of this is that, in the presence of arabinose, AraC can no longer form the loop structure in the DNA to repress the *araBAD* promoter (Carra and Schleif, 1993). Directly upstream of the *araBAD* operon is the *araC* gene, arranged divergently, so the two promoter regions overlap. AraC self regulates by repressing expression of its gene in the absence of arabinose. This is partly achieved by the DNA loop structure in the *araBAD* promoter, but also by a second pair of half sites called O_1 situated between the I_1 and O_2 sites (see figure 4.2) (Schleif, 2000). The O_1 DNA site partially overlaps the RNAP binding site for the *araC* promoter, preventing the initiation of transcription. Activity of the *araC* promoter increases 10 fold upon induction for around 15 minutes, before repression is restored (Ogden *et al.*, 1980). When AraC is not bound at the O_1 DNA site, CRP is able to bind its site to activate transcription of *araBAD*.

AraC also controls genes encoding proteins involved in the transport of arabinose into *E. coli*. There are two independent transport systems for arabinose in *E. coli*. The lower affinity system, encoded by *araE,* has a K_m for arabinose transport of 5×10^{-5} M. The higher affinity

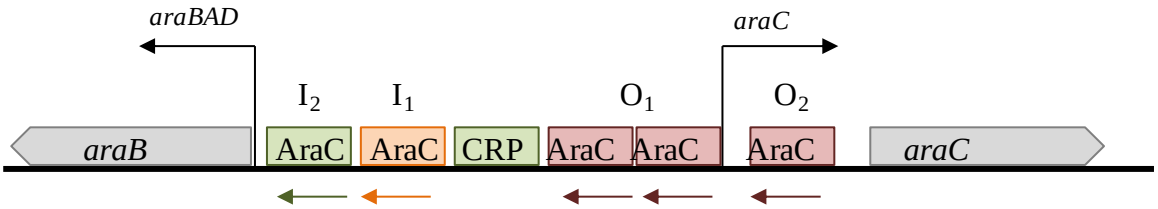


Figure 4.2: Regulation of the *araBAD* and *araC* promoters by AraC

In the presence of arabinose, AraC binds to half sites I_2 and I_1 to activate the *araBAD* promoter along with CRP. In the absence of arabinose AraC binds to the I_1 half site and the O_2 half site to repress both the *araBAD* and *araC* promoters. Large arrows represent the transcription start sites of *araC* and *araBAD*. AraC half sites are shown as coloured rectangles. Green rectangles represent activating sites, red, repressing sites, and orange, sites that are both activating and repressing. The direction of each half site is indicated by an arrow. Not to scale.

system is encoded by the operon *araFGH* and has a K_m of 3 X 10^{-6} M (Kolodrubetz and Schleif, 1981). The *araE* promoter has a similar structure to that of *araBAD* with a pair of AraC DNA sites overlapping the -35 sequence and a CRP site upstream. However, in the *araFGH* promoter, it is the CRP site that overlaps the -35 sequence and two pairs of AraC sites are found upstream (Hendrickson *et al.*, 1990). All of the AraC DNA sites at *araE* and *araFGH* are activatory. The AraC DNA sites in the *araE* promoter are in the same orientation as in the *araBAD* promoter but in the *araFGH* promoter they are in the opposite orientation. It is thought that the different orientation of AraC DNA sites seen at *araFGH* is due to the fact that the CRP DNA site is in a different location to that seen at *araBAD*. This in turn appears to make the *araFGH* promoter more sensitive to the presence of glucose (Hendrickson *et al.*, 1990). The level of induction at *araE* and *araFGH* was found to be approximately 3 fold lower than that of *araBAD*. All 3 promoters reach their maximum level of expression 10 minutes after being induced and the level then starts to fall (Johnson and Schleif, 1995).

The promoter of a gene called *araJ* has also been reported to be activated by AraC and DNA sites of AraC have been identified (Hendrickson *et al.*, 1990). However, the function of *araJ* is unknown and deletion of the gene shows it is not involved in transport of arabinose or arabinose regulation in the cell (Reeder and Schleif, 1991). ChIP-chip data has shown binding of AraC at all the promoters described as well as three novel regions although there is no information about its function when bound (J.T. Wade, personal communication). As with MntR, the promoters bound by AraC are spread throughout the whole chromosome, with *araBAD* and *araJ* found in the non-structured right macrodomain 350 kb apart, *araFGH* in the Ter macrodomain and *araE* in non-structured left (see figure 4.3). Despite the distance

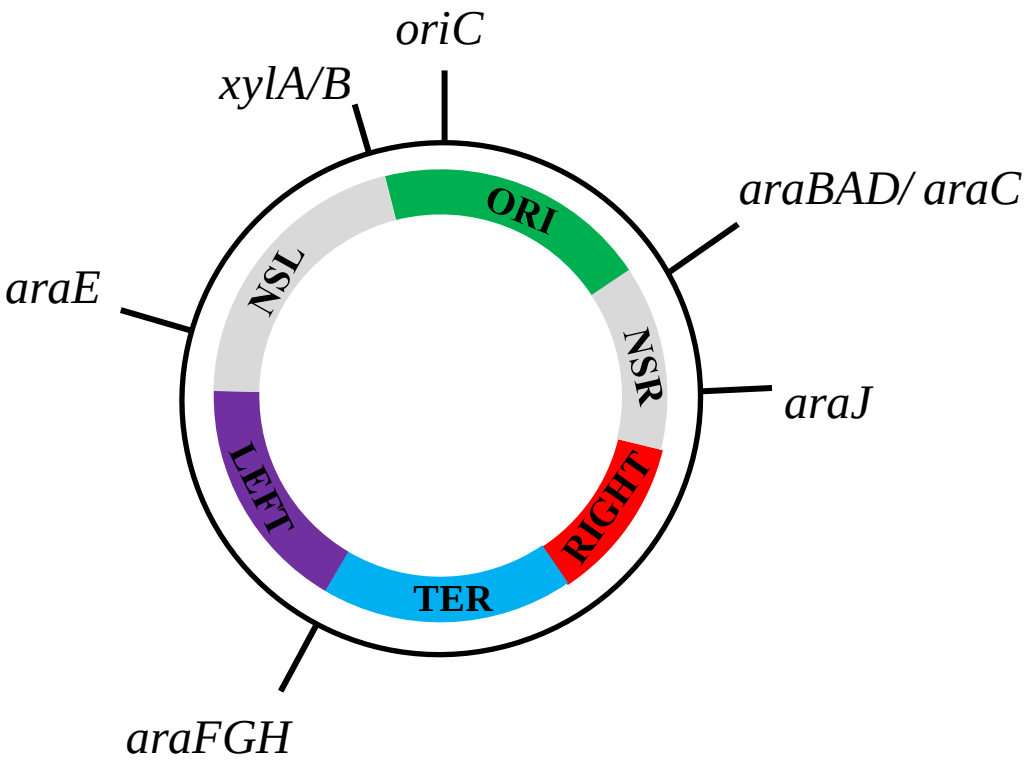


Figure 4.3: Schematic of the *E. coli* chromosome showing locations of AraC controlled genes

The *E. coli* chromosome is represented by a circle with *oriC* marked and macrodomains as described by Valens *et al*. shown by coloured segments. AraC controls 5 promoters across the *E. coli* chromosome, including its own promoter. *araBAD* and *araC* are located in the non-structured right macrodomain positioned at approximately 70 kb. *araJ* is also in the non-structured right macrodomain at position 411kb. *araFGH* is found in the Ter macrodomain at position 1983 kb and *araE* is in the non-structured left macrodomain at position 2980 kb.

between these genes, this study sought to determine if they colocalise in the folded chromosome, due to the binding of their transcription factor, AraC.

4.2 Studying MntR-regulated promoters

As shown by Qian *et al.*, binding by transcription factors can affect chromosome structure. Hence the positions of two promoters regulated by MntR were investigated with and without induction by Mn^{2+} using fluorescent reporter/operator system (FROS). If MntR binding changes chromosome structure, the cellular position of the genes it regulates may change.

Arrays of multiple DNA sites for LacI and MalI were inserted adjacent to two MntR controlled promoters. 20 MalI DNA sites are inserted adjacent to the *mntH* promoter in strain SXB3, which was then transformed with a plasmid encoding a MalI::mCherry fusion, pLER108. This strain containing the plasmid is referred to as SXB3-pLER108. 22 LacI DNA sites are inserted adjacent to the *dps* promoter in strain SXB4, which also has a chromosomal LacI::GFP fusion. Cultures were grown as previously described, with 0.1 μM $MnCl_2$ added for the final 10 minutes of growth (Waters *et al.*, 2011). Slides were prepared as previously described. The experiment was repeated 3 times with at least 100 cells analysed from each repeat. Only single cells were analysed, as cells at the point of division are thought to have a different chromosome structure which will be studied separately. Single cells were distinguished first by brightfield microscopy, and then DAPI was used to view the nucleoid. Although the nucleoid is often lobed in single cells, there is still a connection between the two halves.

4.2.1 Comparison of the number of foci per cell with and without induction by Mn^{2+}

The number of foci visible in each cell was counted to give the percentage of cells containing different numbers of foci. Strain SXB3-pLER108, with a tag adjacent to *mntH*, had one focus in the majority of cells, around 60% of cells, both with and without induction by Mn^{2+} (see figure 4.4a). Around 35% of cells had 2 foci with a small number of cells having 3, 4 or no foci. There was no significant difference ($P > 0.05$) between the number of foci per cell in cultures that had been grown with or without Mn^{2+}. Strain SXB4, with a tag adjacent to *dps*, had one focus in around 65% of cells, 2 foci in 30% of cells and a small number of cells with 3, 4 or no foci (see figure 4.4b). Again, there was no significant difference ($P > 0.05$) between the number of foci per cell when the cultures had been grown with or without Mn^{2+}.

When chromosome replication has passed the point of the chromosomal tag, two foci will be visible in the cell. *E. coli* can initiate multiple rounds of replication, so it is sometimes possible to see up to 4 foci in one cell. If there are 2 copies of the chromosome at the position tagged, but the sister chromatids are still close to each other, the microscope will not be able to distinguish two separate foci, due to the limits of diffraction, so only one focus will be observed. Due to the mechanism of chromosome replication in *E. coli*, cells with 3 foci were not expected. However, if there are 4 copies of the chromosome at the position of the insert and 2 are still close together, having recently been replicated, only 3 foci would be observed. There are slightly more cells with a single focus when the chromosome is tagged adjacent to *dps* because this gene is further from *oriC* than *mntH*, therefore will be replicated later in the cell cycle. The average number of foci per cell was calculated for strains tagged adjacent to *araBAD*, *mntH* and *dps* (strains LR06, SXB3-pLER108 and SXB4), grown in the presence and absence of Mn^{2+}, and this number was plotted against the distance, in kb, from *oriC* to the tag (see figure 4.5). The resulting graph shows that the average number of foci per cell

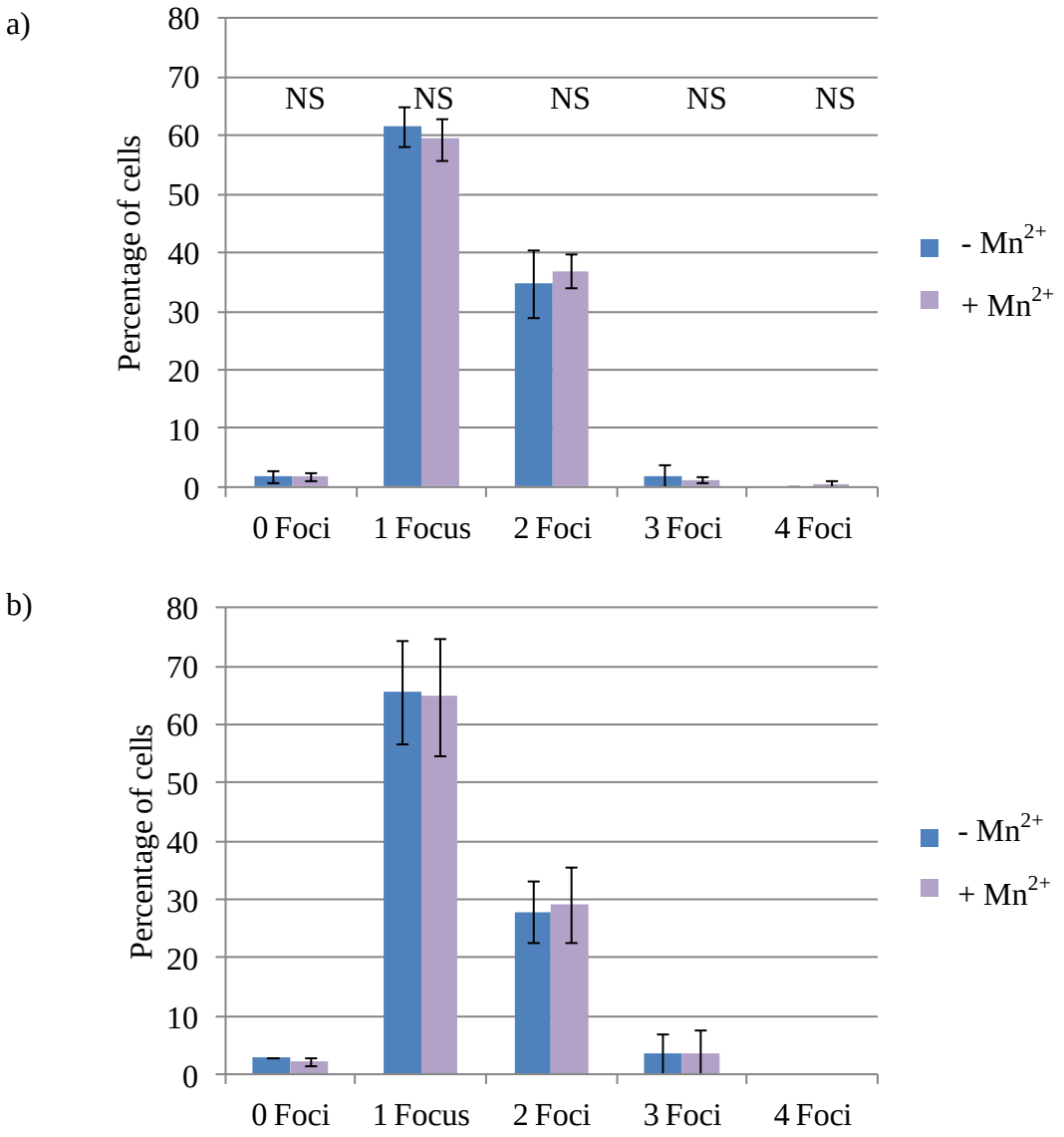


Figure 4.4: Comparison of the number of foci representing MntR controlled promoters seen with and without induction by Mn^{2+}

Bar chart showing the percentage of cells containing different number of foci.

a) Strain SXB3-pLER108, with an insert of 20 MalI DNA sites adjacent to the *mntH* promoter, was transformed with pLER108, carrying MalI::mCherry, grown with and without induction by Mn^{2+}.

b) Strain SXB4, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by Mn^{2+}.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 µM $MnCl_2$ added for the final 10 minutes of growth. The number of foci per cell was recorded. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 350 cells.

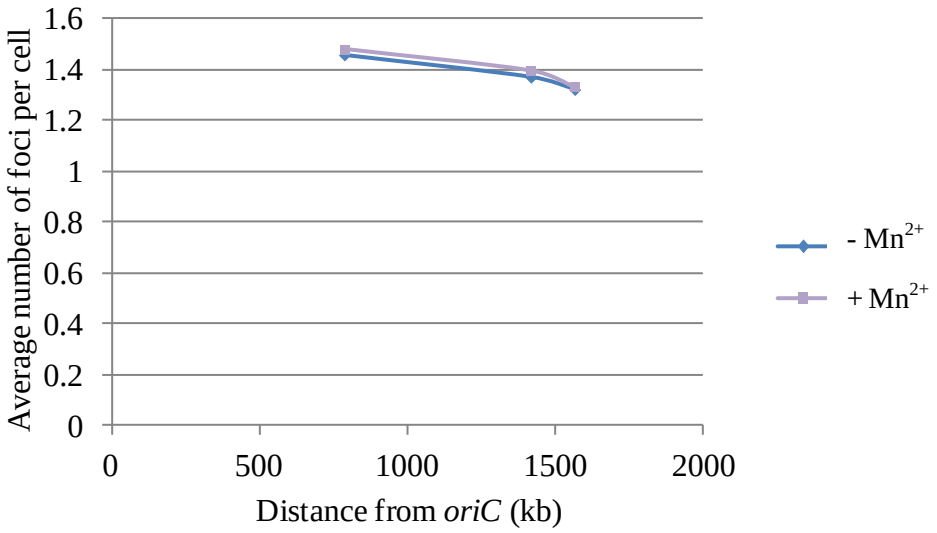


Figure 4.5: Average number of foci per cell relative to distance from tag to *oriC*, in cells with and without Mn^{2+}

The average number of foci per cell was calculated for strains tagged adjacent to *araBAD* (LR06), *mntH* (SXB3-pLER108) and *dps* (SXB4). This number was plotted against the distance from each locus to *oriC*, 785 kb, 1415 kb and 1563 kb respectively.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1µM $MnCl_2$ added for the final 10 minutes of growth. The number of foci per cell was recorded.

decreases with distance from *oriC*, and is not affected by the presence of Mn^{2+}. Any variation in the number of foci per cell between strains tagged at different locations is due to the position of the tag relative to *oriC*. As there is no difference between the number of foci ininduced and uninduced cells, the addition of Mn^{2+} must not affect either the rate of DNA replication or the separation of sister chromatids at MntR controlled promoters.

4.2.2 Comparison of the position of a single focus in cells tagged adjacent to MntR controlled promoters, with and without induction by Mn^{2+}

For each cell with a single focus, the length of the cell and the distance from the centre of the focus to the closest cell pole was measured. As the length of cells varied, the distance from the focus to cell pole was then divided by cell length to give a value representing the position of the focus relative to cell length. As the focus was measured to the closest pole, the furthest distance between a focus and the cell pole is half of cell length so values must be 0.5 or less. The percentage of cells containing a focus centred at each 0.02 of cell length between 0 and 0.5 were calculated. Strain SXB3-pLER108, tagged adjacent to *mntH*, did not show any significant difference ($P > 0.05$) between cultures that had been grown with and without Mn^{2+}, with the bars showing a very similar overall shape for both conditions (see figure 4.6a and b). There was also no significant difference between cultures of SXB4, tagged adjacent to *dps*, that had been grown with and without Mn^{2+} (see figure 4.6 a and c). This suggests that induction by Mn^{2+} does not cause MntR controlled promoters to move to a specific part of the cell. If there is a movement it is too small to be detected by this method.

Figure 4.6: Comparison of the position of foci representing MntR controlled promoters with and without induction by Mn^{2+} in cells with a single focus

Bar charts showing the average position of foci or percentage of foci centred at each 0.02 of cell length in the absence and presence of Mn^{2+}.

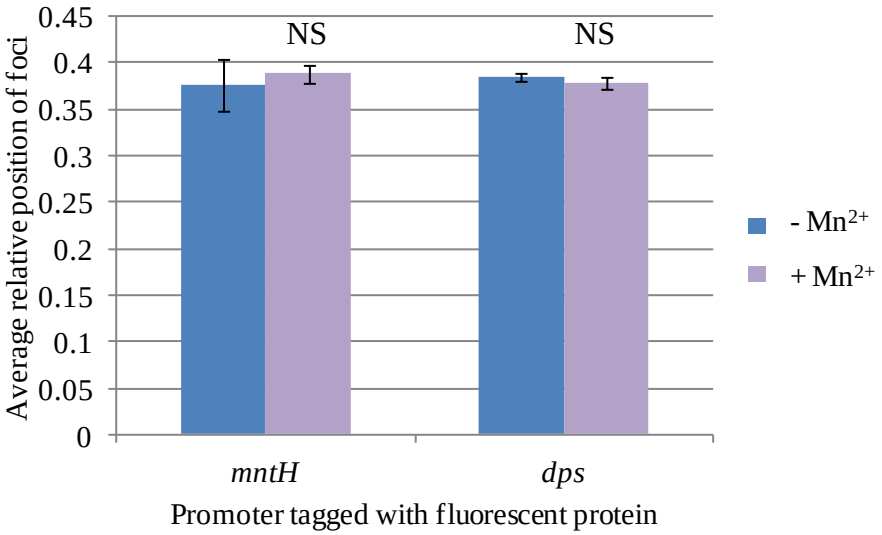
a) Average relative position of foci in cells with a single focus tagged adjacent to *mntH* or *dps*.

b) Strain SXB3-pLER108, with an insert of 20 MalI DNA sites adjacent to the *mntH* promoter, was transformed with pLER108, carrying MalI::mCherry, grown with and without induction by Mn^{2+}.

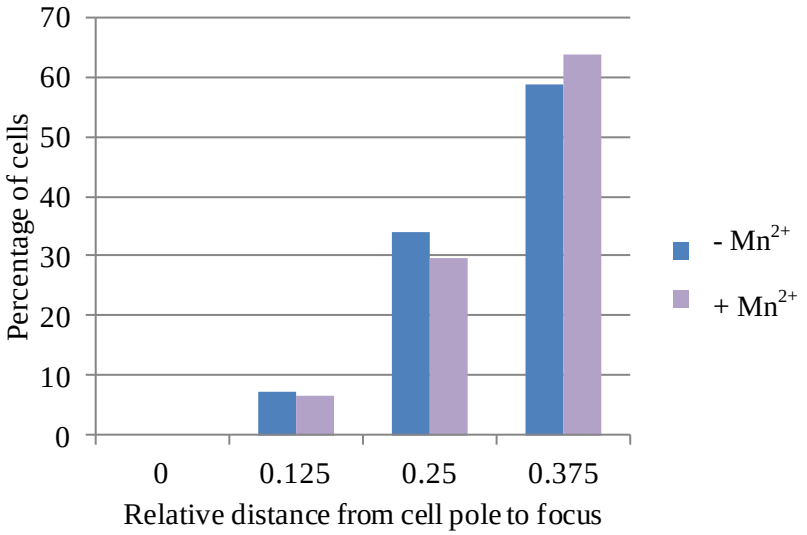
c) Strain SXB4, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by Mn^{2+}.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 μM added $MnCl_2$ for the final 10 minutes of growth. The position of foci in cells with a single focus were measured relative to cell length. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 350.

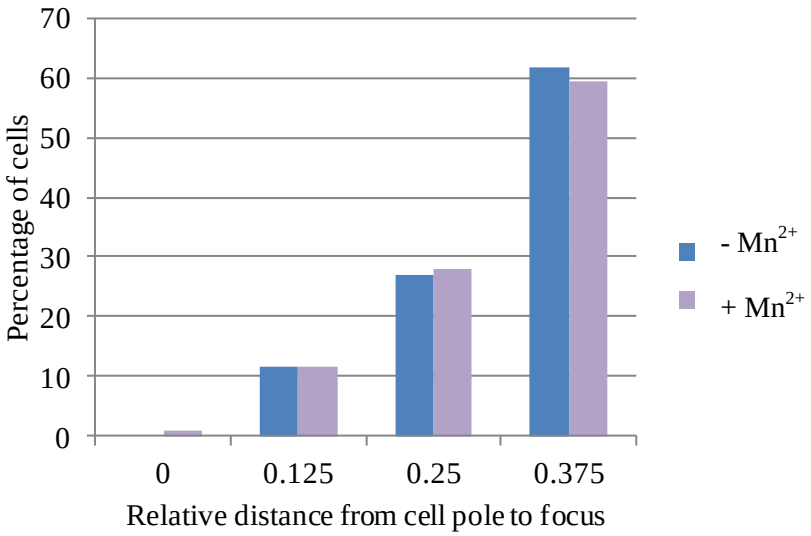
a)


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Average relative position of foci
NS
NS
- Mn2+
+ Mn2+
mntH
dps
Promoter tagged with fluorescent protein

b)


70
60
50
40
30
20
10
0
Percentage of cells
- Mn2+
+ Mn2+
0
0.125
0.25
0.375
Relative distance from cell pole to focus

c)


70
60
50
40
30
20
10
0
Percentage of cells
- Mn2+
+ Mn2+
0
0.125
0.25
0.375
Relative distance from cell pole to focus

4.2.3 Comparison of the positions of 2 foci from cells tagged adjacent to MntR controlled promoters, with and without induction by Mn^{2+}

For cells with 2 foci, the positions of foci were also measured relative to cell length. The focus closest to a cell pole was called "1st of 2 foci" and the distance from the centre of this focus to the cell pole was measured. The position of the "2nd of 2 foci" was then measured to the same cell pole. The position of the first focus, relative to cell length, was plotted on the x axis of a scatter plot against the position of the second focus on the y axis. If the points on the plot localise to a particular area it will indicate that the pair of foci are often found in a certain arrangement. For example, if the foci are mostly positioned at opposite ends of the cell close to the poles, the points will cluster towards the top left corner. If the foci are mostly located in the same half of the cell close together and close to a pole the points will be in the bottom left corner.

SXB3-pLER108 cells, tagged at *mntH*, do not show a significant difference ($P > 0.05$) between supplemented and unsupplemented cultures, with the points covering a similar area of the plot (see figure 4.7a, b and c). SXB4 cells, tagged at *dps*, also do not show a significant difference in the positions of 2 foci with and without induction ($P > 0.05$) (see figure 4.7a, b and d). This confirms that the addition of Mn^{2+} does not cause MntR-regulated promoters to be repositioned within the cell. The pattern on the scatter plot is different depending on whether *mntH* or *dps* is tagged. In the *mntH* tagged strain, the points form a tight bunch in the centre of the graph, indicating the 2 foci are often positioned with one focus near the centre of the cell, around 0.6 – 0.7 relative to cell length, and the other focus closer to a cell pole, around 0.2 – 0.3 relative to cell length (see figure 2.20). In the *dps* tagged strain, the points are spread more evenly across the centre of the graph with no clear localisation. The differences between the two locations may be because they are located in different macrodomains. For

Figure 4.7: Comparison of the position of foci representing MntR controlled promoters seen with and without induction by Mn^{2+} in cells with a two foci

Bar chart showing the average relative position of the 1st of 2 foci and 2nd of 2 foci and scatter plot showing the position of the 1st of 2 foci relative to cell length (x axis) against the position of the 2nd of 2 foci relative to cell length (y axis) in the absence and presence of Mn^{2+}.

a) Average relative positions of the 1st of 2 foci in cells tagged adjacent to *mntH* and *dps*, grown with and without induction by Mn^{2+}.

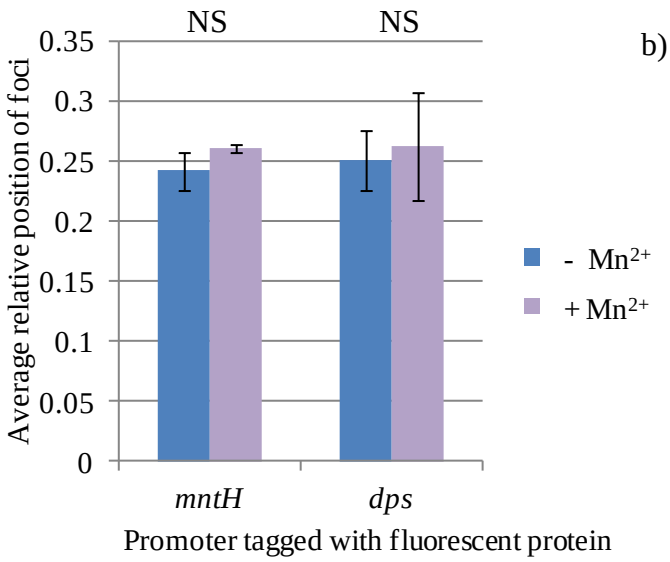
b) Average relative positions of the 2nd of 2 foci in cells tagged adjacent to *mntH* and *dps*, grown with and without induction by Mn^{2+}.

c) Strain SXB3-pLER108, with an insert of 20 MalI DNA sites adjacent to the *mntH* promoter, was transformed with pLER108, carrying MalI::mCherry, grown with and without induction by Mn^{2+}.

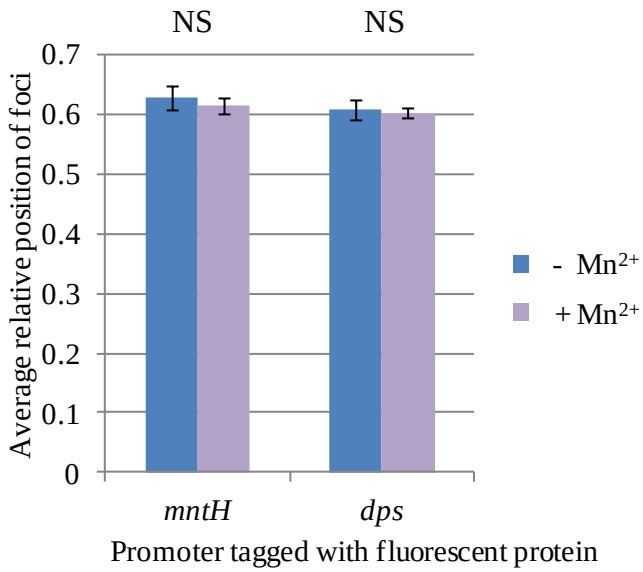
d) Strain SXB4, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by Mn^{2+}.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 μM $MnCl_2$ added for the final 10 minutes of growth. The position of foci in cells with 2 foci were measured relative to cell length. The position of the focus closest to a cell pole ("1st of 2 foci) is plotted on the x axis. The position of the further away focus (2nd of 2 foci) was plotted on the y axis. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = > 350 cells.

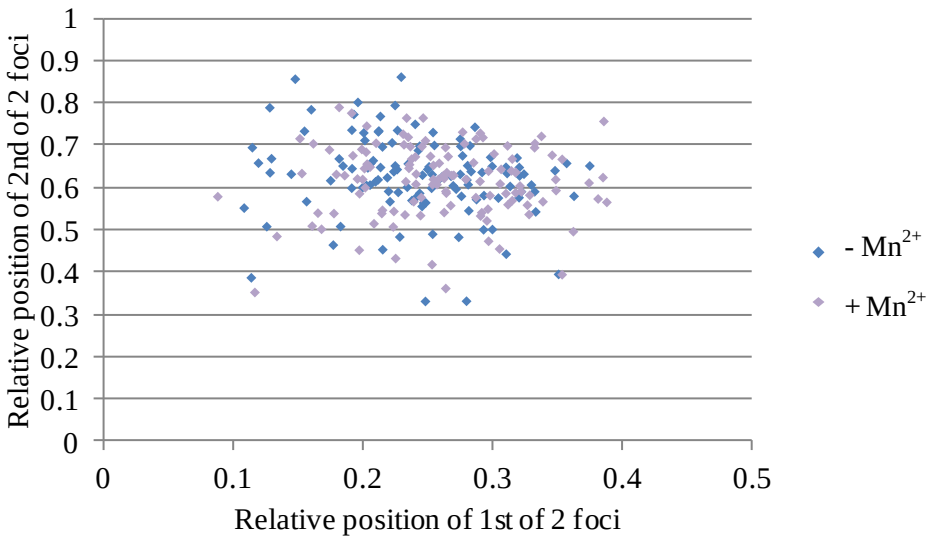
a)


NS
NS
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Average relative position of foci
- Mn2+
+ Mn2+
mntH
dps
Promoter tagged with fluorescent protein

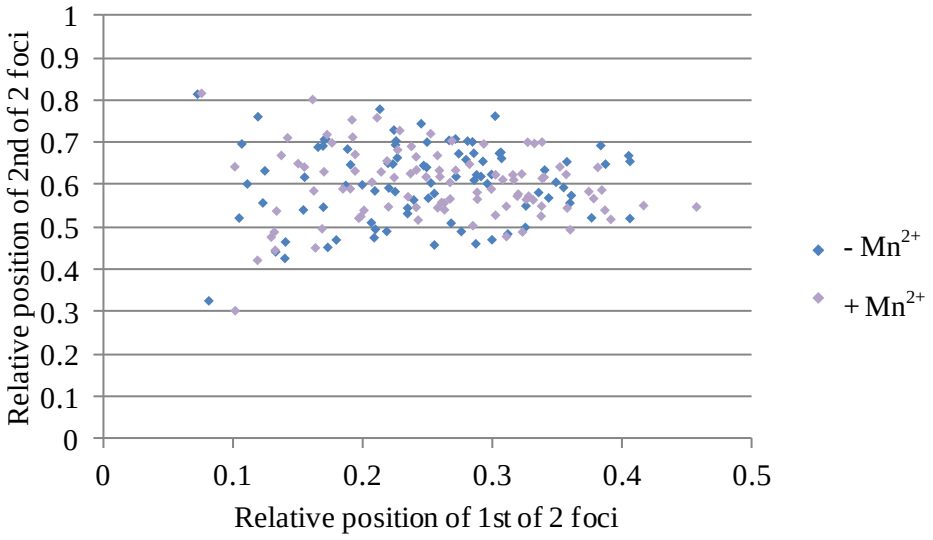
b)


NS
NS
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Average relative position of foci
- Mn2+
+ Mn2+
mntH
dps
Promoter tagged with fluorescent protein

c)


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Relative position of 2nd of 2 foci
0
0.1
0.2
0.3
0.4
0.5
Relative position of 1st of 2 foci
- Mn2+
+ Mn2+

d)


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Relative position of 2nd of 2 foci
0
0.1
0.2
0.3
0.4
0.5
Relative position of 1st of 2 foci
- Mn2+
+ Mn2+

both strains there are very few points in the bottom left quarter of the graph. As sister chromatids of the left and right macrodomains are segregated soon after replicating, this is to be expected, as the two copies of the chromosome are positioned on either side of the cell, ready for when the cell divides so each daughter cell has one copy of the chromosome. The positions of the two foci seem to be asymmetric, with the majority of foci in both strains falling between the positions of 0.2 and 0.3 or 0.6 and 0.7 compared to cell length. If the two copies of the chromosome were segregated equally after replication, the foci would be expected to be equal distance from the centre of the cell, 0.5. This asymmetry can be explained by the positioning of the left and right replichores after replication.

4.3 Defining the threshold concentration of arabinose for induction of transcription at araBAD

AraC acts as an arabinose-dependent switch and has even been referred to as a rheostat. This means that if the concentration of arabinose present is below the optimal concentration, in some cells the target promoters will be fully induced, and, in the remainder, the promoter will be fully repressed. For experiments studying AraC-regulated promoters, it is essential that as many cells as possible are fully induced, as the data collected from microscopy is from individual cells. To find the threshold concentration for full induction, the β-galactosidase activity of cells carrying a plasmid containing *lacZ* under the control of the *araBAD* promoter, pBAD, was measured after induction with various concentrations of arabinose (see figure 4.8). Strain KH001 was used as it carries a deletion of the *lacZ* promoter. The highest levels of β-galactosidase activity were seen with final concentrations of 0.1% and 0.3% arabinose. Increasing the concentration above 0.5% was seen to significantly reduce the β-galactosidase activity ($P < 0.05$). For all further experiments requiring induction by arabinose, 0.3% arabinose was used.

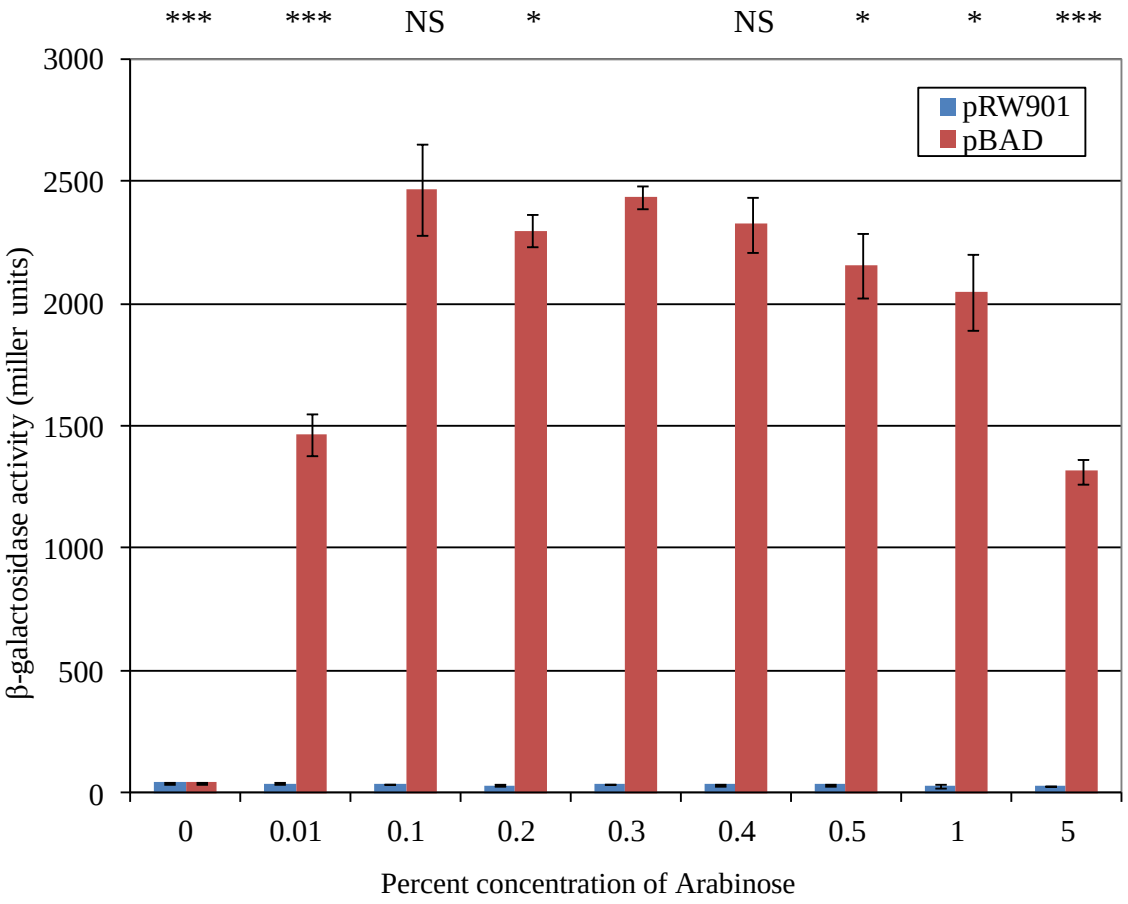


Figure 4.8: Induction of *araBAD* promoter with different concentrations of Arabinose

Expression from the *araBAD* promoter cloned into the *lac* expression vector pRW50 was measured in strain KH001, grown aerobically at 37ºC to mid-exponential phase (OD_{650} 0.3-0.5) in M9 salts medium supplemented with 0.3% fructose and induced with various concentrations of arabinose 2 hours prior to assaying. pRW901 was included as a negative control. Data shown are representative of results from three independent experiments and error bars indicate one standard deviation from the mean. A star above a bar represents data that is significantly different to cultures supplemented with 0.3% arabinose (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = 9.

4.4 Studying AraC controlled promoters

To investigate whether AraC controlled promoters, located at various positions on the chromosome, colocalise due to AraC binding after induction by arabinose, strains were made for use with FROS. 22 LacI DNA sites are inserted adjacent to the *araBAD* promoter in a strain that also had a chromosomal LacI::GFP fusion, LR06. 22 LacI DNA sites are also inserted adjacent to the *araJ* promoter in strain LR39, with a chromosomal LacI::GFP fusion 20 MalI DNA sites are inserted adjacent to the *araFGH* promoter in strain LR38. This was then transformed with plasmid pLER108, encoding a MalI::mCherry fusion. This strain containing the plasmid will be referred to as LR38-pLER108. Cultures were grown as previously described, with 0.3% arabinose or glucose added for 1 minute before slides were prepared. Slides were prepared as previously described. The experiment was repeated 3 times, with at least 100 cells analysed from each repeat. Only single cells were analysed, as cells at the point of division are thought to have a different chromosome structure which will be studied separately. Single cells were distinguished first by brightfield view, and then DAPI was used to view the nucleoid. Although the nucleoid is often lobed in single cells, there is still a connection between the two halves.

4.4.1 Comparison of the number of foci per cell with and without induction by arabinose

The number of foci in each cell was counted to give a percentage of cells containing each number of foci. LR06, tagged at *araBAD*, had around 60% of cells with a single focus and 30% of cells with 2 foci (see figure 4.9a). Cultures grown with arabinose, glucose or with no supplement did not show any significant difference in the number of foci per cell ($P > 0.05$). LR39, tagged at *araJ*, had around 70% of cells with a single focus and 25% with 2 foci, also with no significant difference in the number of foci per cell for cultures grown in the different

Figure 4.9: Comparison of the number of foci representing AraC controlled promoters seen with and without induction by arabinose

Bar chart showing the percentage of cells with different numbers of foci, in unsupplemented cultures and those supplemented with arabinose or glucose.

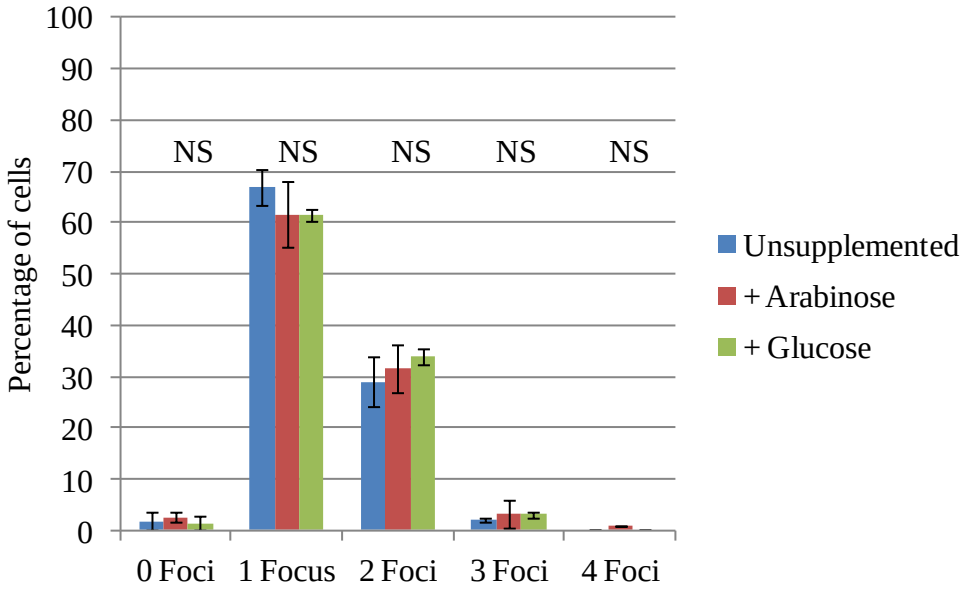
a) Strain LR06, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose.

b) Strain LR39, with an insert of 22 LacI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose.

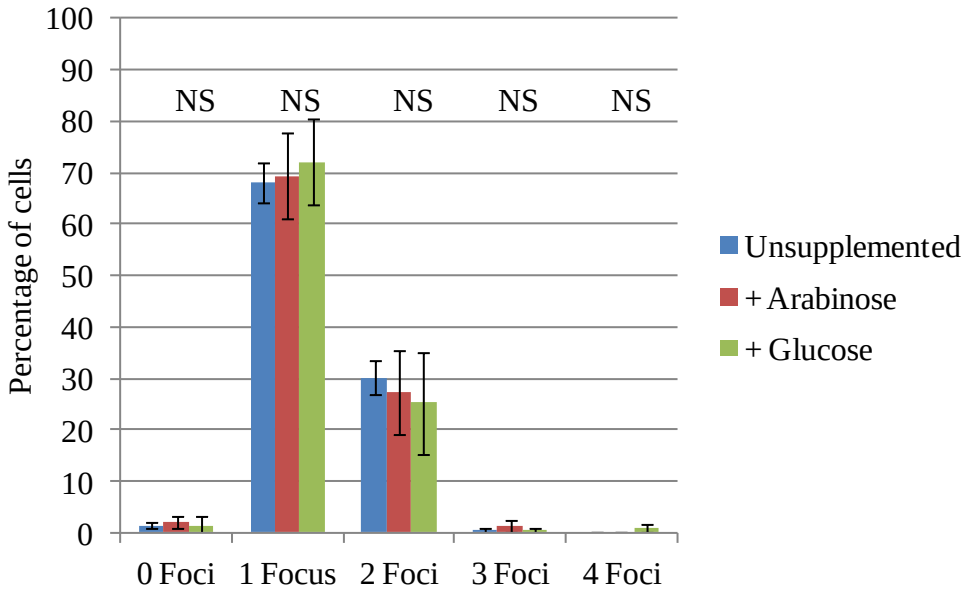
c) Strain LR38-pLER108, with an insert of 20 MalI DNA sites adjacent to the *araFGH* promoter, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. The number of foci per cell was recorded. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 350.

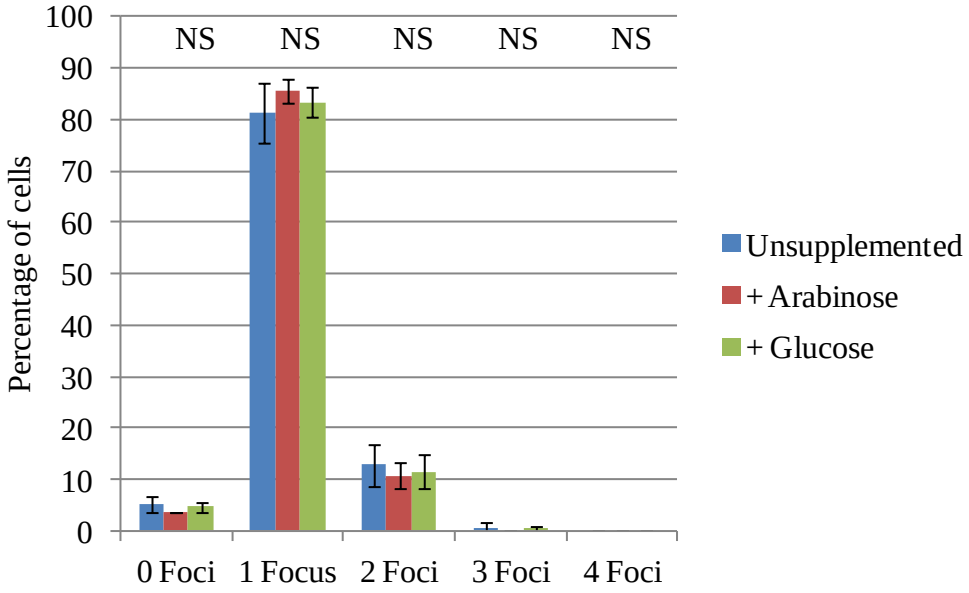
a)


100
90
80
70
60
50
40
30
20
10
0
Percentage of cells
NS
NS
NS
NS
NS
0 Foci
1 Focus
2 Foci
3 Foci
4 Foci
Unsupplemented
+ Arabinose
+ Glucose

b)


100
90
80
70
60
50
40
30
20
10
0
Percentage of cells
NS
NS
NS
NS
NS
0 Foci
1 Focus
2 Foci
3 Foci
4 Foci
Unsupplemented
+ Arabinose
+ Glucose

c)


100
90
80
70
60
50
40
30
20
10
0
Percentage of cells
NS
NS
NS
NS
NS
0 Foci
1 Focus
2 Foci
3 Foci
4 Foci
Unsupplemented
+ Arabinose
+ Glucose

conditions ($P > 0.05$) (see figure 4.9b). LR38-pLER108, tagged at *araFGH*, had around 80% of cells containing a single focus and 10% with 2 foci. Again, there was no significant difference between the cultures grown in the 3 different conditions ($P > 0.05$) (see figure 4.9c). This suggests that induction by arabinose does not cause a change in the rate of DNA replication or the segregation of sister chromatids at positions on the chromosome adjacent to AraC controlled promoters. The percentage of cells containing a single focus rises from 60% to 70% and then 80% in cells tagged adjacent to *araBAD*, *araJ* and *araFGH* respectively. As the percentage of cells with a single focus rises, the percentage of cells with 2 foci falls by approximately the same amount. This is due to the positions of the 3 genes relative to *oriC*, the origin of replication. As *araBAD* is close to the edge of the Ori macrodomain it is replicated early in the cell cycle and as a result a relatively large proportion of cells will have 2 foci. In contrast, *araFGH* is located in the Ter macrodomain and as a result is replicated late in the cell cycle meaning relatively few cells have 2 foci and more having a single focus. The average number of foci per cell was calculated for each strain and strain SXB4, tagged adjacent to *dps*, in each condition. This average was then plotted against distance from the tag to *oriC*, in kb (see figure 4.10). The graph shows that the average number of foci per cell decreases with distance from *oriC*, with little difference between cultures grown with arabinose, glucose or no supplement.

4.4.2 Comparison of the position of a single focus in cells tagged adjacent to AraC controlled promoters, with and without induction by arabinose

For cells containing a single focus, the cell length was measured and then the distance from the centre of the focus to the closest cell pole. Cell length varied so to make the data comparable, the distance from focus to cell pole was divided by cell length to give a position of the focus relative to cell length. The percentage of cells containing a focus centred at each

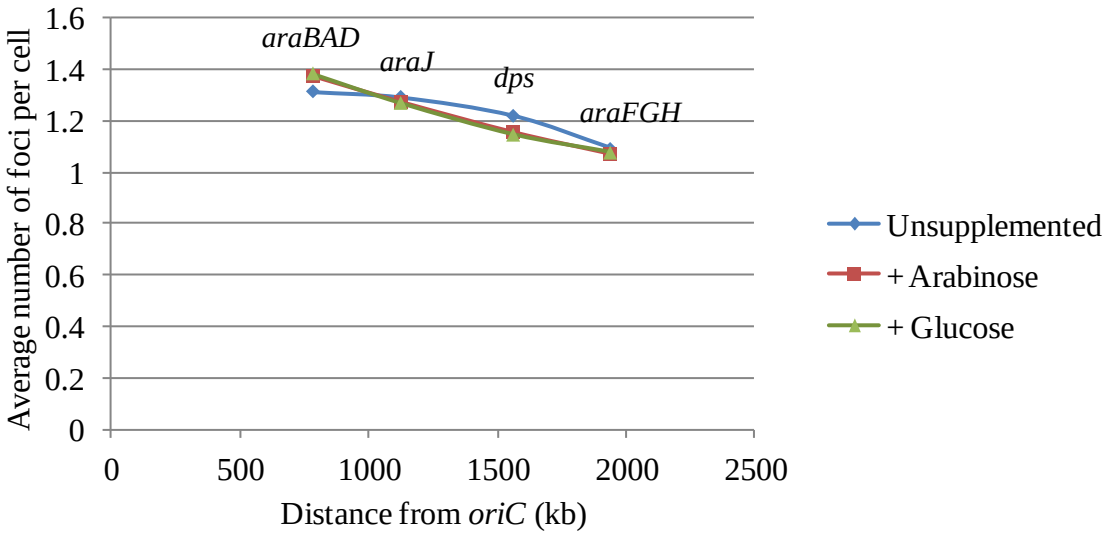


Figure 4.10: Average number of foci per cell relative to distance from tag to *oriC*, in cells unsupplemented or supplemented with arabinose or glucose

The average number of foci per cell was calculated for strains tagged adjacent to *araBAD* (LR06), *araJ* (LR39), *dps* (SXB4) and *araFGH* (LR38-pLER108). This number was plotted against the distance from each gene to *oriC*, 785 kb, 1126 kb, 1563 kb and 1940 kb respectively.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. The number of foci per cell was recorded.

0.02 of cell length between 0 and 0.5 were calculated. Strain LR06, tagged adjacent to *araBAD*, did not show any significant difference ($P > 0.05$) between unsupplemented cultures and those grown with arabinose and glucose (see figure 4.11a and b). Strain LR39, tagged adjacent to *araJ*¸ also did not show any significant difference between cultures grown in the 3 conditions ($P > 0.05$) (see figure 4.11a and c). LR38-pLER108, tagged adjacent to *araFGH*, showed no significant difference between unsupplemented cultures and those grown with glucose ($P > 0.05$) but the cultures grown with arabinose were significantly different ($P < 0.05$) (see figure 4.11a and d). The average position of a focus changing from 0.351 and 0.358, for unsupplemented and glucose supplemented cultures respectively, to 0.375 for arabinose supplemented cultures. This can be seen on the bar chart (see figure 4.11d) as a shift in the bars representing arabinose supplemented cultures towards the right side of the graph. This means that upon induction with arabinose the *araFGH* promoter is more likely to be located towards the centre of the cell (0.5 relative to cell length) than towards the pole of the cell. This suggests that, for some promoters at least, the activity of a promoter can influence its position in the cell.

4.4.3 Comparison of the positions of 2 foci from cells tagged adjacent to AraC controlled promoters, with and without induction by arabinose

The positions of foci in cells containing 2 foci were also measured. The focus closest to a cell pole was called "1st of 2 foci" and the distance was measured from the centre of this focus to the cell pole. The other focus was called "2nd of 2 foci" and the distance was measured from the centre of this focus to the same pole. These distances were represented relative to cell length and plotted on a scatter graph with "1st of 2 foci" on the x axis and "2nd of 2 foci" on the y axis. The data for strain LR06, tagged adjacent to *araBAD*, was spread across a similar area of the graph for cultures grown with arabinose, glucose or without an extra supplement

Figure 4.11: Comparison of the position of foci representing AraC controlled promoters seen with and without induction by arabinose in cells with a single focus

A bar chart showing the average position of foci or percentage of foci centred at each 0.125 of cell length in the absence and presence of arabinose and glucose.

a) Average position of foci in cells with a single focus tagged adjacent to *araBAD*, *araJ*, or *araFGH* , grown with and without induction by arabinose

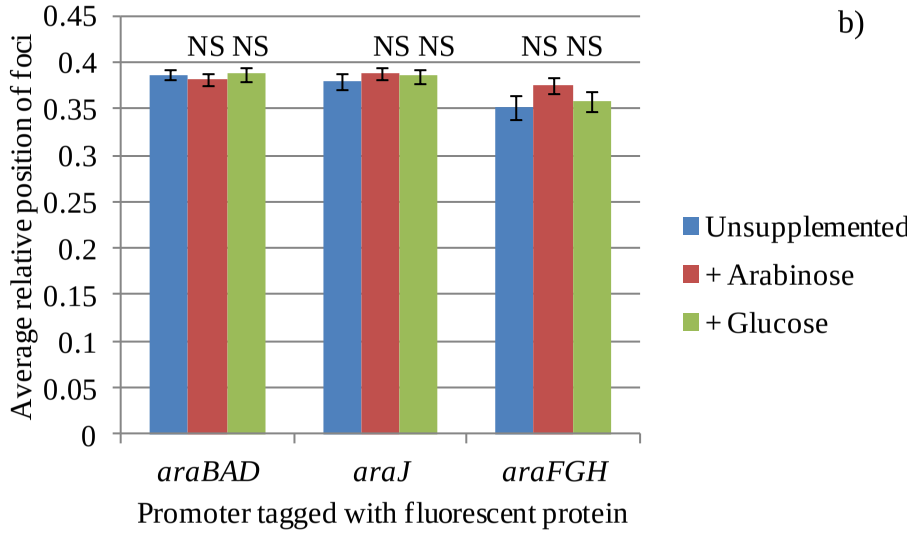
b) Strain LR06, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose.

c) Strain LR39, with an insert of 22 LacI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose

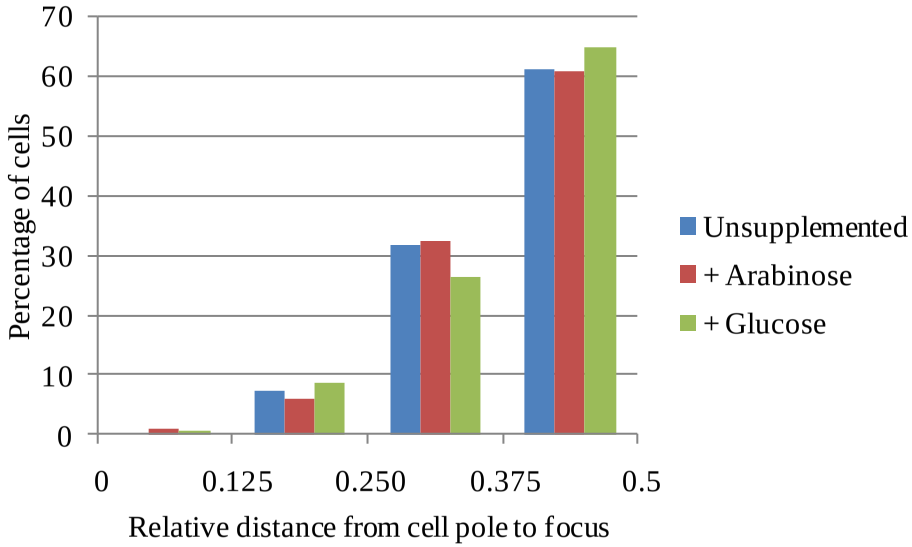
d) Strain LR38-pLER108, with an insert of 20 MalI DNA sites adjacent to the *araFGH* promoter, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. The position of foci in cells with a single focus were measured relative to cell length. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 350 cells.

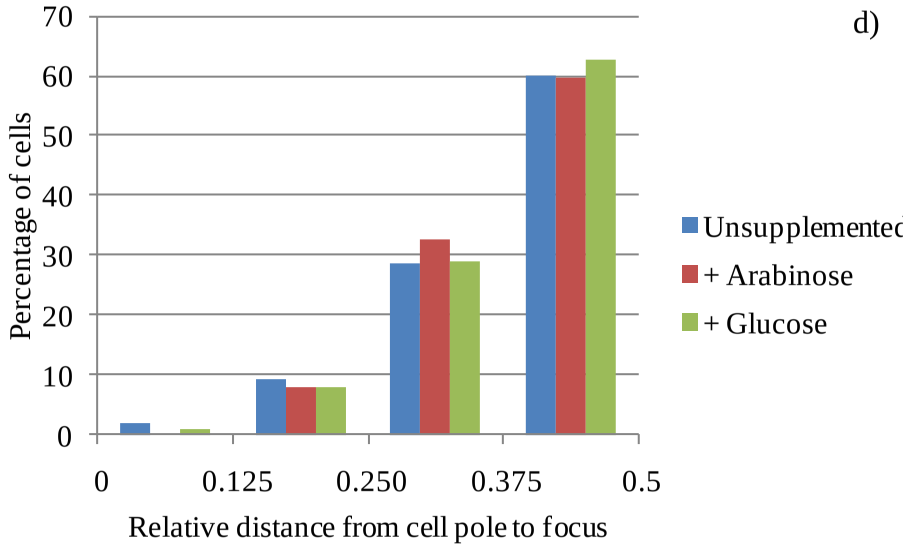
a)


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Average relative position of foci
NS NS
NS NS
NS NS
araBAD
araJ
araFGH
Promoter tagged with fluorescent protein
Unsupplemented
+ Arabinose
+ Glucose

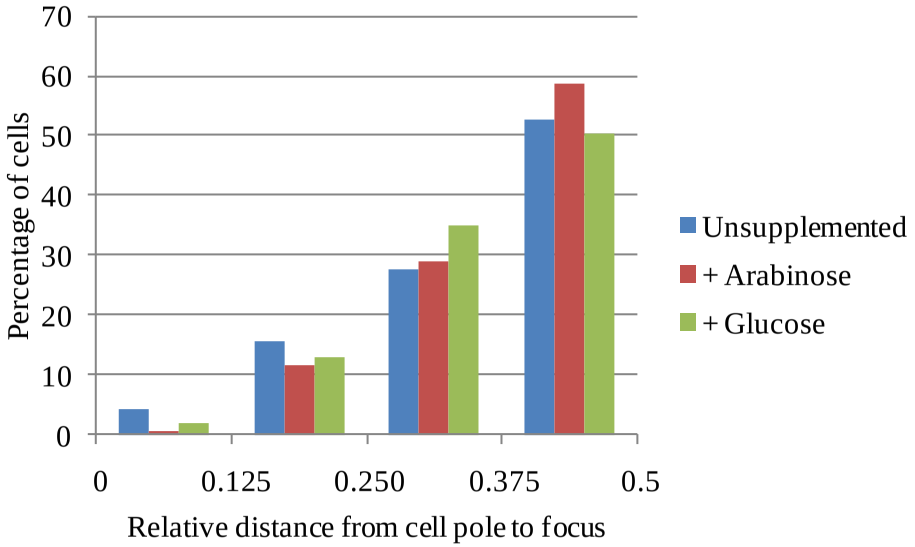
b)


70
60
50
40
30
20
10
0
Percentage of cells
0
0.125
0.250
0.375
0.5
Relative distance from cell pole to focus
Unsupplemented
+ Arabinose
+ Glucose

c)


70
60
50
40
30
20
10
0
Percentage of cells
0
0.125
0.250
0.375
0.5
Relative distance from cell pole to focus
Unsupplemented
+ Arabinose
+ Glucose

d)


70
60
50
40
30
20
10
0
Percentage of cells
0
0.125
0.250
0.375
0.5
Relative distance from cell pole to focus
Unsupplemented
+ Arabinose
+ Glucose

200

($P > 0.05$) (see figure 4.12a and c). Strain LR39, tagged adjacent to *araJ*, was also found to have no significant difference between the positions of foci grown with no extra supplement, arabinose and glucose, with the points from all 3 conditions covering a similar area of the graph ($P > 0.05$) (see figure 4.12a and d). The data from strain LR38-pLER108, tagged at *araFGH*, showed a significant difference between the position of the 1st of 2 foci in cultures grown with arabinose ($P < 0.01$) or glucose ($P < 0.05$) compared to unsupplemented cultures (see figure 4.12a and e). These differences can be seen in figure 4.12e as a small shift towards the right of the chart for the points representing glucose supplemented cultures and a considerable shift towards the right for the points representing arabinose supplemented cultures. The area of the graph containing the majority of the points representing the arabinose and glucose supplemented cultures (0.25-0.4 on the x axis and 0.5-0.7 on the y axis) corresponds to cells with the 2 foci arranged with both foci close to the centre of the cell, with areas around both poles of the cell unoccupied by foci. This is seen by the lack of arabinose or glucose points in the top section of the scatter graph or the left side (see figure 2.20). There was no significant difference between the 3 conditions for the 2nd of 2 foci (see figure 4.12b) indicating that the movement seems to be in the focus designated as the "1st of 2 foci". This may be because it is a movement from an asymmetrical to a more symmetrical arrangement of the 2 foci, and as the cells have already been orientated for measuring purposes so that the focus further from the cell centre is "1st", this will always be the focus that "moves".

The area covered by points in the graphs for strains tagged adjacent to *araBAD* and *araJ* appears to cover a larger area than for the strain tagged adjacent to *araFGH*. This is probably because both *araBAD* and *araJ* are located in the Non-Structured Right (NSR) macrodomain

Figure 4.12: Comparison of the position of foci representing AraC controlled promoters seen with and without induction by arabinose in cells with a two foci

Bar chart showing the average relative position of the 1st of 2 foci and 2nd of 2 foci and scatter plot showing the position of the 1st of 2 foci relative to cell length (x axis) against the position of the 2nd of 2 foci relative to cell length (y axis) in the absence and presence of arabinose and glucose.

a) Average relative positions of the 1st of 2 foci in cells tagged adjacent to *araBAD*, *araJ* or *araFGH*, grown with and without induction by arabinose.

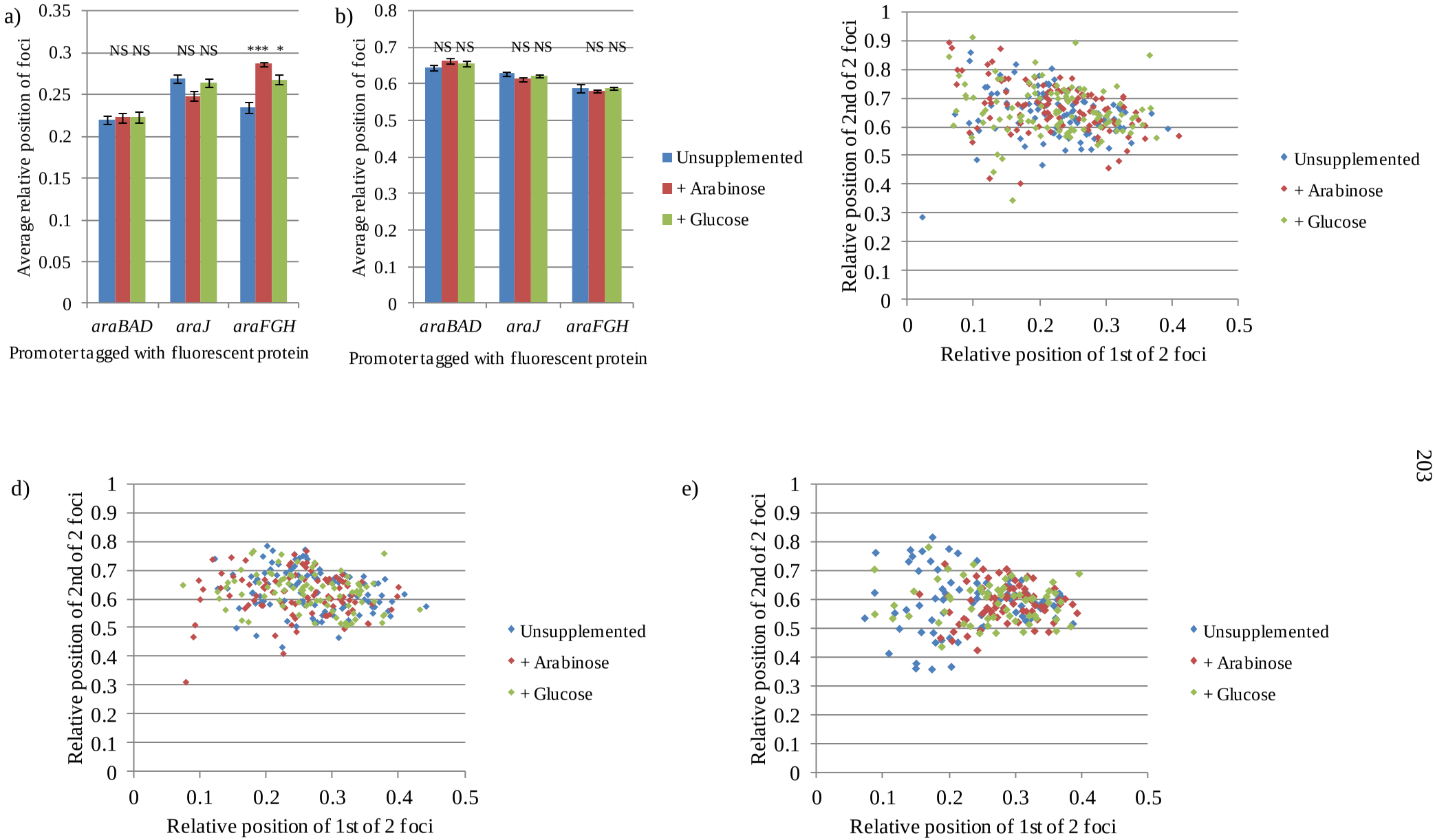
b) Average relative positions of the 2nd of 2 foci in cells tagged adjacent to *araBAD*, *araJ* or *araFGH*, grown with and without induction by arabinose.

c) Strain LR06, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose.

d) Strain LR39, with an insert of 22 LacI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was grown with and without induction by arabinose

e) Strain LR38-pLER108, with an insert of 20 MalI DNA sites adjacent to the *araFGH* promoter, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. The position of foci in cells with 2 foci were measured relative to cell length. The position of the focus closest to a cell pole ("1st of 2 foci) is plotted on the x axis. The position of the further away focus (2nd of 2 foci) was plotted on the y axis. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 350 cells.


a)
Average relative position of foci
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
NS NS
NS NS
*** *
araBAD
araJ
araFGH
Promoter tagged with fluorescent protein
b)
Average relative position of foci
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
NS NS
NS NS
NS NS
araBAD
araJ
araFGH
Promoter tagged with fluorescent protein
Unsupplemented
+ Arabinose
+ Glucose
Relative position of 2nd of 2 foci
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
Relative position of 1st of 2 foci
Unsupplemented
+ Arabinose
+ Glucose
d)
Relative position of 2nd of 2 foci
Relative position of 1st of 2 foci
Unsupplemented
+ Arabinose
+ Glucose
e)
Relative position of 2nd of 2 foci
Relative position of 1st of 2 foci
Unsupplemented
+ Arabinose
+ Glucose

which is less constrained than the Ter domain, where *araFGH* is located. This may allow DNA in the NSR a larger area to diffuse within.

4.5 Colocalisation of foci

4.5.1 Colocalisation of MntR controlled promoters with and without induction by Mn^{2+}

To further investigate the possibility that MntR positions the promoters that it regulates, two MntR controlled promoters are tagged in the same cell. 20 copies of the MalI DNA site are inserted adjacent to the *mntH* promoter in a strain that also has 22 copies of the LacI DNA site adjacent to the *dps* promoter, strain LR47. This strain also has a chromosomal LacI::GFP fusion and was transformed with pLER108, encoding MalI::mCherry. This strain carrying the plasmid will be referred to as LR47-pLER108. A second strain was used in this experiment to compare the localisation of *mntH* to an unrelated promoter. This has 20 copies of the MalI DNA site inserted adjacent to the *mntH* promoter and 22 copies of the LacI DNA site adjacent to the *araBAD* promoter and is called LR42. It also has a chromosomal LacI::GFP fusion and was transformed with plasmid pLER108, encoding a MalI::mCherry fusion. This strain carrying the plasmid will be referred to as LR42-pLER108. Cultures were grown as described previously with 0.1 μM $MnCl_2$ added for the final 10 minutes of growth. Slides were prepared as previously described. The experiment was repeated 3 times, with at least 150 cells analysed from each repeat. An example of an image showing cells tagged adjacent to *dps* and *mntH* is shown in figure 4.13

Cell length was measured for each cell containing 1 GFP and 1 mCherry focus. The distance from the centre of the GFP focus to the closest cell pole was measured, followed by the distance from the centre of the mCherry focus to the same cell pole. Focus to cell pole

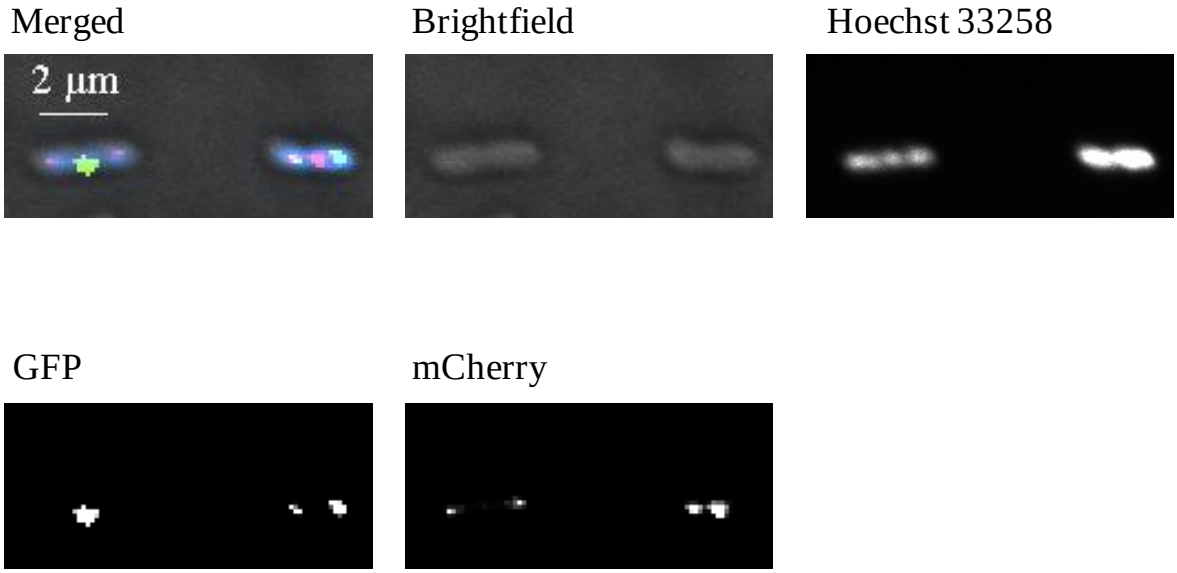


Figure 4.13: Using LacI::GFP and MalI::mCherry to tag *dps* and *mntH* in the same cell

Strain LR47-pLER108 has 22 LacI DNA sites inserted adjacent to the *dps* promoter and 20 MalI DNA sites inserted adjacent to the *mntH* promoter. It also has chromosomally encoded LacI::GFP and is transformed with pLER108, encoding MalI::mCherry. Cells were grown in M9 minimal media supplemented with 0.3% fructose at 23ºC until reaching an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid.

distances were divided by cell length to give a position of each focus relative to cell length. The position of the GFP focus for each cell was plotted on a scatter graph against the position of the mCherry focus. If the two foci colocalise, the points on the graph will fall along a diagonal line, which would go through the origin (as shown in figure 4.14 c and d). The data from LR47-pLER108, tagged adjacent to *mntH* and *dps*, is spread evenly across the graph and there is no significant difference ($P > 0.05$) between cultures grown with and without Mn^{2+} (see figure 4.14a, b and c). This suggests that the induction by Mn^{2+} does not cause MntR controlled promoters to colocalise in the nucleoid. Strain LR42-pLER108, tagged adjacent to *mntH* and *araBAD*, also does not show any significant difference ($P > 0.05$) between cultures grown with and without Mn^{2+} (figure 4.14a, b and d).

4.5.2 Colocalisation of AraC controlled promoters with and without induction by arabinose

To investigate the possibility of AraC playing a role in chromosome structure by bringing the promoters it controls together to a particular location, two AraC controlled promoters were tagged in the same cell to directly compare their positions. 3 strains were made that all had 22 copies of the LacI DNA site are inserted adjacent to the *araBAD* promoter and a chromosomal LacI::GFP fusion. In combinations with these inserts, 20 MalI DNA sites were inserted adjacent to either the *araJ* promoter (LR31), *araFGH* promoter (LR48) or the *mntH* promoter (LR42). These strains were transformed with plasmid pLER108, encoding MalI::mCherry. Strains carrying the plasmid are referred to as LR31-pLER108, LR48-pLER108 and LR42-pLER108, respectively. Cultures were grown as described previously with 0.3% arabinose added for 1 minute before slides were prepared. Slides were prepared as previously described. The experiment was repeated 3 times, with at least 150 cells analysed from each repeat. An

Figure 4.14: Colocalisation of foci representing MntR controlled promoters with and without induction by Mn^{2+}

Bar charts showing average relative positions of foci. Also, scatter plots comparing the positions of two loci on the chromosome, tagged with LacI::GFP (x axis) and MalI::mCherry (y axis). Diagonal line shows the position of colocalising points.

a) Average relative positions of foci in strain LR47-pLER108 with an insert of 22 LacI DNA sites adjacent to the *dps* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

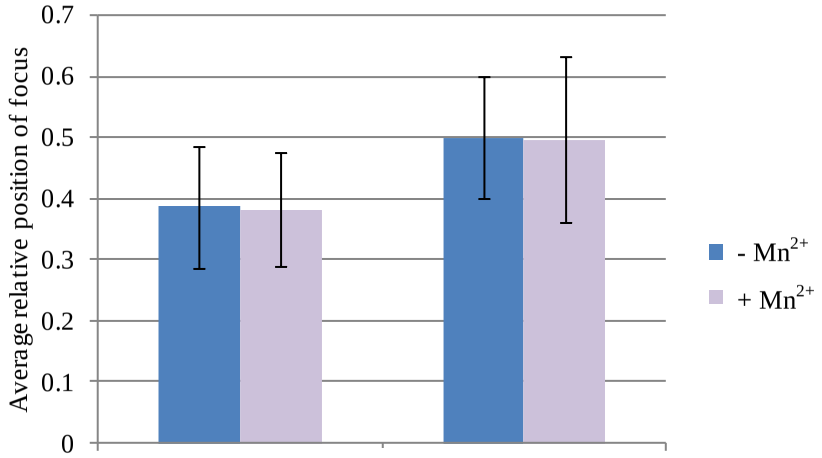
b) Average relative positions of foci in strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

c) Strain LR47-pLER108, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

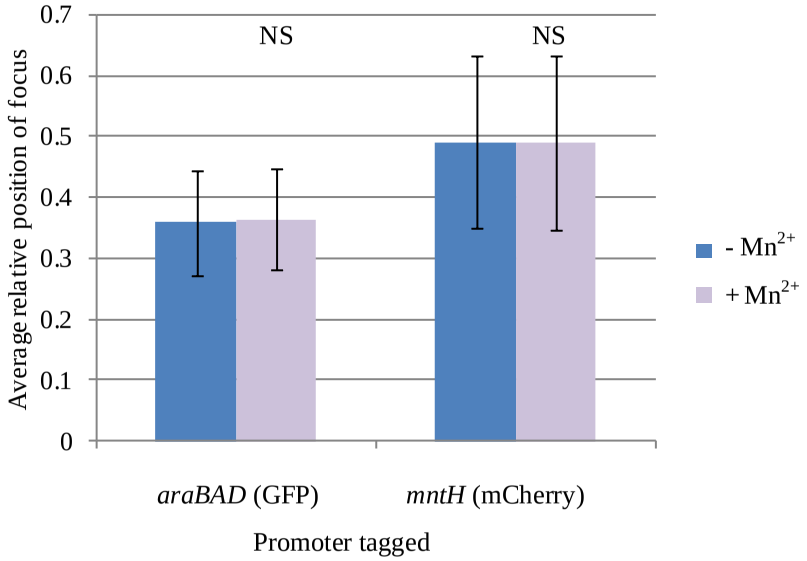
d) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5μg/ml chloramphenicol, if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 μM $MnCl_2$ added for the final 10 minutes of growth. The distance between GFP or mCherry foci and the cell pole were measured in μm. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 500 cells.

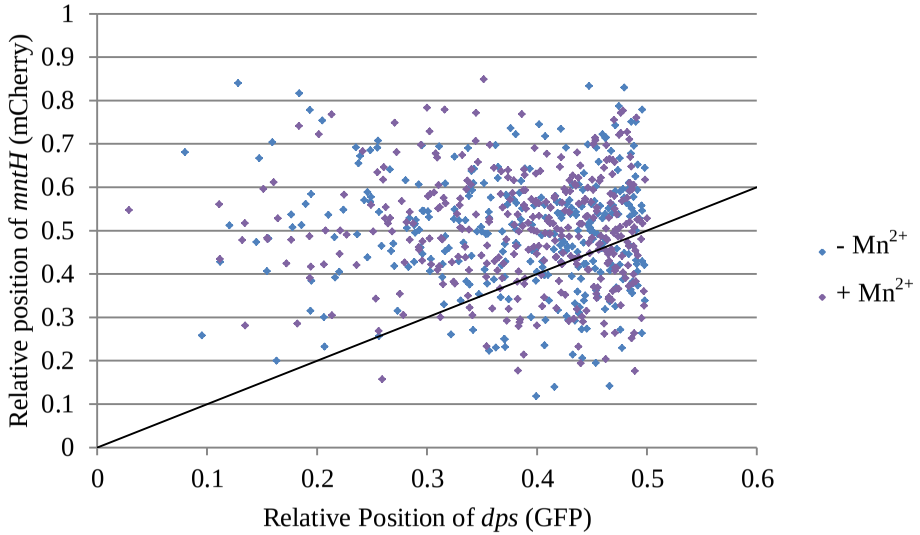
a)


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Average relative position of focus
- Mn2+
+ Mn2+

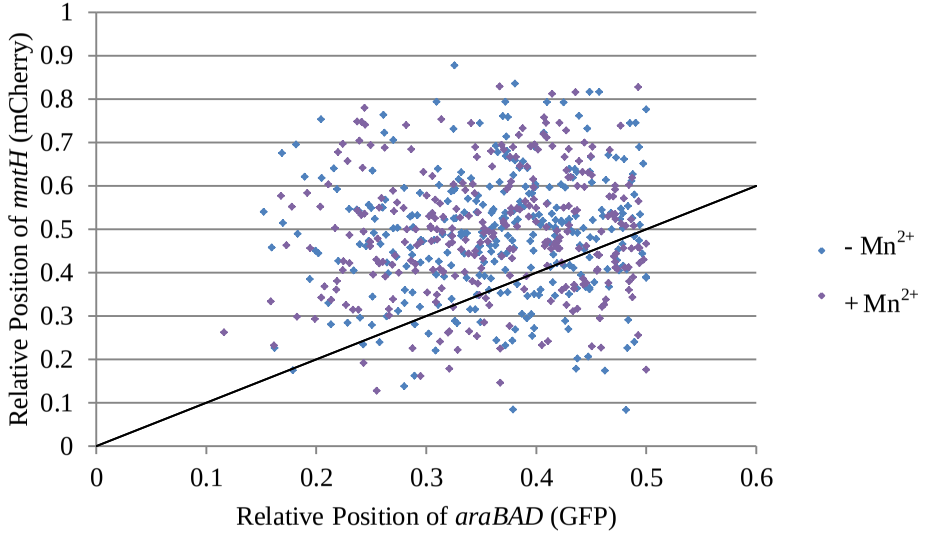
b)


NS
NS
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Average relative position of focus
- Mn2+
+ Mn2+
araBAD (GFP)
mntH (mCherry)
Promoter tagged

c)


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Relative position of mntH (mCherry)
0
0.1
0.2
0.3
0.4
0.5
0.6
Relative Position of dps (GFP)
- Mn2+
+ Mn2+

d)


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Relative Position of mntH (mCherry)
0
0.1
0.2
0.3
0.4
0.5
0.6
Relative Position of araBAD (GFP)
- Mn2+
+ Mn2+

example of an image showing cells tagged adjacent to *araBAD* and *araFGH* is shown in figure 4.15. For cells with one focus each of GFP and mCherry, the distance between the centre of the GFP focus and the cell pole was measured and then the distance between the centre of the mCherry focus and the same cell pole. These distances were represented relative to cell length and plotted on scatter graph, with GFP focus position on the x axis and mCherry focus position on the y axis. Colocalisation would result in the points clustering around a diagonal line, intersecting the origin.

The scatter plot for strain LR31-pLER108, tagging *araBAD* and *araJ*, shows the points are spread out with a no obvious clustering on the diagonal line, suggesting that these promoters do not colocalise (see figure 4.16a and d). Strain LR48-pLER108, tagged adjacent to *araBAD* and *araFGH*, gave data that is very evenly spread suggesting that in many cells these 2 promoters are at distant points. For these 2 strains there was no significant difference between the positions of foci with and without induction by arabinose ($P > 0.05$) (see figure 4.16b and e). The data for strain LR42-pLER108, tagged adjacent to *araBAD* and *mntH*, is also spread out on the graph suggesting there is no colocalisation (see figure 4.16c and f). However, there is a significant difference between the cultures grown with and without arabinose ($P < 0.01$), with the average position of the focus representing the position of *mntH* changing from 0.476 to 0.507 upon addition of arabinose, so moving towards the centre of the cell upon induction. The average position of the focus representing *araBAD* was 0.350 in the absence of arabinose and 0.345 in the presence, with no significant difference. This was unexpected as *mntH* and the surrounding genes are not regulated by AraC and have no known connection to arabinose. The reason for this change is unknown. The points representing foci adjacent to *araBAD* and *araJ* are much tighter together than the corresponding points for *araBAD* and *araFGH*. This

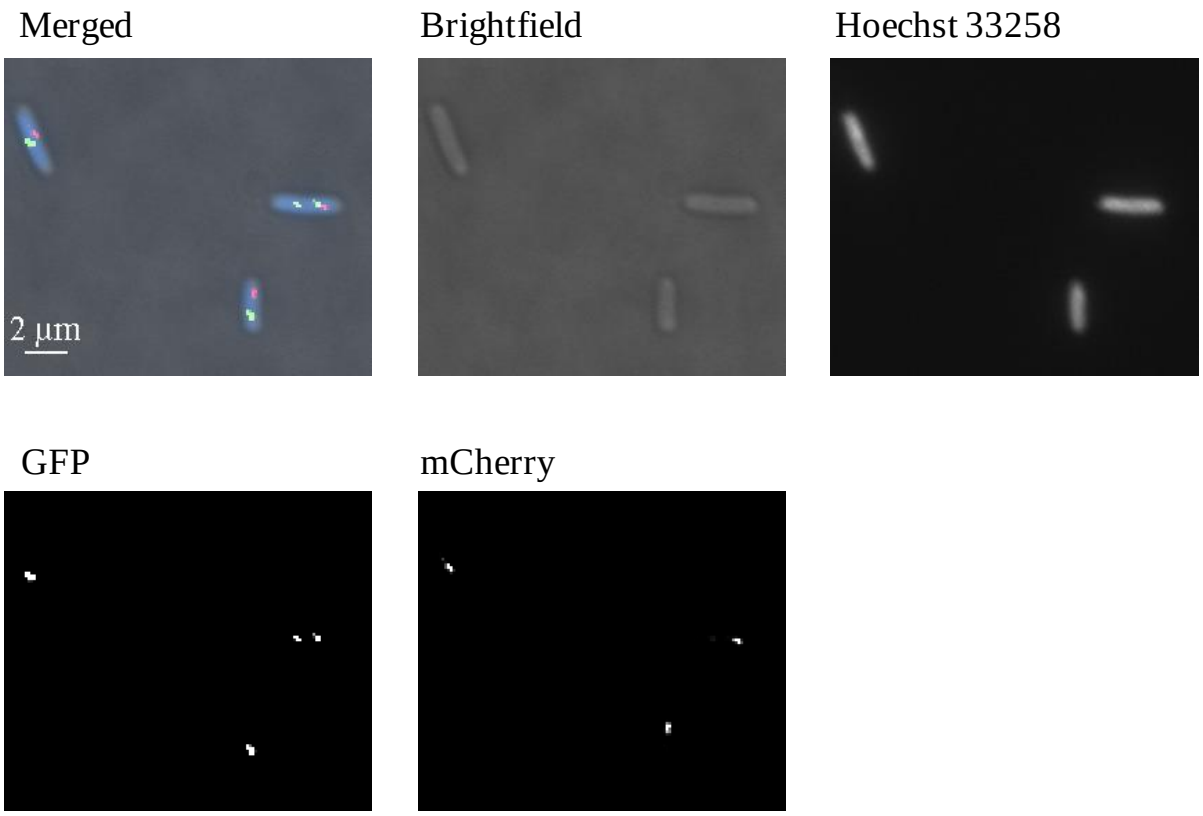


Figure 4.15: Using LacI::GFP and MalI::mCherry to tag *araBAD* and *araFGH* in the same cell

Strain LR48-pLER108 has 22 LacI DNA sites inserted adjacent to the *araBAD* promoter and 20 MalI DNA sites inserted adjacent to the *araFGH* promoter. It also has chromosomally encoded LacI::GFP and is transformed with pLER108, encoding MalI::mCherry. Cells were grown in M9 minimal media supplemented with 0.3% fructose at 23ºC until reaching an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid.

Figure 4.16: Colocalisation of foci representing AraC controlled promoters with and without induction by arabinose

Bar Charts showing the average position of foci. Also, scatter plot comparing the positions of two loci on the chromosome, tagged with LacI::GFP (x axis) and MalI::mCherry (y axis). Diagonal line shows the position of colocalising points.

a) Average positions of foci in strain LR31-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

b) Average positions of foci in strain LR48-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araFGH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

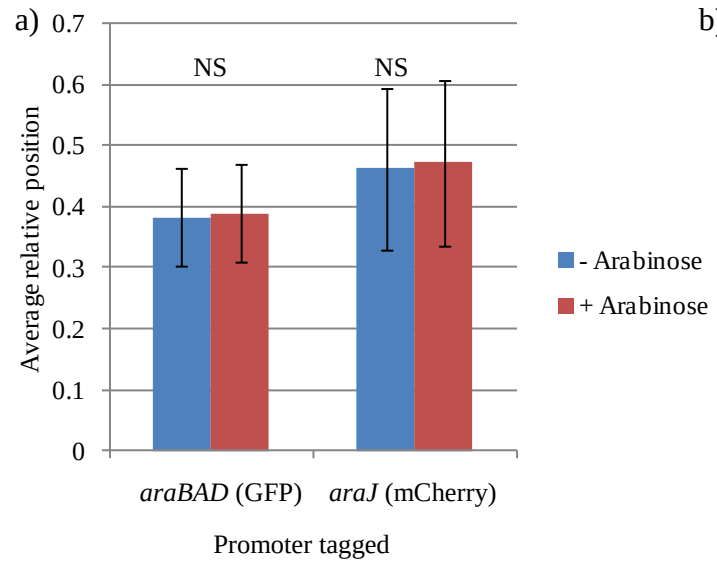
c) Average positions of foci in strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

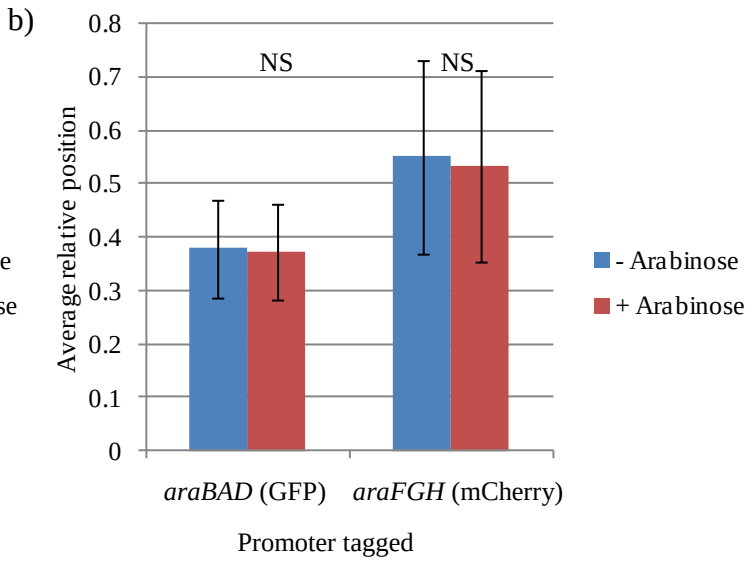
d) Strain LR31-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

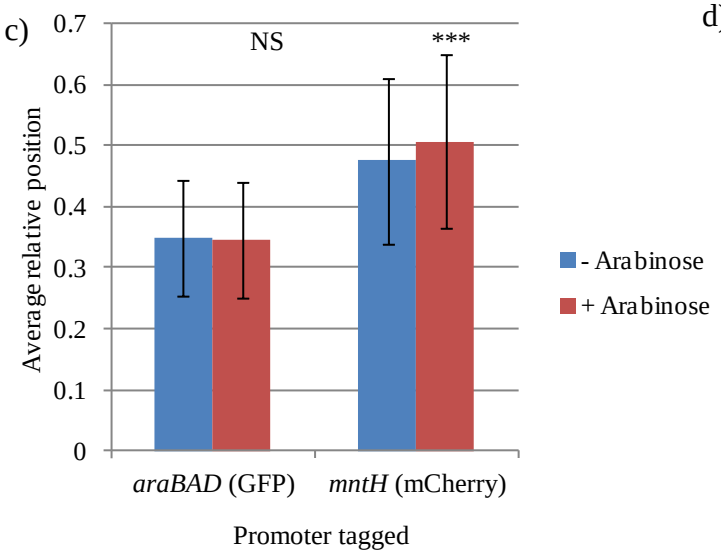
e) Strain LR48-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araFGH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

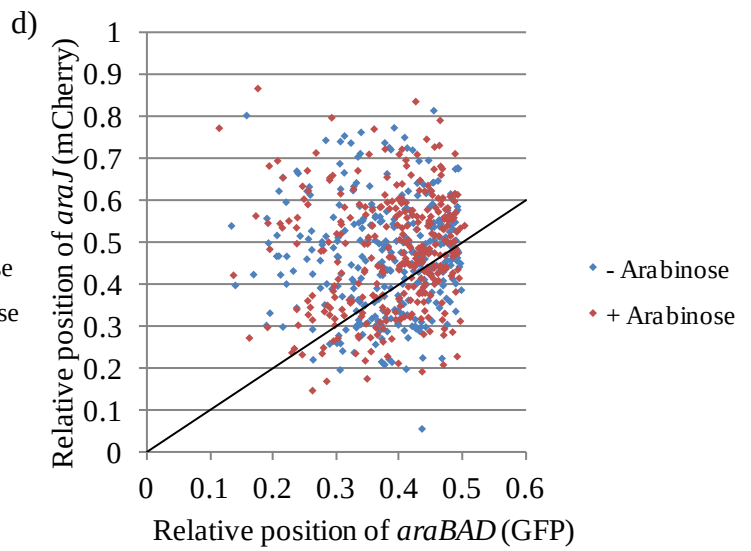
f) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

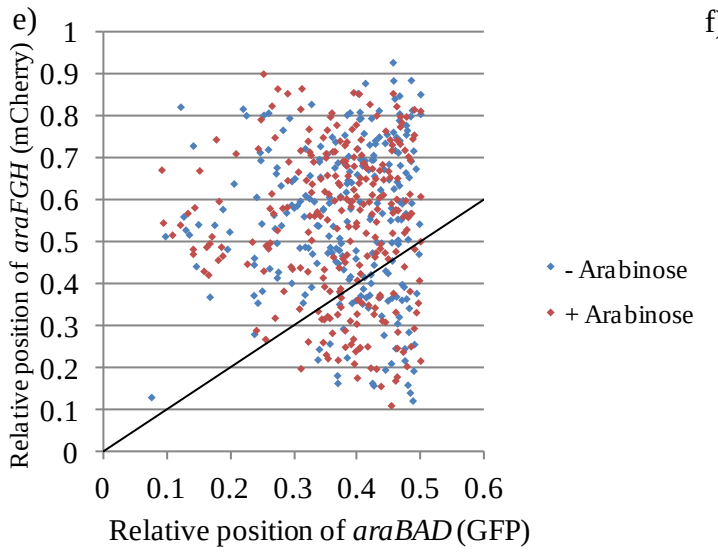
Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose added for 1 minute before slides were prepared. The distance from the GFP focus to the closest cell pole was measured relative to cell length and plotted on the x axis. The distance from the mCherry focus to the same cell pole was measured relative to cell length and plotted on the y axis. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 500 cells.

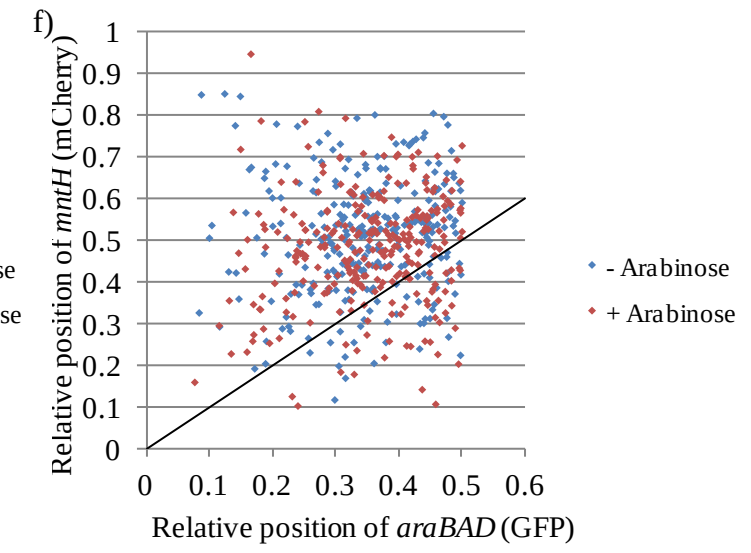

a)
Average relative position
NS
NS
- Arabinose
+ Arabinose
araBAD (GFP)
araJ (mCherry)
Promoter tagged


b)
Average relative position
NS
NS
- Arabinose
+ Arabinose
araBAD (GFP)
araFGH (mCherry)
Promoter tagged


c)
Average relative position
NS
***
- Arabinose
+ Arabinose
araBAD (GFP)
mntH (mCherry)
Promoter tagged


d)
Relative position of araJ (mCherry)
- Arabinose
+ Arabinose
Relative position of araBAD (GFP)


e)
Relative position of araFGH (mCherry)
- Arabinose
+ Arabinose
Relative position of araBAD (GFP)


f)
Relative position of mntH (mCherry)
- Arabinose
+ Arabinose
Relative position of araBAD (GFP)

is to be expected as *araBAD* and *araJ* are both located in the NSR macrodomain whereas *araFGH* is in the Ter domain. Although the NSR domain is less restricted than other macrodomains it would still be more likely to find *araBAD* in a similar area of the cell to another gene in NSR than a gene from a different macrodomain. These results suggest that AraC binding does not change chromosome structure by bringing different AraC-regulated promoters together, supporting the result seen in section 4.4.3.

4.6 Distance between promoters with and without induction

4.6.1 Distance between foci representing MntR controlled promoters with and without induction by Mn^{2+}

To investigate the distance between 2 MntR-regulated promoters, strains with 2 FROS tags were used. Strain LR47 has 22 LacI DNA sites adjacent to *dps* and 20 MalI DNA sites adjacent to *mntH*. The distance between *mntH* and *araBAD* was also measured as *araBAD* is not controlled by MntR so can be used as a reference point. Strain LR42 has 22 LacI DNA sites inserted adjacent to *araBAD* and 20 MalI DNA sites adjacent to *mntH*. Both of these strains also have a chromosomal LacI::GFP fusion and were transformed with a plasmid encoding MalI::mCherry, pLER108. Strains transformed with the plasmid are referred to as the strain name-pLER108. Cells with one focus of each colour were analysed. The distance from the centre of the GFP focus was measured to the closest cell pole before the distance from the centre of the mCherry focus was measured to the same pole. As cell length is not relevant for this analysis, distance was measured in μm rather than relative to cell length to give an actual value to the distance between the 2 loci. Cultures were grown as described previously with 0.1 μM $MnCl_2$ added for the final 10 minutes of growth. Slides were prepared as previously described. The experiment was repeated 3 times, with at least 150 cells analysed from each repeat.

The percentage of cells with the distance between foci at each 0.1 μm from 0 to 1.5 μm was calculated for cells grown with and without Mn^{2+}. Strain LR47-pLER108, tagged at *mntH* and *dps*, did not show any significant difference ($P > 0.05$) between the distance between the 2 locations when cells are grown with or without Mn^{2+} (see figure 4.17a and b). Strain LR42-pLER108, tagged adjacent to *mntH* and *araBAD*, also had no significant difference ($P > 0.05$) in the distance between the 2 locations when grown with and without Mn^{2+} (see figure 4.17a and c). To further investigate, the cells with a distance between the 2 locations tagged that was less that 0.5 μm were split to show the percentage of cells for each 0.02 μm. Again, there was no significant difference (see figure 4.18). This suggests that MntR does not alter the chromosome structure by repositioning promoters that it controls.

4.6.2 Distance between foci representing AraC controlled promoters with and without induction by arabinose

To further investigate the possibility of the induction state of AraC controlled promoters affecting their position within the cell. 22 LacI DNA sites are inserted adjacent to *araBAD* and 20 MalI DNA sites are inserted adjacent to *araJ* in strain LR31. 22 LacI DNA sites are inserted adjacent to *araBAD* and 20 MalI DNA sites are inserted adjacent to *araFGH* in strain LR48. To measure the position of *araBAD* against a position that should be unaffected by the addition of arabinose, a third strain was made with 22 LacI DNA sites adjacent to *araBAD* and 20 MalI DNA sites adjacent to *mntH*, strain LR42. All 3 strains also had a chromosomal LacI::GFP fusion and were transformed with a plasmid encoding MalI::mCherry, pLER108. Strains transformed with the plasmid will be referred to as the strain name-pLER108. Cells with one focus of each colour were analysed. The distance from the centre of the GFP focus was measured to the closest cell pole before the distance from the centre of the mCherry focus was measured to the same pole. As cell length is not relevant for this analysis, distance was

Figure 4.17: Distance between of foci representing MntR controlled promoters with and without induction by Mn^{2+}

Bar chart showing the distance in μm between LacI::GFP and MalI::mCherry foci.

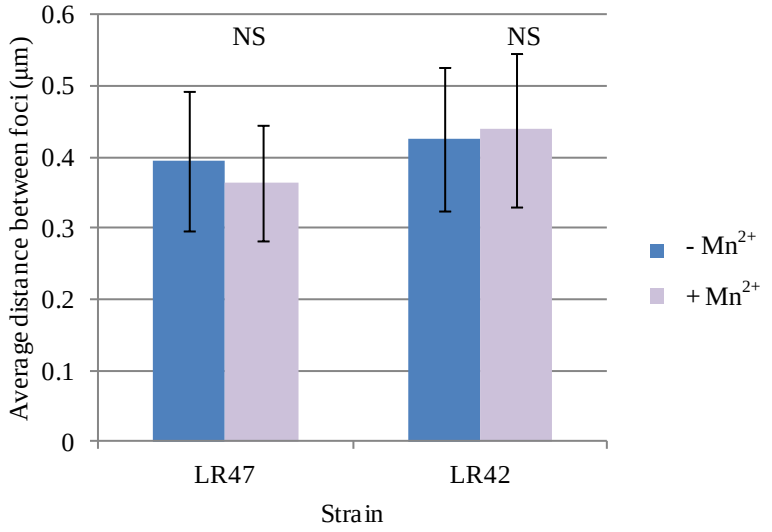
a) Average distance between a LacI::GFP focus a MalI::mCherry focus in strain LR47, tagged adjacent to *mntH* and *dps*, and LR42, tagged adjacent to *araBAD* and *mntH*.

b) Strain LR47-pLER108, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

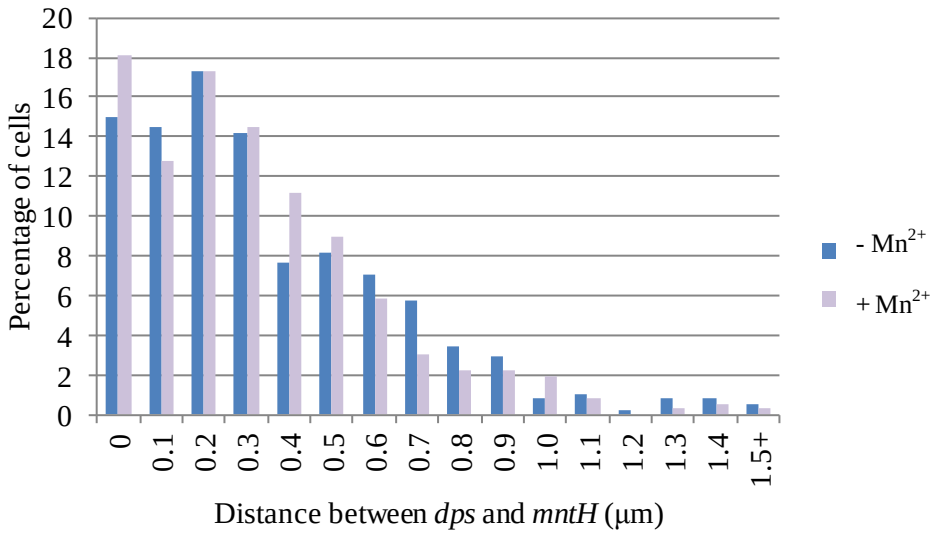
c) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by Mn^{2+}.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 μM $MnCl_2$ added for the final 10 minutes of growth. The distance between GFP and mCherry foci were measured in μm. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 500 cells.

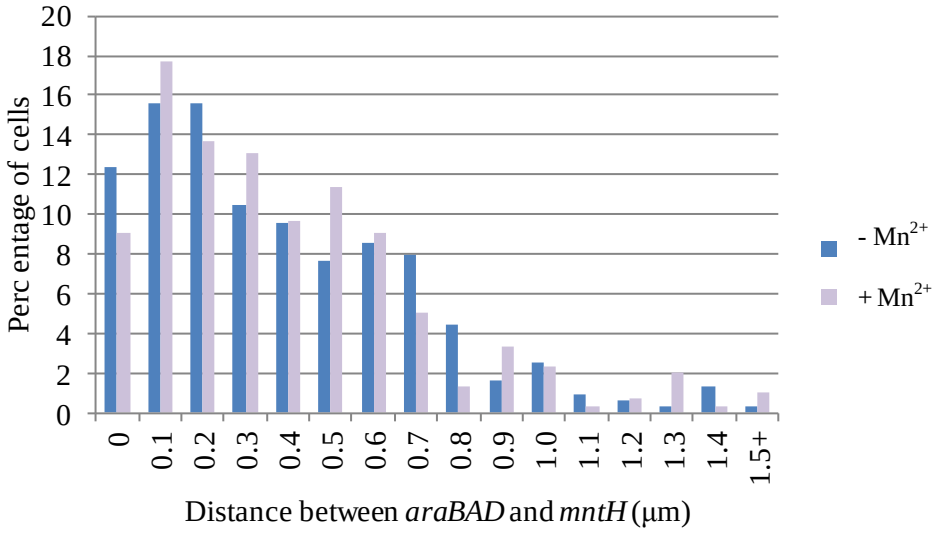
a)


0.6
0.5
0.4
0.3
0.2
0.1
0
NS
NS
Average distance between foci (µm)
LR47
LR42
Strain
- Mn2+
+ Mn2+

b)


20
18
16
14
12
10
8
6
4
2
0
Percentage of cells
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5+
Distance between dps and mntH (µm)
- Mn2+
+ Mn2+

c)


20
18
16
14
12
10
8
6
4
2
0
Perc entage of cells
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5+
Distance between araBAD and mntH (µm)
- Mn2+
+ Mn2+

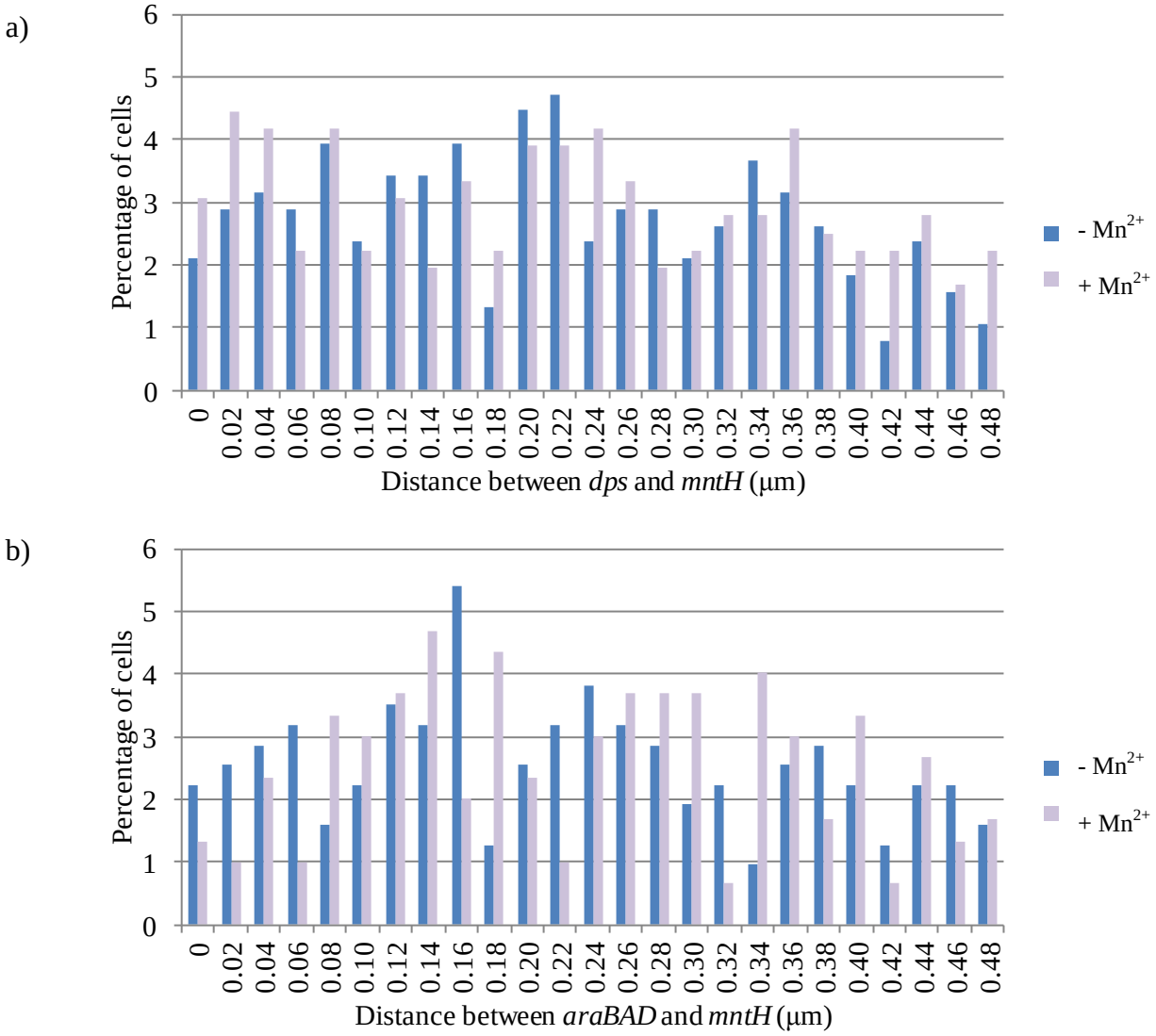


Figure 4.18: Distance between of foci representing MntR controlled promoters with and without induction by manganese

Bar chart showing the distance in µm between LacI::GFP and MalI::mCherry foci

a) Strain LR47-pLER108, with an insert of 22 LacI DNA sites adjacent to the *dps* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by manganese.

b) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by manganese.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.1 µM $MnCl_2$ added for the final 10 minutes of growth. The distance between GFP and mCherry foci were measured in µm. Data shown for cells with foci under 0.5 µm apart. n = > 100

measured in μm rather than relative to cell length to give an actual value to the distance between the 2 loci. Cultures were grown as described previously with 0.3% arabinose added for 1 minute before slides were prepared. Slides were prepared as previously described. The experiment was repeated 3 times, with at least 150 cells analysed from each repeat.

The percentage of cells with the distance between foci at each 0.1 μm from 0 to 1.5 μm was calculated for cells grown with and without arabinose. The distance between foci in strain LR31-pLER108, tagged adjacent to *araBAD* and *araJ*, did not change significantly ($P > 0.05$) with addition of arabinose (see figure 4.19a and b). Strain LR48-pLER108, tagged adjacent to *araBAD* and *araFGH*, also showed no significant difference in the distance between loci with and without arabinose ($P > 0.05$) (see figure 4.19a and c). When the cells with foci located within 0.5 μm of each other were further studied, there was also no significant difference for either of these strains when grown with and without arabinose ($P > 0.05$). Strain LR42-pLER108, tagged adjacent to *araBAD* and *mntH*, was made to investigate the position of the *araBAD* promoter when compared to a promoter that is not regulated by AraC. There was a significant change in the distance between these two positions when the cells were grown with arabinose ($P < 0.05$), with the distance increasing with addition of arabinose (see figure 4.19 a and d). When just cells with foci located within 0.5 μm of each other were studied it was found there was no significant difference ($P > 0.05$) (see figure 4.20). As seen previously (Section 4.5.2) when colocalisation of *araBAD* and *mntH* was studied, there was a change in the position of the mCherry focus representing the position of *mntH*. The average position of *araBAD* was 0.350 and 0.345 in the absence and presence of arabinose. The average position of *mntH* changed from 0.475 to 0.507 with the addition of arabinose, explaining the increase in distance

Figure 4.19: Distance between of foci representing AraC controlled promoters with and without induction by arabinose

Bar chart showing the distance in μm between LacI::GFP and MalI::mCherry foci.

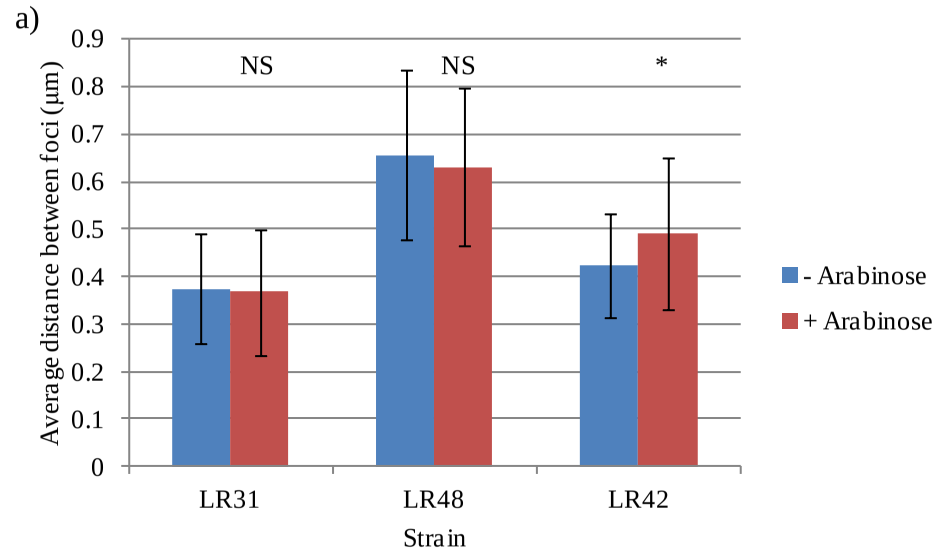
a) Average distance between a LacI::GFP focus a MalI::mCherry focus in strain LR31, tagged adjacent to *araBAD* and *araJ*, LR48, tagged adjacent to *araBAD* and *araFGH*, and LR42, tagged adjacent to *araBAD* and *mntH*.

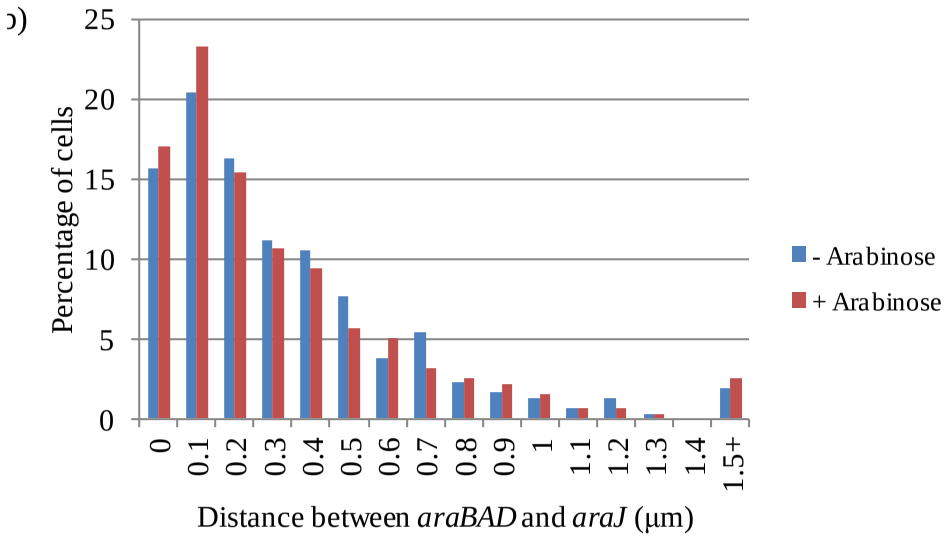
b) Strain LR31-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

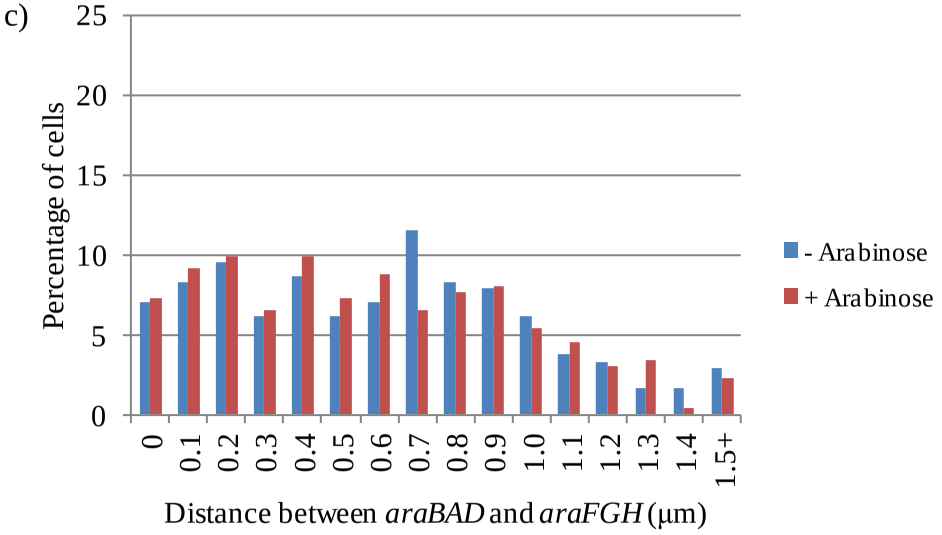
c) Strain LR48-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araFGH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

d) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose added for 1 minute before slides were prepared. The distance between GFP and mCherry foci were measured in μm. A star above a bar represents data that is significantly different to unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 500 cells.


a)
Average distance between foci (µm)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
NS
NS
*
LR31
LR48
LR42
Strain
- Arabinose
+ Arabinose


b)
Percentage of cells
25
20
15
10
5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5+
Distance between araBAD and araJ (µm)
- Arabinose
+ Arabinose


c)
Percentage of cells
25
20
15
10
5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5+
Distance between araBAD and araFGH (µm)
- Arabinose
+ Arabinose

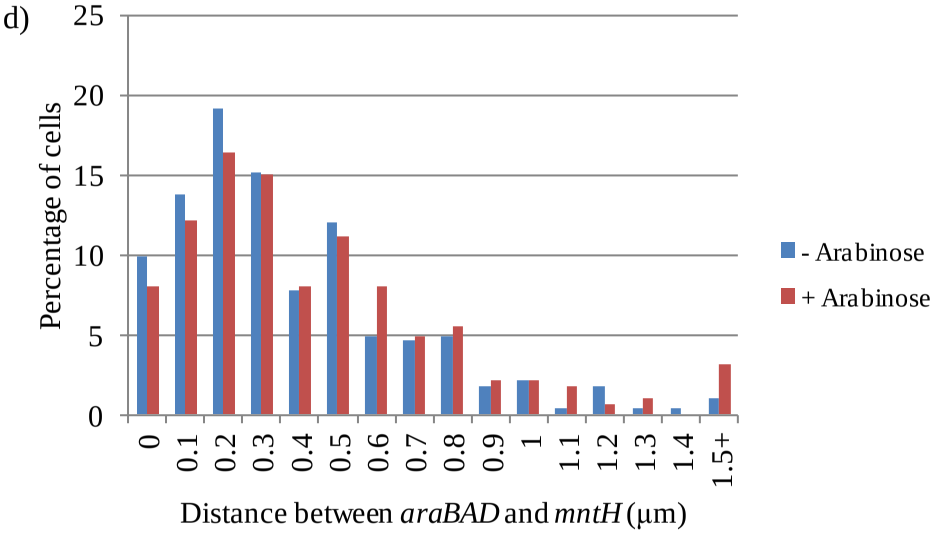

d)
Percentage of cells
25
20
15
10
5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5+
Distance between araBAD and mntH (µm)
- Arabinose
+ Arabinose

Figure 4.20: Distance between of foci representing AraC controlled promoters with and without induction by arabinose

Bar chart showing the distance in μm between LacI::GFP and MalI::mCherry foci.

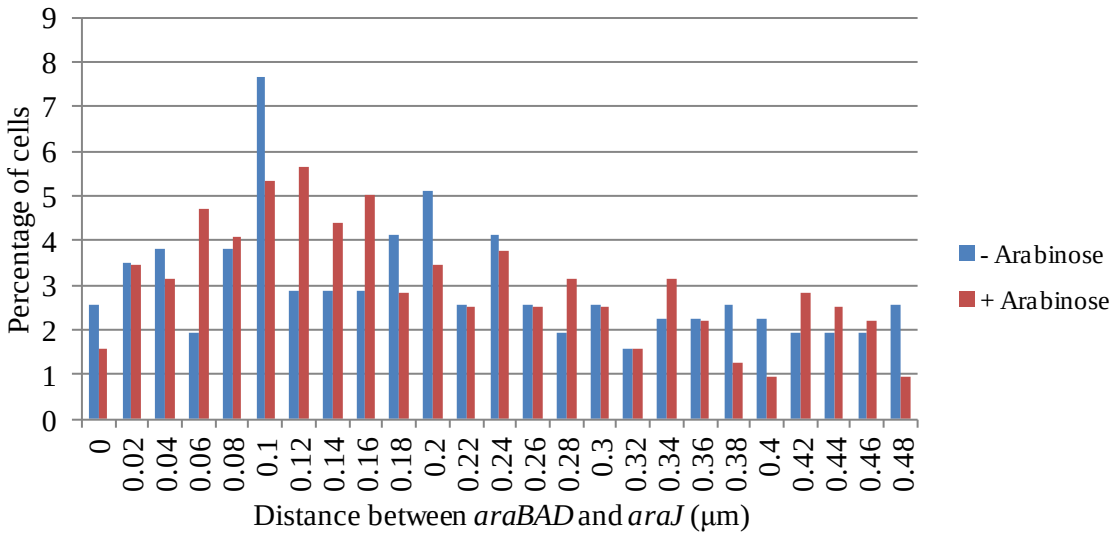
a) Strain LR31-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araJ* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

b) Strain LR48-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *araFGH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

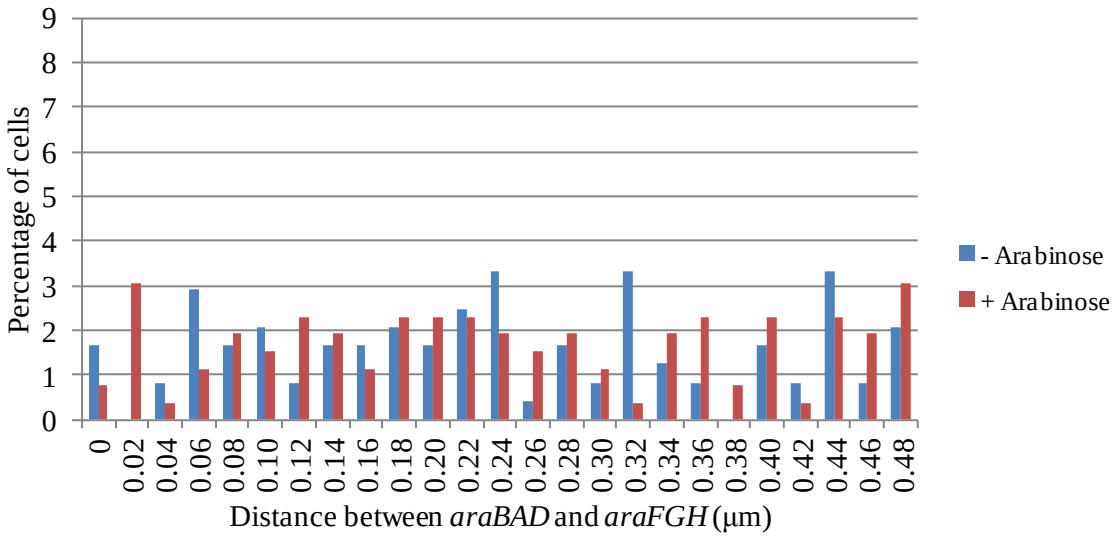
c) Strain LR42-pLER108, with an insert of 22 LacI DNA sites adjacent to the *araBAD* promoter, 20 MalI DNA sites adjacent to the *mntH* promoter and a chromosomal LacI::GFP fusion, was transformed with pLER108, carrying MalI::mCherry, and grown with and without induction by arabinose.

Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose added for 1 minute before slides were prepared. The distance between GFP and mCherry foci were measured in μm. Data shown for cells with foci under 0.5 μm apart. n = > 100

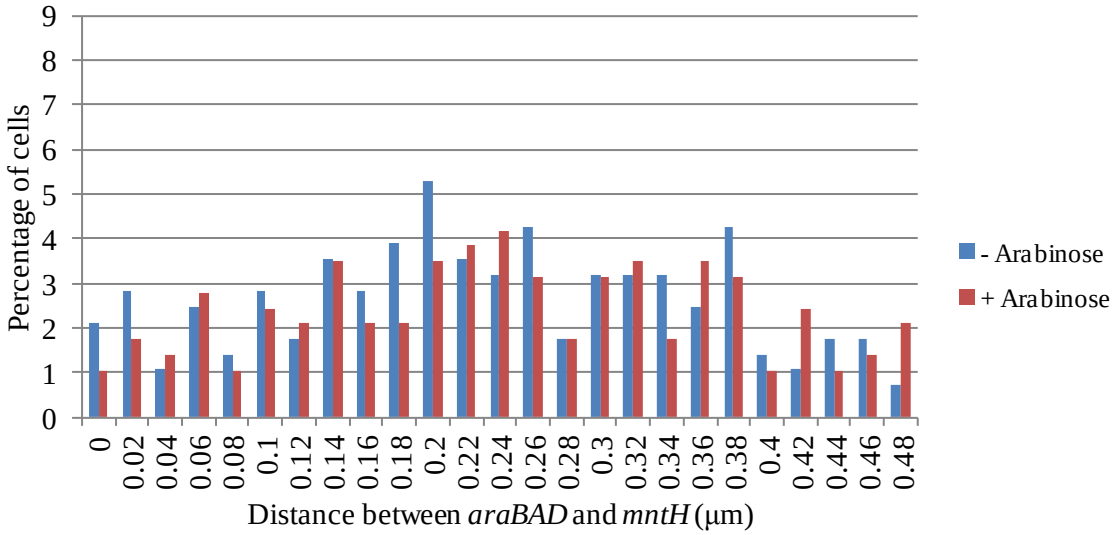
a)


Percentage of cells
0
1
2
3
4
5
6
7
8
9
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24 0.26 0.28 0.3 0.32 0.34 0.36 0.38 0.4 0.42 0.44 0.46 0.48
Distance between *araBAD* and *araJ* (µm)
- Arabinose
+ Arabinose

b)


Percentage of cells
0
1
2
3
4
5
6
7
8
9
0 0.02 0.04 0.06 0.08 0.10 0.12 0.14 0.16 0.18 0.20 0.22 0.24 0.26 0.28 0.30 0.32 0.34 0.36 0.38 0.40 0.42 0.44 0.46 0.48
Distance between *araBAD* and *araFGH* (µm)
- Arabinose
+ Arabinose

c)


Percentage of cells
0
1
2
3
4
5
6
7
8
9
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24 0.26 0.28 0.3 0.32 0.34 0.36 0.38 0.4 0.42 0.44 0.46 0.48
Distance between *araBAD* and *mntH* (µm)
- Arabinose
+ Arabinose

between these two points. However, the expression of *mntH* was not expected to be affected by arabinose.

4.7 Induction of promoters in cells at the point of division and its effects on segregation of sister chromatids

4.7.1 Separation of sister chromatids by induction of the *araFGH* promoter

Cultures are a mixture of single cells and cells at the point of division, so the cultures used to measure the position of promoters could also be used to study the segregation of sister chromatids in dividing cells, at various locations on the chromosome. The percentage of single cells in cultures varied between 85 and 95%, independent of the strain analysed and the conditions used. This means that between 5 and 15% of cells imaged were at the point of division. Cells at the point of division were defined as those that could be seen as two individual cells in a brightfield image, presumably indicating that the septum has formed, and seemed to have two separate nucleoids when viewed with the DAPI filter. FROS was used to study the effect of promoter induction on the segregation of sister chromatids. 20 MalI DNA sites are inserted adjacent to the *araFGH* promoter in strain LR38, which was transformed with a plasmid encoding MalI::mCherry, pLER108. This transformed strain is referred to as LR38-pLER108. 22 LacI DNA sites are inserted adjacent to *araBAD* in strain LR06, which also has a chromosomal LacI::GFP fusion. 22 LacI DNA sites are also inserted adjacent to *dps* in strain SXB4, which also has a chromosomal LacI::GFP fusion. Cultures were grown as described previously, with 0.3% arabinose or glucose added for 1 minute before slides were prepared. Slides were prepared as previously described. Cells at the point of division were analysed, and the percentage of these that had a single focus in the centre of the two new cells were counted (see figure 4.21a). The presence of a single focus between the two new cells indicates that either the chromosome has not yet been replicated at the point that has been

Figure 4.21: Images of cells at the point of division

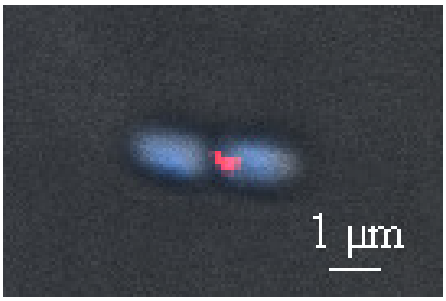
Strain LR38-pLER108 has 20 MalI DNA sites inserted adjacent to the *araFGH* promoter and is transformed with pLER108, encoding MalI::mCherry. Cells were grown in M9 minimal media supplemented with 0.3% fructose at 23ºC until reaching an OD_{650} of approximately 0.1. Hoechst 33258 was used to stain the nucleoid. Cells at the point of division were defined as cells which had separate nucleoids and two separate cells could be seen in a brightfield image.

a) Cells with a single, central focus.

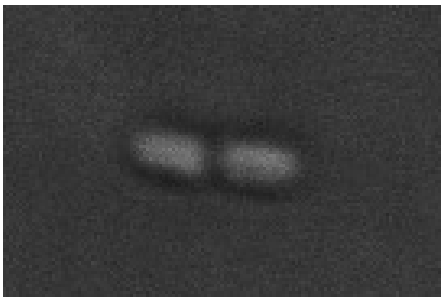
b) Cells with two separate foci.

a)

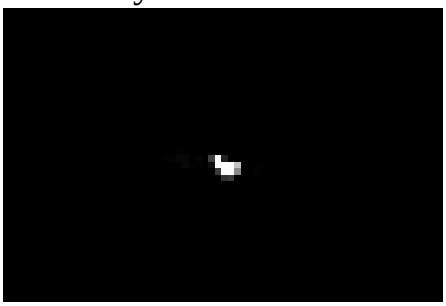
Merged



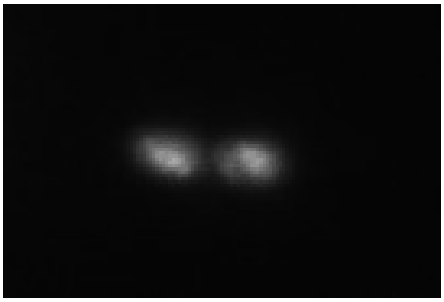
Brightfield



mCherry

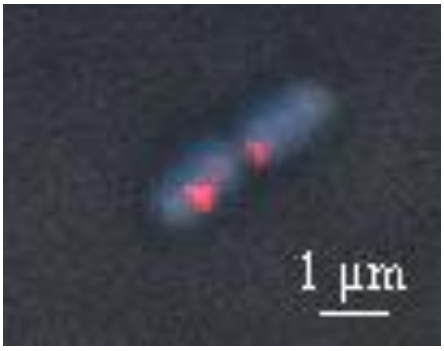


Hoechst 33258



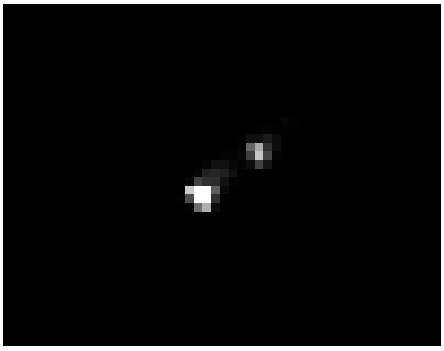
b)

Merged

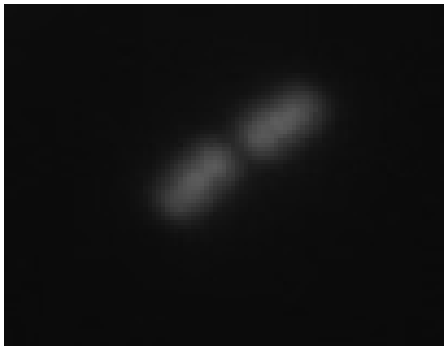


Brightfield

mCherry



Hoechst 33258



tagged, or it has been replicated but the sister chromatids have not segregated. If the sister chromatids have not been segregated, although there are two copies of the DNA site array on the chromosome, they are not far enough apart to be distinguished as separate foci (see figure 4.21b).

The LR38-pLER108 cells, tagged at *araFGH*, had the highest percentage of cells with a single focus, with over 35% of cells from unsupplemented cultures showing this arrangement (see figure 4.22). In strain SXB4, tagged at *dps*, 7% of cells had a single focus, and just 2% in LR06, tagged adjacent to *araBAD*. This difference can be explained by looking at the positions of these three genes on the chromosome. *araBAD* is located in the NSR macrodomain, near to the Ori macrodomain, which means it is replicated early in the cell cycle. Very few pairs of dividing cells in rapidly growing cultures, if any, tagged at this position would be expected to have a single focus. *dps* is in the Right macrodomain which is further away from OriC so would be replicated later in the cell cycle, hence the slightly higher percentage of dividing cells with a single focus. *araFGH* is replicated in the Ter macrodomain, which is last to be replicated. It is also known that sister chromatids in the Ter macrodomain delay before segregating, more than the other macrodomains (Possoz *et al.*, 2012). *araFGH* is located at the very edge of the Ter macrodomain, so it is possible that a site further towards the dif would have a higher percentage of dividing cells with a single focus.

When cultures were supplemented with arabinose or glucose, there was no significant difference ($P > 0.05$) seen in the percentage of dividing cells with a single central focus for strains LR06 and SXB4, tagged adjacent to *araBAD* and *dps*, respectively. However, when

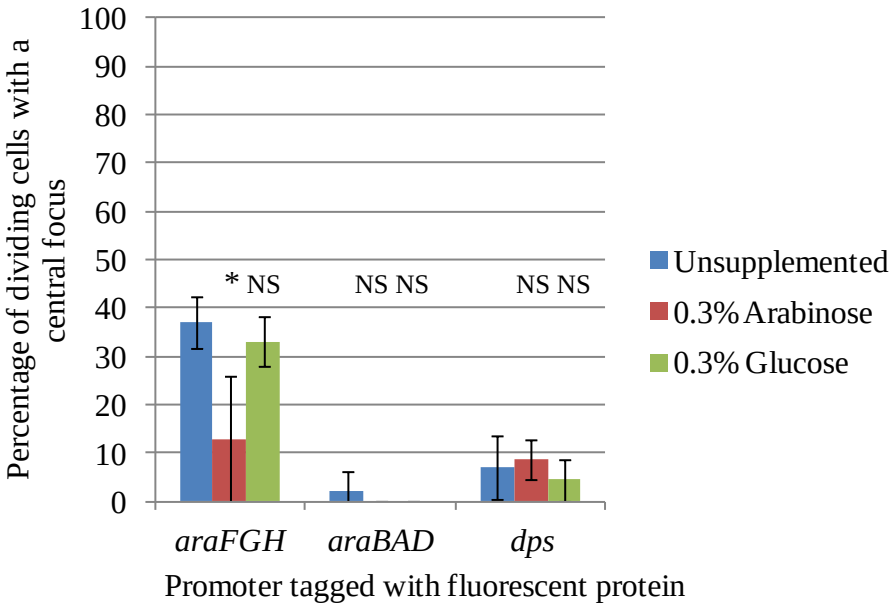


Figure 4.22: Segregation of sister chromatids in cells at the point of division

Bar chart showing the percentage of cells at the point of division had a single central focus (unsegregated sister chromatids at the position tagged).

Strains tagged adjacent to *araFGH* (LR38-pLER108), *araBAD* (LR06) and *dps* (SXB4) were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. Dividing cells were analysed and the percentage of dividing cells which had one focus between the 2 cells was calculated for each strain in each condition. A star above a bar represents data that is significantly different from unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = > 550 cells.

strain LR38-pLER108, tagged adjacent to *araFGH*, was supplemented with arabinose there was a significant drop in the percentage of dividing cells with a single focus ($P < 0.05$). This effect was not seen in cultures supplemented with glucose. This shift from around 35% of cells at the point of division having a single focus to an average of 12.5%, when the *araFGH* promoter is induced, is evidence that induction of a promoter can have an effect on chromosome structure. It is likely that many of the uninduced cells had two copies of the chromosome, at the position of the *araFGH* gene and the tag, but these had not yet segregated. Induction of *araFGH* appears to speed up the segregation process, possibly because the promoter is not accessible to the transcription machinery in the unsegregated sister chromatids. Although the addition of a sugar to the growth medium could increase growth rate, and therefore the rate of chromosome replication, this effect is specific to arabinose and is not seen in the presence of glucose.

4.7.2 Comparison of dividing cells tagged at different points in the Ter macrodomain with and without induction by arabinose

To check whether the addition of arabinose to a culture has an effect on the separation of loci throughout the whole of the Ter macrodomain, or if the effect is specific to the region adjacent to the *araFGH* locus, strains with tags at different positions in the Ter macrodomain were compared. 20 MalI DNA sites are inserted adjacent to *araFGH* in strain LR38, which was transformed with a plasmid encoding MalI::mCherry, pLER108. This strain transformed with the plasmid is referred to as LR38-pLER108. 22 LacI DNA sites are inserted adjacent to the gene *flxA* in strain KC01, which also has a chromosomal LacI::GFP fusion. *flxA* is located in the Ter macrodomain around 40 kb from the *dif*. The insert adjacent to *araFGH* is approximately 384 kb from the *dif*. Cultures were grown as described previously with 0.3% arabinose added for 1 minute before slides were prepared. Slides were prepared as previously

described. Cells at the point of division were analysed and the percentage of these that had a single focus in the centre of the two new cells were counted.

In strain KC01, tagged adjacent to *flxA*, around 75% of cells at the point of division had unsegregated chromosomes, with no significant difference in the presence and absence of arabinose ($P > 0.05$) (see figure 4.23). Strain LR38-pLER108, tagged adjacent to *araFGH*, showed a significant drop in the number of cells at the point of division with unsegregated chromosomes in the presence of arabinose ($P < 0.01$). This confirms that the change in chromosome structure in the Ter macrodomain when cells are induced with arabinose is restricted to the region surrounding the *araFGH* gene. It is not known how far the effect of arabinose induced segregation will extend. The insert of multiple MalI DNA sites is 1.3 kb from the *araFGH* transcription start site, spanning another gene, so this region does not just include the *araFGH* promoter and operon.

4.7.3 Separation of sister chromatids by induction of the *araFGH* promoter in cells treated with rifampicin

To check whether the effect on sister chromatid separation seen when the *araFGH* promoter is induced is dependent on transcription, the experiment was repeated in the presence of rifampicin. Rifampicin works by binding RNA polymerase in the channel where the DNA/RNA hybrid sits and prevents RNAP from extending past 2 or 3 nucleotides (Campbell *et al.*, 2001). 20 MalI DNA sites are inserted adjacent to *araFGH* in strain LR38, which was also transformed with a plasmid encoding MalI::mCherry, pLER108. Strain LR38 transformed with the plasmid is referred to as LR38-pLER108. 22 LacI DNA sites are inserted adjacent to *dps* in strain SXB4, which also has a chromosomal LacI::GFP fusion.

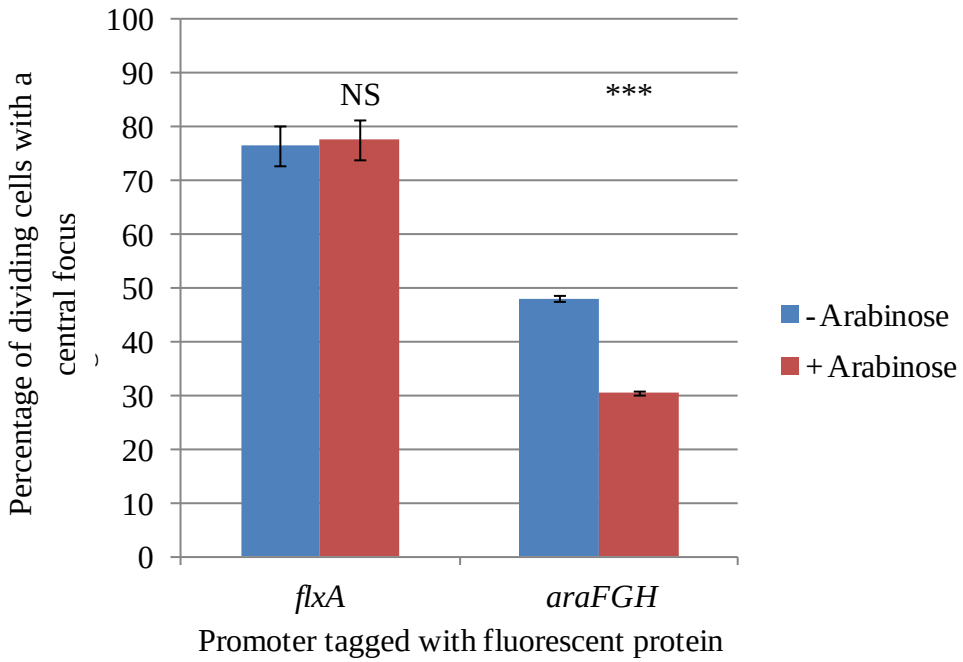


Figure 4.23: Segregation of sister chromatids in cells at the point of division at different locations in the Ter macrodomain

Bar chart showing the percentage of cells at the point of division with a single central focus.

20 MalI binding sites were inserted adjacent to *araFGH* in strain LR38-pLER108, which was transformed with plasmid pLER108, encoding MalI::mCherry. 22 LacI binding sites were inserted adjacent to *flxA* in strain KC01 which also has a chromosomal LacI::GFP fusion. Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and induced cultures had 0.3% arabinose or glucose added for 1 minute before slides were prepared. Dividing cells were analysed and the percentage of dividing cells which had one focus between the 2 cells was calculated for each strain in each condition. A star above a bar represents data that is significantly different from unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS (P > 0.05) represents data that is not significantly different. n = > 1000 cells

Cultures were grown as described previously with 50 μM rifampicin added for the final 15 minutes of growth (Grainger *et al.*, 2005) and 0.3% arabinose added for 1 minute before slides were prepared. Dividing cells were analysed and the percentage of these that had a single focus in the centre of the two new cells were counted.

When cultures had been treated with rifampicin, neither strain showed significant difference ($P > 0.05$) in the percentage of dividing cells with a single focus in response to arabinose (see figure 4.24). Strain LR38-pLER108, tagged adjacent to *araFGH*, had around 40% of dividing cells with a single focus and strain SXB4, tagged adjacent to *dps*, had around 10%. Rifampicin causes RNAP to become trapped at promoters. Since rifampicin was added to the cells 15 minutes before they were induced with arabinose it is likely that the majority of RNAP in the cells was already immobilised and therefore unable to bind the *araFGH* promoter. The effect seen in section 4.7.1, of arabinose inducing chromatid segregation at *araFGH*, was lost in the presence of rifampicin. This suggests that the arabinose induced segregation is due to active transcription, presumably at the *araFGH* promoter, not any other effect of arabinose addition.

4.7.4 Evidence that the segregation of sister chromatids at the position of *araFGH* upon induction by arabinose is dependent on AraC

To investigate whether the increased segregation of sister chromatids seen in cells tagged adjacent to *araFGH* is due to promoter activation by AraC, a strain with an *araC* knock out was constructed. Strain LR38 carries 20 MalI DNA sites inserted adjacent to *araFGH*. LR44 is a Δ*araC* derivative of LR38. Both of these strains were then transformed with plasmid pLER108 encoding MalI::mCherry. Strains containing the plasmid will be referred to as

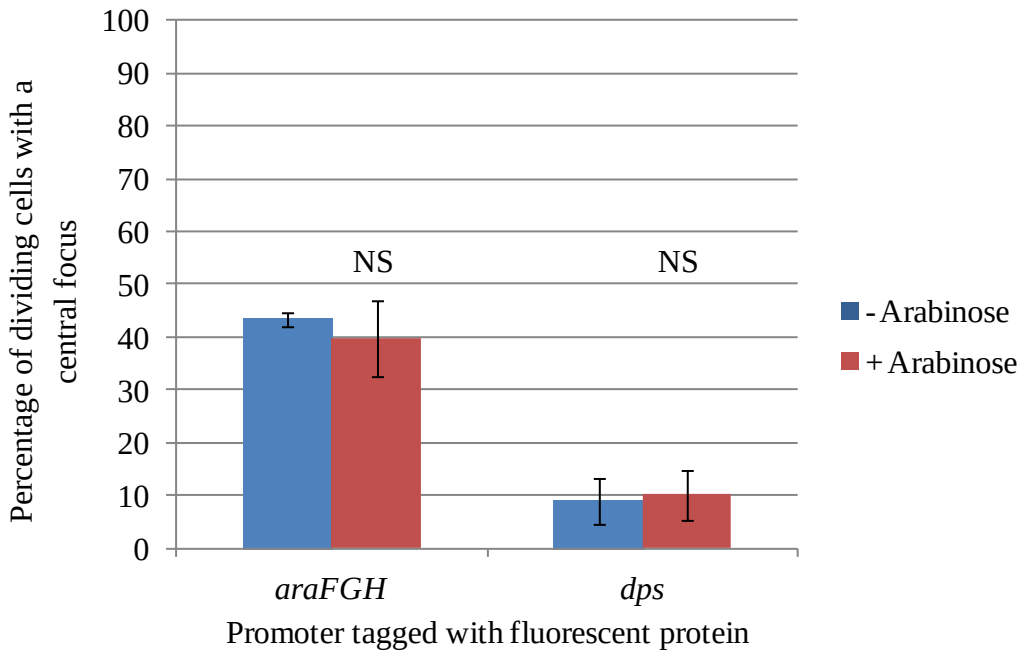


Figure 4.24: The effect of rifampicin on the stimulation of the separation of sister chromatids adjacent to *araFGH* by arabinose

Bar chart showing the percentage of cells at the point of division with a single central focus.

Strains tagged adjacent to *araFGH* (LR38-pLER108) and *dps* (SXB4) were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and rifampicin was added to a final concentration of 50 μM. Induced cultures had 0.3% arabinose added for 1 minute before slides were prepared. Dividing cells were analysed and the percentage of cells at the point of division which had one focus between the 2 cells was calculated for each strain in each condition. A star above a bar represents data that is significantly different from unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 1000 cells

LR38-pLER108 and LR44-pLER108. Cultures were grown as described previously with 0.3% arabinose added for 1 minute before slides were prepared. Cells at the point of division were analysed and the percentage of these that had a single focus in the centre of the two new cells were counted.

Strain LR38-pLER108, with AraC, had a significant reduction ($P < 0.01$) in the percentage of cells at the point of division with a single central focus in cultures induced with arabinose, as previously seen (see figure 4.25). Strain LR44-pLER108, with a chromosomal AraC deletion, the addition of arabinose caused no significant difference ($P > 0.05$) in the percentage of cells at the point of division with a central focus. However, the percentage of cells with a central focus in uninduced cultures was significantly less in the absence of AraC ($P < 0.01$). This could suggest that AraC is needed for the normal chromosome structure, and when AraC has been deleted, cells are in a state more similar to that of arabinose induced wildtype cells. Although AraC activates the *araFGH* promoter in the presence of arabinose, it also binds in the absence of arabinose (Hendrickson *et al.*, 1990). The presence of arabinose changes the function of AraC, instead of causing binding. However, the separation of sister chromatids adjacent to *araFGH* in the presence of arabinose is not seen in a Δ*araC*, suggesting this effect is dependent on AraC.

4.7.5 The effect of inducing a neighbouring promoter on the separation of sister chromatids adjacent to *araFGH*

To investigate whether inducing another promoter in the Ter macrodomain can also stimulate the separation of sister chromatids adjacent to *araFGH*, we searched for an inducible promoter close to *araFGH*. Unfortunately, many of the genes surrounding *araFGH* are of unknown function or are not easily inducible. The closest suitable promoter found was *mntP*,

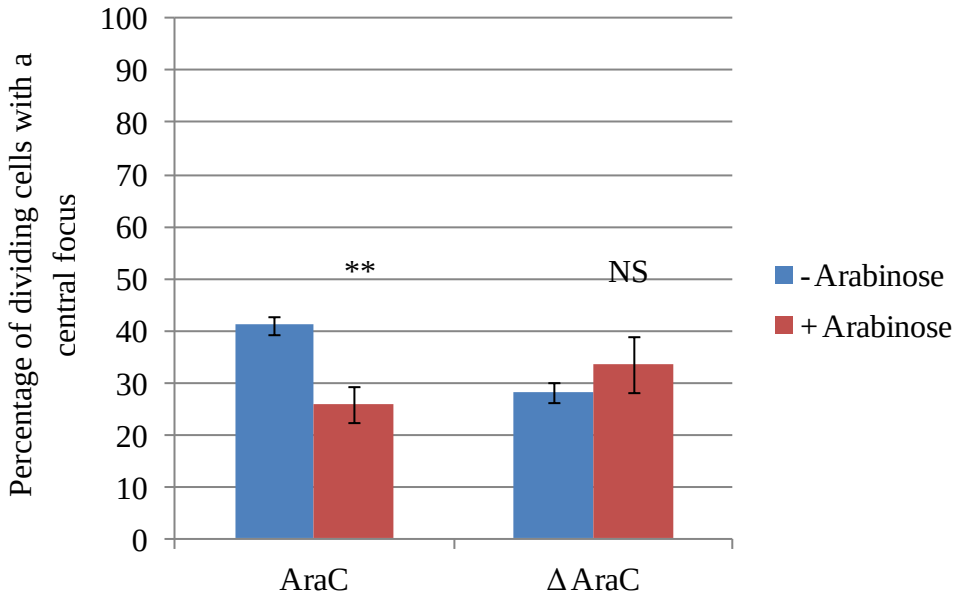


Figure 4.25: The effect of AraC on the stimulation of the separation of sister chromatids adjacent to *araFGH* by arabinose

Bar chart showing the percentage of cells at the point of division with a single central focus.

Strains tagged adjacent to *araFGH* either with *araC* present on the chromosome (LR38-pLER108) or with a mutated *araC* (LR44-108). Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 μg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1. Induced cultures had 0.3% arabinose added for 1 minute before slides were prepared. Cells at the point of division were analysed and the percentage of dividing cells which had one focus between the 2 cells was calculated for each strain in each condition. A star above a bar represents data that is significantly different from unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 1000 cells

around 80 kb from *araFGH*, towards the dif. *mntP* is inducible with 10 μM $MnCl_2$ (Waters *et al.*, 2011). The effect of Mn^{2+} addition on foci in cells at the point of division was observed in strains LR38-pLER108 and SXB4, tagged adjacent to *araFGH* and *dps* respectively. Cultures were grown as previously described, with 0.1 μM $MnCl_2$ added for the final hour of growth. Slides were prepared as previously described. Dividing cells were analysed and the percentage of these that had a single focus in the centre of the two new cells were counted. There was no significant difference for either strain ($P > 0.05$) in the percentage of dividing cells with a single central focus in cultures grown with or without Mn^{2+} (see figure 4.26). This means that either induction of the *mntP* promoter does not cause sister chromatids to segregate, or if induction of the *mntP* promoter does cause segregation of sister chromatids, the effect is localised to a small area around *mntP*.

4.7.6 Comparison of the position of *araBAD* and *araFGH* in cells at the point of division with segregated chromosomes, with and without induction by arabinose

Section 4.5 and 4.6 concluded that, in single cells, there was no colocalisation between *araBAD* and *araFGH* and there was no significant difference in the distance between these two loci in the presence and absence of arabinose (see figures 4.19b and 4.20b). Strain LR48-pLER108 was used for this experiment, with a LacI DNA site insert adjacent to *araBAD*, MalI DNA site insert adjacent to *araFGH*, chromosomal LacI::GFP fusion and a plasmid encoded MalI::mCherry fusion. As studies of cells at the point of division have shown that the addition of arabinose initiates the segregation of sister chromatids adjacent to *araFGH*, this strain was used to see whether, after segregation, *araFGH* moves to a position closer to *araBAD*. Cells at the point of division, around 10% of total cells, were analysed from cultures of LR48-pLER108, either unsupplemented or supplemented with 0.3% arabinose 1 minute before slides were prepared. Data was included from cells at the point of division when each

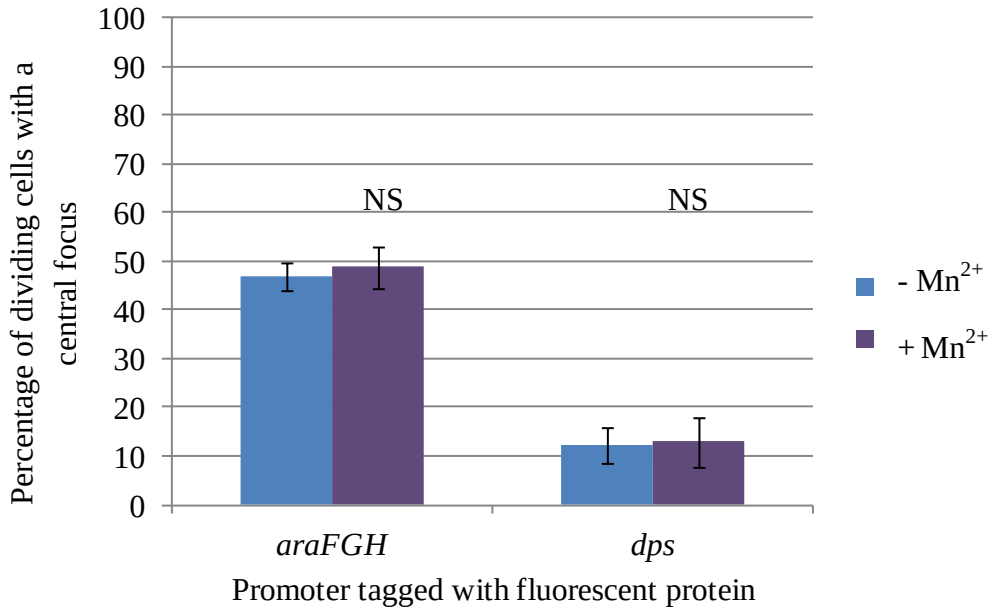


Figure 4.26: The effect of inducing a neighbouring promoter on the separation of sister chromatids adjacent to *araFGH*

Bar chart showing the percentage of cells at the point of division with a single central focus.

Strains tagged adjacent to *araFGH* (LR38-pLER108) and *dps* (SXB4) were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and 0.1 µM $MnCl_2$ was added for the final 10 minutes of growth. Cells at the point of division were analysed and the percentage of dividing cells which had one focus between the 2 cells was calculated for each strain in each condition. A star above a bar represents data that is significantly different from unsupplemented cultures (* = 0.05, ** = 0.01, *** = 0.001), NS ($P > 0.05$) represents data that is not significantly different. n = > 1000 cells

daughter cell had a single focus of GFP and mCherry, to study colocalisation and distance between the two locations. The distance from the centre of the GFP focus to the closest cell pole was measured in μm, and then the distance from the centre of the mCherry focus to the same pole, for both daughter cells. These distances were represented relative to cell length and plotted on a scatter graph, with GFP focus position on the x axis and mCherry focus position on the y axis. If the two points tagged did colocalise, the points on the scatter plot would cluster along the diagonal line. This does not happen for *araBAD* and *araFGH* in cells at the point of division (see figure 4.27a). There is also no difference in the positions of foci in the absence and presence of arabinose, suggesting that when sister chromatids segregate at *araFGH*, the *araFGH* gene moves to a position similar to that which it would occupy when it segregated by normal cell mechanisms.

The distance between *araBAD* and *araFGH* was then calculated. The percentage of cells with foci separated by each 0.1 μm was then calculated for cells that had been grown with and without arabinose (see figure 4.27b). There was no significant difference between the foci representing the positions of the *araBAD* and *araFGH* promoters ($P > 0.05$) suggesting the araFGH sister chromatids do not move to a location close to the AraC controlled araBAD promoter. However, the distance between these two promoters in dividing cells was significantly less than that seen in single cells (see section 4.6.2, figure 4.19b) ($P < 0.01$) for both induced and uninduced cultures, with averages falling from 0.66 μm and 0.63 μm to 0.48 μm and 0.43 μm. This suggests that the chromosome structure in dividing cells may be different to that in single cells. As each daughter cell is likely to be much smaller than a single cell, the distances between various points on the chromosome will also be shorter as the

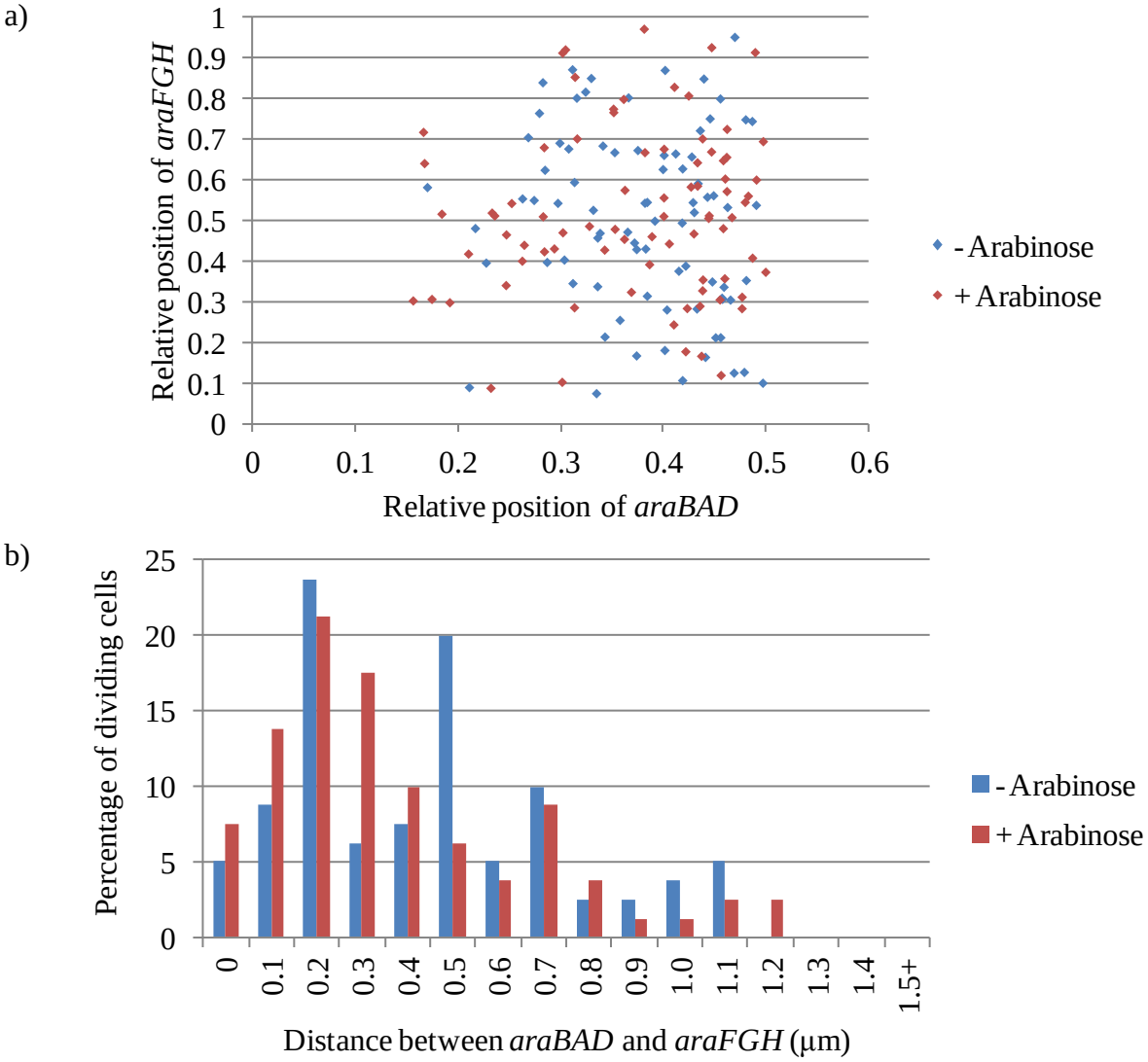


Figure 4.27: Distance between *araBAD* and *araFGH* in cells at the point of division, grown with and without arabinose

a) Scatter plot comparing the positions of two loci on the chromosome, tagged with LacI::GFP (x axis) and MalI::mCherry (y axis).

b) Bar chart showing the distance in µm between LacI::GFP and MalI::mCherry foci.

22 LacI DNA sites were inserted adjacent to *araBAD* and 20 MalI binding sites were inserted adjacent to *araFGH* in strain LR48-pLER108, which also had a chromosomal LacI::GFP fusion and was transformed with a plasmid encoding MalI::mCherry, pLER108.Cells were grown in M9 minimal media supplemented with 0.3% fructose and 17.5 µg/ml chloramphenicol if necessary. Cells were grown to an OD_{650} of approximately 0.1 and 0.1 µM $MnCl_2$ was added for the final 10 minutes of growth. Dividing cells were analysed. The distance between GFP and mCherry foci were measured in µm. n = > 40

nucleoid itself will be smaller. A comparison of relative positions may show there is no difference between single cells and cells at the point of division.

4.8 Discussion

4.8.1 Positions of inducible promoters

It has been shown that GalR tetramers binding at distant DNA sites cause clustering of GalR-regulated genes in the cell (Qian *et al.*, 2012). This chapter aimed to use FROS to study two other transcription factors, AraC and MntR, to see whether their binding affected the cellular position of promoters they control and if this caused clustering. This was done by first tagging locations adjacent to specific promoters, and measuring their positions with and without induction. Then, two locations were tagged in the same cell to give information about their positions relative to each other, and the distance between them. Finally, the same strains were used to study the effects of transcription on chromosome segregation in dividing cells.

No significant differences were seen in strains tagged at a location adjacent to two MntR-regulated genes, *mntH* and *dps*, when the number of foci per cell or the positions of foci were analysed in the presence and absence of Mn^{2+}. In the presence of Mn^{2+}, MntR will bind to both of these genes and repress transcription (Yamamoto *et al.*, 2011). The positions of promoters *araBAD*, *araJ* and *araFGH* were also studied individually, in the presence of arabinose or glucose or in the absence of any supplement. In the presence of arabinose, AraC should bind all three promoters and activate transcription. In the presence of glucose, AraC is known to bind adjacent to, and actively repress, the *araBAD* and *araC* promoters. There was no significant difference seen in strains tagged adjacent to *araBAD* or *araJ* when the number of foci per cell or the positions of foci were analysed in the different conditions.

A strain tagged adjacent to *araFGH* did not show a significant difference in the number of foci per cell, but some differences in the positions of foci were seen in the different conditions. For cells with a single focus, the average position of the focus was shown to be slightly closer to the centre of the cell, on average, when cultures had been grown in the presence of arabinose. A difference was also seen in cells with two foci. Due to the position of the *araFGH* gene relative to *oriC*, there were less cells containing two foci than most other strains, around 10% of all cells. Despite this lower number of cells analysed, the data shown is the result of around 70 cells for each condition, and is statistically significant. In the presence of arabinose, both foci representing the position of the *araFGH* promoter tended to be located close to the centre of the cell. There were very few foci in the fifth of the cell closest to either cell pole in any cells analysed. AraC is thought to bind to the *araFGH* promoter in both the absence and presence of arabinose (Hendrickson *et al.*, 1990), suggesting that binding of arabinose by AraC bound to the *araFGH* promoter leads to other changes.

One possible explanation as to why cells tagged adjacent to *araFGH* are the only ones to show differences in the position of foci, is that the products of the *araFGH* operon are membrane proteins. It has long been known that in bacteria, due to the lack of nuclear membrane, the processes of transcription and translation can be linked in time and space (Miller *et al.*, 1970). In the case of membrane proteins, it is thought that transcription and translation can also occur at the same time as the insertion of the protein into the membrane in a process known as transertion, preventing the hydrophobic membrane proteins from being exposed to the unfavourable conditions of the cytoplasm (Norris, 1995). This involves the gene encoding the membrane protein moving towards the membrane, along with RNAP, ribosomes and any other factors needed. A previous study has used FROS to visualise the

movement of *lacY* towards the membrane upon induction (Libby *et al.*, 2012). Although this study has identified a movement in the *araFGH* operon upon induction as a shift away from the cell pole, it may be that the movement is also towards the side of the cell, towards the membrane. As the size of the focus was very large compared to cell width and measurements may not have been accurate, the position of foci in the width of the cell was not measured and so this has not been seen. It is possible that there is a preferred position on the membrane for transertion, or *araFGH* specific transertion, that is slightly closer to the centre of the cell than the *araFGH* gene is usually found. From this position, the AraFGH transporter could diffuse around the membrane to be more evenly distributed. This explanation also accounts for the change seen in the distribution of foci in cells with two foci. The addition of arabinose in this case also leads to a movement of foci away from the pole of the cell. A model has been suggested where, when there are two copies of a gene in the cell after replication, one copy will be expressed while any others are not (Norris and Madsen, 1995). In the case of *araFGH*, the two foci are arranged symmetrically in the presence of arabinose, not suggesting that one copy of the gene is being expressed and the other silenced. FROS could be used with RNA FISH to show the locations of both the *araFGH* promoter and mRNA, which might answer this question. To confirm the movement of the *araFGH* locus towards the cell membrane, measurements could be taken of the position of the focus relative to cell width. This method has been used to show that expression of *lacY*, encoding a membrane protein, caused the locus to be repositioned closer to the membrane (Libby *et al.*, 2012). However, there are problems associated with this method. Because a bacterial cell is 3D and the microscope produces a 2D image, foci that are adjacent to the membrane on the top or bottom of the cell, as it lies on the slide, will appear to be in the centre of the cell. Also, as cell width is much less than cell length in *E. coli*, the focus produced by FROS is quite large compared to cell width. This could make it difficult to get accurate measurements. A super resolution microscope approach

may give more reliable data. Overall, this result suggests that inducing transcription, and therefore transertion, at *araFGH* causes some changes in the local chromosome structure. Although *mntH* also encodes a membrane protein this effect is not seen in a strain tagged adjacent to *mntH*. This could be for a variety of reasons, for example, transertion may not happen in the cases of all membrane proteins. Also, if there are specific transertion sites, the "home" position of *mntH* may be close to this position, meaning little rearrangement is needed after induction, and the movement was too small to detect by FROS.

The average positions of foci in cells with 2 foci correlate with the information in the literature about segregation of sister chromatids. It is known that, once the majority of the chromosome is replicated, each sister chromatid is segregated to one half of the cell, with the left and right replichores alternating along cell length (see section 1.3) (Nielsen *et al.*, 2006, Reyes-Lamothe *et al.*, 2008b). Of the five locations tagged for experiments in this chapter, *araBAD*, *araJ* and *dps* are found in the right replichore, with *araFGH* and *mntH* located in the left replichore. When the positions of foci in cells with 2 foci was analysed, the position of the focus closest to a cell pole was analysed relative to the closest focus. For this reason, strains with tags at different locations cannot be directly compared, as they may have been orientated using different cell poles. To resolve this issue, the experiment would need to be repeated with a label on either the old or new pole of the cell. However, if it is assumed that strains with tags in the left and right replichores are orientated to opposite poles of the cell, the positions of foci compares to that predicted in other studies. All of the positions are a fraction of cell length, represented as 1, so the average positions of foci in 2 foci cells tagged in the right replichore were subtracted from 1. This gave an average position corrected for the replichore. The average positions of foci from cells with 2 foci, tagged at different positions,

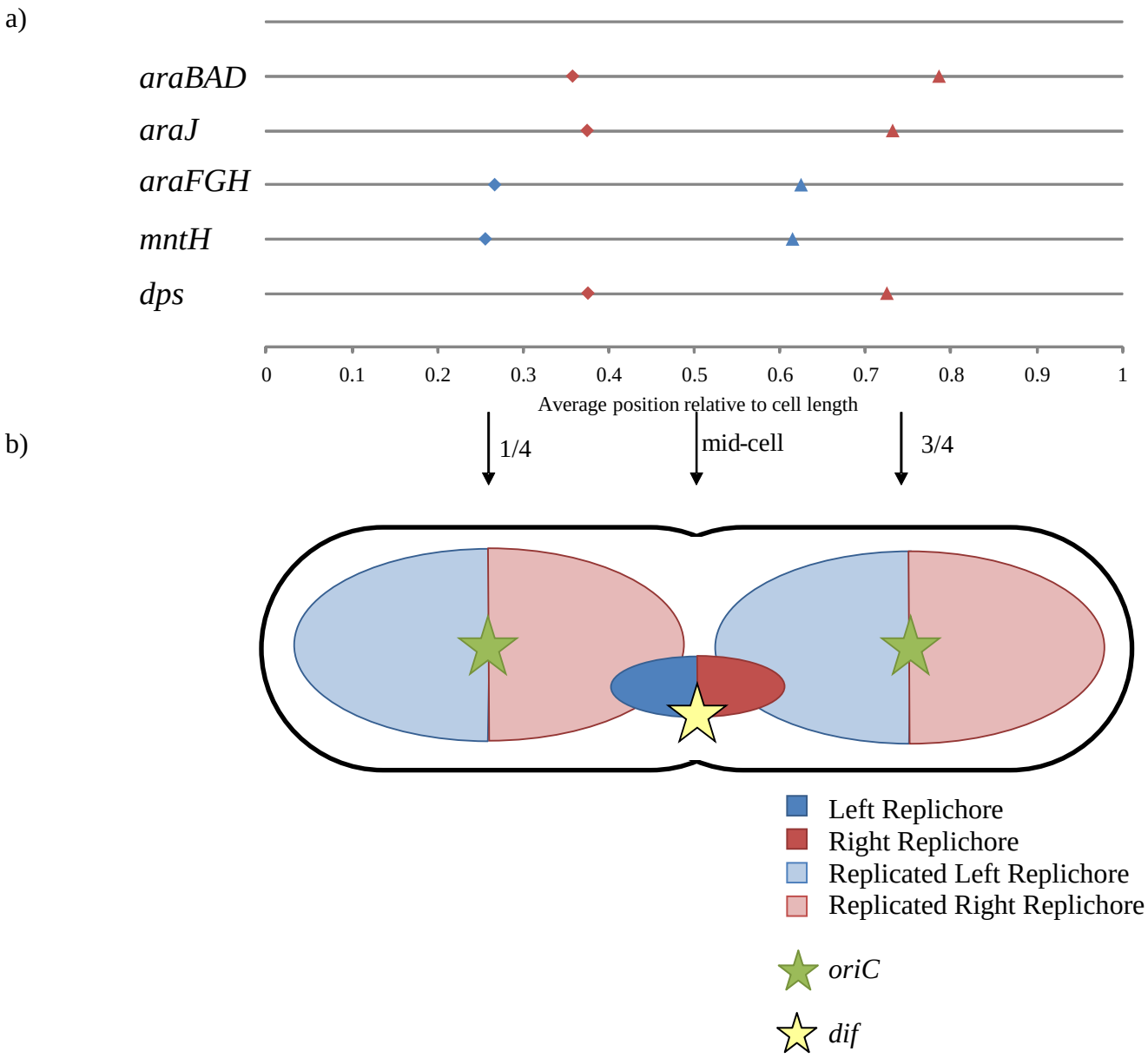


Figure 4.28: Positions of foci representing locations in the left and right replichores

a) Positions of foci in cells containing 2 foci in strains tagged adjacent to *araBAD*, *araJ*, *araFGH*, *mntH* and *dps*. Locations tagged in the Left replichore are show in blue, locations in the right replichore are shown in red. The focus closest to a cell pole (1st of 2 foci) is represented as a diamond, the 2nd of 2 foci is represented as a triangle.

b) A model of a bacterial cell, with the majority of the chromosome replicated, showing the segregation of the left and right replichores into the future daughter cells. The positions of mid-cell and one quarter and three quarters of cell length are shown with arrows. Figure adapted from Nielsen *et al*. 2006.

are shown in figure 4.28a. It can be seen that foci representing locations in the left replichore, blue markers, are found in similar areas of the cell, as are locations in the right replichore. It was expected that the border between the left and right replichores in each copy of the chromosome would be located at approximately the one quarter and three quarter positions. Figure 4.28b shows an estimation of the layout of the chromosome after the majority has been replicated. The average positions of the promoters studied can be compared to this model and it is seen that they are approximately in the positions expected. Some of the promoters have an average position that is slightly closer to mid-cell than expected. This is probably due to the fact that the cells were not synchronised so some of the cells would not have fully segregated chromatids, causing a bias in the average.

The movement, upon induction, of foci representing the position of the *araFGH* promoter in cells with 2 foci goes against this unsymmetrical model. In uninduced cells, the majority of foci in cells with 2 foci are arranged asymmetrically. Upon addition of arabinose, there is a shift towards the majority of cells being more symmetrical, with both foci close to mid-cell (figure 4.12). This is a movement away from the natural arrangement of the chromosome that is brought about by induction of gene expression. Although the results from this chapter cannot explain this movement, it is a clear indication that gene expression can affect chromosome structure.

4.8.2 Clustering of promoters from the same regulon

To investigate whether promoters controlled by MntR and AraC cluster in the cell, like those controlled by GalR, strains were created with two FROS tags labelling different promoters. This allowed the colocalisation of two locations to be studied, as well as measuring the

distance between the two points. When locations adjacent to the *mntH* and *dps* promoters were tagged in the same cell there did not seem to be any colocalisation. A second strain compared the position of *mntH* to *araBAD*, not known to be affected by the addition of Mn^{2+}. *mntH* seemed to colocalise more with the MntR-regulated *dps* than *araBAD* (figure 4.14). This may be due to the fact that *araBAD* is located in the non-structured right (NSR) domain, and so has more range of movement than *dps*, which is located in the right domain. The addition of Mn^{2+} to the cultures did not cause any significant difference in the distribution of foci, suggesting that MntR binding does not cause MntR-regulated promoters to cluster.

There were also no significant differences in the distances between *mntH* and either *dps* or *araBAD*, in the presence and absence of Mn^{2+}. *araBAD* can be used as a reference point in this case as it is assumed not to be affected by Mn^{2+}, and potentially show small movements that could not be detected when measuring focus position compared to cell length. However, there were no significant differences. Again, the data suggested that *mntH* was likely to be found closer to *dps* than *araBAD*, as the graph representing distances between *mntH* and *dps* was shifted slightly to the left, compared to the *mntH* and *araBAD* graph. This could also be due to the increased flexibility of the NSR macrodomain, allowing *araBAD* to be found in a larger area of the cell. This experiment did not find any suggestion that binding by MntR causes distant DNA sites on the chromosome to cluster, unlike GalR binding. This may be because, unlike GalR, MntR is unable to tetramerise (Tanaka *et al.*, 2009). MntR dimers would need another mechanism of interacting with each other in order to confine distant DNA sites to the same area of the cell, but it seems this is not the case.

Pairs of AraC controlled promoters were also tagged in the same cell using LacI and MalI FROS. Tags were inserted adjacent to the *araBAD* promoter in strains also tagged adjacent to the *araJ* or *araFGH* promoters. There was no colocalisation between the *araBAD* and *araFGH* promoters, with the points spread evenly across the graph, with no significant difference between induced and uninduced cultures. These promoters are located in the NSR and Ter domains respectively, and are almost opposite each other on a circular chromosome map. These locations would not be expected to interact, unless AraC binding brought them together. The result for these two promoters suggests that AraC is not capable of bringing distant DNA sites together in the cell. *araBAD* and *araJ* were shown to be more likely to colocalise, with points on the scatter graph in a much tighter group. There is no significant difference between cultures grown with and without arabinose, suggesting that the reason for *araBAD* and *araJ* being located in similar areas of the cell does not involve AraC binding. Although AraC binds at *araBAD* in the presence and absence of arabinose, this is not thought to be the case at *araJ*. Instead it is more likely that *araBAD* and *araJ* colocalise because they are located in the same macrodomain, NSR. Sites in this macrodomain are not as restricted as other parts of the chromosome, and are more able to interact with sites in other macrodomains, there is still a region of the cell where sites within the NSR are largely found, although it is larger than areas for the more structured macrodomains. As a result, *araBAD* and *araJ* will often be located in similar areas of the cell due to macrodomain organisation. When the distance between foci representing the positions of *araBAD* and either *araFGH* or *araJ* were measured in the presence and absence of arabinose there was no significant difference, suggesting AraC binding does not alter the distance between the two sites. Again, *araBAD* and *araJ* were shown to often be found close together, with a peak in the graph, in both conditions, of around 0.1 μm gap. In contrast, the graph showing the distance *araBAD* and *araFGH* does not seem to have a peak, with foci almost as likely to be found 0.1 μm and

nearly 1 μm apart. This suggests there are no specific factors regulating the position of *araFGH* relative to *araBAD*, only the natural variations in the positions of the macrodomains. AraC does not have the property, shown by GalR, of colocalising distant DNA sites. This is probably because AraC cannot tetramerise (Soisson *et al.*, 1997), which has shown to be essential for colocalisation of GalR DNA sites (Qian *et al.*, 2012).

The position of *araBAD* in the presence and absence of arabinose was compared to the position of *mntH*, not controlled by AraC, as a reference point. Unexpectedly, when the colocalisation of foci and distance between foci were measured, it was seen that the position of *mntH* was significantly different when grown with arabinose. In the presence of arabinose, *mntH* was seen to move towards the centre of the cell and away from *araBAD*. As *mntH* was not thought to have any connection to arabinose or AraC, a search was completed of the genes in the surrounding area. An operon consisting of genes *ypdAB* was found within 13 kb of the tag adjacent to *mntH*. *ypdA* is predicted to encode an inner membrane sensor kinase and *ypdB* is predicted to encode a response regulator. A recent study has identified one gene, *yhiX* that is thought to be the sole target of *ypdB*. It was thought that this system was involved in carbon control so a carbon source screen was carried out. The expression of *yhiX* was induced in a few specific conditions, including in the presence of pyruvate (Fried *et al.*, 2013). However, arabinose was not included in this screen and could have some affect on the expression of *ypdAB*. As *ypdAB* encodes proteins linked to the membrane, it is possible that transertion has caused this movement. The region surrounding *mntH* includes several genes of unknown function making it difficult to rule out one of these genes being responsible for the change in chromosome structure in the presence of arabinose, not *mntH*.

4.8.3 Chromosome segregation

Cells at the point of division were studied to investigate whether transcription had any effect on chromosome segregation. It is known that when the chromosome is replicated, the sister chromatids stay together in the region where they were replicated for a short time, before being segregated and moving to opposite halves of the cell. The time before the sister chromatids segregate varies between macrodomains. The non-structured macrodomains segregate quicker than the other macrodomains. The Ter macrodomain has been shown to have the longest time sister chromatids segregate with estimates ranging from 15 to 60 minutes and higher (Possoz *et al.*, 2012). When cells at the point of dividing, tagged adjacent to *araBAD, araFGH* or *dps,* were studied it was seen that some pairs of cells had a single focus located near the middle of the two cells. It is likely that this single focus is actually an unsegregated pair of sister chromatids, too close together for individual foci to be distinguished by microscopy. As expected, these cells were very rarely seen in the strain tagged adjacent to *araBAD*, located in NSR, relatively close to *oriC*. This location will be replicated early in the cell cycle and the sister chromatids will be segregated before the cell begins to divide. Around 8% of cells at the point of division tagged adjacent to *dps* had a single central focus. Since *dps* is located in the right macrodomain, this small number is also expected, as most sister chromatids at this point will have separated before the cells starts to divide. Cells tagged adjacent to *araFGH* had the highest percentage of cells at the point of division with single foci, with 35% of cells. *araFGH* is in the Ter macrodomain, making it more likely that, at the point where the cell begins to divide, this location will either not have been replicated or still have cohesive sister chromatids. This number may be lower than expected for a location in the Ter macrodomain, possibly because *araFGH* is located at the very edge of the macrodomain. The structure of the Ter macrodomain is organised by MatP binding to sites known as *matS*¸ of which there are 21 within the Ter. *araFGH* is located 40

kb outside the last *matS* site (Mercier *et al.*, 2008), possibly meaning the structural properties of the Ter macrodomain will not be seen as clearly at this point.

Cells at the point of division were then studied under conditions that were inducing, supplemented with arabinose, or repressing, supplemented with glucose. No significant difference was seen in cells at the point of division tagged adjacent to *araBAD* or *dps*. However, in cells tagged adjacent to *araFGH*, there was a significant reduction in the percentage of cells at the point of division with a single focus in the presence of arabinose. This effect was not seen in cells grown with glucose. This result suggest that the induction of the *araFGH* promoter stimulates the segregation of sister chromatids at this point. This could be because DNA in unsegregated sister chromatids is not very accessible for transcription machinery, so it moves to an area of the cell where transcription can occur. Alternatively, it could be due to the fact that the proteins encoded by *araFGH* are membrane proteins, and therefore may undergo the combined processes of transcription, translation and insertion in to the membrane, known as transertion. This involves proximity to cell membrane, and the location of the sister chromatids is in the centre of two cells preparing to divide, so much of the membrane at this point is likely to be covered by FtsK and other proteins involved in cell division. *araFGH* may need to move into the two future daughter cells for transertion, to find a suitable location to insert protein into the membrane.

This effect was shown to be specific to the region of the chromosome adjacent to *araFGH*, as a strain with an insert in the Ter around 300 kb from *araFGH* was unaffected by the addition of arabinose. This suggests that the addition of arabinose does not affect the structure of the Ter domain for reasons unrelated to *araFGH* induction. It also shows that the separation of

the sister chromatids at *araFGH* does not lead to the segregation of the whole Ter domain, probably just the region around *araFGH*. A promoter 80 kb from *araFGH*, *mntP*, was induced to investigate whether this would also lead to separation of sister chromatids at *araFGH*. The addition of Mn^{2+} to induce *mntP* did not have any effect on segregation of sister chromatids adjacent to *araFGH*. *mntP* encodes a membrane protein so transcription and translation may be combined with inserting the protein into the membrane, but this is not known. If transertion is the driving force for chromatid segregation at *araFGH* and the same process is occurring at *mntP*, a loop of DNA must be separating, leaving the DNA on either side undisturbed. Alternatively, *mntP* may be transcribed separately to insertion of the protein into the membrane, not altering the structure of the chromosome. If *mntP* does undergo transertion, it may be that a small loop of DNA, including the *mntP* gene, moves away from the rest of the Ter macrodomain. Without further experiments, it is unknown whether the segregation of sister chromatids by transcription is a disruptive process, or if it could even just be a temporary effect for the time of transcription.

When cells tagged adjacent to *araFGH* were grown in the presence of rifampicin, the effect of arabinose on sister chromatid segregation was not seen. Cells were grown with rifampicin before arabinose was added so, at the time of induction, most RNA polymerase in the cell would probably be stuck at other promoters, leaving no free RNA polymerase to transcribe *araFGH*. This proves the segregation of sister chromatids at *araFGH* upon addition of arabinose is dependent on transcription. Deletion of AraC from the strain tagged adjacent to *araFGH* also removed the effect of arabinose on chromatid segregation at *araFGH*. This shows that both RNA polymerase and AraC are required for any effect. Compared to uninduced cultures with AraC, the percentage of cells at the point of division with a central

focus was lower in the absence of AraC, both with and without arabinose. This result suggests AraC may play some role in chromosome organisation.

When the sister chromatids adjacent to the induced *araFGH* gene move, they are moving away from the "home" position of the Ter domain. It was investigated whether, in the absence of a "home" position, *araFGH* moved to a location closer to another AraC controlled gene, *araBAD*. A strain tagged using different FROS reporters at *araBAD* and *araFGH* was used to measure the distances between the two genes in cells at the point of division, in the presence and absence of arabinose. This showed there was no significant difference between the two conditions suggesting that in the presence of arabinose, *araFGH* moves to a similar position to that which it moves to when the sister chromatids separate in uninduced cells. However, the distances between the two genes in cells at the point of division are different to that seen in single cells. This suggests that cells at the point of division may have a different arrangement of macrodomains to that in single cells.

Overall, it has been shown that, unlike GalR, MntR and AraC do not cause DNA sites across the chromosome to colocalise. GalR dimers, bound to distant DNA sites, can interact and form tetramers, which causes the DNA sites to colocalise. Bound MntR and AraC dimers cannot interact, therefore cannot colocalise the promoters they regulate. There is some evidence that the AraC-regulated *araFGH* changes position when it is induced, probably to move closer to the cell membrane to allow the process of transertion to occur. Inducing *araFGH* in cells at the point of division causes an increase in the number of cells with segregated sister chromatids at this point, again, probably due to transertion.

5. Super resolution microscopy studies on Transcription Factors

Molecular constructs and bacterial strains used in this chapter were prepared by Laura Sellars (University of Birmingham). Single-molecule data was collected, analysed and interpreted by Federico Garza de Leon (University of Oxford).

5.1 Introduction

Transcription factors fused to fluorescent proteins have been used in FROS as a tool for tagging the chromosome. However, these labelled transcription factors could also be used to study the transcription factors themselves, to get information about diffusion and binding. Chapter 3 showed that the epifluorescent microscopy set up was not able to show any information about individual molecules of fluorescently tagged transcription factors as, in the absence of an array of DNA sites, transcription factors could not be observed. In order to visualise single transcription factor molecules, improved resolution is needed. Techniques have been developed that are capable of studying individual molecules, including atomic force microscopy (AFM), optical tweezers and patch clamping but none of these would be suitable for for studying individual transcription factors *in vivo*. There are also several super-resolution microscopy techniques which could be used for these experiments.

5.1.1 Super-resolution Microscopy

To get information about the exact position of individual molecules in the cell, improved resolution is needed. Light is diffracted, meaning that two objects closer together than half the wavelength of light cannot be distinguished as separate. Resolution of about 250 nm is the best that can be achieved by conventional light microscopy (Cattoni *et al.*, 2012). Optimal spatial resolution also requires a high signal to noise ratio. This is difficult to achieve in

biological samples and with molecules that may be moving. (Schermelleh *et al.*, 2010). The limited resolution of conventional light microscopy is a particular problem in the study of bacteria, where cells are often only 1-2 μm long, and resolution of 250 nm restricts the information that images can give. New microscopy techniques have been developed that have broken the diffraction limit and made it possible for single molecules to be studied inside the cell. These include derivatives of confocal microscopy, which use a pinhole to block light from points in the sample that are not at the focal point of the lens. These include 4Pi (Gugel *et al.*, 2004) and Stimulated Emisson Depletion (STED) microscopy (Hell, 2007). Although resolution is improved and a thin section of the sample is imaged very clearly, much light is blocked, reducing signal intensity. Other techniques employ the principles of Total Internal Reflection Fluorescence (TIRF) microscopy, which uses the principle of total internal reflection to excite only fluorophores within around 100 nm of the surface without exciting fluorophores further away, and therefore reducing background fluorescence. It is, however, limited to imaging surface structures. TIRF based techniques include Near-field Scanning Optical Microscopy (NSOM) and Photoactivated Localisation Microscopy (PALM) (Adelmann *et al.*, 1999). PALM uses a photoactivatable fluorophore, which can be photobleached, thereby isolating single molecules so the exact position can be calculated statistically. Fluorophores are activated and then photobleached so they do not interfere with subsequent rounds of photoactivation. (Single molecule tracking PALM will be referred to as PALM for the remainder of this thesis).

The technique used in this chapter is a PALM system developed by the Kapanidis lab (University of Oxford) which combines PALM and single molecule tracking, so molecules can be followed within a cell (Uphoff *et al.*, 2013). Near TIRF illumination was used, which

uses light at a angle close to that of TIRF but results in illumination deeper into the sample. This technique was chosen because it had previously been used to track DNA polymerase I, so would be suitable for tracking transcription factors. PALM is used in this chapter to study the diffusion of LacI, with the view to extending this technique to study other transcription factors in the future. PAmCherry is a fluorophore developed for PALM, which exists in a dark state until it is exposed to violet light, when it becomes red fluorescent (Subach *et al.*, 2009). Photoactivation is controlled so less than one fluorophore is activated at any time, per cell. Cells were imaged in 15 ms exposures until all activated fluorophores are photobleached (Uphoff *et al.*, 2013). Single molecule tracking (Manley *et al.*, 2008, Niu and Yu, 2008) was used with PALM to follow molecules and give information about their diffusion (Uphoff *et al.*, 2013) (see section 2.12).

5.1.2 Super-resolution microscopy studies on LacI

Two recent papers have used single molecule microscopy techniques to investigate the diffusion and distribution of LacI. Hammar *et al.*, 2012 investigated the theory that LacI searches for its DNA sites by a combination of 3D diffusion and 1D sliding using wide-field laser microscopy. It was discovered in 1970 that the rate at which LacI associated with its DNA sites was not limited by diffusion (Riggs *et al.*, 1970). The greater than expected rate of LacI association with its DNA sites was explained by the facilitated diffusion theory. This suggests that transcription factors could search for their DNA sites by combining 3D diffusion with non-specific binding to DNA, followed by a brief period of 1D sliding along the DNA (Richter and Eigen, 1974, Berg *et al.*, 1981). Although this theory had previously been investigated *in vitro* (Ruusala and Crothers, 1992), super resolution imaging provided the opportunity to investigate it *in vivo* (Hammar *et al.*, 2012). Strains were made with two copies

of the LacI DNA site inserted onto the *E. coli* chromosome and different spaces between them. If the space between the two sites is greater than the distance LacI slides after binding non-specific DNA, then the sites will be recognised separately. If the space is smaller, the sites may be recognised as one target due to the facilitated diffusion theory. This study found that LacI DNA sites spaced 45 or 25 bp apart were bound slower than a single LacI DNA site, or two sites further apart. This suggests that when the space between two LacI DNA sites is 45 bp or less they act like one site. LacI slides an average of 45 ± 10 bp, and this sliding makes the rate of finding a LacI binding site 40 times faster. More than 90% of the time LacI passes the DNA site without recognising it, indicating that although sliding is a quicker method of searching for DNA sites, some specificity is lost. Hammar *et al.,* also found that the presence of another protein bound adjacent to the LacI DNA site prevent LacI from approaching the DNA site by sliding from that direction. Although this study was able to investigate LacI at the single-molecule level, the technique used involved limiting the number of LacI molecules to three to five molecules per cell (Hammar *et al.*, 2012), a reduction from the usual 20 monomers per cell (Gilbert and Müller-Hill, 1966).

A second study the same year used super resolution imaging to study the distribution of LacI in *E. coli*. LacI was known to be able to diffuse at a rate of 0.4 $\mu m^2/s$ (Elf *et al.*, 2007), suggesting that LacI distribution should be homogenous throughout the cell. Kuhlman and Cox used TIRF microscopy to visualise the distribution of tagged LacI and *lacI-venus* mRNA compared to the position of the *lacI* gene, which was inserted at different positions on the *E. coli* chromosome. Strains had the native *lac* operon deleted so there were no LacI DNA sites on the chromosome. Surprisingly, LacI was found to colocalise with its encoding gene and mRNA, whether the gene was inserted at an origin-proximal or terminus-proximal location.

This localisation was not seen when the DNA binding domain was deleted. Condensation of DNA seemed to have an effect on the distribution of LacI when the DNA binding domain was deleted. The mutant LacI was excluded from the nucleoid in stationary phase cells, where the DNA is tightly packed, compared to its even distribution in slow growing cells. The full LacI protein localised with its gene in both cases. The authors suggest this means *E. coli* could change the distribution of transcription factors with growth phases, as DNA compaction changes (Kuhlman and Cox, 2012).

LacI is one of the best studied transcription factors in *E. coli*, and is part of a large family of transcription factors. This makes LacI a good candidate for further single-molecule studies, as any findings about LacI can probably be applied to other LacI/GalR family members. A system that works for LacI is also likely to be suitable for studying other transcription factors.

5.2 Super resolution microscopy of LacI::PAmCherry

For use in PALM experiments, LacI needed to be tagged with a photoactivatable fluorophore, PAmCherry. Strain LR35 is derived from MG1655 and carries a chromosomal LacI::PAmCherry fusion. Strain LR18 carries 6 LacI DNA sites adjacent to the *araBAD* promoter and a chromosomal LacI::GFP fusion. Strain LR37 is derived from LR18 and has a chromosomal LacI::PAmCherry fusion. In both LR35 and LR37, the 3 LacI DNA sites at the *lacZ* promoter have been removed and the LacI::PAmCherry fusion is under the control of the *lacI* promoter. Fixed cells had previously been imaged using PALM to define the apparent diffusion coefficient of bound LacI.

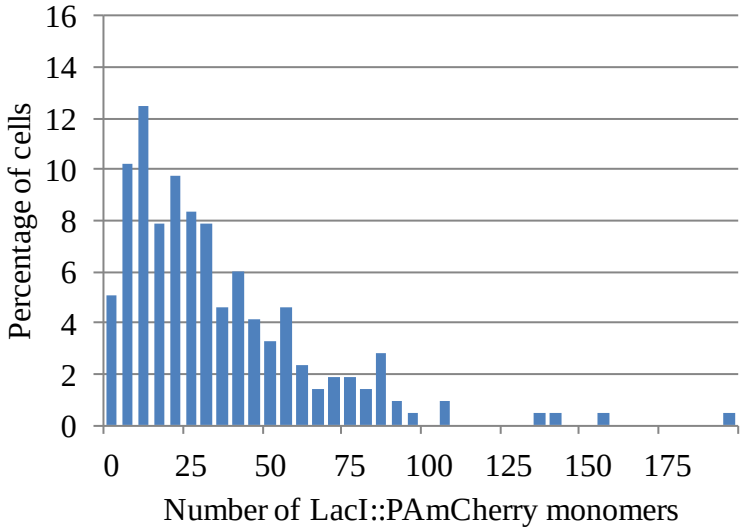
5.2.1 Number of LacI molecules per cell

The number of LacI molecules per cell averaged at 31.5 for strain LR37, carrying 6 LacI DNA sites, and 36.0 for strain LR35, with no LacI DNA sites (figure 5.1). There are thought to be around 20 monomers of LacI per gene copy (Gilbert and Müller-Hill, 1966), suggesting that many of the cells have multiple copies of *lacI*. In both strains more than three quarters of the cells had 50 or less monomers, and less than 3% had over 100 monomers. Although there seems to be more LacI::PAmCherry molecules present in the strain without any LacI DNA sites, LR35, this has been shown not to be significantly different to LR37. The average for strain LR35 may have been skewed by the presence of a few cells with very high numbers of LacI::PAmCherry molecules, but this is not the norm. Overall, this result suggests that the addition of a PAmCherry tag to LacI has not changed the copy number of the protein as it is found in similar numbers as previously reported.

5.2.2 Effect of LacI DNA sites on the apparent diffusion coefficient of LacI::PAmCherry

To investigate whether the presence of LacI DNA sites on the chromosome would change the diffusion of LacI::PAmCherry, strains LR35, with no DNA sites, and LR37, with 6 LacI DNA sites adjacent to *araBAD*, were compared. Molecules with an apparent diffusion coefficient of less than 0.1 $\mu m^2/s$ were defined as being "immobile" or bound to DNA. Molecules with a higher apparent diffusion coefficient are mobile and freely diffusing. In the strain with 6 LacI DNA sites, 36.6% of LacI molecules were in the immobile fraction, shown in the grey are of the graph, compared to 25% in the strain with no LacI DNA sites (see figure 5.2). The immobile molecules in the strain without any DNA sites are probably LacI molecules bound non-specifically to DNA. LacI has been shown to bind DNA and slide to search for its DNA sites (Hammar *et al.*, 2012) (see section 5.1.2). There are only 6 LacI DNA sites per copy of

a)



b)

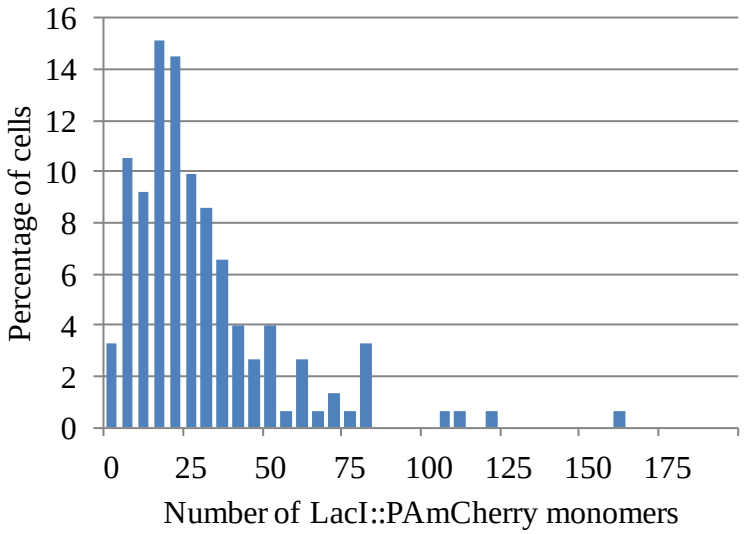


Figure 5.1: Number of LacI::PAmCherry molecules per cell

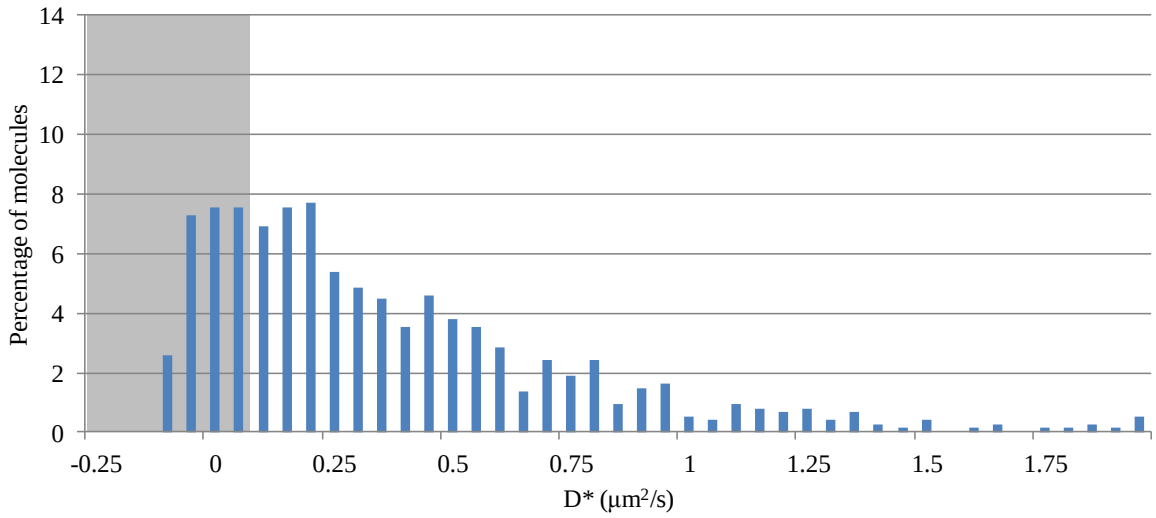
Bar chart to show the number of LacI::PAmCherry molecules per cell.

a) Strain LR35, with a chromosomal fusion of LacI::PAmCherry and all LacI DNA sites removed.

b) Strain LR37, with a chromosomal fusion of LacI::PAmCherry and 6 LacI DNA sites inserted adjacent to *araBAD*.

Cells were grown in M9 minimal media supplemented with 0.2% glycerol to an OD_{600} of approximately 0.1. Localisation analysis was performed using custom written MATLAB software. Microscopy and analysis was carried out by Federico Garza de Leon (University of Oxford). n = >150

a)



b)

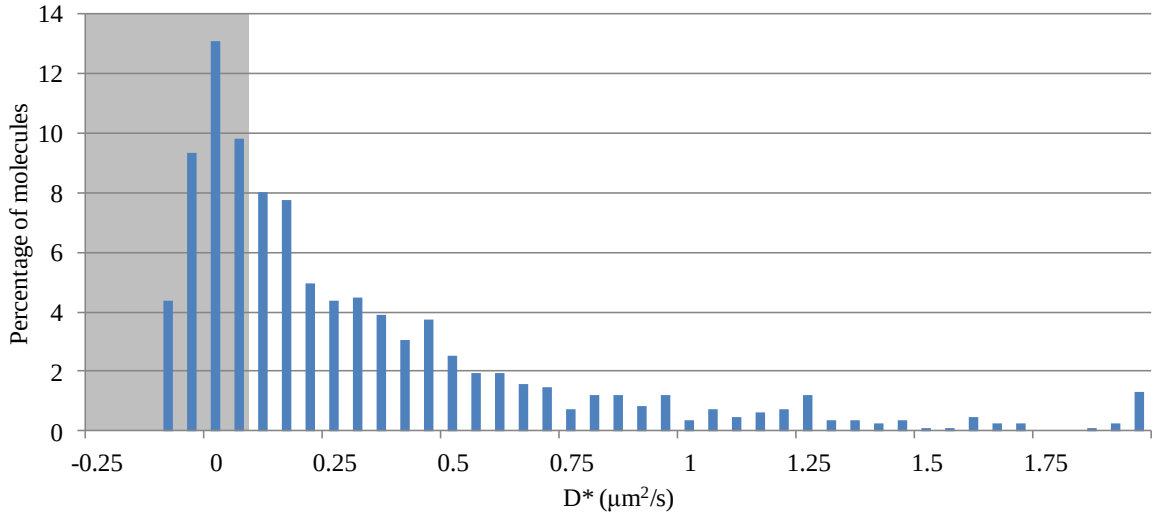


Figure 5.2: Effect of LacI DNA sites on the apparent diffusion coefficient of LacI::PAmCherry

Bar chart to show the apparent diffusion coefficient (D*) of LacI::PAmCherry. Bars in the grey section of the graph represent the immobile fraction of LacI::PAmCherry (D* < 0.1 $\mu m^2/s$).

a) Strain LR35, with a chromosomal fusion of LacI::PAmCherry and all LacI DNA sites removed.

b) Strain LR37, with a chromosomal fusion of LacI::PAmCherry and 6 LacI DNA sites inserted adjacent to *araBAD*.

Cells were grown in M9 minimal media supplemented with 0.2% glycerol to an OD_{600} of approximately 0.1. Localisation analysis was performed using custom written MATLAB software. Microscopy and analysis was carried out by Federico Garza de Leon (University of Oxford). n = > 150

the chromosome and an average of 31 monomers per cell. Each LacI DNA site is bound by one dimer, a maximum of 12 monomers per array, which means there will also be a proportion of LacI::PAmCherry that is not bound to DNA, as the DNA site array will be saturated. A strain with a higher number of DNA sites on the chromosome would probably have a higher percentage of immobile molecules and lower percentage of mobile.

5.3 The effect of IPTG on the apparent diffusion coefficient of LacI::PAmCherry

In the presence of IPTG, LacI should not bind to its DNA sites. To investigate the effect of IPTG on the apparent diffusion coefficient of LacI::PAmCherry, strains LR35, with no LacI DNA sites, and LR37, with 6 LacI DNA sites adjacent to *araBAD*, were imaged in the presence and absence of IPTG. Cells were immobilised in 1 μl wells and, to induce, washed with M9 minimal media supplemented with 1 mM IPTG.

Both of these strains carry a chromosomal LacI::PAmCherry fusion. When cells were induced with IPTG, strain LR35, with no LacI DNA sites, showed no change in the apparent diffusion coefficient of LacI::PAmCherry ($P > 0.05$) (figure 5.3a). This suggests that the bound fraction of LacI in these cells is bound non-specifically, in a way that is not affected by IPTG. Strain LR37 had a significant reduction in the fraction of LacI::PAmCherry molecules in the immobile state when cells were induced with IPTG ($P < 0.01$) (see figure 5.3b). The percentage of mobile molecules dropped from 36.7% to 19.8% when induced with IPTG. When the apparent diffusion coefficient of LacI::PAmCherry in strain LR37 was compared to that in LR35 there was found to be no significant difference ($P > 0.05$), suggesting that the addition of IPTG has eliminated all specific binding of LacI::PAmCherry to LacI DNA sites.

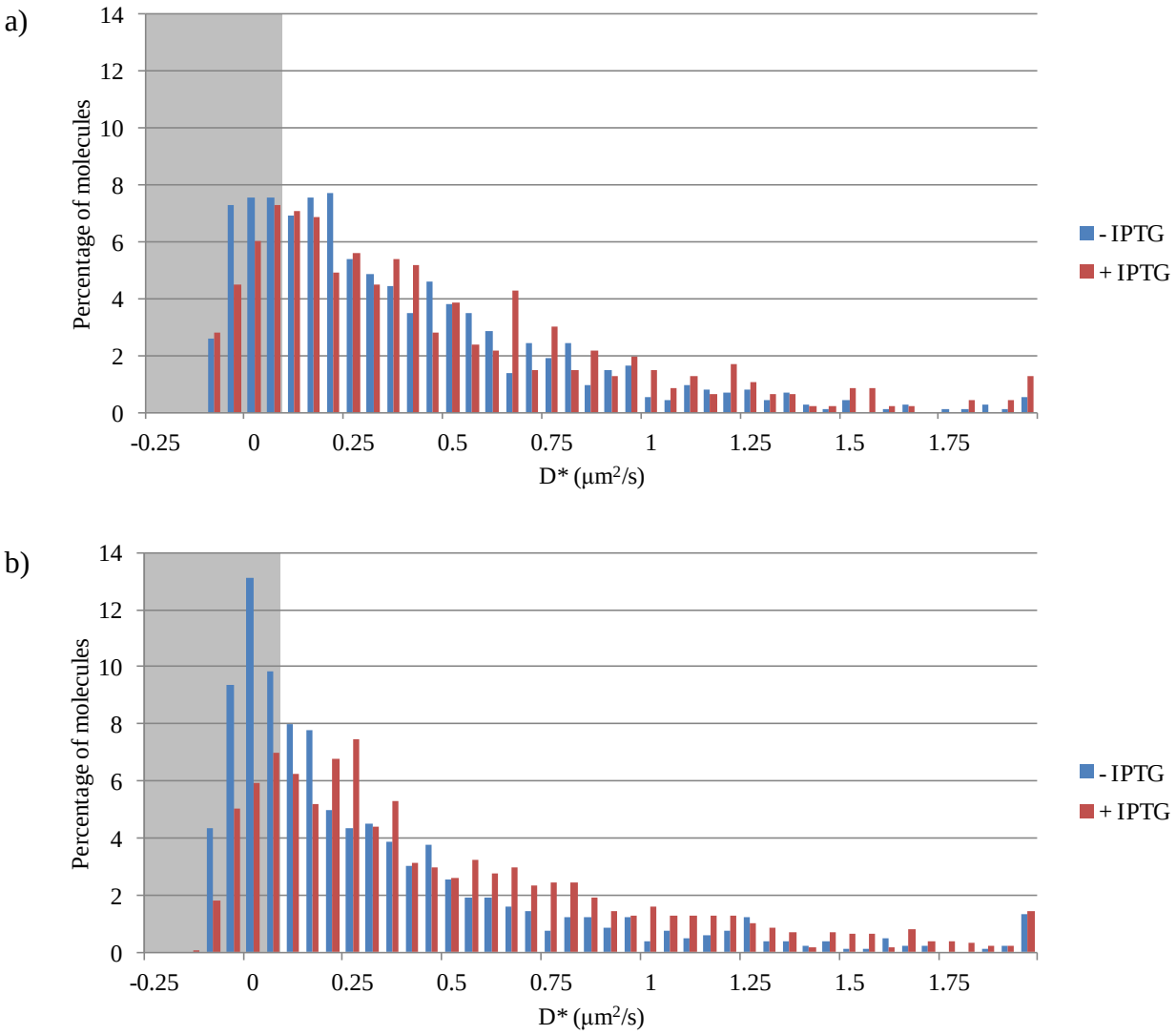


Figure 5.3: Effect of IPTG on the binding of LacI::PAmCherry to LacI DNA sites

Bar chart to show the apparent diffusion coefficient (D*) of LacI::PAmCherry with and without induction by IPTG. Bars in the grey section of the graph represent the immobile fraction of LacI::PAmCherry (D* < 0.1 $\mu m^2/s$).

a) Strain LR35, with a chromosomal fusion of LacI::PAmCherry and all LacI DNA sites removed.

b) Strain LR37, with a chromosomal fusion of LacI::PAmCherry and 6 LacI DNA sites inserted adjacent to *araBAD*.

Cells were grown in M9 minimal media supplemented with 0.2% glycerol to an OD_{600} of approximately 0.1. Cells were immobilised in 1 µl channels, and to induce, washed with M9 minimal media supplemented with 1 mM IPTG. Localisation analysis was performed using custom written MATLAB software. Microscopy and analysis was carried out by Federico Garza de Leon (University of Oxford). n = > 150

This result shows that changing the position of the LacI DNA sites on the chromosome does not change the behaviour of LacI. Also, as it agrees with known behaviours of LacI, it shows that this technique is relevant for studying transcription factors.

Figure 5.4 shows preliminary results of the tracks of LacI::PAmCherry in strain LR37, with 6 LacI DNA sites adjacent to *araBAD*. Red tracks show the diffusion of molecules with an apparent diffusion coefficient of less than 0.1 $\mu m^2/s$, blue tracks show the diffusion of molecules with an apparent diffusion coefficient of between 0.1 and 0.4 $\mu m^2/s$ and green tracks show the diffusion of molecules with an apparent diffusion coefficient of greater than 0.4 $\mu m^2/s$. It appears that when cells are uninduced (figure 5.4a), the tracks of immobile molecules cluster in one or two areas of the cell, sometimes with blue tracks, representing the diffusion of molecules that are mobile, but at the less mobile end of the scale. In cells that have been induced with IPTG the immobile molecules are more spread out (figure 5.4b). Although more cells need analysing, this suggests that the immobile molecules seen in induced cells are not due to LacI::PAmCherry binding at the multiple LacI DNA sites, as this would cause the immobile molecules to clusters.

5.4 The effect of the addition of arabinose on LacI binding to DNA sites adjacent to *araBAD*

Strain LR37 has a LacI::PAmCherry fusion on the chromosome and 6 LacI DNA sites inserted adjacent to *araBAD*. These are the only LacI DNA sites on the chromosome. To investigate whether induction of a neighbour promoter would affect the ability of LacI to bind its DNA sites, cells were induced with arabinose. If induction of *araBAD* changes the local chromosome structure the LacI DNA sites may become less accessible to LacI::PAmCherry,

a)

b)

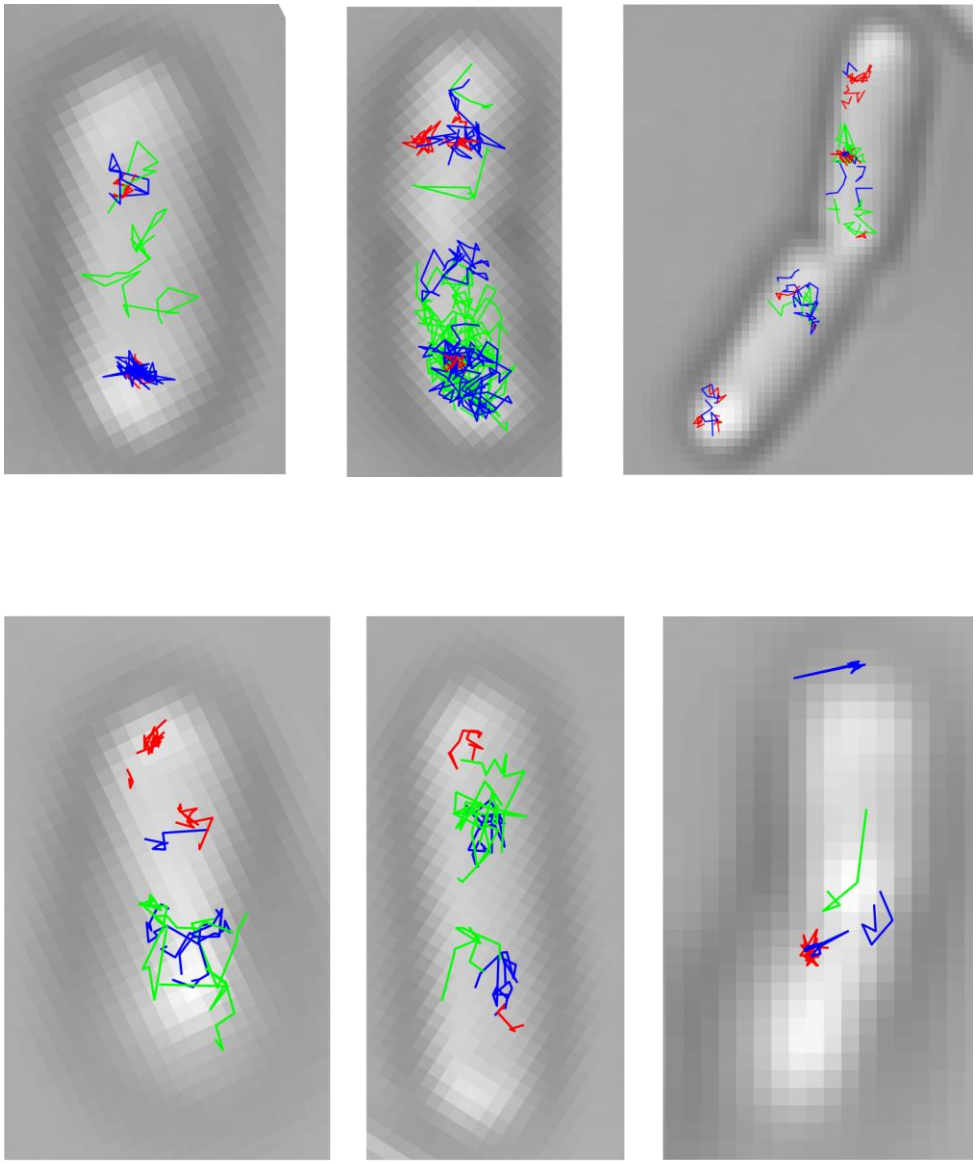


Figure 5.4: Tracks of LacI::PAmCherry

Images showing tracks of LacI::PAmCherry. Red tracks show the diffusion of molecules with an apparent diffusion coefficient of less than 0.1 $\mu m^2/s$, blue tracks show the diffusion of molecules with an apparent diffusion coefficient of between 0.1 and 0.4 $\mu m^2/s$ and green tracks show the diffusion of molecules with an apparent diffusion coefficient of greater than 0.4 $\mu m^2/s$.

a) Uninduced cells.

b) Cells induced with 1 mM IPTG.

Strain LR37, with a chromosomal fusion of LacI::PAmCherry and 6 LacI DNA sites inserted adjacent to *araBAD*. Cells were grown in M9 minimal media supplemented with 0.2% glycerol to an OD_{600} of approximately 0.1. Cells were immobilised in 1 µl channels, and to induce, washed with M9 minimal media supplemented with 1 mM IPTG. Localisation analysis was performed using custom written MATLAB software. Microscopy and analysis was carried out by Federico Garza de Leon (University of Oxford). Pixel size = 0.1145 µm.

which would be reflected in the apparent diffusion coefficient. Strain LR35, carrying a LacI::PAmCherry fusion and no LacI DNA sites, was also induced with arabinose as a control. Cells were immobilised in 15 μl wells containing polyethylenimine, which facilitates attachment of the cells. To induce, cells were washed with M9 minimal media supplemented with 0.2% arabinose.

Strain LR37 did not show any significant difference in the apparent diffusion coefficient of LacI::PAmCherry upon induction with arabinose ($P > 0.05$) (figure 5.5b). Strain LR35 was included as a control where LacI diffusion should not be affected by arabinose and also showed no significant difference ($P > 0.05$) (figure 5.5a). If induction of *araBAD* made the surrounding DNA more accessible to LacI::PAmCherry, there would be an increase in the percentage of molecules in the bound fraction. If it made the LacI DNA sites less accessible, there would be a decrease in the percentage of molecules in the immobile fraction (grey area of the graph). As there was no change, this suggests that induction of gene expression does not change the accessibility of neighbouring DNA sites to DNA binding proteins in this case.

5.5 Discussion

Previous super resolution studies of LacI had shown that the repressor locates its binding sites by a combination of 3D diffusion and non-specific binding to DNA, followed by sliding (Hammar *et al.*, 2012). It has also been shown that the LacI protein colocalised with its encoding gene and mRNA, even when this was not at its native location (Kuhlman and Cox, 2012). This chapter aimed to use another super resolution technique, PALM, to track LacI

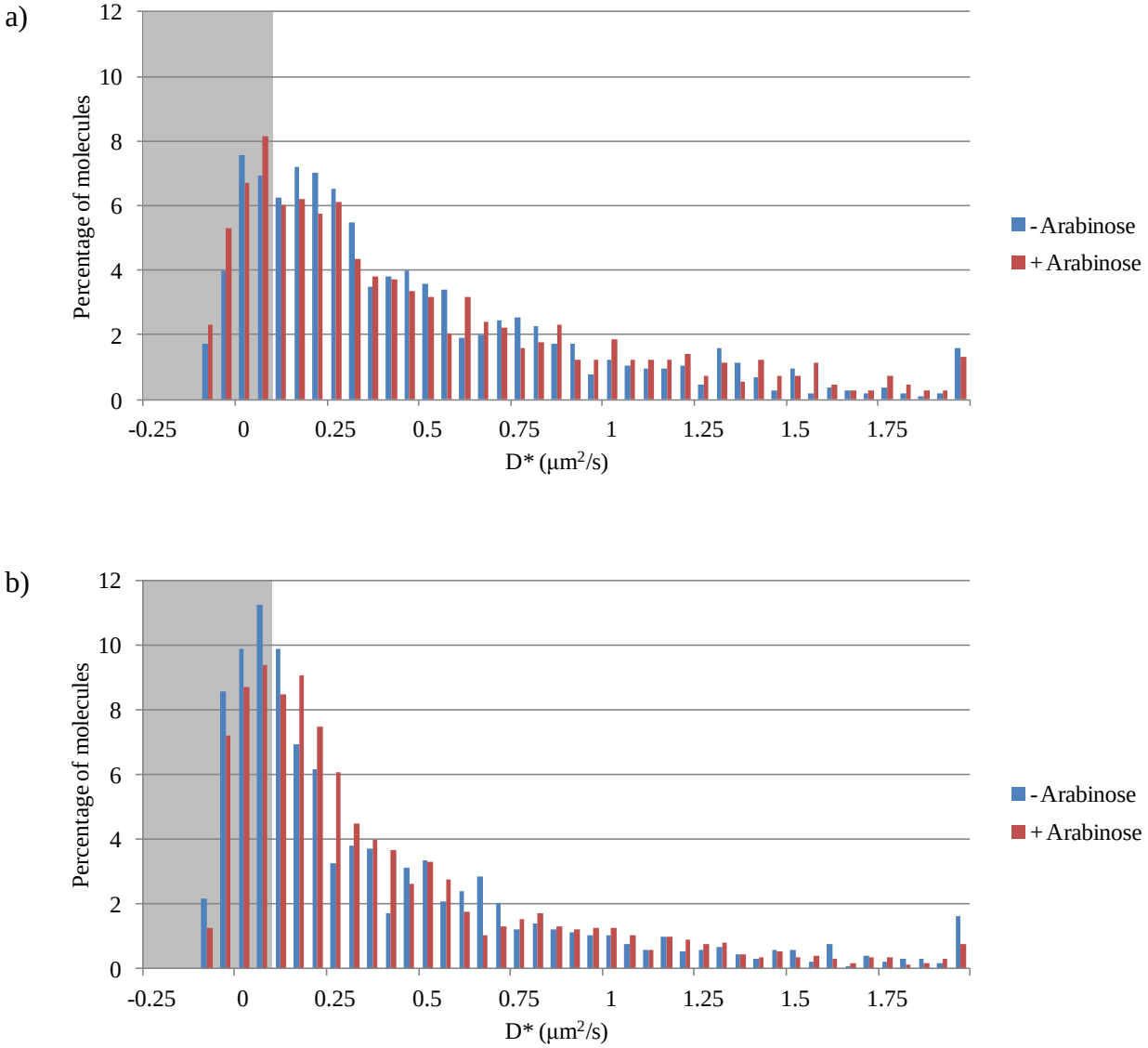


Figure 5.5: Effect of *araBAD* induction on the binding of LacI::PAmCherry to LacI DNA sites

Bar chart to show the apparent diffusion coefficient (D*) of LacI::PAmCherry with and without induction by IPTG. Bars in the grey section of the graph represent the immobile fraction of LacI::PAmCherry (D* < 0.1 $\mu m^2/s$).

a) Strain LR35, with a chromosomal fusion of LacI::PAmCherry and all LacI DNA sites removed.

b) Strain LR37, with a chromosomal fusion of LacI::PAmCherry and 6 LacI DNA sites inserted adjacent to *araBAD.*

Cells were grown in M9 minimal media supplemented with 0.2% glycerol to an OD_{600} of approximately 0.1. Cells were immobilised in 15 µl wells, and to induce, washed with M9 minimal media supplemented with 0.2% arabinose. Localisation analysis was performed using custom written MATLAB software. Microscopy and analysis was carried out by Federico Garza de Leon (University of Oxford). n = > 150

molecules in the cell to understand more about LacI diffusion and develop a technique that could be used to study other transcription factors.

The apparent diffusion coefficient of LacI::PAmCherry in a strain with no LacI DNA sites was compared to that in a strain with 6 LacI DNA sites inserted adjacent to the *araBAD* promoter. In the absence of LacI DNA sites 25% of LacI::PAmCherry was in the immobile fraction, even though there were no specific sites on the chromosome. These molecules are probably bound to DNA non-specifically, perhaps searching for a binding site by facilitated diffusion. This result agrees with the Hammar *et al.*, 2012 paper, and suggests that, when LacI is present in the cell at wild-type levels, about a quarter of the molecules are undergoing facilitated diffusion (Hammar *et al.*, 2012). When 6 LacI DNA sites were present on the chromosome, the percentage of LacI::PAmCherry in the bound fraction increased to 37%. This 11% change is presumed to be due to LacI::PAmCherry binding to the DNA sites. As there are only 6 DNA sites and an average of around 36 monomers of LacI::PAmCherry per cell, it is expected that there would still be a large percentage of molecules in the mobile fraction. If a larger array of DNA sites was inserted, the increase in molecules in the bound fraction would be expected to be larger. If the DNA site array is saturated, 12 LacI::PAmCherry monomers would be bound, one third of the molecules in the average cell. As 37% of molecules were in the bound fraction this suggests that the array may not be saturated, as some of these molecules will be bound non-specifically. There may not be 100% occupancy of the DNA site array due to steric hindrance effects.

When cells with LacI DNA sites on the chromosome were induced with IPTG, the percentage of molecules in the bound fraction decreased to a similar level to that of cells with no LacI DNA sites. This was expected as IPTG prevents LacI from binding specifically to its DNA sites. As 20-25% of molecules were still in the bound fraction this suggests that non-specific binding of LacI is not affected by IPTG. The profile of the apparent diffusion coefficient of LacI::PAmCherry in IPTG induced cells with LacI DNA sites looked very similar to that of uninduced cells with no DNA sites. This suggests that IPTG removes any effect of LacI DNA sites. In Chapter 3, inducing cells with IPTG did not prevent the formation of LacI::GFP foci, implying that IPTG did not have any effect. Although the fusions of LacI to GFP and PAmCherry should behave in the same way, they appear to have different sensitivity to IPTG. It is possible that the different tags have had different effects on the way that LacI responds to IPTG. However, this seems unlikely since GFP and PAmCherry are very similar molecules in both size and structure. A higher number of LacI DNA sites were present in the strain with LacI::GFP, although this does not explain why IPTG does not dissociate LacI::GFP from DNA sites on the chromosome.

Induction of the *araBAD* promoter did not affect the binding of LacI::PAmCherry to LacI DNA sites inserted adjacent to *araBAD*. If induction of *araBAD* caused a change in the local chromosome structure that made the DNA less accessible to LacI, a reduction in the percentage of molecules in the bound fraction would be seen, which was not the case. However, a change in structure due to induction of gene expression would be predicted to make the surrounding DNA more accessible, as the promoter would need to be bound by RNAP and any activating transcription factors. In this experiment, LacI is present in excess so

LacI DNA sites are thought to be saturated, as far as steric hindrance will allow. In this case, an increase in accessibility would not be seen, only a decrease.

These results have shown that PALM can be used to track transcription factors and the results agree with previous experiments, suggesting this method produces accurate data. To further investigate the diffusion of LacI, the LacI::PAmCherry fusion needs to be inserted onto the chromosome of a strain which still has the 3 natural LacI DNA sites. This will give information about LacI diffusion and binding in a system with the same stoichiometry as wild-type *E. coli*. Other transcription factors can then be investigated using this same method. In the case of transcription factors like AraC and MntR, with DNA sites at multiple locations around the chromosome, this will show whether they cluster, or not as suggested by the results of chapter 4.

6. Closing Comments

Our understanding of the bacterial nucleoid has expanded greatly in recent years, and bacteria are no longer assumed to have little or no structure in the way their DNA is arranged. It has been shown that, like eukaryotes, there are proteins dedicated to maintaining chromosome structure. However, unlike in eukaryotes, a number of these proteins involved in structuring chromosomes are also involved in gene regulation. This led to the idea that gene expression and chromosome structure could be linked in bacteria, and that a change in gene expression could change the local chromosome structure, or vice versa.

The technique used to investigate this idea was fluorescent repressor/operator system (FROS). A development of this technique to use smaller inserts has made it more suitable for studying chromosome structure. The development of MalI as a novel FROS reporter has given another option for experiments requiring multiple labels, as an alternative to TetR. Although the transcription factors studied in chapter 4 did not cause clustering of the promoters they control, there is evidence that this is the case for some other transcription factors, for example, GalR (Qian *et al.,* 2012). LacI::GFP and MalI::mCherry could be used to confirm the results already seen by chromosome conformation capture. Although new microscopy techniques, described in chapter 5, have shown that conventional fluorescence microscopy and FROS are limitated by resolution, there is still a place in science for these types of microscopy. Although epifluorescent microscopy does not give the same detail as super-resolution techniques, this study has shown it is sufficient to show large movements and requires much less specialised equipment and operators. One future application for this adapted FROS method could be to combine it with RNA FISH (Fluorescence *in situ* hybridisation) to follow both the promoter and newly synthesised mRNA of an inducible gene. To further study the induction of gene expression, the protein encoded by the gene could be tagged with a

fluorescent protein to give a three colour system able to follow a promoter, mRNA and protein. This would be particularly useful for studying membrane proteins and the possibility of the process of transertion.

The only promoter studied that changed position upon induction was *araFGH*. In all cases, induction of *araFGH* led to a movement away from the poles of the cell. *araFGH* encodes transporters, including a membrane protein, so the process of transertion could be the driving force behind this movement. Transertion would involve a movement towards a membrane, which would be assumed to be a movement along the radius of the cell. A movement along the length of the cell was seen but no radial measurements were taken in this study as resolution was thought to be insufficient for reliable measurements. Figure 6.1 shows how a movement along the length of the cell could be showing a movement towards the chromosome. As the cells are 3D but images are 2D, a movement towards the membrane could involve moving towards or away from the slide, which would not been seen by this type of microscopy, although TIRF microscopy may be able to detect this movement. The fact that a movement away from the cell pole is seen implies that the proteins encoded by *araFGH* require insertion at a specific point on the cell membrane, rather than insertion into the membrane at the closest point to the genes. In this case it seems to be an avoidance of the area of the cells around the pole, changing the usual asymmetrical arrangement of the chromosome to a symmentrical one . Without further experiments it would not be possible to say whether this effect was specific to a few membrane proteins, including AraFGH, or the case for all transerted proteins. The lack of any movement seen when *mntH* was induced suggests that there is no specific location for insertion of the MntH protein, although it is not

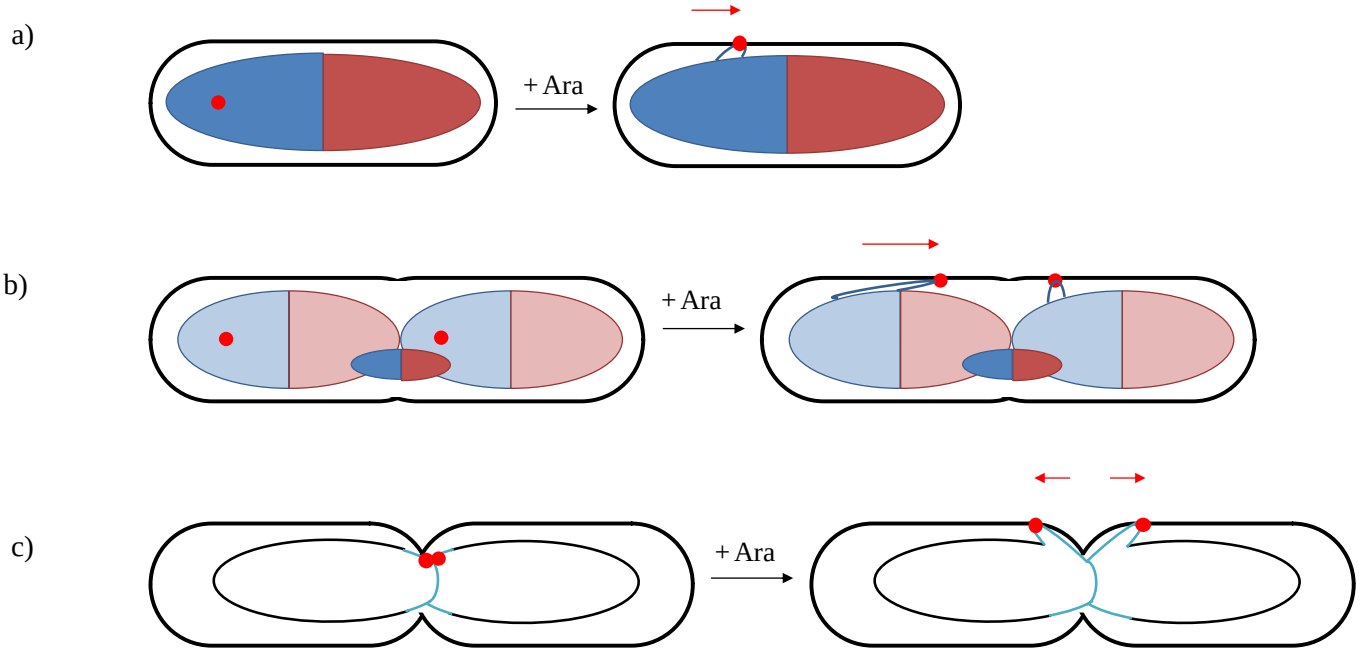


Figure 6.1 Schematic model of the movement of *araFGH* upon induction

Model showing the position of *araFGH* in cells at different stages of the cell cycle.

a) Cells with a single copy of the chromosome and a single focus representing the position of *araFGH*.

b) Cells with most of the chromosome replicated and two foci representing the position of *araFGH*.

c) Cells at the point of division.

Figures a and b show the left and right replichores as blue and red respectively and newly replicated DNA is paler. Figure c shows the chromosome as a black circle with the Ter macrodomain shown in pale blue, unsegregated. The position of *araFGH* is shown as a red circle. Red arrows show the direction of movement along the length of the cell upon induction by arabinose.

yet known whether all membrane proteins are transerted, or if some are inserted into the membrane separately from transcription and translation. A recent study showed that *lacY* moved across the radius of the cell, towards the membrane, upon induction, but no measurements of position related to cell length were made (Libby *et al.*, 2012). The promoters of a selection of other membrane proteins would need to be investigated, in the presence and absence of rifampicin, to see how widespread this effect is. To show the link between the movement and transertion, the membrane protein could be tagged with a fast folding fluorescent protein to show its position in the membrane when inserted, if the membrane protein can be tagged without affecting function. Other promoters of genes located in the Ter macrodomain should also be investigated. The properties of the Ter macrodomain may mean that promoters of genes encoding non-membrane inserted proteins may also move upon induction.

It is the belief of some that chromosome segregation is driven by transertion of membrane proteins. It is thought that as the region of the chromosome containing the gene that encodes the membrane protein moves towards the membrane, it would separate the two sister chromatids into opposite halves of the cell. For this to be the case, the two copies of the gene encoding the membrane protein would need to move towards the cell poles. This is not seen in the case of *araFGH* in single cells, probably because the chromosome is either not replicated or already segregated at the position of *araFGH*. However, when cells at the point of division are studied, the percentage of cells containing cohesive sister chromatids at the location of *araFGH* drops significantly upon addition of arabinose. *araFGH* is then segregated into the two future daughter cells. Alternatively, this could again be seen as a

movement away from the cell poles, although this time away from the septum, where the future cell poles of the daughter cells will be.

From these experiments, it is unclear whether the movement of *araFGH* leads to a permanent segregation of the entire Ter macrodomain. The regions of chromosome surrounding *araFGH* move in opposite directions, away from mid-cell and the rest of the Ter macrodomain, and towards the membrane. After transcription, translation and insertion have occurred, the DNA may remain in the same area of the cell or return to the rest of the Ter macrodomain. Alternatively, after the segregation of *araFGH*, the rest of the Ter macrodomain may follow, meaning that *araFGH* induction initiates chromosome segregation. This may depend on the point of the cell cycle and whether the Ter macrodomain has finished replicating. As figure 6.1 suggests, the movement of the *araFGH* promoter may involve a loop of the chromosome moving towards the chromosome without disturbing the rest of the structure.

It is known that the *E. coli* membrane has microdomains, some areas of the membrane are enriched for certain lipids, and this creates a non-uniform distribution of membrane proteins (Jacoby and Young, 1988). One reason for lipid enrichment is that some lipids have a more curved structure which excludes them from the length of the rod shaped cell. Instead, these lipids are found in clusters at the pole of the cell (Huang *et al.*, 2006). It has also been shown that two differently labelled phospholipids inserted into the *E. coli* membrane do not mix wth each other (Vanounou *et al.*, 2003). These microdomains will have differing properties which may favour insertion of different membrane proteins. Many *E. coli* proteins have been observed to be located at the pole of the cell, including proteins involved chemotaxis,

autotransporters and type III secretion systems (Li and Young, 2012). It is likely that there are some proteins that preferentially locate in the side membranes of the cell, perhaps for their function. It is unknown whether the site of transertion is the same as the final location for the membrane protein, but the AraFGH transporter may be required to be positioned away from the cell poles, so this is the site of insertion. Fluorescently labelled phospholipids used to investigate microdomains could be combined with FROS to see whether promoters move to a particular type of phospholipid.

The only promoter to show any movement within the cell upon induction was *araFGH*. The positions of other promoters controlled by AraC and promoters controlled by MntR showed no significant difference between supplemented and unsupplemented cultures. This suggests that, at these sites, gene expression does not have an effect on chromosome structure that can be detected using this method. It also suggests that neither AraC nor MntR binding causes clustering of their DNA sites. This effect has been shown with the transcription factor GalR (Qian *et al.,* 2012), indicating a possible role in chromosome structuring for transcription factors with DNA sites at multiple locations on the chromosome. Now the PALM system has been established and data from LacI has shown that it has given comparable data to previous studies, other transcription factors will be studied. This will include other transcription factors with a small number of DNA sites at one location on the chromosome, for example MalI, those with DNA sites at a small number of locations around the chromosome, for example MntR and AraC, and those with a large number of DNA sites, for example IHF.

Although the idea of transcription factors structuring the chromosome in a dynamic way that can react to the environment of the cell as the binding of transcription factors change, is attractive, this may not be practical. In *E. coli* there are thought to be around 300 transcription factors, 80% of which regulate more than 2 promoters (Martínez-Antonio and Collado-Vides, 2003). If all of these caused clustering when bound, the chromosome would end up as a knotted mass, where the possibility of other processes, such as replication, taking place seems unlikely. On top of this, 49% of promoters are regulated by more than one transcription factor, meaning that those promoters would need to belong to multiple clusters. Transcription factors also have an effect on chromosome structure when they have a secondary function as NAPs, when binding introduces bends and bridges into the DNA. It seems more likely that a small number of transcription factors could cause clustering of promoters they regulate, but not the majority.

References

Adelmann, C., Hetzler, J., Scheiber, G., Schimmel, T., Wegener, M., Weber, H. and Löhneysen, H. v., (1999) Experiments on the depolarization near-field scanning optical microscope. *Applied Physics Letters* **74**.

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

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

Anjem, A., Varghese, S. and Imlay, J. A., (2009) Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Molecular Microbiology* **72**: 844-858.

Atlung, T. and Hansen, F. G., (2002) Effect of different concentrations of H-NS protein on chromosome replication and the cell cycle in *Escherichia coli*. *Journal of Bacteriology* **184**: 1843-1850.

Aussel, L., Barre, F.-X., Aroyo, M., Stasiak, A., Stasiak, A. Z. and Sherratt, D., (2002) FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* **108**: 195-205.

Autret, S. and Errington, J., (2003) A role for division-site-selection protein MinD in regulation of internucleoid jumping of Soj (ParA) protein in *Bacillus subtilis*. *Molecular Microbiology* **47**: 159-169.

Azam, T. A. and Ishihama, A., (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli* sequence recognition specificity and DNA binding affinity. *Journal of Biological Chemistry* **274**: 33105-33113.

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

Barkley, M. D., Riggs, A. D., Jobe, A. and Bourgeois, S., (1975) Interaction of effecting ligands with *lac* repressor and repressor-operator complex. *Biochemistry* **14**: 1700-1712.

Bates, D. and Kleckner, N., (2005) Chromosome and Replisome Dynamics in *E. coli*: Loss of Sister Cohesion Triggers Global Chromosome Movement and Mediates Chromosome Segregation. *Cell* **121**: 899-911.

Bendich, A. J., (2001) The form of chromosomal DNA molecules in bacterial cells. *Biochimie* **83**: 177-186.

Berg, O. G., Winter, R. B. and Von Hippel, P. H., (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**: 6929-6948.

Bermejo, R., Branzei, D. and Foiani, M., (2008) Cohesion by topology: sister chromatids interlocked by DNA. *Genes & Development* **22**: 2297-2301.

Bernhardt, T. G. and De Boer, P. A., (2005) SlmA, a Nucleoid-Associated, FtsZ Binding Protein Required for Blocking Septal Ring Assembly over Chromosomes in *E. coli*. *Molecular Cell* **18**: 555-564.

Bhattacharyya, P., (1970) Active transport of manganese in isolated membranes of *Escherichia coli*. *Journal of bacteriology* **104**: 1307-1311.

Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. and Mayhew, G. F., (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1462.

Brendler, T., Sawitzke, J., Sergueev, K. and Austin, S., (2000) A case for sliding SeqA tracts at anchored replication forks during *Escherichia coli* chromosome replication and segregation. *The EMBO Journal* **19**: 6249-6258.

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

Browning, D. F., Grainger, D. C. and Busby, S. J. W., (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Current Opinion In Microbiology* **13**: 773-780.

Busby, S. and Ebright, R. H., (1999) Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology* **293**: 199-213.

Butala, M., Busby, S. J. and Lee, D. J., (2009) DNA sampling: a method for probing protein binding at specific loci on bacterial chromosomes. *Nucleic Acids Research* **37**: e37.

Cabrera, J. E. and Jin, D. J., (2003) The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. *Molecular Microbiology* **50**: 1493-1505.

Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A. and Darst, S. A., (2001) Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* **104**: 901-912.

Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L. and Darst, S. A., (2002) Structure of the bacterial RNA polymerase promoter specificity σ subunit. *Molecular Cell* **9**: 527-539.

Carmi, I., Kopczynski, J. B. and Meyer, B. J., (1998) The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. *Nature* **396**: 168-173.

Carra, J. H. and Schleif, R. F., (1993) Variation of half-site organization and DNA looping by AraC protein. *The EMBO Journal* **12**: 35.

Cattoni, D., Fiche, J. and Nöllmann, M., (2012) Single-molecule super-resolution imaging in bacteria. *Current Opinion In Microbiology*.

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

Cherepanov, P. P. and Wackernagel, W., (1995) Gene disruption in *Escherichia coli*: TcR and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**: 9-14.

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

Cho, H., McManus, H. R., Dove, S. L. and Bernhardt, T. G., (2011) Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proceedings of the National Academy of Sciences* **108**: 3773-3778.

Croxen, M. A. and Finlay, B. B., (2009) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews Microbiology* **8**: 26-38.

Daber, R., Stayrook, S., Rosenberg, A. and Lewis, M., (2007) Structural analysis of *lac* repressor bound to allosteric effectors. *Journal of Molecular Biology* **370**: 609-619.

Dame, R. T., Kalmykowa, O. J. and Grainger, D. C., (2011) Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in Gram negative Bacteria. *PLoS Genetics* **7**: e1002123.

Dame, R. T., Wyman, C. and Goosen, N., (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Research* **28**: 3504.

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D. and Possoz, C., (2007) MukB colocalizes with the *oriC* region and is required for organization of the two

Escherichia coli chromosome arms into separate cell halves. *Molecular Microbiology* **65**: 1485-1492.

Darst, S. A., Kubalek, E. W. and Kornberg, R. D., (1989) Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography.

Datsenko, K. A. and Wanner, B. L., (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**: 6640.

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Jannière, L. and Ehrlich, S. D., (2001) Two essential DNA polymerases at the bacterial replication fork. *Science* **294**: 1716-1719.

Dinardo, S., Voelkel, K. A., Sternglanz, R., Reynolds, A. E. and Wright, A., (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. *Cell* **31**: 43-51.

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

Draper, G. C., McLennan, N., Begg, K., Masters, M. and Donachie, W. D., (1998) Only the N-terminal domain of FtsK functions in cell division. *Journal of Bacteriology* **180**: 4621-4627.

Dupaigne, P., Tonthat, N. K., Espéli, O., Whitfill, T., Boccard, F. and Schumacher, M. A., (2012) Molecular Basis for a Protein-Mediated DNA-Bridging Mechanism that Functions in Condensation of the *E. coli* Chromosome. *Molecular Cell*.

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

Ebright, R. H., (2000) RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of Molecular Biology* **304**: 687-698.

Ehrmann, M. and Boos, W., (1987) Identification of endogenous inducers of the *mal* regulon in *Escherichia coli*. *Journal of Bacteriology* **169**: 3539-3545.

Elf, J., Li, G.-W. and Xie, X. S., (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* **316**: 1191-1194.

Englesberg, E., (1961) Enzymatic characterization of 17 L-arabinose negative mutants of *Escherichia coli*. *Journal of Bacteriology* **81**: 996.

Espéli, O., Borne, R., Dupaigne, P., Thiel, A., Gigant, E., Mercier, R. and Boccard, F., (2012) A MatP–divisome interaction coordinates chromosome segregation with cell division in *E. coli*. *The EMBO Journal* **31**: 3198-3211.

Espéli, O., Lee, C. and Marians, K. J., (2003) A physical and functional interaction between *Escherichia coli* FtsK and topoisomerase IV. *Journal of Biological Chemistry* **278**: 44639-44644.

Espéli, O., Mercier, R. and Boccard, F., (2008) DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Molecular Microbiology* **68**: 1418-1427.

Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J. F., Mbabaali, A., Haq, I., Higgins, C. F., Hinton, J. C. and Driscoll, P. C., (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *Journal of Molecular Biology* **324**: 841-850.

Falcon, C. M. and Matthews, K. S., (1999) Glycine insertion in the hinge region of lactose repressor protein alters DNA binding. *Journal of Biological Chemistry* **274**: 30849-30857.

Finkel, S. and Johnson, R., (1992) The Fis protein: it's not just for DNA inversion anymore. *Molecular Microbiology* **6**: 3257-3265.

Finzi, L. and Gelles, J., (1995) Single DNA Molecules. *Science* **267**: 20.

Fried, L., Behr, S. and Jung, K., (2013) Identification of a target gene and activating stimulus for the YpdA/YpdB histidine kinase/response regulator system in *Escherichia coli*. *Journal of Bacteriology* **195**: 807-815.

Friedman, D. I., (1988) Integration host factor: A protein for all reasons. *Cell* **55**: 545-554.

Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K. and Ramos, J. L., (1997) Arac/XylS family of transcriptional regulators. *Microbiology and Molecular Biology Reviews* **61**: 393-410.

Gasser, S. M., (2002) Visualizing chromatin dynamics in interphase nuclei. *Science* **296**: 1412.

Geslin, C., Llanos, J., Prieur, D. and Jeanthon, C., (2001) The manganese and iron superoxide dismutases protect *Escherichia coli* from heavy metal toxicity. *Research in Microbiology* **152**: 901-905.

Ghosh, P., Ishihama, A. and Chatterji, D., (2001) *Escherichia coli* RNA polymerase subunit ω and its N-terminal domain bind full-length β′ to facilitate incorporation into the α2β subassembly. *European Journal of Biochemistry* **268**: 4621-4627.

Gilbert, W. and Müller-Hill, B., (1966) Isolation of the *lac* repressor. *Proceedings of the National Academy of Sciences* **56**: 1891.

Gordon, G. S., Sitnikov, D., Webb, C. D., Teleman, A., Straight, A., Losick, R., Murray, A. W. and Wright, A., (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* **90**: 1113-1121.

Grainger, D. C., Hurd, D., Goldberg, M. D. and Busby, S. J. W., (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Research* **34**: 4642-4652.

Grainger, D. C., Hurd, D., Harrison, M., Holdstock, J. and Busby, S. J., (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* **102**: 17693-17698.

Gugel, H., Bewersdorf, J., Jakobs, S., Engelhardt, J., Storz, R. and Hell, S. W., (2004) Cooperative 4Pi excitation and detection yields sevenfold sharper optical sections in live-cell microscopy. *Biophysical Journal* **87**: 4146-4152.

Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E. G., Berg, O. G. and Elf, J., (2012) The *lac* repressor displays facilitated diffusion in living cells. *Science* **336**: 1595-1598.

Hayward, R. S., Igarashi, K. and Ishihama, A., (1991) Functional specialization within the α-subunit of *Escherichia coli* RNA polymerase. *Journal of Molecular Biology* **221**: 23-29.

Hell, S. W., (2007) Far-Field Optical Nanoscopy. *Science* **316**: 1153-1158.

Henaut, A., Rouxel, T., Gleizes, A., Moszer, I. and Danchin, A., (1996) Uneven Distribution of GATC Motifs in the *Escherichia coli* Chromosome, its Plasmids and its Phages. *Journal of Molecular Biology* **257**: 574-585.

Hendrickson, W., Stoner, C. and Schleif, R., (1990) Characterization of the *Escherichia coli ara FGH* and *ara J* promoters. *Journal of Molecular Biology* **215**: 497-510.

Hengge-Aronis, R., (1993) Survival of hunger and stress: the role of rpoS in early stationary phase gene regulation in *E. coli*. *Cell* **72**: 165-168.

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

Hill, T., Kopp, B. and Kuempel, P., (1988a) Termination of DNA replication in *Escherichia coli* requires a trans-acting factor. *Journal of Bacteriology* **170**: 662-668.

Hill, T. M., Pelletier, A. J., Tecklenburg, M. L. and Kuempel, P. L., (1988b) Identification of the DNA sequence from the *E. coli* terminus region that halts replication forks. *Cell* **55**: 459-466.

Hobot, J. A., Villiger, W., Escaig, J., Maeder, M., Ryter, A. and Kellenberger, E., (1985) Shape and fine structure of nucleoids observed on sections of ultrarapidly frozen and cryosubstituted bacteria. *Journal of Bacteriology* **162**: 960.

Holmes, V. F. and Cozzarelli, N. R., (2000) Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proceedings of the National Academy of Sciences* **97**: 1322.

Hu, Z., Mukherjee, A., Pichoff, S. and Lutkenhaus, J., (1999) The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. *Proceedings of the National Academy of Sciences* **96**: 14819-14824.

Huang, K. C., Meir, Y. and Wingreen, N. S., (2003) Dynamic structures in *Escherichia coli*: spontaneous formation of MinE rings and MinD polar zones. *Proceedings of the National Academy of Sciences* **100**: 12724-12728.

Huang, K. C., Mukhopadhyay, R. and Wingreen, N. S., (2006) A curvature-mediated mechanism for localization of lipids to bacterial poles. *PLoS Computational Biology* **2**: e151.

Hudson, J. M. and Fried, M. G., (1990) Co-operative interactions between the catabolite gene activator protein and the *lac* repressor at the lactose promoter. *Journal of Molecular Biology* **214**: 381-396.

Ip, S. C., Bregu, M., Barre, F.-X. and Sherratt, D. J., (2003) Decatenation of DNA circles by FtsK-dependent Xer site-specific recombination. *The EMBO Journal* **22**: 6399-6407.

Ishihama, A., (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annual Reviews in Microbiology* **54**: 499-518.

Jacoby, G. H. and Young, K. D., (1988) Unequal distribution of penicillin-binding proteins among inner membrane vesicles of *Escherichia coli*. *Journal of Bacteriology* **170**: 3660-3667.

Jakubovics, N. S. and Jenkinson, H. F., (2001) Out of the iron age: new insights into the critical role of manganese homeostasis in bacteria. *Microbiology* **147**: 1709.

Johnson, C. M. and Schleif, R. F., (1995) In vivo induction kinetics of the arabinose promoters in *Escherichia coli*. *Journal of Bacteriology* **177**: 3438-3442.

Kahramanoglou, C., Webster, C. L., El-Robh, M. S., Belyaeva, T. A. and Busby, S. J. W., (2006) Mutational analysis of the *Escherichia coli melR* gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. *Journal of Bacteriology* **188**: 3199-3207.

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

Katayama, T., Ozaki, S., Keyamura, K. and Fujimitsu, K., (2010) Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nature Reviews Microbiology* **8**: 163-170.

Kato, N. and Lam, E., (2001) Detection of chromosomes tagged with green fluorescent protein in live *Arabidopsis thaliana* plants. *Genome Biology* **2**: 1-0045.0010.

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

Kehres, D. G., Zaharik, M. L., Finlay, B. B. and Maguire, M. E., (2000) The NRAMP proteins of *Salmonella typhimurium* and *Escherichia coli* are selective manganese

transporters involved in the response to reactive oxygen. *Molecular Microbiology* **36**: 1085-1100.

Keseler, I. M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñiz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M. and Altman, T., (2011) EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Research* **39**: D583-D590.

Kim, J., Yoshimura, S. H., Hizume, K., Ohniwa, R. L., Ishihama, A. and Takeyasu, K., (2004) Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Research* **32**: 1982-1992.

Knowles, T. J., Scott-Tucker, A., Overduin, M. and Henderson, I. R., (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nature Reviews Microbiology* **7**: 206-214.

Kolodrubetz, D. and Schleif, R., (1981) L-arabinose transport systems in *Escherichia coli* K-12. *Journal of Bacteriology* **148**: 472-479.

Kuhlman, T. E. and Cox, E. C., (2012) Gene location and DNA density determine transcription factor distributions in *Escherichia coli*. *Molecular Systems Biology* **8**.

Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C. O., Mavathur, R., Muskhelishvili, G., Pon, C. L. and Rimsky, S., (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Research* **35**: 6330-6337.

Lau, I. F., Filipe, S. R., Søballe, B., Økstad, O. A., Barre, F. X. and Sherratt, D. J., (2003) Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Molecular Microbiology* **49**: 731-743.

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

Lerman, L., (1971) A transition to a compact form of DNA in polymer solutions. *Proceedings of the National Academy of Sciences* **68**: 1886-1890.

Levine, C., Hiasa, H. and Marians, K. J., (1998) DNA gyrase and topoisomerase IV: biochemical activities, physiological roles during chromosome replication, and drug sensitivities. *Biochimica et Biophysica Acta -Gene Structure and Expression* **1400**: 29-43.

Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. and Lu, P., (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**: 1247.

Lewis, P. J., Thaker, S. D. and Errington, J., (2000) Compartmentalization of transcription and translation in *Bacillus subtilis*. *The EMBO Journal* **19**: 710-718.

Li, G. and Young, K. D., (2012) Isolation and identification of new inner membrane-associated proteins that localize to cell poles in *Escherichia coli*. *Molecular Microbiology* **84**: 276-295.

Li, Z. and Jensen, G. J., (2009) Electron cryotomography: a new view into microbial ultrastructure. *Current Opinion in Microbiology* **12**: 333-340.

Libby, E. A., Roggiani, M. and Goulian, M., (2012) Membrane protein expression triggers chromosomal locus repositioning in bacteria. *Proceedings of the National Academy of Sciences* **109**: 7445-7450.

Liu, L. F., Liu, C.-C. and Alberts, B. M., (1980) Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* **19**: 697-707.

Lloyd, G. S., Godfrey, R. E. and Busby, S. J. W., (2010) Targets for the MalI repressor at the divergent *Escherichia coli* K 12 malX malI promoters. *FEMS Microbiology Letters* **305**: 28-34.

Lloyd, G. S., Hollands, K., Godfrey, R. E. and Busby, S. J. W., (2008) Transcription initiation in the *Escherichia coli* K-12 *malI–malX* intergenic region and the role of the cyclic AMP receptor protein. *FEMS Microbiology Letters* **288**: 250-257.

Löwe, J. and Amos, L. A., (1998) Crystal structure of the bacterial cell-division protein FtsZ. *Nature* **391**: 203-206.

Maeda, H., Fujita, N. and Ishihama, A., (2000) Competition among seven *Escherichia coli* subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Research* **28**: 3497.

Makui, H., Roig, E., Cole, S. T., Helmann, J. D., Gros, P. and Cellier, M. F. M., (2000) Identification of the *Escherichia coli* K-12 Nramp orthologue (MntH) as a selective divalent metal ion transporter. *Molecular Microbiology* **35**: 1065-1078.

Manley, S., Gillette, J. M., Patterson, G. H., Shroff, H., Hess, H. F., Betzig, E. and Lippincott-Schwartz, J., (2008) High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nature Methods* **5**: 155-157.

Männik, J., Wu, F., Hol, F. J., Bisicchia, P., Sherratt, D. J., Keymer, J. E. and Dekker, C., (2012) Robustness and accuracy of cell division in *Escherichia coli* in diverse cell shapes. *Proceedings of the National Academy of Sciences* **109**: 6957-6962.

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

Meister, P., Towbin, B. D., Pike, B. L., Ponti, A. and Gasser, S. M., (2010) The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes & Development* **24**: 766.

Mercier, R., Petit, M. A., Schbath, S., Robin, S., El Karoui, M., Boccard, F. and Espéli, O., (2008) The MatP/*matS* site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135**: 475-485.

Michalet, X., (2010) Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E* **82**: 041914.

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

Miller, O., Hamkalo, B. A. and Thomas, C., (1970) Visualization of bacterial genes in action. *Science* **169**: 392-395.

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

Mukherjee, K., Nagai, H., Shimamoto, N. and Chatterji, D., (1999) GroEL is involved in activation of *Escherichia coli* RNA polymerase devoid of the ω subunit in vivo. *European Journal of Biochemistry* **266**: 228-235.

Mulcair, M. D., Schaeffer, P. M., Oakley, A. J., Cross, H. F., Neylon, C., Hill, T. M. and Dixon, N. E., (2006) A Molecular Mousetrap Determines Polarity of Termination of DNA Replication in *E. coli*. *Cell* **125**: 1309-1319.

Mulugu, S., Potnis, A., Taylor, J., Alexander, K. and Bastia, D., (2001) Mechanism of termination of DNA replication of *Escherichia coli* involves helicase–contrahelicase interaction. *Proceedings of the National Academy of Sciences* **98**: 9569-9574.

Murakami, K. S. and Darst, S. A., (2003) Bacterial RNA polymerases: the wholo story. *Current Opinion in Structural Biology* **13**: 31-39.

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. and Darst, S. A., (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285-1290.

Murphy, L. D. and Zimmerman, S. B., (1995) Condensation and cohesion of λ DNA in cell extracts and other media: Implications for the structure and function of DNA in prokaryotes. *Biophysical Chemistry* **57**: 71-92.

Murray, P. R., Rosenthal, K. S. and Pfaller, M. A., (2009) *Medical Microbiology*. Mosby/Elsevier Philadelphia, PA.

Natarajan, S., Kelley, W. L. and Bastia, D., (1991) Replication terminator protein of *Escherichia coli* is a transcriptional repressor of its own synthesis. *Proceedings of the National Academy of Sciences* **88**: 3867-3871.

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

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

Niu, L. and Yu, J., (2008) Investigating intracellular dynamics of FtsZ cytoskeleton with photoactivation single-molecule tracking. *Biophysical Journal* **95**: 2009-2016.

Norris, V., (1995) Hypothesis: chromosome separation in *Escherichia coli* involves autocatalytic gene expression, transertion and membrane-domain formation. *Molecular Microbiology* **16**: 1051-1057.

Norris, V. and Madsen, M. S., (1995) Autocatalytic Gene Expression Occurs via Transertion and Membrane Domain Formation and Underlies Differentiation in Bacteria: A Model. *Journal of Molecular Biology* **253**: 739-748.

Nurse, P., Levine, C., Hassing, H. and Marians, K. J., (2003) Topoisomerase III can serve as the cellular decatenase in *Escherichia coli*. *Journal of Biological Chemistry* **278**: 8653-8660.

Oehler, S., Eismann, E. R., Krämer, H. and Müller-Hill, B., (1990) The three operators of the *lac* operon cooperate in repression. *The EMBO Journal* **9**: 973.

Ogden, S., Haggerty, D., Stoner, C. M., Kolodrubetz, D. and Schleif, R., (1980) The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. *Proceedings of the National Academy of Sciences* **77**: 3346.

Ohsumi, K., Yamazoe, M. and Hiraga, S., (2001) Different localization of SeqA-bound nascent DNA clusters and MukF–MukE–MukB complex in *Escherichia coli* cells. *Molecular Microbiology* **40**: 835-845.

Pan, C. Q., Finkel, S. E., Cramton, S. E., Feng, J. A., Sigman, D. S. and Johnson, R. C., (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *Journal of Molecular Biology* **264**: 675-695.

Patzer, S. I. and Hantke, K., (2001) Dual repression by Fe^{2+}-Fur and Mn^{2+}-MntR of the *mntH* gene, encoding an NRAMP-like Mn^{2+} transporter in *Escherichia coli*. *Journal of Bacteriology* **183**: 4806-4813.

Pease, P. J., Levy, O., Cost, G. J., Gore, J., Ptacin, J. L., Sherratt, D., Bustamante, C. and Cozzarelli, N. R., (2005) Sequence-directed DNA translocation by purified FtsK. *Science* **307**: 586-590.

Pérals, K., Capiaux, H., Vincourt, J. B., Louarn, J. M., Sherratt, D. J. and Cornet, F., (2001) Interplay between recombination, cell division and chromosome structure during chromosome dimer resolution in *Escherichia coli*. *Molecular Microbiology* **39**: 904-913.

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

Peter, B. J., Arsuaga, J., Breier, A. M., Khodursky, A. B., Brown, P. O. and Cozzarelli, N. R., (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biology* **5**: R87.

Possoz, C., Junier, I. and Espéli, O., (2012) Bacterial chromosome segregation. *Frontiers in Bioscience: a Journal and Virtual Library* **17**: 1020.

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

Qian, Z., Dimitriadis, E. K., Edgar, R., Eswaramoorthy, P. and Adhya, S., (2012) Galactose repressor mediated intersegmental chromosomal connections in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **109**: 11336-11341.

Raskin, D. M. and de Boer, P. A., (1999) Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. *Proceedings of the National Academy of Sciences* **96**: 4971-4976.

Reeder, T. and Schleif, R., (1991) Mapping, sequence, and apparent lack of function of *araJ*, a gene of the *Escherichia coli* arabinose regulon. *Journal of Bacteriology* **173**: 7765-7771.

Reidl, J. and Boos, W., (1991) The *malX malY* operon of *Escherichia coli* encodes a novel enzyme II of the phosphotransferase system recognizing glucose and maltose and an enzyme abolishing the endogenous induction of the maltose system. *Journal of Bacteriology* **173**: 4862-4876.

Reidl, J., Römisch, K., Ehrmann, M. and Boos, W., (1989) MalI, a novel protein involved in regulation of the maltose system of *Escherichia coli*, is highly homologous to the repressor proteins GalR, CytR, and LacI. *Journal of Bacteriology* **171**: 4888-4899.

Reyes-Lamothe, R., Nicolas, E. and Sherratt, D. J., (2012) Chromosome replication and segregation in bacteria. *Annual Review of Genetics* **46**: 121-143.

Reyes-Lamothe, R., Possoz, C., Danilova, O. and Sherratt, D. J., (2008a) Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**: 90-102.

Reyes-Lamothe, R., Wang, X. and Sherratt, D., (2008b) *Escherichia coli* and its chromosome. *Trends in Microbiology* **16**: 238-245.

Rice, P. A., Yang, S.-W., Mizuuchi, K. and Nash, H. A., (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295-1306.

Richter, P. H. and Eigen, M., (1974) Diffusion controlled reaction rates in spheroidal geometry: application to repressor-operator association and membrane bound enzymes. *Biophysical Chemistry* **2**: 255-263.

Riggs, A. D., Bourgeois, S. and Cohn, M., (1970) The *lac* represser-operator interaction: III. Kinetic studies. *Journal of Molecular Biology* **53**: 401-417.

Robinett, C. C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A. and Belmont, A. S., (1996) *In vivo* localization of DNA sequences and visualization of large-scale chromatin organization using *lac* operator/repressor recognition. *The Journal of Cell Biology* **135**: 1685.

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R. L., (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**: 1407-1413.

Russo, T. A. and Johnson, J. R., (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection* **5**: 449-456.

Ruusala, T. and Crothers, D. M., (1992) Sliding and intermolecular transfer of the lac repressor: kinetic perturbation of a reaction intermediate by a distant DNA sequence. *Proceedings of the National Academy of Sciences* **89**: 4903-4907.

Salerno, P., Larsson, J., Bucca, G., Laing, E., Smith, C. P. and Flärdh, K., (2009) One of the two genes encoding nucleoid-associated HU proteins in *Streptomyces coelicolor* is developmentally regulated and specifically involved in spore maturation. *Journal of Bacteriology* **191**: 6489-6500.

Sánchez-Romero, M. A., Lee, D. J., Sánchez-Morán, E. and Busby, S. J. W., (2012) Location and dynamics of an active promoter in *Escherichia coli* K-12. *Biochemical Journal* **441**: 481.

Sanderson, A., Mitchell, J. E., Minchin, S. D. and Busby, S. J., (2003) Substitutions in the *Escherichia coli* RNA polymerase σ^{70} factor that affect recognition of extended−10 elements at promoters. *FEBS Letters* **544**: 199-205.

Schermelleh, L., Heintzmann, R. and Leonhardt, H., (2010) A guide to super-resolution fluorescence microscopy. *The Journal of Cell Biology* **190**: 165-175.

Schleif, R., (2000) Regulation of the arabinose operon of *Escherichia coli*. *Trends in Genetics* **16**: 559-565.

Schleif, R., Hess, W., Finkelstein, S. and Ellis, D., (1973) Induction kinetics of the L-arabinose operon of *Escherichia coli*. *Journal of Bacteriology* **115**: 9-14.

Schneider, D. A., Ross, W. and Gourse, R. L., (2003) Control of rRNA expression in *Escherichia coli*. *Current Opinion in Microbiology* **6**: 151-156.

Schneider, R., Lurz, R., Lüder, G., Tolksdorf, C., Travers, A. and Muskhelishvili, G., (2001) An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Research* **29**: 5107-5114.

Semsey, S., Geanacopoulos, M., Lewis, D. E. and Adhya, S., (2002) Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *The EMBO Journal* **21**: 4349-4356.

Sherratt, D. J., (2003) Bacterial chromosome dynamics. *Science* **301**: 780-785.

Sherratt, D. J., Lau, I. F. and Barre, F.-X., (2001) Chromosome segregation. *Current Opinion in Microbiology* **4**: 653-659.

Shin, M., Kang, S., Hyun, S.-J., Fujita, N., Ishihama, A., Valentin-Hansen, P. and Choy, H. E., (2001) Repression of *deoP2* in *Escherichia coli* by CytR: conversion of a transcription activator into a repressor. *The EMBO Journal* **20**: 5392-5399.

Sivanathan, V., Allen, M. D., de Bekker, C., Baker, R., Arciszewska, L. K., Freund, S. M., Bycroft, M., Löwe, J. and Sherratt, D. J., (2006) The FtsK γ domain directs oriented DNA translocation by interacting with KOPS. *Nature Structural & Molecular Biology* **13**: 965-972.

Soisson, S. M., MacDougall-Shackleton, B., Schleif, R. and Wolberger, C., (1997) Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**: 421-425.

Soppa, J., (2001) Prokaryotic structural maintenance of chromosomes (SMC) proteins: distribution, phylogeny, and comparison with MukBs and additional prokaryotic and eukaryotic coiled-coil proteins. *Gene* **278**: 253-264.

Straight, A. F., Belmont, A. S., Robinett, C. C. and Murray, A. W., (1996) GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. *Current Biology* **6**: 1599-1608.

Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J. and Verkhusha, V. V., (2009) Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Methods* **6**: 153-159.

Swint-Kruse, L. and Matthews, K. S., (2009) Allostery in the LacI/GalR family: variations on a theme. *Current Opinion in Microbiology* **12**: 129-137.

Tanaka, T., Shinkai, A., Bessho, Y., Kumarevel, T. and Yokoyama, S., (2009) Crystal structure of the manganese transport regulatory protein from *Escherichia coli*. *Proteins: Structure, Function, and Bioinformatics* **77**: 741-746.

Thanbichler, M. and Shapiro, L., (2006) Chromosome organization and segregation in bacteria. *Journal of Structural Biology* **156**: 292-303.

Thiel, A., Valens, M., Vallet-Gely, I., Espéli, O. and Boccard, F., (2012) Long-range chromosome organization in *E. coli*: a site-specific system isolates the Ter macrodomain. *PLoS Genetics* **8**: e1002672.

Thomason, L. C. and Costantino, N., (2007) E. coli Genome Manipulation by P1 Transduction. *Current Protocols in Molecular Biology* **1**: 17

Trun, N. J. and Marko, J. F., (1998) Architecture of a bacterial chromosome. *ASM News* **64**: 276-283.

Uphoff, S., Reyes-Lamothe, R., de Leon, F. G., Sherratt, D. J. and Kapanidis, A. N., (2013) Single-molecule DNA repair in live bacteria. *Proceedings of the National Academy of Sciences* **110**: 8063-8068.

Valens, M., Penaud, S., Rossignol, M., Cornet, F. and Boccard, F., (2004) Macrodomain organization of the *Escherichia coli* chromosome. *The EMBO Journal* **23**: 4330-4341.

Vanounou, S., Parola, A. H. and Fishov, I., (2003) Phosphatidylethanolamine and phosphatidylglycerol are segregated into different domains in bacterial membrane. A study with pyrene-labelled phospholipids. *Molecular Microbiology* **49**: 1067-1079.

Wang, M. D., Schnitzer, M. J., Yin, H., Landick, R., Gelles, J. and Block, S. M., (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* **282**: 902-907.

Wang, X., Liu, X., Possoz, C. and Sherratt, D. J., (2006) The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes & Development* **20**: 1727-1731.

Waters, L. S., Sandoval, M. and Storz, G., (2011) The *Escherichia coli* MntR Miniregulon Includes Genes Encoding a Small Protein and an Efflux Pump Required for Manganese Homeostasis. *Journal of Bacteriology* **193**: 5887-5897.

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

Woldringh, C. L., (2002) The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Molecular Microbiology* **45**: 17-29.

Woldringh, C. L., Jensen, P. R. and Westerhoff, H. V., (1995) Structure and partitioning of bacterial DNA: determined by a balance of compaction and expansion forces? *FEMS Microbiology Letters* **131**: 235-242.

Wolf, S. G., Frenkiel, D., Arad, T., Finkel, S. E., Kolter, R. and Minsky, A., (1999) DNA protection by stress-induced biocrystallization. *Nature* **400**: 83-85.

Worcel, A. and Burgi, E., (1972) On the structure of the folded chromosome of *Escherichia coli*. *Journal of Molecular Biology* **71**: 127.

Yamamoto, K., Ishihama, A., Busby, S. J. W. and Grainger, D. C., (2011) The Escherichia coli K-12 MntR miniregulon includes dps, which encodes the major stationary-phase DNA-binding protein. *Journal of bacteriology* **193**: 1477-1480.

Zaritsky, A. and Woldringh, C. L., (2003) Localizing cell division in spherical Escherichia coli by nucleoid occlusion. *FEMS microbiology letters* **226**: 209-214.

Zdych, E., Peist, R., Reidl, J. and Boos, W., (1995) MalY of Escherichia coli is an enzyme with the activity of a beta CS lyase (cystathionase). *Journal of bacteriology* **177**: 5035.

Zechiedrich, E. L., Khodursky, A. B., Bachellier, S., Schneider, R., Chen, D., Lilley, D. M. and Cozzarelli, N. R., (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in Escherichia coli. *Journal of Biological Chemistry* **275**: 8103-8113.

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. and Darst, S. A., (1999) Crystal Structure of< i> Thermus aquaticus</i> Core RNA Polymerase at 3.3 Å Resolution. *Cell* **98**: 811-824.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M. X., Arnold, E. and Ebright, R. H., (2012) Structural basis of transcription initiation. *Science* **338**: 1076-1080.

Zimmerman, S. B., (2006) Cooperative transitions of isolated Escherichia coli nucleoids: implications for the nucleoid as a cellular phase. *Journal of structural biology* **153**: 160-175.

Zimmerman, S. B. and Trach, S. O., (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of molecular biology* **222**: 599-620.

Zuo, Y., Wang, Y. and Steitz, Thomas A., (2013) The Mechanism of E. coli RNA Polymerase Regulation by ppGpp Is Suggested by the Structure of their Complex. *Molecular cell* **50**: 430-436.

Zusman, D. R., Carbonell, A. and Haga, J. Y., (1973) Nucleoid condensation and cell division in Escherichia coli MX74T2 ts52 after inhibition of protein synthesis. *Journal of bacteriology* **115**: 1167-1178.