THE UNIVERSITY
of EDINBURGH

Visualization of Replication-Dependent DNA Double-Strand Break Repair in *Escherichia coli*

Vincent Amarh

Doctor of Philosophy (PhD)

The University of Edinburgh

2017

Declaration

I hereby declare that the content of this thesis is the outcome of the research I conducted under the supervision of Prof. David Leach. Citations from all other authors have been duly acknowledged. This work has not been submitted elsewhere for the award of a degree.

Vincent Amarh

Acknowledgements

I am extremely grateful to Prof. David Leach for offering me the opportunity to conduct this research under his supervision. I am thankful to Dr Meriem El Karoui for discussions and feedbacks during the course of this research. I am also grateful to Prof. Adele Marston for accepting to be a member of my thesis committee.

To Dr Elise Darmon, I am forever grateful for all the assistance rendered to me; it was a wonderful experience working together using the fluorescence microscope. I will also like to acknowledge the immense support I received from the entire members of the Leach lab.

Finally, I want to express my utmost gratitude to the Darwin Trust of Edinburgh for funding this research.

Lay summary

DNA double-strand breaks (DSBs) can occur naturally in a living cell during duplication of the genetic material. These DNA breaks must be accurately repaired to ensure cell survival. In the *E. coli* model organism, the repair of a DSB involves copying genetic information from another undamaged DNA template which is identical to the DNA bearing the DSB. This study investigated the duration and localization of the repair of a DSB which was generated at a specific site on the *E. coli* chromosome (genetic material) in a manner that is dependent on duplication of the genetic material. The data indicate that repair was initiated during the period when DNA were held together by physical linkage following replication. As a result, the search for identical DNA sequences within the undamaged DNA template was rapid during repair of the replication-dependent DSB. The subsequent repair events were relatively prolonged, thereby increasing the duration required for separation of the newly replicated DNA. Repair was initiated at the same localization within the cell where the DSB was formed. This study has shown that the physical linkage between newly replicated DNA is essential for eliminating an extensive search for the undamaged DNA template within the entire interior space of the rod-shaped *E. coli* cell during repair of a DSB arising from duplication of the genetic material.

Abbreviations

ATC	Anhydrotetracycline
ATP	Adenosine Triphosphate
BIR	Break-Induced Replication
ChIP	Chromatin Immunoprecipitation
Cm	Chloramphenicol
DAPI	4',6-diamidino-2-phenylindole
D-loop	Displacement-Loop
DNA	Deoxyribonucleic Acid
DSB	Double-Strand Break
dsDNA	Double-Stranded DNA
ECFP	Enhanced Cyan Fluorescence Protein
EDTA	Ethylenediaminetetraacetic acid
EYFP	Enhanced Yellow Fluorescence Protein
FISH	Fluorescence *In Situ* Hybridization
FRET	Fluorescence Resonance Energy Transfer
GFP	Green Fluorescent Protein
Gm^R	Gentamycin Resistance Gene
HR	Homologous Recombination
Km^R	Kanamycin Resistance Gene
LacI	Lac Inhibitor
LacO	Lac Operator
NHEJ	Non-Homologous End Joining

OD_{600}	Optical Density at 600 nm
OriC	Origin of Replication
Pal	246 bp Interrupted DNA Palindrome
PCR	Polymerase Chain Reaction
PMGR	Plasmid Mediated Gene Replacement
SDSA	Synthesis-Dependent Strand Annealing
SSA	Single-Strand Annealing
SSB	Single-Strand DNA-Binding Protein
ssDNA	Single-Stranded DNA
TetO	Tet Operator
TetR	Tet Repressor
TIRF	Total Internal Reflected Fluorescence
UV-light	Ultraviolet Light
v/v	volume per unit volume
w/v	weight per unit volume

Abstract

Chromosomal replication is a source of spontaneous DNA double-strand breaks (DSBs). In *E. coli*, DSBs are repaired by homologous recombination using an undamaged sister template. During repair, the RecA protein polymerizes on single-stranded DNA generated at the site of the DSB and catalyses the search for sequence homologies on the undamaged sister template. This study utilized fluorescence microscopy to investigate the spatial and temporal dynamics of the RecA protein at the site of a replication-dependent DSB generated at the *lacZ* locus of the *E. coli* chromosome. The DSB was generated by SbcCD-mediated cleavage of a hairpin DNA structure formed on the lagging strand template of the replication fork by a long palindromic sequence. The tandem insertion of a *recA-mCherry* gene with the endogenous *recA* gene at the natural chromosomal locus produced no detectable effect on cell viability in the presence of DSB formation. During repair, the fluorescently-labelled RecA protein formed a transient focus, which was inferred to be the RecA nucleoprotein filament at the site of the replication-dependent DSB. The duration of the RecA focus at the site of the DSB was modestly reduced in a Δ*dinI* mutant and modestly increased in a Δ*uvrD* or Δ*recX* mutant. Most cells underwent a period of extended cohesion of the sister *lacZ* loci after disappearance of the RecA focus. Segregation of the sister *lacZ* loci was followed by cell division, with each daughter cell obtaining a copy of the fluorescently-labelled *lacZ* locus. The RecA focus at the site of the DSB was observed predominantly between the mid-cell and the ¼ position. In the absence of DSB formation, the *lacZ* locus exhibited dynamic movement between the mid-cell and the ¼ position until the onset of segregation. Formation of the DSB and initiation of repair occurred at the spatial localization for replication of the *lacZ* locus while the downstream repair events occurred very close to the mid-cell. Genomic analysis of RecA-DNA interactions by ChIP-seq was used to demonstrate that the RecA focus at the *lacZ* locus was generated by the repair of the palindrome-induced DSB and not the repair of one-ended DSBs emanating from stalled replication forks at the repressor-bound operator arrays. This study has shown that the repair of a replication-dependent DSB occurs exclusively during the period of cohesion of the sister loci and the repair is efficiently completed prior to segregation of the two sister loci.

Table of Contents

Tables and Figures

Chapter 1: INTRODUCTION

1.1 DNA double-strand breaks

DNA breaks are physiologically important because they are threats to the integrity and stability of the genome. As a result, all cells have devised mechanisms for repairing this DNA lesion. In-depth understanding of these repair mechanisms is being used for the development of therapeutic strategies against cancerous cells. Moreover, drugs such as quinolones, which are used for the treatment of some bacterial infections, rely on the formation of irrepairable DNA breaks in bacterial species and not the human host. Hence, the insights that are obtained from DNA double-strand break repair mechanisms have useful applications in medical research and drug development.

A DSB refers to the simultaneous breakage of both complementary strands of the double helix at the same site or at different sites that are in very close proximity, thereby causing a defect in the juxtaposition of the two DNA ends. During normal cellular growth, a progressing replication fork generates a DSB at the site of a DNA nick or gap (Kuzminov, 2001). The collision of a replication fork with a DNA-bound protein complex can also result in the formation of a DSB (Michel *et al*., 2001). Stem-loop structures (DNA hairpin and cruciform) that are produced at the site of long DNA palindromic sequences during replication also generate DSBs following the cleavage of these secondary structures by the SbcCD endonuclease (Leach, 1994; Eykelenboom *et al*., 2008). These cellular events indicate that chromosomal replication is a source of DSBs in living cells (Haber, 1999; Pennington and Rosenberg, 2007).

Although DSBs are major threats to genome stability, they are deliberately generated during specific cellular processes for defined biological purposes. In B- and T-cells of higher eukaryotes, the formation of DSBs is particularly relevant for generating diverse immunoglobin and T-cells receptor genes (Fugmann *et al*., 2000). The diverse cell surface receptors that are encoded by these genes ensure that varieties of pathogens are recognized and destroyed by the immune system of these higher organisms. In diploid organisms, the formation of DSBs during meiosis is physiologically important for the occurrence of genetic recombination between homologous chromosomes (Sun *et al*., 1989). Furthermore, site-specific DSBs that are formed at the *MAT* locus of *Saccharomyces cerevisiae* enable haploid cells to switch mating type (Haber, 1998). Switching of mating type allows the formation of

diploids which can undergo meiosis and spore formation under unfavourable growth conditions (Haber, 2012).

All DSBs that are generated in a cell, either deliberately or by accident, must be repaired efficiently to avoid the formation of deleterious mutations or cell death. Mutations that might be generated following the formation of DSBs include inversions, insertions, deletions, duplications or translocations (Cromie *et al*., 2001). These mutations may cause cell death if they inactivate an essential gene such as a DNA repair gene or a tumour suppressor gene. In vertebrates, chromosomal translocation can lead to elevated expression of a proto-oncogene when the gene is relocated to a DNA region that is under the control of a strong promoter (Nambiar *et al*., 2008). De-regulation of proto-oncogenes and deletion of tumour suppressor genes have been implicated in tumorigenesis (Korsmeyer, 1992) thereby emphasizing the need for efficient repair of all the forms of DSBs that are generated within a cell.

1.2 DNA double-strand break repair

Ideally, two-ended DSBs that have blunt or complementary ends could be repaired by direct ligation using a DNA ligase. However, the presence of exonucleases in a cell may result in the degradation of the ends of the DSBs, thereby leading to loss of genetic information. Moreover, DSBs that are generated at the sites of single-strand interruptions or DNA-bound protein complexes during replication have only one free DNA end (Kuzminov, 2001; Michel *et al*., 2001). Repair of DSBs usually occurs by either the non-homologous end joining (NHEJ) pathway or the homologous recombination (HR) pathway. During DSB repair by NHEJ in eukaryotes, the Ku protein binds to both ends of the DSB and facilitates ligation of the DNA ends in the presence of DNA ligase IV (Lieber, 2010). Prior to ligation, nucleases resect the ends of the DSB and polymerase fills the DNA gaps that are generated (Lieber, 2010). Degradation of the ends of the DSB by exonucleases prior to repair, or the occurrence of resection during NHEJ, result in the formation of a deletion mutation following ligation of the ends of the DSB. Although repair by NHEJ can produce mutations at the site of the DSB, the consequences of these mutations are often benign in mammalian cells because of the presence of numerous non coding regions in the genome. Additionally, if the rare scenario of generating a mutation in the coding sequence of an essential gene occurs, the presence of an intact allele on the homologous chromosome may compensate for the defective allele (Pfeiffer *et al*., 2000). DSB repair by NHEJ is predominant during the G0 or G1 phases of the cell cycle (Rothkamm *et al*., 2003). In mammalian cells, DSBs that are generated during the G1 phase of the cell cycle are mainly repaired by NHEJ while DSBs generated in cells during late S/G2 phase are

occasionally repaired by NHEJ (Takashima *et al*., 2009). Even though NHEJ was originally discovered in eukaryotes, homologues of the components of this repair pathway have also been identified in some prokaryotes (Weller *et al*., 2002).

In addition to NHEJ, DSB repair can occur via pathways that rely on the presence of significant sequence homology. One form of such homology-directed repair mechanism is single-strand annealing (SSA), which can be detected at chromosomal loci that have repeated DNA sequences flanking the site of the DSB (van den Bosch *et al*., 2002). During repair via the SSA pathway, resection of the ends of the DSB, followed by annealing of the complementary repeat sequences that are adjacent to the site of the DSB, leads to the loss of at least one of the repeats and the intervening DNA sequences (Figure 1.1A). Thus, DSB repair by SSA generates a deletion mutation at the site flanking the DNA damage.

Another form of homology-directed DSB repair is the synthesis-dependent strand annealing pathway (SDSA). The SDSA repair mechanism requires a homologous donor template for restoring the DNA sequences that are lost at the site of the DSB (McGill *et al*., 1989; Nassif *et al*., 1994). The initial stages of SDSA involve resection of the ends of the DSB, homologous pairing, strand invasion and the priming of DNA synthesis from the 3' end of the invading single-strand (Figure 1.1B; SDSA model). Following DNA synthesis, the invading single-strand is displaced from the donor template and re-anneals to the other end of the DSB. Unlike NHEJ and SSA, the SDSA pathway does not generate a deletion mutation at the site of the DSB. The canonical model for recombinational repair of two-ended DSBs is shown in Figure 1.1B (Canonical DSBR model). According to this model, the invading single strand is not displaced from the donor template following DNA synthesis. However, the 3' terminal DNA from the other end of the DSB also interacts with the displaced donor template to prime DNA synthesis and seal the break. The resolution of the four-way junctions result in the formation of two intact duplex DNA.

One-ended DSBs are repaired exclusively by homologous recombination in a process referred to as Break-induced replication (BIR). The repair of one-ended DSBs involves extensive DNA synthesis after strand invasion (Figure 1.1B; BIR). Due to the absence of the other end of the DSB, DNA synthesis that is initiated at the site of the DSB progresses till the end of the chromosome (Kogoma, 1997; Morrow *et al*., 1997). In *E. coli*, DSBs are repaired exclusively by homologous recombination (Kowalczykowski *et al*., 1994) using an undamaged sister

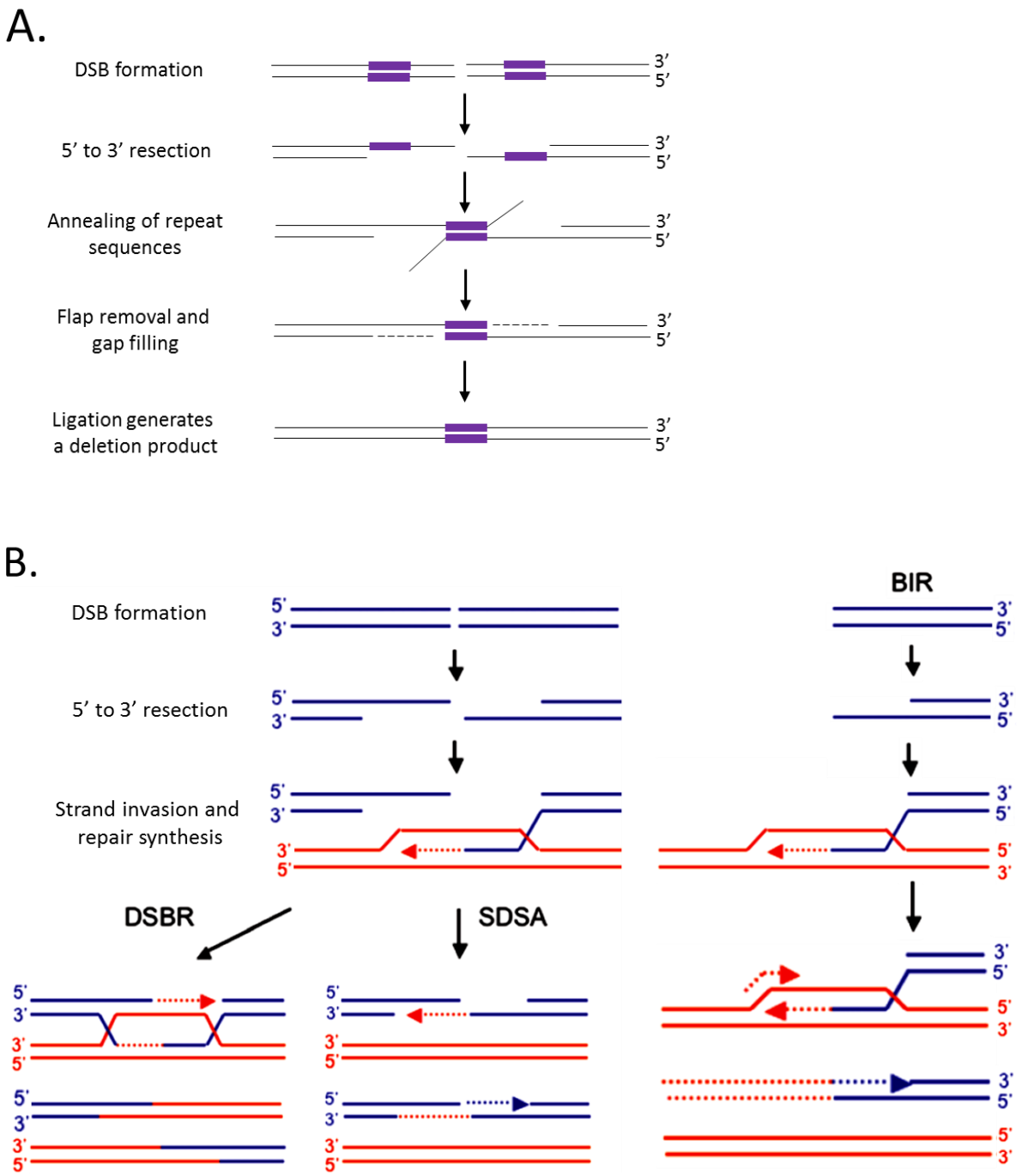


Figure 1.1: Models for homology-directed DSB repair pathways. (A) DSB repair by single-strand annealing (SSA). Resection of the ends of the DSB, followed by annealing of the complementary repeat sequences that are adjacent to the site of the DSB lead to the formation of a deletion mutation. Repeats sequences are shown in purple. (B) Models of DSB repair via mechanisms that require strand invasion. SDSA model: After strand invasion and DNA synthesis, the 3' terminal strand is displaced from the D-loop and re-anneals to the other end of the DSB. BIR model: Strand invasion is followed by extensive DNA synthesis from the 3' end of the invading strand due to the absence of the other end of the DSB. Canonical DSBR model: DNA synthesis from the initial invading 3' terminal strand, and the capture of the other 3' end within the D-loop results in the formation of double Holliday junctions, which are resolved to generate two intact duplex DNA. The figure 1.1B was modified from (Llorente *et al.*, 2008).

template. As a result, *E. coli* is a useful model organism for understanding DSB repair that occurs via the homologous recombination pathway.

1.3 DNA double-strand break repair in *E. coli*

DSB repair in *E. coli* is classified into three steps: (i) processing the ends of the DSB (Pre-synapsis), (ii) homologous pairing and strand exchange (Synapsis), (iii) DNA synthesis, branch migration and resolution of the heteroduplex DNA (Post-synapsis).

1.3.1 Pre-synapsis

The initial stages of homologous recombination are necessary for generating a suitable substrate onto which the RecA protein binds and polymerizes to form a nucleoprotein filament (Wyman *et al*., 2004). During pre-synapsis, the ends of the DSBs are recognized and bound by the RecBCD complex which has both helicase and exonuclease activities (Figure 1.2; Dillingham and Kowalczykowski, 2008; Smith, 2012). These activities unwind the duplex DNA at the site of the DSB and degrade the single-strand DNA (ssDNA) that are generated (Dillingham and Kowalczykowski, 2008). Degradation of the ssDNA is asymmetric because it is continuous and rapid on the 3' terminal strand while the 5' terminal strand is intermittently cleaved (Dixon and Kowalczykowski, 1993). Meanwhile, the RecC subunit of the enzyme complex scans the 3' ssDNA for the presence of an 8-nucleotide Chi sequence (5'-GCTGGTGG-3'; Handa *et al*., 2012). After detection of a correctly-oriented Chi sequence, degradation of the 3' ssDNA is attenuated without an effect on the helicase activity of the enzyme complex. These events lead to the formation of a 3' ssDNA overhang, which is the suitable substrate for RecA-catalysed invasion of the homologous donor template (Dixon and Kowalczykowski, 1991; Dixon and Kowalczykowski, 1993). Recognition of the Chi sequence is very crucial for regulating the inherently destructive capacity of the RecBCD complex while enhancing its role in genetic recombination (Dixon and Kowalczykowski, 1991; Dixon and Kowalczykowski, 1993).

An alternative model suggests that the RecBCD complex unwinds the duplex DNA without exhibiting an exonuclease activity during pre-synapsis (Smith, 2012). Upon encountering and interacting with a correctly-oriented Chi sequence, RecBCD generates a nick at the Chi site on the 3' ssDNA. The continuous unwinding of the duplex DNA beyond the Chi sequence results in the formation of a 3' ssDNA which terminates at the Chi sequence and serves as the template for the formation of the RecA nucleoprotein filament. Thus, in this model the endonuclease activity of RecBCD is utilized for the formation of the pre-synaptic filament.

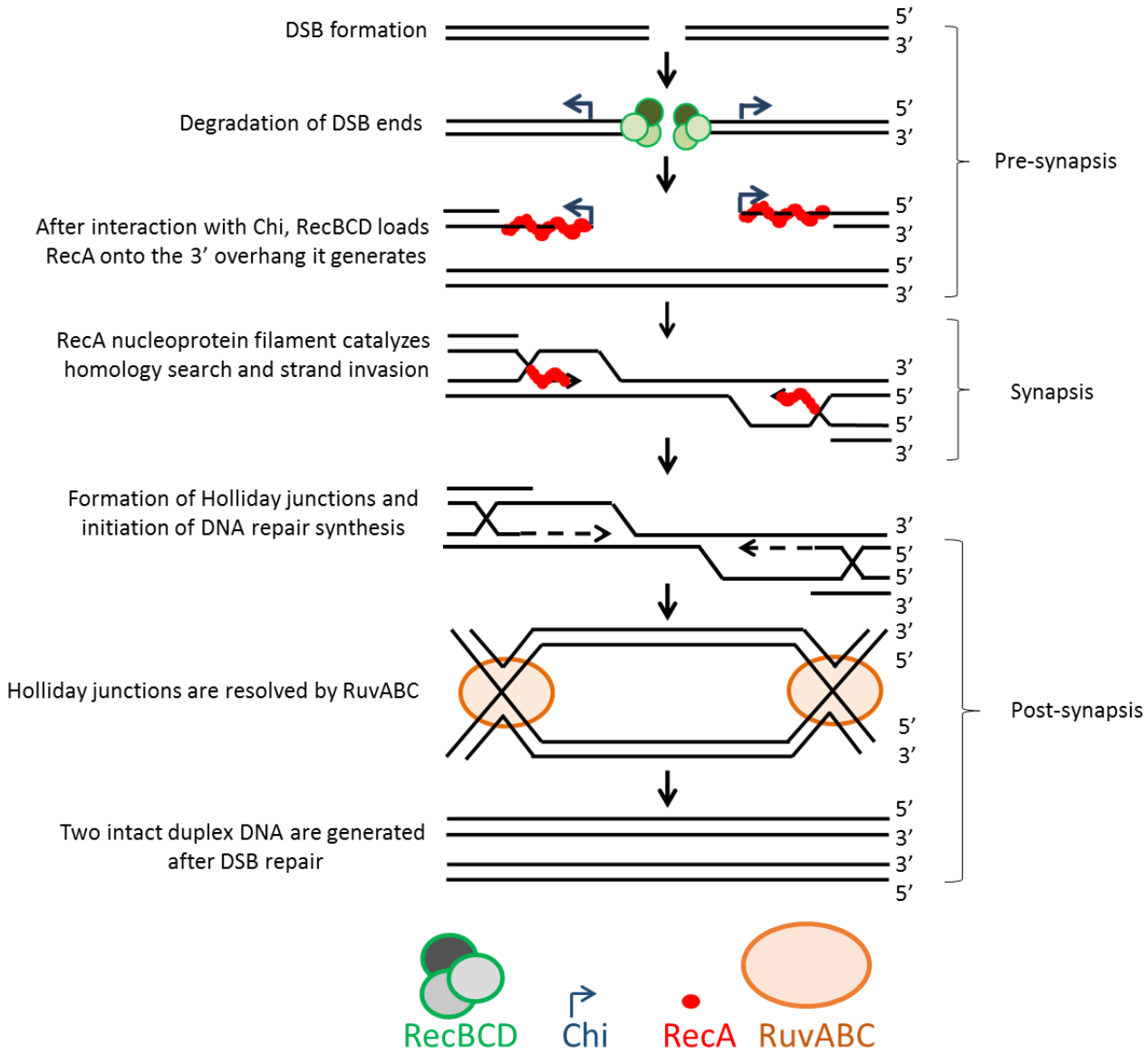


Figure 1.2: Repair of two-ended DSBs by homologous recombination in *E. coli*. After DSB formation, RecBCD processes both ends of the DSB to generate 3' ssDNA overhangs. The loading of RecA onto the 3' ssDNA overhangs leads to the formation of RecA nucleoprotein filaments, which catalyze homology search and strand invasion. The 3' end of the invading ssDNA primes DNA repair synthesis and Holliday junctions that are produced during the repair of DSBs are resolved by the RuvABC complex.

In addition to processing the ends of the DSB, the RecBCD complex loads RecA onto the 3' ssDNA. The active loading of RecA by the RecBCD enzyme is essential for overcoming the competitive binding of the single-stranded DNA-binding protein (SSB) to the 3' ssDNA overhang (Anderson and Kowalczykowski, 1997). Thereafter, the RecA protein polymerizes on the ssDNA preferentially in the 5' to 3' direction, which ultimately results in the formation of a RecA nucleoprotein filament (Figure 1.2; Register and Griffith 1985). During the final phase of pre-synapsis, the RecBCD enzyme dissociates from the DNA and is disassembled into its three inactive subunits (Taylor and Smith, 1999).

1.3.2 Synapsis and post-synapsis

The nucleoprotein filament that is generated during pre-synapsis represents the active form of RecA during DSB repair (Lusetti and Cox, 2002). During synapsis, the RecA nucleoprotein filament interacts with the undamaged sister template (Tsang *et al*., 1985) and performs an efficient search for DNA sequences that are homologous to the DNA region adjacent to the DSB (Kowalczykowski *et al*., 1994). After identifying sequence homologies, the 3' terminal end of the nucleoprotein filament invades the homologous region of the undamaged sister template, leading to the formation of a displacement (D-) loop (Figure 1.2; Szostak *et al*., 1983). During post-synapsis, the 3' end of the invading strand primes DNA repair synthesis (Szostak *et al*., 1983), which restores the DNA sequences that were lost at the site of the DSB. Pairing of the ssDNA at the 5' end with the D-loop generates a Holliday Junction (Kowalczykowski *et al*., 1994) which is stabilised by branch migration (Mawer and Leach, 2014). The completion of repair synthesis and the resolution of the Holliday junction by the RuvABC enzyme complex result in the formation of two intact duplex DNA that are separated from each other.

1.4 The RecA protein

The RecA protein has a molecular mass of approximately 38 kDa and consists of 352 amino acids. It is found in most bacteria (Brendel *et al*., 1997; Roca and Cox, 1990; Roca and Cox 1997), and has a highly conserved polypeptide chain (Lusetti and Cox, 2002). Structural and functional homologues of RecA have also been identified in eukaryotes (Ogawa *et al*., 1993; Sung, 1994). In *E. coli*, RecA catalyzes the pairing of homologous DNA and the strand exchange reactions which occur during genetic recombination and repair of DNA damage by homologous recombination (Kowalczykowski & Eggleston, 1994; Roca & Cox, 1997). Additionally, RecA is required for the induction of the SOS response following the formation of DNA damage in *E. coli*. (Roca & Cox, 1997).

Recombination and induction of the SOS response are dependent on the formation of the RecA nucleoprotein filament. The first step during the formation of the RecA nucleoprotein filament is the nucleation of RecA onto the DNA (Pugh and Cox, 1988; Pugh and Cox 1987). Nucleation of RecA involves the binding of a monomeric RecA protein onto the DNA (Figure 1.3). Even though RecA can bind to both ssDNA and dsDNA, the process of nucleation is faster on ssDNA compared to dsDNA (Lusetti and Cox, 2002). After nucleation, the binding of subsequent monomeric RecA occurs cooperatively and rapidly in the 5' to 3' direction using ATP (Kowalczykowski & Eggleston, 1994; Register and Griffith 1985), whereas the disassembly of the RecA nucleoprotein filament proceeds unidirectionally from the 5' end (Figure 1.3; Shan *et al*., 1997).

In the presence of ATP, the polymerization of RecA monomers on ssDNA results in the formation of a right-handed helical filament, which represents the active form of the protein (the top image of Figure 1.4; Lusetti and Cox, 2002). Electron micrographs of the RecA-DNA complex indicate that when RecA monomers are bound to dsDNA, the nucleoprotein filament is elongated by approximately 60% and is underwound relative to the B-form helix of the DNA alone (Stasiak and Di Capua, 1982). The elongated dsDNA within the nucleoprotein filament consist of 18.6 bp and 6.4 monomers of RecA per turn of the helix (Stasiak *et al*., 1981; Di Capua *et al*., 1982). Recently, the crystal structure of the *E. coli* RecA protein bound to either ssDNA or dsDNA have confirmed that each RecA monomer covers three nucleotides in the nucleoprotein filament (Chen *et al*., 2008). The arrangement of the RecA monomers within the nucleoprotein filament and the number of nucleotides per turn of the helix suggest that the dsDNA is partially exposed and thus can interact with an external DNA (Stasiak and Di Capua, 1982; Egelman and Stasiak, 1986), thereby facilitating the homology search and strand exchange reactions. Moreover, the partial elongation of the DNA within the nucleoprotein filament has been suggested to be consistent with an intercalating mode of binding of RecA (Stasiak *et al*., 1981). In the absence of either single-stranded or double-stranded DNA, monomers of RecA can form aggregate structures (Brenner *et al*., 1988).

The crystal structure of RecA has revealed that the protein consists of three distinct domains: a bigger core domain, which is flanked by smaller domains at the amino (N-) and carboxyl (C-) termini of the polypeptide (the middle image of Figure 1.4; Story *et al*., 1992; Story and Steitz, 1992; Datta *et al*., 2000). The core domain ranges from residues 33 to 269 in *E. coli*, and is the most highly conserved region of the polypeptide across the other bacterial homologues (Lusetti and Cox, 2002).

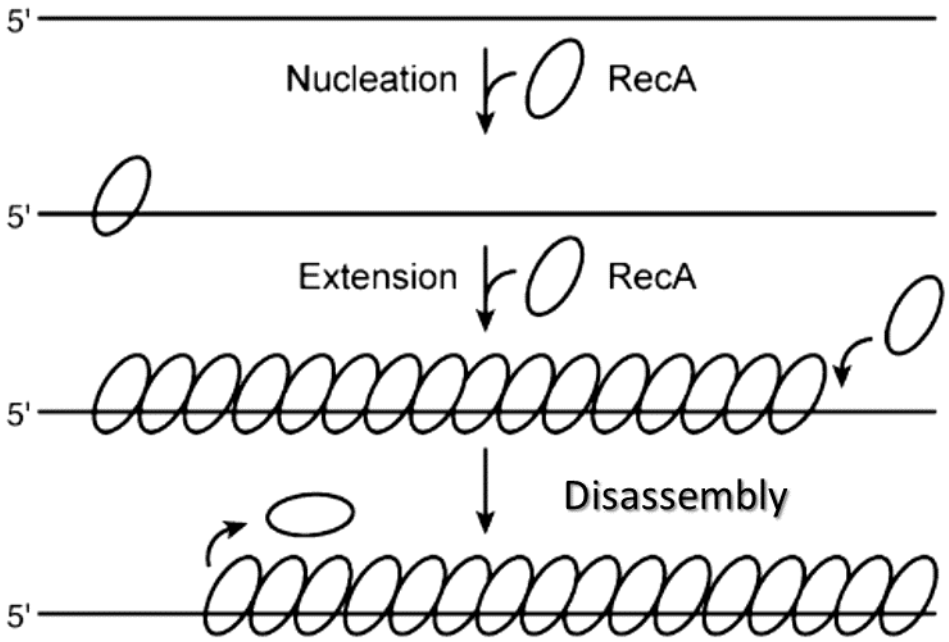


Figure 1.3: Formation and disassembly of the RecA nucleoprotein filament. Nucleation of RecA monomers onto the DNA is the rate-determining step during the formation of the nucleoprotein filament. After nucleation, polymerization of RecA on the DNA is rapid and occurs preferentially in the 5' to 3' direction. Disassembly of the nucleoprotein filament is initiated from the terminal RecA at the 5' end. The figure 1.3 was modified from Drees *et al*., 2004

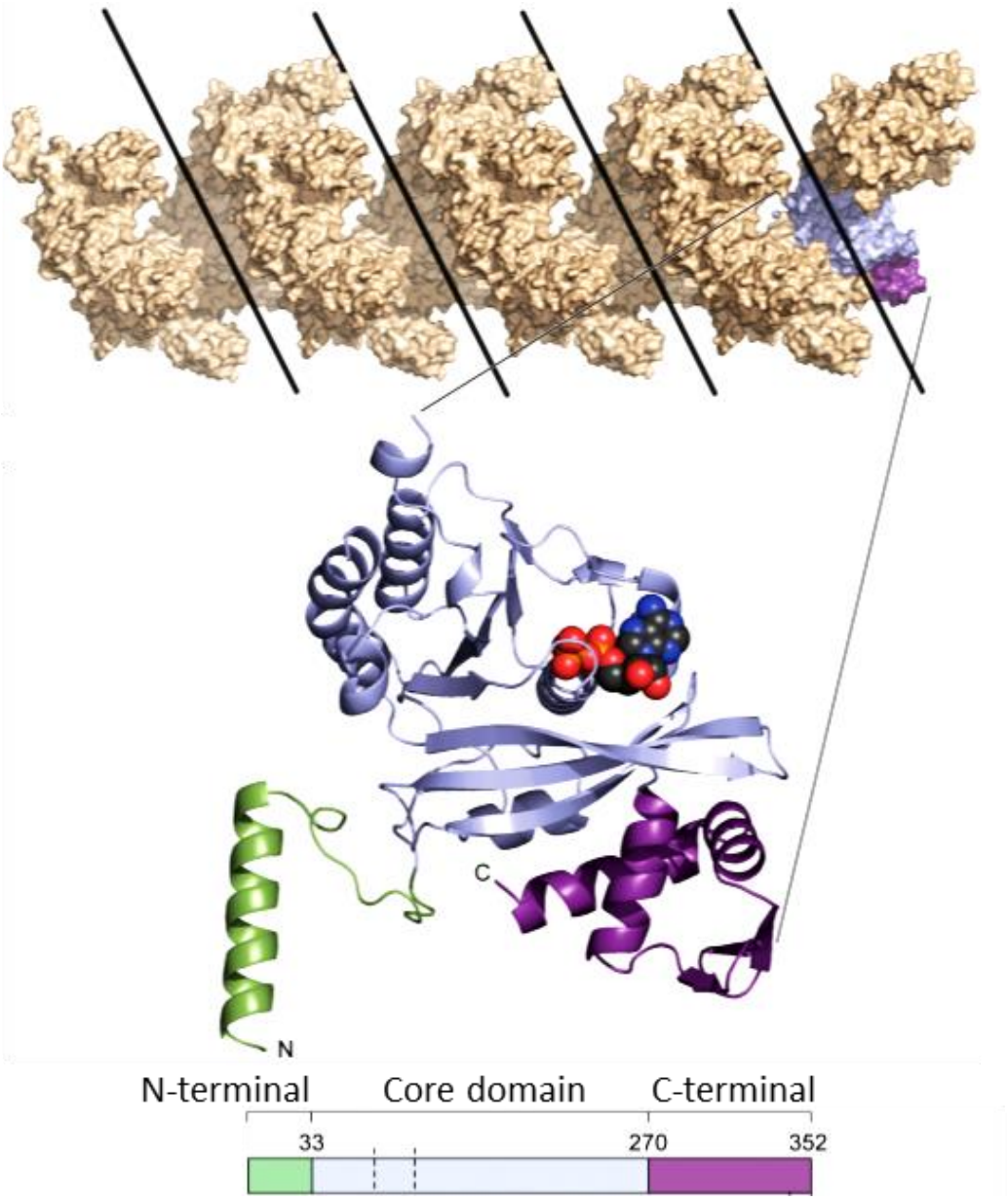


Figure 1.4: Structure of the *E. coli* RecA protein. Top: A RecA nucleoprotein filament containing 24 monomers of the RecA protein. The nucleoprotein filament has four grooves, indicated with black lines. A monomeric RecA is shown below the nucleoprotein filament. Middle: Structure of a monomeric RecA protein. The N-terminal, Core and C-terminal domains of the polypeptide are shown in green, blue and magenta respectively. ADP is shown as a space-filling model within the core domain of the polypeptide. Bottom: Schematic representation of each domain of the polypeptide. N-terminal domain (residues 1 to 32); Core domain (residues 33 to 269); C-terminal domain (residues 270 to 352). The dash lines represent the binding sites for ATP. This figure was modified from Cox, 2007.

The crystal structure also revealed two disordered regions (residues 157 to 164; residues 195 to 209) within the core domain which have been identified as loops. Studies that were performed using protein-DNA crosslinking (Malkov and Camerini-Otero, 1995) have demonstrated that these loops are binding sites for ssDNA and dsDNA. Specifically, the first loop (residues 157 to 164) has been proposed to be the binding site for the incoming homologous dsDNA, whereas the second loop (residues 195 to 209) is the binding site for the ssDNA during the formation of the nucleoprotein filament (Story *et al*., 1992). Moreover, a Walker A box (GPESSGKT; residues 66 to 73) and a Walker B box (VIVVD; residues 140 to 144), which are binding sites for nucleoside triphosphates, are also located within the core domain of the RecA polypeptide chain (Rehrauer and Kowalczykowski, 1993; Shan *et al*., 1996; Lusetti and Cox, 2002). *In vitro* studies have shown that ATP binding alone, without hydrolysis, is required for the formation of the nucleoprotein filament (Rehrauer and Kowalczykowski, 1993; Shan *et al*., 1996). However, binding and hydrolysis of ATP are both essential for the occurrence of Hfr-mediated recombination and efficient repair of DNA damage, as well as induction of the SOS response, following UV irradiation in *E. coli* (Renzette and Sandler, 2008).

The smaller domain at the N-terminal of the polypeptide consists of residues 1 to 32 and is responsible for interactions between neighbouring monomers of the RecA protein in the nucleoprotein filament (Story *et al*., 1992; Lusetti and Cox, 2002). Interactions amongst the neighbouring RecA monomers have been proposed to stabilize the formation of the nucleoprotein filament (Story *et al*., 1992). The second smaller domain, which is found at the C-terminal of the polypeptide, spans from residues 270 to 352 and has a regulatory role during the formation of the nucleoprotein filament (Eggler *et al*., 2003). *In vitro* studies have also shown that the C-terminal domain of RecA modulates the accessibility of external dsDNA to the RecA nucleoprotein filament, thereby modulating the occurrence of homologous pairing and DNA strand exchange (Lusetti *et al*., 2003).

1.5 Induction of the SOS response by the RecA nucleoprotein filament

The SOS response is an inducible mechanism that allows bacteria to repair DNA damage (Figure 1.5; Michel, 2005). The LexA repressor protein and the RecA nucleoprotein filament are the two elements that regulate induction of the SOS response (Friedberg *et al*., 1995). In the absence of DNA damage during cellular growth, LexA suppresses the expression of the

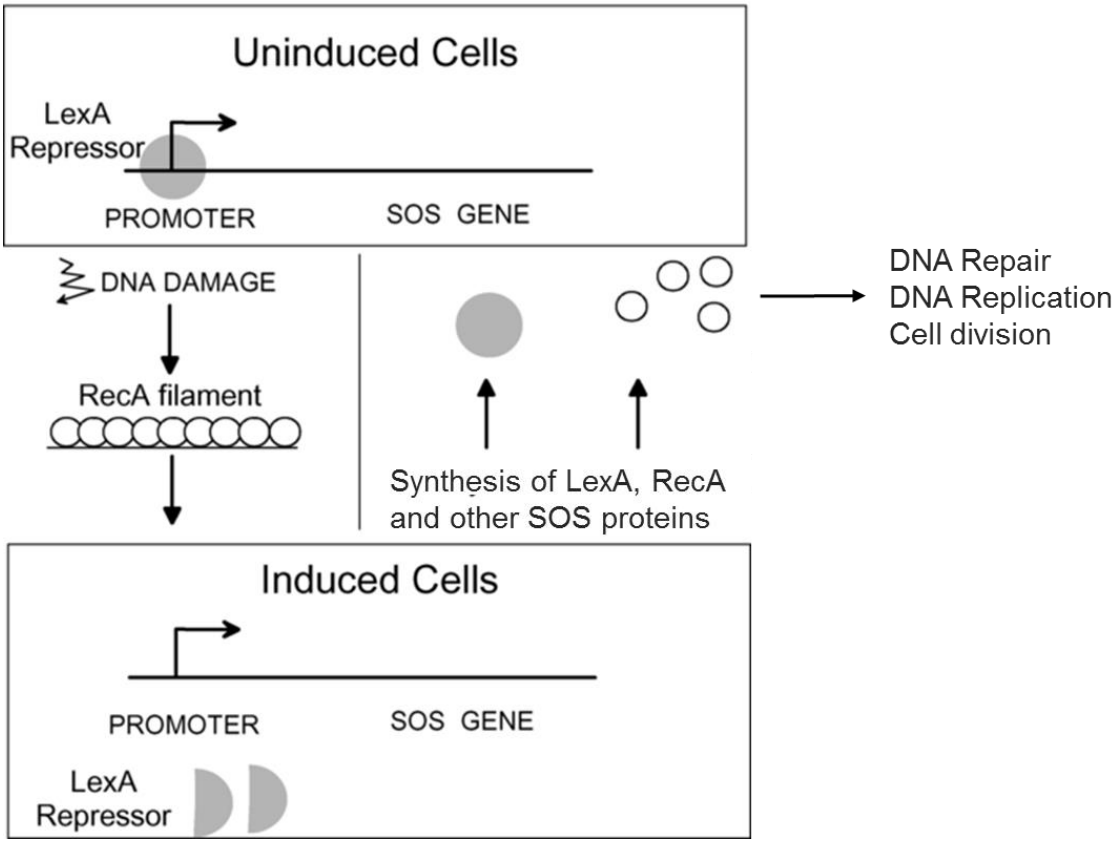


Figure 1.5: The SOS response in *E. coli*. The LexA repressor binds to the SOS box within the promoter region of the SOS genes in the absence of DNA damage. Following the formation of a DNA damage, RecA polymerizes on ssDNA and the RecA nucleoprotein filament stimulates auto-cleavage of the LexA repressor which leads to the expression of the SOS genes including *lexA* and *recA*. The synthesis of the SOS proteins enables the cell to repair the DNA damage. Induction of the SOS response might also affect other cellular processes such as DNA replication and cell division. After repair of the DNA damage and disassembly of the RecA nucleoprotein filament, the LexA protein accumulates and represses the expression of the SOS-inducible genes. This figure was modified from Michel, 2005.

SOS inducible-genes by binding to a 20bp sequence, referred to as the SOS box, which is located within the promoter region of these genes (Brent and Ptashne 1981; Little *et al*., 1981). The degree of repression of each SOS gene is dependent on the nature of the sequence of its SOS box, the strength of the promoter and the position of the SOS box within the promoter region (Michel, 2005). Consequently, the basal level of expression is variable amongst the SOS genes in the absence of SOS induction during normal cellular growth. For example, there are approximately 8,000 molecules of RecA (Boudsocq *et al*., 1997) and an undetectable amount of Pol V during basal gene expression from their respective promoters (Michel, 2005).

When a DNA damage is detected in the cell, expression of the SOS genes is upregulated due to auto-proteolysis of the LexA repressor. Auto-proteolysis of LexA is stimulated by the RecA nucleoprotein filament (Little, 1984). Thus, the RecA nucleoprotein filament exhibits a co-protease activity during induction of the SOS response. The requirement of the RecA nucleoprotein filament for induction of the SOS response indicates that ssDNA or DNA gaps are generated either directly or as a consequence of some forms of DNA damage (Patel *et al*., 2010).

Following auto-proteolysis of LexA, the expression of at least 40 genes is upregulated, including the *recA* and *lexA* genes (Fernandez de Henestrosa *et al*., 2000; Courcelle *et al*., 2001). The proteins encoded by most of the SOS genes are involved in either the repair of the DNA damage, DNA replication or the control of cell division (Friedberg *et al*., 1995; Koch and Woodgate 1998). The function of some of the genes that are upregulated during SOS induction are yet to be discovered (Michel and Leach, 2012). The expression of all the genes that comprises the SOS regulon is not upregulated simultaneously. Genes encoding proteins involved in nucleotide excision repair, base excision repair and recombinational DNA repair are among the first set of genes that are rapidly upregulated by the SOS response (Friedberg *et al*., 2006). Upregulation of the expression of the *sfiA* gene, which encodes an inhibitor of cell division, ensures that repair of the DNA damage is complete before the commencement of cell division (Michel, 2005). If the DNA lesion is not repaired via an error-free pathway, expression of proteins involved in translesion DNA synthesis is upregulated (Goodman, 2002). Translesion DNA synthesis generates several mutations, some of which might be lethal to the cell (Patel *et al*., 2010). Thus, the *umuD* and *umuC* genes (encoding the mutagenic DNA repair polymerase V) are amongst the last set of genes to be induced by the SOS response (Sommer *et al*., 1993). After expression, the UmuD protein dimerizes, undergoes auto-cleavage of the 24 amino acids at the N-terminal of each polypeptide and interacts with the UmuC protein to

generate the stable protein complex, UmuD'$_2$C (Woodgate *et al*., 1989; Bruck *et al*., 1996), which was identified as Pol V (Tang *et al*., 1999). Auto-cleavage of UmuD is facilitated by the RecA nucleoprotein filament and its occurrence is less efficient compared to the auto-cleavage of LexA (Burckhardt *et al*., 1988). These events demonstrate that SOS mutagenesis is tightly regulated and is relegated to the later stages of the SOS response when the DNA damage is severe and persistent (Patel *et al*., 2010). The *lexA* gene is constantly expressed during the SOS response, thereby ensuring that the expression of the SOS genes is repressed immediately after repair of the DNA damage (Michel, 2005). It has been suggested that the SOS response is also controlled by modulation of the stability of the RecA nucleoprotein filament (Yasuda *et al*., 1998; Lusetti *et al*., 2004). Yasuda *et al*., (1998) has also reported that the upregulation of the expression of the DinI protein may play a role in the repression of the SOS response by inhibiting the co-protease activity of RecA.

1.6 Regulation of the activities of RecA during recombination

Expression of the *recA* gene is upregulated by 10-fold in response to the formation of a DNA damage (Boudsocq *et al*., 1997; Sommer *et al*., 1998). The activities of RecA are also subject to autoregulation by the C-terminal region of the polypeptide (Benedict and Kowalczykowski, 1988; Eggler *et al*., 2003; Cox, 2007). Additionally, other proteins have been shown to regulate the activities of RecA by either inhibiting the nucleation and polymerization of RecA on the DNA or enhancing the stability of the nucleoprotein filament, which is the active form of the RecA protein (Cox, 2007). Ultimately, these modulators affect (either directly or indirectly) the pairing of homologous DNA and the DNA strand exchange reactions, which are the core functions of RecA during homologous recombination.

1.6.1 Autoregulation of the activities of RecA by the C-terminal of the polypeptide

The last 17 amino acids at the C-terminal end of the RecA polypeptide consist predominantly of negatively-charged residues (Aspartic acid and Glutamic acid). *In vitro* studies have shown that, by deleting this region of the polypeptide, the truncated RecA protein exhibited an inherent ability to displace single-stranded DNA-binding protein (SSB) from ssDNA during the formation of the RecA nucleoprotein filament (Eggler *et al*., 2003). Due to the high affinity of the truncated RecA protein for ssDNA, its expression also caused a constitutive induction of the SOS response even in the absence of a DNA damage (Tateishi *et al*., 1992). Because the untruncated RecA protein is loaded onto ssDNA by mediator proteins in the presence of SSB, it has been suggested that the C-terminus of RecA is a negative regulator of the RecA nucleoprotein filament (Eggler *et al*., 2003). *In vitro* studies have also shown that the presence

of the C-terminal region of RecA results in the formation of a nucleoprotein filament which is in a ‘‘closed’’ conformation, thereby restricting the access of incoming homologous dsDNA to the nucleoprotein filament (Lusetti *et al*., 2003). Furthermore, the study demonstrated that the ‘‘closed’’ conformation of the nucleoprotein filament changes into an ‘‘open’’ conformation in the presence of excess Mg^{2+} or after deletion of the utmost 17 amino acids at the C-terminal of the RecA polypeptide (Lusetti *et al*., 2003). These observations have been used to infer that the amino acids at the C-terminal of the RecA polypeptide has a regulatory effect on the activities of the RecA nucleoprotein filament (Cox, 2007).

1.6.2 Regulation of the stability of the RecA nucleoprotein filament by other proteins

1.6.2.1 The RecBCD complex

The loading of RecA monomers onto ssDNA during the formation of the nucleoprotein filament is primarily accomplished by either the RecBCD or the RecFOR complex (Kowalczykowski *et al*., 1994). The function of these protein complexes is to displace SSB and simultaneously facilitate the assembly of RecA onto the ssDNA (Cox, 2007). Interestingly, *in vitro* studies have demonstrated that the presence of SSB stimulates the formation of the RecA nucleoprotein filament by destabilizing ssDNA secondary structures (Kowalczykowski *et al*., 1987; Kowalczykowski and Krupp, 1987) that might accumulate after unwinding of dsDNA by RecBCD (Anderson and Kowalczykowski, 1998). Additionally, the strand exchange reaction that is catalysed by the RecA nucleoprotein filament is enhanced when SSB binds to the displaced DNA strand in the joint molecule (Lavery and Kowalczykowski, 1992). These observations demonstrate that the presence of SSB is essential for the occurrence of efficient recombination. It is essential for RecBCD to displace SSB from the ssDNA and actively load RecA at pre-synapsis (Anderson and Kowalczykowski, 1998) to ensure that recombinational repair of DSBs is not inhibited by the presence of SSB. Thus, the dynamic interplay between SSB and RecBCD during the formation of the RecA nucleoprotein filament is intricate.

1.6.2.2 The RecFOR complex

The RecF, RecO and RecR proteins form a complex which loads RecA onto ssDNA during recombinational repair of DNA gaps (Morimatsu and Kowalczykowski, 2003). In the absence of RecF, SSB proteins that are bound to ssDNA are displaced by RecA with the help of the RecO and RecR proteins (Umezu *et al*., 1993; Umezu and Kolodner 1994). Mutations that inactivate any of the *recFOR* genes result in a delay in the induction of the SOS response due to a slower rate of assembly of RecA onto the ssDNA (Madiraju *et al*., 1988; Whitby and Lloyd, 1995). The above-mentioned *recFOR* phenotype is also detected in *E. coli* strains in which

SSB is overexpressed (Moreau, 1988). Therefore the role of the RecFOR complex during DNA gap repair is to facilitate the binding of RecA onto SSB-coated ssDNA. Other proteins that modulate the formation and stability of the presynaptic filament are UvrD, RecX and DinI.

1.6.2.3 The UvrD protein

The UvrD protein is a multifunctional enzyme: it plays a role in nucleotide excision repair (Runyon *et al*., 1990), methyl-directed mismatch repair and recombination (Matson, 1991). UvrD is a helicase and a translocase with a 3' to 5' polarity that displays optimal activity when the dsDNA has a 3' overhang (Matson and George, 1987; Cox, 2007). *In vitro* studies have demonstrated that at least a dimeric form of the UvrD protein is required for processivity during unwinding of the dsDNA (Ali *et al*., 1999; Maluf *et al*., 2003). It was later shown that individual monomers of UvrD can only translocate along ssDNA but are unable to unwind dsDNA, *in vitro* (Fischer *et al*., 2004). During recombinational repair, UvrD facilitates the progression of repair forks through protein-DNA barriers such as the Tus-*Ter* block, which is usually present at the terminus of the *E. coli* chromosome (Bidnenko *et al*., 2006). Additionally, UvrD has a putative antirecombinase activity by which it efficiently inhibits homologous pairing (Morel *et al*., 1993). *In vitro* data from a study that was conducted using electron microscopy have revealed that UvrD is capable of disassembling the RecA nucleoprotein filament (Veaute *et al*., 2005). This study also demonstrated that the translocase activity, and not the helicase activity of UvrD, is responsible for the disassembly of RecA from ssDNA (Veaute *et al*., 2005). Data from *in vivo* studies have been used to postulate that RecA is loaded at the site of blocked replication forks via the RecFOR pathway (Florés *et al*., 2005). In such circumstances, the authors demonstrated that the presence of the UvrD protein is essential for cell viability and the requirement of UvrD is eliminated upon inactivation of the RecFOR pathway. These observations have been used to propose that the translocase activity of UvrD is vital for the removal of RecA at the site of a blocked replication fork, in order to facilitate the restart of replication via replication fork reversal (Florés *et al*., 2005).

1.6.2.4 The RecX protein

The *recX* gene is located immediately downstream of the *recA* gene and is conserved in many bacteria (Sano, 1993; De Mot *et al*., 1994). It has been shown that expression of the *recX* gene occurs by transcriptional read-through from the *recA* promoter in *E. coli* (Pages *et al*., 2003). As a result, expression of the *recX* gene is also upregulated in response to the formation of a DNA damage (Stohl *et al*., 2003). Due to the presence of a transcription terminator sequence between the *E. coli recA* and *recX* genes, expression of *recX* from the *recA* promoter is minimal

(Pages *et al*., 2003). Even though a null *recX* mutant showed no distinct phenotype (Pages *et al*., 2003), overexpression of RecX resulted in a decrease in the induction of the SOS response (Stohl *et al*., 2003). Yeast two-hybrid analysis has been used to demonstrate that the RecA and RecX proteins interacts physically (Stohl *et al*., 2003). *In vitro,* sub-stoichiometric amount of RecX inhibited DNA strand exchange, thereby indicating that RecX acts as a negative modulator of the activities of RecA (Stohl *et al*., 2003). From these observations, the authors proposed that the RecX protein functions as an antirecombinase by binding to free monomeric RecA, thereby minimising the affinity of the latter for ssDNA during the formation of the RecA nucleoprotein filament (Stohl *et al*., 2003).

Electron microscopy has been used to demonstrate that the presence of the RecX protein resulted in partial disassembly of the RecA nucleoprotein filament (Drees *et al*., 2004). The authors also showed that very low concentrations of RecX relative to RecA are sufficient for inhibiting hydrolysis of ATP by the RecA nucleoprotein filament (Drees *et al*., 2004). ATP hydrolysis is usually used as a functional assay of the RecA nucleoprotein filament because the rate of ATP hydrolysis correlates significantly with the amount of RecA that is bound to the DNA (Pugh and Cox, 1988). From these observations, Drees *et al*., (2004) inferred that a physical interaction between RecX and free monomers of RecA could not be attributed to the partial disassembly of the RecA nucleoprotein filament by the RecX protein. An alternative model was suggested, indicating that RecX binds to the terminal RecA at a gap within the nucleoprotein filament and inhibits the subsequent binding of other RecA monomers at the gap (Figure 1.6; Drees *et al*., 2004). The inability of RecA to bind to the ssDNA at the site of the gap, due to the presence of RecX, facilitates the net disassembly of the nucleoprotein filament, thereby implicating RecX as an antirecombinase protein (Cox, 2007). Thus, like UvrD, the presence of RecX results in a decrease in the stability of the RecA nucleoprotein filament.

1.6.2.5 The DinI protein

Expression of the *dinI* gene is upregulated in response to the formation of a DNA damage (Lewis *et al*., 1994). *In vitro* studies have revealed that the concentration of DinI, relative to the concentration of RecA, determines the effect of the DinI protein on the stability of the RecA nucleoprotein filament (Lusetti *et al*., 2004). Thus, DinI stabilizes the RecA nucleoprotein filament when the concentration of DinI is either below, the same as, or moderately above the concentration of RecA in the experimental system (Lusetti *et al*., 2004).

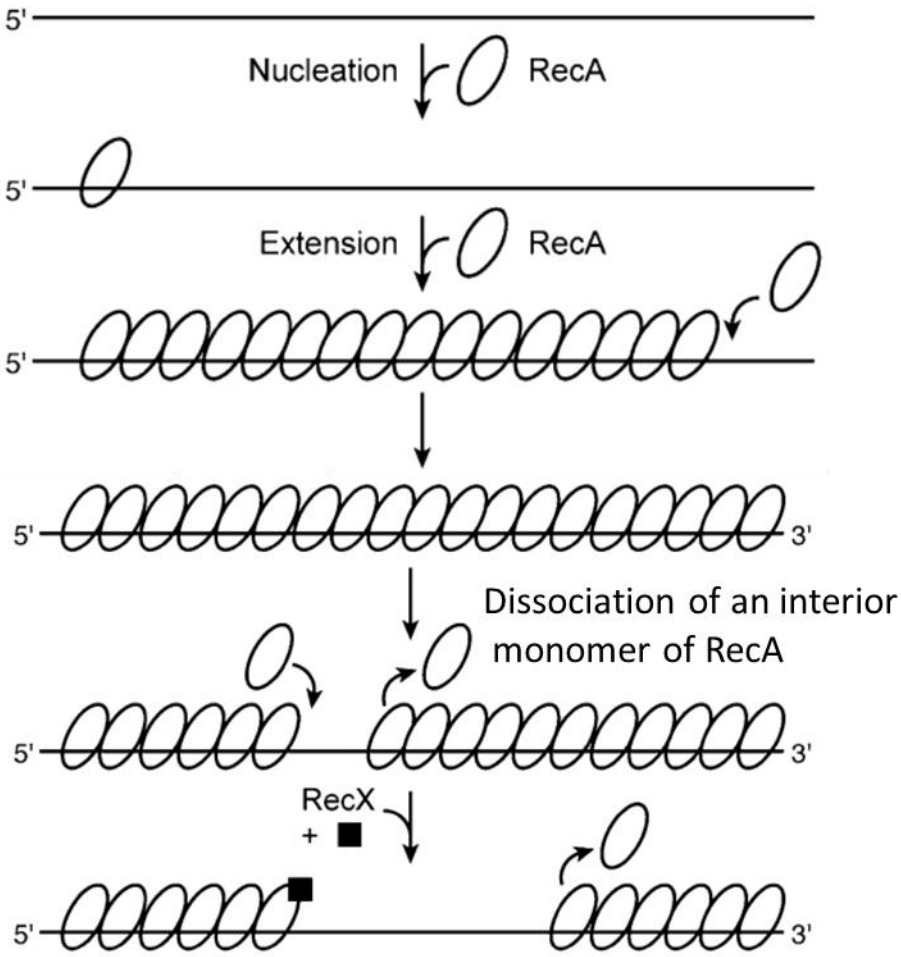


Figure 1.6: Model for the destabilization of the RecA nucleoprotein filament by RecX. The dissociation of an interior monomer of RecA produces a gap within the nucleoprotein filament. In the absence of RecX, the equilibrium between disassembly and assembly of RecA ensures that there is no net disassembly of the nucleoprotein filament. In the presence of RecX, a net disassembly of the nucleoprotein filament occurs due to binding of RecX to the terminal RecA at the gap, which inhibits subsequent binding of RecA monomers onto the gap within the nucleoprotein filament. This figure was modified from Drees *et al*., 2004.

Previous *in vitro* studies had demonstrated that DinI interacts with RecA bound to ssDNA (Yasuda *et al*., 2001). From this information, Lusetti *et al*., (2004) proposed that DinI binds to RecA within the RecA nucleoprotein filament and prevents disassembly of the nucleoprotein filament from the 5' end (Figure 1.3) or suppresses the interchange of RecA between the free monomeric form and those that are bound to the ssDNA. The increase in stability of the RecA nucleoprotein filament, due to the presence of DinI, had no effect on the co-protease activity of RecA (Lusetti *et al*., 2004). When present at very high concentrations relative to RecA, DinI causes a destabilization of the RecA nucleoprotein filament and an inhibition of the SOS response (Yasuda *et al*., 1998 Voloshin *et al*., 2001; Lusetti *et al*., 2004). Following induction of the SOS response, the amount of RecA per *E. coli* cell is approximately 30-fold higher than the amount of DinI (Sommer *et al*., 1998; Yasuda *et al*., 1998). Hence, it has been proposed that under physiological conditions, the expression of the DinI protein results in the stabilization of the RecA nucleoprotein filament following the formation of DNA damage in *E. coli* (Lusetti *et al*., 2004).

1.7 Molecular mechanisms of RecA-mediated homology search

Even though the RecA protein is a DNA-dependent ATPase, data from *in vitro* studies have suggested that ATP hydrolysis is not required for DNA pairing and strand exchange (Menetski and Kowalczykowski, 1985; Kowalczykowski and Krupp, 1995). These observations have been used to infer that homology search does not require active search processes which utilize ATP as a source of energy (Forget and Kowalczykowski, 2012). *In vitro* studies have indicated that one-dimensional sliding of the RecA nucleoprotein filament along the dsDNA is a possible mechanism for homology search prior to homologous pairing and strand exchange (Ragunathan *et al*., 2012). The authors utilized single molecule fluorescence imaging in which the RecA nucleoprotein filament was immobilized on a quartz surface and the dsDNA was added to the experimental system via a flow system. Analysis of these fluorescently-labelled DNA via fluorescence resonance energy transfer (FRET) was used to infer that whenever the RecA nucleoprotein filament encounters the dsDNA, it slides along the dsDNA over a length of 60 - 300 bp within a duration of 0.5 - 10 seconds before dissociation.

Single molecule fluorescence imaging coupled with total internal reflected fluorescence (TIRF) has also been used to demonstrate that the efficiency of homology search is dependent on the three-dimensional conformation of the dsDNA and the length of the RecA nucleoprotein filament (Forget and Kowalczykowski, 2012). The authors observed that homologous pairing between the nucleoprotein filament and a dsDNA was highly efficient when the ends of the

dsDNA were anchored to a microfluidic device in a manner that favoured the three-dimensional folding of the dsDNA. These observations led the authors to propose the 'intersegmental contact sampling model'. According to the model, the nucleoprotein filament exhibits weak and transient non specific contacts with the dsDNA until significant sequence homology is identified, which stabilizes the pairing of the nucleoprotein filament to the dsDNA.

Mara Prentiss and colleagues have also utilized single molecule magnetic tweezers to apply external force at the ends of dsDNA in order to gain indepth understanding on the role of mechanical stress during homology search and the concomitant strand exchange reaction (Danilowicz *et al*., 2012). Outward external forces that were applied to both ends of only one of the two strands of a duplex DNA resulted in a stable binding of the dsDNA to the pre-synaptic filament. On the contrary, forces that were applied to either all four ends of the dsDNA or two out of the four ends which had the same polarity (3'3' or 5'5') did not stabilize the ssDNA-RecA-dsDNA complex. A model was proposed which indicated that the binding of dsDNA to the pre-synaptic filament receives less mechanical support in comparison to the product of the strand exchange reaction. The model also predicted that upon detection of at least 12 nucleotides of sequence homology, strand exchange would occur and the stability of the strand exchange product is independent of heterologous sequences as long as it contained the minimum 12 nucleotide of sequence homology.

1.8 Replication and segregation of the *E. coli* chromosome

The genetic material of a bacterium is encapsulated within a nucleoid. Kleckner and colleagues have reported that the *E. coli* nucleoid exhibits a dynamic helical ellipsoidal structure within the rod-shaped cell (Fisher *et al*., 2013). The authors also showed that, prior to initiation of replication in new born cells, the nucleoid rarely localizes at the regions close to the old cell pole, thereby indicating that the nucleoid does not simply occupy the entire interior space within a cell. Nonetheless, these regions are still accessible to the nucleoid during other phases of the cell cycle. The *E. coli* chromosome is also organised into macro-domains (Ori, Ter, Left, and Right) and two non-structured regions which flank the Ori macro-domain (Valens *et al*., 2004). Interactions are restricted to chromosomal loci within the same macro-domain because of the limited mobility of loci within these macro-domains. However, loci that are within a non-structured region are capable of interacting with each other, as well as with other loci in the neighbouring macro-domains due to their ability to exhibit greater mobility relative to the loci in the macro-domains (Espeli *et al*., 2008).

Replication of the circular *E. coli* chromosome is initiated at *oriC* and proceeds bidirectionally towards the terminus region, which is diametrically opposite to *oriC* (Messer and Weigel, 1996). In slow growth conditions, the cells have a single replicating chromosome and *oriC* is predominantly located at the mid-cell at the onset of DNA replication (Niki *et al*., 2000). Afterwards, the replicated *oriC* loci migrates to the ¼ and ¾ positions, which eventually becomes the mid-cell for the two daughter cells generated from these parents (Niki *et al*., 2000).

A model has suggested that the two sister replisomes originating from *oriC* traverse independently along the left and right arms of the circular *E. coli* chromosome during DNA replication (Brier *et al*., 2005; Reyes-Lamothe *et al*., 2008). An alternative model argues that the sister replisomes are associated with each other to form a centralized compartment at the mid-cell, commonly referred to as the replication factory (Koppes *et al*., 1999; Brendler *et al*., 2000; Molina and Skarstad, 2004). The existence of a replication factory during chromosomal replication have been demonstrated in *Bacillus subtilis* (Lemon and Grossman, 1998; Migocki *et al*,. 2004) and *Caulobacter crescentus* (Jensen *et al*., 2001). In *B. subtilis*, it was shown that the replication factory was highly mobile at the mid-cell without a preferred directionality of movement (Migocki *et al*., 2004). Furthermore, the replisome of each replication fork within the replication factory was dynamic in nature; they underwent transient and reversible separation and re-merging during chromosomal replication (Migocki *et al*., 2004; Berkmen and Grossman, 2006). Interestingly, Reyes *et al*., (2008) also detected that sister replisomes occasionally merged at the mid-cell and re-separated during DNA replication. However, the authors attributed this phenomenon to contraction and re-organization of the nucleoid during chromosomal replication (Reyes *et al*., 2008).

The replication factory model that was proposed by Lemon and Grossman (1998) also suggested that extrusion of the newly replicated loci from the stationary replication factory, and the compaction and refolding of the chromosome may provide the force that is required for physical separation of chromosomal loci following replication. A recent study by Sherrat and colleagues have demonstrated that segregation of the *codA* and *cynR* intergenic loci of the *E. coli* chromosome can occur via a mechanism that is independent of replication during the repair of a site-specific DSB (Lesterlin *et al*., 2014). Moreover, the observation that sister replisomes traversed independently along the two arms of the *E. coli* chromosome did not support the capture extrusion model of segregation that was proposed by Lemon and Grossman (Bates and Kleckner, 2005; Reyes-Lamothe *et al*., 2008). Hence, it has been suggested that chromosomal replication alone is not sufficient for providing the forces that drives segregation

(Badrinarayanan *et al*., 2015a). Unlike eukaryotes where it has been established that mitotic spindles mediate chromosome segregation, the mechanism that is utilized in *E. coli* remains unclear. Nonetheless, segregation in other bacteria such as *V. cholerae*, *C. crescentus* and *B. subtilis* is actively coordinated by the *parABS* system (Badrinarayanan *et al*., 2015a; Gerdes *et al*., 2010) following the recruitment of partition (ParB-*parS*) complexes to high-density regions within the nucleoid via a mechanism that is mediated by the ATPase activity of ParA (Le Gall *et al*., 2016).

As replication progresses from the origin to the terminus of the chromosome, the newly replicated loci undergo sequential segregation into the opposite halves of the *E. coli* cell (Nielsen *et al*., 2006; Wang *et al*., 2006; Nielsen *et al*., 2007). For most loci across the *E. coli* chromosome, replication precedes spatial separation of the sister loci by approximately 10 minutes (Reyes-Lamothe *et al*., 2008; Joshi *et al*., 2011). However, cohesion of newly replicated *oriC* loci and snap loci, which are located on the origin-proximal half of the right chromosomal arm, are relatively extensive compared to the other loci (Bates and Kleckner, 2005; Nielsen *et al*., 2006; Nielsen *et al*., 2007; Lesterlin *et al*., 2012; Joshi *et al*., 2011). As a result, the segregation of these late splitting loci (*oriC* and the snap loci) initiates global re-organization of the genome, leading to the formation of a bilobed nucleoid morphology (Espeli *et al*., 2008; Joshi *et al*., 2011; Kleckner *et al*., 2014). Re-organization of the nucleoid precedes replication of the terminus region of the *E. coli* chromosome. Replication of chromosomal loci at the terminus region occurs at the mid-cell even though these *ter* loci were initially localized at a cell pole in new born cells (Espeli *et al*., 2012). Migration of the *ter* loci from the cell pole to the mid-cell is coordinated by the process of replication and is independent on MatP (Espeli *et al*., 2012), the protein which is responsible for structuring and organizing the terminus region of the *E. coli* chromosome into a macro-domain (Mercier *et al*., 2008). Replication and the concomitant segregation of the Ter macro-domain drives the global segregation of each genome into the daughter cells (Joshi *et al*., 2011; Kleckner *et al*., 2014).

It has been demonstrated that the formation of topological linkages (precatenanes) behind a progressing replication fork is responsible for the cohesion of sister loci after replication (Wang *et al*., 2008). Removal of the linkage between the replicated sister loci is accomplished by topoisomerase IV during decatenation (Espeli and Marians, 2004). Hence, it has been proposed that cohesion of sister loci is modulated by the activity of topoisomerase IV and that decatenation of sister loci facilitates chromosome segregation (Wang *et al*., 2008; Lesterlin *et al*., 2012; Joshi *et al*., 2013). A recent study has shown that the duration of cohesion of sister

loci is not exclusively dependent on the activity of topoisomerase IV during decatenation (Joshi *et al*., 2013). The authors demonstrated that binding of SeqA proteins to GATC sequences of hemimethylated DNA inhibits the immediate resolution of precatenanes by topoisomerase IV (Joshi *et al*., 2013). Genomic analyses also suggested that chromosomal loci which experience longer cohesion periods have higher local amount of GATC sites. From these observations, the authors concluded that the binding of SeqA to GATC sites of hemimethylated DNA is a primary determinant for the duration of cohesion of sister loci after replication. Kleckner and colleagues have also demonstrated the presence of waves of density flux along the longitudinal axis of the nucleoid (Fisher *et al*., 2013). These mechanical waves which diffuses through the longitudinal axis of the nucleoid may also facilitate efficient chromosomal segregation following re-methylation of newly replicated loci and resolution of the topological linkages (Fisher *et al*., 2013, Kleckner *et al*., 2014).

In vivo studies have shown that newly replicated sister loci containing *loxP* sites are not inert during the period of cohesion but can undergo homologous recombination in the presence of the bacteriophage P1 Cre recombinase (Lesterlin *et al*., 2012). It has been suggested that cohesion of sister loci is biologically relevant for the repair of replication-dependent DNA damage (Cox *et al*., 2000). Because DSBs are repaired exclusively by homologous recombination in *E. coli*, it is plausible that cohesion might have evolved to facilitate efficient repair of DSBs arising from chromosomal replication. Thus, the existence of sister loci cohesion might be essential for avoiding deleterious effects of unrepaired DSBs occurring during chromosomal replication.

1.9 Scope of this thesis

Live-cell fluorescence imaging has been used to extensively study the cellular dynamics of the RecA protein during the repair of DSBs in bacteria (Kidane and Graumann, 2005; Lesterlin *et al*., 2014; Badrinarayanan *et al*., 2015b). In these studies, the DSBs were generated using either the HO- or I-SceI- endonuclease systems. The observations from these *in vivo* studies provided very useful insight on the cellular dynamics of RecA in the absence and presence of DSBs. Moreover, these studies demonstrated that an extensive homology search was required during the repair of the DSBs generated by the endonuclease systems. The repair of replication-dependent DSBs would not require an extensive homology search if the repair is initiated during the period of cohesion of the sister loci. Because homology search is catalysed by the RecA nucleoprotein filament, it is possible that the cellular dynamics of fluorescently-labelled RecA might be different during the repair of endonuclease-induced DSBs compared with

replication-dependent DSBs. Thus, the observations from the repair of the endonuclease-induced DSBs might not represent or be applicable to the cellular dynamics of RecA during the repair of replication-dependent DSBs. The ultimate aim of this study was to investigate the spatial and temporal dynamics of repair by recombination of a site-specific DSB that is generated during replication. This objective was achieved via live-cell fluorescence imaging of the RecA protein and the site of the replication-dependent DSB using *E. coli* as a model organism.

It has been shown that the insertion of 246bp interrupted palindrome at the *lacZ* locus of the *E. coli* chromosome results in the formation of a replication-dependent DSB at that locus following expression of the SbcCD endonuclease (Eykelenboom *et al*., 2008). The authors proposed that the lagging strand template forms a DNA hairpin at the site of the interrupted palindrome, which is recognized and cleaved by the SbcCD endonuclease to generate a two-ended DSB (Figure 1.7). Because the replication-dependent DSB at the *lacZ* locus is formed on only one of the sister chromosomes, the unbroken chromosome is used as a template for the repair of the DSB by homologous recombination (Eykelenboom *et al*., 2008). Throughout this study, the SbcCD/palindrome system was used to induce the formation of a repairable DSB at the *lacZ* locus of the *E. coli* chromosome following the onset of DNA replication. The induction of the replication-dependent DSB was accomplished by placing the genes encoding the SbcCD endonuclease under the control of an arabinose-inducible promoter.

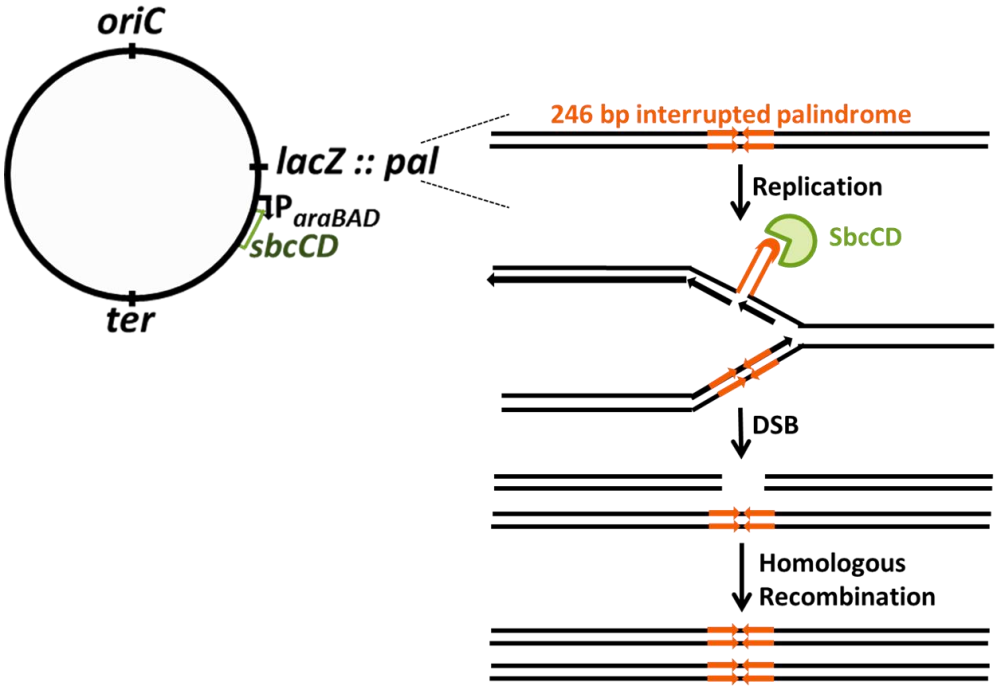


Figure 1.7: The system used for the formation of a replication-dependent DSB at the *lacZ* locus of the *E. coli* chromosome. A 246 bp interrupted palindrome was inserted in the *lacZ* locus. The genes encoding the SbcCD endonuclease were also placed under the control of an arabinose-inducible promoter (P_{araBAD}). During replication, a DNA hairpin is formed on the lagging strand template at the *lacZ* locus. The cleavage of the DNA hairpin by the SbcCD endonuclease generates a two-ended DSB on only one of the sister chromosomes.

Chapter three introduces the tools that were used for fluorescence labelling of RecA and the *lacZ* locus. These tools comprise the tandem insertion of a *recA-mCherry* gene between the endogenous *recA* and *recX* genes at the native chromosomal locus. In addition, *tetO* and *lacO* arrays were inserted at the origin-proximal and origin-distal ends of the *lacZ* locus respectively. The binding of the TetR-YPet and LacI-Cerulean proteins to the *tetO* and *lacO* arrays respectively, was used for visualizing the site of the replication-dependent DSB. The chapter also describes the effect of the formation of replication-dependent DSBs on cell viability and growth rate of *E. coli* strains containing the fluorescently-labelled RecA protein. In order to ensure that most cells have a single replicating chromosome during live-cell imaging, *E. coli* strains were grown in M9-glycerol medium prior to fluorescence imaging. The DNA content of *E. coli* cells grown in M9-glycerol medium was analysed by flow cytometry and was used to demonstrate that the majority of the cell population had a single replicating chromosome under such growth condition. Finally, live-cell imaging was used to obtain preliminary insight on the effect of DSB formation on the duration of cohesion of the sister *lac* loci.

In chapter four live-cell imaging was used to investigate the cellular dynamics of RecA and the *lacZ* locus during the repair of a replication-dependent DSB. Data from time-lapse fluorescence imaging was used to determine whether repair of the replication-dependent DSB occurred during the period of cohesion of the sister *lac* loci or after segregation of these sister loci to the opposite cell halves. The duration of the RecA focus at the site of the DSB was determined in the wild type strain and was compared to that obtained from strains with a deletion mutation for either *uvrD*, *recX* or *dinI* genes. The data obtained from these deletion mutants were used to identify the effect of the absence of the UvrD, RecX or DinI proteins on the stability of the RecA nucleoprotein filament that is formed during the repair of the replication-dependent DSB. Furthermore, time-lapse imaging was used for estimating the duration of cohesion of the sister *lac* loci after the disassembly of the RecA focus at the site of the DSB.

Chapter Five investigates the effect of the formation of a replication-dependent DSB on the spatial localization of the *lac* locus. It was hypothesized that the formation and repair of the DSB could perturb the spatial localization of the *lac* locus relative to the natural localization of that locus if an extensive homology search occurred at pre-synapsis. Moreover, a perturbation of the spatial localization of the *lac* locus in response to the formation of the DSB could be possible if there exist specialized DSB repair centres in *E. coli*. Thus, the chapter also investigated the subcellular position for the repair of the replication-dependent DSBs and identified its relation to the natural localization of the *lac* locus. Furthermore, the spatial

localization for the formation of replication-dependent DSBs was investigated and compared to the subcellular position for the repair of these DSBs. The rationale for the comparison was to identify whether the formation and repair of replication-dependent DSBs occur at the same or different subcellular positions within the cell.

Chapter Six investigated if there was an effect of the fluorescence labelling of the *lac* locus on the binding of RecA at the site of the replication-dependent DSB. Specifically, genomic analysis of RecA-DNA interactions by ChIP-seq was used to ascertain whether binding of the fluorescently-labelled repressor proteins to their operator arrays perturbed the binding of RecA at the site of the DSB. The genomic analysis was also used to identify possible formation of one-ended DSBs due to replication fork collapse at the repressor-bound operator arrays at the *lacZ* locus in the absence of DSB induction.

2. Materials and Methods

2.1 Materials

2.1.1 Growth media

Growth media were prepared using distilled water, autoclaved and stored either at room temperature (liquid LB and M9-glycerol media) or at 55°C (molten LB agar). M9-glycerol agar was always freshly prepared and was used for growing cells on microscope slides during live-cell fluorescence imaging. Antibiotics and inducers were added to the growth media when required.

Table 2.1: Growth media.

Media	Composition
LB	0.01 g/ml Bacto-tryptone, 0.005 g/ml yeast extract, 0.01 g/ml NaCl; pH 7.2
LB Agar	0.01 g/ml Bacto-tryptone, 0.005 g/ml yeast extract, 0.01 g/ml NaCl, 0.015 g/ml Bacto-agar; pH 7.2
M9 Salts (4x)	0.2 M Na_2HPO_4, 0.09 M KH_2PO_4, 0.03 M NaCl, 0.08 M NH_4Cl
M9-Glycerol Medium	1x M9 salts, 0.2% glycerol, 1 mM $MgSO_4$, 20μM $CaCl_2$
M9-Glycerol Agar	1x M9 salts, 0.2% glycerol, 1 mM MgSO4, 20μM $CaCl_2$, 0.015 g/ml agarose

2.1.2 Antibiotics and inducers

Antibiotics and inducers were dissolved in the indicated solvents and stored at -20°C.

Table 2.2: Antibiotics and inducers.

Antibiotic/Inducers	Solvent	Stock concentration	Final concentration
Cephalexin	H_2O	1 mg/ml	10 µg/ml
Chloramphenicol	100% ethanol	50 mg/ml	50 µg/ml
Gentamycin	H_2O	10 mg/ml	10 µg/ml
Kanamycin	H_2O	50 mg/ml	50 µg/ml
Rifampicin	dimethyl formamide	50 mg/ml	150 µg/ml
Anhydrotetracycline	50% ethanol	100 µg/ml	100 ng/ml

2.1.3 Supplements of the growth media

Supplements were added to the growth media when required. They were prepared using distilled water, autoclaved and stored at room temperature.

Table 2.3 Supplements of the growth media.

Supplements	Stock concentration	Final concentration
Arabinose	20%	0.2%
Glucose	20%	0.5%
Glycerol	20%	0.2%
Sucrose	20%	5%

2.1.4 Restriction enzymes

Restriction enzymes were purchased from New England Biolabs (NEB) and were used according to the manufacturer's instructions**.**

2.1.5 Buffers and solutions for Chromatin Immunoprecipitation (ChIP)

ChIP Buffer: 200 mM Tris-HCl (pH 8.0), 600 mM NaCl, 4% Triton X-100, 1x Roche Complete ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor.

Protein G Dynabeads: Supplied in PBS (pH 7.4) and contains 0.1% Tween-20 and 0.02% NaN_3.

Wash Buffer: 1x PBS and 0.02% Tween-20.

TE Buffer: 10 mM Tris (pH 7.4) and 1 mM EDTA.

Crosslinking Agent: 37% Formaldehyde containing 10-15% Methanol to prevent polymerization (Sigma).

Quenching Agent: 2.5 M glycine.

ChIP-Seq Loading Buffer: 50 mM Tris (pH 8.0), 40 mM EDTA and 40% (w/v) sucrose.

2.1.6 Other buffers and solutions

Propidium Iodide Staining Solution

Propidium iodide and RNase A were dissolved in 1x PBS to a final concentration of 50 µg/ml and 500 µg/ml, respectively, to obtain 5x propidium iodide staining solution. The 5x staining solution was diluted to 1x using 1x PBS and filter sterilised prior to storing in the dark at 4˚C.

Tris-acetate (TAE) buffer

A 50x stock solution (1L) was prepared by dissolving 242 g of Tris base in distilled water, followed by the addition of 57.1 mL of glacial acetic acid and 100 mL of 500 mM EDTA (pH 8.0) solution. The solution was stored at room temperature and was used at a working concentration of 1x for agarose gel electrophoresis.

2.2 Methods

2.2.1 Molecular Biology Methods

2.2.1.1 Isolation of plasmids from *E. coli* strains

The QIAprep® Spin Miniprep Kit (Qiagen) was used for all plasmid extractions according to the manufacturer's instructions. The plasmid DNA was eluted with sterile water and stored at -20°C.

2.2.1.2 Genomic DNA extraction

The Wizard® Genomic DNA purification kit (Promega) was used for isolating genomic DNA from the desired *E. coli* strains. Genomic DNA was eluted with sterile water and stored at -20 °C.

2.2.1.3 Polymerase chain reaction

Polymerase chain reactions (PCR) were performed in a thermo-cycler from PeQLab Biotechnology. GoTaq® Flexi DNA polymerase (Promega) was used for routine PCR reactions that were performed mainly for confirming the presence or absence of a gene or DNA fragment of interest. The Herculase II Fusion DNA Polymerase (Agilent) was used when the PCR reaction was performed for cloning of a DNA fragment into a plasmid vector. A 25 µL reaction in sterile water consisted of 500 nM of deoxyribonucleotides, 400 nM of reverse and forward primers each, 1x of reaction buffer containing 2mM of Mg^{2+}, 1 unit of DNA Polymerase and 2 µL of boiled-cell DNA extract or 0.5 µl of purified DNA template. The program for performing a PCR was usually:

Initial template denaturation	95°C	5 mins	
Denaturation	95°C	30 sec	30 cycles
Annealing	(Tm-5)°C	30 sec	
Extension	72°C	1 min/kb	
Final extension	72°C	10 mins	

2.2.1.4 Agarose gel electrophoresis

Gel electrophoresis was carried out at 80 volts for 30 minutes using 1% (w/v) agarose gels in 1x TAE buffer. The nucleic acid stain, safeview from NBS Biologicals, was added (to a final concentration of 0.005%) during the preparation of agarose gels to enable the visualisation of the separated DNA fragments. Gels were viewed under a UV lamp (UVIdoc, by UVItec) and

the size of the DNA fragments of interest was estimated using DNA ladders from New England Biolabs.

2.2.1.5 Sanger Sequencing of DNA

The complete DNA sequences of the *tetO* and *lacO* arrays were determined via Sanger sequencing of PCR products amplified from the chromosomal loci bearing these operator arrays. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) prior to sequencing. The Sanger sequencing reaction was set-up using the BigDye® Terminator v3.1 Cycle-Sequencing Kit (Applied Biosystems) by following the manufacturer's guidelines. The samples were sent to the Edinburgh Genomic Facility for analysis using the ABI 3730XL capillary sequencing instrument.

2.2.2 Bacterial Methods

2.2.2.1 Storage of bacteria

Overnight cultures of *E. coli* strains were suspended in an equal volume of 80% sterile glycerol (v/v) and stored at -80°C for long-term. For routine purposes, *E. coli* strains of interest were stored at 4°C on LB agar plates containing the required antibiotics.

2.2.2.2 Transformation of *E. coli* strains

The *E. coli* strain to be transformed was grown overnight at 37°C in liquid LB medium containing the required antibiotics. The overnight culture was diluted (50-fold) in LB medium and grown for 2 hours at 37°C. Cells were harvested from 1 ml of the culture and re-suspended in 0.5 ml of 0.1 M $CaCl_2$. The cells in suspension were left on ice for 30 minutes. Afterwards, the cells were harvested, re-suspended in 0.1 ml of 0.1 M $CaCl_2$ and 4 µl of plasmid DNA were added. The cells were left on ice for another 30 minutes, heat-shocked at 37°C for 5 minutes and incubated in 400 ml of LB medium for 1 hour at 30°C. Finally, 100 µl of these transformed cells were incubated for two overnights at 30°C on LB agar plates containing the required antibiotics. The 0.1 M $CaCl_2$ was freshly prepared and was always kept on ice during the experiment.

2.2.2.3 Plasmid mediated gene replacement

Derivatives of the pTOF24 plasmid were used for introducing mutations (insertions or deletions) at specific loci in the *E. coli* chromosome. In order to construct derivatives of the pTOF24 plasmid, primers were designed to amplify the relevant DNA sequences from a plasmid or the chromosomal DNA. Crossover PCR was used for generating a final product which had the desired mutation flanked by DNA sequences containing approximately 400 bp

homology to the chromosomal locus where the mutation was to be introduced. The original pTOF24 plasmid and the final product of the crossover PCR were digested with suitable restriction enzymes and ligated together to obtain a derivative of the pTOF24 plasmid that would be used to introduce the mutation at the desired chromosomal locus.

The pTOF24 plasmid has three unique features which permits the introduction of mutations into the *E. coli* chromosome by plasmid mediated gene replacement (PMGR). These features comprise a gene which encodes a temperature sensitive RepA protein, a chloramphenicol (Cm) resistance gene and a gene from *B. subtilis* which encodes the levansucrase enzyme (Merlin *et al*., 2002). The temperature sensitive RepA protein is required for replication of the plasmid and is functional at 30°C but not at 42°C (Link *et al*., 1997). Levansucrase catalyzes the hydrolysis of sucrose and synthesizes levan polysaccharide (Merlin *et al*., 2002). The levan polysaccharide is toxic in *E. coli*, hence cells containing the pTOF24 plasmid are sensitive to sucrose.

The procedure for PMGR requires that the *E. coli* strain of interest was transformed with the pTOF24 derivative plasmid containing the desired mutation. Cells that acquired the plasmid during transformation were observed as distinct colonies on LB plates, containing chloramphenicol, following incubation for two overnights at 30°C. Single colonies were streaked onto Cm plates and incubated overnight at 42°C in order to select for cells that had integrated the plasmid into the chromosome because the plasmid is unable to autonomously replicate at 42°C. Integration of the plasmid into the chromosome occurs by homologous recombination using either one of the two homology arms flanking the site of the desired mutation. Large single colonies that were obtained on the Cm plates were purified by re-streaking onto fresh Cm plates and incubated overnight at 42 °C. The plasmid that had integrated into the chromosome was allowed to excise by growing a colony of cells from the latter Cm plate in liquid LB medium at 30°C, overnight. The LB medium was not supplemented with Cm to avoid selecting for cells which contained the plasmid integrated into the chromosome. The overnight culture was diluted serially, inoculated onto LB agar plates supplemented with 5% sucrose and incubated overnight at 37°C to select for cells that had lost the plasmid. The desired mutation was transferred from the plasmid to the chromosome only if the integration and excision occurred via different homology arms. Otherwise, the original genotype of the chromosome is retained if the integration and excision of the plasmid occurred using the same homology arm. Levansucrase is not a perfect negative selection marker, hence

the sucrose-resistant colonies were screened for Cm sensitivity to confirm the absence of the plasmid. Cells that were resistant to sucrose but sensitive to Cm were tested for the presence of the desired mutation by PCR. In addition, spot test and ultraviolet light sensitivity assays were used to confirm the phenotype of recombination-deficient mutants.

2.2.2.4 Spot test and ultraviolet light sensitivity assay

A colony of the *E. coli* strain of interest was grown overnight in liquid LB medium at 37˚C. The OD_{600} of the overnight culture was adjusted to 1.0 and serially diluted 10-fold. Aliquots of 4 µl of the serially diluted cultures were spotted onto LB agar plates containing either 0.2% of arabinose or 0.5% of glucose. The LB agar plates were incubated overnight at 37˚C. For ultraviolet light sensitivity assay, the LB-glucose agar plates were exposed to the required dose of the radiation prior to incubation at 37˚C.

2.2.2.5 Growth of bacterial strains for flow cytometry, microscopy and chromatin immunoprecipitation

Unless otherwise stated, all the strains used in this study were derivatives of the *E. coli* BW27784 strain. This background strain enables homogenous expression of the SbcCD endonuclease from the arabinose-inducible promoter (Khlebnikov *et al*., 2001). The *E. coli* strain of interest was grown overnight in M9-minimal medium supplemented with 0.2% of glycerol. The overnight culture was diluted to an OD_{600} = 0.09 and grown for 3 hours (OD_{600} = 0.3 – 0.35). The bacterial culture was diluted again in the same medium, and 0.2% of arabinose was added for induction of the expression of the SbcCD endonuclease. The diluted culture was grown for 2 hours prior to microscopy, flow cytometry or chromatin immunoprecipitation.

E. coli cells were also grown in LB medium for chromatin immunoprecipitation. Under this growth condition, overnight cultures of cells grown in LB medium were diluted to an OD_{600} = 0.02. The diluted cultures were grown for 90 minutes (OD_{600} = 0.2 – 0.25) and re-diluted to OD_{600} = 0.02. Arabinose was added at a final concentration of 0.2% and the diluted cultures were grown for a further 75 minutes (OD_{600} = 0.25). Afterwards, the cells were used for the chromatin immunoprecipitation experiment.

All bacterial cultures were grown in a shaking incubator at 37 ˚C. The M9-glycerol medium or LB medium was supplemented with 100 ng/ml of anhydrotetracycline (ATC) when growing strains that contained the *tetO* array and the TetR-YPet proteins. Additionally, liquid LB

medium was always supplemented with 0.2% of glucose when growing overnight cultures of the desired *E. coli* strains.

2.2.3 Analysis of DNA content by flow cytometry

Cells at exponential phase of growth were treated with cephalexin and rifampicin. Cephalexin and rifampicin were added at a final concentration of 10 µg/ml and 150 µg/ml, respectively, and the cells were grown for a further 3 hours. Aliquots of 1 ml of the overnight, exponential phase and cephalexin/rifampicin-treated cultures were separately fixed in 70% ethanol and stored overnight at 4°C. The cells in 70% ethanol were harvested, washed in 1x PBS and the DNA of these cells were stained with 1x propidium iodide solution for 1 hour in the dark at room temperature. An A50 Micro flow cytometer (Apogee Flow systems) was used for recording the fluorescence signal generated by the stained cells after excitation with the blue laser (488nm). The data obtained were analysed using the Apogee Histogram Software v3.1.

2.2.4 Wide-field microscopy

Conventional wide-field fluorescence microscopy was performed using a Zeiss Axiovert 200 fluorescence microscope equipped with a 100X/1.4 oil Plan Apochromat objective (Phase or DIC), dual OptoLED light source (Cairn Research), a MS-2000 piezo z stage (ASI) and an Evolve™ 512 EMCCD camera (Photometrics). The microscope is enclosed in an incubation chamber in order to control the temperature during live-cell fluorescence imaging. Prior to either snapshot or time-lapse microscopy, the incubation chamber of the microscope was kept at 37°C. The Continuous Reflection Interface Sampling and Positioning (CRISP) system was used to ensure that cells remained accurately focused with extremely minimal drift during time-lapse imaging.

Molten agarose (0.015 g/ml) was prepared in M9-glycerol medium and was mounted within a 1.5 x 1.6 cm Gene Frame (Thermo Scientific) that was sealed onto a microscope slide. An aliquot of 6 µl of the cells growing in M9-glycerol medium was spread onto the solidified agarose and sealed with a cover slide. The Metamorph software (Molecular Devices) was used for image acquisition. Phase contrast or DIC images were acquired concurrently with the fluorescence images. For snapshot microscopy, each fluorescence image consisted of 11 z-sections with 200 nm z-distance. The fluorescence images that were acquired during time-lapse microscopy consisted of 6 z-sections with 350nm z-distance in order to minimize photobleaching. The following settings were used during image acquisition: an exposure time of 50 ms and a gain of 1 for phase and DIC; an exposure time of 50 ms and a gain of 150 for

RecA-mCherry; an exposure time of 100 ms and a gain of 250 for LacI-Cerulean, TetR-YPet and DnaN-YPet images. 100 % transmission was used for all filters during image acquisition.

The images obtained from microscopy were analysed using either the ImageJ (http://rsbweb.nih.gov/ij/) or the Metamorph software. During analysis, the z-sections of each fluorescence image were converted to a single image that corresponded to the sum of the stacked images. Afterwards, the image was deconvolved using the *AutoQuant* X software. Deconvolution was carried out to reduce noise and improve the resolution of the fluorescence image. During analysis of the cellular localization of fluorescence spots within the cell, the Oufti software (Paintdakhi *et al*., 2016) was used for generating demographs of the images obtained from either snapshot or time-lapse microscopy.

2.2.5 Chromatin Immunoprecipitation (ChIP)

Proteins were crosslinked to DNA by the addition of formaldehyde (Sigma) to 50 ml of the *E. coli* cultures at exponential phase of growth. Formaldehyde was added at a final concentration of 1% and crosslinking was performed for 10 minutes at room temperature in a shaking incubator. The crosslinking was quenched by the addition of glycine at a final concentration of 0.5 M for 10 minutes at room temperature in a shaking incubator. After crosslinking, the cells were harvested by centrifugation at 4,000 x g for 7 minutes at 4°C and were washed three times in ice-cold 1x PBS. The cell pellet was re-suspended in 250 µl of ChIP buffer and sonicated at 30 second intervals for 10 minutes at high amplitude using a Diagenode Bioruptor. During sonication, the samples were kept at 5°C using a temperature-controlled water bath. After sonication, 350 µl of ChIP buffer was added to each sample and mixed gently by pipetting. An aliquot of 100 µl of the cell lysate was stored as ‘’input’’ DNA at -20°C. The remaining 500 µl of each cell lysate was denoted as ‘’IP’’ and was incubated overnight at 4°C with 5 µl of anti-RecA antibody. Afterwards, the IP samples were incubated with Protein G Dynabeads (Invitrogen) for 2 hours at room temperature. This step was essential for the binding of the antibody-RecA protein-DNA complex to the Dynabeads via the Fc region of the antibody. The supernatant was discarded and the Dynabeads were washed three times with 1x PBS supplemented with 0.002% of Tween-20. The Dynabeads were re-suspended in 200 µl of TE buffer containing 1% of SDS and 0.1 mg/ml of RNase A. An aliquot of 100 µl of TE buffer containing 1% of SDS and 0.2 mg/ml of RNase A was also added to each of the input samples. The input and IP samples were incubated at 65°C for 10 hours to reverse the chemical crosslinks between the DNA and the RecA proteins. DNA were isolated from the input and IP

samples using the MinElute PCR purification kit (Qiagen). DNA were eluted in 100 µl of TE buffer and stored at -20°C.

2.2.6 Preparation of DNA libraries for high-throughput DNA sequencing

DNA samples obtained from the RecA-ChIP experiments were used for DNA library preparation following the manufacturer's guidelines of the NEBNext® ChIP-Seq Library preparation kit for Illumina sequencing (Figure 2.1). During DNA library preparation, dsDNA with overhangs were filled to obtain blunt-ended dsDNA. The blunt-ended dsDNA were phosphorylated at the 5' ends and were also adenylated at the 3' ends to obtain dA-tailed DNA. The dA-tailed DNA samples were ligated to NEBNext DNA Adaptors and the uracil-specific excision reagent (USER) enzyme was used for cleaving the DNA adaptors at the site of a Uracil nucleotide. DNA barcodes bearing the P7 primers was annealed to one of the ends of the dsDNA and the P5 primer was annealed to the other end of the dsDNA during amplification of the Adaptor-ligated DNA samples by PCR. The DNA samples were purified after each step ('Clean Up' as indicated in Figure 2.1) using the Qiagen MinElute PCR purification kit following the manufacturer's guidelines. The products of the PCR were ran on a 2% agarose gel at 100 v for 90 minutes to select for DNA fragments whose sizes were between 300 to 400 bp. The required DNA fragments were excised and purified from the gel using the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's guidelines. The purified DNA was quantified using a Bioanalyzer (Agilent) and sent to the Edinburgh Genomics Facility for sequencing on the Illumina HiSeq 4000 sequencer.

2.2.7 Analysis of ChIP-seq data

High-throughput sequencing of DNA libraries by the Illumina Hiseq 4000 sequencer generated compressed fastq files with a characteristic length of each DNA read of 75 bp. These files were decompressed using the gunzip program and mapped to the modified *E. coli* BW27784 reference genome using the Bowtie 2 program by applying the default parameters. The SAM files that were generated by the mapping of the DNA reads to the reference genome were converted to sorted BAM files using the Samtools program. The sorted BAM files were plotted using the R software by applying the in-lab scripts.

For statistical comparison of RecA enrichment between two strains, the sorted BAM files were used to generate count data using the HTS count program. Fold change was calculated using in-lab scripts of the R software which utilised the DESeq program. Analyses of the ChIP-seq data were performed by Dr. Mahedi Hasan.

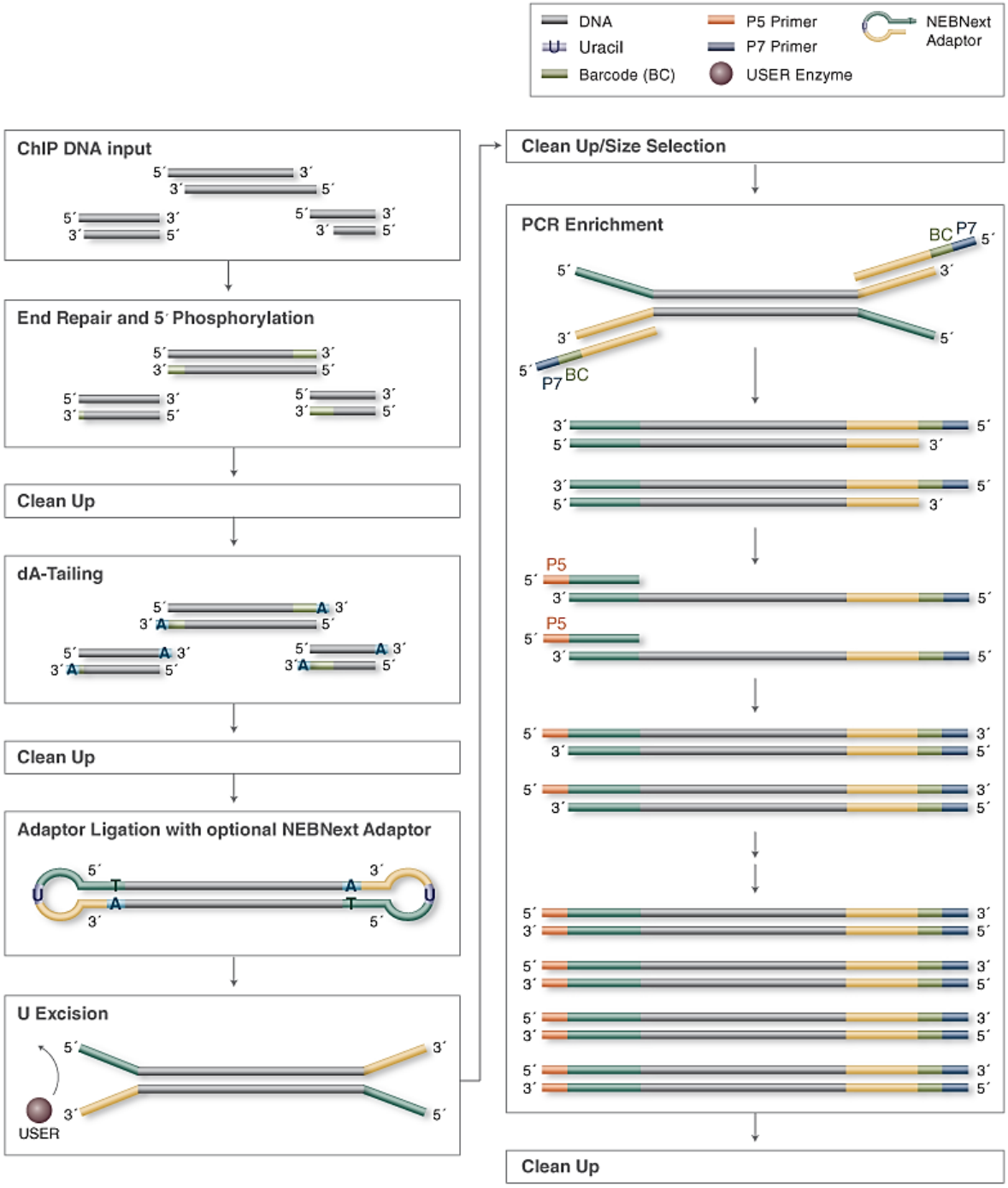


Figure 2.1: Schematic representation for preparation of DNA libraries. The figure was adapted from the webpage of the manufacturer of the NEBNext® ChIP-Seq Library preparation kit for Illumina sequencing (https://www.neb.com/products/e6200-nebnext-chipseq-sample-prep-reagent-set-1).

2.2.8 *Escherichia coli* strains, plasmids and oligonucleotides.

The *E. coli* strains, plasmids and oligonucleotides that were used in this study are listed in Tables 2.2.1, 2.2.2 and 2.2.3 respectively.

Table 2.4: *E. coli* strains

Strain	Genotype	Source
DL4695	BW27784 *hupA-mCherry rph+ ykgC::Pmw1-lacI-Cerulean,tetR-YPet* P_{BAD}*-sbcDC lacZ::pal246 cynX::[(240xtetO)::Gm^R] mhpC::[(240xlacO)::Km^R] ΔlacI lacZχ- mhpA::χχχ lacZY::χχχ*	Dr Martin White, The Leach laboratory
DL4696	DL4695 *lacZ+*	
DL5516	DL4695 *hupA+*	PMGR using pDL5146
DL5521	DL4696 *hupA+*	
DL5525	DL5521 *recA-recX::recAmwg-mCherry*	PMGR using pDL5196
DL5528	DL5516 *recA-recX::recAmwg-mCherry*	
DL5865	DL5528 Δ*uvrD*	PMGR using pDL2391 (Aikaterini Kopsia)
DL5866	DL5528 Δ*recA*	PMGR using pDL5855
DL5901	DL5525 Δ*uvrD*	PMGR using pDL2391
DL5924	DL5521 Δ*recA*	PMGR using pDL2711
DL5941	DL5516 Δ*recA*	
DL5942	DL5525 Δ*recA*	PMGR using pDL5855
DL6107	DL5528 Δ*lacI-Cerulean* Δ*tetR-YPet*	PMGR using pDL2802
DL6123	DL6107 *ykgC::Pmw1-lacI-Cerulean*	PMGR using pDL4580
DL6155	DL5525 Δ*recX*	PMGR using pDL6140
DL6156	DL5528 Δ*recX*	
DL6306	DL6123 *YPet-dnaN*	PMGR using pDL6255
DL6311	DL6107 *lacZ+*	PMGR using pDL1823
DL6312	DL6123 *lacZ+*	PMGR using pDL1823
DL6326	DL6312 *YPet-dnaN*	PMGR using pDL6255
DL6443	DL6107 transformed with pDL2856	This study

Strain	Genotype	Source
DL6444	DL6311 transformed with pDL2856	This study
DL6446	DL5528 Δ*dinI*	PMGR using pDL4423 (Dr Mahedi Hasan)
DL6465	DL5525 Δ*dinI*	PMGR using pDL4423 (Ewa Okely)
DL6486	DL5516 Δ*lacI-Cerulean* Δ*tetR-YPet*	PMGR using pDL2802
DL6517	DL6486 *lacZ+*	PMGR using pDL1823

χχχ represents an array of three chi sites inserted at the *mhpA* and l*acZY* loci. *lacZY* represents the intergenic region between *lacZ* and *lacY*. *lacZχ-* represents an allele of *lacZ* containing a deletion of one endogenous Chi site.

Table 2.5: Plasmids

Plasmid	Description	Source
pDL1605	pTOF24 plasmid	Merlin *et al*., 2002
pDL1823	pTOF24 derivative for converting *lacZ::pal246* to *lacZ+*	Zahra *et al*., 2007
pDL2391	pTOF24 derivative for generating *ΔuvrD*	John Blackwood
pDL2711	pTOF24 derivative for generating *ΔrecA* in strains without the *recAmwg-mCherry* gene integrated into the chromosome	Ewa Okely
pDL2802	pTOF24 derivative for generating *ΔlacI-Cerulean ΔtetR-YPet* in strains with these genes integrated into the chromosome	
pDL2856	Plasmid which contains the *gfp* gene under the control of the *sfiA* promoter	Gary Blakely
pDL4423	pTOF24 derivative for generating *ΔdinI*	Charlie Cockram
pDL4580	pTOF24 derivative for integrating *Pmw1-lacI-cerulean* into the chromosomal *ykgC* locus	Martin White
pDL5146	pTOF24 derivative for converting *hupA-mCherry* to *hupA+*	
pDL5196	pTOF24 derivative for integrating *recAmwg-mCherry* between the endogenous *recA* and *recX* genes	
pDL5855	pTOF24 derivative for deleting the endogenous *recA* gene in strains containing the *recAmwg-mCherry* gene integrated into the chromosome	This study
pDL6140	pTOF24 derivative for generating *ΔrecX* in strains containing the *recAmwg-mCherry* gene integrated into the chromosome	Benura Azeroglu
pDL6255	pTOF24 plasmid for inserting *YPet* and a 33 bp DNA linker to the *ATG* of the endogenous *dnaN* gene	This study

Table 2.6: Oligonucleotides

Primer	Sequence (5' – 3')	Purpose
DinI.F1	AAAAACTGCAGCATCGCAAAGAGAGCAGTTG	Confirm *dinI* deletion by pDL4423
DinI.R2	AAAAAGTCGACGTGCTGATGTCCCACCCTAC	
Ex-test_F	TTATGCTTCCGGCTCGTATG	Confirm the presence of the 246 bp interrupted palindrome inserted at *lacZ*
Ex-test_R	GGCGATTAAGTTGGGTAACG	
HupA-mCherry_F2	AAAAACTCGAGCCCGTCTGGTCTACATTTGG	Confirm the conversion of *hupA-mCherry* into *hupA+* by pDL5146
HupA-mCherry_R2	AAAAAGCGGCCGCTGGTCGTTAGAAAGCTGCTG	
pKO F	AGGGCAGGGTCGTTAAATAGC	Sequencing fragments inserted into the cloning site of the pTOF24 plasmid
pKO R	AGGGAAGAAAGCGAAAGGAG	
RecA-KO-F1	AAAAACTGCAGAACGCGGATTTGTCACCTAC	Confirm *recA* deletion by pDL2711
RecA-KO-R2	AAAAAGTCGACCGCGGGAAATACCTTTCTG	
RecA-mCherry_F1	AAAAAGCGGCCGCCGCTGAAATTCTACGCCTCT	Confirm the integration of *recAmwg-mCherry* between *recA* and *recX* by pDL5196
RecA-mCherry_R4	AAAAACTCGAGATTCACGCATCGCTTTTTCT	
recX2.F1	AAAAACTGCAGCGCTGAAAGGCGAAATAAAA	Generating a *ΔrecX* fragment for pDL6140
RecX.R1	AACGCTGGATCTGTCGCAGTTCTTGCTCACTG	
RecX.F2	AAGAACTGCGACAGATCCAGCGTTTTCTGCTC	
RecX.R2	AAAAAGTCGACGTAGCGCGGGAATAATTACG	
uvrD.F1	AAAACTGCAGTGACCTCGCTGATATAATCA	Confirm *uvrD* deletion by pDL2391
uvrD.R2	AAAAGTCGACTCAGATACTGAAGATGGCGC	
ykgC-F1	AAAAACTGCAGGCATCAACAAACGGCTAAGG	Confirm the presence of the DNA insert at *ykgC* by either pDL2802 or pDL4580
ykgC-R2	AAAAAGTCGACTGTTCTGGCGTCTGATTTTG	
delta recA-mCherry F1	AAAAACTGCAGGGTGAGTGAACCCGTCGT	Generating a *ΔrecA* fragment for pDL5855
delta recA-mCherry R1	AAAAATCTTCGTTAGTTTCTTTGTTTTCGTCGATAGCC	
delta recA-mCherry F2	GGCTATCGACGAAAACAAAGAAACTAACGAAGATTTTT	
delta recA-mCherry R2	AAAAAGTCGACCTGGGAACACAGGAGGTTGT	
YPet-dnaN F1	AAAAACTGCAGAACGTACGTGAGCTGGAAGG	For constructing pDL6255
Same RBS R1	GAATATCCTCCTTAGAAGCAGCTCCAGCCTACAAAGGT TTACGATGACAATGTTCTGATTAAA	
Same RBS F2	TTGTAGGCTGGAGCTGCTTCTAAGGAGGATATTCATGT CTAAAGGTGAAGAATTATTCACTG	
YPet-dnaN R2	AAAAAGTCGACTTCAATCAGACGCTTCATCG	

Primer	Sequence (5’ – 3’)	Purpose
seq_tetO-1.F	CAGTGATAGAGAAGACGAACCG	For sequencing the *tetO* array
seq_tetO-1.R	GTAGGGACCATTGCTAGAGTTG	
seq_tetO-2.F	GTAATCGCTCTAGCAAGCG	
seq_tetO-2.R	CTGCTAGAGCCAGCTCG	
seq_tetO-3.F	CAACTCTAGCAATGGTCCCTAC	
seq_tetO-3.R	CATGCTAGAGCCTCCTTAC	
seq_tetO-4.F	CGAGCTGGCTCTAGCAG	
seq_tetO-4.R	CTCTCTGCTAGAGTACCTTC	
seq_tetO-5.F	GTAAGGAGGCTCTAGCATG	
seq_tetO-5.R	GCTCCCCTGCTAGAGATTATAAG	
seq_tetO-6.F	GAAGGTACTCTAGCAGA	
seq_tetO-6.R	CTCTATCACTGGTAGGGACG	
seq_tetO-7.F	CTTATAATCTCTAGCAGGGGAGC	
seq_tetO-7.R	CGTTTGGCCCTGCTAGAG	
seq_tetO-8.F	CGTCCCTACCAGTGATAGAG	
seq_tetO-8.R	CCTTCTGCTAGAGCTACTATTC	
seq_tetO-9.F	CTCTAGCAGGGCCAAACG	
seq_tetO-9.R	GACAGCTACCACTGCTAG	
seq_tetO-10.F	GAATAGTAGCTCTAGCAGAAGG	
seq_tetO-10.R	GTTTCGTTGCTAGAGCACTAG	
seq_tetO-11.F	CTAGCAGTGGTAGCTGTC	
seq_tetO-11.R	CTCGCTCTCTTGCTAGAGTG	
seq_tetO-12.F	CTAGTGCTCTAGCAACGAAAC	
seq_tetO-12.R	CTCTATCACGATAGGGAAGTGC	
seq_tetO-13.F	CACGTGATAGAGATGGAGTTGGA	
seq_tetO-13.R	CTATCACTGATAGGGCCTC	
seq_tetO-14.F	CAGAGGCTCTAGCAGAGAC	
seq_tetO-14.R	CTCTACCACTGATAGGGAAC	
seq_tetO-15.F	GAGGCCCTATCAGTGATAG	
seq_tetO-15.R	CTGCTTGCTAGAGCTGG	
seq_tetO-16.F	GTTCCCTATCAGTGGTAGAG	

Primer	Sequence (5' – 3')	Purpose
seq_tetO-16.R	CTCTATCACTGATAGGAGGCC	For sequencing the *tetO* array
seq_tetO-17.F	CCAGCTCTAGCAAGCAG	
seq_tetO-17.R	CGTCCTTCTCTATCACTGAGTC	
seq_tetO-18.F	GGCCTCCTATCAGTGATAGAG	
seq_tetO-18.R	CGCACTTCTGCTAGAGTTCAAACG	
seq_tetO-19.F	GACTCAGTGATAGAGAAGGACG	
seq_tetO-19.R	CCTCATGCTTGCTAGAGATTG	
seq_tetO-20.F	CGTTTGAACTCTAGCAGAAGTGCG	
seq_tetO-20.R	CTGCTATGCTAGAGTTTAGACC	
seq_tetO-21.F	CAATCTCTAGCAAGCATGAGG	
seq_tetO-21.R	CTTGTCACGATGCTAGAGTC	
seq_tetO-22.F	GGTCTAAACTCTAGCATAGCAG	
seq_tetO-22.R	CTGATAGGGAGATGCACAG	
seq_tetO-23.F	GACTCTAGCATCGTGACAAG	
seq_tetO-23.R	GCTCTCTATCACAGATAGGG	
seq_tetO-24.F	CTGTGCATCTCCCTATCAG	
seq_tetO-24.R	CTGACAGGGACCTCTCTTC	
seq_tetO-25.F	CCCTATCTGTGATAGAGAGC	
seq_tetO-25.R	GTCGACTCTAGAGTACGTC	
Gentamicin 1	AGAGCGTATTACCTTC	For sequencing the gentamycin resistance gene within *tetO* array
Gentamicin 2	GTAGCCACCTACTCCCAACAT	
Gentamicin 3	AGATGGGGGCCTCTAGCAA	
seq_lacO-1.F	ACACTTTATGCTTCCGGCTC	For sequencing the *lacO* array
seq_lacO-2.F	ACGGACGGAAAATTGTGAGC	
seq_lacO-3.F	GAACCAGGAACTCTAGCAAG	
seq_lacO-4.F	AGTAGGAAGCTCTAGCAGAC	
seq_lacO-5.F	AGCGGATAACTATTGGGTCG	
seq_lacO-6.F	CGGCTCTAGCATGTTAGCAT	
seq_lacO-7.F	GAAGAGTAGCTCTAGCAAGG	
seq_lacO-8.F	CAAGGAGCTCTAGCAAAAGG	
seq_lacO-9.F	GCACATACCCCGAAATTGTG	

Primer	Sequence (5’ – 3’)	Purpose
seq_lacO-10.F	GCGGAAACAATTGGGAGACA	For sequencing the *lacO* array
seq_lacO-11.F	CACGCGCATAGGAATTGTGA	
seq_lacO-12.F	GAGCAAAGCTCTAGCACG	
seq_lacO-13.F	TGACCTCTGGAATTGTGAGC	
seq_lacO-14.F	TGATAACAGCGCTCTAGCAG	
seq_lacO-14.R	CGCTCACAATTCCGATATGC	
seq_lacO-15.F	TCTAGCAAAGGTGCCAA	
seq_lacO-15.R	TCTTTCCTTCTGCTAGAGGG	
seq_lacO-16.F	GAATCCAGGACTCTAGCAAC	
seq_lacO-17.F	CCCTCTAGCAGAAGGAAAGA	
seq_lacO-17.R	GTGATGCTAGAGGCTACTTG	
seq_lacO-18.F	CAGTCACTGCTAGAGCATTG	
seq_lacO-19.F	CAAGTAGCCTCTAGCATCAC	
seq_lacO-20.F	ATGGCTCTAGCAATGGAGAG	
seq_lacO-21.F	TGGGGATAAGTCTCTAGCAG	
seq_lacO-22.F	TGGCCAAAAGTCTCTAGCAG	
seq_lacO-23.F	TGAGAAGCGGGAATTGTGAG	
seq_lacO-24.F	TGGCCCCGGATAATTATGAG	
seq_lacO-25.F	AGGCTACGAACTCTAGCAAG	
seq_lacO-26.F	CTCTAGCAGCGGGAAGTTAA	
seq_lacO-27.F	AGGACCTCTAGCAAAGATGG	
seq_lacO-28.F	CAAATGGGCCTCTAGCAATC	
seq_lacO-28.R	TCCGCTCATTTATTACGCCC	For sequencing the *lacO* array
seq_lacO-1.R	CTTGCTAGAGTTCCTGGTTC	
seq_lacO-2.R	GTCTGCTAGAGCTTCCTACT	
seq_lacO-3.R	CGACCCAATAGTTATCCGCT	
seq_lacO-4.R	ATGCTAACATGCTAGAGCCG	
seq_lacO-5.R	CCTTGCTAGAGCTACTCTTC	
seq_lacO-6.R	CCTTTTGCTAGAGCTCCTTG	
seq_lacO-7.R	CACAATTTCGGGGTATGTGC	
seq_lacO-8.R	TGTCTCCCAATTGTTTCCGC	

Primer	**Sequence (5’ – 3’)**	**Purpose**
seq_lacO-9.R	TCACAATTCCTATGCGCGTG	For sequencing the *lacO* array
seq_lacO-10.R	CGTGCTAGAGCTTTGCTC	
seq_lacO-11.R	GCTCACAATTCCAGAGGTCA	
seq_lacO-12.R	CTGCTAGAGCGCTGTTATCA	
seq_lacO-13.R	CTGAATTGTTATCCGCTCGC	
seq_lacO-16.R	AATGCTCTAGCAGTGACTGG	
seq_lacO-17.R	GTGATGCTAGAGGCTACTTG	
seq_lacO-18.R	CTCTCCATTGCTAGAGCCAT	
seq_lacO-19.R	CTGCTAGAGACTTATCCCCA	
seq_lacO-20.R	CTGCTAGAGACTTTTGGCCA	
seq_lacO-21.R	CTCACAATTCCCGCTTCTCA	
seq_lacO-22.R	CTCATAATTATCCGGGGCCA	
seq_lacO-23.R	CTTGCTAGAGTTCGTAGCCT	
seq_lacO-24.R	TTAACTTCCCGCTGCTAGAG	
seq_lacO-25.R	CCATCTTTGCTAGAGGTCCT	
seq_lacO-26.R	GATTGCTAGAGGCCCATTTG	
seq_lacO-27.R	TGCAGTGCTAGAGTAAACCC	
LacO H1	TTACGCCGATGCCAACCGAA	
LacO H2	CGTTCACCCTTAAATGGCCG	
seq_lacO-29.R	CTCGAATTGTGCTCACAATTG	
LacO H3	GCTAGAGCAATCTTGCCAATTG	For sequencing the *lacO* array
seq_lacO-30.R	GTTATCCGCTCACAATTTCTCC	

Chapter 3: Effect of a replication-dependent DSB on cell viability, growth rate and duration of *lacZ* locus cohesion in *E. coli* strains with a fluorescently labelled RecA protein

3.1 Introduction

The cellular behaviour of the RecA protein during repair of DSBs has been extensively studied by live-cell fluorescence imaging in *Bacillus subtilis, Escherichia coli* and *Caulobacter crescentus* (Kidane and Graumann, 2005; Lesterlin *et al*., 2014; Badrinarayanan *et al*., 2015b). These studies primarily used fluorescently labelled RecA fusion proteins for visualizing repair of an induced DSB. RecA is the preferred protein for studying DSB repair by conventional fluorescence microscopy because of its tendency to polymerise on the single-stranded DNA that is generated at the site of the DSB after resection. Ultimately, the RecA nucleoprotein filament that is formed at the site of the DSB is visible as a fluorescent focus within the cell (Renzette *et al*., 2005).

It has been reported that the fusion of a *gfp* allele to the C- terminal region of the endogenous *recA* gene resulted in a three-fold reduction in the ability of *E. coli* to survive UV-irradiation, when compared to the *recA*$^{+}$ strain (Renzette *et al*., 2005). Expression of an untagged RecA protein from an ectopic locus or a low copy number plasmid resulted in the complementation of the ability of these RecA-GFP strains to survive UV-irradiation (Lesterlin *et al*., 2014). These data imply that the RecA-GFP protein is not fully functional for the repair of UV-induced DNA damage. Fluorescence labelling of RecA may perturb recombination if the fluorescence tag prevents RecA from interacting with the appropriate presynaptic proteins or impedes the binding of RecA onto single-stranded DNA during the formation of the RecA nucleoprotein filament (Renzette *et al*., 2005). Moreover, a deficiency in recombination may occur if the fluorescence tag causes an anomaly in the interaction between the RecA nucleoprotein filament and the LexA protein during induction of the SOS response (Renzette *et al*., 2005). Therefore, it is essential that the procedure used for fluorescence labelling of the RecA protein does not compromise recombination proficiency, especially if it is to be used as a tool for imaging DSB repair.

Florescence imaging of specific cellular events in *E. coli* is usually conducted in slow growth conditions. These events include the cellular dynamics and localization of the replisome (Reyes-Lamothe *et al*., 2008), organization of sister chromosomes during segregation (White *et al*., 2008), global organization of the nucleoid (Joshi *et al*., 2011) and the physical interaction

of sister chromosomes prior to segregation (Lesterlin *et al*., 2012). The slow growth of *E. coli* in minimal medium prevents the multiple re-initiation of chromosomal replication in most cells. An advantage of the single replication initiation event is that it simplifies the interpretation of data obtained from fluorescence imaging, especially if it involves the visualization of a specific chromosomal locus.

In *E. coli*, the initiation of chromosomal replication at *oriC* results in a bidirectional duplication of the circular chromosome by two replisomes. As the replication forks progress along the chromosome, segregation of sister loci occurs sequentially after a period of cohesion (Nielsen *et al*., 2006). The duration of cohesion for most replicated loci of the *E. coli* chromosome has been reported to be approximately 10 minutes (Reyes Lamothe *et al*., 2008; Joshi *et al*., 2011). Segregation of sister loci to each half of the cell ensures that daughter cells obtain a copy of the genome after completion of cell division.

This chapter describes the procedure that was used for generating a fluorescently labelled RecA protein while ensuring that recombination proficiency was not compromised. Thereafter, the effect of the presence of the fluorescently labelled RecA on viability and growth rate during repair of the replication-dependent DSB was investigated. Additionally, the effect of the DSB on initiation of chromosomal replication for cells with the fluorescently-labelled RecA protein in slow growth conditions was also explored. Finally, data obtained from live-cell imaging of the *lac* locus was used to gain preliminary insight on the effect of the presence of the DSB on cohesion of sister *lacZ* loci.

3.2 Fluorescence labelling of RecA and the site of the replication-dependent DSB at the *lacZ* locus.

3.2.1 A Carboxy-terminal RecA-mCherry fusion protein

Prior to the start of this study, a synthetic plasmid containing a codon diversified *recA* gene was purchased from Eurofins Genomics. The synthetic plasmid was used as a template for amplifying the codon diversified *recA* gene. The gene encoding mCherry was also amplified from a *mCherry*-containing pTOF24 plasmid and was inserted before the stop codon of the synthetic *recA* by crossover PCR. The primers that were used for the crossover PCR were designed such that a three-amino acid linker sequence was situated between the penultimate codon of the synthetic *recA* gene and the first codon of the *mCherry* gene (Renzette *et al*., 2005). Afterwards, the *recA-mCherry* gene was flanked by DNA sequences that had homology to the endogenous *recA* and *recX* genes. The significance of the homology arms was to allow

integration of the *recA-mCherry* gene in tandem with the endogenous *recA* gene, which is immediately upstream of the *recX* gene (Figure 3.1A). The *recA-mCherry* gene flanked by the two homology arms, and the original pTOF24 plasmid (Merlin *et al*., 2002), were digested with XhoI and NotI and ligated together to obtain pDL5196. The plasmid pDL5196 was constructed by Dr Martin White (unpublished) and was used for integrating the *recA-mCherry* gene into the *E. coli* chromosome by plasmid mediated gene replacement. A Codon diversified *recA* gene was used for constructing the *recA-mCherry* fusion gene in order to prevent intergenic recombination with the endogenous gene.

3.2.2 The construct for visualizing the site of the replication-dependent DSB at *lacZ*

The tool that was used for visualizing the site of the replication-dependent DSB at the *lac* locus has previously been reported (White *et al*., 2008; Figure 3.1B). It consists of arrays of *tetO* and *lacO* sequences integrated approximately 6 kb away from the site of the interrupted palindrome at *lacZ*. The arrays of the *tet* and *lac* operator sequences were inserted at the origin-proximal (*cynX* locus) and origin-distal (*mhpC* locus) ends of *lacZ*, respectively. One of the features of this construction is the insertion of an array of three Chi sites on both sides of the interrupted palindrome at *lacZ*. The purpose of the Chi arrays was to minimise the probability of the RecBCD complex degrading the *lac* and *tet* operator arrays during resection. Additionally, the *lacI-Cerulean* and *tetR-YPet* genes were inserted at the *ykgC* locus in the *E. coli* chromosome, with both genes under the control of a strong constitutive promoter (P_{mw1}).

The P_{mw1} promoter was derived from P_{ftsKi} (a weak constitutive promoter; Wang *et al*., 2005) by mutating the -10 and -35 elements to their respective consensus sequences using site-directed mutagenesis. The expression and binding of TetR-YPet and LacI-Cerulean proteins to their respective operator arrays generates twin YPet-Cerulean foci, which was used as a fluorescence label for visualising both ends of the DSB induced at *lacZ*. The endogenous *lacI* gene was deleted from the *E. coli* chromosome to prevent the LacI protein from competing with LacI-Cerulean for the binding of the *lac* operator array.

3.3 Effect of the presence of the *recA-mCherry* gene on cell viability and growth rate following formation of the replication-dependent DSB

3.3.1 Effect of the presence of the *recA-mCherry* gene on cell viability following formation of the replication-dependent DSB

The *recA-mCherry* gene was integrated into the chromosome of the strain containing a 246 bp interrupted palindrome at *lacZ* and the corresponding control strain which did not have the palindrome inserted at *lacZ*. For both strains, the *lacZ* gene was flanked by the *lac* and *tet*

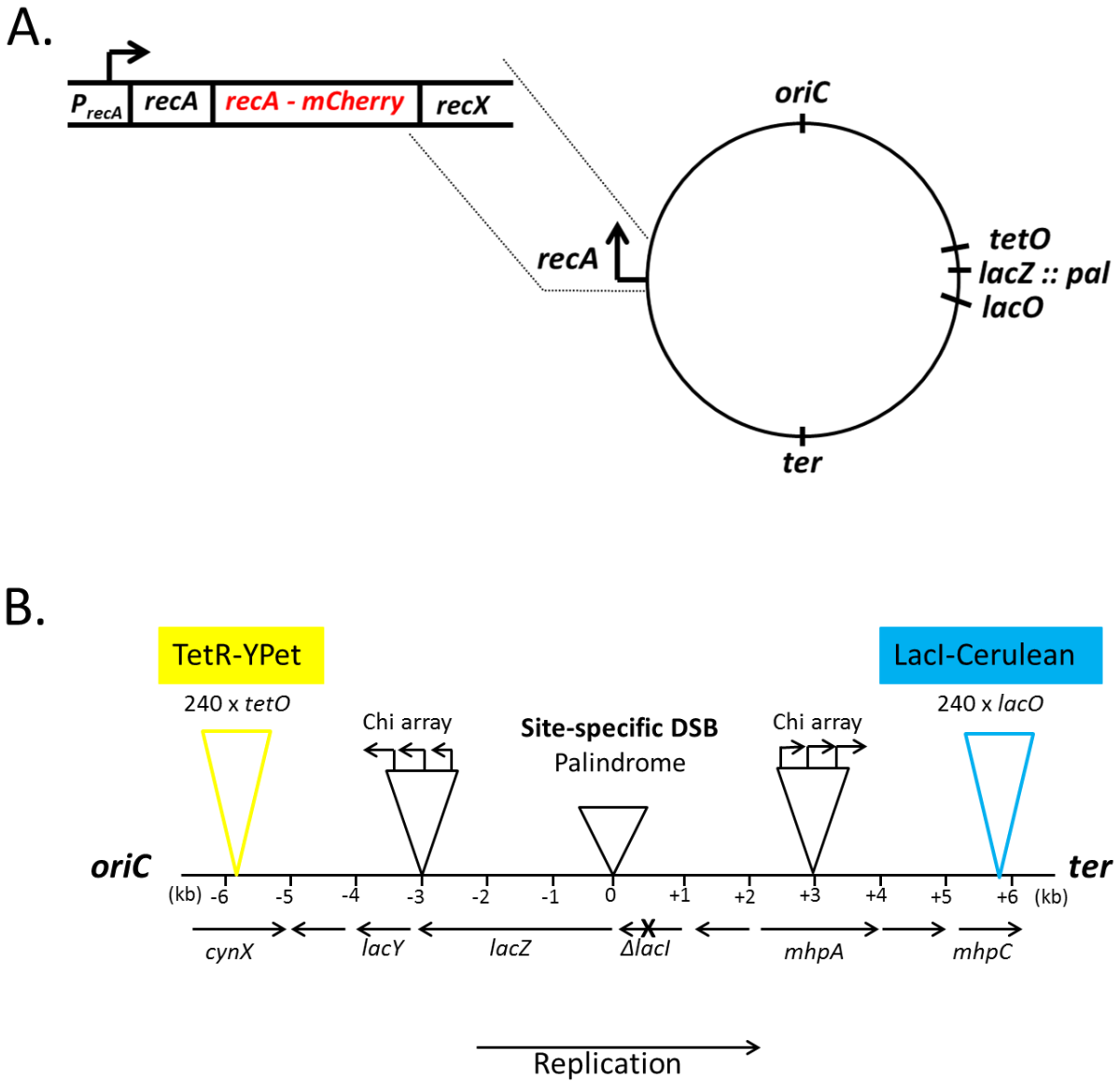


Figure 3.1. Tools for visualizing RecA and the site of the replication-dependent DSB at *lacZ*. (A) Schematic representation of the modified *recA* operon. The *mCherry* gene was fused to the C-terminal of a codon diversified *recA* gene and inserted in tandem to the endogenous *recA* gene of the *E. coli* chromosome. (B) Tool for visualising the site of the replication-dependent DSB at *lacZ* (White *et al*., 2008). Arrays of *tetO* and *lacO* were inserted approximately 6kb away from the site of an interrupted palindrome at *lacZ*. The *tetR-YPet* and *lacI-Cerulean* genes were expressed from a strong constitutive promoter (P_{mw1}) at the *ykgC* locus. Arrays of three chi sites was also inserted 3kb away from the interrupted palindrome at *lacZ* to minimise the probability of degradation of the *tetO* and *lacO* arrays by the RecBCD complex. The endogenous *lacI* was deleted to prevent LacI from competing with LacI-Cerulean for binding to the *lacO* array.

operator arrays while the LacI-Cerulean and TetR-YPet fluorescence proteins were constitutively expressed from the P_{mw1} promoter in the *ykgC* locus. In addition, both strains had the genes encoding the SbcCD endonuclease placed under the control of an arabinose inducible promoter. The effect of the presence of the *recA-mCherry* gene on cell viability following the formation of the replication-dependent DSB at *lacZ* was investigated by a spot test assay. The appropriate control strains were constructed to enable an explicit interpretation of the data obtained from this assay. These controls include strains with only the endogenous *recA* gene (DL5516 and DL5521), strains with only the integrated *recA-mCherry* gene (without the endogenous *recA*; DL5866 and DL5942) and strains with none of these *recA* genes (DL5941 and DL5924).

The spot test assay showed that the strains containing the tandem insertion of the *recA-mCherry* gene with the endogenous *recA* gene (DL5525 and DL5528) had similar viability as the wild type strains (DL5516 and DL5521) in the presence or absence of DSBs (Figure 3.2, Glucose/Arabinose plate). Following chronic expression of the SbcCD endonuclease (Figure 3.2, Arabinose plate), it was observed that the palindrome-containing strain with only the *recA-mCherry* gene (DL5866) exhibited a profound loss of viability, which was similar to the loss of viability of the *recA*$^-$ strain (DL5941). These observations indicate that a C-terminal fusion of *mCherry* to *recA* generates a protein that has a compromised functionality for repair of the DSB that was formed at *lacZ*. However, the presence of the untagged RecA protein restores recombination proficiency in the strain with the *recA-mCherry* gene. In the absence of the palindrome-induced DSBs, the strains with both alleles of *recA* (DL5525 and DL5528) exhibited no loss of viability after exposure to UV-light, unlike the strains with only the *recA-mCherry* gene (DL5888 and DL5942; Figure 3.2, Glucose/UV plate). The results from this spot test assay demonstrate that the tandem insertion of the *recA-mCherry* gene with the endogenous *recA* gene has no detectable effect on viability following formation of the replication-dependent DSBs and the UV-induced DNA damage.

3.3.2 Effect of the presence of the *recA-mCherry* gene on growth rate following formation of the replication-dependent DSB

The growth profiles of the *E. coli* strains that were used for the spot test assay are presented in Figure 3.3. These strains were grown in M9-glycerol minimal media and the expression of the SbcCD endonuclease was induced by the addition of 0.2% arabinose. The growth profiles revealed that the tandem insertion of the *recA-mCherry* gene with the endogenous *recA* gene

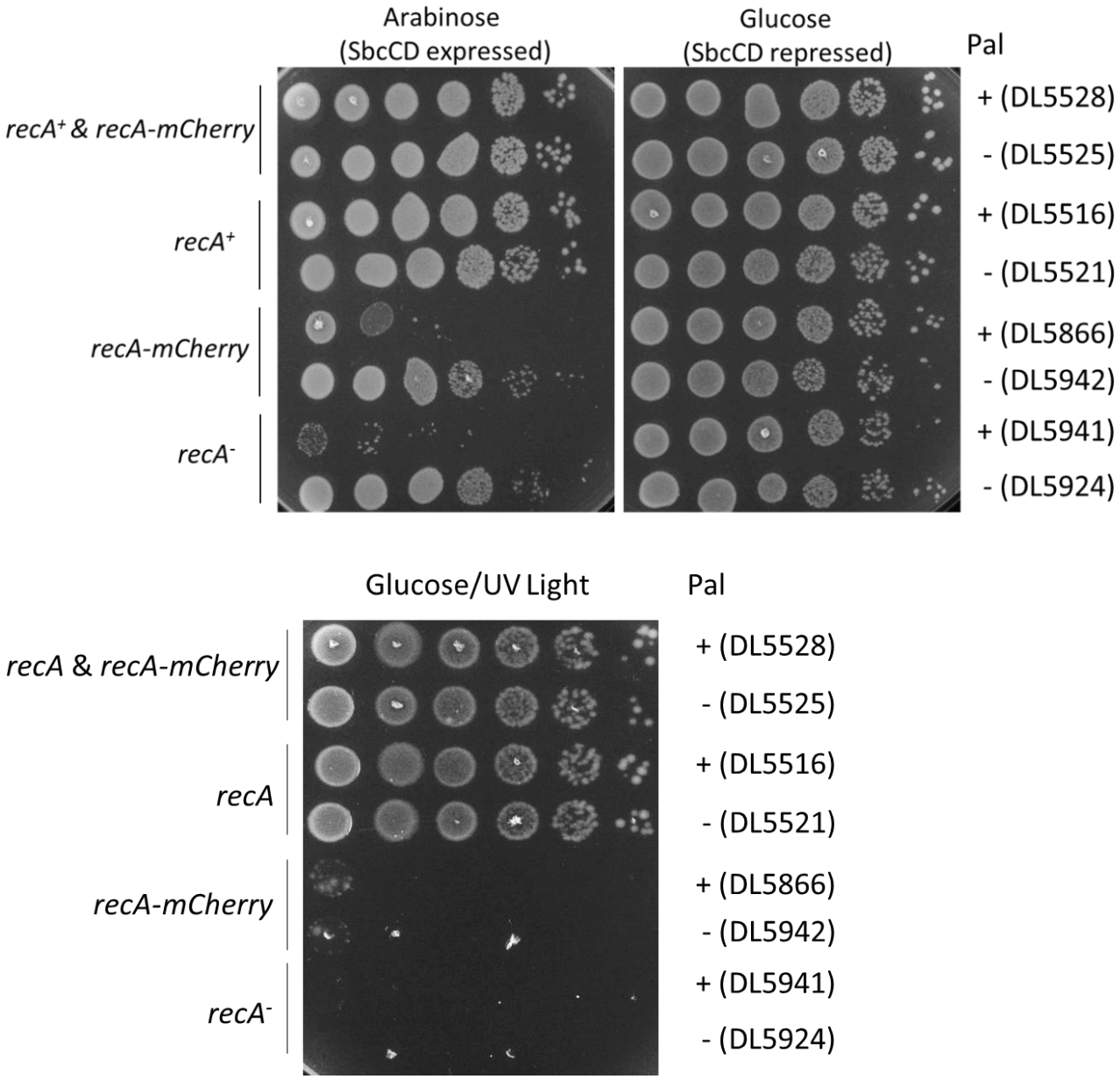


Figure 3.2. Spot test assay for detecting the effect of the *recA-mCherry* gene on cell viability following formation of the replication-dependent DSB. Overnight cultures grown in L broth were normalised to an OD_{600}=1 and diluted serially (10-fold) before being spotted on LB plates supplemented with either 0.5% of glucose or 0.2% of arabinose. One of the LB/Glucose plates was irradiated with UV-light (0.01J) and all plates were incubated overnight at 37˚C. The strains that were used for the assay had both the endogenous *recA* and the integrated *recA-mCherry* genes (DL5525 and DL5528), or only the endogenous *recA* gene (DL5516 and DL5521), or only the integrated *recA-mCherry* gene (DL5866 and DL5942) or none of the *recA* alleles (DL5941 and DL5924). All strains had both the *tetO* and *lacO* arrays and the expression of the LacI-Cerulean and TetR-YPet proteins were under the control of P_{mw1} .Pal represents the presence (+) or absence (-) of the interrupted palindrome at *lacZ*.

had no detectable effect on the growth rate, when compared to the *recA*$^{+}$ strain, following DSB formation (Figure 3.3, *lacZ::pal*). However, the palindrome-containing strain with only the *recA-mCherry* gene (DL5866) grew at a slower rate compared to the *recA*$^{+}$ strain (DL5516). In the absence of the interrupted palindrome at *lacZ*, the growth rate of the strain with only the *recA-mCherry* gene (DL5942) was comparable to that of the strain with both alleles of *recA* (DL5525; Figure 3.3, *lacZ*$^{+}$). These observations, together with the results from the spot test assay, demonstrate that the functionality of the RecA-mCherry protein was compromised for the repair of the DSB that was formed at *lacZ.* Additionally, the growth profiles indicate that the RecA-mCherry protein is not entirely non-functional for repair of the palindrome-induced DSBs during the 7 hours of SbcCD expression because the strain with only the *recA-mCherry* gene had a faster growth rate compared to the *recA*$^{-}$ mutant (Figure 3.3, *lacZ::pal*).

The data from the growth profiles also revealed that the growth rate was slower in the recombination-proficient strains that were subjected to DSBs (DL5516 and DL5528) compared to the corresponding control strains in which the palindrome-induced DSB formation was absent (DL5521 and DL5525). The estimated generation time increased by approximately 10 minutes for the palindrome-containing strain with only the endogenous *recA* gene, while the corresponding strain with both alleles of the *recA* gene had an increase in generation time of 7 minutes following expression of the SbcCD endonuclease (Table 3.1). These results indicate that the 7 hours of SbcCD expression in the palindrome-containing strains had an effect on the growth rate of the recombination-proficient strains even though the cell viability was not affected by the chronic SbcCD expression (Figure 3.2, arabinose plate).

3.4 Effect of the replication-dependent DSB on initiation of chromosomal replication in slow growth conditions

For the purpose of live-cell imaging, the *E. coli* strains were grown in minimal medium to ensure the occurrence of a single replication-dependent DSB at *lacZ* per cell for every generation following the induction of SbcCD expression. This is because *E. coli* cells have a single round of initiation of chromosomal replication per cell cycle when grown at 37˚C in M9 minimal medium supplemented with 0.2% glycerol as carbon source (Wang *et al*., 2005).

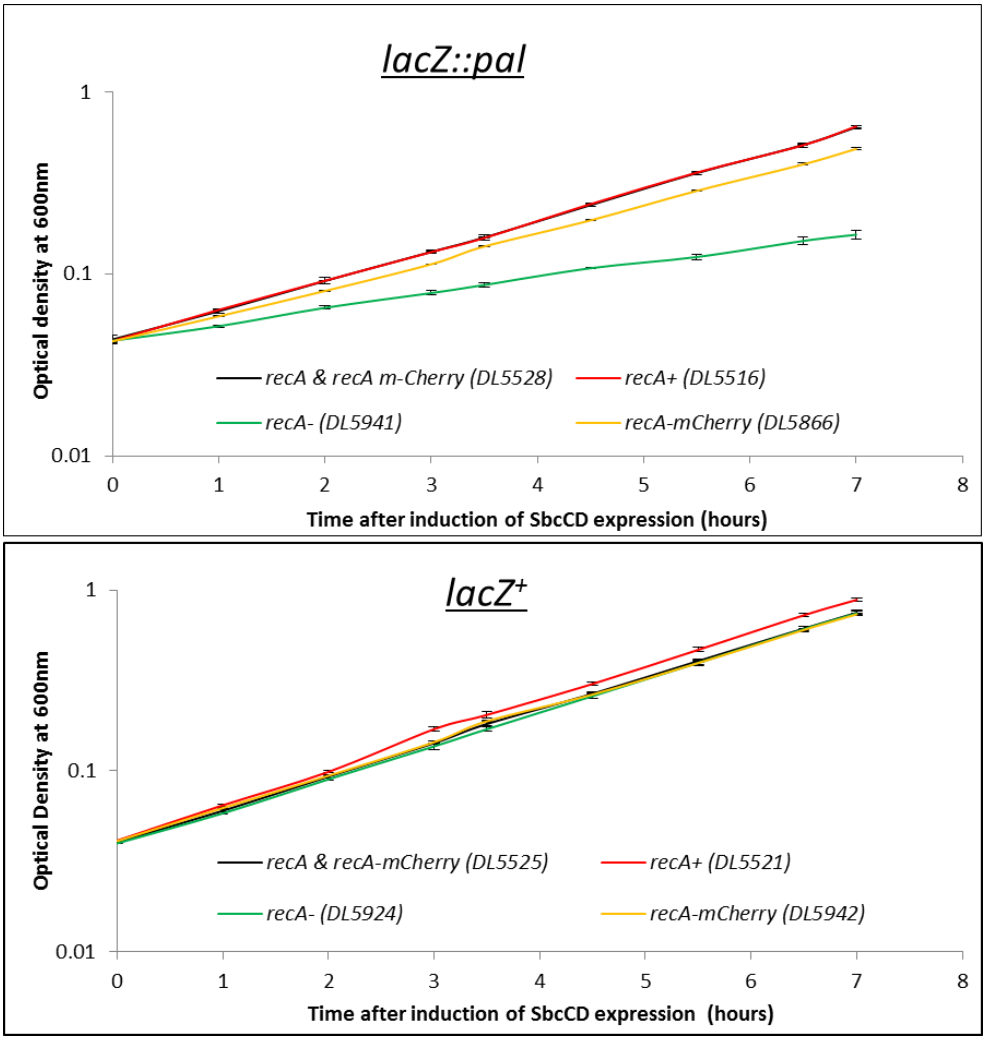


Figure 3.3. Effect of the presence of the *recA-mCherry* gene on *E. coli* growth rate following formation of the replication-dependent DSB. Overnight cultures were grown at 37°C in M9 minimal media supplemented with 0.2% glycerol. The overnight cultures were diluted and grown to exponential phase for 3hrs (OD_{600} < 0.3). The exponentially growing cultures were diluted to an OD_{600}= 0.04 in the same growth medium with the addition of 0.2% arabinose. The diluted cultures were maintained in exponential phase by dilution while measuring the OD_{600} at the indicated time intervals for 7 hours. Error bars represent the standard error of the mean of three independent experiments.

Table 3.1: Effect of the replication-dependent DSB on generation time of the recombination-proficient strains.

Genotype of the *recA* allele of the recombination-proficient strains	**Generation time (minutes)**	
	lacZ::pal	*lacZ*$^+$
recA$^+$	106.8 $\pm$ 0.6	96.6 $\pm$ 0.4
recA & recA-mCherry	107.2 $\pm$ 0.9	100.0 $\pm$ 0.7

In this study, the number of replication initiation events was estimated by measuring the DNA content of cells at exponential phase following treatment with rifampicin and cephalexin. Rifampicin is a drug that inhibits the initiation of RNA synthesis (Wehrli and Staehelin, 1971) and cephalexin inhibits cell division. Treatment with rifampicin allows the existing replication forks to proceed with chromosomal replication while preventing re-initiation of replication because the latter process requires transcription (Lark, 1972; Messer, 1972; von Meyenburg *et al*., 1979; Ogawa *et al*., 1985) and *de novo* protein synthesis (Maaløe and Hanawalt, 1961; Lark *et al*., 1963). After treatment with both drugs, the DNA was stained with propidium iodide and the number of chromosomes in the cell population was estimated by flow cytometry. The number of chromosomes per cell provides an indication of the number of chromosomal origins (*oriC*) in the cell population during exponential phase of growth, which also signifies the number of replication initiation events per cell.

The DNA contents of the *E. coli* strains containing both alleles of the *recA* gene (DL5525 and DL5528), grown in M9-glycerol minimal medium, are presented in Figure 3.4 and Table 3.2. Expression of the SbcCD endonuclease was induced by the addition of 0.2% arabinose to cells in exponential phase of growth. In the absence of the interrupted palindrome at *lacZ*, 87% of the population had two genome equivalents (chromosomes) after treatment of the cells in exponential phase with cephalexin and rifampicin (Run-out of Figure 3.4, Table 3.2). This observation demonstrates that most of the cell population had two copies of *oriC* when grown in M9-glycerol minimal medium, thereby confirming the existence of a single replication initiation event in the majority of the cell population. Following expression of the SbcCD endonuclease in the strain with the interrupted palindrome inserted at *lacZ* (DL5528), the cell population with two chromosomes reduced to approximately 83% (Table 3.2). Statistical analysis (*t-test*) revealed that this reduction in percentage of cells with two chromosomes was significant.

Though very rare, DSB formation also resulted in an increase (by three-fold) in asynchronous initiation of chromosomal replication, which was denoted by the presence of three chromosomes in the cell after treatment with cephalexin and rifampicin. The percentage of cells with either a single chromosome or four chromosomes was not significantly affected by the DSB formation. This indicates that the increase in asynchrony, following DSB formation, originated predominantly from the cell population with two copies of *oriC*.

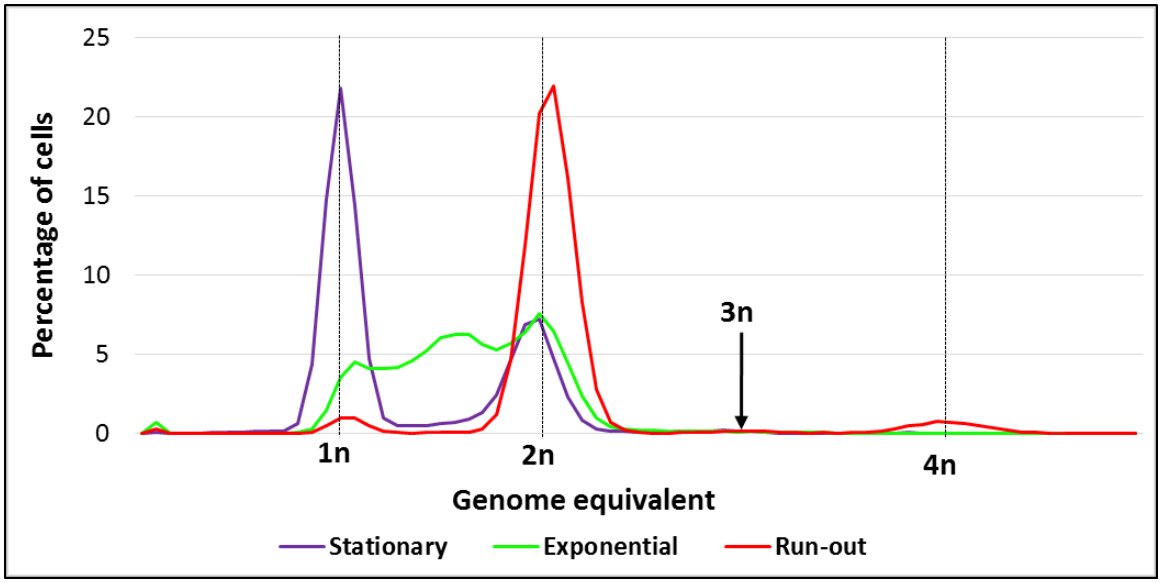


Figure 3.4. DNA content of strain DL5525 during exponential growth in minimal medium. Overnight cultures (stationary phase) were diluted and grown to exponential phase in M9-glycerol minimal medium at 37˚C. The cells in exponential phase were re-diluted in the same medium supplemented with 0.2% arabinose to enable the expression of the SbcCD endonuclease. After 2 hours of growth, rifampicin and cephalexin were added to the cells and were grown for a further 3 hours to allow completion of chromosomal replication by the progressing replication forks (Run-out). Aliquots of the cells during stationary phase, exponential phase and after the run-out were fixed in ethanol, and the DNA was stained with propidium iodide prior to analysis by flow cytometry.

Table 3.2. Effect of the replication-dependent DSB on the number of chromosomal origins (*oriC*) during exponential growth in minimal medium.

Number of *oriC* (Run-out)	**Percentage of cells**	
	DL5525 (DSB$^-$)	**DL5528 (DSB$^+$)**
1 *oriC*	3.8 (± 0.5)	3.9 (± 0.5)
2 *oriC*$^{\#}$	87 (± 1)	82.9 (± 0.6)
3 *oriC*$^{\#}$	1.13 (± 0.04)	3.7 (± 0.3)
4 *oriC*	5.8 (± 0.2)	5.6 (± 0.2)
Average *oriC*/cell	**2.091 (± 0.004)**	**2.11 (± 0.01)**

The standard error of the mean of three and four independent experiments are indicated in brackets for DSB$^-$ and DSB$^+$, respectively. # represents a statistically significant difference (t-test; p-value < 0.05) between DSB$^-$ and DSB$^+$ for the indicated number of *oriC*.

From Table 3.2, it can be deduced that the duration of the ‘’B phase’’ (eukaryotic G1 phase) of the cell cycle was very transient because less than 4% of the cell population in exponential phase of growth had one genome equivalent after treatment with rifampicin and cephalexin. Furthermore, in M9-glycerol medium, approximately 6% of the cell population re-initiated replication in both chromosomes prior to cell division, hence the presence of four genome equivalents after treatment with the drugs. Primarily, these data confirm that the strains that would be used for live-cell fluorescence imaging (DL5525 and DL5528) undergo a single replication initiation event in most of the cell population during exponential growth in M9-glycerol minimal medium.

3.5 Visualization of the site of the replication-dependent DSB in *E. coli* cells at exponential phase of growth.

3.5.1 Number of YPet_Cerulean foci per cell in slow growth condition

After confirming that most of the cells grown in M9-glycerol medium had a single replication initiation event per cell cycle, this study proceeded with the visualization of the fluorescently labelled *lac* locus, which is the site of DSB formation in the palindrome-containing strain. The purpose for visualization of the fluorescently-labelled *lac* locus was to ascertain if the presence of the replication-dependent DSB had an effect on the cohesion of sister *lac* loci. It has previously been reported that the fluorescently-labelled *lac* locus was visualized as a twin YPet_Cerulean focus by conventional fluorescence microscopy (White *et al*., 2008). Therefore in *E. coli* cells with a single replication initiation event per cell cycle, the presence of one YPet_Cerulean focus in a cell (Figure 3.5A, red arrow) could represent either a new-born cell that is yet to commence chromosomal replication, or a cell that has initiated replication and has not yet replicated the *lac* locus, or a cell with a replicated *lac* locus but has not yet segregated the sister loci due to cohesion. Two YPet_Cerulean foci (Figure 3.5A, white arrow) in a cell correspond to a cell that has replicated and segregated the sister *lac* loci. Hence the estimation of the number of YPet_Cerulean foci per cell could provide useful insight on the effect of DSB formation on cohesion of the sister *lac* loci.

In the absence of DSB formation, approximately 59% of the cells had a single YPet_Cerulean focus while 40% of the cells had two YPet_Cerulean foci per cell (Figure 3.5B). From Table 3.2, it can be deduced that less than 4% of the cells with a single YPet_Cerulean focus are yet to initiate chromosomal replication. Hence, most of the cells with a single YPet_Cerulean focus have a replicating chromosome which may or may not have replicated the *lac* locus. Following replication of the *lac* locus and cleavage of the DNA hairpin structure in the palindrome-

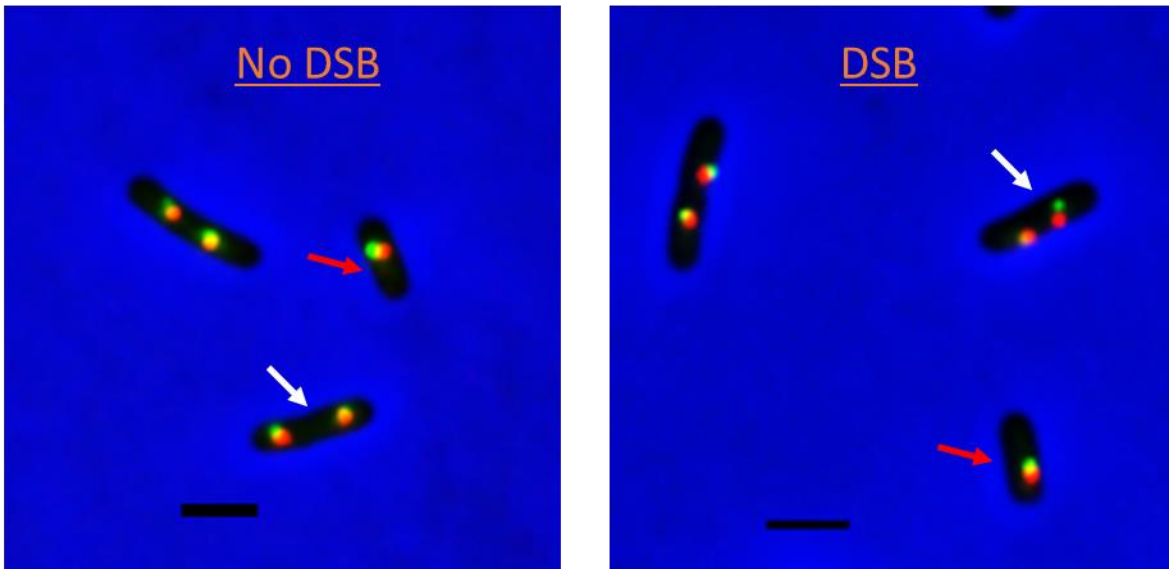
containing strain (DL5528), it is anticipated that the cell will initiate DSB repair during the period of cohesion of the sister *lac* loci. This is because the need for an extensive homology search will be eliminated during pre-synapsis. In fact, the modest increase in generation time following DSB formation (Table 3.1) might suggest that the repair of the DSB is rapid, thereby favouring the initiation of the repair event during the period of cohesion of the sister *lac* loci. If repair of the replication-dependent DSB was initiated during the period of cohesion of the sister *lac* loci, it should be completed before segregation of that locus. Consequently, the duration of cohesion of the sister *lac* loci might be extended due to the repair of the replication-dependent DSB. Hence, the percentage of cells with one YPet_Cerulean focus would increase whilst those with two YPet_Cerulean foci would decrease due to formation and repair of the replication-dependent DSB at *lacZ.*

From Figure 3.5B, it was observed that DSB formation resulted in a significant reduction in the percentage of cells with one YPet_Cerulean focus. In addition, there was an increase in the percentage of cells without the YPet_Cerulean focus. The percentage of cells with either two or three YPet_Cerulean foci were both not significantly affected by DSB formation, as confirmed by a *t-test* statistical analysis. These observations indicate that the reduction in the percentage of cells with a single YPet_Cerulean focus, in the presence of DSB formation, was due to the increase in percentage of cells without the YPet_Cerulean focus. Therefore, the percentage of cells with a single YPet_Cerulean focus could not be used to predict the effect of the replication-dependent DSB on the duration of cohesion of sister *lac* loci. However, the data from the percentage of cells with two YPet_Cerulean foci may suggest that the duration of cohesion might be unaffected by the presence of the replication-dependent DSBs.

3.5.2 Number of YPet_Cerulean foci per cell length in slow growth conditions

Cell division in Enterobacteriaceae, which includes *E. coli*, was previously reported to occur at a well-defined cell mass (Schaechter *et al*., 1958). However, the initiation of cell division at the expected cell mass can be inhibited or delayed by the formation of DNA damage. Inhibition of cell division is essential for efficient repair of the DNA damage and segregation of intact genomes into the daughter cells (Huisman and D'Ari, 1981). Accordingly, it has been shown that DSB formation during three generations resulted in an increase in cell length by approximately 11% (White, unpublished). In addition, chronic DSB formation caused an increase in cell length by approximately 34% compared to control strains that were not subjected to DSB formation (Darmon *et al*., 2014). The dramatic increase in average cell length during chronic DSB formation was due to an overall increase in cell length for most of the cells

A.



B.

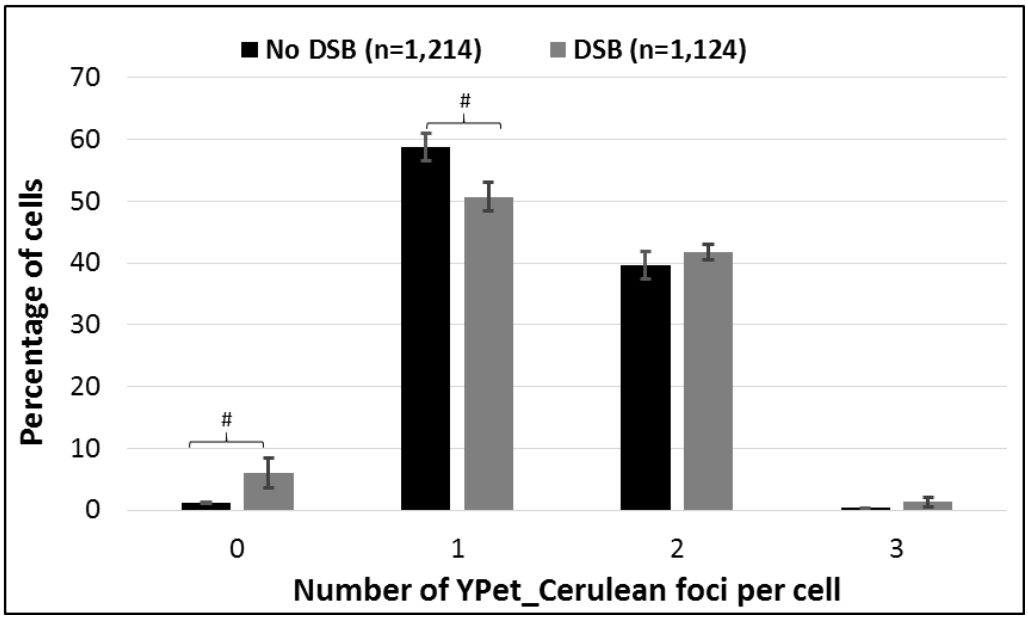


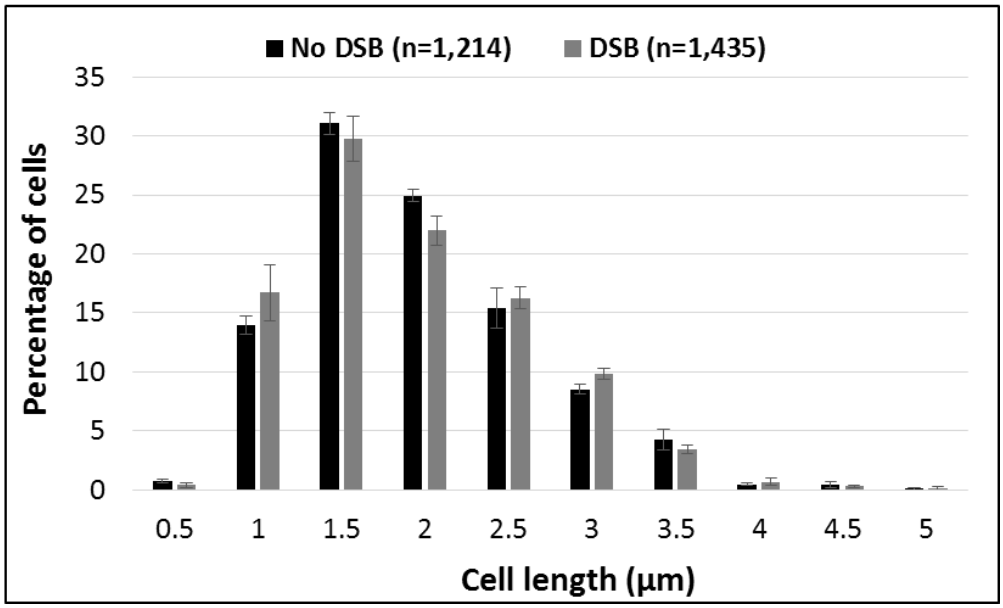
Figure 3.5. Effect of the replication-dependent DSB on the number of YPet_Cerulean foci per cell in slow growth conditions. Strains DL5525 (*lacZ*$^{+}$) and DL5528 (*lacZ::pal*) were grown in M9-glycerol minimal medium at 37˚C. The SbcCD endonuclease was expressed for 2 hours by the addition of 0.2% arabinose to the growth medium. An aliquot of 10 µl of cells was added to an agarose pad for imaging by snapshot microscopy. (A) Visualization of the *lac* locus in *E. coli* cells grown in minimal medium. The YPet and Cerulean foci are indicated as green and red spots respectively. Scale bar represents 2 µm. (B) Number of YPet_Ceulean foci per cell in slow growth conditions. Error bars represent the standard error of the mean of three independent experiments. # represents a statistically significant difference (t-test; p-value < 0.05) between No DSB (DL5525) and DSB (DL5528) for the indicated number of YPet_Cerulean foci per cell.

that were analysed and the filamentation of approximately 17% of the cell population. These data on the increase in cell length following formation of the replication-dependent DSB were obtained from studies conducted in fast growth conditions (LB medium). Based on this information, chronic replication-dependent DSBs were induced in cells growing in M9-glycerol medium in order to investigate the effect of the DSB on the number of YPet_Cerulean foci per cell length. The formation of chronic DSBs was chosen for this specific purpose because chronic DSBs had a larger effect on delaying cell division in fast growth conditions (Darmon *et al*., 2014). An advantage of determining the number of YPet_Cerulean foci per cell length is that it accounts for the delay of cell division during repair of the replication-dependent DSB. If repair of the replication-dependent DSB was initiated prior to segregation of the sister *lac* loci, and caused a substantial extension in the duration of cohesion, then cells with a single YPet_Cerulean focus would be expected to be longer in size following DSB formation. Ultimately, the calculated mean of YPet_Cerulean foci per cell length would be expected to be smaller for the cell population subjected to the replication-dependent DSB, compared to the corresponding cells without the induced DSB.

The effect of chronic DSB formation on cell length, for slow growth conditions, is presented in Figure 3.6. The graph in Figure 3.6A shows that most of the cell population had a cell length from 1 µm to 3.5 µm. The distribution of cell length following chronic DSB formation was identical to that obtained in the absence of DSBs. Notably, filamentation of cells was very rare following chronic DSB formation. Statistical (*t-test*) analysis was used to demonstrate that there was no significant difference between the mean cell lengths in the absence and presence of DSBs when cells were grown in M9-glycerol medium at 37˚C (Figure 3.6B).

This study also investigated the effect of DSB formation on the induction of the SOS response in cells grown in M9-glycerol medium at 37˚C. The genes encoding the TetR-YPet and LacI-Ceulean proteins were deleted from the strains that were used for generating Figure 3.6 (DL5525 and DL5528). Afterwards, the strains obtained (DL6107 and DL6311) were transformed with a plasmid (pDL2856) which contained a *gfp* allele under the control of the *sfiA* promoter. Following a DNA damage, expression of the SfiA protein from its native promoter is elevated via induction of the SOS response, thereby ensuring that cell division is inhibited until completion of the repair of the DNA damage. Thus the strains containing a *gfp* allele under the control of the *sfiA* promoter was used to investigate whether expression of the GFP protein from the *sfiA* promoter was elevated in cells that were grown in M9-glycerol

A.



B.

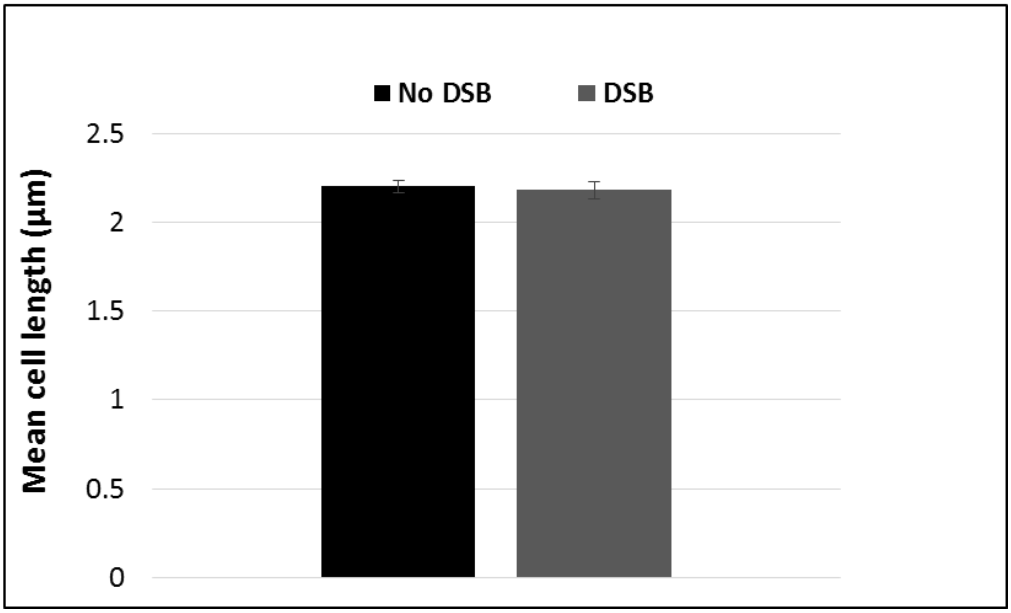


Figure 3.6. Effect of replication-dependent DSB formation on cell length in slow growth conditions. (A) Distribution of cell length in the absence (DL5525) and presence (DL5528) of replication-dependent DSB formation. Error bars represent the standard error of the mean of three independent experiments. (B) Mean cell length in the absence (DL5525) and presence (DL5528) of replication-dependent DSB formation. Error bars represent the standard error of the mean of three independent experiments. Statistical analysis (t-test) revealed that there was no significant difference between the mean cell length in the absence and presence of DSB formation (p-value > 0.05).

medium following DSB formation. The genes encoding the TetR-YPet and LacI-Cerulean proteins were deleted because the emission spectra of the fluorescent proteins encoded by these genes overlap with the GFP emission spectrum. The data obtained from the flow cytometric analysis of the GFP fluorescence signals of these strains are shown in Figure 3.7. The data indicated that expression of the GFP protein from the *sfiA* promoter was elevated in response to DSB formation. Hence, formation of the replication-dependent DSB at the *lacZ* locus, in cells grown in M9-glycerol medium, resulted in the induction of the SOS response even though the average cell length was not significantly affected by the presence of the DSB (Figure 3.6B).

Box plots displaying the effect of DSB formation on the number of YPet_Cerulean foci in relation to the cell length are shown in Figure 3.8A. Following DSB formation, the box plots revealed that there was no detectable increase in the median cell length of cells with one YPet_Cerulean focus, when compared to the corresponding cells that were not subjected to DSBs. The same observation was also detected for cells with two YPet_Cerulean foci. In contrast, the median cell length increased by two-fold in response to DSB formation for cells with no YPet-Cerulean foci. Statistical analysis (*t-test*) of the mean YPet_Cerulean foci per cell length indicated that there was no significant difference for the values obtained in the absence and presence of the replication-dependent DSB (Figure 3.8B). Ultimately, these observations demonstrate that the formation of chronic replication-dependent DSBs resulted in an increase in the length of cells lacking YPet_Cerulean foci, possibly because these cells might have generated a DSB on both the leading and lagging strand templates at the *lacZ* locus. The data presented in this chapter was insufficient to deduce the effect of the presence of DSBs on the duration of cohesion of sister *lac* loci for cells grown in M9-glycerol medium at 37˚C.

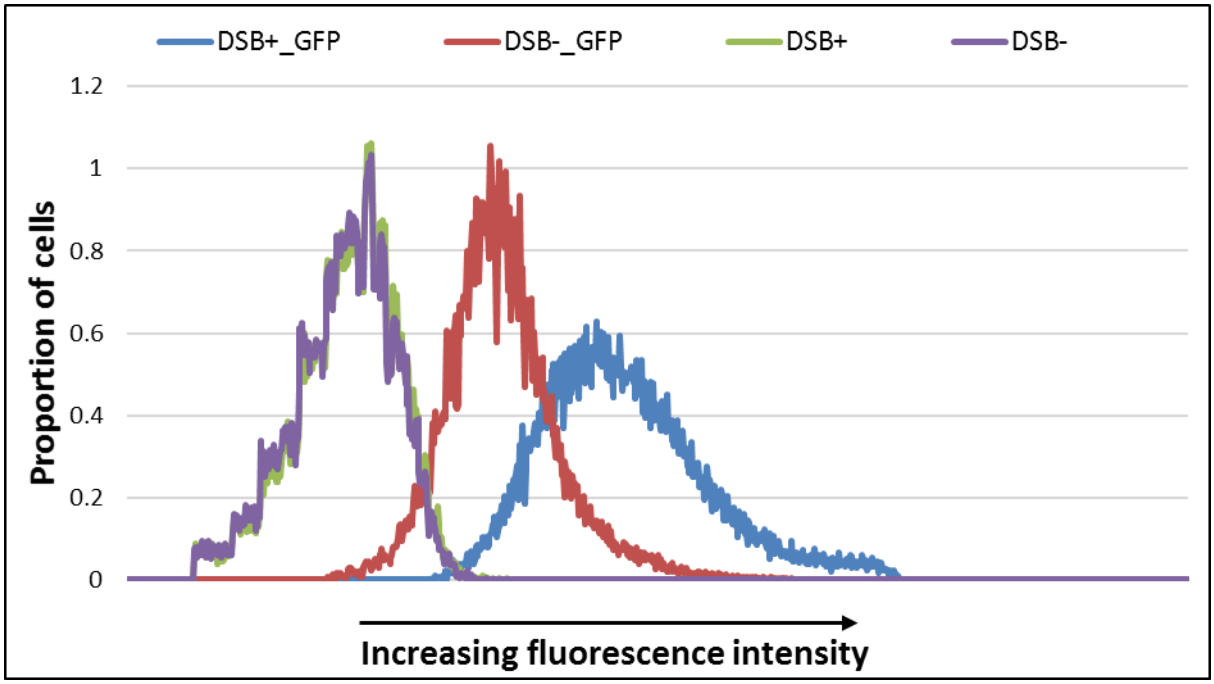
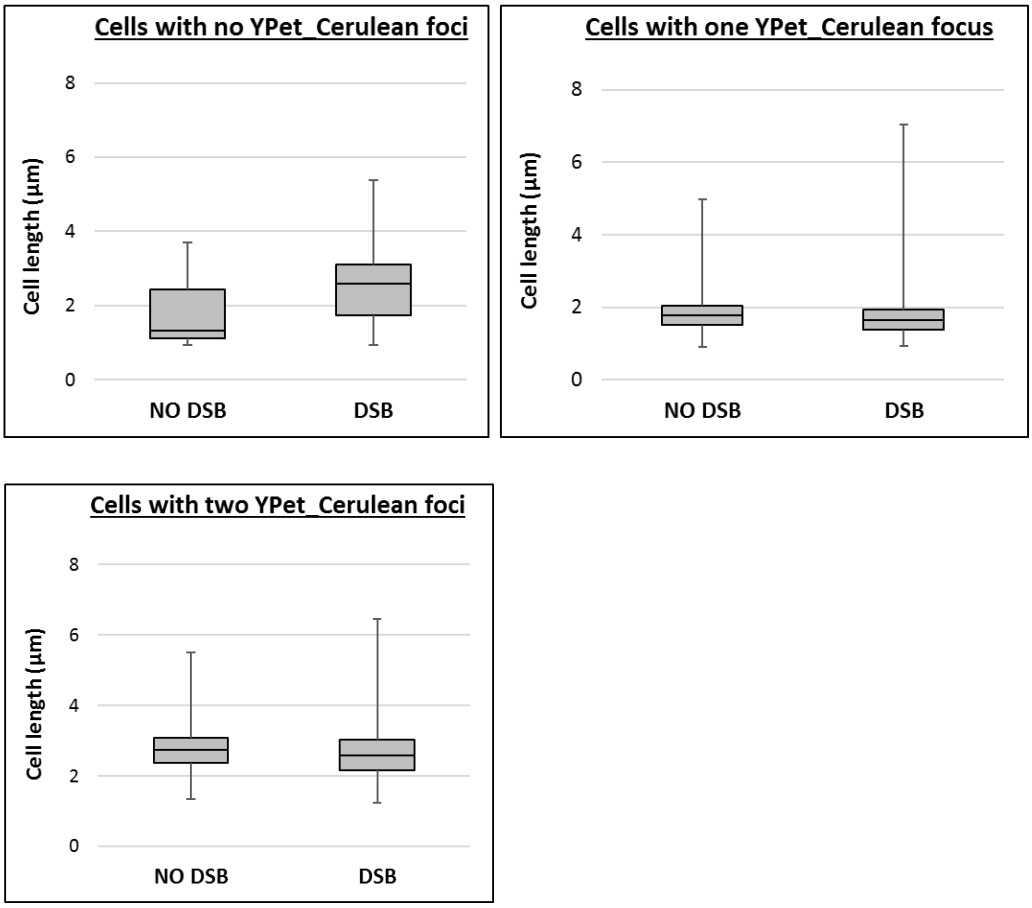


Figure 3.7. Effect of replication-dependent DSB formation on the induction of the SOS response in slow growth conditions. Overnight cultures were diluted and grown to exponential phase in M9-glycerol minimal medium at 37˚C. The cells in exponential phase were re-diluted in the same medium supplemented with 0.2% arabinose to enable the expression of the SbcCD endonuclease. After 2 hours of growth, the cell pellets were obtained via centrifugation and were re-suspended in 1X PBS followed by analysis by flow cytometry. The plasmid pDL2856, which contained a *gfp* allele under the control of the *sfiA* promoter was either present or absent in the the strains that were used for this analysis. The strains that were used are DL6107 (DSB^+) and DL6311 (DSB^-), which did not contain the plasmid pDL2856, and DL6443 (DSB^+) and DL6444 (DSB^-), which contained the plasmid pDL2856. DSB^+ represents the presence of DSB induction while DSB^- represents the absence of DSB induction.

A.



B.

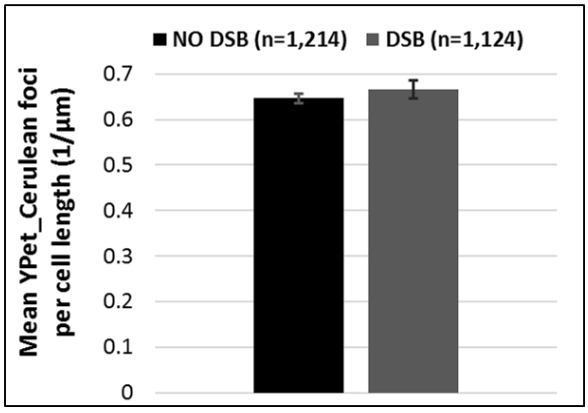


Figure 3.8. Effect of replication-dependent DSB formation on the number of YPet_Cerulean foci in relation to cell length. The cell length and the number of YPet_Cerulean foci for each cell was determined for DL5525 (DSB^-) and DL5528 (DSB^+). (A) Box plots for identifying the effect of DSB formation on the correlation between cell length and number of YPet_Cerulean foci. (B) The mean YPet_Cerulean foci per cell length. Error bar represents the standard error of the mean of three independent experiments. Statistical analysis (t-test) revealed that there was no significant difference between the mean foci per cell length in the absence and presence of DSB formation (p-value > 0.05).

3.6 DISCUSSION

The insertion of a 246 bp interrupted palindrome at the *lacZ* locus results in the formation of a replication-dependent DSB on the lagging strand template in the presence of the SbcCD endonuclease (Eykelenboom *et al*., 2008). Repair of the replication-dependent DSB occurs via homologous recombination using the undamaged sister chromosome which is derived from the leading strand template. The data shown in this chapter indicate that the presence of a C-terminal *recA-mCherry* fusion gene, instead of the endogenous *recA* gene, result in a defect in the repair of the DSB. Nonetheless, the tandem insertion of the *recA-mCherry* gene with the endogenous *recA* gene restores recombination proficiency during repair of the DSB by homologous recombination. The observation that the RecA-mCherry fusion protein could not act as a substitute for the native RecA protein implies that the fluorescently-labelled protein is not capable of performing either some or all the functions of the native RecA protein during DSB repair. These functions include the interaction of RecA with the RecBCD complex which leads to formation of the RecA nucleoprotein filament, the interaction of the RecA nucleoprotein filament with the LexA protein during SOS induction, and the homology search and strand invasion reactions that occur after resection.

It has previously been shown that *E. coli* strains containing a C-terminal *recA-gfp* gene, instead of the endogenous *recA* gene, can induce the SOS response and form nucleoid-associated RecA-GFP foci after exposure to UV-light (Renzette *et al*., 2005). Even though the RecA-GFP protein had a compromised recombination activity, its ability to form nucleoid associated foci demonstrates that the fluorescently-labelled protein is capable of nucleating at the sites of UV-induced DSBs during formation of the nucleoprotein filament. Therefore, it can be inferred that the defect in repair of the replication-dependent DSB in the strain containing only the *recA-mCherry* gene may not be due to the inability to form the nucleoprotein filament nor induce the SOS response. An anomaly in the strand exchange reaction could be a plausible explanation for the defective DSB repair in the strain encoding only the *recA-mCherry* gene, probably because the presence of successive mCherry proteins within the nucleoprotein filament could sterically inhibit efficient binding and interaction of RecA with the homologous dsDNA template. In the presence of both the native RecA and the RecA-mCherry proteins, the recombination proficiency was not compromised following formation of the replication-dependent DSB. From these observations, it can be inferred that the binding of the RecA-mCherry protein did not sterically inhibit the binding of the native RecA protein during formation of the RecA nucleoprotein filament. Moreover, the presence of the native RecA

protein did not competitively inhibit binding of the RecA-mCherry protein to the site of the replication-dependent DSB during formation of the nucleoprotein filament because the latter generated a fluorescence focus at the site of the DSB during repair (discussed in Chapter 4). Thus, within the nucleoprotein filament, the presence of the RecA-mCherry protein does not interfere with the activities of the native RecA protein which are essential for repair of the replication-dependent DSB.

In order to generate a single replication-dependent DSB in the cell per generation, the *E. coli* strains were grown in M9-glycrerol medium, which eliminates multiple re-initiation of chromosomal replication in most of the cell population. For the slow growth condition that was used in this study, it was observed that the formation of the replication-dependent DSB resulted in a 7% increase in generation time for the strains expressing both the native and fluorescently-labelled RecA proteins. It can be inferred that the increase in generation time in response to DSB formation was predominantly due to the slow growth condition instead of the presence of the *recA-mCherry* gene in strains containing the endogenous *recA* gene. This is because, it was shown in this study that there was a substantial increase in the generation time of strains containing only the endogenous *recA* gene following DSB formation. Moreover, previous studies have reported that there was no significant increase in the generation time of cells which encode only the endogenous RecA protein following palindrome-induced DSB formation under fast growth conditions (Darmon *et al*., 2014).

For slow growing *E. coli* cells, the cell cycle comprises of the B-phase (duration from completion of cell division to initiation of replication in the new-born cell), C-phase (duration for replication of the chromosome) and D-phase (duration for occurrence of cell division after completion of replication). Analysis of DNA content by flow cytometry indicated that the formation of the replication-dependent DSB had no significant effect on the percentage of cells with one genome equivalent following treatment with cephalexin and rifampicin. As a result, it was deduced that the presence of the DSB has no detectable effect on the duration of the B-phase of the cell cycle. Moreover, the formation of the DSB behind the progressing replication fork might suggest that the duration required for chromosomal replication is not affected by the presence of the DSB (data provided in chapter 5). Hence, the increase in generation time following formation of the replication-dependent DSB might solely be due to an increase in the duration of the D-phase of the cell cycle, which may ensure that the DSB is efficiently repaired before segregation of intact genomes into the daughter cells.

In *E. coli*, segregation of sister loci occurs progressively from the origin to the terminus during chromosomal replication (Nielson *et al*., 2006). The duration of cohesion for most newly replicated loci has been estimated to be approximately 7 to 10 minutes (Joshi *et al*., 2011; Reyes-Lamothe *et al*., 2008). It has been suggested that the cohesion of sister loci is essential for repair of spontaneous DSB that might arise during chromosomal replication (Sjoegren and Nasmyth, 2001; Cortes-Ledesma and Aguilera, 2006). During repair of the replication-dependent DSB, the homology search and strand exchange reactions are expected to be rapid because the undamaged sister template is near the chromosomal locus bearing the DSB. Hence the rate-determining step for the repair of the DSB might possibly be either the re-synthesis of the lost DNA sequences or the resolution of the joint molecules or both reactions. If these downstream reactions are also rapid, then the repair of the replication-dependent DSB at the *lacZ* locus should be completed before termination of chromosomal replication. However, the increase in generation time during DSB formation at the *lacZ* locus might indicate that, under slow growth condition, the presence of the DSB might have resulted in an increase in the D-phase of the cell cycle to enable efficient repair. The data provided in this chapter is not sufficient for identifying the specific DSB repair reaction which resulted in an increase in the generation time. The significant loss of YPet_Cerulean foci in the presence of the replication-dependent DSB was a major drawback for the use of the mean foci per cell length in identifying the effect of the DSB on the duration of cohesion of sister *lac* loci.

Chapter 4: Live-cell imaging of RecA and the site of a replication-dependent DSB at the *lacZ* locus of the *E. coli* chromosome

4.1 Introduction

The cellular dynamics of fluorescently labelled RecA during DSB repair have been studied in *Bacillus subtilis*, *Caulobacter crescentus* and *Escherichia coli* by live-cell imaging (Kidane and Graumann, 2005; Badrinarayanan *et al*., 2015b; Lesterlin *et al*., 2014). The formation of a repairable DSB in the above studies was achieved using either mitomycin C or an endonuclease (I-SceI or homothallic) system. Following DSB induction, the authors monitored the cellular dynamics of the fluorescently labelled RecA and concluded that RecA forms bundles or thread-like structures during repair of the induced DSB.

Mitomycin C generates DNA interstrand crosslinks which can be converted into DSBs during the excision repair process of the DNA adduct. After exposure of *B. subtilis* to a non-lethal dose of mitomycin C, Kidane and Graumann (2005) observed the formation of RecA threads which extended towards the opposite half of the cell. In addition, they observed that when a cell had two separated nucleoids, these nucleoids began to fuse in mostly one-half of the cell. Even though mitomycin C generates random DSBs across the chromosome, the formation of RecA threads and the fusion of separated nucleoid were inferred to represent the search for the undamaged sister template and strand invasion reactions during DSB repair. This inference was confirmed by the formation of RecA threads following induction of a site-specific DSB using the HO-endonuclease system (Kidane and Graumann, 2005).

More recently, the induction of a site-specific DSB in *E. coli* using the I-SceI endonuclease system also demonstrated the formation of RecA bundles during repair (Lesterlin *et al*., 2014). By selecting for cells with distant sister loci that have segregated into the opposite cell halves, and controlling the amount of I-SceI enzyme in the cell, the authors inferred that a single repairable DSB was generated in each cell they analysed. Following expression of I-SceI endonuclease, they observed that RecA initially formed a focus at the site of the induced DSB. Thereafter, the RecA focus developed into a bundle that extended towards the undamaged sister locus and mediated pairing of the sister loci. Eventually, the RecA bundle was disassembled prior to re-segregation of the sister loci into the opposite halves of the cell. These cellular events have also been observed in *C. crescentus* during the repair of I-SceI-induced DSBs (Badrinarayanan *et al*., 2015b).

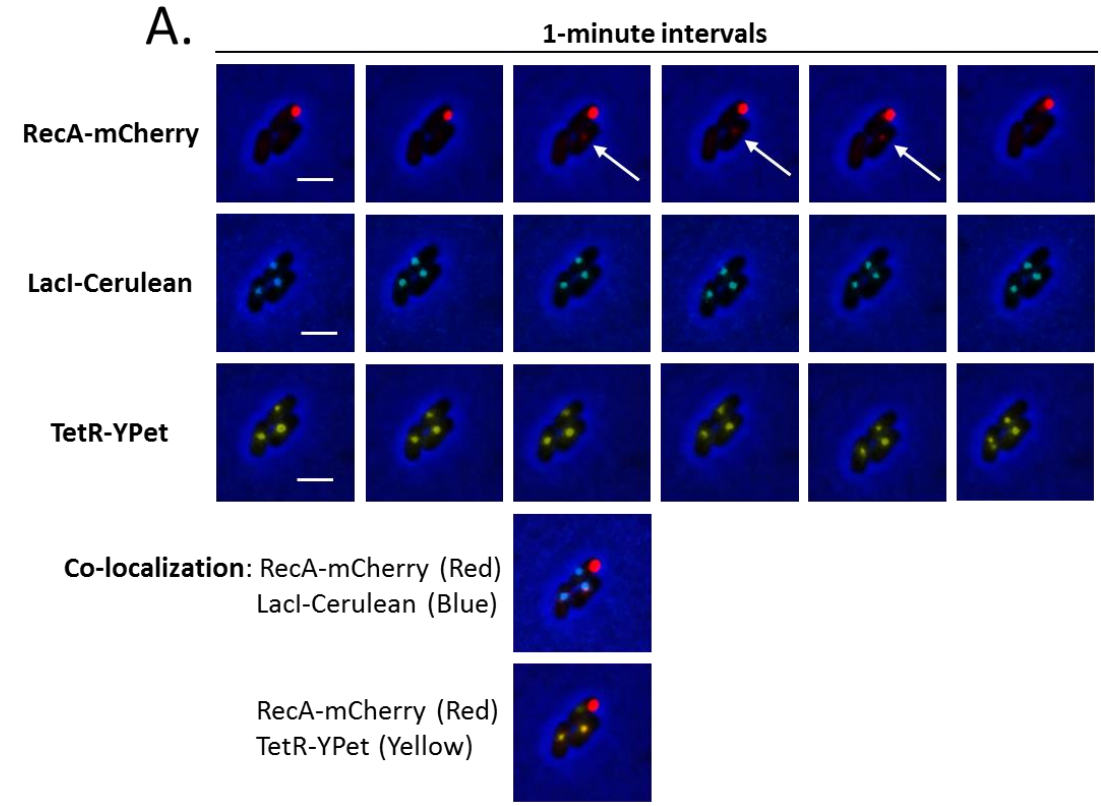
During normal cellular growth, the formation of spontaneous DSBs are usually dependent on chromosomal replication (Pennington and Rosenberg, 2007). A systematic live-cell imaging study of fluorescently labelled RecA during the repair of a site-specific DSB emanating from replication in bacteria has not yet been performed. This chapter describes the cellular dynamics of RecA during the repair of replication-dependent DSBs at the *lacZ* locus. DSBs arising from replication of the chromosome would not require an extensive homology search if repair is initiated during cohesion of the sister loci. Thus, this study utilized time-lapse fluorescence imaging to investigate whether DSBs emanating from replication are repaired during cohesion or after segregation of the sister loci. In addition, time-lapse and snapshot imaging was used to ascertain whether RecA bundle or thread is formed during the repair of a replication-dependent DSB. Furthermore, single deletion mutants for proteins that regulate the activity of RecA were used to identify the effect of the absence of these regulators on the stability of the RecA nucleoprotein filament that is formed during repair of a replication-dependent DSB.

4.2 Live-cell imaging of RecA-mCherry and the site of a replication-dependent DSB at the *lacZ* locus.

4.2.1 Transient co-localization of a RecA-mCherry focus with the site of the replication-dependent DSB.

The strain containing the interrupted palindrome at *lacZ* (DL5528) was grown overnight in M9-glycerol medium at 37˚C. The overnight culture was diluted and grown for three hours in the same liquid medium. The dilution was repeated, and 0.2% arabinose was added to the diluted culture, after which it was grown for an additional two hours. An aliquot of the cells growing in the liquid medium was added to an agarose pad prepared in the same growth medium and supplemented with arabinose. The addition of arabinose to the growth medium resulted in the expression of the SbcCD endonuclease from the arabinose-inducible promoter in the strain containing the interrupted palindrome at *lacZ* (DL5528). Moreover, the addition of arabinose to the agarose pad ensured the continual induction of the SbcCD endonuclease, and consequently, the formation of a replication-dependent DSB at every replication cycle during the duration of the time-lapse imaging. Images were acquired at 1-minute intervals using a fluorescence microscope enclosed in an incubation chamber at 37 ˚C.

Following DSB induction, it was observed that RecA-mCherry proteins assembled from background levels to form a distinct focus which co-localized with both the TetR-YPet and LacI-Cerulean foci (Figure 4.1A, white arrows). Subsequently, the RecA-mCherry focus that



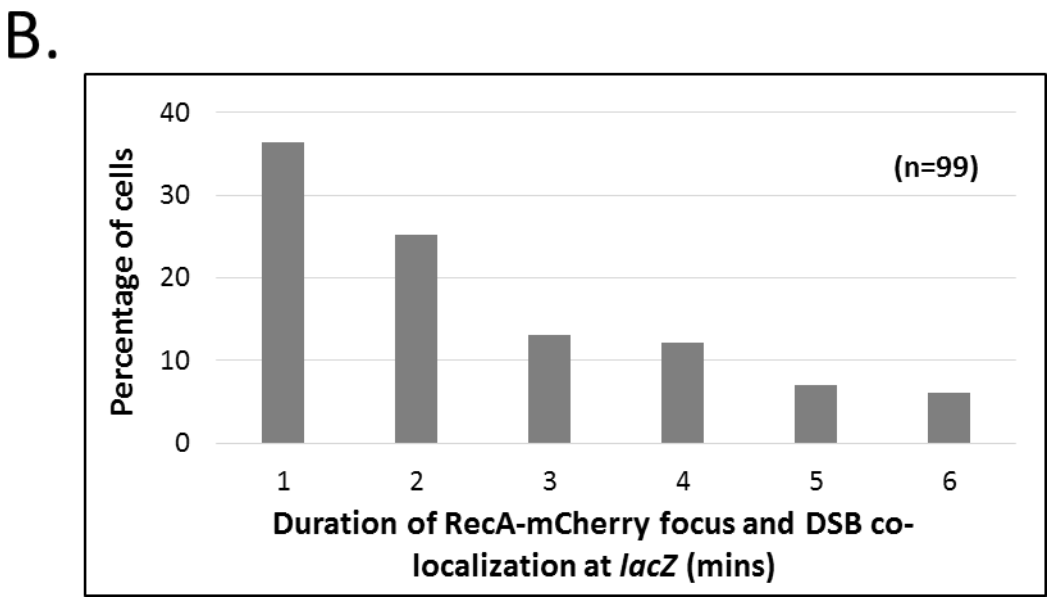


Figure 4.1. Formation of a transient RecA-mCherry focus at the site of the replication-dependent DSB. (A) Time-lapse fluorescence imaging of RecA-mCherry and the site of the replication-dependent DSB in a microcolony of three *E. coli* cells. Following DSB induction, RecA-mCherry assembled to form a distinct focus (white arrow) that co-localized with both the TetR-YPet (origin-proximal end) and LacI-Cerulean (origin-distal end) foci. Scale bar is 2µm. (B) Duration of the RecA-mCherry focus at the site of the replication-dependent DSB. Number of cells that were analysed (n) = 99.

co-localized with the YPet_Cerulean foci was disassembled to background fluorescence within the cell. The formation and disassembly of the RecA-mCherry focus at the site of the DSB was rapid and reversible. Analysis of 99 cells revealed that the range for the duration of the RecA-mCherry focus co-localizing with the site of the DSB was from 1 to 6 minutes (Figure 4.1B), with an average of (2.5 $\pm$ 1.6) minutes. Because the RecA-mCherry focus co-localized with both the TetR-YPet and LacI-Cerulean foci, it was inferred that the RecA-mCherry focus corresponded to the RecA nucleoprotein filament that was formed during the repair of the DSB at *lacZ*. RecA-mCherry bundles, filaments or thread-like structures were not detected during this time-lapse imaging. The absence of maturation of RecA-mCherry foci into bundles, filaments or threads is consistent with the expectation that repair of a DSB arising during replication does not require an extensive search for the undamaged homologous template. The short average duration of the RecA-mCherry focus also suggests that homology search and strand exchange are rapid events during the repair of a replication-dependent DSB.

4.2.2 Segregation of sister *lac* loci occur after repair of the replication-dependent DSB

Time-lapse imaging was performed for an extended duration (3 hours) in order to monitor the behaviour of the YPet_Cerulean foci after the disappearance of the transient RecA-mCherry focus that co-localized with the site of the DSB. Images were acquired at 2- or 3-minute intervals to minimise photo-bleaching and photo-toxicity during this extended period of image acquisition. The data revealed that the disassembly of the RecA-mCherry focus was followed by segregation of the sister *lac* loci into the opposite cell halves, and then cell division (Figure 4.2A). The average duration between disassembly of the RecA-mCherry focus and segregation at the *lac* loci was approximately (23 $\pm$ 12) minutes (Figure 4.2B). Surprisingly, this duration of cohesion of sister *lac* loci was 3 minutes or less in some (2.5%) of the cells, suggesting that the cell is capable of rapidly performing the downstream events of DSB repair following homology search and strand exchange. Moreover, the data showed that approximately 20% of the cells that were analysed required at least 30 minutes to segregate their sister *lac* loci after disassembly of the transient RecA-mCherry focus. Because the RecA-mCherry focus at the site of the DSB was formed before segregation of the sister *lac* loci, and re-pairing of *lac* loci was not detected after the disappearance of the RecA-mCherry focus and the concomitant segregation, it was deduced that repair of the replication-dependent DSB occurred during the period of cohesion of the *lac* loci. Moreover, if the normal duration of cohesion at the *lac* locus in the absence of DSB repair, as reported by Bates and collaborators (Joshi *et al*., 2011), is approximately 7 to 10 minutes, then it can be inferred that the formation and repair of the

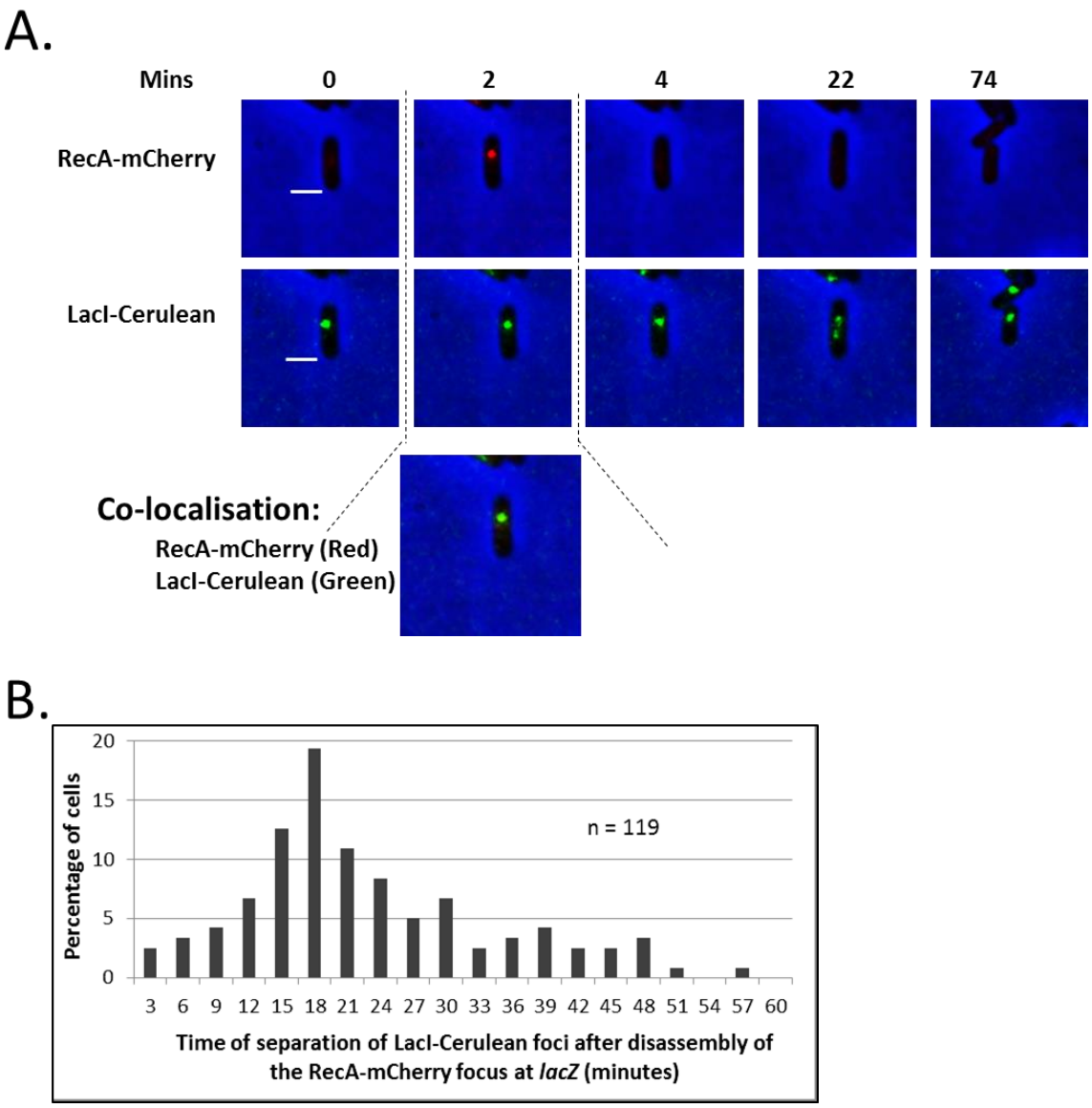


Figure 4.2. Repair of the replication-dependent DSB occurs prior to segregation at the *lac* locus. (A) Time-lapse fluorescence imaging of the RecA-mCherry focus and the LacI-Cerulean focus during DSB repair, segregation and cell division. Scale bar is 2µm. (B) Time of cohesion of sister *lac* loci during the repair of the replication-dependent DSB. The time of cohesion of sister *lac* loci was estimated by measuring the duration between the disassembly of the RecA-mCherry focus at the site of the DSB and the separation of the LacI-Cerulean foci (origin-distal end). The average duration was 23 minutes. Number of cells that were analysed (n) = 119.

replication-dependent DSB generally result in extension of the duration of cohesion at the *lac* locus.

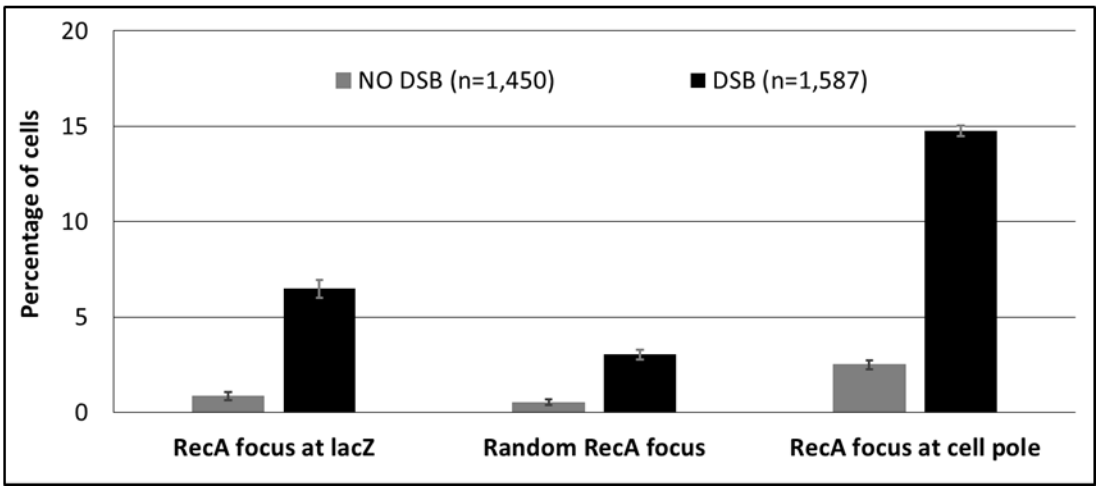
4.2.3 Quantification of the RecA-mCherry focus at the site of the replication-dependent DSB and other forms of RecA-mCherry foci in the cell

The proportion of cells with a RecA-mCherry focus at the site of the DSB was quantified by snapshot microscopy. Following expression of the SbcCD endonuclease in the palindrome-containing strain, the proportion of cells with a RecA-mCherry focus that co-localized with the YPet_Cerulean foci was 7.7-fold higher than the corresponding control strain without the palindrome at *lacZ* (Figure 4.3A). In the absence of DSB formation in the control strain, it was observed that approximately 0.8% of the cell population had a RecA-mCherry focus that co-localized with the YPet_Cerulean foci (Figure 4.3A). In this control strain, the RecA-mCherry focus might possibly represent the repair of DSBs arising from replication fork collapse at the site of the operator arrays bound by their respective repressor proteins.

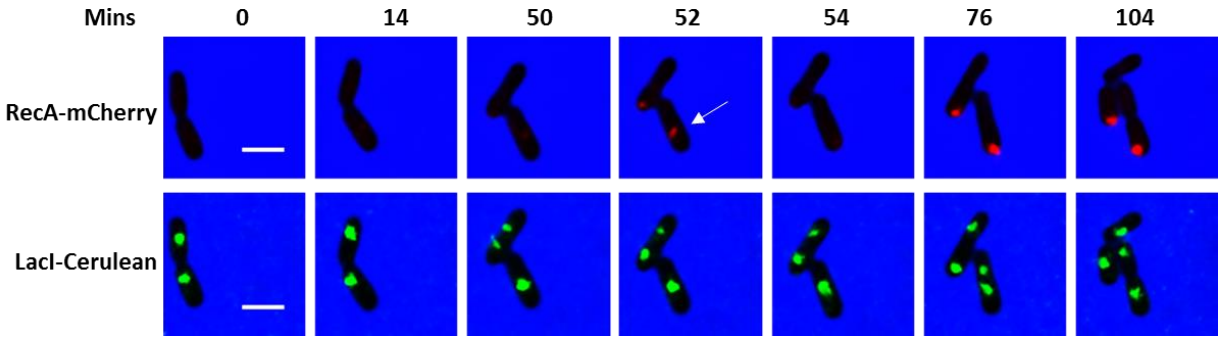
Time-lapse imaging and snapshot microscopy also revealed two other forms of RecA-mCherry foci. These were RecA-mCherry foci that were located at a cell pole (Figure 4.3B) and RecA-mCherry foci that did not co-localize with the YPet_Cerulean foci. The latter RecA-mCherry foci can be attributed to spontaneous DSBs that were generated during normal cellular growth. Surprisingly, the proportion of cells with a RecA-mCherry focus that did not co-localize with the YPet_Cerulean foci increased by 5.7-fold in response to DSB formation (Figure 4.3A). It has been reported that the induction of the replication-dependent DSB at *lacZ* results in the formation of DSBs in the terminus region of the *E. coli* chromosome (Cockram *et al*., 2015). The repair of these DSBs at the terminus might be responsible for the increase in the proportion of cells with a RecA-mCherry focus that did not co-localize with the YPet_Cerulean foci following DSB formation at the *lacZ* locus.

Moreover, the proportion of cells with a RecA-mCherry focus at a cell pole also increased by approximately 6-fold in the presence of palindrome-induced DSBs (Figure 4.3A). Formation of a RecA-mCherry focus at a cell pole occurred usually after the disassembly of the RecA-mCherry focus that co-localized with the site of the DSB (Figure 4.3B, the white arrow indicates a RecA-mCherry focus that co-localized with the Cerulean focus). However, the disassembly of a RecA-mCherry focus at the site of the DSB was not always followed by the formation of a RecA-mCherry focus at a cell pole (Figure 4.2A). Data from time-lapse imaging also revealed that RecA-mCherry foci at cell poles were persistent with time, and occasionally,

A.


20
15
10
5
0
Percentage of cells
NO DSB (n=1,450)
DSB (n=1,587)
RecA focus at lacZ
Random RecA focus
RecA focus at cell pole

B.


Mins
0
14
50
52
54
76
104
RecA-mCherry
LacI-Cerulean

C.

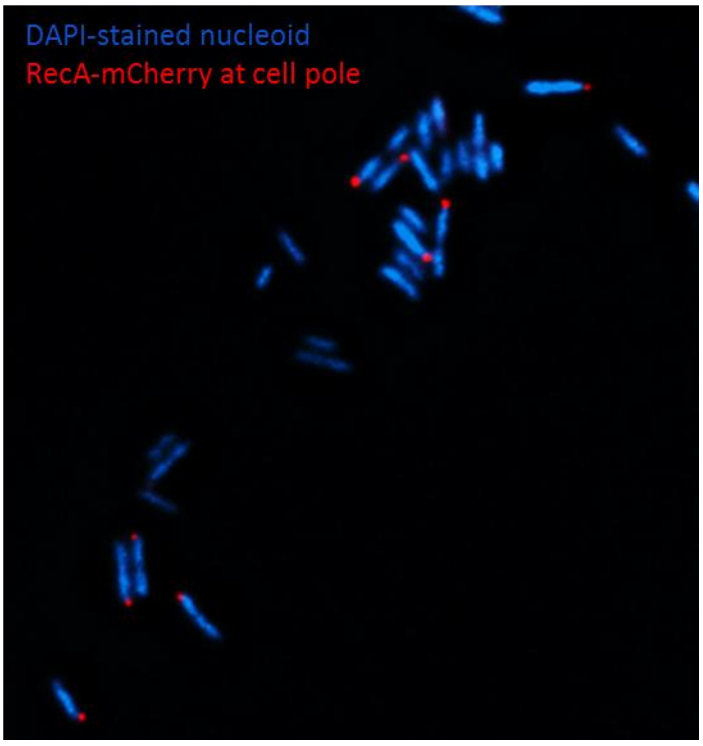

DAPI-stained nucleoid
RecA-mCherry at cell pole

Figure 4.3. The proportion of cells containing a RecA-mCherry focus increases in response to DSB formation. (A) Percentage of cells containing any of the three forms of RecA-mCherry foci. (B) Time-lapse images of cells with a persistent RecA-mCherry focus at a cell pole. The arrow indicates a cell with a RecA-mCherry focus that co-localized with the LacI-Cerulean focus. Scale bar is 2μm. (C) Localization of DAPI-stained nucleoids in relation to RecA-mCherry foci at cell poles.

these foci were still present after cell division (Figure 4.3B). In rare occasions, a RecA-mCherry focus that co-localized (or did not co-localize) with the site of the DSB was generated in a cell already showing a persistent RecA-mCherry focus at a cell pole. Simultaneous imaging of DAPI-stained nucleoid and RecA-mCherry revealed that RecA-mCherry foci at the cell pole were sequestered from the nucleoid (Figure 4.3C).

A RecA-mCherry focus at a cell pole might possibly correspond to aggregation of RecA-mCherry proteins (Renzette *et al*., 2005) after the occurrence of homology search and strand exchange during DSB repair. In general, the data obtained by snapshot microscopy indicate that the formation of replication-dependent DSBs result in an increase in the proportion of cells with any of the three forms of the RecA-mCherry focus, suggesting that the SOS response was induced following DSB formation. RecA bundles, filaments or thread-like structures were not detected during analysis of the images obtained from snapshot microscopy.

4.3 Live-cell imaging of RecA-mCherry and the site of the replication-dependent DSB in Δ*uvrD*, Δ*recX* and Δ*dinI* mutants

4.3.1 Viability and growth profiles of Δ*uvrD*, Δ*recX* and Δ*dinI* mutants following replication-dependent DSB induction

In vitro studies have demonstrated that the UvrD protein inhibits DNA strand exchange via removal of RecA from the RecA nucleoprotein filament during recombination (Veaute *et al*., 2005). As a result, UvrD is classified as an antirecombinase in *E. coli* (Cox, 2007). Purified RecX protein also inhibits DNA strand exchange *in vitro* (Stohl *et al*., 2003). However, RecX functions by binding to RecA at gaps within the nucleoprotein filament, thereby facilitating the net disassembly of RecA from the end of the gap (Drees *et al*., 2004). Unlike UvrD and RecX, the DinI protein stabilizes the RecA nucleoprotein filament (Lusseti *et al*., 2004) by interacting with the core domain of RecA (Yoshimasu *et al*., 2003). Thus UvrD, RecX and DinI are proteins which modulate the stability, and consequently, the activity of the RecA nucleoprotein filament (Cox, 2007).

In order to ascertain the role of these proteins during the repair of replication-dependent DSBs, the genes encoding these proteins were deleted using plasmid mediated gene replacement. The deletion of these genes was carried out in the strain containing the interrupted palindrome at *lacZ* (DL5528) and the corresponding control strain (DL5525). The single deletion mutants that were obtained were used for investigating the effect of the absence of these modulators on the duration of the RecA-mCherry focus at the site of the replication-dependent DSB. Before proceeding with time-lapse fluorescence imaging, the viability and growth profiles of these

single deletion mutants were determined following induction of the replication-dependent DSB.

In the absence of DSB induction, the spot test assay revealed that there was no detectable loss in viability of either the Δ*uvrD*, Δ*recX* or Δ*dinI* mutants (glucose plate and pal- strains on the arabinose plate; Figure 4.4). However, a small colony phenotype was observed during chronic induction of SbcCD expression in the palindrome-containing Δ*uvrD* mutant (arabinose plate; Figure 4.4). In addition, the Δ*uvrD* mutants with (DL5865) and without (DL5901) the interrupted palindrome showed no visible colony on the glucose plate that was irradiated with UV light, thereby confirming the requirement of the UvrD protein for the repair of UV-induced DNA damage (Goosen and Moolenaar, 2008). Like the wild type strains, the Δ*recX* and Δ*dinI* mutants showed no loss of viability on the UV-irradiated glucose plate indicating that *recX* and *dinI* are not essential for the repair of UV-induced DNA damage.

The growth profiles of Δ*uvrD*, Δ*recX* and Δ*dinI* mutants following induction of SbcCD expression are presented in Figure 4.5. All the single deletion mutants without the interrupted palindrome at *lacZ* showed a growth profile that was identical to the corresponding wild type strain (DL5525). The growth profiles also revealed that the induction of SbcCD expression in the palindrome-containing Δ*recX* or Δ*dinI* mutants (DL6156 and DL6446) did not affect their growth rate when compared to the strain without the deletion mutation (DL5528). However, the palindrome-containing Δ*uvrD* mutant had a slower growth rate compared to the corresponding wild type strain. The slow growth rate, and the small colony phenotype of the Δ*uvrD* mutant in the presence of DSB formation imply that the UvrD protein has an important role for efficient repair of the replication-dependent DSB.

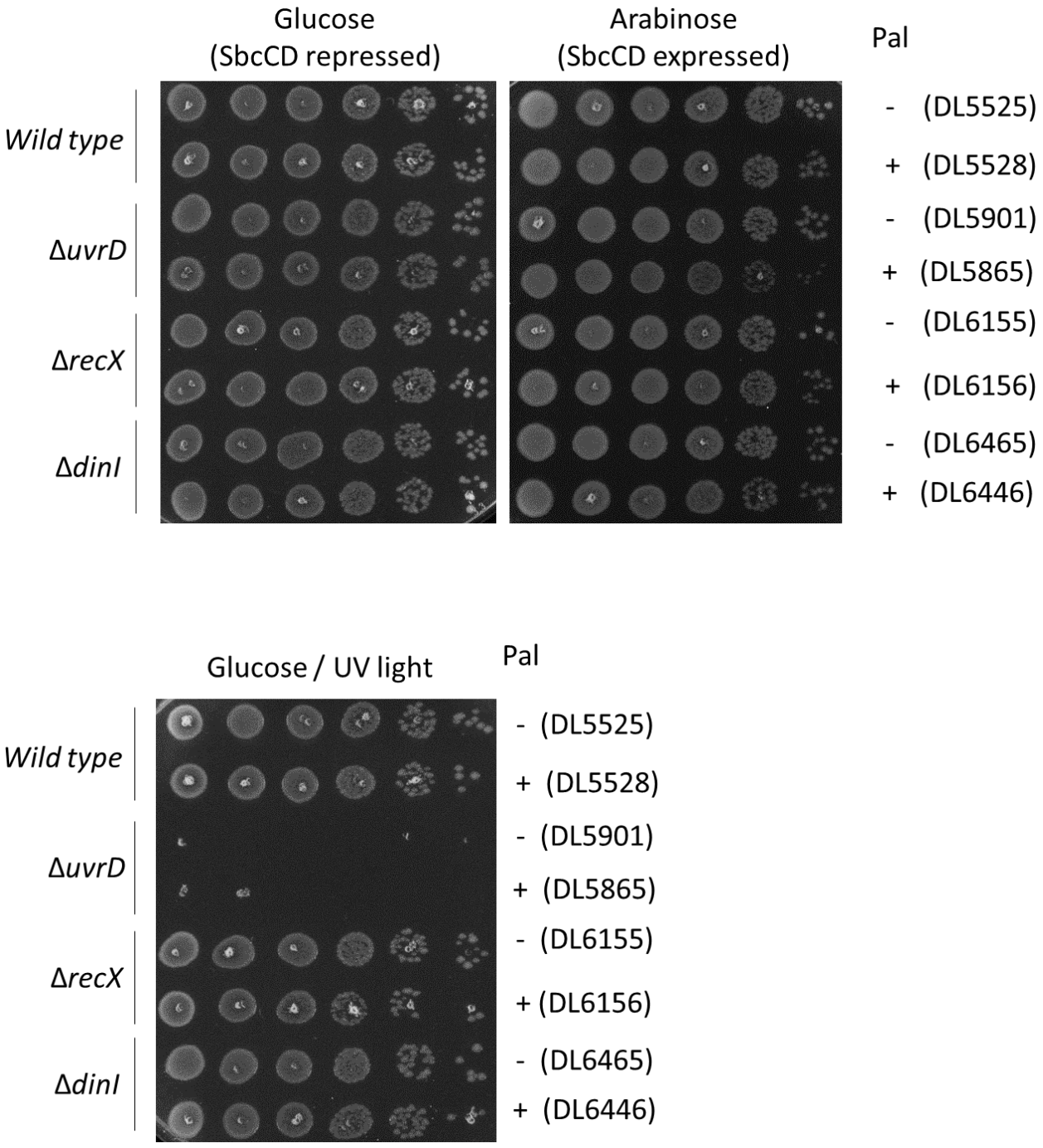


Figure 4.4. Viability test by spot test assay for Δ*uvrD*, Δ*recX* and Δ*dinI* mutants. Overnight cultures were grown in liquid LB, normalised to an OD_{600}=1 and diluted serially (10-fold) before been spotted on LB plates supplemented with either 0.5% glucose or 0.2% arabinose. One of the LB/Glucose plates was exposed to UV-light (0.01J). Thereafter, all the plates were incubated overnight at 37˚C. The deletion mutant strains that were used for this assay are all derivatives of either DL5528 or DL5525. Pal represents the presence (+) or absence (-) of the interrupted palindrome at *lacZ.*

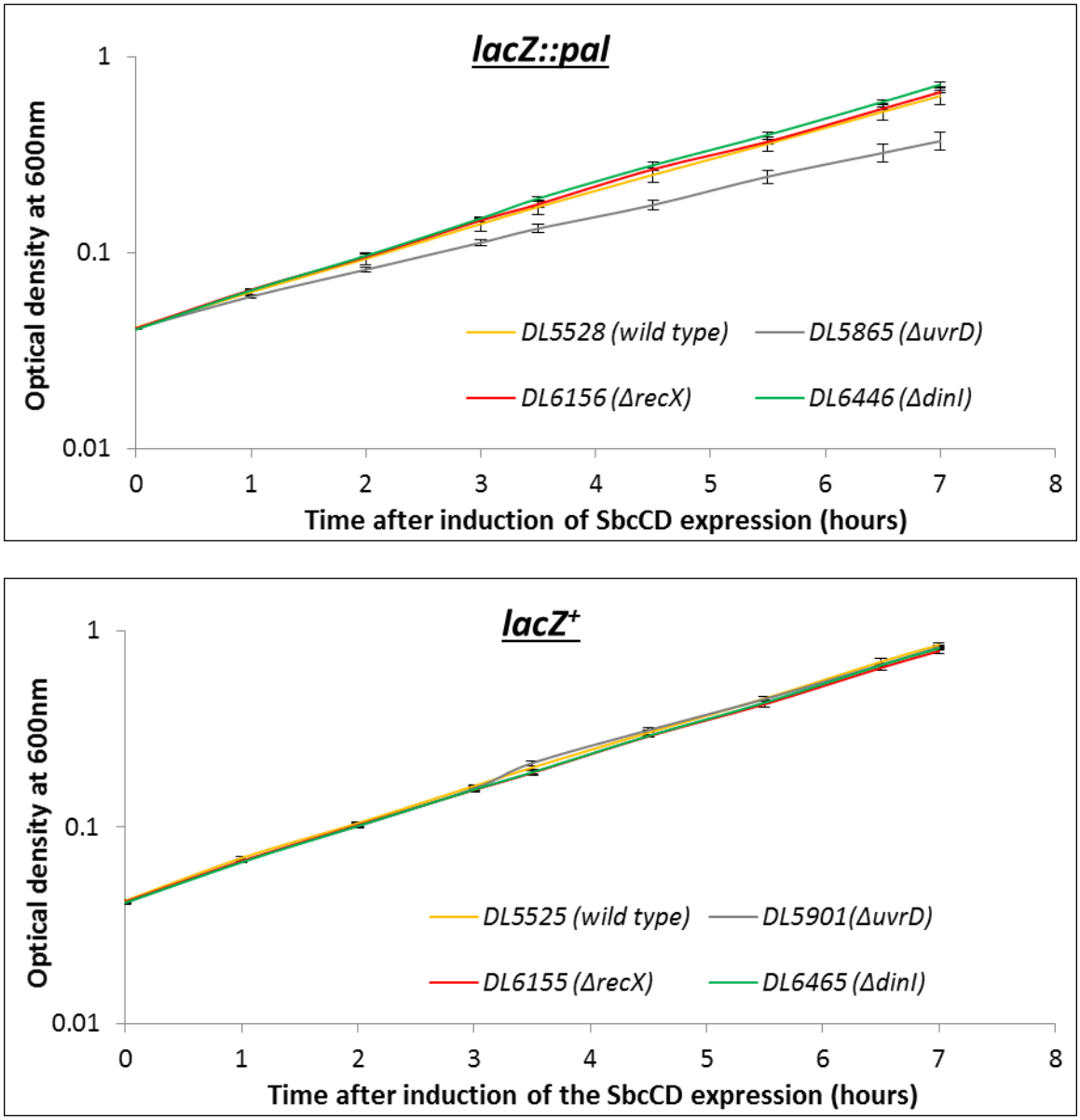


Figure 4.5. Growth profiles of Δ*uvrD*, Δ*recX* and Δ*dinI* mutants. Overnight cultures were grown at 37°C in M9 minimal medium supplemented with 0.2% glycerol. The overnight cultures were diluted and grown to exponential phase for 3 hours (OD_{600} < 0.3). The cultures at exponential phase of growth were diluted to an OD_{600}= 0.04 in the same growth medium, with the addition of 0.2% arabinose. The diluted cultures were grown and maintained in exponential phase by dilution while measuring the OD_{600} at the indicated time intervals for 7 hours. Error bars represent the standard error of the mean of three independent experiments.

4.3.2 Time-lapse imaging of Δ*uvrD*, Δ*recX* and Δ*dinI* mutants during replication-dependent DSB repair

The deletion mutants containing the interrupted palindrome at *lacZ* were used for time-lapse imaging following induction of SbcCD expression by the addition of arabinose. Each deletion mutant was grown to exponential phase in M9-glycerol minimal medium and re-diluted in the same medium supplemented with 0.2 % arabinose. The re-diluted cells were grown for 2 hours at 37˚C and an aliquot was placed onto the agarose pad for time-lapse imaging.

For each of the three deletion mutants and the wild type strain, the distributions of the duration of RecA-mCherry focus at the site of the DSB are presented in Figure 4.6A. Left-skewed distributions were obtained for all the three deletion mutants and the wild type strain. However, the modal duration of RecA-mCherry focus at the site of the DSB was 1 minute for the Δ*dinI* mutant, Δ*uvrD* mutant, and the wild type strain, whereas the Δ*recX* mutant had a modal duration of 2 minutes. Additionally, there were outliers for the duration of RecA-mCherry focus at the site of the DSB for the Δ*uvrD* and Δ*recX* mutants (duration > 6 minutes). From the distributions shown in Figure 4.6A, the mean durations were 2.16 minutes, 2.72 minutes, 2.99 minutes and 2.47 minutes for the Δ*dinI* mutant, Δ*uvrD* mutant, Δ*recX* mutant and the wild type strain respectively (Figure 4.6B). Statistical analysis (t-test) revealed that there was no significant difference between the mean duration of RecA-mCherry focus for the wild type strain and each of the three deletion mutants ($p > 0.05$).

Due to the presence of outliers in the distributions for the Δ*recX* and Δ*uvrD* mutants (Figure 4.6A), as well as the high standard deviations that were obtained for each of the four strains (Figure 4.6B), the median was used, instead of the mean, for comparing the average duration of RecA-mCherry foci at the *lacZ* locus for each of the three deletion mutants, and for the wild type strain. Accordingly, cumulative frequency distributions were generated for the wild type strain and each deletion mutant (Figure 4.6B). From the cumulative frequency distributions, the median duration of RecA-mCherry focus at the site of the DSB for Δ*dinI*, Δ*uvrD* and Δ*recX* mutants were 1.2 minutes, 1.6 minutes and 2.1 minutes respectively, while the median duration for the wild type strain was 1.5 minutes. The median durations indicate that deletion of the *uvrD* gene had the least effect on the duration of RecA-mCherry focus at the site of the DSB, even though it has been reported that the UvrD protein has a putative role in the disassembly of the RecA nucleoprotein filament (Veaute *et al*., 2005). The deletion of either *dinI* or *recX* generated opposite effects on the duration of RecA-mCherry focus at the

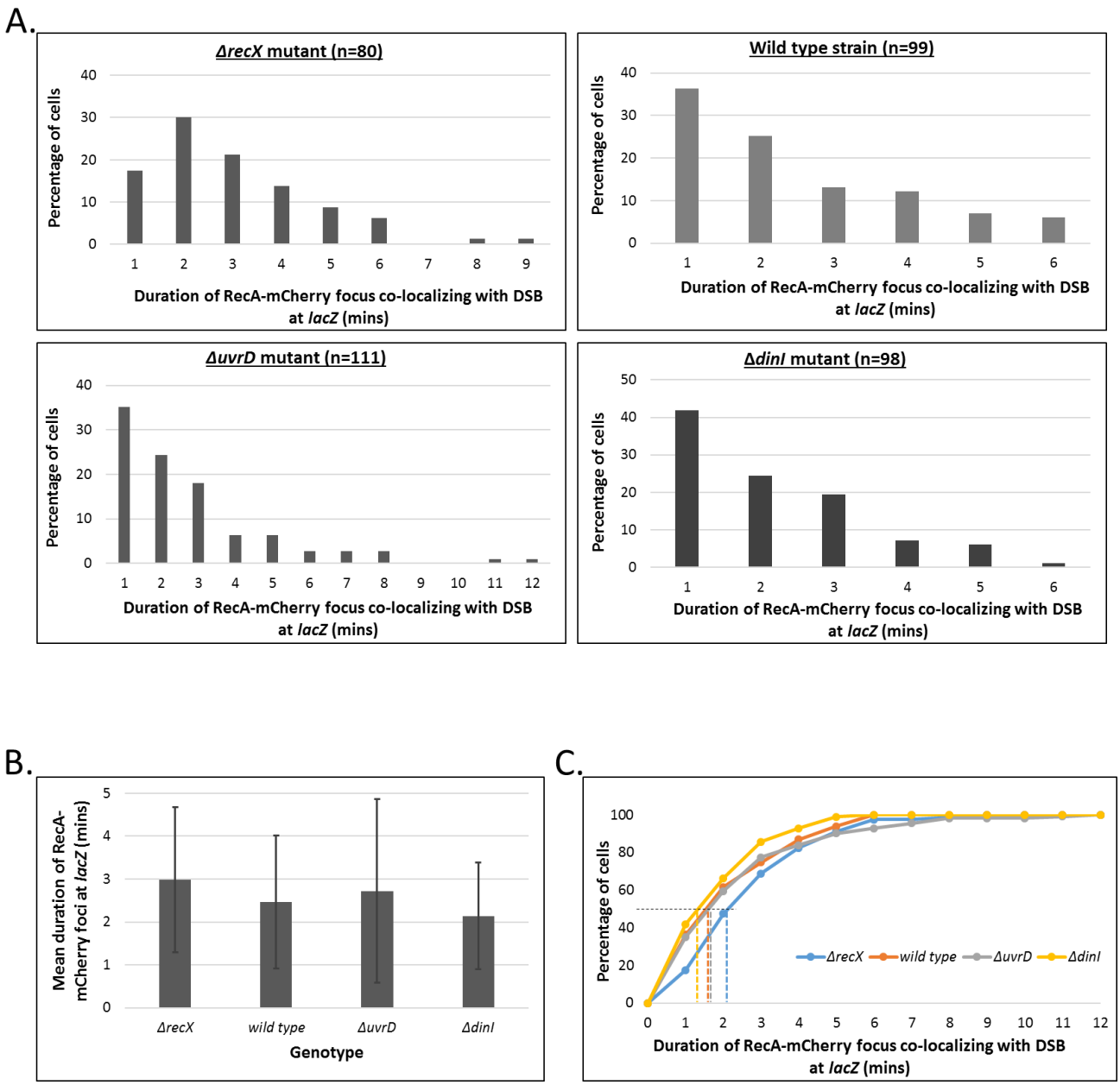


Figure 4.6. Effect of the absence of the RecX, UvrD or DinI protein on the duration of a RecA-mCherry focus at the site of the replication-dependent DSB. (A) Duration of RecA-mCherry focus at the site of the replication-dependent DSB for the single deletion mutants and the wild type strain. Time-lapse imaging was performed at 1-minute interval for each strain following induction of the replication-dependent DSB. (B) Mean duration of RecA-mCherry focus at the site of the replication-dependent DSB for the single deletion mutants and the wild type strain. Error bars represent the standard deviation. Statistical analysis (t-test) revealed that there was no significant difference between the mean duration of each deletion mutant and the wild type strain (p-value > 0.05). (C) Cumulative frequency distribution for the duration of RecA-mCherry focus at the site of the replication-dependent DSB for the single deletion mutants and DL5528. Dotted lines represent the extrapolation of the median values for DL5528 and each deletion mutant. The strains that were used for the time-lapse imaging were DL6156 (Δ*recX*), DL5528 (wild type), DL5865 (Δ*uvrD*) and DL6446 (Δ*dinI*).

site of the DSB. These observations correlate with the opposing effects of the RecX and DinI proteins on the RecA pre-synaptic filament (Lusetti *et al*., 2004). Thus, the data obtained for Δ*recX* and Δ*dinI* mutants are in accordance with RecX protein having a destabilising effect on the RecA nucleoprotein filament, while the presence of the DinI protein results in an increase in the stability of the nucleoprotein filament. Even though the median duration of the RecA-mCherry focus in the Δ*uvrD* mutant was higher than that obtained for the wild type strain, the modest difference between these median durations suggests that, generally, the presence of the UvrD protein has a minimal effect on the stability of the RecA nucleoprotein filament that is formed during the repair of a replication-dependent DSB.

4.4 DISCUSSION

Replication forks originating from *oriC* may encounter an unrepaired DNA lesion during normal cellular growth (Cox *et al*., 2000). These DNA lesions include DNA nicks or gaps, and are potential sources of spontaneous DSBs during chromosomal replication (Kuzminov, 1999; Kuzminov, 2001; Vilenchik and Knudson, 2003). The collision of a progressing replication fork with a protein-DNA complex may also generate DSBs in the cell (Michel *et al*., 2001). Furthermore, spontaneous DSBs can be formed at the site of long palindromic sequences on the lagging strand template in the presence of the structure-specific SbcCD endonuclease (Leach, 1994). These occurrences demonstrate that chromosomal replication is a source of spontaneous DSBs (Habber, 1999; Pennington and Rosenberg, 2007).

The data presented in this chapter showed that RecA-mCherry proteins formed a distinct and transient focus at the site of a replication-dependent DSB during repair. The transient RecA-mCherry focus at the site of the DSB was disassembled prior to segregation of the sister *lac* loci. After segregation, re-pairing of the sister *lac* loci was not observed in any of the cells that were studied by time-lapse fluorescence imaging. Cell division occurred after segregation, with each daughter cell acquiring a copy of the *lac* locus. These observations indicate that the repair of a site-specific DSB emanating from replication is initiated and completed during cohesion of the newly replicated loci. Moreover, the RecA-mCherry focus that co-localized with the site of the replication-dependent DSB did not mature into a filament or thread or a bundle-like structure, consistent with the expectation that the repair of a site-specific DSB arising from chromosomal replication does not require an extensive homology search. By using the I-SceI endonuclease system to generate a site-specific DSB, Sherratt and colleagues reported that a RecA-GFP focus co-localized transiently with the site of the DSB prior to maturation to form a bundle which mediated pairing of distant sister loci (Lesterlin *et al*., 2014). Hence, a possible interpretation is that a RecA focus at the site of a DSB matures to form a bundle only when an extensive homology search is required at pre-synapsis.

Kidane and Graumann (2005) also reported that GFP-RecA formed threads or filament-like structures during the repair of DSBs that were generated using mitomycin C in *B. subtilis*. Even though DSBs that are generated by mitomycin C are dependent on replication, the GFP-RecA threads or filaments that were observed by the authors might represent the repair of several clustered DSBs on the chromosome. Even though Sherratt and colleagues used anti-RecA antibodies to demonstrate that RecA bundles that were formed during the repair of I-SceI-induced DSBs are not artefacts, Kidane and Graumann did not provide such data in their report.

If the GFP-RecA threads are not artefacts of the fluorescent labelling system, it can be inferred that the cellular dynamics of RecA during repair of several clustered DSBs might be different from its behaviour during repair of a site-specific DSB, even though both forms of DBSs are replication-dependent.

The data presented in this chapter reveal that homology search is a rapid event during the repair of a site-specific DSB arising from replication. Moreover, the data also show that the downstream events (after occurrence of homology search and strand exchange) might possibly represent the rate-determining step in the repair of a site-specific DSB that is generated during chromosomal replication. Sherratt and colleagues reported that RecA bundles were wholly assembled approximately 15 minutes after the formation of a RecA-GFP focus at the site of an I-SceI-induced DSB. Thereafter, pairing of the distant sister loci occurred 40 minutes after assembly of the RecA bundle. Thus, the homology search and strand exchange reactions occurred approximately 55 minutes after the formation of a RecA focus at the site of an I-SceI-induced DSB. Segregation of the sister loci into the opposite cell halves occurred 30 minutes after homology search was completed. These observations suggest that homology search is the rate-determining step during the repair of a DSB that require an extensive search for the undamaged sister template. After the occurrence of homology search during the repair of an I-SceI-induced DSB, the duration that was required for the segregation of sister loci into the opposite cell halves was fairly similar to that obtained in this study. It can therefore be postulated that cohesion of sister loci might have evolved to ensure that repair of a DSB emanating from chromosomal replication occur rapidly by eliminating the need for an extensive homology search.

The duration of cohesion of several loci across the *E. coli* chromosome have been investigated in the absence of DSB induction (Reyes-Lamothe *et al*., 2008; Joshi *et al*., 2011; Lesterlin *et al*., 2012). Amongst these reports, only Bates and colleagues studied the duration of cohesion of sister *lac* loci prior to segregation (Joshi *et al*., 2011). According to the authors, sister *lac* loci are cohered for approximately 7 to 10 minutes before segregation into the opposite cell halves (Joshi *et al*., 2011). In this current study, a direct estimation of the duration of cohesion of sister *lac* loci, in the absence of DSB induction, was not provided. However, the data presented here indicate that approximately 90% of the cells that were analysed required at least 12 minutes for segregation of sister *lac* loci after disassembly of the RecA-mCherry focus at the *lacZ* locus. Additionally, the data showed that the average duration for segregation of sister *lac* loci, after homology search and possibly strand exchange, was 23 minutes. Consequently,

it can be postulated that the duration of cohesion at the *lac* loci is extended during repair of a site-specific DSB arising from chromosomal replication.

Time-lapse imaging of mutants with deletions of either the *uvrD*, *dinI* or *recX* genes were used to investigate the effect of these gene products (modulators) on the stability of the RecA nucleoprotein filament that is formed during the repair of a replication-dependent DSB. The Δ*uvrD* mutant showed a small colony phenotype and slower growth rate in the presence of DSB formation, indicating that the UvrD protein has an important role in efficient repair of a replication-dependent DSB. The 3' to 5' translocase activity of the UvrD protein has been shown to dismantle the RecA nucleoprotein filament, *in vitro* (Veaute *et al*., 2005). As a result, it was expected that the RecA-mCherry focus at the site of the DSB would have a longer duration in the Δ*uvrD* mutant. However, the median duration of a RecA-mCherry focus at the site of the DSB was not greatly increased by the Δ*uvrD* mutation. Nonetheless, the data did reveal a small percentage of cells (less than 8 %) that had a longer duration (> 6 minutes) of the RecA-mCherry focus at the site of the DSB. These observations might suggest that the absence of the UvrD protein could have an effect on the stability of long-lasting RecA nucleoprotein filaments formed during the repair of a replication-dependent DSB. The data also suggest that the absence of UvrD had little or no impact on the majority of RecA foci implying that another helicase might be involved in the disassembly of the RecA nucleoprotein filament in the absence of UvrD. Importantly, the small colony phenotype and the slower growth rate in the presence of DSB formation demonstrate that this putative alternative helicase cannot completely substitute for UvrD during the repair of a replication-dependent DSB.

It has been postulated that, if a monomer of RecA dissociates from within the nucleoprotein filament, RecX binds to the terminal RecA at the gap that is generated within the nucleoprotein filament (Drees *et al*., 2004). The presence of RecX at the gap prevents the re-filling of the gap by RecA, and facilitates the disassembly of RecA from the end of the gap that was generated within the nucleoprotein filament. Time-lapse imaging revealed that the Δ*recX* mutant showed a 40% higher median duration for the RecA-mCherry focus at the site of the DSB when compared to the *recX*$^{+}$ strain. This observation suggests that the absence of the RecX protein resulted in an increase in the stability of the RecA nucleoprotein filament that was formed during the repair of the replication-dependent DSB. The increase in stability of the RecA nucleoprotein filament in a Δ*recX* mutant did not result in a significant growth defect during the repair of the replication-dependent DSB.

Previous studies on the regulation of the stability of the RecA nucleoprotein filament by DinI revealed that the stability of the nucleoprotein filament is dependent on the concentration of DinI (Lusetti *et al*., 2004). Specifically, the authors used data from a RecA-mediated ATPase assay and electron microscopy to infer that at very high concentrations of DinI compared to RecA, the RecA nucleoprotein filament is destabilized, and vice versa (Lusetti *et al*., 2004). The amount of RecA in a cell is approximately 30-fold higher compared to DinI following induction of the SOS response during DSB repair (Sommer *et al*., 1998; Yasuda *et al*., 1998). Thus, the presence of the DinI protein in a cell is not expected to decrease the stability of the RecA nucleoprotein filament *in vivo*. The data obtained from time-lapse imaging illustrate that deletion of *dinI* resulted in a 20% decrease in the median duration of the RecA-mCherry focus at the site of the DSB when compared to the *dinI*$^{+}$ strain. Hence, it can be inferred that the absence of DinI caused a decrease in the stability of the RecA nucleoprotein filament that was formed during the repair of the replication-dependent DSB. Moreover, the decrease in stability of the RecA nucleoprotein filament in the absence of DinI did not generate a detectable growth defect during repair of the replication-dependent DSB.

The distributions for the duration of RecA-mCherry foci at the site of the DSB revealed the presence of outliers for the Δ*uvrD* mutant and Δ*recX* mutant but not the Δ*dinI* mutant and the wild type strain. The proportion of cells that showed these outliers might correspond to a subpopulation that had a substantially stable RecA nucleoprotein filament in the absence of either UvrD or RecX. Therefore, it can be deduced that the accessory helicase, if present, does not always disassemble the RecA nucleoprotein filament efficiently, in the absence of either UvrD or RecX.

In conclusion, the data obtained from this study demonstrated that the repair of a site-specific DSB arising from replication is rapid because the repair does not require an extensive homology search and occurs during cohesion of the sister loci. Additionally, the repair of a replication-dependent DSB generally results in an increase in the duration of cohesion of the sister loci. Furthermore, this study demonstrated that the stability of the RecA nucleoprotein filament is not extensively altered by the absence of UvrD. The absence of RecX or DinI proteins stabilised or destabilised the RecA filament as expected.

Chapter 5: Cellular localization of the *lac* locus during the formation and repair of a replication-dependent DSB

5.1 Introduction

Unlike eukaryotes, the genetic material of a bacterial cell is not enclosed within a nuclear membrane. However, the bacterial chromosome is not amorphous because it is subjected to some degree of compression due to supercoiling (Trun and Marko, 1998), macromolecular crowding (de Vries, 2010), and interactions with nucleoid-associated proteins (Dillon and Dorman, 2010). It has also been shown that the *E. coli* chromosome is organised into four macro-domains and two non-structured regions (Valens *et al*., 2004). According to the authors, chromosomal loci within the same macro-domain have a higher tendency to interact with each other compared to loci in different macro-domains.

Fluorescence *in situ* hybridization (FISH) has been used to reveal the cellular localisation of 23 different loci that are separated by approximately 230 kb on the *E. coli* chromosome (Niki *et al*., 2000). Analysis of the data obtained from the FISH studies indicated that *oriC* is located at the mid-cell at the onset of chromosomal replication in *E. coli* cells growing slowly in minimal medium. The *lac* locus, which is located within the non-structured region of the right chromosomal arm, was not included amongst the loci that were studied by FISH analysis. However, the regions flanking the *lac* locus were observed to be located preferentially between the mid-cell and ¼ positions of the *E. coli* cell.

Live-cell fluorescence imaging has revealed that the *cynR* locus, which is immediately upstream of the *lac* locus, localizes between the ¼ position and the mid-cell in *E. coli* (Lesterlin *et al*., 2014). Following the induction of a repairable DSB using the I-SceI endonuclease system, the authors demonstrated that the natural localization of the *cynR* locus was disrupted due to the extensive search for sequence homologies on the undamaged sister chromosome (Lesterlin *et al*., 2014). Additionally, they observed that the *cynR* locus bearing the DSB traversed approximately 84 % of the original distance separating the two sister loci during re-pairing of the distant loci. In *C. cresentus*, time-lapse fluorescence imaging was used to demonstrate that a chromosomal locus bearing an I-SceI-induced DSB at a cell pole traversed the entire cell length to pair with the undamaged sister locus at the opposite cell pole during repair of the DSB (Badrinarayanan *et al*., 2015b). These observations suggest that increase in mobility and re-positioning of chromosomal loci are cellular events which occur during the repair of a site-specific DSB that requires extensive homology search in bacteria.

In *S. cerevisiae*, the formation of an I-SceI-induced DSB at the *URA3* locus of chromosome V resulted in an increase in motility of that locus, which facilitated the exploration of a greater nuclear volume during homology search (Mine-Hattab and Rothstein, 2012). The authors also observed that the presence of the DSB led to an increase in motility of the uncut homologous *URA3* locus. Furthermore, the repair of multiple DSBs that were generated by ɣ-irradiation have been shown to occur within specialised repair centres in *S. cerevisiae* (Lisby *et al*., 2001; Lisby *et al*., 2003). These observations have led to the hypothesis that the increase in motility of both the cleaved and uncleaved chromosomal loci during DSB repair may be essential for scanning the entire nucleus in order to locate specialised repair centres (Lemaitre and Soutoglou, 2014). In mammalian cells, the two ends of an I-SceI-induced DSB are positionally stable and exhibit increased mobility only upon knockdown of the Ku80 protein; a protein involved in DSB repair via the non-homologous end joining pathway (Soutoglou *et al*., 2007). This implies that the increase in mobility of chromosomal loci in response to the formation of a DSB is a fundamental feature that underlies repair via the homologous recombination pathway.

This chapter discusses the effect of the formation of replication-dependent DSBs on the spatial localization of the *lac* locus in *E. coli*. This study also investigated whether there was a preferred subcellular position for the formation and the repair of replication-dependent DSBs. The spatial localizations of these cellular events were used to identify whether formation and repair of the DSBs occurred at the same or different subcellular positions within the cell.

5.2 Cellular position of the *lac* locus and the RecA focus at the site of the replication-dependent DSBs

5.2.1 Effect of the formation of replication-dependent DSBs on the spatial localization of the *lac* locus

Following 2 hours of SbcCD expression from the arabinose-inducible promoter, an aliquot of cells at exponential phase of growth in liquid M9-glycerol medium was added to an agarose pad. Afterwards, phase contrast and fluorescence images were acquired by snapshot microscopy. The cellular localization of the LacI-Cerulean focus/foci in the strains containing or not containing the interrupted palindrome at the *lacZ* locus are presented in the demographs shown in Figure 5.1. The demographs were obtained by sorting the cells according their lengths; cells with the smallest length were displayed at the top. The demographs indicate that the cells with smaller lengths had a single LacI-Cerulean focus while the longer cells had two foci. For cells with a single LacI-Cerulean focus, it was observed that the focus was

preferentially closer to the mid-cell than the cell pole. The cells with two LacI-Cerulean foci had a focus localized in each of the cell halves.

In order to identify whether there was an effect of DSB formation on the cellular localization of the *lac* locus, the distance between the centroid of a LacI-Cerulean focus and the closer cell pole was determined for the cells containing one LacI-Cerulean focus. The closer cell pole was chosen during this analysis because the two cell poles of *E. coli* are indistinguishable. Consequently, the data that were obtained for a particular cell half were also shown as a mimic for the other cell half. For cells with two LacI-Cerulean foci, the cell pole which was closer to both foci was used for the analysis. The data obtained revealed that for cells with a single LacI-Cerulean focus, the focus localized predominantly from the ¼ position to the mid-cell, with a preference for the mid-cell (Figure 5.2A). Additionally, the localization of each LacI-Cerulean focus in cells with two foci was predominant between the ¼ position and the mid-cell. The LacI-Cerulean focus was rarely localized at the cell pole. Importantly, Figure 5.2A indicates that the formation of replication-dependent DSBs did not greatly affect the cellular localization of the *lac* locus for the cells with a single LacI-Cerulean focus or those with two foci.

5.2.2 Cellular localization of the RecA focus at the site of the replication-dependent DSB

Analysis of the images obtained by snapshot microscopy revealed that 6.5% of the cell population showed a RecA focus at the site of the DSB following expression of the SbcCD endonuclease in the palindrome-containing strain (Figure 4.3A of Chapter 4). The cellular position of the RecA focus at the site of the DSB is presented in Figure 5.2B. The data revealed that the RecA focus was preferably located at the mid-cell or between the ¼ position and the mid-cell. However, the RecA focus was also localized at the ¼ position or between the cell pole and the ¼ position in 15% of the cells that were analysed.

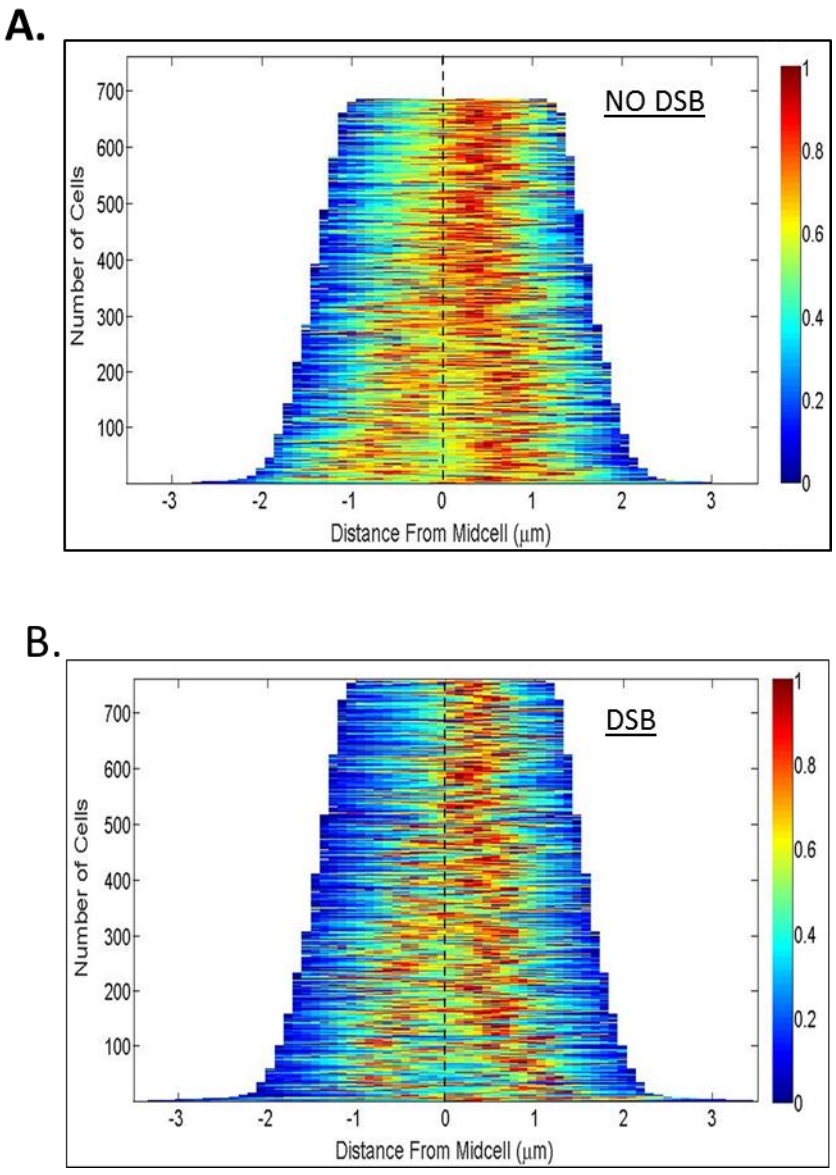


Figure 5.1: Cellular localization of the *lac* locus in the absence and presence of replication-dependent DSBs. Phase contrast and fluorescence (LacI-Cerulean) images were acquired by snapshot microscopy after 2 hours of SbcCD expression in the strains not containing (A) or containing (B) the interrupted palindrome at the *lacZ* locus. The demographs were obtained by sorting the cells that were obtained by snapshot microscopy according their lengths; cells with the smallest length were displayed at the top. The demographs were generated using the OUFTI software (Paintdakhi *et al*., 2016). The black dash lines on the demographs represent the mid-cell. The strains which were used for this analysis were DL5525 (*lacZ*$^{+}$) and DL5528 (*lacZ::pal*).

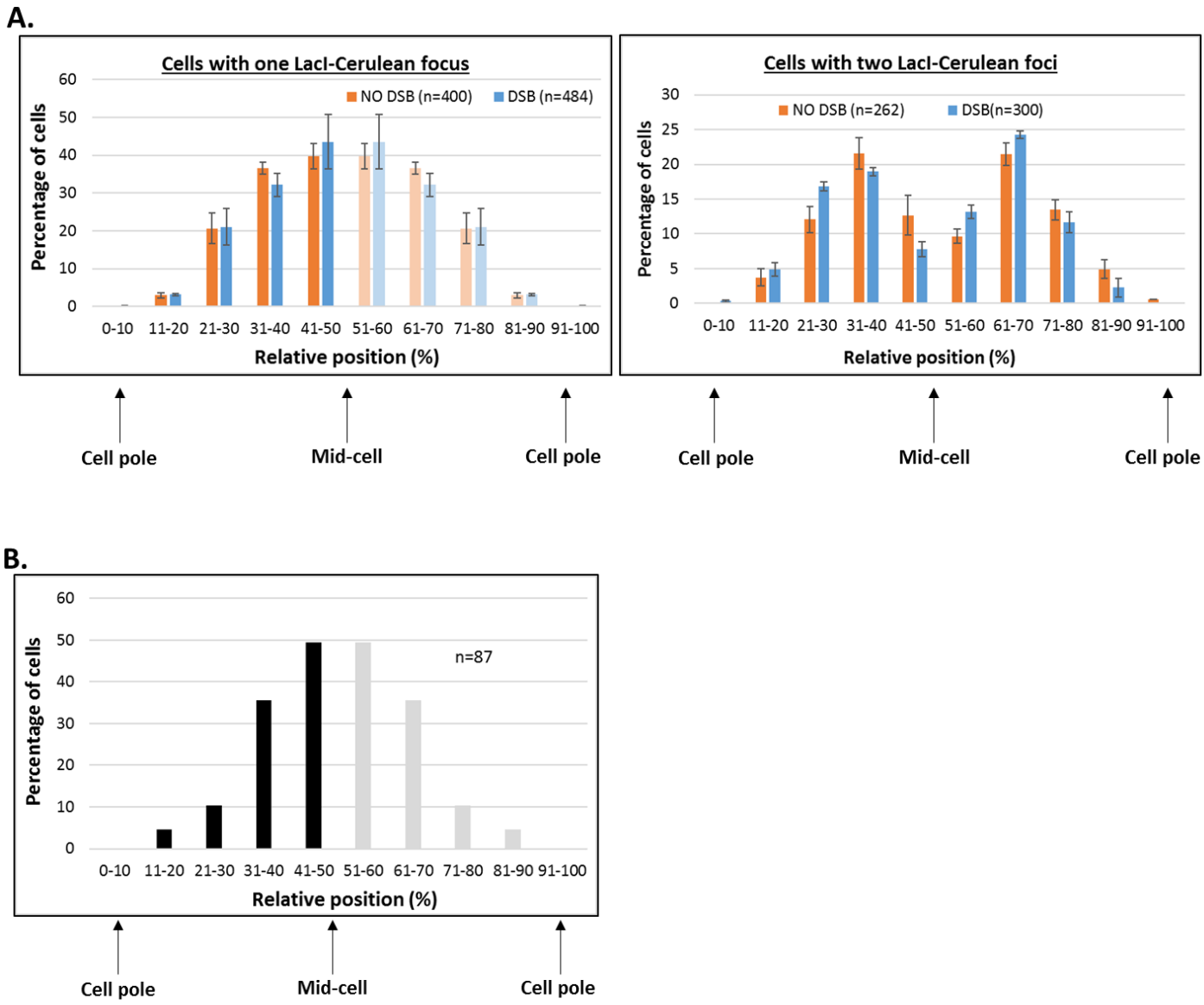


Figure 5.2: Subcellular positions of the *lac* locus in cells at exponential phase of growth and RecA focus at the site of the replication-dependent DSB. (A) Subcellular position of the *lac* locus in cells with either one or two LacI-cerulean foci in the absence or presence of DSBs. (B) Subcellular position of the RecA focus at the site of the replication-dependent DSBs. All images used for these analyses were acquired by snapshot microscopy. The phase contrast and fluorescence images were acquired following 2 hours of SbcCD expression in the strains not containing (NO DSB) or containing (DSB) the interrupted palindrome at the *lacZ* locus. The OUFTI software (Paintdakhi *et al.*, 2016) was used for measuring the distance between the centroid of a focus and the closest cell pole. For cells containing one LacI-Cerulean focus, the light-coloured histograms were used to represent the mimic of the deep-coloured histograms since the two poles of the *E. coli* cell are indistinguishable. The strains which were used for this analysis were DL5525 (*lacZ*$^{+}$) and DL5528 (*lacZ::pal*).

5.2.3 Spatial dynamics of the *lac* locus in the absence and presence of a replication-dependent DSB

The effect of DSB formation on the spatial localization of the *lac* locus was also investigated using images obtained from time-lapse microscopy since most of the cells analysed from the snapshot images might not be undergoing DSB repair during image acquisition. Following 2 hours of SbcCD expression in strains containing or not containing the interrupted palindrome at *lacZ*, phase contrast and fluorescence images were acquired at 45-second intervals for 42 minutes during time-lapse imaging. In the absence of DSB induction, the kymographs revealed that cells with a single LacI-Cerulean focus exhibited dynamic movement of the *lac* locus around the mid-cell until segregation of the sister loci (Figure 5.3A). Following DSB induction, the dynamic movement of the *lac* locus around the mid-cell was constrained after the formation and disassembly of the RecA focus at the site of the DSB (Figure 5.3B). These observations suggest that the heteroduplex DNA, or some other feature of the repair complex, that is formed after homology search and strand invasion restrains the dynamic movement of the *lac* locus during DSB repair. Additionally, the kymographs revealed that the restrained movement of the *lac* locus occurred very close to the mid-cell during DSB repair. From these kymographs, it can be inferred that, even though pre-synapsis and synapsis could occur between the ¼ position and the mid-cell, post-synapsis usually occurs very close to the mid-cell.

5.3 Fluorescence labelling of the β-sliding clamp of the *E. coli* replisome

5.3.1 Construction of *E. coli* strains that express a N-terminal YPet-DnaN fusion protein

This study also investigated whether there was a preferred cellular position for the replication of the *lac* locus, which precedes the formation and repair of the DSBs in the palindrome-containing strain. If the DNA hairpin that is formed at the site of the interrupted palindrome on the lagging strand template is cleaved immediately after replication of that locus, then the spatial localization for replication of the *lac* locus could also represent the spatial localization for DSB formation in the palindrome-containing strain. In order to ascertain the spatial localization for replication and formation of the DSB, the endogenous *dnaN* gene was tagged with the *YPet* gene. Thus, *E. coli* strains encoding the YPet-DnaN, LacI-Cerulean and RecA-mCherry proteins were used to investigate the spatial and temporal dynamics of the replisome and the *lac* locus following replication and formation of the DSB until the onset of segregation.

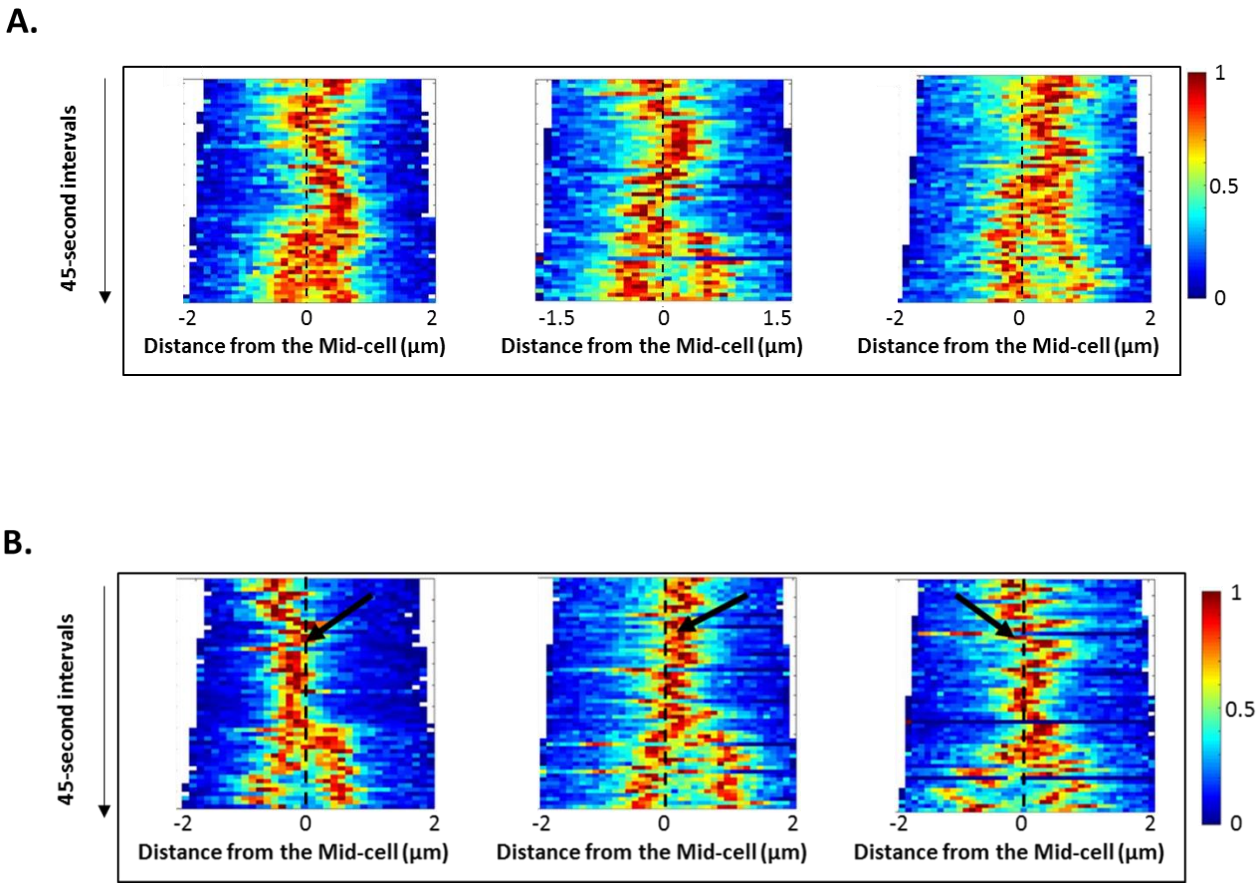


Figure 5.3: Spatial dynamics of the *lac* locus in the absence and presence of a replication-dependent DSB. Phase contrast and fluorescence (LacI-Cerulean) images were acquired by time-lapse microscopy after 2 hours of SbcCD expression in the strains not containing (A) or containing (B) the interrupted palindrome at the *lacZ* locus. The time-lapse images were acquired at 45-second intervals for 42 minutes. The black dash lines on the kymographs represent the mid-cell. The black arrows represent the formation of a RecA focus at the *lacZ* locus for cells undergoing DSB repair. Kymographs were generated using the OUFTI software (Paintdakhi *et al*., 2016). The strains which were used for this analysis were DL5525 (*lacZ*$^{+}$) and DL5528 (*lacZ::pal*).

In *E. coli*, the *dnaN* gene encodes the β-sliding clamp, which is responsible for the high processivity of DNA polymerase III during chromosomal replication (Kurth and O'Donnell, 2013). Prior to constructing the strains that express the YPet-DnaN fusion protein, the *tetR-YPet* and *lacI-Cerulean* genes were deleted from the *ykgC* chromosomal locus by plasmid mediated gene replacement using the plasmid, pDL2802. Afterwards, the *lacI-Cerulean* gene was re-inserted into the *ykgC* locus using the plasmid, pDL4580. The plasmid, pDL6255, was constructed purposely to generate *YPet-dnaN* gene at the native locus of the *dnaN* gene. During the construction of pDL6255, a ribosomal binding site (TAAGGAGGA) was inserted upstream of the start codon of the *YPet* gene, while ensuring that there was an optimal spacer (5 bp) between these DNA sequences (Reyes-Lamothe *et al*., 2010). In addition, a 33 bp linker sequence was inserted between the penultimate codon of the *YPet* gene and the start codon of the *dnaN* gene (Reyes-Lamothe *et al*., 2010). The *YPet* gene, together with the ribosomal binding site and the 33 bp linker sequence, was flanked by approximately 400 bp of homology sequences which allowed the integration of the construct immediately upstream of the start codon of the endogenous *dnaN* gene.

5.3.2 Effect of the presence of the *YPet-dnaN* gene on cell viability and growth rate following induction of replication-dependent DSBs

The plasmid, pDL6255, was used for generating a *YPet-dnaN* gene at the native *dnaN* locus of the strain with the 246 bp interrupted palindrome inserted at *lacZ* (*lacZ::pal*) and the corresponding control strain (*lacZ*$^{+}$). Prior to fluorescence imaging, the effect of the presence of the *YPet-dnaN* gene on cell viability and growth rate of these strains was investigated in the absence and presence DSB induction.

The spot test assay and growth profiles revealed that the presence of the *YPet-dnaN* gene at the native *dnaN* chromosomal locus had no detectable effect on viability and growth rate in the absence of DSB induction (Figures 5.4A and 5.4B). This observation confirmed that the presence of the N-terminal YPet-DnaN fusion protein did not perturb normal chromosomal replication (Reyes-Lamothe *et al*., 2010). However, the palindrome-containing strain with the *YPet-dnaN* gene (DL6306) showed a small colony phenotype following chronic DSB formation (Figure 5.4A). The small colony phenotype on the LB agar plate suggests that there was a loss of viability in a subpopulation of the palindrome-containing strain encoding the

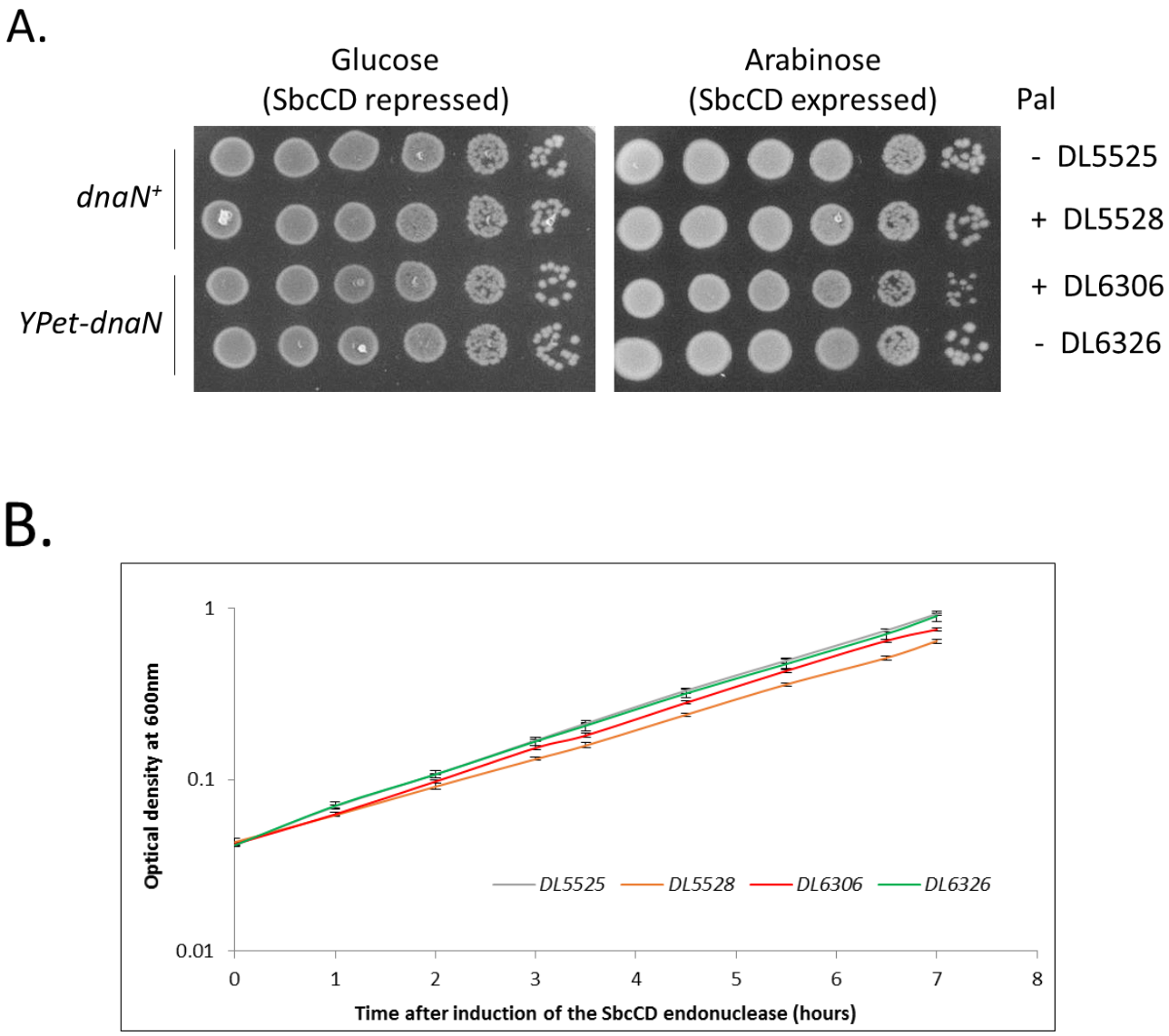


Figure 5.4. Effect of the presence of the *YPet-dnaN* gene on cell viability and growth rate following DSB induction. (A) Viability test by spot test assay for the strains with either the endogenous *dnaN* or the *YPet-dnaN* gene in the absence or presence of DSBs. Overnight cultures grown in liquid LB medium were normalised to an OD_{600}=1 and diluted serially (10-fold) before spotted on LB plates supplemented with either 0.5% glucose or 0.2% arabinose. All the plates were incubated overnight at 37˚C. (B) Growth profiles of the strains with either the endogenous *dnaN* or *YPet-dnaN* gene in the absence or presence of DSBs. Overnight cultures were grown at 37°C in M9-glycerol medium. The overnight cultures were diluted to an OD_{600}=0.09 and grown to exponential phase for 3hrs (OD_{600} < 0.3). The cultures in exponential phase of growth were diluted to an OD_{600}= 0.04 in the same growth medium, with the addition of 0.2% arabinose. The diluted culture was grown and maintained in exponential phase while measuring the OD_{600} at the indicated time intervals for 7 hours. Error bars represent the standard error of the mean of three independent experiments. The strains containing the endogenous *dnaN* gene were DL5525 (*lacZ*$^+$) and DL5528 (*lacZ::pal*) while the strains containing the *YPet-dnaN* gene were DL6326 (*lacZ*$^+$) and DL6306 (*lacZ::pal*).

YPet-dnaN gene following chronic DSB formation. Surprisingly, the palindrome-containing strain which encodes the YPet-DnaN protein showed a faster growth rate when compared to the corresponding strain containing the endogenous *dnaN* gene (DL6306 and DL5528; Figure 5.4B). Most single cells of either the *lacZ*$^{+}$ (DL6326) or *lacZ::pal* (DL6306) strains did not undergo filamentation during time-lapse imaging following 2 hours of SbcCD expression in M9-glycerol medium (Figure 5.5). As a result, the palindrome-containing strain encoding the YPet-DnaN protein was used for time-lapse fluorescence imaging despite the small colony phenotype that was detected by the spot test assay following chronic DSB induction on the LB agar plates.

5.3.3 Time-lapse imaging of the YPet-DnaN protein during chromosomal replication

It has been reported that YPet-DnaN molecules assemble to form a focus at the onset of chromosomal replication in *E. coli* (Wang *et al*., 2011). Reyes *et al*. (2008) also showed that foci that were formed by fluorescently-labelled components of the replisome were disassembled prior to segregation of the *ter* loci. Hence, the presence of YPet-DnaN foci in cells at exponential phase of growth has been used as a fluorescent marker for chromosomal replication in *E. coli* (Wang *et al*., 2011).

In this study, time-lapse imaging confirmed that YPet-DnaN proteins were assembled to form a fluorescent focus, which was later disassembled to the background fluorescence of the protein within the cell (Figure 5.6A). In the absence of DSBs, the duration of the YPet-DnaN focus per replication cycle, in cells at exponential phase of growth, ranged from 52 to 92 minutes (Figure 5.6B). The distribution of the duration of the YPet-DnaN focus was not greatly affected by the formation of DSBs in the palindrome-containing strain (Figure 5.6B). Statistical analysis demonstrated that the presence of DSBs had no significant effect on the average duration of the YPet-DnaN focus per replication cycle in cells at exponential phase of growth (Figure 5.6C). From these observations, it was inferred that the formation of the replication-dependent DSB at the *lacZ* locus had no significant effect on the duration of chromosomal replication for cells at exponential phase of growth in M9-glycerol medium.

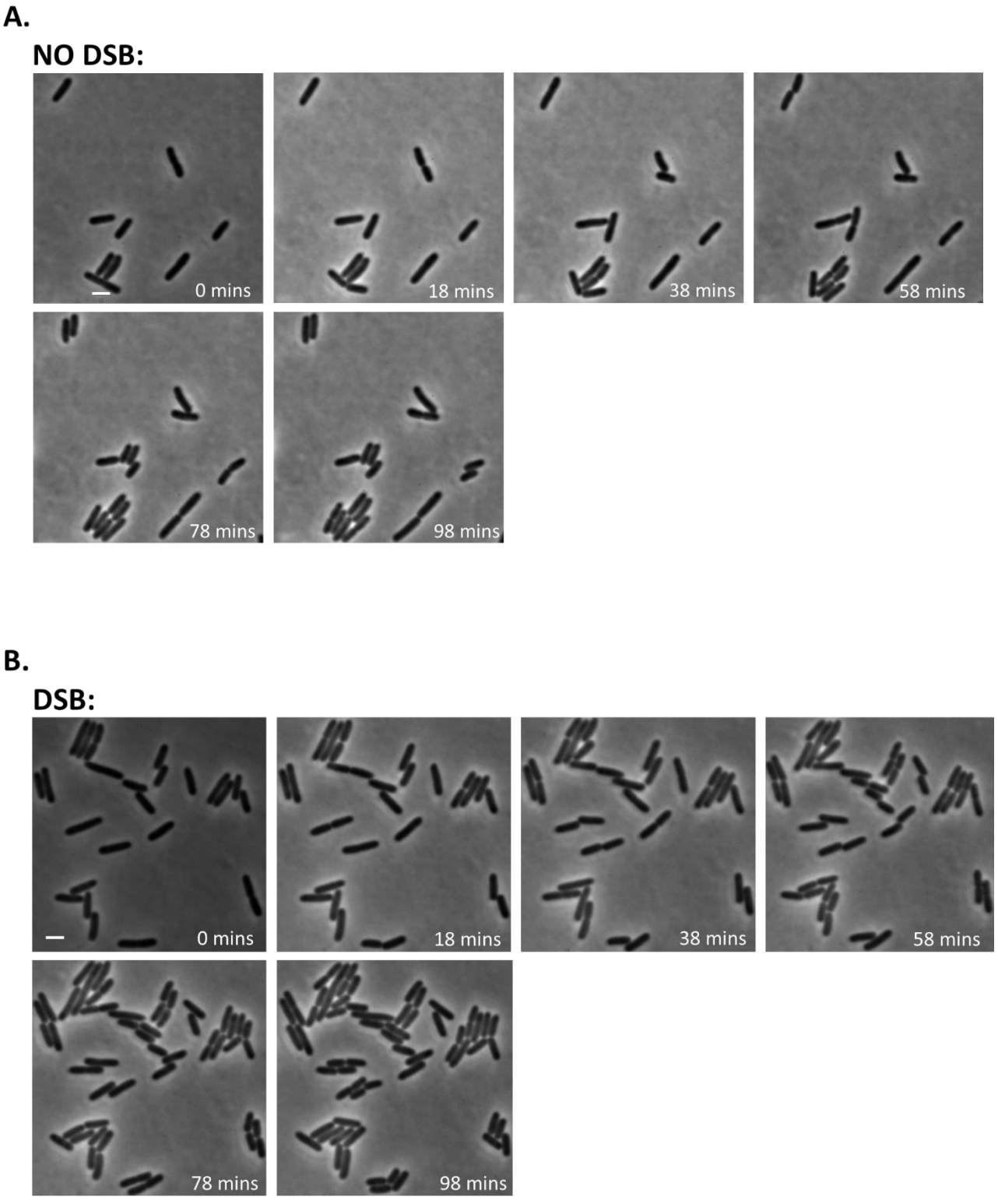


Figure 5.5. Time-lapse imaging of strains encoding the YPet-DnaN protein in the absence and presence of DSB induction. Phase-contrast images were acquired at 2-minute intervals following 2 hours of SbcCD expression in the *lacZ*$^{+}$ strain (A) and the *lacZ::pal* strain (B) which were growing exponentially in M9-glycerol medium. Scale bar is 2µm. The strains which were used for the time-lapse imaging were DL6326 (*lacZ*$^{+}$) and DL6306 (*lacZ::pal*).

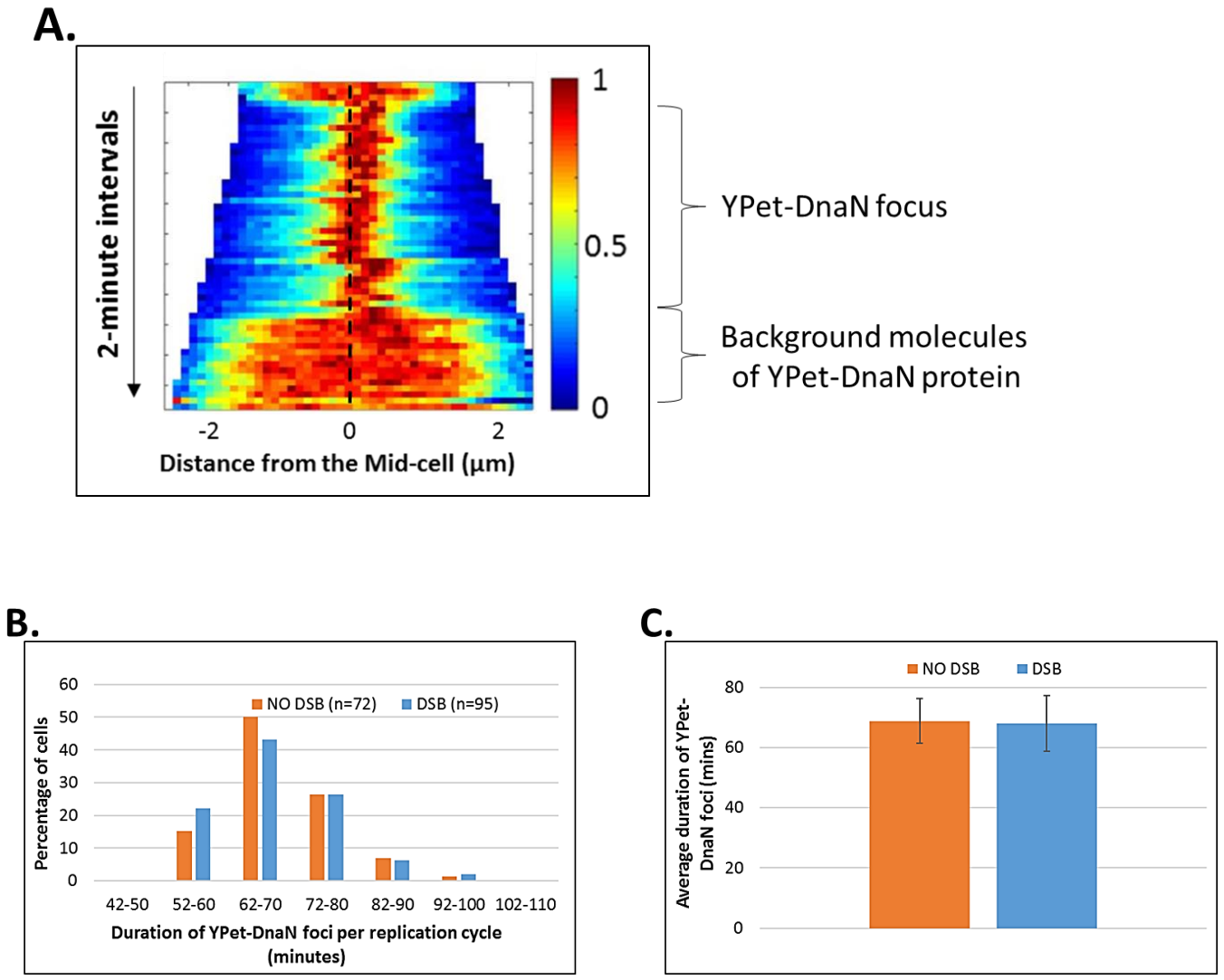


Figure 5.6. Formation and duration of YPet-DnaN foci in cells at exponential phase of growth in M9-glycerol medium. (A) Formation and disassembly of YPet-DnaN focus in a cell at exponential phase of growth. Phase contrast and fluorescence (YPet-DnaN) images were acquired by time-lapse microscopy at 2-minute intervals for 120 minutes. The black dash line on the kymograph represent the mid-cell. The kymograph was generated using the OUFTI software (Paintdakhi *et al*., 2016). (B) Distribution of the duration of YPet-DnaN foci per replication cycle in the absence and presence of DSBs. The SbcCD endonuclease was expressed for 2 hours in the *lacZ*$^{+}$ and *lacZ::pal* strains prior to time-lapse imaging. (C) Average duration of YPet-DnaN foci per replication cycle in the absence and presence of DSBs. Statistical analysis (t-test) revealed that there was no significant difference between the average duration of the YPet-DnaN foci in the absence and presence of DSBs (p-value > 0.05). The strains which were used for these analyses were DL6326 (*lacZ*$^{+}$) and DL6306 (*lacZ::pal*).

Further analyses were performed to identify the spatial dynamics of YPet-DnaN foci during chromosomal replication. Time-lapse images were acquired at 1-minute intervals for 55 minutes and were used for generating the kymographs presented in Figure 5.7. From these kymographs, it was observed that cells had either one or two YPet-DnaN foci during chromosomal replication. For cells that exhibited a single predominant YPet-DnaN focus during the replication cycle (28 cells out of 41 cells), the focus was localized at or near the mid-cell (Figure 5.7A). In these cells, the sister replisomes occasionally underwent transient separation to generate two YPet-DnaN foci which were very close to each other at the mid-cell. The two YPet-DnaN foci that were generated by spatial separation of the sister replisomes were not short-lived in a subpopulation of the cells that were analysed (13 cells out of 41 cells; Figure 5.7B). Thus, in that subpopulation, the single YPet-DnaN focus that was formed at the onset of replication was not usually predominant during the replication cycle. From these observations, it can be inferred that the sister replisomes which duplicate each arm of the circular *E. coli* chromosome are not entirely independent but are capable of co-existing preferably at the mid-cell.

5.4 Spatial and temporal dynamics of the replisome and the *lac* locus during DSB formation and repair

Simultaneous time-lapse imaging of the LacI-Cerulean, YPet-DnaN and RecA-mCherry proteins was performed following 2 hours of SbcCD expression in the palindrome-containing strain. The co-localization of a YPet-DnaN focus with the LacI-Cerulean focus might indicate replication of the *lac* locus, which was postulated to correspond to DSB formation in the palindrome-containing strain following expression of the SbcCD endonuclease. Thus, the data which were obtained from the time-lapse imaging were used to investigate the spatial localization of the *lac* locus during DSB formation (co-localization of a YPet-DnaN focus with the LacI-Cerulean focus). The data were also used to identify whether repair was initiated immediately after DSB formation by estimating the duration from DSB formation (co-localization of a YPet-DnaN focus with the LacI-Cerulean focus) to the formation of a RecA focus at the site of the DSB.

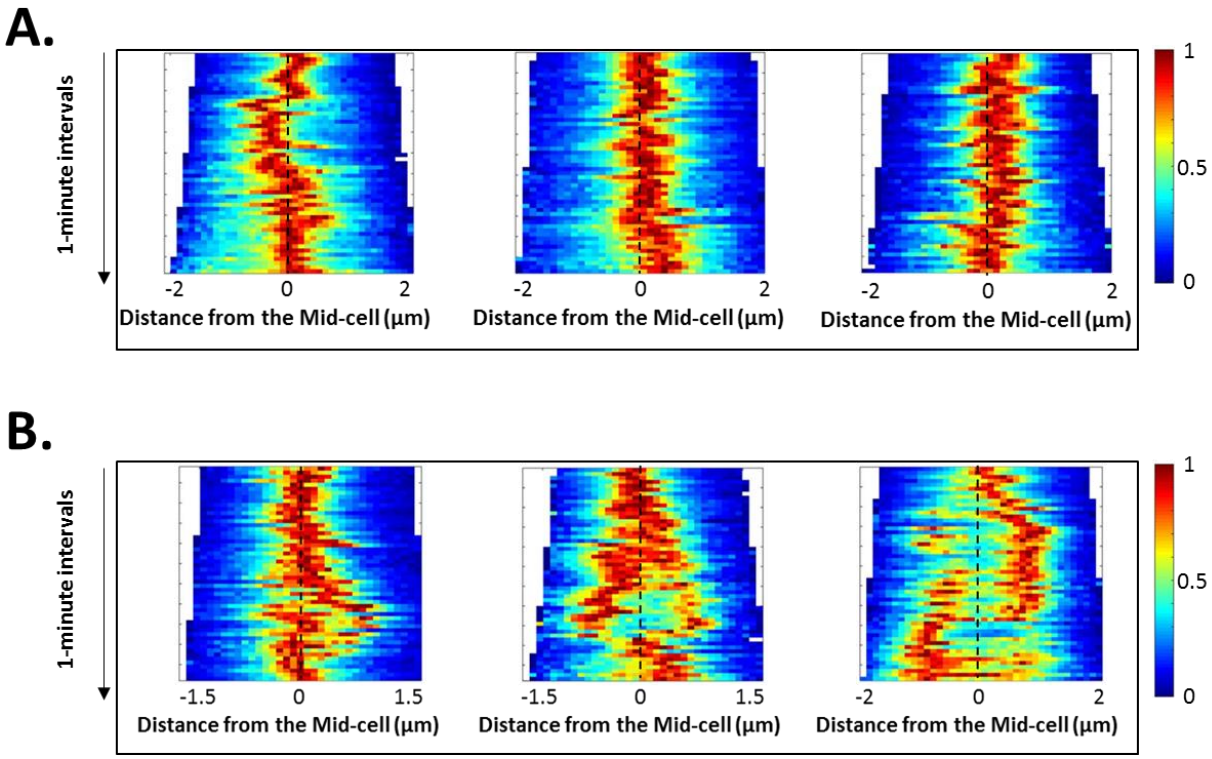


Figure 5.7: Spatial dynamics of the replisome in cells at exponential phase of growth in M9-glycerol medium. (A) Cells showing a single predominant YPet-DnaN focus during chromosomal replication. (B) Cells showing long-lived two YPet-DnaN foci during chromosomal replication. Phase contrast and fluorescence (YPet-DnaN) images were acquired by time-lapse microscopy after 2 hours of SbcCD expression in the strains (*lacZ*$^+$ and *lacZ::pal*) containing the *YPet-dnaN* gene. The time-lapse images were acquired at 1-minute intervals for 55 minutes. The black dash lines on the kymographs represent the mid-cell. Kymographs were generated using the OUFTI software (Paintdakhi *et al*., 2016). The strains that were used for this analysis were DL6306 (*lacZ::pal*) and DL6326 (*lacZ*$^+$).

Co-localization of a YPet-DnaN focus with the LacI-Cerulean focus occurred either at the mid-cell or between the mid-cell and the ¼ position (black circles on the kymographs presented in Figure 5.8). Out of the 20 cells that were analysed, 15 cells showed only one co-localization of a YPet-DnaN focus with the LacI-Cerulean focus prior to formation of the RecA focus at the site of the DSB. The remaining 5 cells showed either two or three co-localizations of a YPet-DnaN focus with the LacI-Cerulean focus prior to formation of the RecA focus at the site of the DSB. The duration of co-localization of a YPet-DnaN focus with the LacI-Cerulean focus ranged from 5 to 10 minutes. For the 15 cells that showed only one co-localization event (a YPet-DnaN focus with the LacI-Cerulean focus), 12 of these cells showed a RecA focus at the site of the DSB in less than 3 minutes following the co-localization event. Thus, for these 12 cells, the LacI-Cerulean focus, YPet-DnaN focus and the RecA focus at the site of the DSB were all co-localized in the cell when the latter focus was formed. For the remaining 3 cells that also showed only one co-localization event (a YPet-DnaN focus with the LacI-Cerulean focus), the RecA focus was formed at the site of the DSB within 5 to 15 minutes following the co-localization event.

For the 5 cells which showed either two or three co-localization events (a YPet-DnaN focus with the LacI-Cerulean focus), the last co-localization event always preceded formation of the RecA focus at the site of the DSB by less than 3 minutes. Thus, the data obtained from the time-lapse imaging indicated that 17 out of the 20 cells that were analysed showed a RecA focus at the site of the DSB within 3 minutes following the co-localization of a YPet-DnaN focus with the LacI-Cerulean focus. If the co-localization event which preceded formation of the RecA focus by 3 minutes (or less) represented the replication of the *lacZ* locus, it can be inferred that repair is initiated soon after DSB formation. The data also indicate that formation of the DSB and the initiation of repair occurred at the spatial localization for replication of the *lacZ* locus (Figure 5.8).

The kymographs presented in Figures 5.3B and 5.8 show that the dynamic movement of the *lacZ* locus around the mid-cell is constrained after the formation of the RecA focus at the site of the DSB. The data obtained from time-lapse imaging also indicate that the focus formed by YPet-DnaN was not always localized at the mid-cell after the formation of the RecA focus at the site of the DSB (the second YPet-DnaN kymograph of Figure 5.8). These observations demonstrate that the constraint on the dynamic movement of the *lacZ* locus after the formation of the RecA focus can be independent on the spatial localization of the replisome.

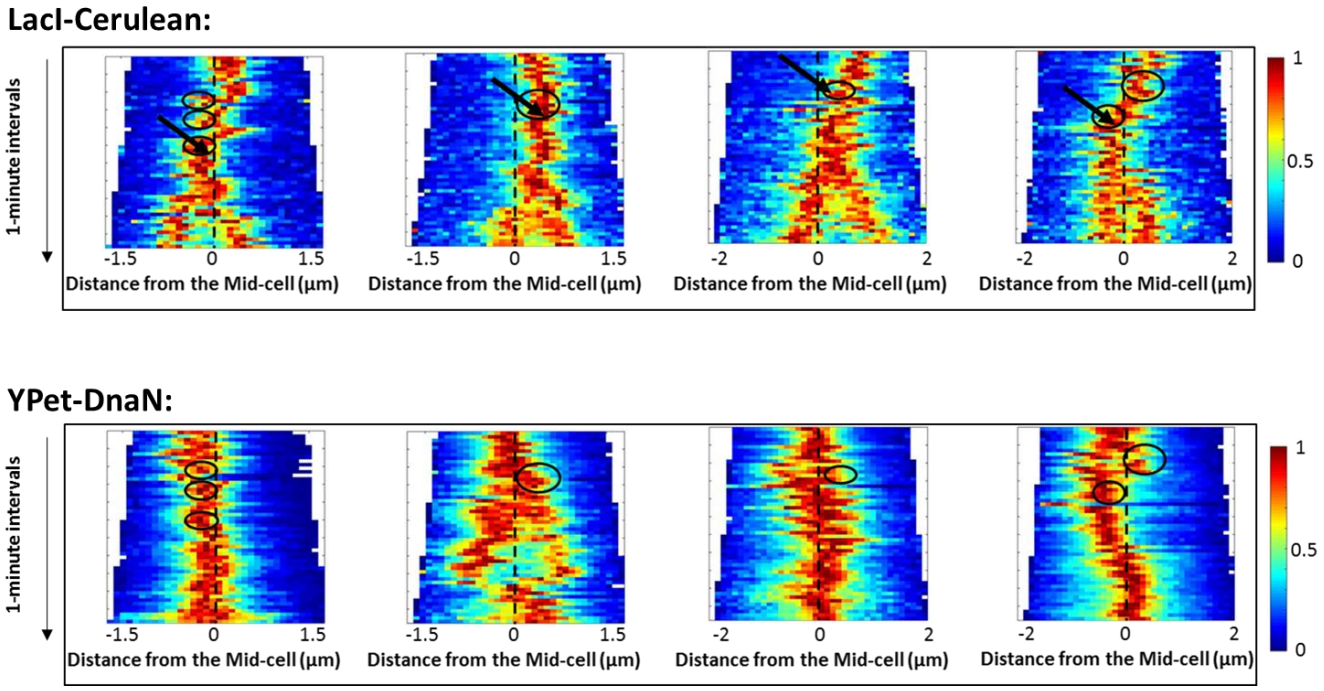


Figure 5.8: Spatial localization of the *lacZ* locus during formation and repair of the replication-dependent DSB. Phase contrast and fluorescence (LacI-Cerulean and YPet-DnaN) images were acquired by time-lapse microscopy after 2 hours of SbcCD expression in the strains containing the interrupted palindrome at the *lacZ* locus. The time-lapse images were acquired at 1-minute intervals. The black dash lines on the kymographs represent the mid-cell. The black arrows represent the formation of a RecA focus at the *lacZ* locus for cells undergoing DSB repair. Black circles represent time of co-localization of LacI-Cerulean focus with a YPet-DnaN focus. Kymographs were generated using the OUFTI software (Paintdakhi *et al*., 2016). The stain that was used for this analysis was DL6306 (*lacZ::pal*).

5.5 DISCUSSION

It has previously been reported that the repair of an I-SceI-induced DSB at the *cynR* locus caused a dramatic change in the localization of that locus when compared to its native localization (Lesterlin *et al*., 2014). The authors proposed that homology search at pre-synapsis might be responsible for the change in localization of the *cynR* locus during repair of the I-SceI-induced DSB at that locus (Lesterlin *et al*., 2014). In the present study, snapshot and time-lapse fluorescence imaging were used to investigate the spatial localization of the *lacZ* locus during replication, DSB formation and repair. Data from snapshot microscopy indicated that the presence of DSBs did not greatly affect the distribution of the subcellular localization of the *lacZ* locus. In Chapter 4 of this thesis, it was shown that the repair of these replication-dependent DSBs does not require an extensive homology search at pre-synapsis. Hence, it can be postulated that the absence of an extensive homology search at pre-synapsis underlies the similar subcellular distribution of the *lacZ* locus that was observed in the absence and presence of replication-dependent DSB repair.

In chapter 4 of this thesis, it was shown that only 6.5% of the cells that were analysed by snapshot microscopy showed a RecA focus that co-localized with the LacI-Cerulean focus following DSB induction. This observation could imply that the majority of the remaining 93.5% of the cells might not be undergoing DSB repair at the time of snapshot image acquisition. Thus, the similar distribution for the subcellular localization of the *lacZ* locus in the absence and presence of DSBs could also be attributed to the smaller proportion of cells that were undergoing DSB repair during snapshot image acquisition.

In the absence of DSBs, time-lapse imaging was used to demonstrate that the *lacZ* locus exhibited dynamic mobility around the mid-cell. Cohesion of the sister *lacZ* loci, which preceded segregation, did not restrain the dynamic movement of that locus. The dynamic movement of the *lacZ* locus might indicate that chromosomal loci are not entirely static at defined subcellular positions under normal growth conditions, in the absence of DSBs. This observation has previously been reported for other chromosomal loci in *E. coli* (Reyes-Lamothe *et al*., 2008 and Lesterlin *et al*., 2014). During repair of the replication-dependent DSBs, the dynamic movement of the *lacZ* locus was restrained close to the mid-cell after the formation and disassembly of the RecA focus at the site of the DSB. If the RecA focus at the site of the DSB is a fluorescent marker for both homology search and strand invasion, it could be inferred that the heteroduplex DNA, or some other feature of the repair complex, that is generated at the *lacZ* locus during synapsis restrains the dynamic movement of that locus.

Previous studies in *E. coli* had suggested that the re-pairing of distant sister chromosomes, after an extensive homology search, did not restrain the dynamic movement of the *cynR* locus during repair of I-SceI-induced DSBs (Lesterlin *et al*., 2014). Thus, the data obtained from that study did not favour the existence of a DSB repair centre in *E. coli*. In *C. crescentus*, a chromosomal locus bearing an I-SceI-induced DSB at one of the cell poles traversed the entire cell length to pair with the undamaged sister locus at the other cell pole (Badrinarayanan *et al*., 2015b). Furthermore, the authors suggested that repair of the DSB occurred at or very near to the localization of the undamaged sister template, which was at one of the cell poles. In the present study, the restrained movement of the *lacZ* locus close to the mid-cell, which occurred after the disassembly of the RecA focus, implied that DSB formation and the initiation of repair occurred at the spatial localization for replication of the *lacZ* locus. Moreover, the subcellular position of the RecA focus at the site of the DSB was reasonably consistent with the spatial localization of the *lacZ* locus prior to formation of the heteroduplex DNA.

Time-lapse imaging of the fluorescently-labelled β-sliding clamp was used to demonstrate that the two sister replisomes generally co-exist at the mid-cell, and occasionally undergo separation to the subcellular positions between the mid-cell and the ¼ position. This observation is consistent with previous studies that showed that sister replisomes generally associate with each other to form a centralized compartment at the mid-cell during chromosomal replication in *E. coli* (Koppes *et al*., 1999; Brendler *et al*., 2000; Molina and Skarstad, 2004). The duration of the YPet-DnaN foci during chromosomal replication was not significantly affected by the presence or absence of DSBs, demonstrating that formation of the replication-dependent DSB occurred behind the progressing replication fork at the *lacZ* locus.

The data from this study indicated that the co-localization of a YPet-DnaN focus with the LacI-Cerulean focus mostly preceded formation of the RecA focus by 3 minutes, or less, and was inferred to represented DSB formation in the presence of the interrupted palindrome and the SbcCD endonuclease. Because the RecA focus was usually formed by 3 minutes after DSB formation, it was deduced that repair was initiated soon after DSB formation. It also implied that the RecBCD enzyme rapidly processed the ends of the DSB during initiation of repair *in vivo*. Together with the data obtained in chapter 4 of this thesis, it can be inferred that processing of the ends of the DSB, formation of the RecA nucleoprotein filament and the concomitant homology search and strand invasion reactions were rapid events during repair of the replication-dependent DSB that was generated at the *lacZ* locus.

Chapter 6: Genomic analysis of RecA binding at fluorescently-labelled *lac* locus during the repair of replication-dependent DSBs

6.1 Introduction

The binding of fluorescently-labelled repressors (LacI-ECFP and TetR-EYFP) to arrays of the *E. coli lac* and *Tn10 tet* operator sequences have been used as a tool for simultaneous visualization of two chromosomal loci by fluorescence microscopy (Lau *et al*., 2003). This system has also been used to visualize both ends of a replication-dependent DSB that was generated at the *lacZ* locus of the *E. coli* chromosome (White *et al*., 2008). Unlike the fluorescent *in situ* hybridization technique, the fluorescent repressor operator system (FROS) does not require cell fixation for the visualization of chromosomal loci. Thus, FROS is widely used for visualization of cellular events as they occur in live cells. A drawback associated with the use of FROS is that the fluorescently-labelled repressors, bound to their cognate operator array, have a tendency to stall a progressing replication fork (Possoz *et al*., 2006). In wild type *E. coli* strains, stalled replication forks stimulate homologous recombination via the RecBCD pathway (Michel *et al*., 2004), suggesting that FROS might induce the formation of one-ended DSBs.

In this study, arrays of *tetO* and *lacO* were inserted at the origin-proximal and origin-distal ends of the *lacZ* locus respectively. Each operator array had a size of approximately 10kb and was integrated 6kb away from the site of the interrupted palindrome at *lacZ* (Figure 6.1). The *tetO* and *lacO* arrays contained a gentamicin and a kanamycin resistance gene, respectively, to ensure that the operator arrays could be selected in these strains. A synthetic array of three Chi sites (the red lines in Figure 6.1) was also inserted between each operator array and the site of the interrupted palindrome at *lacZ*. The genes encoding the TetR-YPet and LacI-Cerulean proteins were inserted in the *ykgC* locus of the *E. coli* chromosome.

Previous genomic analysis has demonstrated that RecA enrichment initiates from the position of a correctly-oriented Chi site at the *lac* locus during the repair of palindrome-induced DSBs (Figure 6.2; Cockram *et al*., 2015). Data from chromatin immunoprecipitation, followed by high-throughput sequencing (ChIP-seq), revealed that RecA enrichment was

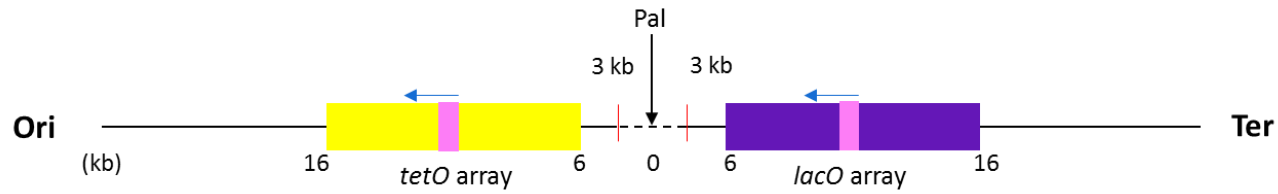


Figure 6.1: Schematic representation of the fluorescent repressor operator system that was used for the visualization of the *lac* locus during DSB repair. Pal represents the site of the interrupted palindrome at *lacZ*. The synthetic Chi arrays are shown in red lines. The pink boxes indicate the gentamycin and kanamycin resistance genes within the *tetO* and *lacO* arrays, respectively. The directionality of transcription of the antibiotic resistance genes within the operator arrays are shown in blue arrows. Distances of DNA sequences from the site of the interrupted palindrome are shown in kilobases.

maximum within a 10kb to 15kb DNA region from the site of a synthetic Chi array (Figure 6.2). As a result, it was anticipated that in the *E. coli* strains with the repressor-bound operator arrays, the RecBCD complex would initiate the loading of RecA from the position of the synthetic Chi array, which is 3kb away from the site of the DSB. Upon encountering the repressor-bound operator arrays, it is unknown whether the helicase activity of RecBCD can displace the repressors from their binding sites and load RecA within the array of the operator sequences. Moreover, the binding of RecA to DNA sequences beyond the operator arrays might be perturbed by the binding of the repressors to their cognate operator sequences.

In this Chapter, ChIP-seq was used to characterize the profile of RecA binding at the *lac* locus, and the DNA regions surrounding that locus, following the formation of replication-dependent DSBs in strains containing the repressor-bound operator arrays. ChIP-seq was also used to ascertain the effect of the presence of the operator arrays, without their cognate repressors, on the RecA enrichment at the *lac* locus during the repair of the DSBs. Thus, the data from these ChIP-seq experiments were used to determine whether the binding of the repressors to their cognate operator arrays perturbs the binding of RecA at the *lac* locus during the repair of the replication-dependent DSBs.

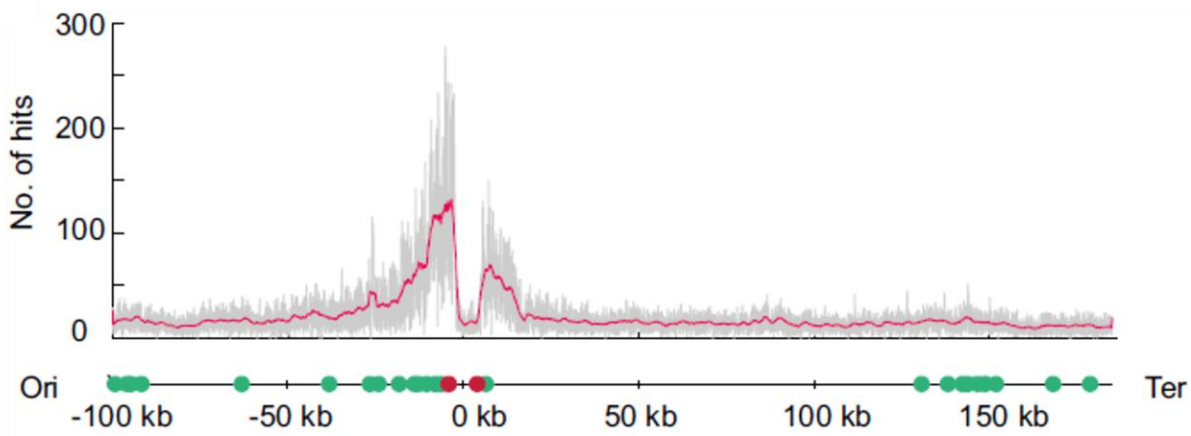


Figure 6.2: Chi-dependent RecA enrichment at the *lac* locus during the repair of replication-dependent DSBs. The site of the interrupted palindrome was designated 0 kb. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The *E. coli* strain was grown in M9-glucose medium supplemented with casamino acids. This figure was obtained from Cockram *et al*., 2015.

6.2 Effect of the presence of the operator arrays and their cognate repressors on the profile of the binding of RecA at the *lac* locus during replication-dependent DSB repair.

The data which demonstrated a Chi-dependent RecA enrichment at the *lac* locus during DSB repair were obtained from *E. coli* strains expressing only the native RecA protein (Figure 6.2; Cockram *et al*., 2015). In this study, RecA-ChIP was also performed using *E. coli* strains (*lacZ*$^+$ and *lacZ::pal*) which expressed only the native RecA protein. Additionally, these strains had the *tetO* and *lacO* arrays and the synthetic Chi arrays, as illustrated in Figure 6.1. Finally, the *lacI-Cerulean* and *tetR-YPet* genes were either absent or inserted in the *ykgC* locus of these strains.

Prior to the RecA-ChIP experiment, overnight cultures of these strains, grown in M9-glycerol medium at 37°C, were diluted in the same medium and grown to exponential phase. The cultures were re-diluted and the expression of the SbcCD endonuclease was induced for 2 hours by the addition of arabinose. Afterwards, the cells were harvested and the chromosomal loci that were bound to RecA were obtained by following the procedure for chromatin immunoprecipitation using anti-RecA antibodies (Chapter 2.2.5). The DNA samples that were obtained from the RecA-ChIP experiment were used for DNA library preparation (described in Chapter 2.2.6). The DNA libraries were sent to the Edinburgh Genomics facility for high-throughput DNA sequencing.

Prior to analysing the DNA reads that were obtained from high-throughput sequencing, the DNA sequences of the *tetO* and *lacO* arrays were determined by Sanger sequencing and incorporated into the reference genome of the *E. coli* BW27784 strain. In order to determine the DNA sequences of both operator arrays, primers were initially designed within the 300 bp DNA sequences of the *cynX* and *mhpC* genes that are flanking the *tetO* and *lacO* arrays, respectively. The DNA reads that were obtained by Sanger sequencing using these initial primers were used to design subsequent primers, which annealed to DNA sequences within the operator arrays. The procedure was repeated systematically until the entire DNA sequences of both operator arrays were determined. Moreover, the known DNA sequences of the gentamycin and kanamycin resistance genes were also used for designing primers to facilitate the process of obtaining the complete DNA sequences of the *lacO* and *tetO* arrays by Sanger sequencing. The reference genome of the *E. coli* BW27784 strain was modified by deleting the coding sequences of the endogenous *lacI* gene and incorporating the complete DNA sequences of the two operator arrays. The *lacI* gene is a 1,083bp DNA sequence which is located between the site of the interrupted palindrome and the *lacO* array. During strain construction, the *lacI* gene was deleted to prevent competition between the endogenous LacI proteins and the LacI-Cerulean proteins for binding to the *lacO* array.

The primers that were used for sequencing the operator arrays were designed by the author of this thesis and Mahedi Hasan (a post-doctoral research fellow). Mapping of the DNA reads, obtained from the high-throughput sequencing, to the modified reference genome of the *E. coli* BW27784 strain was performed by Mahedi Hasan.

6.2.1 DSB-independent RecA enrichment at the site of the antibiotic resistance genes in the operator arrays

In the absence of DSB induction, the ChIP-seq data revealed that there is RecA binding within the *tetO* and *lacO* arrays, even though there is no Chi site in these operator arrays. DSB-independent RecA enrichment within the *lacO* array was confined to the kanamycin resistance gene in the presence of LacI-Cerulean repressors (Figure 6.3A) and extended beyond the gene into the operator array in the absence of the repressors (Figures 6.3B). In contrast, DSB-independent RecA enrichment within the *tetO* array extended from the gentamycin resistance gene to the operator array in the absence or presence of TetR-YPet

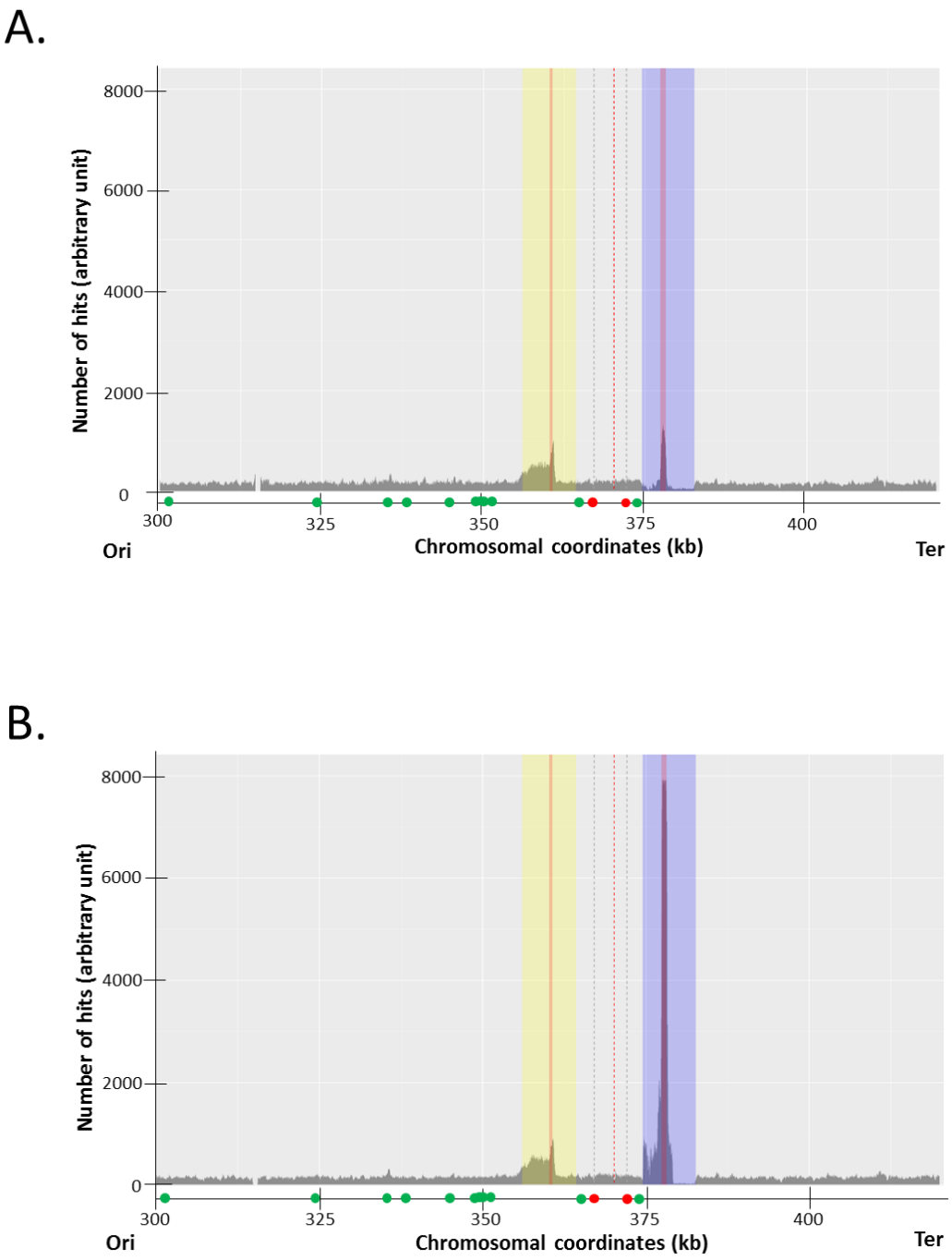


Figure 6.3: DSB-independent RecA enrichment at the site of the operator arrays in the absence and presence of the TetR-YPet and LacI-Cerulean proteins under slow growth condition. (A) Profile of RecA binding within the repressor-bound operator arrays. (B) Profile of RecA binding within the operator arrays in the absence of the TetR-YPet and LacI-Cerulean proteins. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The red dash lines represent the site of the interrupted palindrome at *lacZ*. The grey dash lines represent the positions of the synthetic Chi arrays on the graph. The yellow and blue shaded regions correspond to the *tetO* and *lacO* arrays, respectively. The pink stripe within the yellow and blue shaded regions represents the gentamycin and kanamycin resistance genes, respectively.

proteins (Figure 6.3). Furthermore, the peak of DSB-independent RecA enrichment at the kanamycin resistance gene within the *lacO* array was higher in the absence of LacI-Cerulean repressors compared to the presence of these repressors (Figure 6.3). The peak of DSB-independent RecA enrichment within the *tetO* array was similar in the presence of TetR-YPet proteins compared to the absence of the repressors (Figure 6.3). Previous studies have shown that DSB-independent RecA enrichment was detected at the site of highly transcribed genes in *E. coli* (Cockram *et al*., 2015). Based on the profile of the DSB-independent RecA enrichment within the operator arrays, it was inferred that the binding of RecA in the *tetO* and *lacO* arrays were associated with the transcription of the gentamycin and kanamycin resistance genes, respectively (the blue arrows of Figure 6.1). The presence of transcription-dependent RecA enrichment indicates that, following the formation of palindrome-induced DSBs, RecA binding within the operator arrays might not be entirely due to the repair of the DSBs.

6.2.2 Chi-dependent RecA enrichment at the *lac* locus of *E. coli* strains containing the *tetO* and *lacO* arrays in the absence of the TetR-YPet and LacI-Cerulean proteins

In the absence of DSB formation in the *lacZ*$^+$ strain, the ChIP-seq data showed that Chi-dependent RecA enrichment was not detected at the *lacZ* locus and the 40kb DNA region surrounding that locus (Figure 6.4, DSB$^-$). Following the induction of SbcCD expression in the palindrome-containing strain, RecA enrichment was detected from the synthetic Chi arrays at the origin-proximal and origin-distal ends of the DSB (Figure 6.4, DSB$^+$). Additionally, there was RecA enrichment at the endogenous Chi sites at the origin-proximal end of the DSB, which were located approximately 20kb away from the site of the synthetic Chi array (Figure 6.4, DSB$^+$). The Chi-dependent RecA enrichment covered approximately 70kb of DNA on the origin-proximal end of the DSB. The presence of DSB-dependent RecA enrichment at the endogenous Chi sites beyond the synthetic Chi array was previously reported by Cockram *et al*. (2015).

The profile of Chi-dependent RecA enrichment at the origin-distal end of the DSB was dissimilar to that observed at the origin-proximal end. Firstly, the peak of the RecA enrichment originating from the synthetic Chi array at the origin-distal end of the DSB was eight-fold lower compared to the corresponding peak at the origin-proximal end (Figure 6.4, DSB$^+$). Secondly, the RecA enrichment starting at the synthetic Chi array at the origin-distal end of the DSB was only detected within a 7kb DNA region (Figure 6.4, DSB$^+$). DSB-dependent RecA enrichment

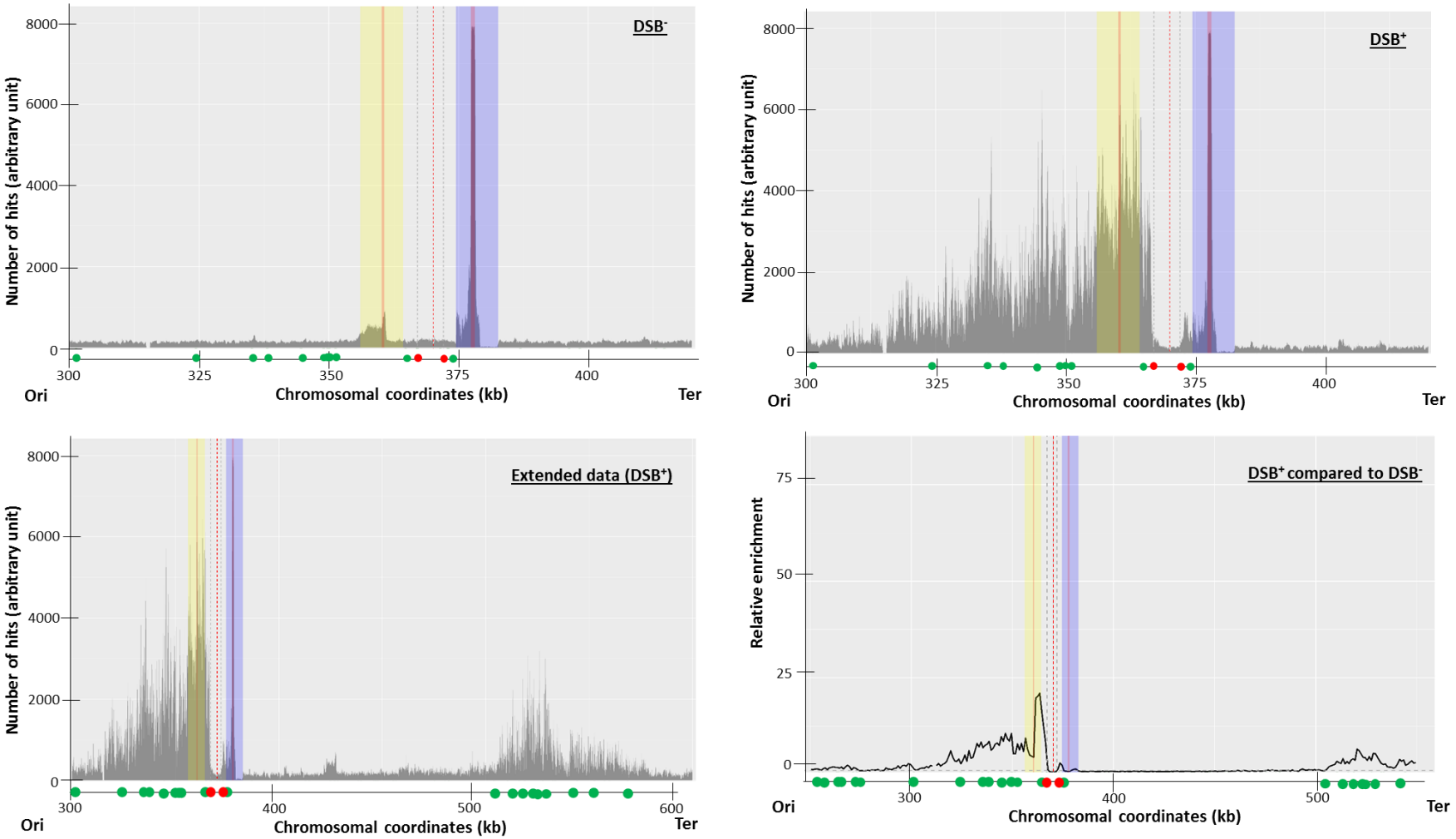


Figure 6.4: RecA enrichment at the *lac* locus and the DNA regions surrounding that locus in the absence of the TetR-YPet and LacI-Cerulean proteins under slow growth condition. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The red dash lines represent the site of the interrupted palindrome at *lacZ*. The grey dash lines represent the positions of the synthetic Chi arrays on the graph. The yellow and blue shaded regions correspond to the *tetO* and *lacO* arrays, respectively. The pink stripe within the yellow and blue shaded regions represent the gentamycin and kanamycin resistance genes, respectively. DSB- represents the absence of DSB induction. DSB+ represents the presence of DSB induction.

was also detected at the endogenous Chi sites which are located at 120kb downstream of the *lacO* array (Figure 6.4, Extended data). Comparison of the ChIP-seq data obtained from the palindrome-containing strain and the strain without the interrupted palindrome at *lacZ* revealed that the RecA enrichment within the *lacO* array is mainly independent of the formation of the DSB (Figure 6.4, Relative enrichment). The observations from this ChIP-seq data illustrate that the presence of the *lacO* array perturbed the widespread DSB-dependent binding of RecA from the synthetic Chi array at the origin-distal end of the DSB, compared to previously reported ChIP-seq data by Cockram *et al*. (2015). Nonetheless, the presence of the *tetO* array did not inhibit the extensive binding of RecA at the origin-proximal end of the DSB.

The ChIP-seq experiment was repeated by growing the same strains in LB (rich) medium to ascertain whether the marginal RecA enrichment at the synthetic Chi array at the origin-distal end of the DSB was the result of the slow growth condition. The ChIP-seq data revealed that the Chi-dependent RecA enrichment at both ends of the DSB were higher for the palindrome-containing strain grown in LB medium (Figure 6.5, DSB^+) compared to the same strain grown in M9-glycerol medium (Figure 6.4, DSB^+). Notably, the peak of the Chi-dependent RecA enrichment at the origin-distal end of the DSB was two-fold to three-fold lower than the corresponding RecA enrichment at the origin-proximal end when the strain was grown in LB medium(Figure 6.5, DSB^+). These observations demonstrate that the marginal binding of RecA at the synthetic Chi array at the origin-distal end of the DSB was a consequence of growing the cells slowly in M9-glycerol medium.

Despite higher Chi-dependent RecA binding on the origin-distal end of the DSB (Figure 6.5, DSB^+), RecA enrichment was still restricted to the 7kb DNA region from the site of the synthetic Chi array, as was initially observed when the strain was grown in M9-glycerol medium (Figure 6.4, DSB^+). Moreover, RecA enrichment was also detected at the endogenous Chi sites that were downstream of the *lacO* array on the origin-distal end of the DSB (Figure 6.5, Extended data). The ChIP-seq data revealed that DSB-dependent binding of RecA at the origin-proximal end was broadly distributed across the Chi sites when the strain was grown in M9-glycerol medium (Figure 6.4, DSB^+). Under fast growth condition (LB medium), DSB-dependent RecA binding increased abruptly from the synthetic Chi array at the origin-proximal end of the DSB and was reduced by three-fold at the subsequent endogenous Chi sites (Figure 6.5, DSB^+).

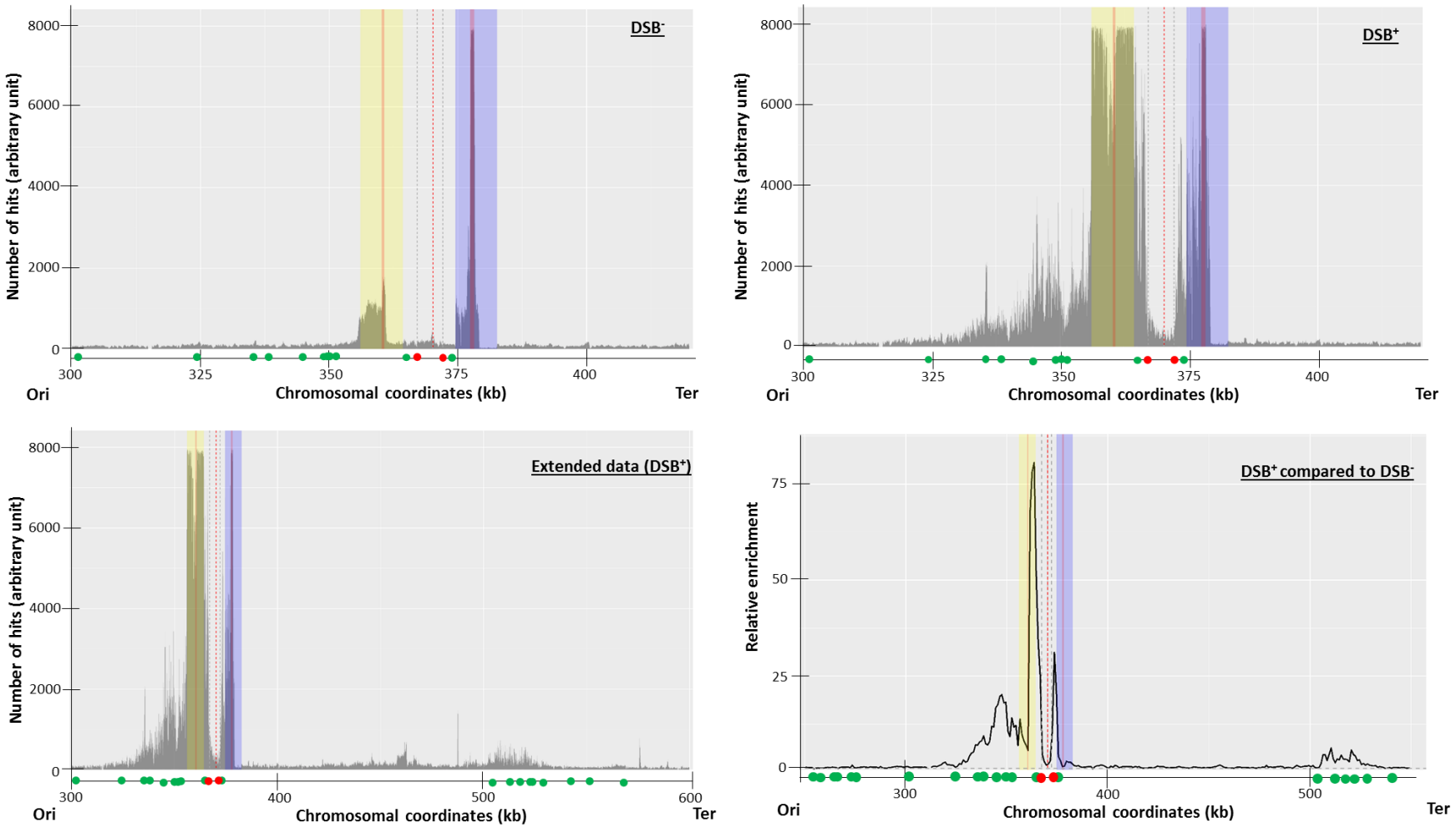


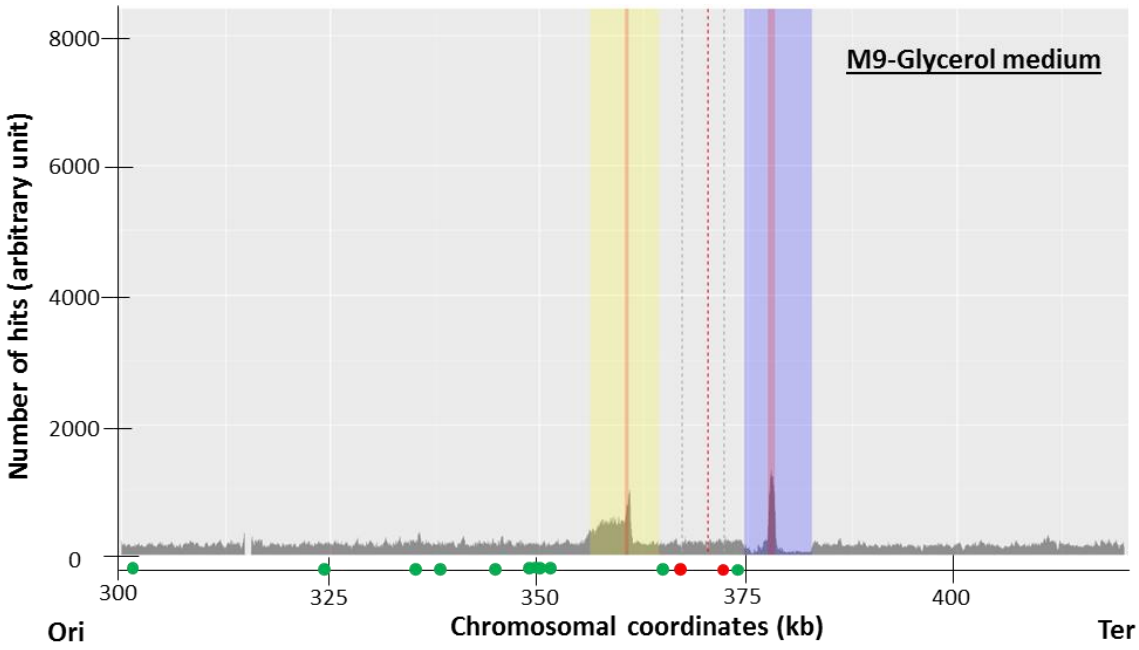
Figure 6.5: RecA enrichment at the *lac* locus and the DNA regions surrounding that locus in the absence of the TetR-YPet and LacI-Cerulean proteins for the fast growth condition. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The red dash lines represent the site of the interrupted palindrome at *lacZ*. The grey dash lines represent the positions of the synthetic Chi arrays on the graph. The yellow and blue shaded regions correspond to the *tetO* and *lacO* arrays, respectively. The pink stripe within the yellow and blue shaded regions represent the gentamycin and kanamycin resistance genes, respectively. DSB- represents the absence of DSB induction. DSB+ represents the presence of DSB induction.

As a result, it can be deduced that the RecBCD complex recognizes the synthetic Chi array on the origin-proximal end of the DSB more efficiently when the palindrome-containing strain was grown in LB medium compared to M9-glycerol medium. Additionally, the synthetic Chi arrays on both ends of the DSB were not recognized with the same efficiency by the RecBCD complex under either slow or fast growth conditions.

6.2.3 Effect of the presence of repressor-bound *tetO* and *lacO* arrays on Chi-dependent RecA enrichment at the *lac* locus

The ChIP-seq data for the *lacZ*$^{+}$ strain containing the repressor-bound operator arrays revealed that there was no Chi-dependent RecA enrichment at the *lacZ* locus and the 40kb DNA region surrounding that locus (Figure 6.6). This observation suggests that the formation of one-ended DSBs due to stalling of replication forks at the repressor-bound operator arrays was either absent or undetected. Thus, the Chi-dependent RecA enrichment detected at the origin-proximal and origin-distal ends of the DSB were entirely generated by the repair of the palindrome-induced DSBs.

The expression of both the LacI-Cerulean and the TetR-YPet proteins had no detectable effect on the profile of the binding of RecA at the origin-proximal and origin-distal ends of the DSB in the palindrome-containing strain grown in LB medium (Figure 6.7). When the strain was grown in M9-glycerol medium, the RecA enrichment at the origin-proximal end of the DSB and at the synthetic Chi array on the origin-distal end of the DSB were not greatly affected by the presence of the TetR-YPet and LacI-Cerulean repressors (Figure 6.7). However, the RecA enrichment at the endogenous Chi sites which are located at 120kb downstream of the *lacO* array was lower in the presence of the LacI-Cerulean repressors (Figure 6.7A) compared to the absence of the repressors (Figure 6.7B) when the cells were in slow growth condition. These observations demonstrate that the repressor-bound *lacO* array perturbs the initiation of DSB repair from the endogenous Chi sites, which are located beyond the *lacO* array, in a subpopulation of cells in the slow growth condition.



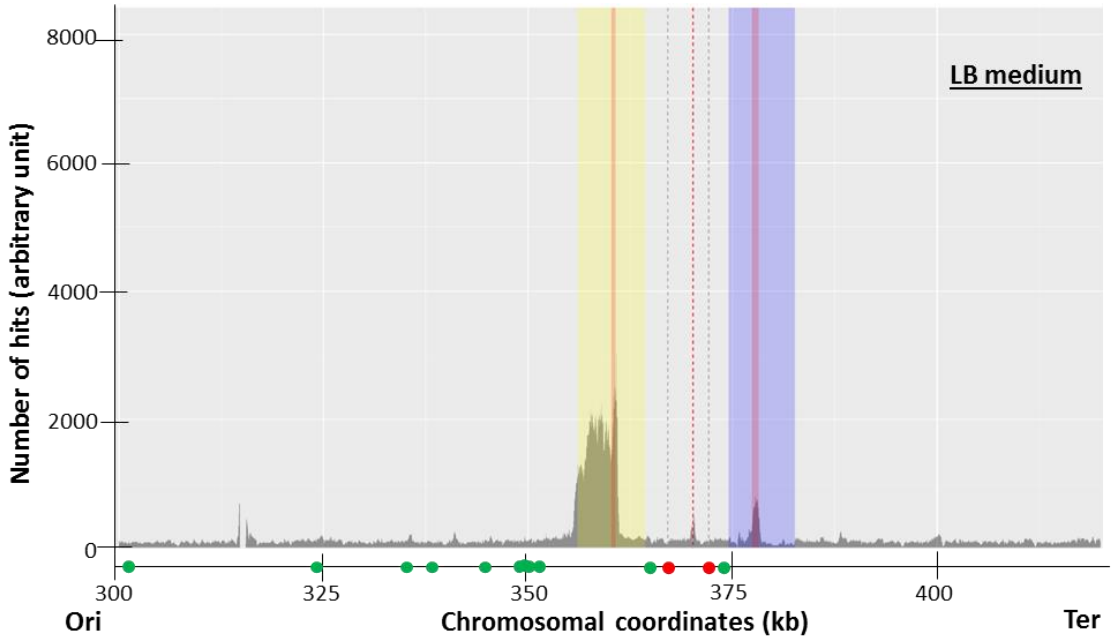


Figure 6.6: DSB-independent RecA enrichment at the *lac* locus and the DNA regions surrounding that locus in the presence of the repressor-bound operator arrays. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The red dash lines represent the site of the interrupted palindrome at *lacZ*. The grey dash lines represent the synthetic Chi arrays on the origin-proximal and origin-distal ends of the DSB. The yellow and blue shaded regions correspond to the *tetO* and *lacO* arrays respectively. The pink stripe within the yellow and blue shaded regions represent the gentamycin and kanamycin resistance genes respectively.

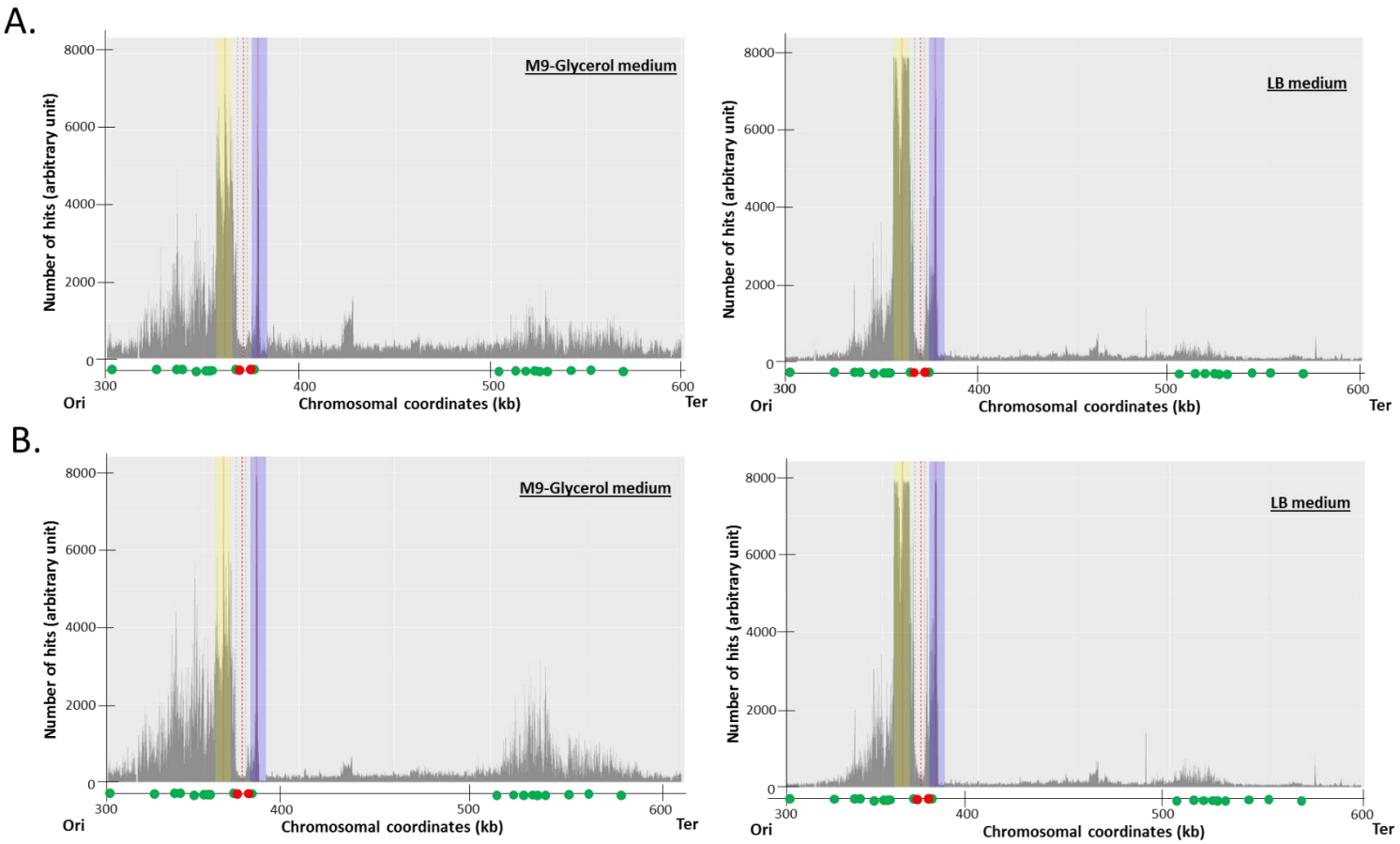


Figure 6.7: Chi-dependent RecA enrichment at the *lac* locus and the DNA regions surrounding that locus in the absence and presence of the repressors. (A) The profile of RecA binding in the presence of the repressor-bound operator arrays. (B) The profile of RecA binding in the absence of the TetR-YPet and LacI-Cerulean proteins. The red circles represent the synthetic Chi arrays and the green circles represent the correctly-oriented endogenous Chi sites. The red dash lines represent the site of the interrupted palindrome at *lacZ*. The grey dash lines represent the synthetic Chi arrays on the origin-proximal and origin-distal ends of the DSB. The yellow and blue shaded regions correspond to the *tetO* and *lacO* arrays respectively. The pink stripe within the yellow and blue shaded regions represent the gentamycin and kanamycin resistance genes respectively.

6.3 DISCUSSION

ChIP-seq was used to identify the profile of the binding of RecA at the *lac* locus and DNA regions surrounding that locus in *E. coli* strains containing the *tetO* and *lacO* arrays and synthetic Chi arrays flanking the site of the interrupted palindrome. In the absence of DSB induction, Chi-dependent RecA enrichment was not detected at the 40kb DNA region surrounding the *lacZ* locus despite the presence of the repressor-bound operator arrays. From this observation, it can be inferred that the binding of the repressors to their operator arrays did not stall progressing replication forks. If replication forks were stalled at the site of the repressor-bound operator arrays, they would have generated one-ended DSBs by replication-fork reversal. Therefore, it can be deduced that in the absence of DSB induction, the RecA foci that co-localized with the YPet-Cerulean foci during live-cell fluorescence imaging (Chapter 4) were possibly not generated by the repair of one-ended DSBs originating from the repressor-bound operator arrays.

The ChIP-seq data also revealed the presence of DSB-independent RecA enrichment within the *tetO* and *lacO* arrays. The profile of the DSB-independent RecA enrichment was consistent with the directionality of transcription of the antibiotic resistance genes within the operator arrays. Consequently, it can be inferred that the DSB-independent RecA enrichment within the *lacO* and *tetO* arrays were associated to the transcription of the gentamycin and kanamycin resistance genes. DSB-independent RecA enrichment has been detected at highly transcribed genes across the *E. coli* chromosome (Cockram *et al*., 2015). The presence of DSB-independent RecA enrichment at the site of the gentamycin and kanamycin resistance genes may indicate that these genes were highly transcribed in the *E. coli* strains grown in either LB or M9-glycerol medium.

A search for a palindromic GC-rich DNA sequence followed by a polyT-tail revealed that there was no transcription terminator sequence within the *tetO* and *lacO* arrays. Because the DSB-independent RecA enrichment within the operator arrays always extended beyond the site of the antibiotic resistance genes in the absence of the repressors, it can be postulated that the termination of transcription of these antibiotic resistance genes occurred possibly beyond the operator arrays. In the presence of the LacI-Cerulean repressors, the repressor-bound *lacO* array might have hindered the progression of the RNA polymerase beyond the kanamycin resistance gene and thus confined the DSB-independent RecA enrichment to the site of the gene (Figure 6.3B). However, the binding of the TetR-YPet proteins to the *tetO* array did not confine the DSB-independent RecA enrichment to the site of the gentamycin resistance gene.

Thus, the transcription-dependent binding of RecA extended from the site of the gentamycin resistance gene to the repressor-bound *tetO* array.

It has previously been shown that binding of TetR-YPet proteins to the *tetO* array caused a detectable loss of cell viability (Martin White, unpublished). This loss of viability was alleviated by the addition of anhydrotetracycline (ATC) to the growth medium (Lau *et al*., 2003). In the present study, the growth medium was always supplemented with ATC when growing *E. coli* strains containing the *tetO* array and the TetR-YPet proteins. ATC functions by binding to the Tet repressors and reducing their affinity for the *tetO* sequences via a change in conformation of these repressors (Gossen and Bujard, 1993). Consequently, the amount of TetR-YPet proteins that are bound to the *tetO* array is reduced in the presence of ATC.

Due to the absence of a transcription terminator sequence within the *tetO* array, the RNA polymerase which transcribed the gentamycin resistance gene could possibly displace the TetR-YPet proteins that were bound to the *tetO* array because lower amount of these repressors were bound to the operator array in the presence of ATC. Thus, the DSB-independent RecA enrichment within the *tetO* array was not confined to the gentamycin resistance gene in the presence of the TetR-YPet proteins because of the addition of ATC to the growth medium. The presence of ATC does not affect the binding of LacI-Cerulean repressors to the *lacO* array. Therefore, it can be suggested that the repressor-bound *lacO* array was a stronger complex which inhibited the further progression of the RNA polymerase after transcription of the kanamycin resistance gene. Thus, the DSB-independent RecA enrichment was confined to the kanamycin resistance gene within the *lacO* array in the presence of the LacI-Cerulean repressors. The transcription-dependent RecA enrichment within the operator arrays might possibly correspond to the RecA foci that co-localized with the YPet-Cerulean foci in the absence of DSB induction during live-cell fluorescence imaging (Chapter 4).

Following SbcCD expression in the palindrome-containing strains, RecA enrichment was shown to originate from the position of the synthetic Chi arrays at both ends of the DSB. This observation has been inferred to correspond to the RecA nucleoprotein filament that is formed during the repair of the DSB (Cockram *et al*., 2015). In the present study, the DSB-dependent RecA binding increased abruptly from the synthetic Chi array at the origin-proximal end of the DSB and reduced by three-fold at the subsequent endogenous Chi sites when the strains were grown in LB medium. However, the binding of RecA at the origin-proximal end of the DSB was broadly distributed across the Chi sites when the strains were grown in M9-glycerol

medium. These observations demonstrate that the recognition of the synthetic Chi array at the origin-proximal end of the DSB was more efficient in cells under fast growth condition compared to slow growth condition. Cockram *et al*. (2015) reported the existence of two populations of RecBCD molecules during the repair of the palindrome-induced DSBs. These populations were RecBCD molecules with either low or high probability of recognizing correctly-oriented Chi sites. The authors proposed that RecBCD molecules with a low probability of recognizing Chi sites might have to travel further before recognizing a Chi site. In contrast, RecBCD molecules with a high probability of recognizing Chi sites would not have to travel long distances on the DNA because they could recognize the synthetic Chi arrays on both ends of the DSB. From these information, it can be postulated that RecBCD molecules with a low probability of recognizing Chi sites would be prevalent in the cells under slow growth condition compared to the fast growth condition.

Under slow growth condition, the DSB-dependent RecA enrichment was marginal at the synthetic Chi array on the origin-distal end but was abundant at subsequent Chi sites beyond the *lacO* array. This observation supports the hypothesis that RecBCD molecules with low probability of recognizing Chi sites would be prevalent under the slow growth condition compared to the fast growth condition. Thus, it was not surprising that RecA enrichment at the synthetic Chi array on the origin-distal end of the DSB was higher under fast growth condition with a corresponding decrease in RecA binding at the endogenous Chi sites beyond the *lacO* array. Consequently, it can be inferred that resection of the ends of the DSBs by the RecBCD complex was more extensive under slow growth condition compared to the fast growth condition.

Under the slow growth condition, DSB-dependent RecA enrichment at the endogenous Chi sites beyond the *lacO* array decreased in the presence of the LacI-Cerulean repressors relative to the corresponding RecA enrichment that was detected in the absence of the repressors. Thus, the extensive movement of the RecBCD molecules with a low probability of recognizing Chi sites could be obstructed by the presence of the repressor-bound *lacO* array in a subpopulation of cells during the repair of the DSB. This observation was not detected when the strains were grown in LB medium possibly because the RecBCD molecules with a low probability of recognizing Chi sites were not prevalent under the fast growth condition.

RecA enrichment on the origin-distal end of the DSB was not detected beyond a 7kb DNA region from the site of the synthetic Chi array. This observation was unaffected by the presence

or absence of the LacI-Cerulean repressors. In the absence of the *lacO* array and the LacI-Cerulean repressor, Cockram *et al*. (2015) reported that Chi-dependent RecA enrichment on the origin-distal end of the DSB covered a 15kb DNA region from the site of the synthetic Chi array (Figure 6.2). The data from this present study indicate that the presence of the *lacO* array, without LacI-Cerulean repressors, resulted in a decrease in the size of the DNA region that was covered by RecA binding from the synthetic Chi array on the origin-distal end of the DSB. Because this observation was not detected on the *tetO* array, it can be hypothesized that the nature of the *lacO* sequence within the *lacO* array might have decreased the distance travelled by the RecBCD complex after recognizing the Chi array. The *lacO* sequence is moderately palindromic whiles the *tetO* sequence is not palindromic. The moderately-palindromic nature of the *lacO* sequence might be a possible explanation for the absence of widespread DSB-dependent RecA enrichment from the synthetic Chi array at the origin-distal end of the DSB.

Cockram *et al*. (2015) reported that Chi-dependent RecA enrichment at the origin-proximal end of the DSB was two-fold higher than the corresponding RecA enrichment at the origin-distal end when the strain was grown in M9-glucose medium supplemented with casamino acids. The authors also observed that Chi-dependent RecA enrichment was not detected at the endogenous Chi sites beyond the synthetic Chi array on the origin-distal end of the DSB. Thus, it was proposed that approximately 50% of the palindrome-induced DSBs were converted to one-ended DSBs by the action of the RecBCD complex (Cockram *et al*., 2015). In the present study, the ChIP-seq data also demonstrated that the overall Chi-dependent RecA enrichment at the origin-proximal end of the DSB was always higher relative to the origin-distal end. This might suggest that a proportion of the palindrome-induced DSBs were converted to one-ended DSBs in a subpopulation of the cells grown in either M9-glycerol medium or LB medium.

This study utilized ChIP-seq to identify the profile of the binding of RecA at the *lac* locus, and the DNA regions surrounding that locus, in the recombination-proficient *E. coli* strains that were used for live-cell fluorescence imaging presented in Chapter 4 of this thesis. The only modification to these strains was the absence of the *recA-mCherry* gene. The data obtained from ChIP-seq provided useful insight on the origin of the RecA foci that co-localized with the YPet-Cerulean foci in the absence and presence of DSBs. It was deduced that in the absence of DSB induction, the RecA foci that co-localized with the YPet-Cerulean foci originated primarily from the transcription of the antibiotic resistance genes within the operator arrays. Moreover, ChIP-seq data did not provide evidence for the formation of one-ended DSBs via replication-fork reversal at the site of the repressor-bound operator arrays. Following SbcCD

expression in the palindrome-containing strains, it was observed that the two-ends of the DSB underwent extensive resection in a subpopulation of cells grown in M9-glycerol medium compared to LB medium. Notably, the ChIP-seq data revealed that the binding of the TetR-YPet proteins to the *tetO* array did not perturb the binding of RecA at the origin-proximal end of the DSB, possibly because of the addition of ATC to the growth medium. However, the binding of the LacI-Cerulean proteins to the *lacO* array perturbed the binding of RecA at the endogenous Chi sites beyond the *lacO* array for the strains grown in M9-glycerol medium. In addition, the presence of the *lacO* array perturbed the widespread binding of RecA from the synthetic Chi array on the origin-distal end of the DSB under either the slow or fast growth condition.

7. CONCLUSIONS AND FUTURE WORK

7.1 CONCLUSIONS

The insertion of a 246 bp interrupted palindrome at the *lacZ* locus and the expression of the SbcCD endonuclease from an arabinose-inducible promoter were used to generate a replication-dependent DSB at the *lacZ* locus in the *E. coli* chromosome. The site of the DSB was visualized via the binding of TetR-YPet and LacI-Cerulean to the *tetO* and *lacO* arrays, respectively, which were flanking the *lacZ* locus. Additionally, a codon diversified *recA-mCherry* gene was inserted next to the endogenous *recA* gene, with both alleles under the control of the native *recA* promoter. By using these systems, it was shown that the RecA-mCherry protein alone is not fully functional for the repair of the DSB that was generated at the *lacZ* locus. However, the presence of both the endogenous and the fluorescently-labelled RecA proteins restore recombination proficiency during DSB repair.

For live-cell imaging, the strains were grown in M9-glycerol medium supplemented with arabinose. The addition of arabinose to the growth medium allowed homogenous expression of the SbcCD endonuclease from the arabinose-inducible promoter. Analysis of the DNA content of cells which were growing exponentially in M9-glycerol medium was used to demonstrate that more than 80% of the cells had a single replicating chromosome in the absence and presence of DSBs. Additionally, the data from the DNA content of these cells revealed that the formation of DSBs did not perturb the B-phase of the cell cycle (duration from birth to initiation of chromosomal replication in new-born cells). Time-lapse imaging of YPet-DnaN revealed that the duration of chromosomal replication (C-phase) was not perturbed by the presence of DSBs. Consequently, the increase in generation time by approximately 7 minutes in response to DSB formation was attributed to an increase in the duration of the D-phase (termination of replication to cell division) of the cell cycle. Despite the increase in generation time in the presence of DSB formation, the distribution of cell lengths and the mean cell length were not significantly affected in the absence and presence of DSBs.

Data from snapshot microscopy showed that 50% of the cells had a single YPet_Cerulean focus while 42% of the cells had two of these twin foci in the presence of DSB formation. Time-lapse imaging was used to demonstrate that DSB repair was initiated in cells with a single YPet_Cerulean focus and that the repair was completed before segregation of the sister *lacZ* loci to generate two YPet_Cerulean foci. During repair, the duration of the RecA focus at the site of the DSB was transient in relation to the generation time. This study did not provide data

for the direct estimation of the duration of cohesion of sister *lacZ* loci in the absence and presence of DSBs. By comparing the duration of cohesion of sister *lacZ* loci after disassembly of the RecA focus with previously estimated duration of cohesion of that locus in the absence of DSBs (Joshi *et al*., 2011), it was postulated that DSB repair might have prolonged the average duration of cohesion of the sister *lacZ* loci by approximately 3-fold.

Time-lapse imaging also revealed that the fluorescently-labelled RecA protein formed three kinds of foci in the absence or presence of DSBs. These were a RecA focus at the site of the DSB, a RecA focus at other loci within the cell and a RecA focus at a cell pole. Following DSB formation, the proportion of cells with either of these three forms of RecA foci was increased, indicating that expression of the RecA protein was elevated via induction of the SOS response. Importantly, it was shown that the proportion of cells with a RecA focus at the *lacZ* locus increased by 7.7-fold in response to DSB formation, suggesting that the palindrome-induced DSBs were repaired by homologous recombination. Moreover, the RecA foci that localized at a cell pole were sequestered from the nucleoid and these foci were inferred to correspond to the accumulation of RecA proteins prior to their degradation.

The RecA focus at the site of the DSB was inferred to represent the RecA nucleoprotein filament that was generated during DSB repair. *In vitro*, the DinI protein have been reported to stabilize the RecA nucleoprotein filament while the RecX and UvrD proteins destabilize the nucleoprotein filament (Cox, 2007). In the present study, it was shown that deletion of either the *recX* or *uvrD* gene resulted in a modest increase in the median duration of a RecA focus at the site of the DSB while deletion of the *dinI* gene caused a modest decrease in the median duration of the RecA focus. These observations demonstrated that the duration of the RecA focus at the site of the DSB was modulated by proteins that have been reported to regulate the stability of the RecA nucleoprotein filament.

Analysis of the subcellular position of the LacI-Cerulean focus showed that the native localization of the *lacZ* locus was between the ¼ position and the mid-cell, with a preference for the mid-cell. In the absence of DSBs, the *lacZ* locus was shown to exhibit dynamic movement around the mid-cell until the onset of segregation of that locus. During DSB repair, the dynamic movement of the *lacZ* locus was constrained to the mid-cell after the formation and disassembly of the RecA focus at the site of the DSB. Consequently, it was inferred that the heteroduplex DNA, or some other feature of the repair complex, that was formed after strand invasion restrained the dynamic movement of the *lacZ* locus. Interestingly, in the

absence of DSBs, cohesion of the newly replicated *lacZ* loci was not sufficient to restrain the dynamic movement of that locus. Further analysis demonstrated that replication of the *lacZ* locus, which was inferred to correspond to DSB formation in the presence of the palindrome and the SbcCD endonuclease, occurred at a similar spatial localization as the RecA focus at the site of the DSB. Because DSB formation and the formation of the RecA focus at the *lacZ* locus preceded the restriction of the dynamic movement of the *lacZ* locus to the mid-cell, it was deduced that DSB formation and the initiation of repair occurred at the native localization of the *lacZ* locus. Data from time-lapse imaging also revealed that repair was usually initiated by 3 minutes after DSB formation.

Chromatin immunoprecipitation, followed by high-throughput DNA sequencing, was used to demonstrate that binding of the TetR-YPet and LacI-Cerulean proteins to the *tetO* and *lacO* arrays, respectively, had no detectable effect on DSB-dependent RecA binding at the *lac* locus and its surrounding loci. The presence of the *lacO* array, but not the *tetO* array, restricted the widespread Chi-dependent binding of RecA at *lac* locus during DSB repair, possibly because of the moderately palindromic nature of the *lacO* sequence. The ChIP-seq data also revealed the presence of DSB-independent RecA binding within the *tetO* and *lacO* array in the absence and presence of the repressor proteins. The profile of the binding of RecA within these operator arrays suggested that the RecA binding was associated with the transcription of the gentamycin and kanamycin resistance genes within the *tetO* and *lacO* arrays, respectively. In the absence of DSB induction, Chi-dependent RecA binding was not detected at the site of the repressor-bound *tetO* and *lacO* arrays. This observation was used to infer that the DSB-dependent RecA enrichment at the *lac* locus was generated by the repair of the palindrome-induced DSBs and not the repair of one-ended DSBs arising from stalled replication forks at the site of the repressor-bound operator arrays.

This study has provided useful insight on the dynamics of repair of a site-specific DSB generated during chromosomal replication in *E. coli*. It has demonstrated that cohesion of sister loci is biologically relevant for eliminating the need for an extensive homology search during repair of DSBs arising from chromosomal replication. This strategy enables a cell to efficiently repair a DSB that is generated during chromosomal replication, thereby minimizing the likelihood of the cell abrogating DSB repair due to inability to identify the undamaged sister template. Moreover, the repair of these DSBs is faster, compared to repair of I-SceI-induced DSBs, and thus, minimizes the effect of the repair of replication-dependent DSBs on cell growth and viability.

7.2 FUTURE WORK

Time-lapse imaging of the RecA protein and the site of the DSB was usually performed at 1-minute intervals and occasionally at 45-seconds intervals. The data obtained from these time-lapse experiments revealed that approximately 50% of the cells showed a RecA focus at the site of the DSB in a cell cycle, and the distribution of the duration of the RecA focus in these cells was left-skewed. This observation might imply that a substantial proportion of the cells undergoing DSB repair had a duration of the RecA focus to be less than 1 minute if the palindrome-induced DSB was generated every replication cycle. Thus, the subpopulation of cells with a very short duration of RecA focus at the site of the DSB was not captured by the time-lapse imaging performed at 1-minute intervals. It would be useful to perform the time-lapse imaging at very short time intervals in order to ascertain if there was a subpopulation of cells which had a duration of RecA focus at the site of the DSB to be less than 1 minute. Moreover, the data obtained from time-lapse imaging at very short time intervals might generate a non-skewed distribution of the duration of the RecA focus at the site of the DSB. In that case, the arithmetic mean could be useful for comparing the duration of these RecA foci in the wild type cells and the deletion mutants that were investigated in chapter 4. A drawback associated with performing time-lapse imaging at very short time intervals is that the imaging cannot be performed for the entire cell cycle due to photo-toxicity and photo-bleaching of the fluorescence foci.

Throughout this study, the fluorescently-labelled RecA protein was detected as a focus or spot at the site of the DSB during live-cell imaging by conventional widefield microscopy. Because conventional fluorescence microscopy is limited to a resolution of 200nm (the diffraction limit of light), it is possible that this study could not detect specific RecA structures at the site of the DSB whose resolution were below 200nm. Hence, super-resolution imaging of the fluorescently-labelled RecA protein and the site of the DSB would allow detection of RecA structures at the site of the DSB, if some structures were formed by the RecA protein during DSB repair.

This study showed that repair of the replication-dependent DSB occurred during the period of cohesion of the sister *lacZ* loci, prior to the onset of segregation of that locus. The effect of DSB repair on the duration of cohesion at the *lacZ* locus can be determined by estimating the duration from initiation of replication (appearance of YPet-DnaN focus) to segregation of the sister *lacZ* loci in the absence and presence of DSBs. Thus, the difference between the observed durations in the absence and presence of DSBs would provide insight on whether DSB repair

increased the duration of cohesion of sister *lacZ* loci relative to cells that are not undergoing DSB repair.

The present study did not investigate the dynamics of the proteins involved in the repair events occurring after homology search and strand invasion reactions. Time-lapse fluorescence imaging of proteins that play a role at the post-synaptic stage of DSB repair would identify whether repair DNA synthesis or resolution of Holliday junctions could be responsible for the prolonged cohesion of sister *lacZ* loci after disassembly of the RecA focus at the site of the DSB. Moreover, the estimation of the duration of cohesion of the sister *lacZ* loci in a *recN* deletion mutant or temperature sensitive mutants of the topoisomerase IV could provide insight on whether the prolonged duration of cohesion is dependent on the RecN or topoisomerase IV proteins instead of reactions occurring at the post-synaptic stage of DSB repair.

Finally, it would be useful to quantify the relative movement of the *lacZ* locus with respect to the mid-cell in the absence and presence of DSB repair. The data obtained can be used for generating a model of the spatial motility of the *lacZ* locus in the absence and presence of DSB repair. Moreover, the diffusion coefficient of the *lacZ* locus can be estimated for each cell in order to numerically compare the relative movement of the *lacZ* locus in the absence and presence of DSB repair.

REFERENCES

- Ali, J. A., Maluf, N. K. and Lohman, T. M. (1999). An oligomeric form of *E. coli* UvrD is required for optimal helicase activity. *J. Mol. Biol.* **293**(4): 815-834.
- Anderson, D. G. and Kowalczykowski, S. C. (1997). The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a χ-regulated manner. *Cell* **90**: 77–86.
- Anderson, D. G. and Kowalczykowski, S. C. (1998). SSB protein controls RecBCD enzyme nuclease activity during unwinding: a new role for looped intermediates. *J. Mol. Biol.* **282**: 275-285.
- Badrinarayanan, A., Le, T. B. K. and Laub, M. T. (2015a). Bacterial chromosome organization and segregation. *Annu. Rev. Cell Dev. Biol.* **31**: 171–199.
- Badrinarayanan, A., Le, T. B. K. and Laub, M. T. (2015b). Rapid pairing and resegregation of distant homologous loci enables double-strand break repair in bacteria. *J. Cell Biol.* **210**: 385–400.
- 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.
- Benedict, R. C. and Kowalczykowski, S. C. (1988). Increase of the DNA strand assimilation activity of RecA protein by removal of the C-terminus and structure-function studies of the resulting protein fragment. *J Biol Chem* **263**: 15513-15520.
- Berkmen, M. B. and Grossman, A. D. (2006). Spatial and temporal organization of the *Bacillus subtilis* replication cycle. *Mol. Microbiol*. **62**: 57–71.
- Bidnenko, V., Lestini, R. and Michel, B. (2006). The *Escherichia coli* UvrD helicase is essential for Tus removal during recombination-dependent replication restart from *Ter* sites. *Mol. Microbiol.* **62**: 382–396.
- Boudsocq, F., Campbell, M., Devoret, R. and Bailone, A. (1997). Quantitation of the inhibition of Hfr × F− recombination by the mutagenesis complex UmuD′C. *J Mol Biol* **270**: 201–211.
- Breier, A. M., Weier, H. -U. G. and Cozzarelli, N. R. (2005). Independence of replisomes in *Escherichia coli* chromosomal replication. *PNAS* **102**: 3942–3947.
- Brendel, V., Brocchieri, L., Sandler, S. J., Clark, A. J. and Karlin, S. (1997). Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms. *J Mol Evol* **44**: 528–541.

- Brendler, T., Sawitzke, J., Sergueev, K. and Austin, S. (2000). A case for sliding SeqA tracts at anchored replication forks during *E. coli* chromosome replication and segregation. *EMBO J.* **19**: 6249–6258.
- Brenner, S. L., Zlotnick, A. and Griffith, J. D. (1988). RecA protein self-assembly multiple discrete aggregation states. *J. Mol. Biol.* **204**: 959-972.
- Brent, R. and Ptashne, M. (1981). Mechanism of action of the *lexA* gene product. *PNAS* **78**: 4204–4208.
- Bruck, I., Woodgate, R., McEntee, K. and Goodman, M. F. (1996). Purification of a soluble UmuD′C complex from *Escherichia coli*: Cooperative binding of UmuD′C to single-stranded DNA. *J Biol Chem* **271**: 10767–10774.
- Burckhardt, S. E., Woodgate, R., Scheuermann, R. H. and Echols, H. (1988). UmuD mutagenesis protein of *Escherichia coli*: overproduction, purification, and cleavage by RecA. *PNAS* **85**: 1811–1815.
- Chen, Z., Yang, H. and Pavletich, N. P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**: 489–484.
- Cockram, C. A., Filatenkova, M., Danos, V., El Karoui, M. and Leach, D. R. (2015). Quantitative genomic analysis of RecA protein binding during DNA double-strand break repair reveals RecBCD action *in vivo*. *PNAS*. **112**(34): 4735-4742.
- Cortes-Ledesma, F. and Aguilera, A. (2006). Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. *EMBO Rep*. **7**: 919–926
- Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. and Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild type and SOS deficient *Escherichia coli*. *Genetics* **158**: 41-64.
- Cox, M. M. (2000). Recombinational DNA repair in bacteria and the RecA protein. In *Progress in Nucleic Acid Research and Molecular Biology*. (*Academic Press*, Inc., San Diego, CA), **63**: 311-366.
- Cox, M. M. (2007). Regulation of bacterial RecA protein function. *Crit Rev Biochem Mol Biol.* **42**: 41-63.
- Cromie, G. A., Connelly, J. C. and Leach, D. R. (2001). Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol. Cell* **8**: 1163–1174.

- Danilowicz, C., Feinstein, E., Conover, A., Coljee, V. W., Vlassakis, J., Chan, Y.-L., Bishop, D. K., and Prentiss, M. (2011). RecA homology search is promoted by mechanical stress along the scanned duplex DNA. *Nucleic Acids Res.* **40**: 1717–1727.
- Darmon, E., Eykelenboom, J. K., Lopez-Vernaza, M. A., White, M. A. and Leach, D. R. (2014). Repair on the go: *E. coli* maintains a high proliferation rate while repairing a chronic DNA double-strand break. *PLoS One* **9**(10): e110784.
- Datta, S., Prabu, M. M., Vaze, M. B., Ganesh, N., Chandra, N. R., *et al*. (2000). Crystal structures of Mycobacterium tuberculosis RecA and its complex with ADP–AlF4: implications for decreased ATPase activity and molecular aggregation. *Nucleic Acids Res*. **28**: 4964–4973.
- De Mot, R., Schoofs, G. and Vanderleyden, J. (1994). A putative regulatory gene downstream of *recA* is conserved in Gram-negative and Gram-positive bacteria. *Nucleic Acids Res.* **22**: 1313–1314.
- de Vries, R. (2010). DNA condensation in bacteria: interplay between macromolecular crowding and nucleoid proteins. *Biochimie* **92**: 1715–1721.
- Di Capua, E., Engel, A., Stasiak, A. and Koller, T. (1982). Characterization of complexes between recA protein and duplex DNA by electron microscopy. *J. Mol. Biol.* **157**: 87–103.
- Dillingham, M. S. and Kowalczykowski, S. C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* **72**(4): 642–671.
- Dillon, S. C. and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Rev. Microbiol.* **8**: 185–195.
- Dixon, D. A. and Kowalczykowski, S. C. (1991). Homologous pairing *in vitro* stimulated by the recombination hotspot, Chi. *Cell* **66**: 361-371.
- Dixon, D. A. and Kowalczykowski, S. C. (1993). The recombination hotspot Chi is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell* **73**(1): 87–96.
- Drees, J. C., Lusetti, S. L., Chitteni-Pattu, S., Inman, R. B. and Cox. M. M. (2004). A RecA filament capping mechanism for RecX protein. *Mol. Cell.* **15**: 789-798.
- Egelman, E. H. and Stasiak, A. (1986). Structure of helical RecA-DNA complexes: Complexes formed in the presence of ATP-gamma-S or ATP. *J. Mol. Biol.* **191**: 677–697.

- Eggler, A. L., Lusetti, S. L. and Cox, M. M. (2003). The C terminus of the *Escherichia coli* RecA protein modulates the DNA binding competition with single-stranded DNA-binding protein. *J Biol Chem.* **278**: 16389–16396.
- Espeli, O. and Marians, K. J. (2004). Untangling intracellular DNA topology. *Mol. Microbiol.* **52**: 925-931.
- Eykelenboom, J. K., Blackwood, J. K., Okely, E. and Leach, D. R. (2008). SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. *Molecular Cell*. **29**(5): 644-651.
- Fernández De Henestrosa, A. R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J. J., Ohmori, H. and Woodgate, R. (2000). Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol. Microbiol.* **35**: 1560-1572.
- Fischer, C. J, Maluf, N. K. and Lohman, T. M. (2004). Mechanism of ATP-dependent translocation of *E.coli* UvrD monomers along single-stranded DNA. *J. Mol. Biol.* **344**: 1287–1309.
- Fisher, J. K., Bourniquel, A., Witz, G., Weiner, B., Prentiss, M. and Kleckner, N. (2013). Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell* **153**: 882-895.
- Flores, M. J., Sanchez, N. and Michel, B. (2005). A fork-clearing role for UvrD. *Mol. Microbiol.* **57**: 1664–1675.
- Forget, A. L. and Kowalczykowski, S. C. (2012). Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature* **482**: 423–427.
- Friedberg, E. C., Walker, G. C. and Siede, W. (1995). DNA repair and mutagenesis. *ASM Press*. Washington, D.C.
- Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A. and Ellenberger, T. (2006). DNA Repair and Mutagenesis. *ASM Press* Washington, DC.
- Fugmann, S. D., Lee, A. I., Shockett, P. E., Villey, I. J. and Schatz, D. G. (2000). The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu. Rev. Immunol.* **18**: 495–527.
- Gerdes, K., Howard, M. and Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell* **141**: 927–942.
- Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* **71**:17–50.

- Goosen, N. and Moolenaar, G. F. (2008). Repair of UV damage in bacteria. *DNA Repair* vol. **7**(3): 353–379.
- Gossen, M. and Bujard, H. (1993). Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells. *Nucleic Acids Res.* **21**: 4411-4412.
- Haber, J. E. (1998). Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu. Rev. Genet*. **32**: 561-99.
- Haber, J. E. (1999). DNA recombination: the replication connection. *Trends Biochem Sci*. **24**(7): 271-275.
- Haber, J. E. (2012). Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* **191**(1): 33-64.
- Handa, N., *et al*. (2012). Molecular determinants responsible for recognition of the single-stranded DNA regulatory sequence, χ, by RecBCD enzyme. *PNAS* **109**(23): 8901–8906
- Huisman, O. and D'Ari, R. (1981). An inducible DNA replication-cell division coupling mechanism in *E. coli*. *Nature*. **290**: 797-799.
- Jensen, R. B., Wang, S. C. and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *EMBO J.* **20**: 4952-4963.
- Joshi, M. C., Bourniquel, A., Fisher, J., Ho, B. T., Magnan, D., Kleckner, N. and Bates, D. (2011). *Escherichia coli* sister chromosome separation includes an abrupt global transition with concomitant release of late-splitting intersister snaps. *PNAS* **108**:2765–2770.
- Joshi, M. C., Magnan, D., Montminy, T. P., Lies, M., Stepankiw, N. and Bates, D. (2013). Regulation of sister chromosome cohesion by the replication fork tracking protein SeqA. *PLoS Genet.* **9**:e1003673.
- Khlebnikov, A., Datsenko, K. A., Skaug, T., Wanner, B. L. and Keasling, J. D. (2001). Homogeneous expression of the P_{BAD} promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiol.* **147**: 3241-3247.
- Kidane, D. and Graumann, P. L. (2005). Dynamic formation of RecA filaments at DNA double strand break repair centres in live cells. *J. Cell. Biol.* **170**: 357-366.

- Kleckner, N., Fisher, J. K., Stouf, M., White, M. A., Bates, D. and Witz, G. (2014). The bacterial nucleoid: nature, dynamics and sister segregation. *Curr. Opin. Microbiol.* **22**: 127–37.
- Koch, W. H. and Woodgate, R. (1998). The SOS response. In DNA Damage and Repair: DNA Repair in Prokaryotes and Lower Eukaryotes. Nickoloff, J. A. and Hoekstra, M. F. (eds). Totowa, NJ: *Humana Press*, pp. 107 – 134.
- Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* **61**: 212-38.
- Koppes, L., Woldringh, C. L. and Nanninga, N. (1999). *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* **81**: 803–810.
- Korsmeyer, S. J. (1992). Chromosomal translocations in lymphoid malignancies reveal novel proto-oncogenes. *Annu. Rev. Immunol.* **10**: 785–807.
- Kowalczykowski, S. C., Clow, J. C., Somani, R. and Varghese, A. (1987). Effects of the *Escherichia coli* SSB protein on the binding of *Escherichia coli* RecA protein to single-stranded DNA: demonstration of competitive binding and the lack of a specific protein-protein interaction. *J. Mol. Biol.* **193**: 81-95.
- Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. and Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**: 401-465
- Kowalczykowski, S. C. and Eggleston, A. K. (1994). Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem*. **63**: 991–1043.
- Kowalczykowski, S. C. and Krupp, R. A. (1987). Effects of the *Escherichia coli* SSB protein on the single-stranded DNA-dependent ATPase activity of *Escherichia coli* RecA protein: evidence that SSB protein facilitates the binding of RecA protein to regions of secondary structure within single-stranded DNA. *J. Mol. Biol.* **193**: 97 - 113.
- Kowalczykowski, S. C. and Krupp, R. A. (1995). DNA-strand exchange promoted by RecA protein in the absence of ATP: implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions. *PNAS* **92**: 3478–3482.
- Kurth, I. and O'Donnell, M. (2013). New insights into replisome fluidity during chromosome replication. *Trends Biochem. Sci.* **38**: 195-203.
- Kuzminov, A. (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* 63: 751–813.

- Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. *PNAS* **98**(15): 8241-8246.
- Lark, K. G. (1972). Evidence for the direct involvement of RNA in the initiation of DNA replication in *Escherichia coli* 15T⁻. *J Mol Biol*. **64**(1): 47–60.
- Lark, K. G., Repko, T. and Hoffman, E. J. (1963). The effect of amino acid deprivation on subsequent deoxyribonucleic acid replication. *Biochim Biophys Acta*. **76**: 9–24.
- Lau, I. F., Filipe, S. R., Soballe, B., Okstad, O. A., Barre, F. X. and Sherratt, D. J. (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* **49**: 731–743.
- Lavery, P. E. and Kowalczykowski, S. C. (1992). A postsynaptic role for single-stranded DNA-binding protein in recA protein-promoted DNA strand exchange. *J. Biol. Chem.* **267**: 9315-9320.
- Leach, D. R. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **16**: 893–900.
- Le Gall, A., Cattoni, D. I., Guilhas, B., Mathieu-Demaziere, C., Oudjedi, L., Fiche, J.-B., Rech, J., Abrahamsson, S., Murray, H., Bouet, J.-Y., Nollmann, M. (2016). Bacterial partition complexes segregate within the volume of the nucleoid. *Nat. Commun.* **7**: 12107.
- Lemaître, C. and Soutoglou, E. (2014). Double strand break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair. *DNA Repair (Amst)* **19**: 163-168.
- Lemon, K. P. and Grossman, A. D. (1998). Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* **282**:1516–1519.
- Lesterlin, C., Ball, G., Schermelleh, L. and Sherratt, D. J. (2014). RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature*. **506**(7487): 249-253.
- Lesterlin, C., Gigant, E., Boccard, F. and Espeli, O. (2012). Sister chromatid interactions in bacteria revealed by a site-specific recombination assay. *EMBO J.* **31**: 3468–3479.
- Lewis, L. K., Harlow, G. R., Gregg-Jolly, L. A. and Mount, D. W. (1994). Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in *Escherichia coli*. *J. Mol. Biol.* **241**: 507–523.

- Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem*. **79**: 181–211.
- Link, A. Phillips, J. D. and Church, G. M. (1997). Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J. Bacteriol.* **179**: 6228–6237.
- Lisby, M., Mortensen, U. H. and Rothstein, R. (2003). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat. Cell Biol*. **5**: 572–577.
- Lisby, M., Rothstein, R. and Mortensen, U. H. (2001). Rad52 forms DNA repair and recombination centers during S phase. *PNAS* **98**: 8276-8282.
- Little, J. W. (1984). Autodigestion of lexA and phage lambda repressors. *PNAS* **81**: 1375–1379.
- Little, J. W., Mount, D. W. and Yanisch-Perron, C. R. (1981). Purified lexA protein is a repressor of the *recA* and *lexA* genes. *PNAS* **78**: 4199–4203.
- Llorente, B., Smith, C. E. and Symington, L. S. (2008). Break-induced replication: what is it and what is it for? *Cell Cycle* **7**: 859-864.
- Lusetti, S. L. and Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. In *Annual Review of Biochemistry* (Charles Richardson, ed.) (*Annual Reviews*, Palo Alto, CA), **71**: 71-100.
- Lusetti, S. L., Drees, J. C. and Cox. M. M. (2004). The DinI and RecX proteins are competing modulators of RecA function. *J. Biol. Chem*. **279**: 55073 - 55079.
- Lusetti, S. L., Shaw, J. J. and Cox, M. M. (2003). Magnesium ion-dependent activation of the RecA protein involves the C terminus. *J Biol Chem.* **278**:16381–16388.
- Lusetti, S. L., Voloshin, O. N., Inman, R. B., Camerini-Otero, R. D. and Cox. M. M. (2004). The DinI protein stabilizes RecA protein filaments. *J. Biol. Chem.* **279**: 30037-30046.
- Maaløe, O. and Hanawalt, P. C. (1961). Thymine deficiency and the normal DNA replication cycle. I. *J. Mol. Biol.* **3**: 144–155.
- Madiraju, M. V., Templin, A. and Clark, A. J. (1988). Properties of a mutant recA-encoded protein reveal a possible role for *Escherichia coli* recF-encoded protein in genetic recombination. *PNAS* **85**: 6592-6596.
- Malkov, V. A. and Camerini-Otero, R. D. (1995). Photocross-links between single-stranded DNA and *Escherichia coli* RecA protein map to loops LI (amino acid residues 157-164) and L2 (amino acid residues 195-209). *J. Biol. Chem*. **270**: 30230–30233.

- Maluf, N. K., Fischer, C. J. and Lohman, T. M. (2003). A dimer of *Escherichia coli* UvrD is the active form of the helicase *in vitro*. *J. Mol. Biol.* **325**(5): 913-935.
- Matson, S. W. (1991). DNA helicases of *Escherichia coli*. *Prog. Nucleic Acids Res. Mol. Biol.* **40**: 289–326.
- Matson, S. W. and George, J. W. (1987). DNA helicase II of *Escherichia coli*. Characterization of the single-stranded DNA-dependent NTPase and helicase activities. *J. Biol. Chem.* **262**: 2066–2076.
- Mawer, J. S. and Leach, D. R. (2014). Branch migration prevents DNA loss during double-strand break repair. *PLoS Genet.* **10**(8): e1004485.
- McGill, C., Shafer, B. and Strathern, J. (1989). Co-conversion of flanking sequences with homothallic switching. *Cell* **57**: 459–467.
- Menetski, J. P. and Kowalczykowski, S. C. (1985). Interaction of recA protein with single-stranded DNA. Quantitative aspects of binding affinity modulation by nucleotide cofactors. *J. Mol. Biol.* **181**: 281–295.
- Mercier, R., Petit, M. A., Schbath, S., Robin, S., El Karoui, M., Boccard, F. and Espeli, O. (2008). The MatP/matS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135**: 475–485.
- Merlin, C., McAteer, S. and Masters, M. (2002). Tools for characterization of *Escherichia coli* genes of unknown function. *J. Bacteriol*. 184: 4573–4581.
- Messer, W. (1972). Initiation of deoxyribonucleic acid replication in *Escherichia coli* B-r: chronology of events and transcriptional control of initiation. *J. Bacteriol*. **112**(1): 7–12.
- Messer, W. and Weigel, C. (1996). Initiation of chromosome replication. In *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. Neidhardt, F. C., Curtiss, R. I., Ingraham, J., Lin, E. C. C., Low, K. B., Magasanik, B., *et al.* (eds). Washington, DC: *American Society for Biology Press* **2**: 1579–1601.
- Michel, B. (2005). After 30 years of study, the bacterial SOS response still surprises us. *PLOS Biology* **3**(7): e255.
- Michel, B., Grompone, G., Flores, M. J. and Bidnenko, V. (2004). Multiple pathways process stalled replication forks. *PNAS* **101**: 12783–12788.
- Michel, B., Flores, M. J., Viguera, E., Grompone, G., Seigneur, M., and Bidnenko, V. (2001). Rescue of arrested replication forks by homologous recombination. *PNAS* **98**: 8181– 8188.

- Michel, B. and Leach, D. (2012). Chapter 7.2.7. Homologous recombination—enzymes and pathways. In Curtiss R III, Kaper J B, Squires C L, Karp P D, Neidhardt F C, Slauch J M (ed). EcoSal—*Escherichia coli* and *Salmonella*: cellular and molecular biology. *ASM Press*, Washington, D C.
- Migocki, M. D., Lewis, P. J., Wake, R. G. and Harry, E. J. (2004). The midcell replication factory in *Bacillus subtilis* is highly mobile: implications for coordinating chromosome replication with other cell cycle events. *Mol. Microbiol*. **54**: 452–463.
- Miné-Hattab, J. and Rothstein, R. (2012). Increased chromosome mobility facilitates homology search during recombination. *Nature Cell Biol*. **14**: 510-517.
- Molina, F. and Skarstad, K. (2004). Replication fork and SeqA focus distributions in *Escherichia coli* suggest a replication hyperstructure dependent on nucleotide metabolism. *Mol. Microbiol*. **52**: 1597-1612.
- Moreau, P. L. (1988). Overproduction of single-stranded-DNA-binding protein specifically inhibits recombination of UV-irradiated bacteriophage DNA in *Escherichia coli*. *J. Bacteriol*. **170**: 2493-2500.
- Morel, P., Hejna, J. A., Ehrlich, S. D. and Cassuto, E. (1993). Antipairing and strand transferase activities of *E. coli* helicase II (UvrD). *Nucleic Acids Res.* **21**: 3205–3209.
- Morimatsu, K. and Kowalczykowski, S. C. (2003). RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: A universal step of recombinational repair. *Mol. Cell.* **11**: 1337–1347.
- Morrow, D. M., Connelly, C and Hieter, P. (1997). Break copy duplication: a model for chromosome fragment formation in *Saccharomyces cerevisiae*. *Genetics* **147**: 371-382.
- Nambiar, M., Kari, V. and Raghavan, S. C. (2008). Chromosomal translocations in cancer. *Biochim Biophys Acta* **1786**: 139–152.
- Nassif, N., Penney, J., Pal, S., Engels, W. R. and Gloor, G. B. (1994). Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.* **14**: 1613–1625.
- Nielsen, H. J., Li, Y., Youngren, B., Hansen, F. G. and Austin, S. (2006). Progressive segregation of the *Escherichia coli* chromosome. *Mol Microbiol* **61**: 383–393.
- Nielsen, H. J., Youngren, B., Hansen, F. G. and Austin, S. (2007) Dynamics of *Escherichia coli* chromosome segregation during multifork replication. *J. Bacteriol.* **189**: 8660–8666.

- Niki H., Yamanichi, Y. and Hiraga, S. (2000). Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev.* **14**: 212–223.
- Ogawa, T., Shinohara, A., Nabetani, A., *et al*. (1993). RecA-like recombination proteins in eukaryotes: functions and structures of RAD51 genes. *Cold Spring Harbor Sympos Quant Biol* **68**: 567–57.
- Pages, V., Koffel-Schwartz, N. and Fuchs, R. P. (2003). *recX*, a new SOS gene that is co-transcribed with the *recA* gene in *Escherichia coli*. *DNA Repair* **2**: 273-284.
- Paintdakhi, A. *et al*. (2016). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol. Microbiol.* **99**: 767–777.
- Patel, M., Jiang, Q., Woodgate, R., Cox, M. M. and Goodman, M. F. (2010). A new model for SOS-induced mutagenesis: How RecA protein activates DNA polymerase V. *Crit Rev Biochem Mol Biol.* **45**: 171–184.
- Pennington, J. M. and Rosenberg, S. M. (2007). Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat. Genet*. **39**(6):797-802.
- Pfeiffer, P., Goedecke, W. and Obe, G. (2000). Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis* **15**: 289–302.
- Possoz, C., Filipe, S. R., Grainge, I. and Sherratt, D. J. (2006). Tracking of controlled *Escherichia coli* replication fork stalling and restart at repressor-bound DNA *in vivo*. *EMBO J*. **25**: 2596–2604.
- Pugh, B. F. and Cox, M. M. (1987). Stable binding of recA protein to duplex DNA: Unraveling a paradox. *J. Biol. Chem.* **262**: 1326-1336.
- Pugh, B. F. and Cox, M. M. (1988). General mechanism for RecA protein binding to duplex DNA. *J. Mol. Biol.* **203**: 479-493.
- Ragunathan, K., Liu, C. and Ha, T. (2012). RecA filament sliding on DNA facilitates homology search. *eLife* **1**: e00067.
- Register III, J. C. and Griffith, J. (1985). The direction of RecA protein assembly onto single strand DNA is the same as the direction of strand assimilation during strand exchange. *J. Biol. Chem*. **260**: 12308–12312
- Rehrauer, W. M. and Kowalczykowski, S. C. (1993). Alteration of the nucleoside triphosphate (NTP) catalytic domain within *Escherichia coli* recA protein attenuates NTP hydrolysis but not joint molecule formation. *J. Biol. Chem.* **268**: 1292–1297.

- Renzette, N., Gumlaw, N., Nordman, J. T., Krieger, M., Yeh, S.-P., Long, E., Centore, R., Boonsombat, R. and Sandler, S. J. (2005). Localization of RecA in *Escherichia coli* K-12 using RecA-GFP. *Mol. Microbiol.* **57**(4): 1074-85.
- Renzette, N. and Sandler, S. J. (2008). Requirements for ATP binding and hydrolysis in RecA function in *Escherichia coli*. *Mol Microbiol* **67**: 1347–1359.
- Reyes-Lamothe, R., Possoz, C., Danilova, O. and Sherratt, D. J. (2008). Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* **133**: 90–102.
- Reyes-Lamothe, R., Sherratt, D. J. and Leake, M. C. (2010). Stoichiometry and Architecture of Active DNA Replication Machinery in *Escherichia coli*. *Science* **328**: 498–501.
- Roca, A. I. and Cox, M. M. (1990). The RecA protein: Structure and function. *Crit. Rev. Biochem. Mol. Biol.* **25**: 415-456.
- Roca, A. L. and Cox, M. M. (1997). RecA protein: Structure, function, and role in recombinational DNA repair. In *Progress in Nucleic Acid Research and Molecular Biology*. W. E. Cohn and K. Moldave, eds. (Academic Press, Inc., San Diego, CA), **56**: 129-223.
- Rothkamm, K., Kruger, I., Thompson, L. H. and Lobrich, M. (2003). Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell Biol.* **23**: 5706–5715.
- Runyon, G. T., Bear, D. G. and Lohman, T. M. (1990). *Escherichia coli* helicase II (UvrD) protein initiates DNA unwinding at nicks and blunt ends. *PNAS* **87**: 6383–6387.
- Sano, Y. (1993). Role of the *recA*-related gene adjacent to the *recA* gene in *Pseudomonas aeruginosa*. *J. Bacteriol.* **175**: 2451–2454.
- Schaechter, M., Maaøe, O. and Kjeldgaard, N. O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol*. **19**: 592–606.
- Shan, Q., Bork, J. M., Webb, B. L., Inman, R. B. and Cox, M. M. (1997). RecA protein filaments: End-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. *J. Mol. Biol.* **265**: 519-540.
- Shan, Q., Cox, M. M. and Inman, R. B. (1996). DNA strand exchange promoted by RecA K72R. Two reaction phases with different Mg^{2+} requirements. *J. Biol. Chem.* **271**: 5712–5724.

- Sjoegren, C. and Nasmyth, K. (2001). Sister chromatid cohesion is required for post-replicative double strand break repair in *Saccharomyces cerevisiae*. *Curr. Biol*. **11**: 991–995.
- Smith, G. R. (2012). How RecBCD enzyme and Chi promote DNA break repair and recombination: A molecular biologist's view. *Microbiol. Mol. Biol. Rev.* **76**(2): 217–228.
- Sommer, S., Bailone, A. and Devoret, R. (1993). The appearance of the UmuD′C protein complex in *Escherichia coli* switches repair from homologous recombination to SOS mutagenesis. *Mol. Microbiol.* **10**: 963–971.
- Sommer, S., Boudsocq, F., Devoret, R. and Bailone, A. (1998). Specific RecA amino acid changes affect RecA-UmuD′C interaction. *Mol Microbiol*. **28**: 281–91.
- Soutoglou, E., Dorn, J. F., Sengupta, K., Jasin, M., Nussenzweig, A., Ried, T., Danuser, G. and Misteli, T. (2007). Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* **9**: 675-682.
- Stasiak, A and Di Capua, E. (1982). The helicity of DNA in complexes with RecA protein. *Nature* **299**: 185–186.
- Stasiak, A., Di Capua, E. and Koller, T. (1981). Elongation of duplex DNA by recA protein. *J. Mol. Biol.* **151**: 557−564.
- Stohl, E. A., Brockman, J. P., Burkle, K. L., Morimatsu, K., Kowalczykowski, S.C. and Siefert, H. S. (2003). *Escherichia coli* RecX inhibits RecA recombinase and coprotease activities *in vitro* and *in vivo*. *J. Biol. Chem.* **278**: 2278-2285.
- Story, R. M. and Steitz, T. A. (1992). Structure of the recA protein-ADP complex. *Nature* **355**: 374–376.
- Story, R. M., Weber, I. T. and Steitz, T. A. (1992). The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**: 318–325.
- Sun, H., Treco, D., Schultes, N. P. and Szostak, J. W. (1989). Double-strand breaks at an initiation site for meiotic gene conversion. *Nature* **338**: 87–90.
- Sung, P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265**: 1241–1243.
- Szostak, J. W., Orr Weaver, T. L., Rothstein, R. J. and Stahl, F. W. (1983). The double-strand-break repair model for recombination. **Cell 33**: 25-35.

- Takashima, Y., Sakuraba, M., Koizumi, T., Sakamoto, H., Hayashi, M. and Honma, M. (2009). Dependence of DNA double strand break repair pathways on cell cycle phase in human lymphoblastoid cells. *Environ. Mol. Mutagen.* **50**: 815–822.
- Tang, M., Shen, X., Frank, E. G., O'Donnell, M., Woodgate, R. and Goodman, M. F. (1999). UmuD′2C is an error-prone DNA polymerase, *Escherichia coli*, DNA pol V. *PNAS* **96**: 8919–8924.
- Tateishi, S., Horii, T., Ogawa, T. and Ogawa, H. (1992). C-terminal truncated *Escherichia coli* RecA protein RecA5327 has enhanced binding affinities to single- and double-stranded DNAs. *J Mol Biol* **223**: 115-129.
- Taylor, A. F. and Smith, G. R. (1999). Regulation of homologous recombination: Chi inactivates RecBCD enzyme by disassembly of the three subunits. *Genes Dev.* **13**(7): 890–900.
- Trun, N. J. and Marko, J. F. (1998). Architecture of a bacterial chromosome. *ASM News* **64**: 276–283.
- Tsang, S. S., Muniyappa, K., Azhderian, E., Gonda, D. K., Radding, C. M., Flory, J. and Chase, J. W. (1985). Intermediates in homologous pairing promoted by recA protein. Isolation and characterization of active presynaptic complexes. *J Mol Biol.* **185**(2): 295–309.
- Umezu, K., Chi, N. W. and Kolodner, R. D. (1993). Biochemical interaction of the *Escherichia coli* RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *PNAS* **90**: 3875–3879.
- Umezu, K. and Kolodner, R. D. (1994). Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J. Biol. Chem.* **269**: 30005–30013.
- Valens, M., Penaud, S., Rossignol, M., Cornet, F. and Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* **23**: 4330–4341.
- van den Bosch, M., Lohman, P. H. M. and Pastink, A. (2002). DNA double-strand break repair by homologous recombination. *Biol. Chem*. **383**: 873–892
- Veaute, X., Delmas, P., Selva, M., Jeusset, J., Le Cam, E., Matic, I., Fabre, F. and Petit, M. A. (2005). UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in *Escherichia coli*. *EMBO J*. **24**: 180–189.

- Vilenchik, M. M. and Knudson, A. G. (2003). Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *PNAS* **100**: 12871–12876.
- Voloshin, O. N., Ramirez, B. E., Bax, A. and Camerini-Otero, R. D. (2001). A model for the abrogation of the SOS response by an SOS protein: a negatively charged helix in DinI mimics DNA in its interaction with RecA. *Genes Dev.* **15**: 415–427.
- von Meyenburg, K., Hansen, F. G., Riise, E., Bergmans, H. E., Meijer, M. and Messer, W. (1979). Origin of replication, *oriC*, of the *Escherichia coli* K12 chromosome: genetic mapping and minichromosome replication. *Cold Spring Harb Symp Quant Biol.* **43**: 121–128.
- Wang, X., Lesterlin, C., Reyes-Lamothe, R., Ball, G. and Sherratt, D. J. (2011). Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *PNAS* **108**: 243–50.
- Wang, X., Liu, X., Possoz, C. and Sherratt, D. J. (2006). The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes Dev.* **20**: 1727-1731.
- Wang, X., Possoz, C. and Sherratt, D. J. (2005). Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes Dev*. **19**: 2367–2377.
- Wang, X., Reyes-Lamothe, R. and Sherratt, D. J. (2008). Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. *Genes Dev* 22: 2426–2433.
- Wehrli, W. and Staehelin, M. (1971). Actions of the rifamycins. *Bacteriol Rev*. **35**(3): 290–309.
- Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., di d'Adda Fagagna F., *et al*. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**: 1686–1689.
- Whitby, M. C. and Lloyd, R. G. (1995). Altered SOS induction associated with mutations in recF, recO and recR. *Mol. General Genet.* **246**: 174-179.
- White, M. A., Eykelenboom, J. K., Lopez-Vernaza, M. A., Wilson, E. and Leach, D. R. (2008). Non-Random Segregation of Sister Chromosomes in *Escherichia coli*. *Nature*. **455**(7217): 1248-1250.
- Woodgate, R., Rajagopalan, M., Lu, C. and Echols, H. (1989). UmuC mutagenesis protein of *Escherichia coli*: Purification and interaction with UmuD and UmuD′. *PNAS* **86**: 7301–7305.
- Wyman, C., Ristic, D. and Kanaar, R. (2004). Homologous recombination-mediated double-strand break repair. *DNA Repair* **3**(8-9): 827-833.

- Yasuda, T., Morimatsu, K., Horii, T., Nagata, T. and Ohmori, H. (1998). Inhibition of *Escherichia coli* RecA coprotease activities by DinI. *EMBO J.* **17**: 3207.
- Yasuda, T., Morimatsu, K., Kato, R., Usukura, J., Takahashi, M., and Ohmori, H. (2001). Physical interactions between DinI and RecA nucleoprotein filament for the regulation of SOS mutagenesis. *EMBO J* **20**(5): 1192–1202.
- Yoshimasu, M., Aihara, H., Ito, Y., Rajesh, S., Ishibe, S., Mikawa, T., Yokoyama, S. and Shibata, T. (2003). An NMR study on the interaction of *Escherichia coli* DinI with RecA-ssDNA complexes. *Nucleic Acids Research* **31**: 1735-1743.
- Zahra, R., Blackwood, J. K., Sales, J. and Leach, D. R. (2007). Proofreading and secondary structure processing determine the orientation dependence of CAG.CTG trinucleotide repeat instability in *Escherichia coli*. *Genetics* **176**(1): 27-41.