THE UNIVERSITY *of* EDINBURGH

Intermediates of DNA Double Strand Break Repair in *Escherichia coli*

Julia S. P. Mawer

Thesis presented for the degree of Doctor of Philosophy

Institute of Cell Biology

The University of Edinburgh

September 2011

Declaration

I hereby declare that this thesis was composed by me, and the research presented is my own, except where otherwise stated. This work has not been submitted for any other degree or professional qualification.

Julia S. P. Mawer

September 2011

Acknowledgements

I would like to thank David Leach for giving me the opportunity to carry out the research for my Ph.D. in his lab and under his supervision. His guidance and support were indispensable. Working in the Leach lab has been great fun thanks to all the wonderful members, both past and present. A special thanks goes out to Elise Darmon who had to put up with various revised and very long drafts of this thesis, which she very kindly reviewed.

I would also like to thank the whole of the 4th floor of the Darwin building. When the Interthal and the Makovets labs joined the floor they helped to create a great research environment. I would like to specifically thank Svetlana Makovets whose input in my research was fundamental for my Ph.D.

My time in Edinburgh would not have been the same without the amazing friends I made. A big hug goes out to Roberta Carloni and her associated "Italian mob" who helped me feel at home by bringing a little bit of Italy into everyday life.

The support of my amazing boyfriend, Rodrigo, was also indispensable. Sometimes I really think I would not have made it this far without him...despite the fact that he disappeared on holiday to Mexico two weeks before my submission date.

A big hug goes out to Simon, Connie, and Matthew Mawer, as without their support and love I would not be where I am today. Thank you!

Finally, I would like to thank the MRC who granted me the scholarship that allowed me to study for my Ph.D. at the University of Edinburgh.

Abstract

A DNA double-strand break (DSB) is a severe form of DNA damage. In fast-growing cells, DSBs are commonly repaired by homologous recombination (HR) and in *E. coli* they are exclusively repaired by this mechanism. Failure to accurately repair DSBs can lead to genomic instability. Characterising the DNA intermediates formed during DSB repair by HR is key to understanding this process. A system for inducing a site-specific DSB in the *E. coli* chromosome has previously been described (Eykelenboom et al., 2008). Here, this system has been used to determine the nature of the intermediates of the repair. It was shown that in a Rec^+ background the repair process is rapid and efficient. By contrast, in a *ruvAB* mutant, which is defective for the Holliday junction (HJ) migration and cleavage complex, RuvABC, HJs are accumulated on both sides of the breakpoint. Replication forks also accumulate at defined positions from the DSB, indicating that unresolved HJs are a barrier to efficient replication that is associated with the repair. This suggests that the resolution of HJs needs to occur prior to the establishment of DNA synthesis. Despite the accumulation of HJs in a *ruvAB* mutant, cell survival occurs when DSBs are induced for short periods, suggesting that HJs can be resolved in a RuvAB-independent manner. In contrast, the RecG helicase is essential for survival. In a *recG* mutant, replication forks but not HJs are detected in the region of DSB repair. In a *ruvAB recG* mutant, intermediates in this region are lost. These observations are consistent with a role of RecG in the stabilisation and maturation of D-loops and not the resolution of Holliday junctions. Nevertheless, an additional role for RecG in later stages of repair cannot yet be excluded. This work provides a solid framework for the further study of DSB repair in *E. coli*.

Abbreviations

ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BIR	Break-induced replication
bp	Base-pair
CFU	Colony forming unit
Cm	Centimetre
Cm^R	Chloramphenicol resistance cassette
°C	Degrees Celcius
DIG	Digoxygenin
DNA	Deoxyribonucleic acid
DSB(s)	Double-strand break(s)
DSBR	Double-strand break repair
dsDNA	Double-stranded DNA
g	Gram
HJ	Holliday junction
HR	Homologous recombination
JM	Joint molecule
Kb	Kilobase-pair
l	Litre
lacZ::246	Allele of *lacZ* containing the 246 bp palindrome
*lacZ*χ-	Allele of *lacZ* with deleted endogenous χ site
M	Molar
MCS	Multiple cloning site
µg	Microgram
µl	Microlitre
µm	Micrometre
µM	Micromolar
mg	Milligram
ml	Millilitre
mM	Millimolar

MRN	Mre11/Rad50/Nbs1
MRX	Mre11/Rad50/Xrs2
ng	Nanogram
NHEJ	Non-homologous end joining
nm	Nanometre
OD_{600nm}	Optical density at 600 nm
PCR	Polymerase chain reaction
PFGE	Pulsed-field gel electrophoresis
pH	Power of hydrogen
PMGR	Plasmid-mediated gene replacement
RFR	Replication fork reversal
ROS	Reactive oxygen species
SDSA	Synthesis dependent strand annealing
SEM	Standard error of the mean
ssDNA	Single-stranded DNA
SSG	(DNA) Single-strand gap
Tm	Melting temperature
UV light	Ultra-violet light
v/v	Volume per unit volume
w/v	Weight per unit volume

Table of Contents

Table of Figures

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Table of Tables

Chapter 2:

Chapter 4:

Chapter 1

Introduction

1.1 DNA damage

During the life of a cell the DNA is subject to damage. This may be endogenous, caused by cellular housekeeping activities such as replication or transcription, or exogenous, caused by extra-cellular agents such as ultra-violet light (UV light) or ionising radiation (IR). Left unrepaired, DNA damage could lead to mutation, which could compromise the health of the cell and lead to cell death.

There are many different forms of DNA damage and it is estimated that each human cell incurs between 1,000-1,000,000 DNA lesions per day (Lodish, 2004). One of the most common forms of damage is the chemical alteration of the DNA bases. Causes include hydrolysis of bases, which can deaminate, depurinate, or depyrimidinate the DNA, and alkylation of bases (Demple and Harrison, 1994; Norbury and Hickson, 2001; Pegg, 2000; Pegg and Byers, 1992). These changes can generate mismatches in the DNA sequence and lead to mutation. Chemical alteration of bases alone is not the sole cause of mutations. Replication itself can spontaneously introduce a wrong base, skip a base, or add an extra base in the DNA sequence (Friedberg, 2003). These mutations can alter the coding sequence of a

protein potentially rendering it non-functional. If the gene affected is essential, this will lead to cell death. In multi-cellular organisms, mutations in tumour suppressor genes, oncogenes or genes involved in a DNA repair pathway, can lead to the development of cancer.

Other types of chemical alteration of bases can distort the structure of the DNA, affecting its replication. One example is inter-strand crosslinks, which results from the formation of a chemical bond between bases from complementary DNA strands (Deans and West). Other distortions are caused by polycyclic aromatic hydrocarbons. These bulky, complex, aromatic molecules are found as environmental pollutants that are generated by combustion of a variety of substances such as of fossil fuels, foods, and tobacco in cigarettes (Farmer et al., 2003). The persistence of such lesions can cause cellular replication to arrest. Similar stalling of the replication machinery can be induced by proteins bound to the chromosome (Rothstein et al., 2000). A variety of known human disease conditions are caused by distortions in the DNA template that stall the replication fork (de Boer and Hoeijmakers, 2000).

As well as damage to the DNA's bases, the chemistry of the backbone can also be altered. IR, such as radiation from cosmic rays, X-rays, or radioactive elements, can physically break covalent bonds in the DNA backbone. Depending on the intensity of the ionizing radiation, this can lead to DNA single-strand breaks (SSBs), or double strand breaks (DSBs). DSBs are amongst the most severe forms of damage as they result in chromosomes being broken into two or more pieces. If the chromosome is not pieced back together again, segments that do not contain an origin of replication get degraded and all the genetic information within becomes

lost. Conversely, if a broken chromosome fragment is not pieced back onto a fragment to which it was originally joined to, gross chromosomal rearrangements can occur. This can cause chromosomal translocations, which in eukaryotes can lead to aneuploidy. Both these phenomena have been implicated in a variety of cancers (Pellman, 2007).

1.2 Causes of DNA Double Strand Breaks

As mentioned above, DSBs can be generated by exposure to IR, although they are more commonly generated by processes endogenous to the cell. Cellular metabolism, including DNA replication and repair, frequently results in the formation of SSBs and single-strand gaps (Demple and Harrison, 1994; Lindahl, 1993; Truglio et al., 2006; Wallace, 1998). If these persist and are encountered by a replication fork they cause the fork to collapse, leaving a partially replicated chromosome terminating in a one-ended DSB (Rothstein et al., 2000) (Fig. 1.1). These are known as endogenous DSBs and arise spontaneously in all organisms, from prokaryotes to eukaryotes (Kuzminov, 2001; Lisby et al., 2001; Zou and Rothstein, 1997). There is striking evidence suggesting that up to 1% of SSBs in humans are converted to DSBs, resulting in about 50 DSBs in every cell, each cell cycle (Vilenchik and Knudson, 2003). A blocked replication fork can also give rise to

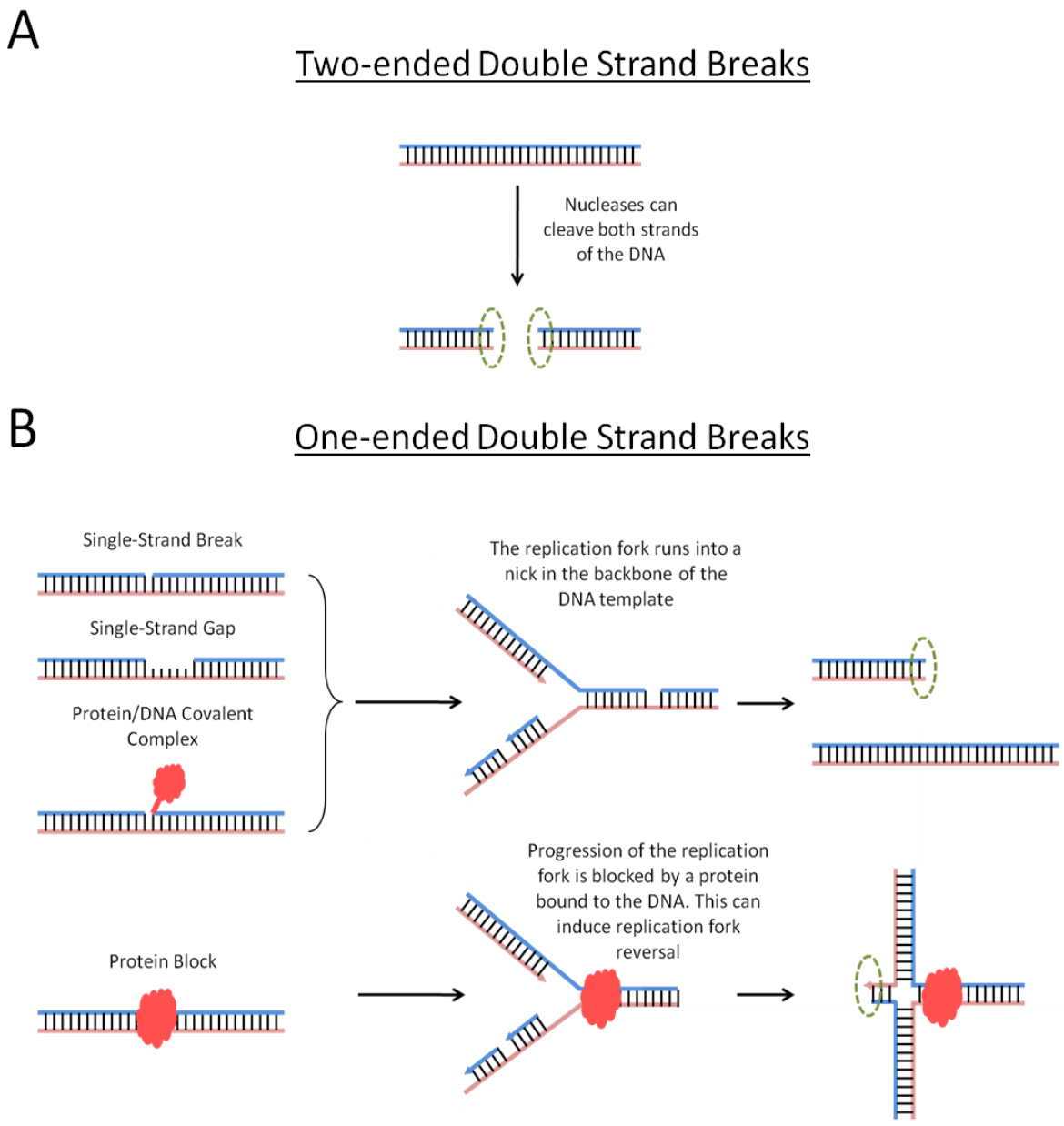

Figure 1. 1 Causes of endogenous DNA double strand breaks

(A) Formation of two-ended DNA double strand breaks are specifically generated by processes such as meiotic recombination in diploid organisms and V(D)J recombination in the vertebrate immune system, which induce the action of endonucleases that target and cleave both strands of the DNA. (B) Formation of one-ended DNA double strand breaks can be caused by a variety of mechanisms. Single-strand breaks, single-strand gaps, and proteins covalently bound to the DNA, generate one-ended DNA double strand breaks when replicated. Additionally, a protein bound to the DNA duplex can induce replication fork reversal. Double strand breaks are indicated by the green dotted circle.

endogenous one-ended DSBs. This can happen when a replication fork encounters a physical block on the DNA template, such as a protein. In this instance an event called replication fork reversal (RFR) may take place (Michel et al., 1997; Pommier et al., 2003). DSBs are also specifically induced by cellular endonucleases during processes such as meiosis in diploid organisms, mating-type switching in the yeast *Saccharomyces cerevisiae* and V(D)J recombination (variable, diverse, and joining gene segments) in antibody production of the vertebrate immune system (Liu et al., 1995; Strathern et al., 1982; Weterings and Chen, 2008). These processes result in the formation of two-ended DSBs. Whatever the nature surrounding DSB formation, these need to be faithfully repaired to ensure survival of the organism.

1.3 Homologous recombination in *E. coli*

In yeasts and higher eukaryotes DNA DSB repair can occur by either non-homologous end joining (NHEJ) or homologous recombination (HR) (van Gent et al., 2001; Weterings and Chen, 2008). In fact, these processes may be coordinated to optimise repair (Takata et al., 1998). Prokaryotes rely mainly on HR for DSB repair, although evidence of repair pathways that require prokaryotic homologues of proteins present in NHEJ have been found in some bacterial species (Gupta et al., 2011; Kobayashi et al., 2008; Stephanou et al., 2007). Nevertheless, no evidence of an NHEJ-like pathway has been found in the enterobacteria, which includes the *Escherichia* genus. HR remains a major survival pathway in this family of prokaryotes (Bowater and Doherty, 2006).

The *E. coli* HR system has been extensively studied, is well understood, and is used as a model for HR (Fig. 1.2) (Kowalczykowski et al., 1994). For repair of a DSB by HR, an undamaged copy of the broken chromosome, which is used as a template for repair DNA synthesis, is required. This copy can be either a homologous chromosome (in diploid organisms) or a sister chromatid (in haploid and diploid organisms during, or following, chromosome replication). HR can be divided into three stages: pre-synapsis, synapsis, and post-synapsis. During pre-synapsis, extensive processing of the broken duplex takes place. In *E. coli* this is driven by the RecBCD enzyme complex and results in the exposure of 3′ single-stranded DNA (ssDNA) ends. These ends are rapidly coated with RecA to form a nucleoprotein filament, which serves both to protect the DNA by preventing intra-strand secondary structures, and to promote the next phase of repair, synapsis. During synapsis, one or both (if the DSB is two-ended) of the RecA-coated 3′ ends invade the intact homologous duplex to form D-loops. DNA synthesis is re-established from these D-loops, initiated by PriA, and the information lost as a result of the DSB is restored. During post-synapsis, the two chromosomes, which have now formed a joint molecule, need to be separated. This involves the resolution of one or more four-way DNA junctions, known as Holliday junctions (HJs), and is dependent on RuvABC and, possibly, RecG. Resolution of two or more HJs can lead to the formation of crossovers, resulting in chromosome dimers, which need further processing to allow for complete chromosome segregation. This processing requires the XerCD site-specific recombination system (Barre et al., 2001).

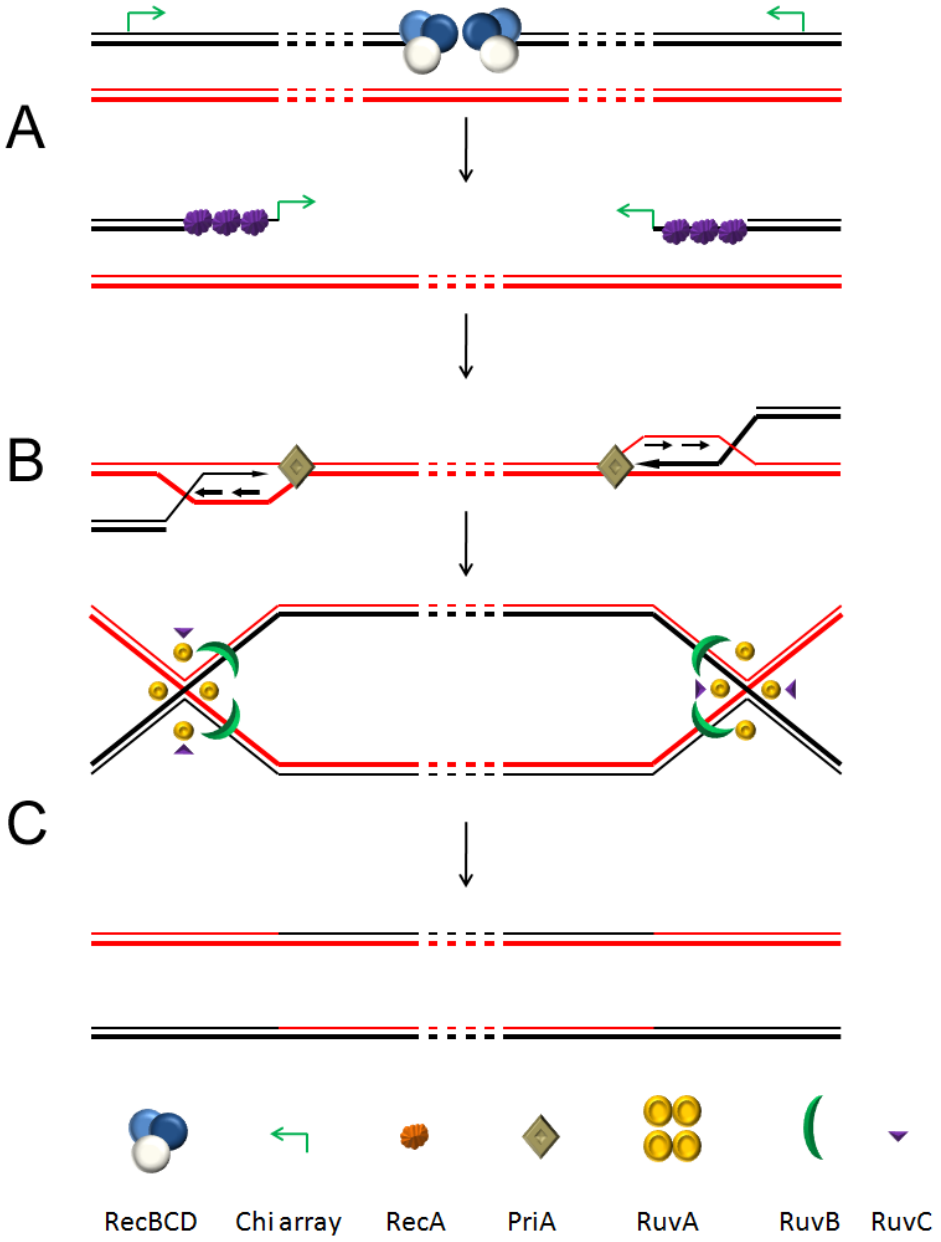

Figure 1. 2 DNA double strand break repair by homologous recombination in *E. coli*

Dashed line indicates DNA of variable length. Proteins are indicated by the key at the bottom of the figure. (A) Pre-synapsis. The free ends are processed by RecBCD to expose a 3' ssDNA overhang. This overhang is coated with RecA, which searches the genome for a homologous sequence. (B) Synapsis. Once homology is found, RecA catalyses strand invasion to form a D-loops, which can be used by PriA to prime DNA synthesis. Synthesis proceeds to restore the information lost as a result of the DSB and the RecBCD-mediated degradation. This forms two sister-chromosomes that are tied together by Holliday junctions. (C) Post-synapsis. The Holliday junctions are resolved by RuvABC, and/or RecG, and the two sister-chromosomes become free to migrate to opposite poles of the cell.

Proteins involved in HR are highly conserved across all domains of life, illustrating the importance of this pathway (Cromie et al., 2001).

1.3.1 Pre-synapsis

RecBCD is a 330 kDa, hetero-trimeric, ATP-dependent, bipolar helicase and nuclease complex also known as exonuclease V. It has high affinity for blunt DNA and it rapidly degrades foreign DNA, such as phage DNA, but can also process chromosomal DSBs for DSB repair by HR. The RecB and RecC genes were originally identified in recombination deficient mutants that had dramatically reduced growth rates (Capaldo-Kimball and Barbour, 1971). The identification of the third and final sub-unit of exonuclease V, RecD, came later due to the recombination proficient phenotype of its mutants (Amundsen et al., 1986).

Each component of the RecBCD complex has a distinct biochemical role (Fig. 1.3). Both RecB and RecD have helicase activities and belong to the SF1 superfamily of helicases (Gorbalenya and Koonin, 1993). RecC also belongs to this superfamily but the helicase activity has been lost. RecB has a slow 3′→5′ directionality while RecD, unconventionally for an SF1 helicase, unwinds in a 5′→3′ orientation with greater processivity (Dillingham et al., 2003; Taylor and Smith, 2003). UV light cross-linking experiments have shown that RecB loads onto the 3′ end of the break and RecD loads onto the 5′ end. This allows the two helicases to work cooperatively to maximise the processivity of exonuclease V (Ganesan and Smith, 1993). The 30 kDa C-terminus of RecB makes up the nuclease domain of the RecBCD complex. This domain is distinct from the helicase domain of RecB, and can be disrupted

without interfering with the helicase activity (Yu et al., 1998a; Yu et al., 1998b). As RecB and RecD translocate along a DNA duplex, the complementary strands of DNA are pulled across a "pin", located in RecC, and fed through separate channels in the protein (Singleton et al., 2004). This serves to separate the complementary strands of the DNA. Following these events, there are two major hypotheses that describe how RecBCD generates a 3′ ssDNA overhang. The first hypothesis predicts extensive degradation of both the 5′ and 3′ strands until a specific sequence, known as a crossover hotspot instigator sequence or Chi (χ 5′-GCTGGTGG-3′), is encountered (Fig. 1.3) (Lam et al., 1974; Smith et al., 1981). The second hypothesis predicts that RecBCD translocates along the broken chromosome, separating the 5′ and 3′ strands, without degrading them. When χ is encountered, RecBCD nicks the 3′ strand, five nucleotides to the 3′ side of χ, and unwinds it to generate a 3′ ssDNA overhang onto which RecA can be loaded (Fig. 1.4) (Ponticelli et al., 1985; Taylor et al., 1985).

1.3.1.1 RecBCD degrades the broken chromosome until χ

In-vitro experiments that analysed the activity of RecBCD on a χ-containing DNA fragment in conditions of excess Mg^{2+} to ATP, showed that the DNA located to the right of the χ sequence was degraded. These observations generated a model in which the broken DNA is degraded until χ (Dixon and Kowalczykowski, 1993).

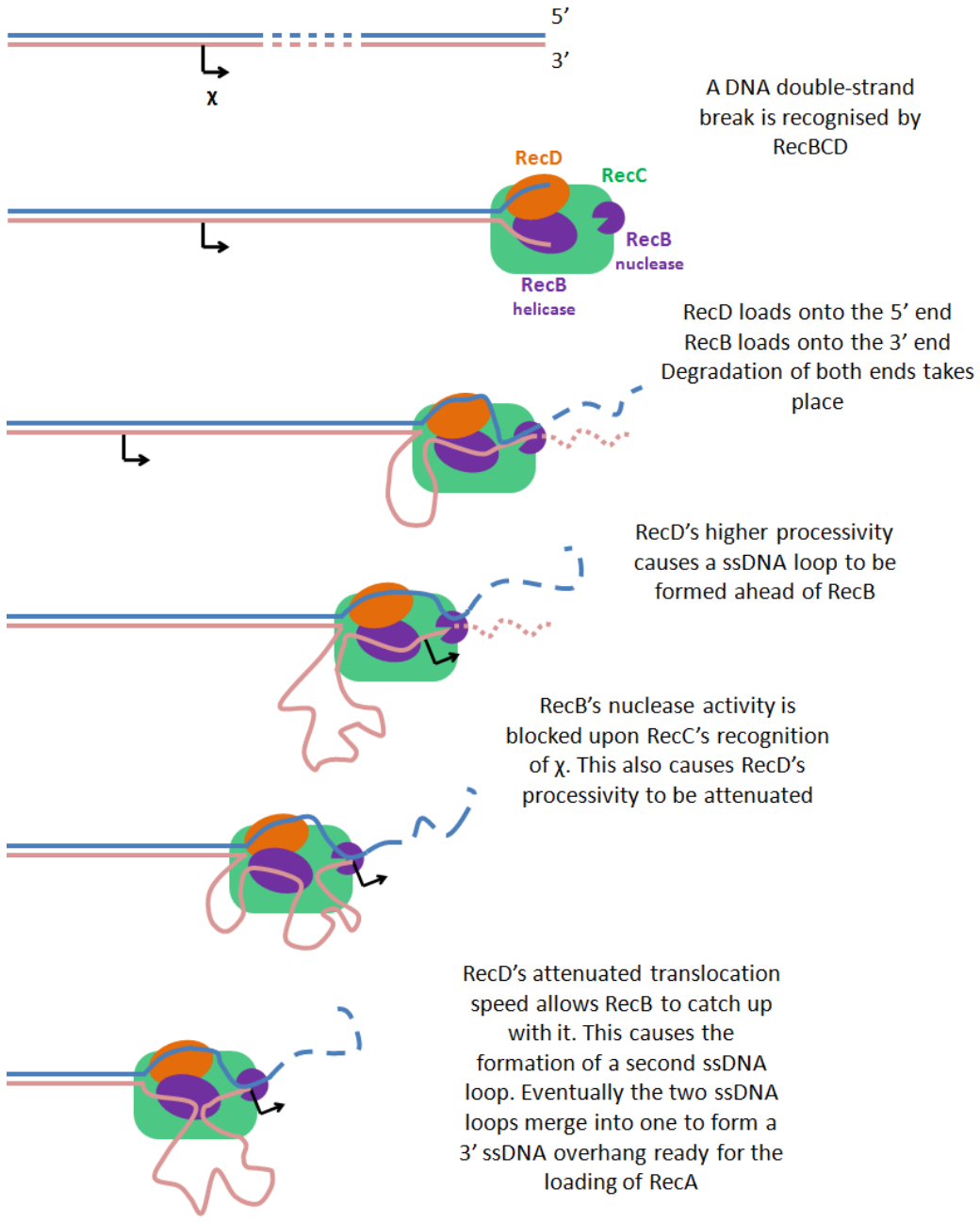

Figure 1. 3 Processing of a DNA double strand break by RecBCD; dsDNA degradation to χ

Upon recognising a DNA double strand break, RecB and RecD load onto the 3' and 5' ends of the DNA, respectively. Both proteins begin to translocate along the DNA duplex, degrading it. The faster translocation speed of RecD causes a ssDNA loop to form ahead of the 3' ended strand of RecB. When a χ sequence is recognised by RecC, RecC binds tightly to the 3' strand preventing its further degradation by RecB. Additionally, the translocation speed of RecD is attenuated, allowing RecB to catch up with it. This generates a ssDNA loop to form, which becomes a substrate for the loading of RecA.

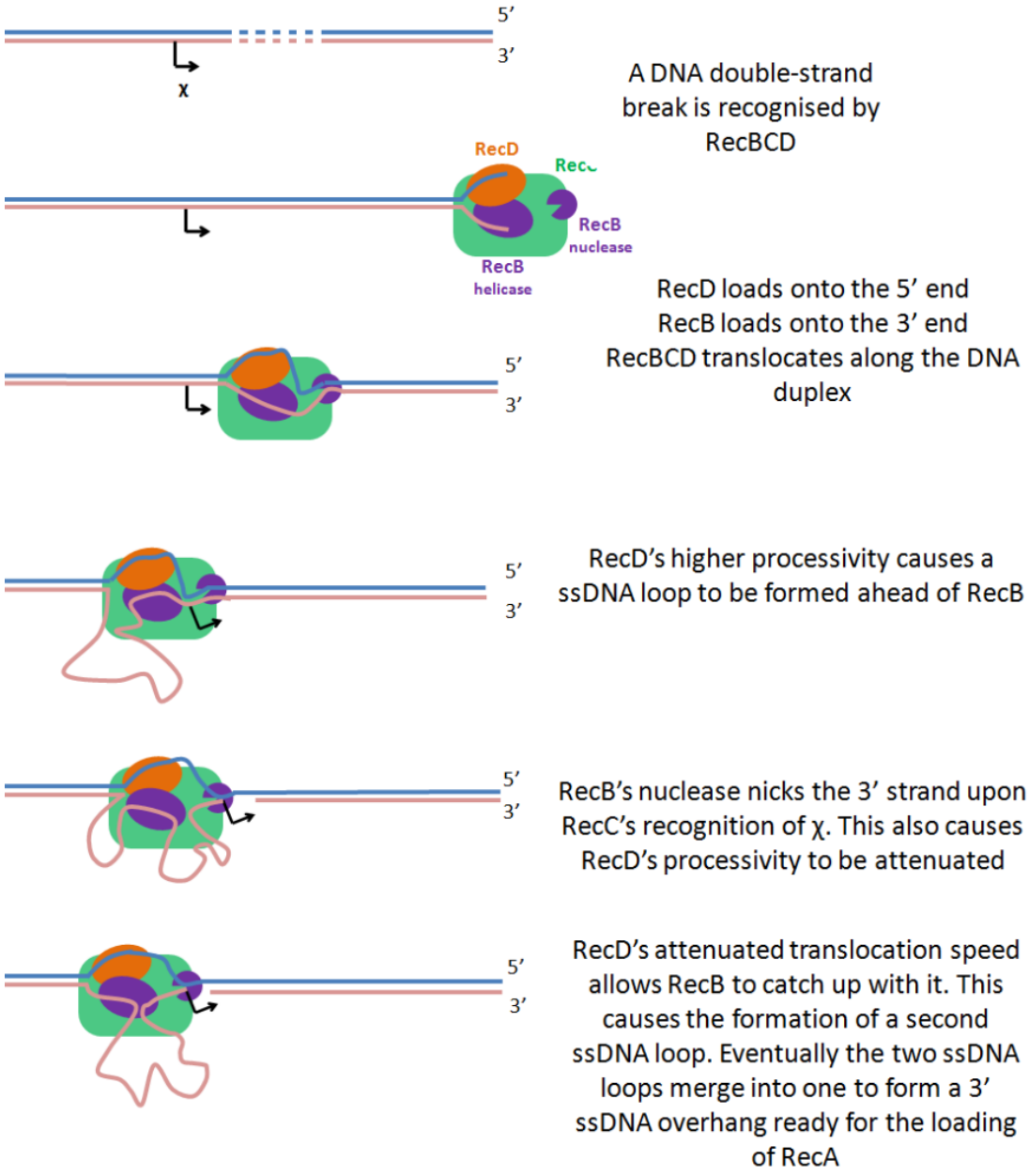

Figure 1. 4 Processing of a DNA double strand break by RecBCD; nick at χ

Upon recognising a DNA double strand break, RecB and RecD load onto the 3' and 5' ends of the DNA, respectively. Both proteins begin to translocate along the DNA. The faster translocation speed of RecD causes a ssDNA loop to form ahead of the 3' ended strand of RecB. When a χ sequence is recognised by RecC, the nuclease domain of RecB nicks the 3' DNA strand, which is then bound tightly by RecC. The translocation speed of RecD is attenuated, allowing RecB to catch up with it. This generates a ssDNA loop to form, which becomes a substrate for the loading of RecA.

Following the melting of the broken chromosome, the 3′ DNA end exits RecC close to the nuclease domain of RecB and is degraded at a high frequency. The 5′end exits RecC further away from the nuclease domain and is therefore cleaved less frequently. During degradation, the difference in translocation speed of RecB and RecD creates a ssDNA loop on the 3′ ended strand in front of RecB as it struggles to keep up with RecD (Dillingham et al., 2003; Spies et al., 2007; Taylor and Smith, 2003). In order to make the switch from degradation of the dsDNA to formation of a 3′ ssDNA overhang and RecA loading, RecC, which is intimately associated with RecB, has to be able to recognise a χ sequence, which is an 8-bp ssDNA sequence located on the 3′ ended strand (Ponticelli et al., 1985; Taylor et al., 1985). As the 3′ ssDNA is extruded from the pin of RecC, it passes through the protein and comes into contact with a region thought to have evolved from the active site of a helicase. It is this region in RecC that has been implicated in recognising χ. Indeed, the ssDNA binding site of SF1 helicases, and thus the same region in RecC, has been reported to recognise 8 bp sequences and this is thought to be the reason why the sequence of χ is also 8 bp long (Korolev et al., 1997; Singleton et al., 2004; Velankar et al., 1999). Upon recognising χ, RecC binds tightly to the 3′ ended tail of the DSB and prevents further degradation/translocation on this strand, generating a second ssDNA loop. This allows the 5′ ended tail of the break to access the nuclease domain of RecB more frequently and degradation of this strand increases. At the same time, the overall processivity of the complex is reduced as the translocation speed of RecD is attenuated (Boehmer and Emmerson, 1992; Dillingham et al., 2003; Singleton et al., 2004). This results in a switch in the lead motor subunit from RecD

to RecB (Spies et al., 2007). RecB is now faster than RecD so the two ssDNA loops that have formed on the 3′ ssDNA end merge into one. This new DNA loop becomes a substrate for RecA loading (Anderson and Kowalczykowski, 1997; Churchill and Kowalczykowski, 2000; Singleton et al., 2004; Spies et al., 2007). A complex "signal cascade" between the three subunits of RecBCD has been proposed to govern these events (Amundsen et al., 2007).

1.3.1.2 RecBCD translocates along the DNA until χ

In-vitro experiments that analysed the activity of RecBCD on a χ-containing DNA fragment in conditions of excess ATP to Mg^{2+}, showed that the DNA located between the 3′ end of the fragment and the χ sequence remained intact after nicking at χ (Taylor et al., 1985). These observations generated a model in which the broken DNA was not degraded by RecBCD (Fig. 1.4) (Taylor et al., 1985). This model was supported by EM data showing that in conditions of limited Mg^{2+}, RecBCD loaded onto the end of a DNA fragment and generated ssDNA loop-tail structures (Taylor and Smith, 1980). All other events are proposed to occur as described in section 1.3.1.1.

1.3.2 Synapsis

1.3.2.1 The formation of a RecA nucleoprotein filament

RecA, responsible for synapsis, is the driving force of DSB repair (Cox, 2007; Lusetti and Cox, 2002). Although a RecA-independent pathway has been reported in *E. coli* its physiological importance is still under question (Dutra et al., 2007). The understanding that RecA played an important role in the life cycle of *E. coli* came in 1965 when it was discovered to be mutated in strains that were unable to undergo conjugational recombination and repair damage induced by UV light (Clark and Margulies, 1965). RecA is a DNA-dependent ATPase and an ATP-dependent DNA binding protein. It consists of 352 amino acids and has a molecular weight of about 37.8 kDa (Sancar et al., 1980). RecA is made of a main core domain, a C-terminal domain, and an N-terminal domain. The core domain is highly conserved across a wide range of organisms, illustrating the importance of this protein (Lusetti and Cox, 2002). Within the core domain there is a Walker A box and the C-terminal domain is rich in negatively charged residues. These features are typical of most DNA binding proteins (Benedict and Kowalczykowski, 1988; Tateishi et al., 1992; Yu and Egelman, 1990).

RecA polymerises on RecBCD-processed 3′ ssDNA ends. This prevents the formation of secondary structures, which have been shown to inhibit the polymerisation reaction (Bar-Ziv and Libchaber, 2001), and promotes strand exchange between homologous sequences. Certain GT-rich sequences, reminiscent

of χ recombinational hotspots, have also been shown to promote this process (Tracy and Kowalczykowski, 1996). In parallel, RecA induces the SOS response by acting as a coprotease through the cleavage of the SOS regulon repressor, LexA (Ivancic-Bace et al., 2006). LexA degradation works as a positive feedback, increasing the expression of many repair proteins including RecA itself. This ensures that DSBs are repaired as efficiently as possible.

RecA polymerisation on the 3′ ssDNA end occurs in two main steps. Initially there is a slow nucleation event, which is then followed by a rapid and cooperative extension with 5′→ 3′ directionality (Register and Griffith, 1985; Sattin and Goh, 2004). As a result of the coating of a strand of DNA by RecA, globally the DNA filament is extended in length by about 50%. This is thought to aid the search for homology (Egelman and Stasiak, 1986; Klapstein et al., 2004; Nishinaka et al., 1997). The crystal structure of RecA bound to ssDNA confirms this, but also highlights that locally the DNA retains a B-DNA-like conformation, which allows for precise Watson-Crick base pairing between DNA strands during the search for homology (Chen et al., 2008). RecA, in a complex with the damaged DNA, searches the genome for a homologous sequence from which to promote repair. *In vitro*, the pre-synaptic nucleoprotein filament can compare DNA sequences by scanning approximately 10^2-10^3 DNA segments per second (Yancey-Wrona and Camerini-Otero, 1995). *In vitro* experiments also reported that during the initial stages of strand exchange the nucleoprotein filament shows high sensitivity to single mismatches, which inhibit the exchange reaction. The sensitivity to single mismatches is higher when they are proximal to the 3′ end of the invading strand, illustrating the importance of homology (Sagi et al., 2006). The crystal structure also

showed that the global elongation of the DNA was very important in the search for homology. The binding of the donor DNA duplex to the RecA nucleoprotein filament was shown to be stabilised primarily by Watson-Crick interactions, explaining the high fidelity of RecA-mediated strand exchange and its sensitivity to single mismatches (Chen et al., 2008). Successful strand exchange promotes the formation of a DNA junction, known as a D-loop, from which repair DNA synthesis, using the unbroken chromosome as a template, can be initiated. This process is dependent on a large array of recombination and replication proteins. The protein that initiates this origin-independent DNA replication is PriA (Gabbai and Marians, 2010).

1.3.2.2 The establishment of origin-independent replication by PriA

PriA is a 3′ → 5′ DNA helicase that is required for origin-independent replication re-start at a variety of DNA junctions (Gabbai and Marians, 2010). PriA, along with six other primosomal proteins, PriB, PriC, DnaT, DnaC, DnaB, and DnaG, were first identified during the *in vitro* replication of ΦX174 bacteriophage (Liu et al., 1996; Ng and Marians, 1996a; Ng and Marians, 1996b). *In vitro* experiments have reported that PriA can bind D-loop structures and it was suggested that the ΦX174 primosome assembly site (PAS) onto which PriA loads during ΦX174 replication evolved to mimic the structure of an *E. coli* chromosomal D-loop, one of the natural substrates for the action of PriA (McGlynn et al., 1997). At the PAS of ΦX174, PriA is the first protein to bind to the SSB-coated DNA and

functions as a scaffold for the recruitment of at least other five proteins (Ng and Marians, 1996a; Ng and Marians, 1996b). PriB is the first protein to recognise and bind to the PriA-ssDNA complex and its binding aids the loading of DnaT (Liu et al., 1996). In some circumstances PriC also binds to the primosome, but the binding of this protein is not always required. Following these events, DnaC, which forms a complex with DnaB, targets DnaB onto the PriAB(C)-DnaT-ssDNA complex. Finally, the primase DnaG, can associate with the PriAB(C)-DnaBT-ssDNA complex and synthesize RNA primers that are used by the DNA polymerase III holoenzyme to synthesize DNA (Marians, 2000). Δ*priA* mutants are recombination deficient and as a result are sensitive to DNA damaging agents (Kogoma et al., 1996). The requirement for the activity of PriA is illustrated by the propensity of Δ*priA* mutants to acquire suppressor mutations in *dnaC* (Sandler et al., 1999; Xu and Marians, 2000). These mutations allow for the loading of DnaB in the absence of PriA, emphasizing the loading of DnaB as being the crucial step in the re-establishment of origin-independent replication (Sandler, 1996).

Recent studies have shown that the N-terminal of PriA (which does not contain the helicase motif, as this is found in the C-terminal domain) has a small pocket that accommodates the 3′ terminal nucleotide of a leading strand of DNA without disrupting the base-pairing of this nucleotide to its complementary strand. This binding is base-non-specific as all of the four bases are recognised by the binding pocket (Sasaki et al., 2007). These results are in accordance with the action of PriA at stalled replication forks or D-loops, as both of these structures are predicted to contain a DNA leading strand with a 3′ terminal nucleotide that is base-paired to a complementary strand of DNA. The correct recognition of the 3′

terminal nucleotide is thought to allow PriA to bind the junction in an orientation that promotes primosome assembly and does not allow the C-terminal helicase domain to unwind the junction, which would de-stabilise the fork or D-loop. In a situation where incorrect binding of PriA occurs, RecG has been postulated to bind to the junction and prevent it from being de-stabilised (Al-Deib et al., 1996; Tanaka and Masai, 2006). Indeed, the helicase and primosome assembly activities of PriA can be uncoupled without resulting in loss of viability or recombination deficiency, strengthening the hypothesis that the helicase activity is not required for primosome assembly (Sandler et al., 1996). In accordance with this, a strain harbouring the helicase deficient *priA300* mutant is proficient in primosome assembly and does not require the activity of RecG (Al-Deib et al., 1996).

1.3.3 Post-synapsis

1.3.3.1 The Holliday junction resolvase, RuvABC

In post-synapsis, PriA-mediated DNA replication has closed the gap generated by the DSB, resulting in the formation of sister-chromosomes that are physically attached to one another by structures known as Holliday junctions (HJs) (Fig. 1.2) (Holliday, 1974; Kobayashi and Ikeda, 1983). These structures are dynamic four-way DNA junctions that lead to the formation of heteroduplex DNA (hDNA) (Duckett et al., 1988). The length of the hDNA can vary through branch migration, a

process first proposed by Holliday and later revisited by Meselson and Radding (Holliday, 1974; Meselson and Radding, 1975).

Otsuji and collaborators identified the first strains of *E. coli* deficient in Holliday junction resolution (Otsuji et al., 1974). These strains were sensitive to UV radiation, after which cells grew as non-septate multinucleate filaments. They were originally thought to be deficient in cell division and were recognised as recombination deficient only ten years later (Lloyd et al., 1984). Since then, our understanding of Holliday junction resolution in *E. coli*, and of the proteins involved, has greatly increased. Similar efforts in eukaryotes have only just started to uncover the details of Holliday junction processing in these organisms (Gaskell et al., 2007; Geuting et al., 2009; Ip et al., 2008; Jessop and Lichten, 2008; Oh et al., 2008).

The RuvABC complex is the major HJ resolvase in *E. coli*. *ruvC* is located upstream of *ruvA* and *ruvB* and sits in its own open reading frame (ORF) while *ruvAB* are in an operon, as shown by polar effects of transposon (Tn) insertion mutagenesis (Sharples et al., 1990). Transcription of *ruvAB* is SOS-inducible, while the transcription of *ruvC*, which is poor, is not (Connolly et al., 1991). RuvA and RuvB work cooperatively to promote branch migration of HJs and/or secondary structures, such as cruciforms and hairpins (Shiba et al., 1991). The crystal structure of *Mycobacterium leprae* RuvA has shown that two tetramers can bind to the DNA junction and sandwich the DNA, keeping the junction in an open square planar conformation (Roe et al., 1998). This recruits RuvB, which interacts with the RuvA-DNA complex and provides the motor force for branch migration of the Holliday junction (Dickman et al., 2002). Indeed, when over-expressed, RuvB can migrate

junctions in the absence of RuvA. Additionally, branch migration can also occur spontaneously *in vitro*, in the absence of both RuvA and RuvB (Panyutin and Hsieh, 1994). RuvC is the subunit responsible for the resolution of the junction and it requires both RuvA and RuvB in order to function. RuvC inserts nicks at, or near, the junction that are then sealed by DNA ligase, forming two recombinant duplexes (Dunderdale et al., 1991; West, 1997).

The eukaryotic Mus81/Eme1 complex has recently been shown to have HJ processing activity both *in vitro* and *in vivo* (Gaskell et al., 2007). Other important complexes required for the resolution of joint molecules in eukaryotes are the Sgs1-TopIII-Rmi1 HJ dissolution complex and the Slx1-Slx4 endonucleases (Ashton et al., 2011; Fricke and Brill, 2003). More recently, a nuclease that resolves HJs in a fashion analogous to that of RuvC has been identified in *Saccharomyces cerevisiae* (Yen1) and humans (Gen1) (Ip et al., 2008). It is apparent that HJ resolution in eukaryotes is a very complicated process that can occur via many different enzymatically-driven pathways.

1.3.3.2 The recombination helicase, RecG

recG was first identified by Storm and collaborators when isolating recombination deficient mutant of *E. coli* K12 (Storm et al., 1971). The mutation *recG162* conferred sensitivity to UV, mitomycin C and ionising radiation, and reduced conjugational and P1 transductional efficiency. A few years later, using transposon mutagenesis, Lloyd & Buckman reported that a mutation in *recG* also

caused a reduction in HR (Lloyd and Buckman, 1991). From the beginning, it appeared that RecG was involved in DNA repair and that it was required to repair a variety of DNA lesions. This characteristic distinguished RecG from other recombination proteins, such as RecBCD, RecFOR or RecET, which fell into specific epistatic groups. Still today the exact role that RecG plays in recombination remains elusive.

recG is the last gene in the *spo* operon (Kalman et al., 1992). The basal expression of *recG* is low and, unlike *ruvAB*, it is not SOS-inducible (Lloyd and Sharples, 1991). The primary amino acid sequence reveals that RecG has similarities to DNA helicases and contains a well-conserved ATP-binding motif (Kalman et al., 1992; Lloyd and Sharples, 1991). However, *in vitro*, RecG is not able to act on conventional helicase substrates, suggesting that RecG targets other DNA junctions (Lloyd and Sharples, 1993b). Following the over-expression and purification of the protein, RecG was found to be a DNA-dependent ATPase, able to bind to and dissociate synthetic HJs in a fashion analogous to that of RuvAB (Lloyd and Sharples, 1993a). Indeed, *recG ruv* mutants had previously been reported to have a 100-fold decrease in efficiency of recombination when compared to the individual mutations. This suggested that RuvABC and RecG were redundant and implicated RecG in the resolution of HJs (Lloyd, 1991). Nevertheless, studies of RecG activity gave no direct evidence that it could resolve HJs by endonucleolytic cleavage, as was the case for HJ resolution mediated by RuvABC (Lloyd and Sharples, 1993b). Experiments that compared the ability of RecG and RuvAB to process four-way (X) and three-way (Y) DNA junctions, both thought to be intermediates present in RecA-mediated HR, showed that both proteins could bind X and Y-junctions,

although RecG appeared to bind Y-junctions with more stability than RuvAB. Unwinding of these junctions was also possible using either protein but RecG showed a higher specific activity for junctions, and a lower concentration of RecG was required to dissociate all junctions efficiently. It was particularly striking that the dissociation of 50% of the X-junctions required 1000-fold more RuvAB than RecG. RecG was also found to drive reactions to completion much faster than RuvAB (Lloyd and Sharples, 1993a). Following these observations it was proposed that RecG can resolve HJs by branch migration of two junctions into each other or by branch migration of a HJ past a replication fork (Wardrope et al., 2009; Whitby et al., 1993). This mechanism of HJ processing allows for their resolution without the requirement for a nuclease.

Another proposed role for RecG *in vivo* is the stabilisation of D-loops generated by RecA-mediated strand invasion (McGlynn et al., 1997; Whitby and Lloyd, 1995). This model proposed that RecG was required to turn D-loops into HJs, which RuvABC could then resolve. This hypothesis is supported by evidence that RecG and SSB directly interact with each other, placing RecG at the location of a replication fork or a replication fork-like structure, such as a D-loop (Buss et al., 2008). This is strengthened by the functional relationship that RecG has with PriA, the helicase required for the assembly of the primosome at stalled replication forks and D-loops (Section 1.3.2.2) (Al-Deib et al., 1996; Tanaka and Masai, 2006). It is interesting to note that RecG can also act at R-loops, which are structures similar to D-loops (Fukuoh et al., 1997; Ohsato et al., 1999).

More recently, it has been suggested that RecG prevents over-replication of the chromosome through PriA-mediated primosome assembly at the location of

converging replication forks (Rudolph et al., 2009a; Rudolph et al., 2009b). This led Rudolph and collaborators to propose that RecG acts as a general guardian of the genome that protects the cell from pathological DNA replication (Rudolph et al., 2010).

1.4 Intermediates of DNA double strand break repair

Gene conversion is the physical manifestation of genetic recombination. When gene conversion was first described it was accepted that some form of intermediate, which required the physical joining of chromatids to allow for the exchange of genetic information, had to be generated (Holliday, 1964; Pritchard, 1955). It was noted that exposing an organism to low doses of a DNA damaging agent, such as UV light, increased the frequency of gene conversion events (Holliday, 1966). This observation highlighted for the first time the intimate relationship between DNA repair and genetic recombination.

Evidence for the existence of intermediates of gene conversion, in the form of branched DNA structures that were generated by events of DNA exchange, was obtained by electron microscopy (EM) of *E. coli* infected with bacteriophage T4 (Broker and Lehman, 1971). It was not until 1984 that enzymes capable of acting on these intermediates were discovered (de Massy et al., 1984; Hsu and Landy, 1984; Kemper et al., 1984). Nearly 40 years later, our understanding of the genetic recombination that occurs during the repair of DNA damage has much improved (Paques and Haber, 1999).

1.4.1 Different pathways for DNA double strand break repair

Much of our understanding of DSB repair has come from the study of the budding yeast *S. cerevisiae*, and more recently the fission yeast *S. pombe*, in which DSBs can be specifically induced by endonucleases such as the HO endonuclease or the rare-cutting I-SceI endonuclease (Davis and Smith, 2001; Osman and Subramani, 1998; Paques and Haber, 1999). Eukaryotic systems for studying the intermediates of DSB repair have prevailed over prokaryotic systems due to the advantage of eukaryotes possessing homologous chromosomes, which has allowed DSB formation to be homologue specific and the genetic exchange that occurs during repair to be between non-sister chromatids, allowing for their detection. As genetic exchange between sister chromatids remains difficult to detect, to date, the analysis of DSB repair intermediates in *E. coli* or other prokaryotes has been less fruitful. Three main mechanisms for DSB repair by HR have been described in eukaryotes; the canonical double strand break repair (DSBR) pathway, synthesis-dependent strand annealing (SDSA), and break-induced replication (BIR).

1.4.1.1 Canonical double strand break repair

The canonical DSBR pathway as we know it today was proposed by Szostak and collaborators following the alteration of previous models (Holliday, 1974; Meselson and Radding, 1975; Resnick and Martin, 1976; Szostak et al., 1983). This

model predicts that the ends of a DSB are processed to generate two ends with 3' ssDNA overhangs, both of which can invade an intact homologous chromosome by strand-invasion. This process would lead to the formation of two HJs, one on either side of the DSB, which would then be resolved by a HJ resolvase to give either crossover or non-crossover products for flanking markers (Figure 1.5).

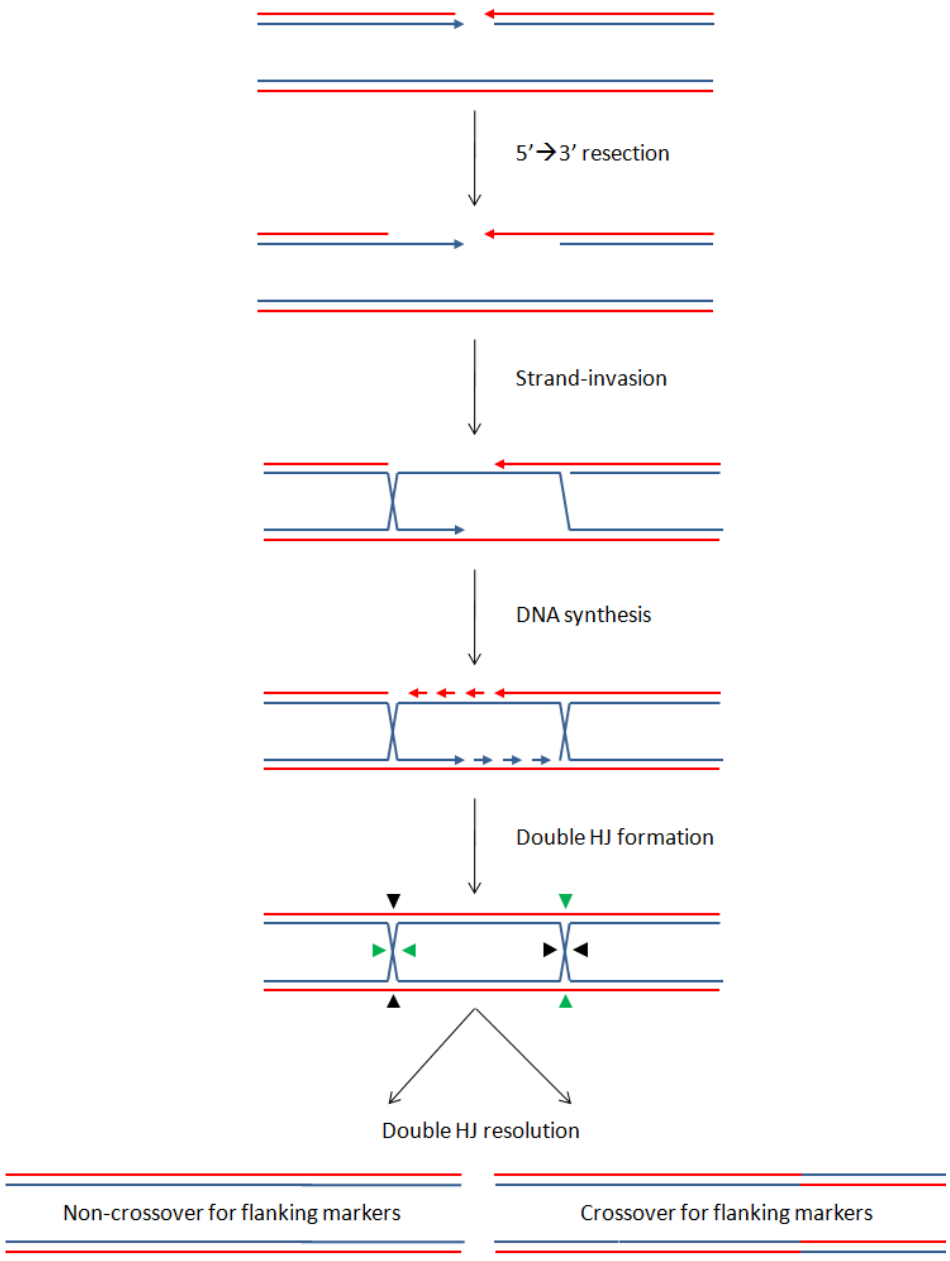

Figure 1. 5 Repair of a DNA double strand break by the canonical double strand break repair pathway

The DNA free ends are processed by a nuclease to expose 3' ssDNA overhangs, which strand-invade an unbroken homologue to generate a D-loop. A primosome is assembled onto the D-loop and new leading and lagging strand DNA synthesis is established. DNA synthesis closes the gap generated by the double strand break to form a joint molecule where a Holliday junction is located either side of the original breakpoint. This double Holliday junction structure is then resolved through the action of a Holliday junction resolvase. The orientation in which the two junctions are resolved with respect to one another determines whether the products of recombination will be crossovers or non-crossovers for flanking markers.

1.4.1.2 Synthesis dependent strand annealing

SDSA was a pathway originally devised to explain gene conversion events in yeast that were not associated with the crossover of flanking markers (Hastings, 1988; McGill et al., 1989; Nassif et al., 1994). Additionally, SDSA seemed to explain recombination events of bateriophage T4 that were observed in *E. coli* (Kreuzer et al., 1995; Mueller et al., 1996). Strong evidence for SDSA was obtained when genetic information from two different loci on two different chromosomes were incorporated into a single broken DNA fragment (Paques et al., 1998; Silberman and Kupiec, 1994). The only way to explain this was that the DNA strands that invaded these two different loci were eventually displaced and re-annealed to one another in order to terminate repair. Indeed, the major difference between SDSA and canonical DSBR is that SDSA predicts the dismantling of the D-loop once the DNA synthesis that is associated with repair has occurred. This dismantling may either be an active event, mediated by the action of a helicase that unwinds the D-loop and extrudes the invading strand, or a passive event such as occurs in the "bubble migration model" of SDSA. As the intermediates of recombination are resolved by the extrusion of the invading strand, SDSA does not require the activity of a HJ resolvase in the way that the canonical DSBR pathway does. Additionally, SDSA does not depend on both DNA free ends engaging with the unbroken chromatid. In the "bubble migration model" only one DNA free end invades the homologue (Figure 1.6). It is important to note that versions of SDSA that do require the resolution of HJs and do result in crossovers for flanking markers have also been proposed (Paques and Haber, 1999). This model differs from canonical DSBR in that

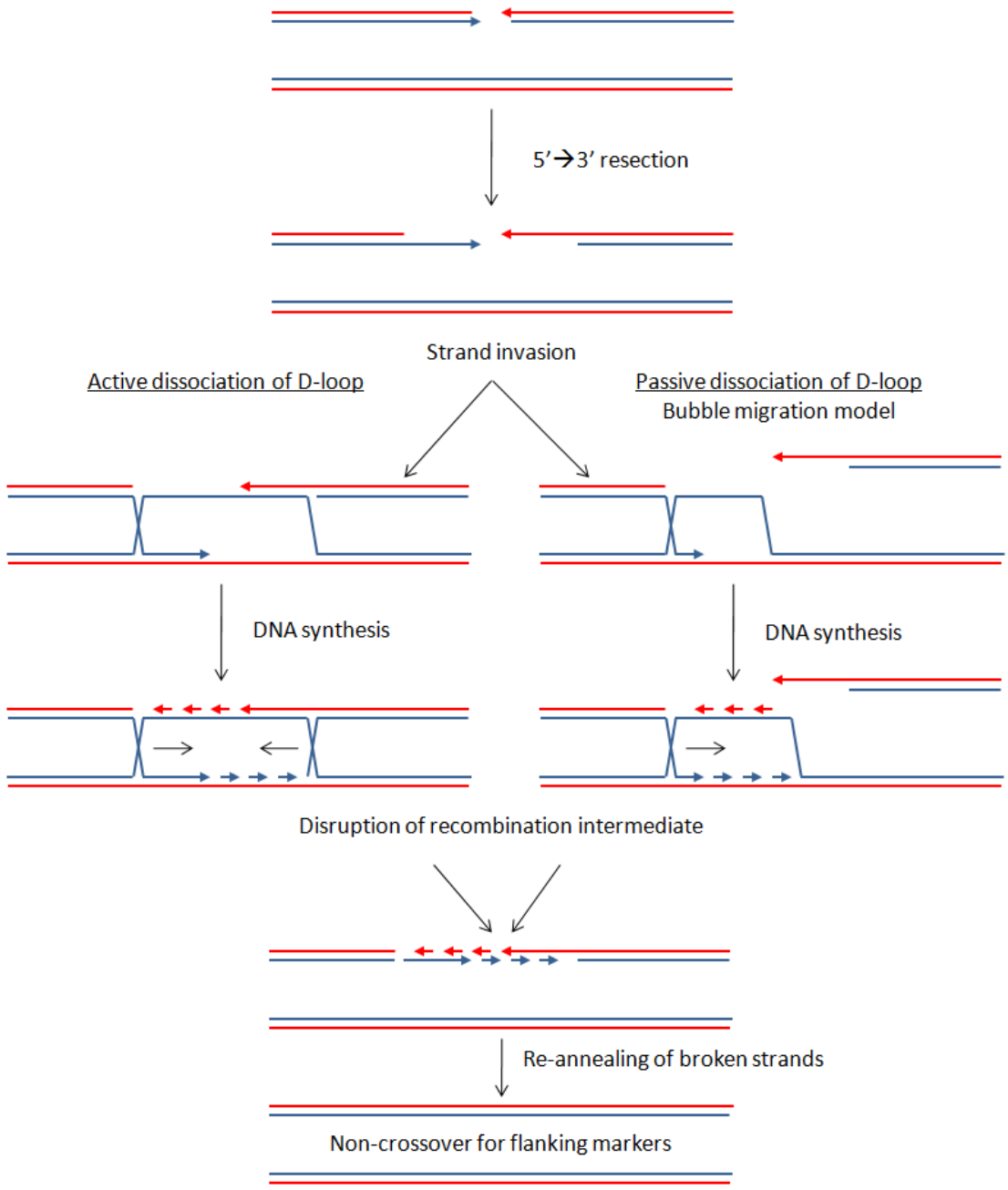

Figure 1. 6 Repair of a DNA double strand break by synthesis dependent strand annealing

The DNA free ends are processed by a nuclease to expose 3' ssDNA overhangs, which strand-invade an unbroken homologue to generate a D-loop. Following this, the second DNA free end is captured. In the bubble migration model the second DNA free end never engages with the homologous chromosome. A primosome is assembled onto the D-loop and new leading and lagging strand DNA synthesis is established. DNA synthesis closes the gap generated by the double strand break and the D-loop is dissociated, either actively or passively. The dissociated strands re-anneal and repair is terminated.

a single HJ is predicted to arise on one side of the DSB, generating an asymmetrical intermediate of repair (Ferguson and Holloman, 1996).

1.4.1.3 Break induced replication

Break induced replication was a term coined to describe gene conversion events that were detected across very large chromosomal distances of up to 400 Kb (Esposito, 1978). This model of DSB repair is very similar to replication events described in bacteriophage T4 and in *E. coli*, termed recombination induced replication (Formosa and Alberts, 1986; Kogoma, 1997; Mosig, 1987). The model predicts that a single DNA free end invades the unbroken chromosome and establishes DNA synthesis. A single HJ, located upstream of the invasion point, then needs to be resolved by a HJ resolvase. This synthesis proceeds to the end of the chromosome (Fig. 1.7). In E. coli, this mechanism of DSB repair would be predicted to occur following chromosomal replication through a ssDNA nick in the template (Fig.1.1 B).

1.4.2 The use of two-dimensional agarose gel electrophoresis in studying intermediates of DNA metabolism

Studying the products of recombination events alone can only yield a limited amount of information. In order to validate a model, it is necessary to gain physical evidence of the intermediates generated by that model. Native two-dimensional

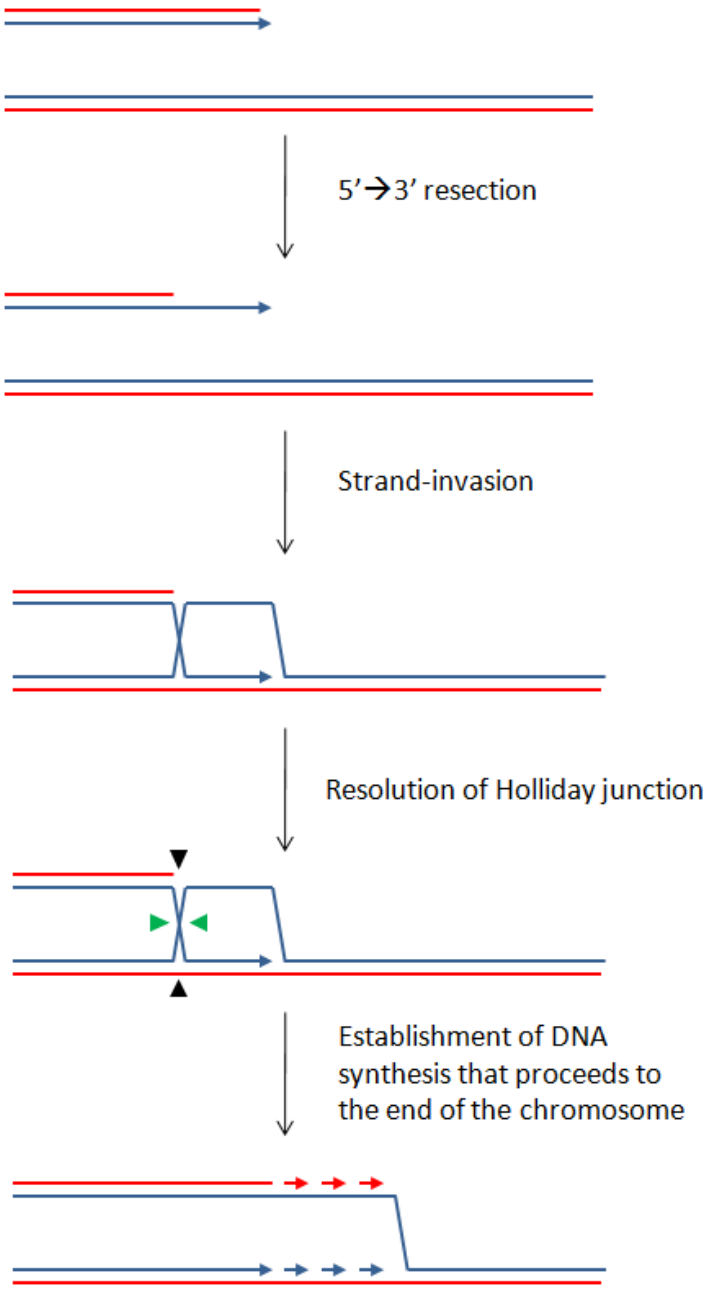

Figure 1. 7 Repair of a one-ended DNA double strand break by break induced replication.

The single DNA free end is processed by a nuclease to expose a 3' ssDNA overhang, which strand-invades an unbroken homologue to generate a D-loop. A primosome is assembled onto the D-loop and new leading and lagging strand DNA synthesis is established. The single Holliday junction that arises upstream of the invasion point is resolved by the action of a Holliday junction resolvase and DNA synthesis proceeds to the end of the chromosome.

agarose gel electrophoresis (2-D agarose gels), performed on restriction digests of chromosomal DNA, is a technique that allows for the visualisation of recombination intermediates (Bell and Byers, 1983). In the first dimension the DNA is separated at low voltage at 4 °C. These conditions minimise damage to the branched intermediates, maintaining them as native as possible, and separates the DNA fragments based on molecular weight. In the second dimension, 0.3 μg of ethidium bromide, which intercalate with the DNA, are added to maximise the differences between the three-dimensional structures of the different DNA species being separated. This, coupled with a high voltage, serves to maximise separation in the second dimension based on molecular shape. The product of this step-wise separation technique is the characteristic spot and arc pattern formed upon Southern blot of a DNA restriction fragment of interest (figure 1.8). An alternative to native 2-D agarose gel electrophoresis is alkaline 2-D agarose electrophoresis, where the second dimension is run in denaturing conditions (Huberman et al., 1987). This allows for the separation of individual strands of a DNA molecule, which can then be detected using strand-specific probes.

Separation of DNA in 2-D was first described in 1981 by Ariella Oppenheim (Oppenheim, 1981). By separating plasmid DNA in a second dimension, it became possible to distinguish between closed circular, nicked circular, and linear DNA species. This technique was then coupled to Southern blotting of the DNA, and further optimised in order to analyse replication and recombination intermediates of DNA metabolism (Bell and Byers, 1983). Following the optimisation of this DNA separation technique, extensive plasmid-based work was carried out by a number of independent groups who were interested in understanding both the replication and

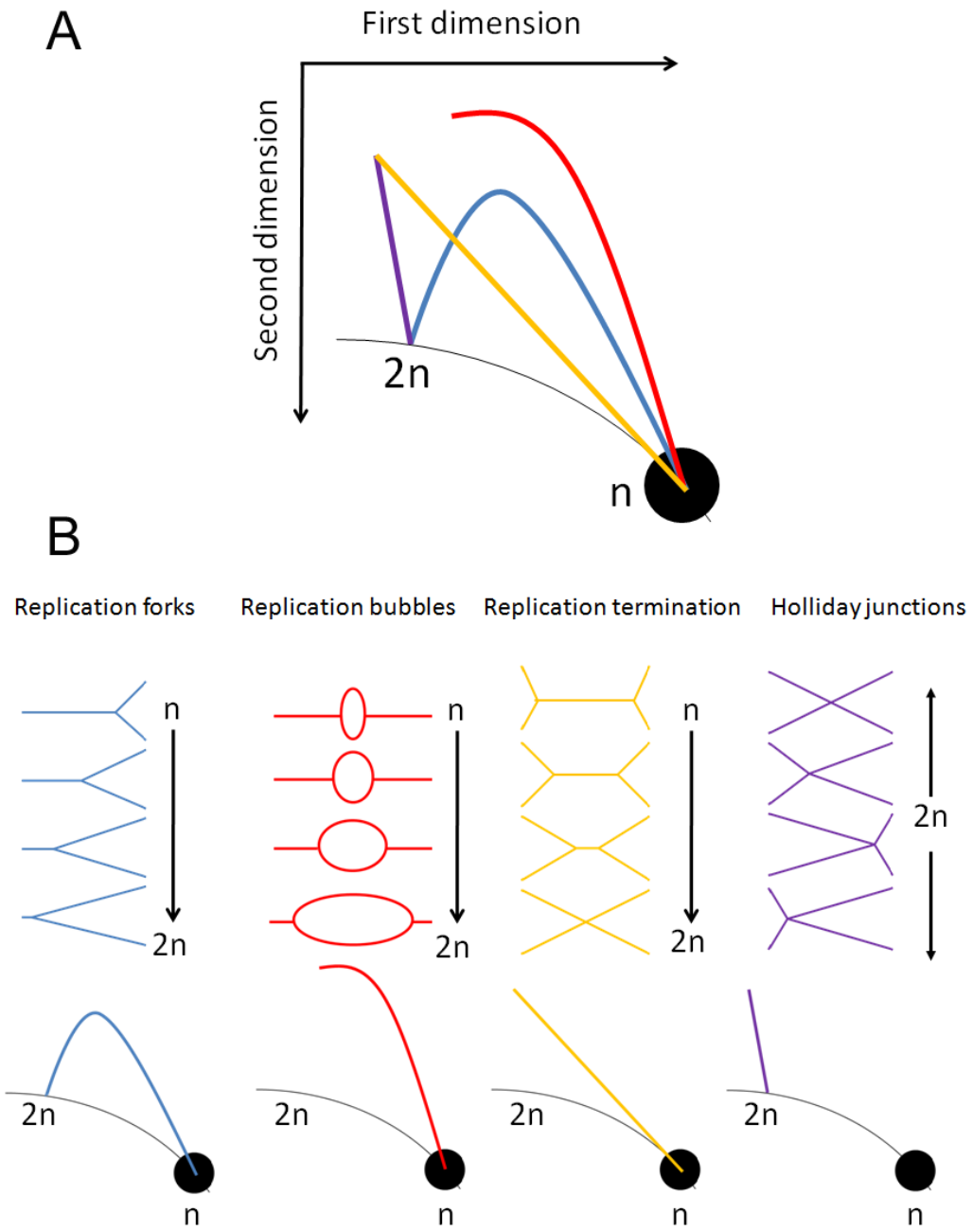

Figure 1. 8 Overview of 2-D agarose gel electrophoresis migration patterns.

The large spots designated "n" indicate the positions of the abundant linear species of the restriction fragments. "2n" indicates the location of this linear species just prior to completion of replication. Accumulation of a particular structure, such as a blocked replication fork at a specific location along a restriction fragment, generates a spot along the relevant migration line. (A) Overview of the most common migration patterns observed. (B) Breakdown of the different molecular shapes placed above the migration pattern they generate. Figure modified from: (Friedman and Brewer, 1995).

recombination of DNA molecules. Brewer and Fangman showed that autonomous replication sequences (ARS) isolated from yeast chromosomes participated in replication initiation *in vivo* by isolating replication bubbles by native 2-D agarose gel electrophoresis (Brewer and Fangman, 1987). A few years later, alkaline 2-D agarose gel electrophoresis was used to provide evidence that 5′ DNA ends of recombining DNA were degraded during recombination in *Xenopus laevis* cell extracts. Due to the nature of the plasmid substrates used in this experiment, 3′ strand invasion and D-loop formation was not required for recombination to occur and therefore evidence of DNA replication, a fundamental step in most recombination events, was not detected. Nevertheless, it was concluded that during recombination DNA ends required processing that involved the extensive degradation of 5′ DNA ends. This would give rise to a DNA duplex containing a 3′ ssDNA overhang, a structure now present in all models of recombination (Maryon and Carroll, 1991).

Studies of DNA replication and recombination moved from using plasmid DNA as substrates to using chromosomal DNA. Resolution of *E. coli* chromosome dimers, products of homologous recombination that arise during chromosome replication, was shown to occur via a HJ intermediate and required the action of the C-terminal domain of cell division protein FtsK for resolution. This indicated that chromosome dimer resolution and cell division are tightly coupled (Barre et al., 2000). A year later, evidence that homologous chromosomes became physically linked to one another to generate a region of heteroduplex DNA, was obtained by native 2-D agarose gel electrophoresis of DNA fragments isolated during yeast meiotic recombination (Allers and Lichten, 2001). More recently, native 2-D agarose

gel electrophoresis has been used to prove that double HJs are intermediates of meiotic and mitotic DSB repair in the budding yeast *S. cerevisiae*, but only single HJs are generated by the repair of a meiotic DSB in the fission yeast *S. pombe* (Bzymek et al., 2010; Cromie et al., 2006; Schwacha and Kleckner, 1995). In the latter study, the data obtained by 2-D agarose gel was confirmed by EM. The observation that different recombination intermediates are accumulated in different yeast strains undergoing DSB repair, emphasises the importance of studying these events in as many different systems as possible in order to gain a complete picture of the events that arise during recombination.

1.5 Palindromes as a cause of genomic instability

Intra-strand secondary structures formed by inverted repeats have been implicated in genomic instability and DSB formation (Leach, 1994). The repair of these DSBs, known as secondary structure repair, is thought to play a major role in genome stability in higher organisms. The instability that arises from ill-processed inverted repeat sequences has been associated with severe diseases such as Ataxia telangiectasia-like disorder or Nijmegan breakage syndrome. Many of these conditions are also linked to cancer predisposition (McKinnon and Caldecott, 2007). The *S. cerevisiae* Mre11/Rad50/Xrs2 (MRX) and human Mre11/Rad50/Nbs1 (MRN) protein complexes are required for secondary structure repair (Lewis and Cote, 2006). Mre11, a 3′→5′ double stranded exonuclease that is essential for the correct processing of hairpin capped dsDNA breaks thought to form as a result of nicked

palindromes, is particularly important in secondary structure repair (Lobachev et al., 2002).

Until recently, studying secondary structure repair using the model organism *E. coli* has proved difficult primarily because of the intrinsic instability of repeat sequences in this organism (Collins, 1981; Lilley, 1981). A study that analysed the fate of a 246 bp interrupted palindrome that was introduced into *E. coli* by λ phage lysogenisation, identified SbcCD, the *E. coli* homologue of the Mre11/Rad50 protein complex, as being responsible for generating a DSB at the site of the palindrome, which was then repaired by HR (Cromie et al., 2000). *In vitro* work indicated that SbcCD could cleave a hairpin structure near the 5′ junction of the loop and the dsDNA hairpin stem (Connelly et al., 1998). Based on this data it was suggested that during replication the palindrome was extruded into a hairpin that was recognised and processed by SbcCD. It was proposed that a blunt DNA free end would be generated downstream of the palindrome and a DNA free end with a single-strand tail would be generated upstream (Cromie et al., 2000). It was also suggested that these two ends might be processed differently as survival depended on both RecBCD and RecFOR. In 2008 the 246 bp interrupted palindrome was stably inserted into the *E. coli* chromosomal *lacZ* gene (Eykelenboom et al., 2008). As expected, SbcCD-mediated cleavage of the palindrome generates a two-ended DNA DSB. It was confirmed that this cleavage was dependent on chromosomal replication, and survival on RecBCD-mediated HR. PriA was also shown to be essential, indicating that the establishment of origin-independent replication was also required for repair. Contrary to what was observed in the λ phage system, RecFOR was dispensable (Fig. 1.9).

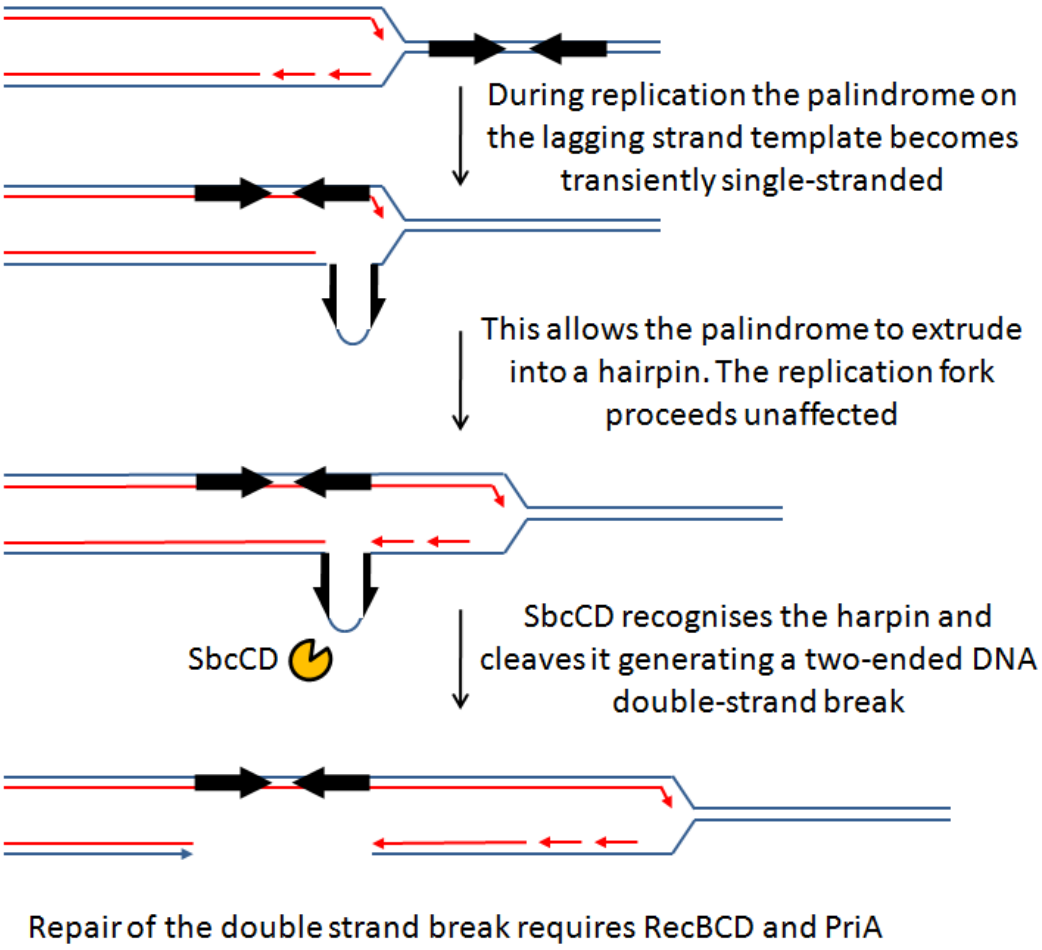

Figure 1. 9 SbcCD-mediated cleavage of a 246 bp interrupted palindrome in *E. coli*

During chromosomal replication, the palindrome on the lagging strand template becomes transiently single-stranded and is extruded into a hairpin. This structure is recognised by SbcCD, which cleaves it to generate a two-ended DNA DSB. Repair of this DSB is dependent on RecBCD-mediated HR and requires the establishment of origin-independent replication initiated by PriA.

1.6 Scope of this thesis

The work presented in this thesis addresses the nature of the recombination intermediates that arise during the repair of a chromosomal DSB in *E. coli*. The unique SbcCD/palindrome system of DSB formation was used to generate a chromosomal site-specific DSB in only one sister chromosome (Eykelenboom et al., 2008). The recombination intermediates generated by the repair of this DSB were analysed using a variety of DNA electrophoretic techniques. It is hoped that this system will serve as a model for mitotic DSB repair in eukaryotes thereby shedding further light on this process. Additionally, the work attempts to understand what role is fulfilled by the *E. coli* recombination helicase, RecG.

Chapter three presents data regarding the viability of Δ*ruvAB*, Δ*recG*, Δ*ruvAB* Δ*recG*, and Δ*ruvC* mutants, which are strains unable to resolve recombination intermediates, and highlights some of the generic requirements for surviving a chromosomal DSB in *E. coli*. In chapter four, chromosomal DNA from Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains was analysed by pulsed-field gel electrophoresis where the abundance and distribution of branched DNA (recombination intermediates) was studied. Chapter five focuses on understanding the three-dimensional structure of these recombination intermediates by isolating them by native 2-D agarose gel electrophoresis.

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Stock solutions

20 % (w/v) Arabinose

Made up to 20 % (w/v) in distilled water and autoclaved. Stored at room temperature.

20 % (w/v) Glucose

Made up to 20 % (w/v) in distilled water and autoclaved. Stored at room temperature.

20 % (w/v) Sucrose

Made up to 20 % (w/v) in distilled water and autoclaved. Stored at room temperature.

80 % (v/v) Glycerol

Made up to 80 % (v/v) in sterile Milli-Q water and autoclaved. Stored at room temperature.

2.5 M $CaCl_2$

Made up to 2.5 M in sterile Milli-Q water. Sterilised using a 0.2 µm syringe filter. Stored at room temperature. Used at 2.5 mM and 0.1 M.

Ethidium bromide (EthBr) 10mg ml^{-1} (Fulka Biochemika)

Stored in the dark at room temperature. Diluted to 0.5 µg ml^{-1} in sterile Milli-Q water when used for staining of electrophoresis gels for the visualisation of DNA under a UV light source, unless otherwise stated.

Antibiotics

All antibiotics (Table 2.1) were dissolved in Milli-Q water, unless otherwise stated, and stored at -20 °C.

Table 2. 1 Antibiotics

Antibiotic/Abbrev.	Stock concentration (mg ml^{-1})	Working concentration (µg ml^{-1})	Source

Chloramphenicol (Cm)*	50	50	Calbiochem
Kanamycin (Km)	50	50	Sigma
Spectinomycin (Spec)	25	50	Sigma

* Stored in 95 % ethanol

2.1.2 Culture media

All solutions were made up to the required volume in distilled water and autoclaved.

4 X M9 salts	**1 Litre**
88.5 mM KH_2PO_4	12g
197 mM Na_2HPO_4	28g
34 mM NaCl	2g
75 mM NH_4Cl	4g

L-broth	**1 Litre**
Bacto-tryptone (Difco)	10g
yeast extract (Difco)	5g
NaCl	10g

pH adjusted to 7.5 with NaOH.

LB agar

15 g Bacto-agar was added to 1 L of L-broth prior to autoclaving.

LC agar	**1 Litre**
Tryptone	10g
Yeast extract	5g
NaCl	5g
Difco-agar	10g

pH adjusted to 7.2 with NaOH.

LC top agar	**1 Litre**
Tryptone	10g
yeast extract	5g
NaCl	5g
Difco-agar	7g

pH adjusted to 7.2 with NaOH.

Phage buffer	**1 Litre**
22 mM KH_2PO_4	3g
49 mM Na_2HPO_4	7g
85 mM NaCl	5g
1 mM $MgSO_4$	0.1 g
1 mM $CaCl_2$	0.1 g

1% (w/v) gelatine	10g

2.1.3 Buffers

5 X Tris-borate (TBE)	**1Lt**
0.89 M Tris Base	53 g
0.89 M Boric acid	27.5g
10 mM EDTA	3.4g
pH adjusted to 8.0 with HCl	

20 X SSC	**1Litre**
3 M NaCl	175.2 g
300 mM Tri-sodium citrate	77.4 g
pH adjusted to 7.0 with HCl	

20 X SSPE	**500 ml**
3 M NaCl	87 g
200 mM NaH_2PO_4	12 g
20 mM EDTA	20 ml of 0.5 M solution at pH 8.0

50 X Tris-acetate (TAE)	**1 Litre**
2 M Tris-base	242 g
0.95 M Glacial acetic acid	57.1 ml

0.05 M EDTA	14.6 g

Alkaline transfer buffer	**1 Litre**
0.5 M NaOH	20 g
10 X SSC	500 ml of 20 X SSC

Church-Gilbert buffer	**20 ml**
7% SDS	14 ml of a 10% solution
0.5 M NaH_2PO_4	5 ml of a 2 M solution at pH 7.2
1 mM EDTA	40 µl of a 0.5 M Solution at pH 8.0
1% BSA	0.2 g

The NaH_2PO_4, EDTA, BSA, and 960 µl of distilled water were mixed until the BSA was completely dissolved. The SDS was then added and the whole was heated to allow for easy mixing. The warm solution was filter sterilised using a filter with an 0.4 µm pore size.

Depurination solution	**500 ml**
0.25 M HCl	12.5 ml of a 37 % solution

Detection buffer	**500 ml**
0.1 M Tris-HCl	6.055 g
0.1 M NaCl	2.922 g

The pH was adjusted to 9.5 with HCl

EDTA	**1 L**
0.5 M EDTA	186.12 g
The pH was adjusted to 8.0 with NaOH	

High stringency buffer	**500 ml**
0.1 X SSC	2.5 ml of 20 X SSC
0.1 % SDS	5 ml of 10 % SDS

Low stringency buffer	**500 ml**
2 X SSC	50 ml of 20 X SSC
0.1 % SDS	5 ml of 10 % SDS

Maleic acid buffer	**500 ml**
0.1 M Maleic acid	5.8 g
0.15 M NaCl	4.35 g
pH adjusted to 7.5 with NaOH	

Maleic acid washing buffer	**500 ml**
Maleic acid buffer	499.7 ml
0.3 % Tween	300 µl of 100 % Tween

NDS buffer	**500 ml**
0.5 M Na_2.EDTA	93 g

10 mM Tris-Base	0.6 g
0.6 mM NaOH	11 g
34 mM N-lauroyl sarcosine	5 g

The Na_2.EDTA, Tris-Base, and NaOH were dissolved in 350 ml of distilled water. Separately, the N-lauroyl sarcosine was completely dissolved in 50 ml of distilled water and then added to the main solution. The pH was adjusted to 8.0 with NaOH and the total volume was brought up to 500 ml.

Stringency washing solution	**500 ml**
0.5 X SSC	12.5 ml of 20 X SSC
0.1% SDS	5 ml of 10% SDS

Stripping buffer	**50 ml**
50% Formamide	25 ml of 100% Formamide
5 X SSPE	12.5 ml of 20 X SSPE

TE buffer	**1 Litre**
10 mM Tris	1.2 g
1 mM EDTA	0.3 g
pH adjusted to 7.4 with HCl	

TEN buffer	**1 Litre**

50 mM Tris	6.1 g
50 mM EDTA	14.6 g
100 mM NaCl	5.8 g
pH adjusted to 8.0 with HCl	

Washing solution	**500 ml**
2 X SSC	50 ml of 20 X SSC
0.1% SDS	5 ml of 10% SDS

2.2 Methods

2.2.1 Bacterial methods

Bacterial stocks stored at -80 °C

0.75 ml of an overnight culture was mixed with 0.75 ml of 80 % (v/v) glycerol and placed in a 1.5 ml Eppendorf tube. The tube was vortexed, sealed with a strip of parafilm, and stored at -80 °C.

Overnight cultures

5 ml of L-broth with the required additives were inoculated with a single colony derived from the -80 °C stock. Cultures were incubated overnight at 37 °C with shaking (120 rpm).

Serial dilutions

A culture was diluted in 1X M9 salts to an optical density (OD_{600nm}) of 0.4. This concentration was considered as the 10^0 samples, which were further diluted in 1X M9 salts in steps of 10^{-1} until a dilution of 10^{-5}, unless otherwise stated.

UV light irradiation of cultures

In order to test whether strains were UV sensitive, overnight cultures were serially diluted (from 10^0 to 10^{-5}) and 5 µl of each dilution was spotted onto LB + 0.5% glucose agar plates. This was repeated 4 times. Plates were exposed to a UV light dose of 0, 100, 200 or 300 J/m^2 using a Stratagene UV Stratalinker™ 1800. Plates were then covered in tin foil to keep out of the light, as UV light irradiated cells can repair the damage by photoreactivation in the presence of light, and placed at 37 °C overnight.

Transformation of *E. coli* by $CaCl_2$ treatment followed by a heat shock

An overnight culture of the strain to be transformed was made. This was diluted 50 times in 25 ml L-broth + 0.5 % glucose and put at 37 °C with shaking (120 rpm) for 2 hours. 2 ml of culture were then spun down in a table-top centrifuge at maximum speed for 2 minutes. The L-broth was removed and 800 µl of 0.1 M $CaCl_2$, previously chilled on ice, were added. The cell pellet was re-suspended and the solution was put on ice for 30 minutes. The culture was centrifuged as previously described and the 800 µl of 0.1 M $CaCl_2$ were replaced with 100 µl of fresh, chilled, 0.1 M $CaCl_2$, and 0.5 µl of DNA (when using a plasmid isolated by mini-prep) or 4 µl of DNA (when transforming with a newly ligated plasmid obtained from an *in vitro* ligation

mix). The pellet was re-suspended in the 0.1M $CaCl_2$ + DNA mix and the tube placed on ice for a further 30 minutes. The cells were then heat shocked in a water bath at 37 °C for 5 minutes, briefly placed back on ice, and then mixed with 400 µl of fresh L-broth with no selection. The cells were allowed to recover at 37 °C (30 °C when a temperature sensitive plasmid was used for the transformation) with shaking (120 rpm) for 1 hour and then plated on LB + 0.5 % glucose agar plates with the appropriate selection added. For every transformation, a no DNA control was prepared in parallel.

Plasmid mediated gene replacement

In order to generate targeted chromosomal alterations, a plasmid-based technique, plasmid-mediated gene replacement (PMGR), first described by Link and collaborators, was used (Link et al., 1997) (Fig. 2.1). The plasmid, pTOF24, has a temperature sensitive replication initiator protein (*repA*101^{TS}), two positive selection markers in the form of a chloramphenicol resistance gene (*cat*) and a kanamycin resistance gene (*aph*), and a gene encoding for a levan sucrase (*sacB*) that can act as negative selection as *E. coli* strains expressing this gene become sensitive to sucrose (Merlin et al., 2002). The kanamycin resistance gene is flanked by PstI/SalI

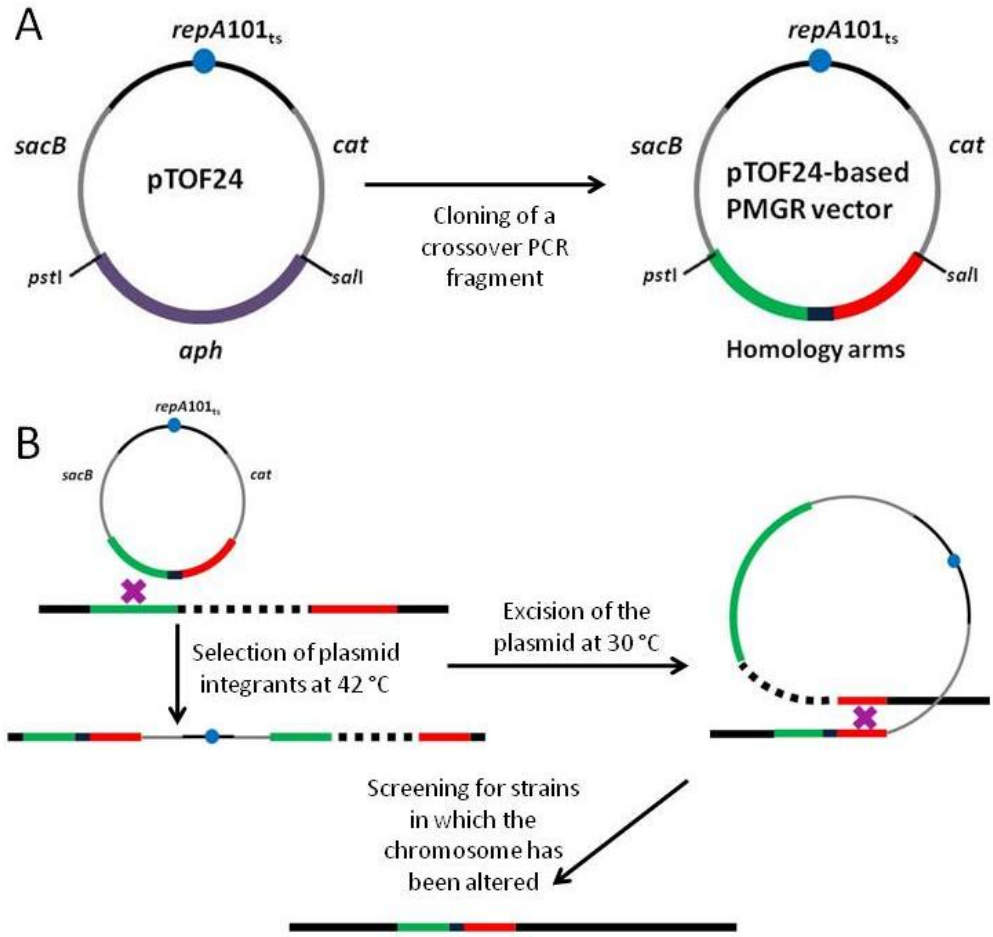

Figure 2. 1 Plasmid mediated gene replacement (PMGR).

(A) Construction of pTOF24-derivatives containing products from a crossover PCR that are cloned into the PstI/SalI locus of the plasmid. *repA101*$_{TS}$ codes for a temperature sensitive replication initiator protein (the permissive temperature being 30 °C and the non-permissive temperature being 42 °C). *aph* and *cat* code for kanamycin and chloramphenicol resistance, respectively. *sacB* codes for a levan sucrase, which converts sucrose into a toxic product for *E. coli* and is used as a negative selection marker. (B) Utilisation of these pTOF24-derivatives for targeted chromosomal modifications in E. coli. At 42 °C the plasmid can only be replicated if it integrates into the chromosome. pTOF24-derivatives are designed to contain two regions of homology to the chromosome. Integration at 42 °C will occur by RecA-mediated homologous recombination between one of the two regions of homology and the same region in the chromosome. This will result in integration of the entire plasmid sequence. Growing the integrant at 30 °C in liquid culture will allow the plasmid to excise from the chromosome. This also occurs by RecA-mediated homologous recombination. If integration occurs at the first region of homology (red) and excision occurs at the second (green), the wild type region of the chromosome is replaced with the modified DNA insert from the plasmid.

restriction sites, which allows for the replacement of the *aph* gene with a cloning fragment of interest. Precise gene alterations, flanked by ~ 400 bp of homology to the surrounding chromosome, generated by cross-over PCR (see section 2.2.2, Fig. 2.2), were cloned into pTOF24 using the PstI/SalI cloning sites. Transformants were screened for loss of *aph* by the acquisition of kanamycin sensitivity. The strain to be modified was transformed with pTOF24-derivatives using chloramphenicol resistance as a selectable marker. Following the transformation, transformans were recovered by incubating plates at 30 °C so as to ensure autonomous replication of the plasmid. Successful transformants were then streaked on fresh LB-agar + chloramphenicol and placed at 42 °C to select for strains in which the plasmid had integrated into the chromosome at one of the homology arms. This was repeated a second time to ensure purity of the integrants. In order to complete PMGR, the plasmid had to be excised from the chromosome. This was done by culturing individual integrants in 5 ml LB with no selection, overnight, at 30 °C with shaking. The culture obtained, containing a mixture of cells that had either lost or retained the plasmid, was serially diluted from 10^{-1} to 10^{-6} in 1 X M9 media. 100 µl of dilutions 10^{-4}, 10^{-5}, and 10^{-6} were plated onto LB-agar containing 5 % sucrose to select against all cells that had retained the plasmid. Colonies were checked for chloramphenicol sensitivity to confirm the loss of the plasmid. Sucrose resistant/chloramphenicol sensitive colonies were checked for integration of the cloning fragment of interest by boiled colony PCR. It was expected that about 50 % of the colonies would retain the wild type DNA sequence and 50 % would acquir the DNA alteration. Once colonies that generated the expected PCR product size were identified, the chromosomal modification was confirmed by sequencing and

by exposure to UV light (if the alteration conferred sensitivity to DNA damage induced by UV light).

P1 phage general transduction

An alternative to PMGR is P1 phage general transduction. The P1 bacteriophage is a general transducing phage and *E. coli* can act as host for its propagation. During its lytic growth cycle, P1 has the tendency to miss-package host chromosomal DNA into its head. This ability can be used to transfer regions of the *E. coli* chromosome from one strain to another (~ 100 Kb). The miss-packaged *E. coli* chromosomal fragment enters the recipient strain as a linear fragment. It can then replace the homologous region on the chromosome by RecA-mediated homologous recombination. The acquisition of the new DNA can be selected for. If it is not possible to directly select for the region of interest, it can be linked to a selectable marker, such as an antibiotic resistance gene, as co-transduction of two linked markers is a frequent event (Lennox, 1955).

P1 lysate preparation

An overnight culture of the *E. coli* strain with the desired marker to be transduced was prepared. This was diluted 10 times in 10 ml L-broth containing 2.5 mM $CaCl_2$ and allowed to grow at 37 °C with shaking (120 rpm) for 2 hours. 200 µl samples of the mid-Log phase cells were taken and mixed with 100 µl of P1 lysate serially diluted in phage buffer, in dilution steps of 10^{-1}, from a concentration of 10^{0} to 10^{-5}. These cultures were put back at 37 °C with shaking for a further 30 minutes to allow for phage adsorption. 2.5 ml of melted LC Top agar + 5 mM $CaCl_2$ were then added

to each sample. These solutions were poured onto LC agar plates + 5 mM $CaCl_2$. The plates were incubated overnight at 37 °C without inversion. The top layer of agar was then covered in 5 ml phage buffer and scraped off into a sterile, detergent-free, 5 ml bottle using a sterile glass pipet. 100 µl of chloroform was added and the whole was vortexed and left at room temperature for 30 minutes. The samples were centrifuged at top speed for 5 minutes and the supernatant was transferred to into a sterile, detergent-free, 5 ml bottle. Lysates were stored at 4 °C.

P1 transduction

An overnight culture, supplemented with 2.5 mM $CaCl_2$, of the strain to be genetically altered was prepared. Four 1 ml samples of the cultures were spun down using a table-top micro centrifuge for 1 minute at top speed. The pellets were re-suspended in 100 µl L-broth + 2.5 mM $CaCl_2$ and mixed with 0 µl, 1 µl, 10 µl and 100 µl of the phage lysate prepared on the desired donor. Samples were incubated at 37 °C for 20 minutes in order to allow for phage adsorption. 800 µl of L-broth + 2 mM Sodium Citrate were added to stop further infection by the P1 phage. Incubation at 37 °C for 1 hour allowed for expression of the selection marker, after which 100 µl of the cultures were plated out onto LB + 0.5 % glucose agar plates containing the appropriate selection.

Growth curves

Overnight cultures were diluted 50 times in fresh L-broth and allowed to grow at 37 °C with shaking to an OD_{600nm} of 0.1. This was considered as time point 0 for the growth curve. Cultures were then split and either 0.5 % glucose or 0.2 % arabinose

was added and the cultures were put back at 37 °C. The OD_{600nm} was measured every hour and cultures were kept at an OD_{600nm} below 1.0 by regular dilution (1/5) in fresh, pre-warmed to 37°C, L-broth (containing the appropriate sugar) to keep cells in exponential phase.

Full viability test

Overnight cultures were diluted 50 times in fresh L-broth and allowed to grow at 37°C with shaking until they reached early exponential phase (OD_{600nm} 0.2-0.4). This was considered as time point 0. Cultures were split and either 0.5 % glucose or 0.2 % arabinose was added and the cultures were put back at 37°C with shaking for 1 hour. At time point 0, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, cultures were serially diluted from 10^{0}-10^{-5}. 50 µl of the appropriate dilution were plated onto LB + 0.5 % glucose. This was repeated 3 times for each condition. Plates were put at 37 °C overnight and the next day colonies were scored.

2.2.2 DNA Cloning techniques

Genomic DNA extraction for PCR (Promega kit)

The Promega Wizard® Genomic DNA purification kit was used following the manufacturer's instructions. DNA was re-hydrated in MQ-water at 65 °C for 1 hour or at 4 °C overnight, and stored at -20 °C.

Genomic DNA preparation from a single colony – Boiled colony PCR

A single colony was picked from a plate and re-suspended in 30 µl of MQ-water in a PCR tube. This was boiled for 10 minutes in a thermocycler to lyse the cells and release the DNA. The whole was centrifuged for 3 minutes at maximum speed in order to precipitate cell debris. 2 µl of the supernatant was used per 50 µl of a PCR reaction.

Plasmid DNA preparation for PCR (QIAGEN kit)

The QIAGEN QIAprep® Spin Miniprep Kit was used following the manufacturer's instructions. DNA was eluted in 30 µl of MQ-water and stored at -20 °C.

Polymerase Chain Reaction (PCR)

Finnzymes Phusion® High-Fidelity DNA Polymerase Cat. No. F-530 is a highly processive and extremely accurate DNA polymerase (with an error rate of 4.4×10^{-7}). Because of these qualities it was chosen as the polymerase for all PCR reactions when the product was required for cloning or as a template for labelling with ^{32}P. When PCR reactions were carried out for checking fragment sizes, for example following PMGR, Promega GoTaq® Flexi DNA Polymerase, Cat. No. M829, was used instead. Reactions were carried out using a PeqLab Biotechnologie GmbH peqSTAR 96 Universal Gradient PCR machine. Primer annealing temperatures were dependent on primer sequence and were altered accordingly. Typically, the annealing temperature increases as the primer length and % GC content increase. The extension time was determined by the polymerase of choice and the length of the template to be amplified. Typically, the longer the template, the longer the extension time. Phusion® High-Fidelity DNA Polymerase can extend a 1Kb

fragment in 30 seconds, while GoTaq® Flexi DNA Polymerase takes twice as long, requiring 1 minute per 1Kb of template.

A typical cycle programme was as follows:

Initial template denaturation :	95 °C	5 min	
Template denaturation :	95 °C	30 sec	} 30 X
Primer annealing :	50-65 °C	30 sec	
Extension:	72 °C	30 sec-3 min	
Final extension:	72 °C	10 min	
Storage:	8 °C	indefinite	

Crossover PCR

Crossover PCR (Fig. 2.2) was used in order to join two separate fragments of DNA without the need for restriction and ligation. For this technique, four primers were required, two for amplifying the first DNA fragment and two for amplifying the second DNA fragment. Internal primers, the reverse primer for the first DNA fragment and the forward primer for the second DNA fragment, were designed to have 20-25 base pairs of homology to each other. Initially, the two DNA fragments were amplified separately, creating two products that at one extremity contained 20-25 base pairs of homology to each other. Finally, to join the two DNA products, a crossover PCR was set up where the two products from the initial PCR reactions were used as template. During melting and cooling of the templates, the region of homology would bring the two DNA fragments together to create a single new fragment for amplification using the external primers, the forward primer for the

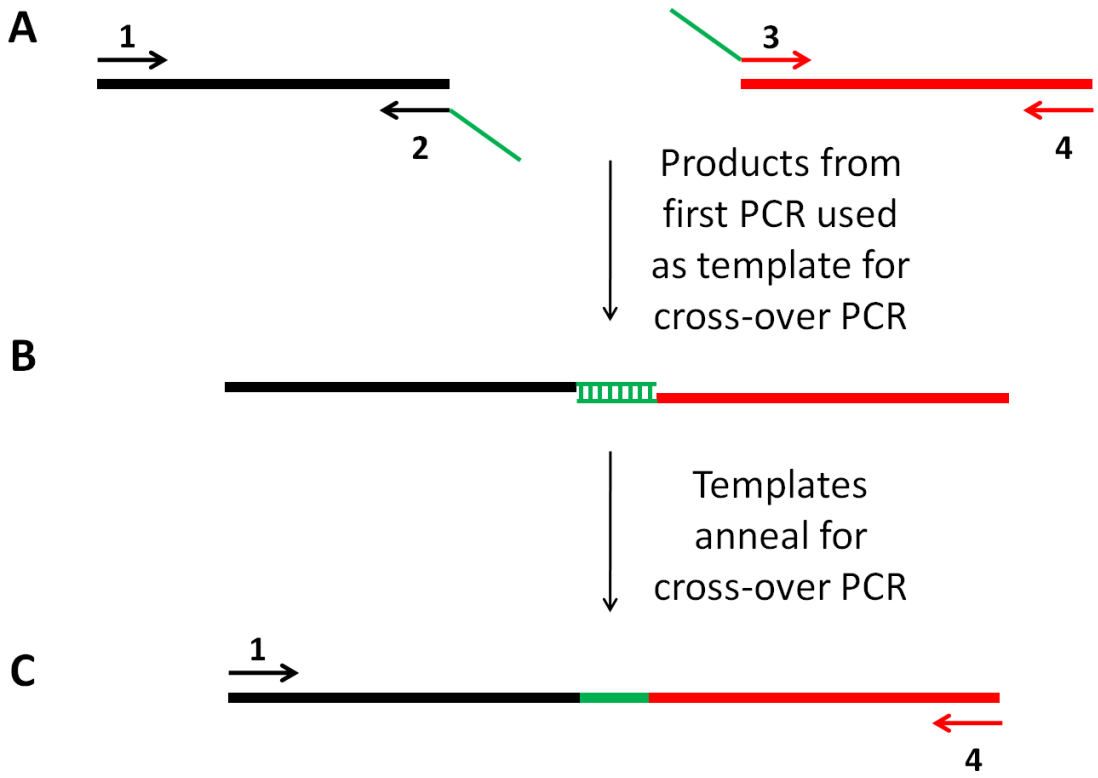

Figure 2. 2 Crossover PCR.

(A) Initially, the two DNA fragments (shown in black and red) are amplified separately using primers 1-4. Primers 2 and 3 are designed to have a 20-25 bp region of homology to each other. (B) In the crossover PCR, the products from the first PCR reaction are used as template. When these melt and re-anneal, the region of homology between them will bring them together. (C) This will create a new, single, template for amplification using primers 1 and 4.

first DNA fragment and the reverse primer for the second DNA fragment. The resulting PCR product would be a fusion of the two DNA fragments of interest.

Sequencing of DNA (Applied Biosystems kit)

Sequencing was carried out using the Applied Biosystems BigDye® Terminator v3.1 Cycle-Sequencing Kit following manufacturer's instructions. Template DNA consisted of a purified PCR product. Sequencing reactions were analysed by the SBS Sequencing Service, Ashworth Laboratories, University of Edinburgh, using an ABI PRISM® 3100-Avant Genetic Analyser.

PCR product purification for cloning (QIAGEN kit)

The QIAGEN QIAquick® PCR purification kit (Cat. No. 28104) or The QIAquick® Gel Extraction Kit (Cat. No. 28704) was used for cleaning DNA fragments for cloning. The manufacturer's instructions were followed. DNA was eluted in 30 µl of MQ-water and stored at -20 °C.

Restriction digestion of PCR purified DNA

Restriction enzymes and buffers were obtained from New England Biolabs (NEB). 30 µl of either PCR purified DNA or plasmid DNA were digested following manufacturer's instructions. Samples were incubated at the optimum digestion temperature (37 °C for all enzymes used in this work) for 2-4 hours.

Ligation of DNA fragments

To ligate fragments of DNA, the New England Biolabs Quick Ligation™ Kit was used following manufacturer's instructions. Ligation reactions were performed in a total of 20 µl per reaction volume.

Agarose gel electrophoresis of PCR products or plasmid DNA

DNA fragments from PCR reactions or digested plasmids were separated on either a 1% (w/v) agarose gel. The appropriate amount of agarose (MELFORD agarose electrophoresis grade Cat. No. MB1200) was dissolved in 1 X TAE and allowed to cool to 55 °C. Safeview (NBS Biologicals Ltd, SafeView, Cat. No. NBS-SV1; 5 µl in 100 ml of liquid agarose) was added to visualise the DNA under UV light. Gels were run from 90-140 V for up to 2 hours and DNA was visualised using a UV box (BioRad). The size of fragments was checked using DNA ladders (New England Biolabs, 1 Kb DNA Ladder Cat. No. N3232, 100 bp DNA Ladder Cat. No. N3231). When necessary, DNA was quantified using a Nanodrop (ND-1000 v3.5).

2.2.3 Agarose gel electrophoresis of Chromosomal DNA

2.2.3.1 Preparing *E. coli* DNA in agarose plugs

An overnight culture of the desired strain was diluted 50 times in L-broth and allowed to grow at 37 °C with shaking to an OD_{600nm} of 0.4-0.8. Cultures were then split in two and 0.5 % glucose was added to one flask (for the repression of the

arabinose-indusible promoter), whereas 0.2 % arabinose was added to the other flask (for the induction of the arabinose-inducible promoter). Flasks were put back at 37 °C for the desired period. Cultures were harvested and washed twice in TEN buffer by spinning and re-suspending. After the final wash cells were re-suspended in TEN buffer to give an OD_{600nm} of 4 (for samples to be analysed by pulsed-field gel electrophoresis) or OD_{600nm} of 80 (for samples to be used in 2-D gel analysis). The cell suspension was briefly warmed to 37 °C and then mixed with an equal volume of agarose (invitrogen UltraPure™ LMP Agarose Cat. no. 16520-050) at 2% in 1 X TEN for pulsed-field gel electrophoresis samples and 0.8% in 1X TEN for 2-D gel samples, giving a final agarose concentration of 1% and 0.4%, respectively. The cell and agarose mix was then poured into plug moulds (Bio-Rad Cat. no. 1703706), and allowed to set at 4 °C for 30 minutes. Once set, all 10 plugs were extruded from the moulds into a falcon tube and 1 ml of NDS + proteinase K (1 mg ml^{-1}) was added per plug of sample. Tubes were left overnight at 37 °C with gentle rocking. The buffer was replaced with fresh 1 ml NDS + proteinase K (1 mg ml^{-1}) per plug of sample for a second night of incubation at 37 °C with gentle rocking. Plugs were stored at 4 °C in fresh NDS without added proteinase K.

2.2.3.2 Digestion of DNA set in agarose plugs

In order to remove any remaining proteinase K and NDS solution, which would inhibit the restriction enzyme, plugs were washed thoroughly in 1.5 ml per plug of

1 X restriction buffer (without BSA and DTT) for 6 hours, making sure to add fresh 1 X restriction buffer every hour. Once washed, the plugs were digested in 500 µl 1 X restriction buffer + BSA + DTT, using between 10U-100U of enzyme per plug (for pulsed-field gel electrophoresis samples) or 500U of enzyme per plug (for 2D-agarose gel electrophoresis samples). Digestions were left overnight at 37 °C with gentle rocking. Plugs were briefly cooled to 4 °C before quickly washing them in 1.5 ml TE just prior to loading onto the gels.

2.2.3.3 Pulsed-field gel electrophoresis (PFGE)

PFGE was used for the separation of large chromosomal DNA fragments as it can resolve fragments above 30 Kb and as big as 1.5 Mb. Chromosomal DNA was prepared in agarose plugs and digested with the relevant restriction enzyme. Plugs and markers were then attached to the PFGE comb using 10 µl of liquid agarose (1% (w/v) high-strength agarose gel (AquaPor™ ES; Fisher catalogue number ELR-300-040F) in 0.5 X TBE) and allowed to set at 4 °C for 30 minutes. The remainder of the agarose was then poured around the plugs and allowed to set at 4 °C for 30-60 minutes. The gel was run in 0.5 X TBE at 6 V/cm, 4 °C using a CHEF-DR™ II PFGE. Switch time was set to 5-30 seconds with an included angle of 120°. Running times started at 8 hours for SacI digested plugs and increased to 10 hours for SalI digested plugs and 17 hours for NotI and IsceI digested plugs. After running, gels were stained with 0.5 µg ml^{-1} EthBr for 30 minutes and viewed under UV light and then processed for Southern blotting.

2.2.3.4 Native two dimensional agarose gel electrophoresis

In order to analyse intermediates of double strand break repair, native two-dimensional agarose gel electrophoresis (2-D agarose gel) was used to separate branched DNA structures from their linear counterpart. Initially, the chromosomal DNA was separated in conditions that minimised the structural difference between fragments but maximised the difference in size. Thereafter, the lane containing the DNA was sliced out and turned 90°, placing the wells to the left, and run in a second dimension in the presence of 0.3 μg ml^{-1} EthBr, so as to maximise the difference in shape between different structures. For separation in the first dimension, chromosomal DNA was prepared in agarose plugs and digested with the relevant restriction enzyme. Plugs were then attached to the gel comb, making sure to leave at least 1 lane gap between samples, using 10 μl of liquid agarose (0.4% (w/v) MELFORD agarose electrophoresis grade (Cat. No. MB1200) in 1 X TBE). This was allowed to set at 4 °C for 30 minutes. The remainder of the agarose was then poured around the plugs and allowed to set at 4 °C for 30-60 minutes. 2.5 μg of NEB 1 Kb DNA Ladder (Cat. No. N3232) was loaded onto the gel, which was run in 1 X TBE at 1 V/cm and 4 °C for 26 hours. The Marker lane was cut out of the gel and stained with 0.5 μg ml^{-1} EthBr for viewing. Intermediates run at a higher molecular weight than their linear counterparts and therefore the lane in which the DNA was run was sliced out and cut 1 cm below where the linear species of the DNA of interest was expected to have migrated to. This gel slice was turned 90°, placing the

wells to the left, and placed into the casting tray for the second dimension. The second dimension agarose (1% (w/v) MELFORD agarose electrophoresis grade Cat. No. MB1200 in 1 X TBE + 0.3 μg ml^{-1} EthBr) was then poured over the 1st dimension slices so as to cover them completely. The whole was allowed to set at 4 °C for 30-60 minutes and was then run for 10 hours in 1 X TBE + 0.3 μg ml^{-1} EthBr, at 6 V/cm and 4 °C. During this time the running buffer was re-circulated so as to maintain a constant concentration of EthBr across the gel and tank. The gel was then exposed to UV light to view the DNA arcs and then processed for Southern blotting.

2.2.4 Southern blotting of DNA

2.2.4.1 Alkaline transfer of DNA to a positively charged nylon membrane

Once viewed under UV light, the gel was washed in depurination solution for 30 minutes to fragment the DNA for easier transfer (pulsed-field gels were not depurinated as this process was previously shown to reduce the transfer of DNA separated using this technique). The gel was then washed in alkaline transfer buffer for 1 hour and the transfer stack was set up and the DNA was allowed to transfer overnight (Fig. 2.3). After transfer the nylon membrane was allowed to dry, completely, at room temperature, and then was UV cross-linked, using a Stratagene UV Stratalinker™ 1800. Membranes obtained from Pulsed-field gel electrophoresis were crosslinked using 1200 J/m^2 while membranes derived from 2-D agarose gel

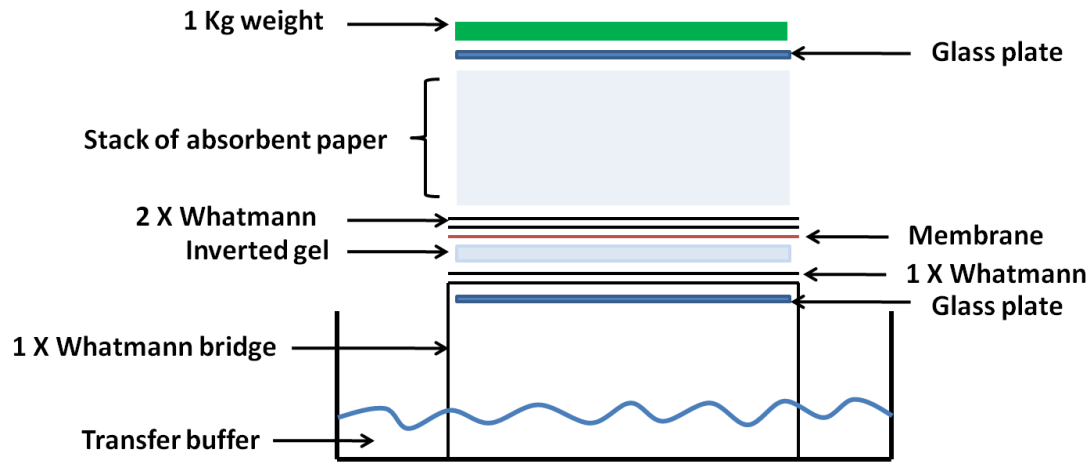

Figure 2. 3 Southern blot transfer stack.

A glass plate was placed over a plastic container filled with transfer buffer and a strip of Whatmann paper was placed over this like a bridge, making sure both ends of the strip came into contact with the buffer. A piece of Whatmann paper, the size of the gel, was placed on top of the Whatmann bridge an+d the inverted gel was stacked next, followed by the membrane, pre-wetted in transfer buffer for 5 minutes, and two more pieces of Whatmann paper. Finally a stack of absorbent paper was placed on top followed by another glass plate. The whole stack was placed under a 1 Kg weight. The DNA transferred from the gel to the membrane overnight by capillary action.

electrophoresis were cross-linked using 1000 J/m^2. Once crosslinked, membranes were washed in distilled water, dried, and sealed in a plastic envelope between two pieces of whatmann paper for storing at 4 °C.

2.2.4.2 Labelling of a probe with digoxygenin-11-dUTP

A Roche PCR DIG probe synthesis kit (Cat. No. 11636090910) was used for DNA labelling with digoxygenin-11-dUTP. This was carried out by PCR where one of the nucleotides available to the polymerase (UTP) is labelled, with a steroid hapten, which is incorporated into the PCR product. The template for the reaction was the product of a previous PCR, obtained by amplifying chromosomal DNA using the same primers as those used for the labelling PCR. This was purified by PCR purification prior to being used as a template in the labelling reaction. In parallel, a standard PCR with un-labelled dNTPs was also set up as a control. Once both reactions were finished, their migration patterns were compared by agarose gel electrophoresis as incorporation of digoxygenin-11-dUTP causes DNA to migrate slower. When not in use, probes were stored at -20 °C.

2.2.4.3 Chemiluminescent detection of a DIG-labelled probe

A Roche DIG Luminescent Detection Kit (Cat. No. 11363514910) was used to detect a digoxygenin-labelled probe bound to DNA on a nylon membrane. The crosslinked membrane was placed in a hybridisation bottle, in a hybridisation chamber, at 47

°C. 12 ml of pre-heated DIG easy hybridisation buffer, at 47 °C, were then added and the membrane was pre-hybridised for 30 minutes. During this time, 50 µl of MQ-H_2O were added to 7 µl of DIG-labelled probe and the probe was boiled in a thermocycler for 5 minutes and then put directly on ice to snap-freeze. Once pre-hybridisation was over, the pre-hybridisation buffer was replaced by 5 ml of fresh pre-heated DIG easy hybridisation buffer to which the 57 µl of snap-frozen probe were added. The probe was allowed to hybridise to the membrane overnight. The membrane was then removed from the bottle and placed in a plastic container. It was washed 2 X for 5 minutes in low stringency buffer at room temperature on a rocker. During this time the hybridisation oven was set to 68 °C and the hybridisation bottle rinsed and placed back in the oven. The membrane was put back in the oven and washed 2 X for 15 minutes with high stringency buffer. The membrane was then placed back in the plastic container and washed for 2 minutes in maleic acid washing buffer. 100 ml of blocking solution (diluted 1/10 in maleic acid) replaced the maleic acid washing buffer. The membrane was allowed to block, with rocking, for 30 minutes to 3 hours to prevent unspecific binding of the anti-digoxygenin-AP. Once blocking was over, 100 ml of fresh blocking solution with added 3 µl of anti-digoxygenin-AP (the antibody against the digoxygenin is fused to alkaline phosphatise) were added to the tray. The antibody was allowed to bind for 30 minutes with rocking. The membrane was then washed in 100 ml of maleic acid washing buffer for 2 X 15 minutes and finally in 40 ml of detection buffer for 3 minutes. A 1/100 dilution of a chemiluminescence substrate for alkaline phosphatase (CSPD) (Cat. No. 11655884001) was then layered over the membrane and the whole was allowed to sit at room temperature for 5 minutes. The excess

liquid was then squeezed out and the membrane was sealed and placed at 37 °C for 10 minutes in order to allow the alkaline phosphatise to react with the CSPD and produce light. The membrane was exposed to X-ray film to detect the emitted light.

2.2.4.4 Labelling of a probe with ^{32}P

A Stratagene Prime-It II random primer labelling kit (Cat. No. 300385) was used to radioactively label probes with ^{32}P α-dATP following manufacturer's guidance. Templates for the reaction were obtained by PCR from chromosomal DNA as previously described. Labelling reactions were allowed to sit at 37 °C for 30 minutes. To clean probes, GE Healthcare illustra Microspin™ G-25 columns were used (Ca. No. 27-5325-01) following manufacturer's guidance.

2.2.4.5 Detection of a ^{32}P labelled probe

The cross-linked membrane was placed into a hybridisation bottle and re-hydrated in 2 X SSC. 10-15 ml (depending on the size of the membrane) of Church-Gilbert buffer, pre-warmed to 65 °C, were added and the bottle was put into a hybridisation oven at 65 °C for 2-6 hours to pre-hybridise. Once pre-hybridisation was over, 20 µl or 100 µl of probe were diluted with 100 µl or 20 µl sterile MQ-water, respectively, and boiled in a thermocycler for 5 minutes. In the meantime, the 10 ml of pre-hybridisation Church-Gilbert buffer were replaced with 10 ml of fresh, pre-warmed

at 65 °C, Church-Gilbert buffer. Once fully denatured, the probe was immediately added to the hybridisation bottle, which was placed back into the oven to hybridise overnight. The hybridisation buffer + probe was removed and a 15 minute wash in 200 ml of washing solution was performed to remove excess, unbound, probe, followed by a 30 minute wash in 200 ml of stringency washing solution, to remove unspecific hybridisation of the probe. The membrane was then wrapped in cling film and exposed to GE healthcare storage phosphor screens (Cat. No. 63-0034-86 and 63-0034-79). Exposure lasted between 1 and 7 days. Screens were scanned using a Molecular Dynamics Storm 860 phosphor imager scanner. When needed, membranes were stripped by washing for 1 hour in 50 ml of stripping buffer and then washed for 30 minutes in 200 ml of stringency washing solution in order to wash away all the formamide. Membranes were re-exposed to check for the degree of stripping, when needed, and pre-hybridised again for additional probing.

2.3 DNA primers, bacterial strains, and plasmids

2.3.1 DNA Primers for PCR

All DNA primers used in this work, listed in table 2.2, were generated using Primer3. Restriction fragment sites are underlined.

Table 2. 2 Primers

Name	Sequence (5′ to 3′)	Purpose
***araJ* F**	CAG CGT CAG CAT CAT ACC TC	Generates a 999 base pair fragment used to make a probe for Southern blot
***araJ* R**	GCC GAA TTT GGC ATT ATG G	
***cysN* F**	CGG TTG ATT GAC AAA TGC AC	Generates an 995 base pair fragment used to make a probe for Southern blot
***cysN* R**	ATC GCG TCA ATG TAC CCT TC	
***lacZ* F**	CTG GCG TAA TAG CGA AGA GG	Generates a base pair fragment used to make a probe for Southern blot
***lacZ* R**	CAT GAC CTG ACC ATG CAG AG	
***lacZ*Chi 1**	AAA AA**C** **TGC** **AG**G TCG GCA AAG ACC AGA CC	Generates a 915 base pair PCR cross-over fragment containing a 44 base pair Chi array at the centre and PstI/SalI restriction
***lacZ*Chi 2**	CAG CGC GTG TCC ACC AGC TCA GCA TCG ACC ACC AGC GTC ACG ACG CGC TGT ATC GCT G	

***lacZ*Chi 3**	TCG ATG CTG AGC TGG TGG ACA CGC GCT GGC TGG TGG TTA GCG CCG TGG CCT GAT TC	fragments at the extremities for cloning into pTOF24 for plasmid-mediated gene replacement. Allows for the insertion of the Chi array at 1.5 Kb upstream of the palindrome
***lacZ*Chi 4**	AAA AA**G** **TCG** **AC**C ACG CTG ATT GAA GCA GAA G	
***lacZ*.prox+1 F**	GGT GTG TGG GTT AGG TCT GG	Generates a 3100 base pair fragment used to make a probe for Southern blot
***lacZ*.prox+1 R**	GTG CAC GGC AGA TACACT TG	
***lacZ*.dist F**	ATC GTC GTA TCC CAC TAC CG	Generates a 2932 base pair fragment used to make a probe for Southern blot
***lacZ*.dist R**	TTT CCA TGC GAG GTT AAA GG	
***lacZ*.dist+1 F**	AGG GAC GCA TAC AGG AAC TG	Generates a 3224 base pair fragment used to make a probe for Southern blot
***lacZ*.dist+1 R**	TCC AGC GAA TAC TGA TGA CG	
***lepA* F**	TGG ATT ATC CGG AAC TGC TC	Generates a 729 base pair fragment used to make a probe for Southern blot
***lepA* R**	CCT TAT CCT GGA GGG CTT C	
***mhpR*Chi 1**	AAA AA**C** **TGC** **AG**C TGG CAC CCA GTT GAT CG	Generates a 915 base pair PCR cross-over fragment containing a 44 base pair Chi
***mhpR*Chi 2**	GCG CAG ACT CGC TGG TGG TCA CAT GGC GGC	

***mhpR*Chi 3**	TGG TGG CAA CGG GTA GCA AAA CAG ATC	array at the centre and PstI/SalI restriction fragments at the extremities for cloning into pTOF24 for plasmid-mediated gene replacement. Allows for the insertion of the Chi array at 1.5 Kb downstream of the palindrome
***mhpR*Chi 4**	CGC CAT GTG ACC ACC AGC GAG TCT GCG CCC ACC AGC TAG CGC CGG AAG ATG CTT TTC	
	AAA AA**G** **TCG** **AC**A GTG GTA TGG CCG ACA GAT G	
***nadA* F**	CCG CGA GAA GAT AAA ACG TC	Generates an 948 base pair fragment used to make a probe for Southern blot
***nadA* R**	AGT GTA GCC GCA AAA TCC AG	
***nagC* F**	AAC CAA AAT TAC GCG TCA GC	Generates an 915 base pair fragment used to make a probe for Southern blot
***nagC* R**	GCA GGG AGC AGC ACT TTA TC	
pKO F	AGG GCA GGG TCG TTA AAT AGC	Amplifies across pTOF24 cloning site to allow for sequencing of inserts
pKO R	AGG GAA GAA AGC GAA AGG AG	
P*recG* 1	AAA AAC TGC AGG ACC GGT GTC GAT GTT TTT C	Generates a PCR cross-over fragment containing recG preceded by its putative promoter, which can be digested with PstI and SalI
P*recG* 2	TTC ATA CTT AAA AAG CTC CAC AGG TGA AGA AAT G	
P*recG* 3	TTT TAA GTA TGA AAG	

P*recG* 4	GTC GCC TGT TAG ATG CTG	for cloning into pGB2
	AAA AAG TCG ACA AAT GGC GGT CTT CTC ACT G	
***purA* F**	ACA CCT TTC CAG TCG TCA GC	Generates a 928 base pair fragment used to make a probe for Southern blot
***purA* R**	TTC TCC GCG AGA ATG TAA CC	
***yagV* F**	AAA GCC CAT CGT TACAGG TG	Generates a 998 base pair fragment used to make a probe for Southern blot
***yagV* R**	ATG ATA GCT GGC GGG ATA TG	

2.3.2 Escherichia coli strains

All *E. coli* strains used in this work are listed in table 2.3. *lacZ*χ- represents an allele of *lacZ*, in which a single endogenous χ site has been removed. χ χ χ indicates a χ array, which is a sequence containing three repeats of the crossover hotspot instigator (Chi; χ).

Table 2. 3 E. coli strains

Strain	Genotype of interest	Background	Source
AB1155	*his4 proA2 argE3 thr1 leuB6 ara14 galK2 lacYi xyl5 mtl1 str*^{R}		
BW27784	*DE(araFGH) (*ΔP_{araE} P_{CP18}*-araE)*		Khlebnikov et al. (2001)
MG1655	*F- lambda- ilvG- rfb-50 rph-1*		
DL1077	Δ*recG*263::Km^R	AB1155	Mahdi et al. (1996)
DL2075	*recA*::Cm^R (P_{BAD}-*sbcDC lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2006 (BW27784)	Eykelenboom et al. (2008)
DL2573	P_{BAD}-*sbcDC lacZ* $^+$ *cynX*::Gm^R *lacI*^{q} *lacZ*χ-	BW27784	Eykelenboom et al. (2008)
DL2605	*recA*::Cm^R (P_{BAD}-*sbcDC lacZ* $^+$ *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2573 (BW27784)	Eykelenboom et al. (2008)
DL2155	Δ*recG*263::Kan^R (P_{BAD}-*sbcDC lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2006 (BW27784)	Eykelenboom et al. (2008)
DL2610	Δ*recG*263::Kan^R (P_{BAD}-*sbcDC lacZ* $^+$	DL2573	Eykelenboom et al.

	cynX::Gm^R *lacI*^{q} *lacZ*χ-)	(BW27784)	(2008)
DL2800	Δ*ruvAB* (P_{BAD}-*sbcDC* *lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2006 (BW27784)	Eykelenboom et al. (2008)
DL2801	Δ*ruvAB* (P_{BAD}-*sbcDC lacZ* $^+$ *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2573 (BW27784)	Eykelenboom et al. (2008)
DL4184	*mhpR*::χ χ χ *lacZ*::χ χ χ (*proA*::ISceI$_{cs}$ *tsx*::ISceI$_{cs}$ P_{BAD} -*sbcDC* *lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	BW27784	This work
DL4464	Δ*recG*263::Kan^R Δ*ruvAB* (P_{BAD}-*sbcDC* *lacZ* $^+$ *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2801 (BW27784)	P1 from DL1077
DL4465	Δ*recG*263::Kan^R Δ*ruvAB* (P_{BAD}-*sbcDC* *lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2800 (BW27784)	P1 from DL1077
DL4724	Δ*ruvC* (P_{BAD}-*sbcDC* *lacZ*::pal246 *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2006 (BW27784)	pTOF from pDL2731
DL4725	Δ*ruvC* (P_{BAD}-*sbcDC lacZ* $^+$ *cynX*::Gm^R *lacI*^{q} *lacZ*χ-)	DL2573 (BW27784)	pTOF from pDL2731

2.3.3 Plasmids

All plasmids used in this work are listed in table 2.4

Table 2. 4 Plasmids

Strain	Plasmid	Background/Alternate	Source
DL1605	pTOF24	XL1 Blue	(Merlin et al., 2002)
DL3596	pGBruvAB	JJC842	Gift from B. Michel
DL3673	pGB2	JJC261	Gift from B. Michel (Churchward et al., 1984)
DL4137	pTOF*lacZ*::χ χ χ	XL1 Blue	This work
DL4138	pTOF*mhpA*::χ χ χ	XL1 Blue	This work
DL4569	pGBrecG	XL1 Blue	This work

Chapter 3

Effects of inducing a DNA double strand break at a 246 base pair interrupted palindrome in mutants unable to resolve recombination intermediates

3.1 Introduction

In this study, a method previously described (see Chapter 1) (Eykelenboom et al., 2008), which induces the formation of a DNA double strand break (DSB) at a 246 base pair (bp) interrupted palindrome in the *Escherichia coli* chromosome, was used to investigate the physical nature of DNA intermediates formed during chromosomal DSB repair and the role of the RecG helicase in this process. One of the advantages of using this system is that the DSB is site-specific, which allows for the analysis of the DNA directly surrounding the break point. In addition, only the lagging strand is cleaved, leaving the leading strand template and its newly synthesised complement available as a template for repair. Indeed, under conditions of DSB formation using this system, recombination proficient (Rec^+) strains do not have reduced viability indicating that the breaks are successfully

repaired. In contrast, strains deficient in RecBCD-mediated homologous recombination (HR) do not survive this type of damage (Eykelenboom et al., 2008). RecG is a protein that has been implicated in the resolution of intermediates formed during DSB repair by RecBCD-mediated HR and has been suggested to be redundant with RuvABC. Therefore, in the absence of both RuvABC and RecG DSB repair intermediates should accumulate in the chromosome. This chapter describes the growth profiles of Δ*ruvAB*, Δ*recG* and Δ*ruvAB* Δ*recG* mutant strains using a combination of growth curves and viability tests under the conditions of acute DSB formation, where DSBs are induced for a short time period, or chronic DSB formation, where DSBs are induced overnight. Complementation experiments are also shown alongside with preliminary data addressing the sensitivity of Δ*ruvC* strains to DSBs.

3.2 Growth profiles

3.2.1 Growth curves of Rec$^+$, Δ*recA*, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains

The growth patterns of Rec$^+$, Δ*recA*, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains, either containing the 246 bp interrupted palindrome in *lacZ* (*lacZ*::246), or not (*lacZ*$^+$), are presented in Figures 3.1 and 3.2. Growth in the presence of 0.5 % glucose results in the silencing of *sbcCD* expression (Fig. 3.1), whereas growth in the presence of 0.2 % arabinose results in induction of *sbcCD* expression (Fig. 3.2).

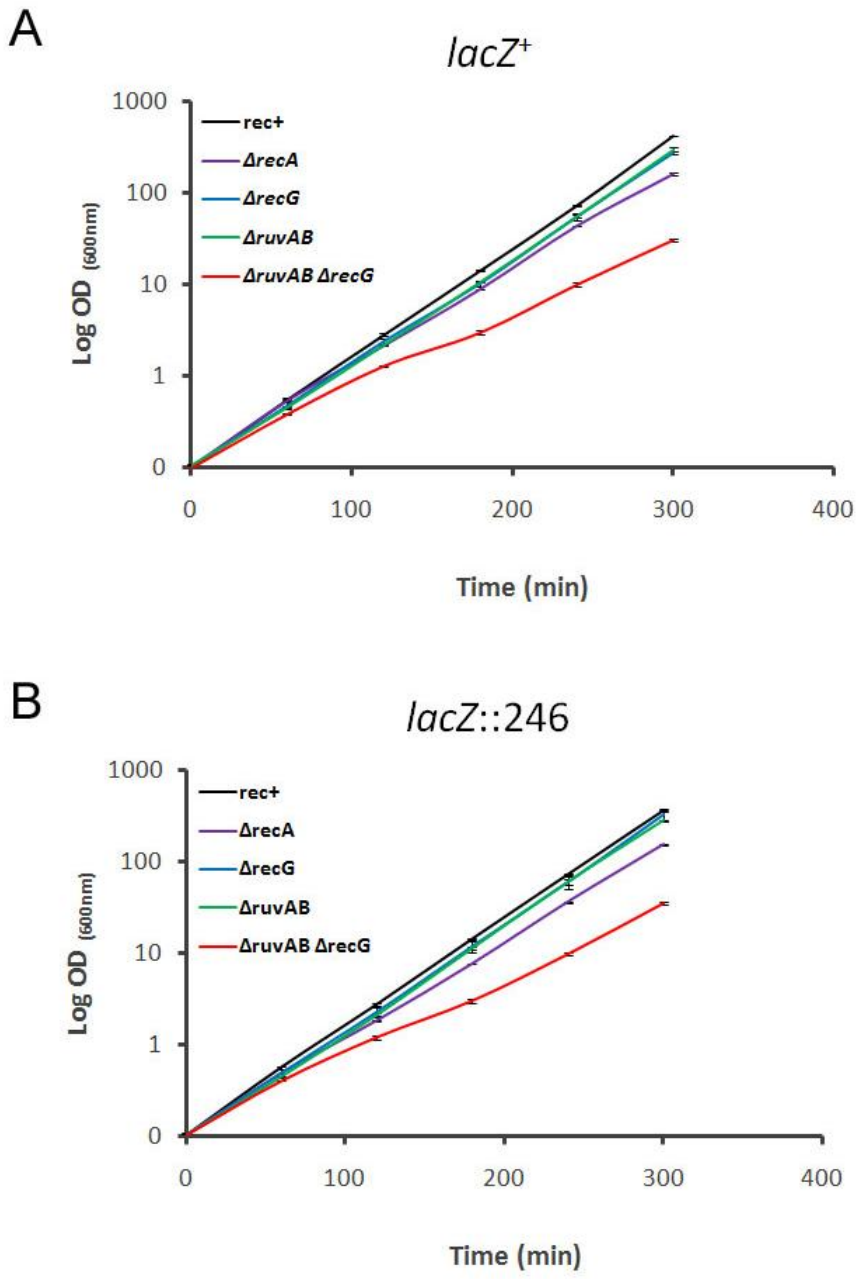

Figure 3. 1 Growth of recombination deficient mutants grown in the presence of 0.5 % glucose (SbcCD$^-$)

All cultures were maintained in exponential growth phase by regularly being diluted in fresh L-broth, supplemented with the relevant sugar, in order to maintain the OD600nm below 1.0. (A) Growth of *lacZ*$^+$ strains. Strains used were DL2573 (Rec$^+$ *lacZ*$^+$), DL2605 (Δ*recA* *lacZ*+), DL2610 (Δ*recG* *lacZ*$^+$), DL2800 (Δ*ruvAB* *lacZ*$^+$), and DL4464 (Δ*ruvAB* Δ*recG* *lacZ*$^+$). (B) Growth of *lacZ*::246 strains. Strains used were DL2006 (Rec$^+$ *lacZ*::246), DL2075 (Δ*recA* *lacZ*::246), DL2155 (Δ*recG* *lacZ*::246), DL2801 (Δ*ruvAB* *lacZ*::246), and DL4465 (Δr*uvAB* Δ*recG* *lacZ*::246). Error bars represent standard error of the mean where n = 3.

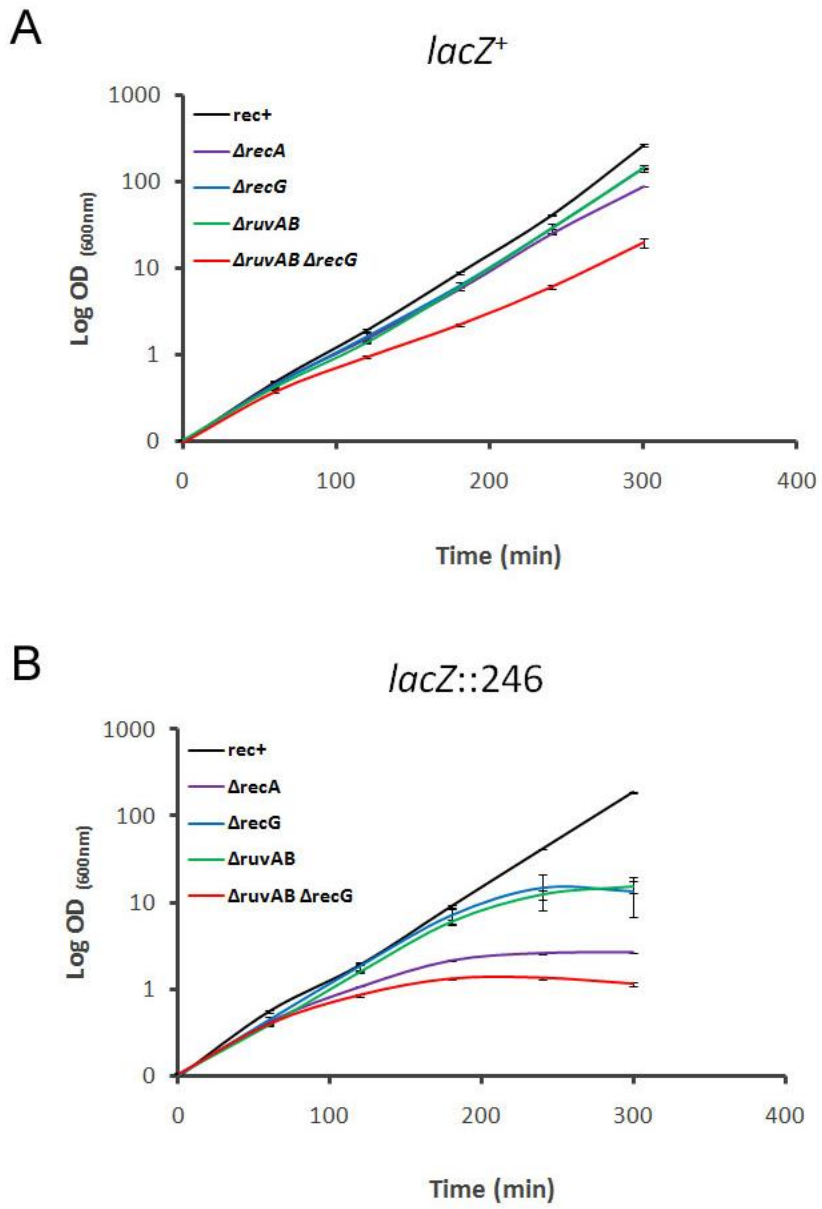

Figure 3. 2 Growth of recombination deficient mutants grown in the presence of 0.2 % arabinose (SbcCD^{+})

All cultures were maintained in exponential growth phase by regularly being diluted in fresh L-broth, supplemented with the relevant sugar, in order to maintain the OD600nm below 1.0. (A) Growth of *lacZ*$^{+}$ strains. Strains used were DL2573 (Rec^{+} *lacZ*$^{+}$), DL2605 (Δ*recA* *lacZ*$^{+}$), DL2610 (Δ*recG* *lacZ*$^{+}$), DL2800 (Δ*ruvAB* *lacZ*$^{+}$), and DL4464 (Δ*ruvAB* Δ*recG* *lacZ*$^{+}$). (B) Growth of *lacZ*::246 strains. Strains used were DL2006 (Rec^{+} *lacZ*::246), DL2075 (Δ*recA* *lacZ*::246), DL2155 (Δ*recG* *lacZ*::246), DL2801 (Δ*ruvAB* *lacZ*::246), and DL4465 (Δ*ruvAB* Δ*recG* *lacZ*::246). Error bars represent standard error of the mean where n = 3.

In the absence of the palindrome, the growth rate of all strains was very similar when cultures were grown in the presence of either 0.5 % glucose or 0.2 % arabinose (panels A from Figs. 3.1 and 3.2). This result suggests that in the absence of the palindrome, the induced expression of *sbcCD* has no deleterious effect on growth rate. The growth rate is marginally faster when cells are grown in 0.5 % glucose than in 0.2 % arabinose. The reduced growth rate in arabinose may be an effect of the increased levels of SbcCD, but another possible explanation is that strains grown in glucose are metabolising the sugar whereas strains grown in arabinose are not. All strains used here are derived from BW27784 where the genes required for arabinose metabolism have been altered to ensure that all arabinose present remains available for the induction of the P_{araBAD} promoter (Khlebnikov et al., 2001). As a result, strains grown solely on arabinose have less sugar available for growth than strains grown on glucose.

As previously mentioned, a recombination proficient strain is able to successfully repair a DSB formed at the 246 bp interrupted palindrome. This ability is also illustrated in Figure 3.2, where the Rec$^+$ *lacZ*::246 strain grew at a rate comparable to the Rec$^+$ *lacZ*$^+$ strain. In contrast, all recombination deficient *lacZ*::246 strains displayed a very different growth profile from their *lacZ*$^+$ counterpart when the expression of *sbcCD* was induced (in presence of 0.2% arabinose). The growth rates of Δ*ruvAB* and Δ*recG* mutant strains began to decline by 180 minutes after the induction of *sbcCD*. A more dramatic fall was seen in the Δ*recA* and Δ*ruvAB* Δ*recG* strains, whose grow rates began to fall just after 100 minutes following the induction of SbcCD expression. Nevertheless, when grown in 0.5 % glucose (Fig. 3.1), these strains had the same growth rate as their equivalent *lacZ*$^+$ strain, which indicates

that the reduction in optical density at 600 nm (OD_{600nm}) observed in 0.2 % arabinose was specifically due to the action of SbcCD at the palindrome.

3.2.2 Acute DSB formation vs chronic DSB formation; requirements for RuvAB and RecG

In order to gain further insight into the viability of the Δ*recG*, Δ*ruvAB*, and Δ*ruvAB* Δ*recG* strains, 10-fold serial dilutions of cultures grown in the specified conditions were spotted onto LB plates. In order to compare different strains, cultures were diluted to the same OD_{600nm} (0.4) before being serially diluted. For looking at the effect of chronic DSB formation (Fig. 3.3), where DSBs are repeatedly induced for an extended time period, overnight cultures were diluted, spotted directly onto LB plates supplemented with either 0.5 % Glucose or 0.2 % Arabinose and incubated at 37°C overnight. For looking at the effect of acute DSB formation (Fig. 3.4), where DSBs were induced for a short period of time resulting in a small number of breaks (one to two), either 0.5 % Glucose or 0.2 % Arabinose was added to exponential cultures for a given period of time. Cultures were then spotted onto LB plates supplemented with 0.5 % Glucose and incubated at 37 °C overnight.

In conditions of acute, chronic or no DSB formation, Rec^+ *lacZ*::246 retained full viability in the presence or absence of SbcCD. This result is consistent with the

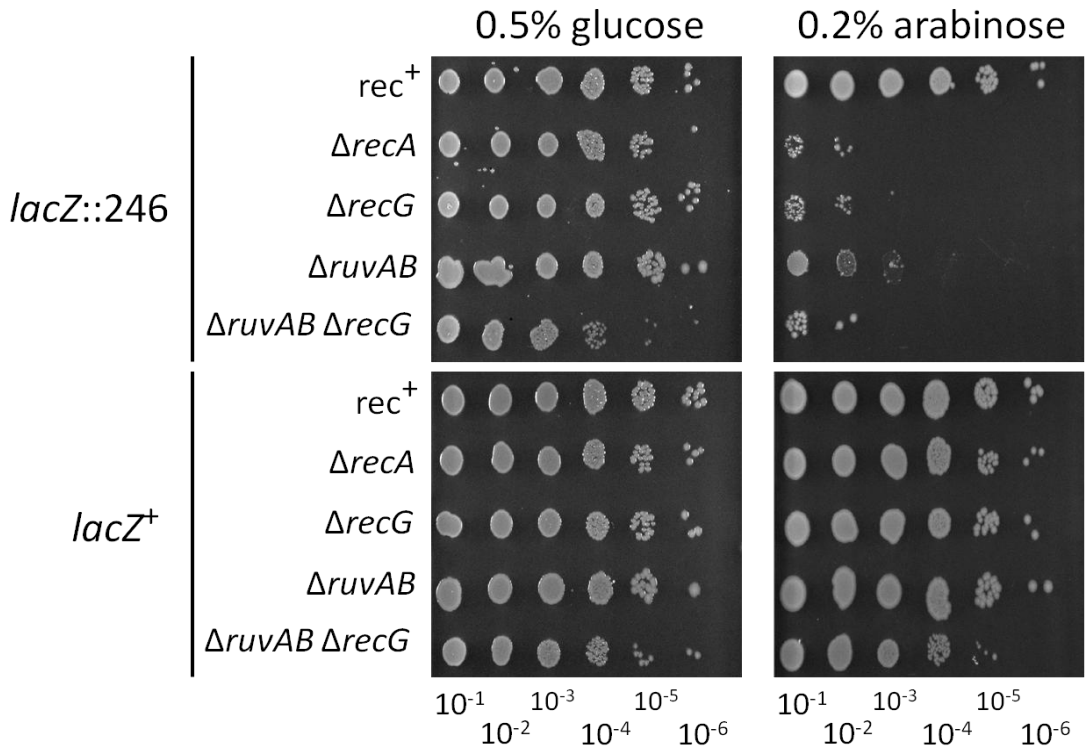

Figure 3. 3 Requirements for RuvAB and RecG under conditions of chronic double strand break formation

10-fold serial dilutions were spotted onto either 0.5 % glucose or 0.2 % arabinose and incubated overnight at 37 °C. Strains used were DL2006 (Rec$^+$ *lacZ*::246), DL2573 (Rec$^+$ *lacZ*$^+$), DL2075 (Δ*recA lacZ*::246), DL2605 (Δ*recA lacZ*$^+$), DL2155 (Δ*recG lacZ*::246), DL2610 (Δ*recG lacZ*$^+$), DL2801 (Δ*ruvAB lacZ*::246), DL2800 (Δ*ruvAB lacZ*$^+$), DL4465 (Δ*ruvAB* Δ*recG lacZ*::246), and DL4464 (Δ*ruvAB* Δ*recG lacZ*$^+$).

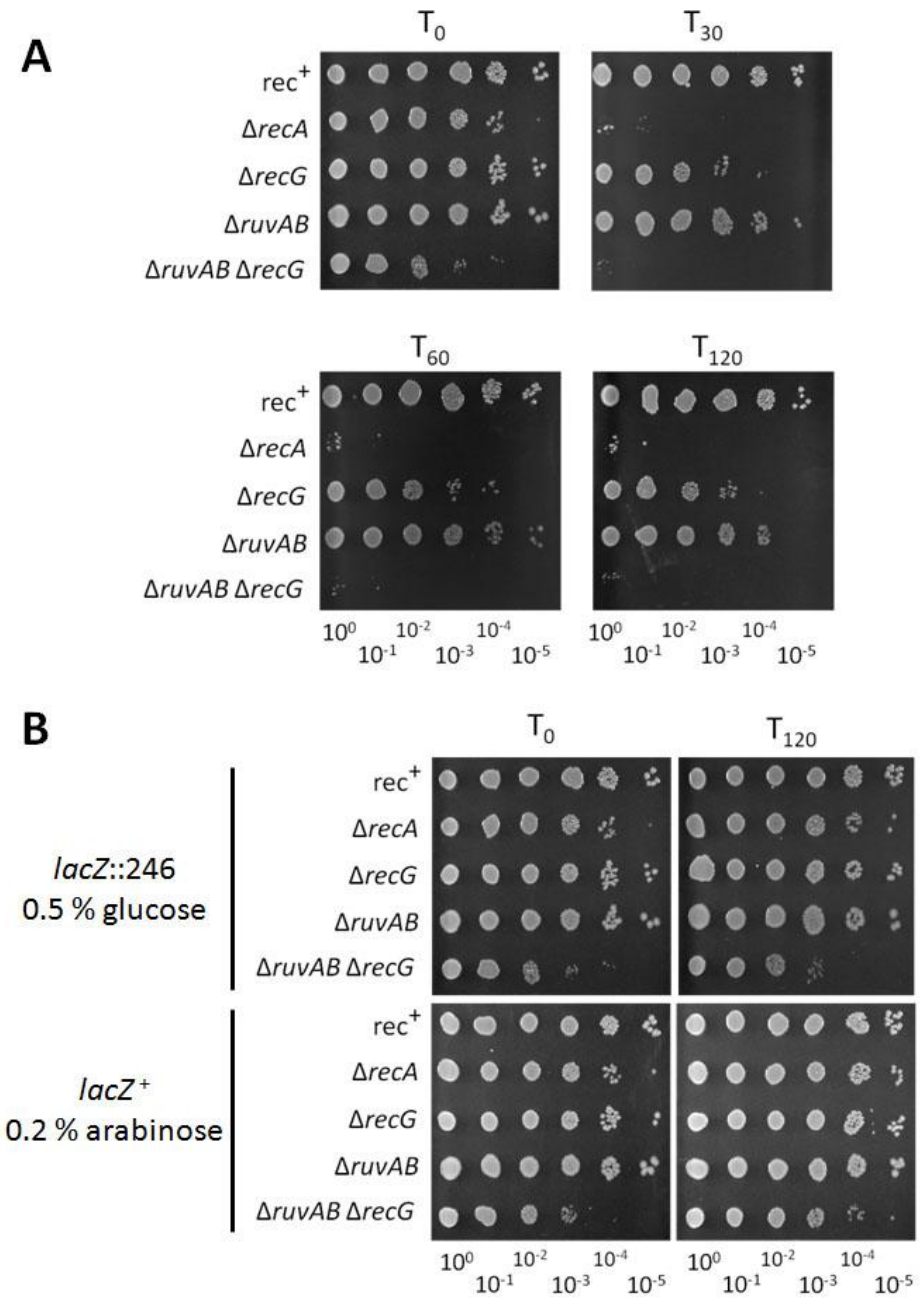

Figure 3. 4 Requirements for RuvABC and RecG under conditions of acute double strand break formation

(A) Viability spot tests from lacZ::246 cultures grown in 0.2 % arabinose for different time periods; from 0 minute (T0) to 120 minutes (T120). (B) Control viability spot tests of *lacZ*::246 cultures grown in 0.5 % glucose for 120 minutes and *lacZ*$^+$ cultures grown in 0.2 % arabinose for 120 minutes. Strains used were DL2006 (Rec$^+$ *lacZ*::246), DL2573 (Rec$^+$ *lacZ*$^+$), DL2075 (Δ*recA lacZ*::246), DL2605 (Δ*recA lacZ*$^+$), DL2155 (Δ*recG lacZ*::246), DL2610 (Δ*recG lacZ*$^+$), DL2801 (Δ*ruvAB lacZ*::246), DL2800 (Δ*ruvAB lacZ*$^+$), DL4465 (Δ*ruvAB* Δ*recG lacZ*::246), and DL4464 (Δ*ruvAB* Δ*recG lacZ*$^+$).

growth rate observed in Figure 3.2 and with the conclusion that recombination proficient cells successfully repair all DSBs formed at the palindrome.

As seen in Figure 3.3, Δ*recA lacZ*::246 has a 10,000-fold reduction in viability under the conditions of DSB formation, illustrating the importance of RecA in DSB repair. This finding is also in accordance with the reduced growth rate of this strain as shown in Figure 3.2. Additionally, the Δ*recA* strain showed the same loss of viability in conditions of acute DSB formation after as little as 30 minutes of exposure to 0.2 % Arabinose (Fig. 3.4). This result suggests that once a single DSB is formed in a strain lacking RecA, plating on glucose and therefore preventing any further breaks from arising does not rescue the viability of the strain and indicates that no RecA-independent pathway for DSB repair is available.

Δ*ruvAB* and Δ*recG* mutants appeared completely viable in conditions of no DSB formation. In conditions of chronic DSB formation, both mutants showed a high sensitivity to DSBs (Fig. 3.3). However, the ability of these mutants to deal with acute DSB formation differed. No reduction in viability was observed for the Δ*ruvAB* mutant for as long as 120 minutes of exposure to arabinose. This is consistent with the growth rate shown in Figure 3.2, which shows that the optical density starts to decline after 180 minutes of exposure to arabinose. This finding can be interpreted in two ways. One possibility could be that all intermediates formed during this period can be successfully resolved, presumably by a RuvAB-independent pathway. Alternatively it is possible that Holliday junctions (HJs) are not generated when following acute DSB formation. This would suggests that in these conditions DSBs are repaired by a pathway analogous to the synthesis-dependent strand annealing (SDSA) pathway that has been described to act in

eukaryotes (Paques and Haber, 1999). In contrast, the Δ*recG* mutant displayed a 10-fold decrease in viability after only 30 minutes of SbcCD expression (Fig. 3.4). This is an interesting observation as the optical density of the Δ*recG* mutant does not start to decrease until about 180 minutes of exposure to arabinose (Fig. 3.2). This could be an indication that the cells filament in response to DSB formation and would account for a stable increase in OD_{600nm} accompanied by a decrease in viability. This result also suggests that the repair of acute DSBs requires the activity of RecG and emphasises that RuvAB and RecG have different roles in the repair of a DSB.

The Δ*ruvAB* Δ*recG* mutant is the sickest strain of all, both in the presence and absence of either acute or chronic DSBs, and in the presence and absence of SbcCD (Figs. 3.3 and 3.4). When DSBs are induced for as little as 30 minutes, there is a complete loss of viability indicating that the absence of both RuvAB and RecG is lethal when as little as one DSB is formed (Fig. 3.4 A). In the absence of DSBs, this mutant displayed at least a 10–fold decrease in viability compared to wild-type strains or strains harbouring the individual mutations. This result means that at least 90 % of the cells in a Δ*ruvAB* Δ*recG* culture are not viable under conditions of no DNA damage, suggesting that events that require the action of these proteins arise regularly in the cell. The fact that the single mutants did not show this decrease in viability indicates that under certain circumstances the functions of RuvAB and RecG may be redundant. Nevertheless, the difference in viability of the Δ*ruvAB* and the Δ*recG* mutants in conditions of acute DSB formation indicates that these two proteins play different roles in the repair of a limited number of DSBs.

To quantify more accurately the differences in viability between the Δ*ruvAB* *lacZ*::246 and the Δ*recG* *lacZ*::246 strains following the formation of acute DSBs, a full

viability test, which quantifies the colony forming unit (CFU) potential of the different strains grown in different conditions for 60 minutes, was carried out (Fig. 3.5). The Rec$^+$ strain formed the largest number of colonies both in the presence and absence of DSBs. A two-sample T-test comparing the values obtained for time point 60 minutes in 0.5 % Glucose and time point 60 minutes in 0.2 % Arabinose shows that there is no significant difference between the viabilities ($p = 0.817$). The Δ*ruvAB lacZ*::246 strain as well as the Δ*recG lacZ*$^+$ strain had a 2-fold reduction in CFU potential in all conditions when compared to the Rec$^+$ strain. Nonetheless, a two-sample T-test for the Δ*ruvAB lacZ*::246 strain shows that the CFU potential remained constant irrespective of whether *sbcCD* was expressed (0.5 % Glucose compared to 0.2 % Arabinose; $p = 0.196$). For the Δ*recG lacZ*$^+$ strain it was not possible to carry out a two-sample T-test as the variances of the two samples being compared were not equal. As a result, the values obtained for 0.5 % Glucose and 0.2 % Arabinose were compared using the Mann-Whitney test for non-parametric data. The P-value obtained with this test showed no significant difference ($p = 0.081$). It is important to note that less than 5 repeats of this experiment were performed and therefore the results of the test ought to be interpreted with caution since a significant difference may be masked by high random variation of the small sample. In contrast to this, and consistent with the viability spot tests shown in Figure 3.4, Δ*recG lacZ*::246 displayed a 14-fold reduction in colony forming unit potential when *sbcCD* was expressed (in 0.2 % Arabinose). This was supported by a two-sample T-test that showed that the difference between the values obtained from the two samples tested (0.5 % Glucose and 0.2 % Arabinose) was highly significant ($p < 0.01$).

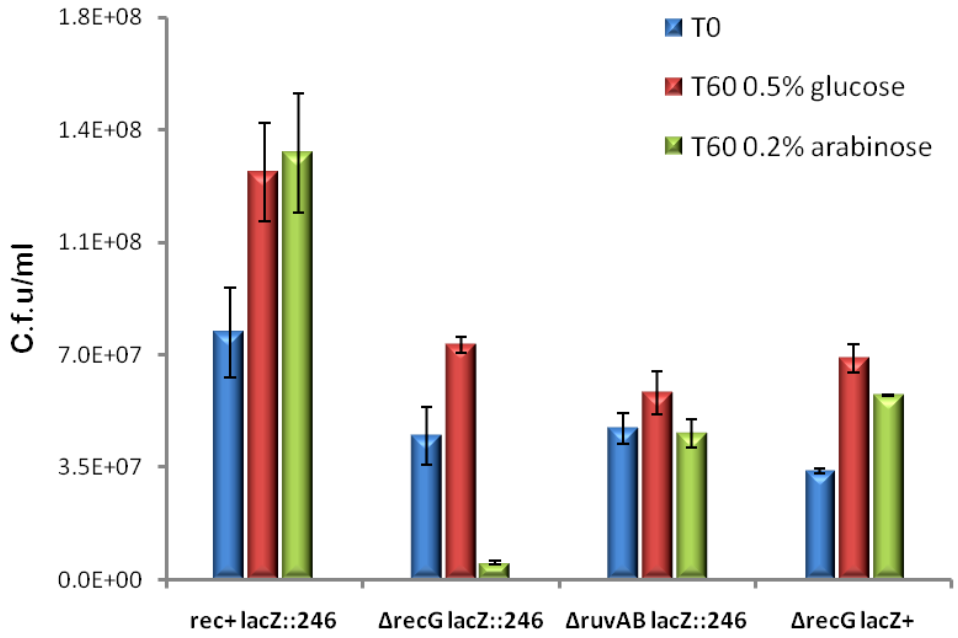

Figure 3. 5 Colony forming unit potential of strains grown in either 0.5 % glucose or 0.2% arabinose for 60 minutes

Strains used were DL2006 (Rec$^+$ *lacZ*::246), DL2155 (Δ*recG lacZ*::246), DL2610 (Δ*recG lacZ*$^+$), and DL2801 (Δ*ruvAB lacZ*::246). Error bars are standard error of the mean where n = 3.

3.3 Complementing the Δ*ruvAB* and Δ*recG* phenotypes

The exact role of RecG in double strand break repair has been a point of discussion ever since the gene was identified in 1971 (Storm et al., 1971). There is *in vitro* evidence suggesting that RecG can act on HJs, in a fashion analogous to that of RuvAB. This would argue that these proteins are redundant (McGlynn and Lloyd, 1999). Another suggested substrate for RecG is a three-way junction such as a D-loop, which would be expected to arise from RecA-mediated strand invasion. It has been postulated that RecG can either unwind these three-way junctions, essentially aborting the strand-invasion event, or wind (migrate) them in the opposite direction, which would result in the maturation of a HJ from the D-loop (Lloyd and Sharples, 1993a; Lloyd and Sharples, 1993b). More recently, a role has been attributed to RecG in the control of replication fork collision, another event expected to occur during DSB repair (Rudolph et al., 2010). In order to determine whether RecG and RuvAB have redundant functions a complementation experiment was carried out using two pGB2-based plasmids, pGB*ruvAB* (gift from Benedict Michel) and pGB*recG* (this work).

3.3.1 Construction of pGB*recG*

pGB*recG* is a pGB2-based plasmid, the construction of which was based on pGB*ruvAB* (Baharoglu et al., 2008; Seigneur et al., 1998). pGB2 is a pSC101-derived plasmid, which has a copy number of about 10-12 molecules per cell, as shown by

monitoring the activity of the fluorescent reporter protein, luciferase, when expressed from either pSC101 or the *E. coli* chromosome (Churchward et al., 1984; Lutz and Bujard, 1997). The expression of a gene placed in pGB2, under the regulation of its native promoter, is therefore expected to be 10-12 times higher than when expressed from the endogenous chromosomal locus.

recG is the last gene of the *spo* operon (Kalman et al., 1992; Lloyd and Sharples, 1991). Transcription of RecG is driven by two promoters located at the beginning of the operon, as well as weaker secondary promoters located within the *spoU* gene region (Gentry and Burgess, 1989; Lloyd and Sharples, 1991). A DNA fragment containing *recG*, preceded by the *spo* operon promoters P1 and P2, was artificially synthesised by crossover PCR. Primers P*recG* 1 and P*recG* 2 (table 2.2) were used to amplify a 389 bp fragment containing the promoters, P1 and P2, from MG1655 chromosomal DNA. Using the same chromosomal template, PCR primers P*recG* 3 and P*recG* 4 (table 2.2) were used to amplify a 2.104 Kb region containing *recG*. For the cross-over PCR, primers P*recG* 1 and P*recG* 4 and the products from the first two PCR reactions were used to join the promoter fragment to the *recG* fragment, producing a final PCR product of 2.506 bp. The promoter-*recG* fragment was then digested with PstI (located at the 5′ end of P*recG* 1) and SalI (located at the 5′ end of P*recG* 4) and ligated into PstI/SalI-digested pGB2, which resulted in the insertion of the fragment in the pGB2 multiple cloning site (MCS), generating pGB*recG*. Correct cloning of the insert was checked by restriction digestion and sequencing.

3.3.2 Complementation experiment

To assess whether active proteins were expressed *in vivo* by both pGB*ruvAB* and pGB*recG*, the ability of the plasmids to rescue the lethality of Δ*ruvAB lacZ*::246 and Δ*recG lacZ*::246 strains in conditions of chronic DSB formation was assessed (Fig. 3.6 A). Upon expression of SbcCD, both pGB*ruvAB* and pGB*recG* rescued the death phenotype of Δ*ruvAB lacZ*::246 and Δ*recG lacZ*::246 strains, respectively. Interestingly, pGB*ruvAB* was not able to fully rescue Δ*recG lacZ*::246 and pGB*recG* may even be causing additional death in Δ*ruvAB lacZ*::246. These results strongly indicate that under conditions of prolonged exposure to DSBs, RuvAB and RecG are not able to substitute for one another and therefore that their function underthese conditions is not redundant. As expected, no rescue was seen with pGB2, the empty vector, in both backgrounds.

Despite this inability of RuvAB to complement for the absence of RecG under conditions of chronic DSB formation, its ability to complement for the Δ*recG* lethality under conditions of acute DSB formation was investigated (Fig. 3.6 B). After 60 minutes of *sbcCD* induction, the viability of Δ*recG lacZ*::246 is 10-fold lower than the viability of both Rec$^+$ *lacZ*::246 and Δ*ruvAB lacZ*::246. As seen in Figure 3.6 B, pGB*ruvAB* rescued the lethal phenotype of the Δ*recG* mutant following the formation of acute DSBs as successfully as pGB*recG* did. This result suggests that the cellular environment of a strain subject to a limited number of DSBs may be different from that of a strain that has been subjected to multiple rounds of break formation. It also suggests that the intermediates generated in a Δ*recG* background

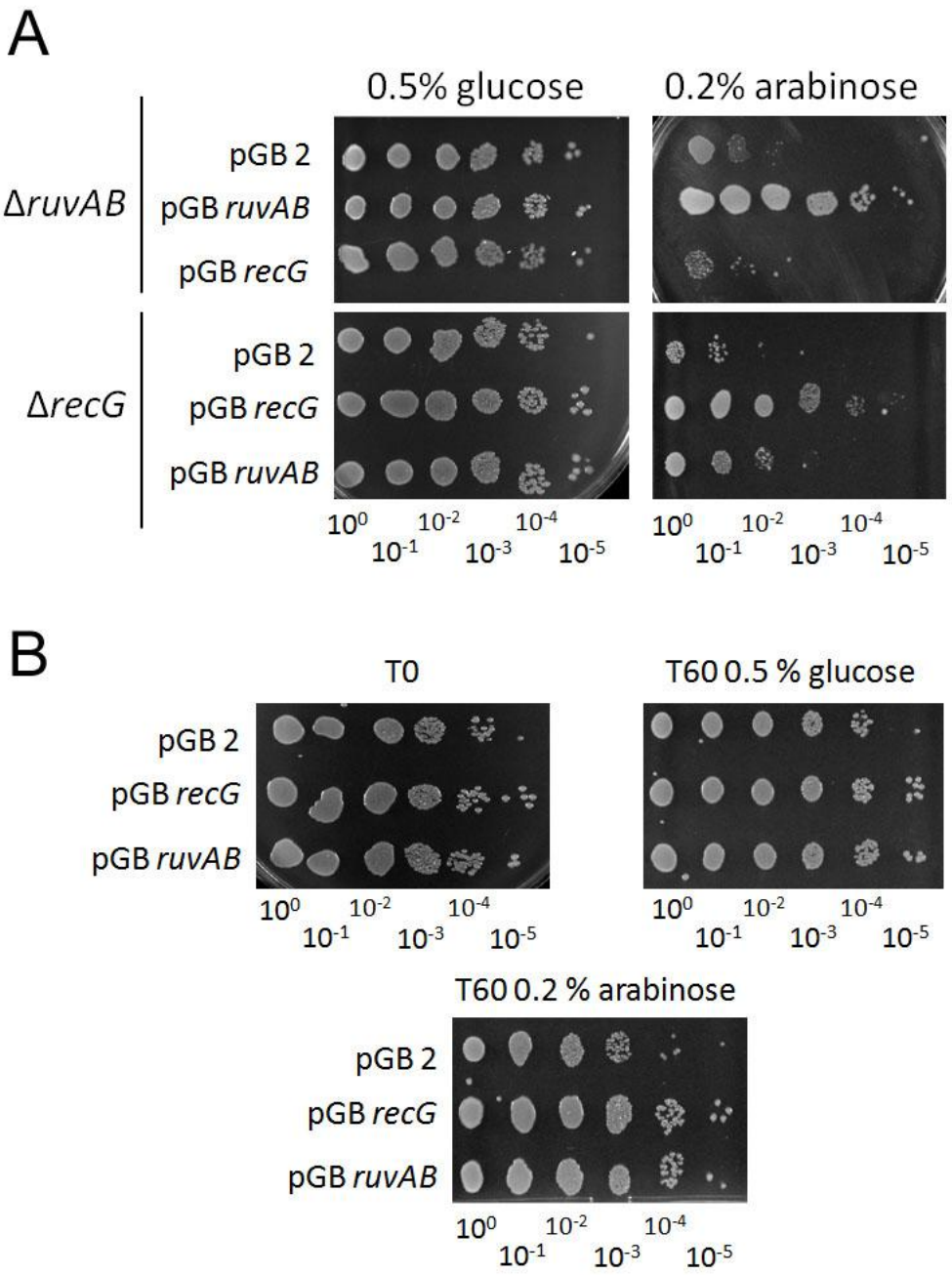

Figure 3. 6 Complementation of Δ*ruvAB* and Δ*recG* strains through the exogenous expression of either RuvAB or RecG

(A) Complementation in conditions of chronic double strand break formation both in Δ*ruvAB* and Δ*recG* backgrounds. Strains used were *lacZ*::246; DL4642 (Δ*recG* + pGB2), DL4643 (Δ*recG* + pGB-*ruvAB*), DL4644 (Δ*recG* + pG-*recG*), DL4645 (Δ*ruvAB* + pGB2), DL4646 (Δ*ruvAB* + pGB-*ruvAB*), DL4647 (Δ*ruvAB* + pG-*recG*). (B) Complementation in a Δ*recG lacZ*::246 background in conditions of acute double strand break formation. Strains used were; DL4642 (Δ*recG* + pGB2), DL4643 (Δ*recG* + pGB-*ruvAB*), and DL4644 (Δ*recG* + pG-*recG*).

in acute DSB formation are affected by elevated levels of RuvAB. These intermediates are either resolved by RuvAB or never generated in this context.

3.4 Requirements for RuvC; viability of Δ*ruvC* strains

The ability of Δ*ruvAB* strains to survive conditions of acute DSB formation lead to the suggestion that an alternative, RuvAB-independent mechanism for HJ resolution was available. It is possible that RuvC can resolve a HJs in the absence of RuvAB as it has been shown *in vitro* that elevated levels of RuvC can resolve four-way junctions without the requirement for RuvAB (Dunderdale et al., 1991). Nevertheless, the intra-cellular levels of RuvC may not be high enough for this activity to be physiologically relevant. Furthermore, RuvC is not induced by the SOS response, which means that the concentration of this nuclease would not increase even in response to DSB formation (Sharples and Lloyd, 1991). Resolution at the hands of another nuclease is a possibility. It is know that activation of RusA, a cryptic prophage encoded by the *E. coli* chromosome, can rescue the lethality of either Δ*ruvAB* or Δ*ruvC* mutations (Mandal et al., 1993; Sharples et al., 1994). Nevertheless, the *E. coli* strains used in this work do not have an active *rusA* gene. To date, the search for an alternative nuclease has not been successful.

If it is assumed that the loss of any of the components of the RuvABC HJ resolvase complex completely abolishes resolution by nuclease action, a more likely mechanism of resolution would rely on a helicase to migrate the junction back to the breakpoint or to migrate two junctions towards each other. In this situation, the

position of the HJs with respect to the breakpoint, or each other, might be important. HJs that have not moved far from the location in which they were first generated could be resolved with more ease than ones that would have migrated into the rest of the chromosome.

In both scenarios, a strain lacking RuvC (but maintaining wild type RuvAB) ought to be more sensitive to acute DSB formation than a strain lacking RuvAB alone. If a Δ*ruvAB* mutant survives because RuvC can resolve a limited number of HJs in the absence of RuvAB, then removing the RuvC activity is clearly going to cause cell death. If resolution of HJs occurs through the action of a helicase, a Δ*ruvC* strain may die because the wild type RuvAB HJ branch migration complex would move the junctions away from the breakpoint, making it more difficult for a helicase to resolve them.

Viability spot tests of Δ*ruvC* strains in conditions of chronic and acute DSB formation confirm the requirement for RuvC and are presented in Figure 3.7. Formation of chronic DSBs in the Δ*ruvC lacZ*::246 spotted on 0.2 % Arabinose decreased cell viability 10,000-fold, like the Δ*ruvAB lacZ*::246 and Δ*recG lacZ*::246 strains grown in the same conditions (Fig. 3.3). This finding confirms that, in conditions where persistent DSBs are being formed, it is not only the branch

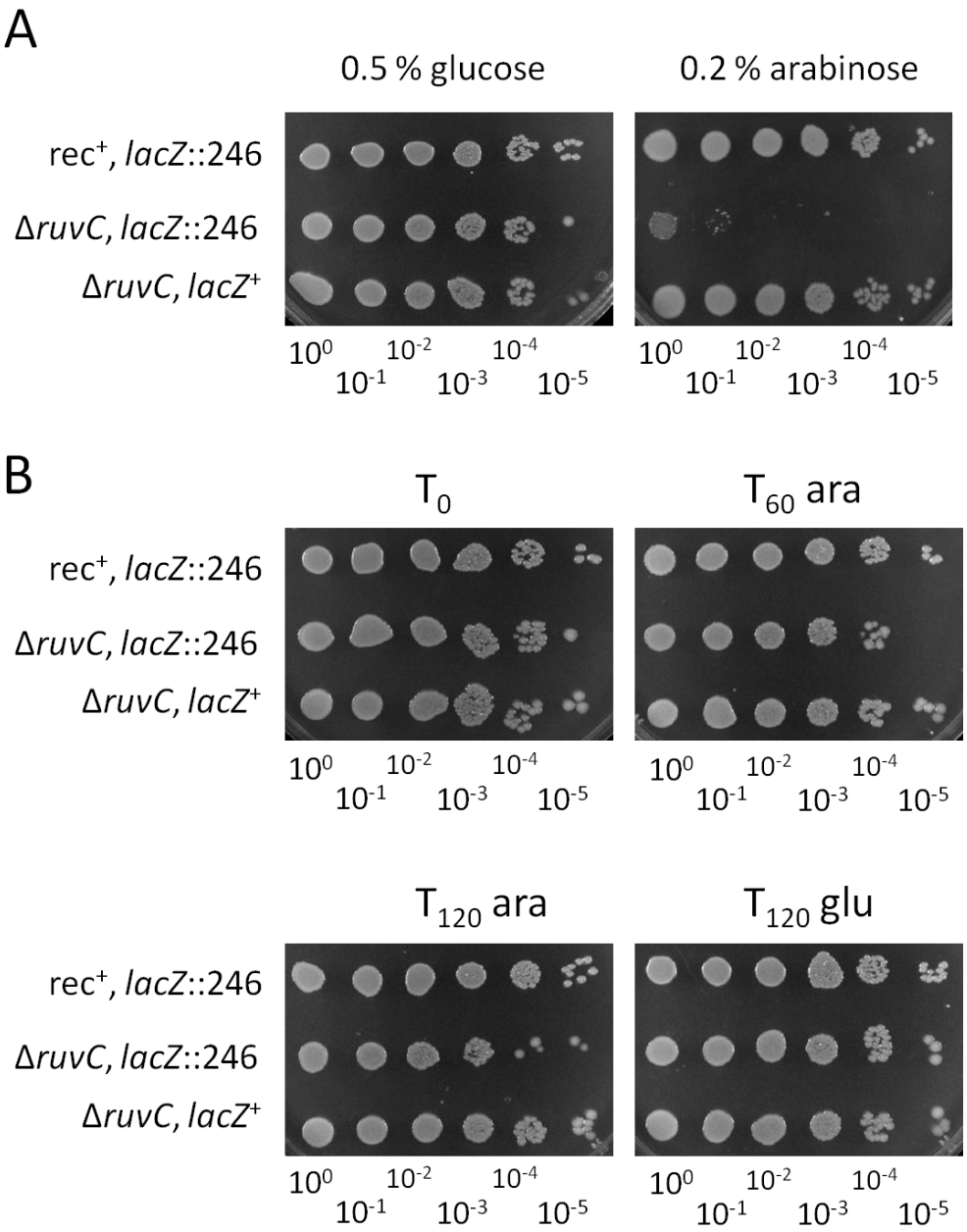

Figure 3. 7 Requirements for RuvC in conditions of chronic and acute double strand break formation

(A) Viability spot tests of Δ*ruvC* strains in conditions of chronic DSB formation. (B) Viability spot tests of Δ*ruvC* strains in conditions of acute DSB formation. Strains used were DL2006 (Rec$^+$ *lacZ*::246), DL4724 (Δ*ruvC lacZ*::246), and DL4725 (Δ*ruvC lacZ*$^+$).

migration activity of RuvAB that is required for survival but also the nuclease activity of RuvC. However, the Δ*ruvC lacZ*::246 strain grown in 0.2% Arabinose for short exposure times behaves differently to the Δ*ruvAB lacZ*::246 strain grown under the same conditions (Fig. 3.4 A). Whereas the latter strain did not seem to lose any viability, the viability of Δ*ruvC lacZ*::246 decreased 10-fold after 120 minutes of DSB formation (Fig. 3.7). These observations are preliminary as the full CFU potential of this culture under these conditions has not yet been tested for. Nevertheless, these initial observations support the idea that a limited number of HJs can be resolved in a RuvAB-independent manner in a context where branch migration has likely been abolished.

3.5 Discussion

The work presented here illustrates the essential requirement for RecA-mediated HR in the repair of DSBs formed by SbcCD at a 246 bp interrupted palindrome in the chromosomal *lacZ* gene of *E. coli* (Figs. 3.1 and 3.2). It also shows that under conditions of prolonged (chronic) DSB formation, the HJ branch migration complex RuvAB, the HJ nuclease RuvC, and the recombination helicase RecG are indispensable for the survival of *E. coli* (Figs. 3.3 and 3.7 A). Nevertheless, there is no requirement for RuvAB in repairing a limited number of DSBs (acute DSB formation), whereas a requirement for RecG remains as 90 % of cells in a population that lacks RecG died when subjected to one or two DSBs (Figs. 3.4 and 3.5). This result suggests that when one or two DSBs are introduced in a RuvAB

mutant, HJs are either never formed, or are resolved in a RuvAB-independent manner. In a culture lacking RuvC but maintaining RuvAB, sensitivity to a limited number of DSBs is also acquired (Fig. 3.7 B). This phenotype might be a direct effect from having lost the nuclease activity of RuvC or an indirect effect from having retained the branch migration activity of RuvAB. In order to distinguish between these two scenarios, a Δ*ruvABC* mutant would need to be studied. If the former event was occurring, a Δ*ruvABC* mutant ought to be as sensitive to DSB formation as a Δ*ruvC* mutant. If the latter scenario were correct, a Δ*ruvABC* mutant would retain full viability after 120 minutes of exposure to 0.2 % Arabinose, just like a Δ*ruvAB* strain. In addition to studying the genetics of a Δ*ruvABC* mutant, physical analysis of the DNA surrounding the breakpoint would help to distinguish between the two suggested mechanisms for RuvAB-independent survival under conditions of acute DSB formation, as these experiments would permit the investigation of the physical abundance of HJs. If RuvC is required for the resolution of HJs in the absence of RuvAB, the abundance of HJs in Δ*ruvABC* and Δ*ruvAB* mutants ought to be different, with more being detected in a Δ*ruvABC* strain. If the alternative scenario applies, and a Δ*ruvABC* mutant is as viable as a Δ*ruvAB* mutant, the abundance of HJs in these two strains ought to be the same.

If the analysis of a Δ*ruvABC* mutant points to the conclusion that HJs can be dissolved by branch migration through the action of a helicase, it would be interesting to determine which helicase is responsible for this. RecG, which has been shown to act on HJs *in vitro*, is a possible candidate. Indeed, a Δ*ruvAB* Δ*recG* strain is more sensitive to DSBs than a Δ*ruvAB* strain. Determining whether this amplified death is due to the accumulation of unresolved HJs alone or whether it is

due to the accumulation of HJs and a different kind of intermediate (generated by the absence of RecG) could be achieved by native two dimensional gel electrophoresis. An alternative candidate for dissolving HJs independently of RuvAB is RecQ, as some of the eukaryotic homologues of this helicase have been implicated in having this role in DSB repair (Ashton et al., 2011; Dayani et al., 2011; Karow et al., 2000).

The data presented in this chapter also indicates that the recombination intermediates accumulated in a Δ*recG* mutant cannot be resolved by the action of any mechanism that is not dependent on RecG. Nevertheless, the complementation experiment shows that increasing the levels of RuvAB beyond those that are physiologically relevant can rescue the death of a strain lacking RecG (Fig. 3.6). Increased levels of RuvAB could allow for the RecG-independent resolution of the intermediates accumulated in a Δ*recG* background. Alternatively, these intermediates may never be formed when elevated levels of RuvAB are present. If the former hypothesis were true, it is not yet possible to determine whether it is the branch migration activity of RuvAB that is necessary for survival or whether this survival is dependent on endogenous RuvC and therefore on the nuclease activity. Constructing a Δ*recG* Δ*ruvC* mutant, and elevating the levels of RuvAB using pGB*ruvAB* to investigate whether the death phenotype of this mutant is rescued, would help to address this question. However, being able to distinguish between intermediates being resolved or never having been formed in the first place remains a difficult task.

Finally, it is interesting to note that under conditions of no DSB formation, both in the presence and absence of the palindrome, the Δ*ruvAB* Δ*recG* mutant is

sicker than the Δ*recA* mutant (Figs. 3.1 and 3.3). This is a surprising result as it suggests that for normal growth the RuvAB and RecG proteins together are more essential than RecA. One possible explanation is that during unchallenged growth certain events occur that require the activity of RuvAB and RecG to allow for cell survival in a RecA-independent manner. Alternatively, it is plausible that in a chromosomal environment where both RuvAB and RecG are not present, forms of DNA damage, which would be reliant on RecA-mediated recombination for repair, arise at a higher frequency. This would increase the cell's requirement for RecA-mediated recombination and therefore for the activity of both RuvAB and RecG. It would be interesting to construct a Δ*recA* Δ*ruvAB* Δ*recG* mutant. If death in a Δ*ruvAB* Δ*recG* mutant is caused by events that are not dependent on RecA-mediated recombination then a Δ*recA* Δ*ruvAB* Δ*recG* mutant should be as sick as a Δ*ruvAB* Δ*recG* mutant. If, on the other hand, a Δ*ruvAB* Δ*recG* mutant is accumulating forms of damage that require RecA-mediated recombination for repair, and initiating RecA-mediated recombination results in death because repair cannot be completed in a context where RuvAB and RecG are not present, then introducing the Δ*recA* mutation in a Δ*ruvAB* Δ*recG* background may restore the death phenotype to that of a Δ*recA* single mutant.

Chapter 4

Locating DNA recombination intermediates by pulsed-field gel electrophoresis

4.1 Introduction

Repair of a site-specific DNA DSB involves the formation of recombination intermediates. In order to analyse the structure of these intermediates, their localisation with respect to the breakpoint needs to be identified. Locating intermediates can be achieved using pulsed-field gel electrophoresis, a technique used to separate large chromosomal DNA fragments (>30 Kb). This technique differs from standard agarose gel electrophoresis in that the voltage is regularly switched between two directions that run at an angle of 120° either side of the gel. At each change of direction, the DNA fragments need to be re-oriented before restarting their migration. Larger fragments require more time to be re-oriented than smaller fragments, and it is this re-orientation time that allows for their separation. In addition to size, the shape of a DNA molecule also affects its electrophoretic mobility. Only linear, un-branched DNA migrates into a pulsed-field gel, meaning that this technique can be used to determine whether different

chromosomal fragments contain branched recombination intermediates upon inducing DNA damage.

The first part of this chapter describes the construction of strains designed to concentrate recombination intermediates close to the breakpoint in order to facilitate their detection. This is achieved by inserting three repeats of the Chi cross-over hotspot sequence (5′-GCTGGTGG-3′), termed a χ array, 1.5 Kb either side of a 246 bp interrupted palindrome. The chapter goes on to describe the localisation of recombination intermediates generated by the repair of DSBs formed by SbcCD-mediated cleavage of the palindrome in the *E. coli* chromosomal *lacZ* gene. Strains used contained the χ array on both sides of the palindrome and were unable to resolve recombination intermediates (Δ*ruvAB*, Δ*recG* and Δ*ruvAB* Δ*recG* mutants). The chromosomes of these strains were digested with either NotI, I-SceI, SalI, or SacI restriction enzymes to release different sized fragments that were analysed using pulsed-field gel electrophoresis followed by ethidium bromide (EtBr) staining and Southern blotting. DNA was detected using either DIG-labelled probes or ^{32}P-labelled probes.

4.2 Insertion of a χ array in *lacZ* and *mhpR*; a modification that maximises events of homologous recombination in close proximity of the palindrome

When repair of an SbcCD-induced DSB was first characterised, stimulation of recombination either side of the breakpoint was measured using a zeocin resistance reporter cassette (Eykelenboom et al., 2008). It was shown that recombination was stimulated on both sides of the break, suggesting that both free ends were actively involved in the repair process. Furthermore, this stimulation was increased when a single χ site was introduced immediately upstream of the zeocin resistance reporter cassette. For the purpose of this work, in order to maximise events of recombination in close proximity of the palindrome, a χ array was placed 1.5 Kb from the centre of the palindrome on both the origin proximal and origin distal sides, in the *lacZ* and *mhpR* genes, respectively. A schematic representation of this modified region of the chromosome is depicted in Figure 4.1.

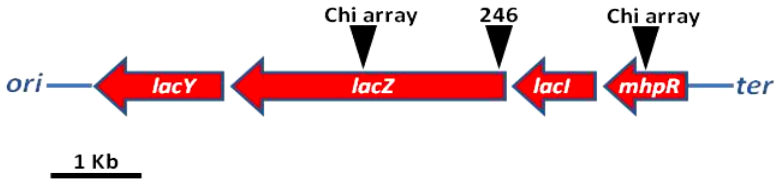

Figure 4. 1 Chromosomal locus of Chi (χ) arrays inserted 1.5 Kb either side of the 246 bp interrupted palindrome

Genes are represented by red arrows. The 246 base pair interrupted palindrome and the χ arrays are marked by black triangles. Origin and terminus are marked in blue as *ori* and *ter*, respectively.

4.2.1 Inserting the χ array into *lacZ*; construction of pDL4137

pDL4137 is a pTOF24-derived plasmid containing an 871 bp sequence homologous to part of the *lacZ* gene with an added χ array around its centre so that, when inserted into the genome, the χ array is placed 1.5 Kb upstream of the centre of the palindrome (Table 2.4). The 871 bp sequence was synthesised by crossover PCR using genomic DNA isolated from *E. coli* MG1655 as template. Primers *lacZ*chi 1 and *lacZ*chi 2 were used to amplify a 430 bp region immediately upstream of the χ array insertion point, while primers *lacZ*chi 3 and *lacZ*chi 4 were used to amplify a 441 bp region immediately downstream of this point (Table 2.2). The χ array was added to the 5′ ends of the *lacZ*chi 2 and *lacZ*chi 3 primers, forming a 23 bp region of homology between these two oligonucleotides. The cross-over PCR, using primers *lacZ*chi 1 and *lacZ*chi 4 and both the 430 bp and 441 bp products from the first two PCR reactions as templates, generated a final PCR product of 871 bp containing a full χ array of 44 bp. The 871 bp fragment was digested with PstI and SalI (restriction sites located at the 5′ ends of primers *lacZ*chi 1 and *lacZ*chi 4, respectively) and ligated into PstI/SalI digested pTOF24. The resulting pDL4137 plasmid was verified by sequencing using primers pKO F and pKO R (Table 2.2). This pTOF24 derivative was used to carry out plasmid-mediated gene replacement (PMGR) in the background strain of choice (Chapter 2 Fig. 2.1). All strains successfully modified by PMGR were checked by sequencing.

4.2.2 Inserting the χ array into *mhpR*; construction of pDL4138

pDL4138 is a pTOF24-derived plasmid containing an 880 bp sequence homologous to a region upstream, and including the beginning of, the *mhpR* gene and encoding around its centre an added χ array, which when inserted into the genome is placed 1.5 Kb downstream from the centre of the palindrome (Table 2.4). The 880 bp sequence was synthesised by cross-over PCR using similar principles to the ones used for the construction of pDL4137. Primers used were *mhpR*chi 1, *mhpR*chi 2, *mhpR*chi 3 and *mhpR*chi 4 (Table 2.2).

4.3 Analysis of large chromosomal fragments; digestion of the chromosome with NotI

Strains of interest were allowed to grow in L-broth at 37°C with shaking, until they reached mid-exponential growth phase. Cultures were then split in two, and one sample was exposed to 0.5 % glucose (to repress the expression of *sbcCD*), while the other was exposed to 0.2 % arabinose (to induce the expression of *sbcCD*). Cultures were then put back at 37 °C, with shaking, and allowed to grow for the desired time, after which chromosomal DNA was prepared for analysis. Chromosomal DNA was prepared by mixing bacterial cultures with liquid agarose. This mix was pipetted into moulds and allowed to set, forming well-shaped plugs containing 1% agarose. Once set, the cells within the plugs were lysed and treated

with proteinase K to release the DNA and digest cellular proteins. DNA was extracted by this method in order to minimise shearing and maximise the retention of branched DNA species and has been shown to be as effective in trapping branched DNA as intra-strand DNA crosslinking (Cromie et al., 2006). In order to gain a global picture of the integrity of the chromosome, plugs containing chromosomal DNA were digested *in vitro* using the NotI restriction enzyme, which, under normal conditions, released 23 linear fragments. These fragments were separated by pulsed-field gel electrophoresis (Fig. 4.2 A shows a map of NotI sites on the *E. coli* chromosome).

4.3.1 Analysis of chromosome integrity; ethidium bromide staining of NotI digested chromosomal DNA of Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* mutants

Due to the large sizes of the fragments generated after digestion using NotI, it was possible to study the banding pattern generated upon DSB formation after staining of the pulsed-field gel with ethidium bromide (EtBr). Indeed, a recent study of EcoKI-mediated DSB formation in *E. coli* showed that strains unable to resolve recombination intermediates accumulated intertwined chromosomal fragments, which remained in the wells of pulsed-field gels and failed to produce a NotI banding pattern (Wardrope et al., 2009). In the work presented here, a clear NotI banding pattern was seen for all samples, including those from cells in which DSBs were induced (Fig. 4.2 B lanes 3, 4, 7, 8, 11, and 12). This result suggests that

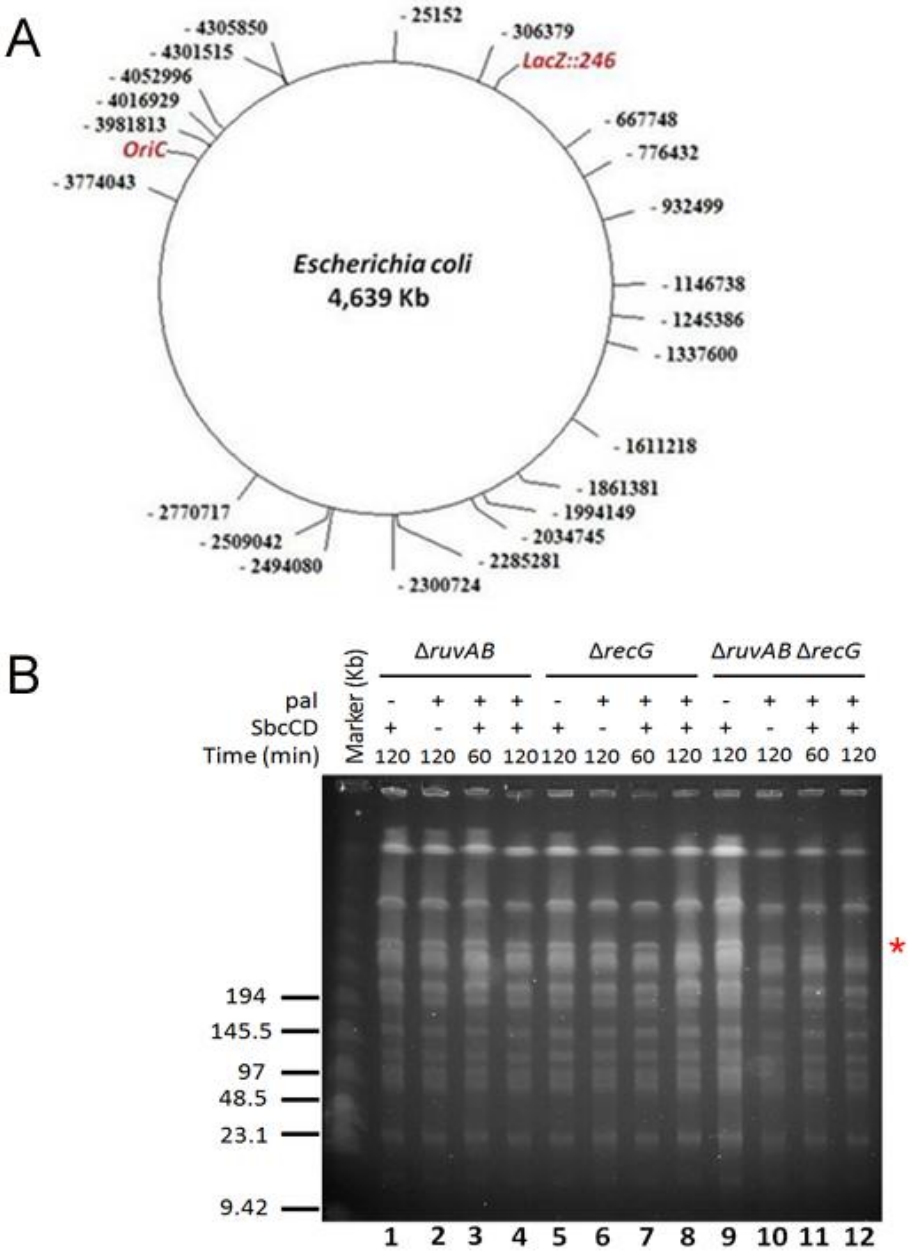

Figure 4. 2 Ethidium Bromide staining of NotI digested chromosomal DNA

(A) NotI restriction map of the *E. coli* chromosome. Restriction sites are shown in base pairs. The origin of replication (OriC; shown in red) is located at 3,923,640 base pairs and the palindrome (*lacZ*::246 shown in red) at 365,519 base pairs. (B) Ethidium bromide staining of NotI digested chromosomal DNA separabinoseted by pulsed-field gel electrophoresis. Marker used: NEB low range Pulsed-field gel marker. Time (min) represents the length of time in which the cultures were grown in the presence or absence of SbcCD. Strains used were DL4243 (Δ*ruvAB lacZ*::246), DL4257 (Δ*ruvAB lacZ*$^+$), DL4311 (Δ*recG lacZ*::246), DL4312 (Δ*recG lacZ*$^+$), DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), DL4313 (Δ*ruvAB* Δ*recG lacZ*$^+$).

upon formation of a site-specific DSB the majority of the chromosome does not accumulate branched DNA species. Nevertheless, there seemed to be a reduction in the amount of a fragment of high molecular weight in the Δ*ruvAB* Δ*recG* mutant following DSB formation (Fig. 4.2 B lane 12; the position of the band is marked by a red *). As staining of DNA with EtBr is unspecific, it was not possible to conclude whether the disappearing fragment, which disappeared upon DSB formation, was located close to, or contained, the palindrome. In order to increase the resolution of the pulsed-field gel analysis and look at the region surrounding the breakpoint in more detail, Southern blotting using probes specific for the NotI fragments surrounding the palindrome was carried out.

4.3.2 Southern blots of NotI fragments surrounding *lacZ*

Δ*ruvAB* Δ*recG* mutants were used for this experiment, as the disappearance of a fragment of high molecular weight was seen in this mutant background when DSBs were induced (Fig. 4.2 B lane 12; the position of the band is marked by a red*). Following digestion of plugs with NotI and separation of the DNA fragments by pulsed-field gel electrophoresis, as shown in Figure 4.2, the DNA was transferred to a positively charged nylon membrane by Southern blotting. The DNA was fixed to the membrane by UV light cross-linking, and DIG-labelled probes specific for the NotI fragments containing, or surrounding, the palindrome (*lacZ*) were used to probe the membrane. Figure 4.3 A displays the four NotI fragments, their respective sizes (Kb), and the probes used to detect them.

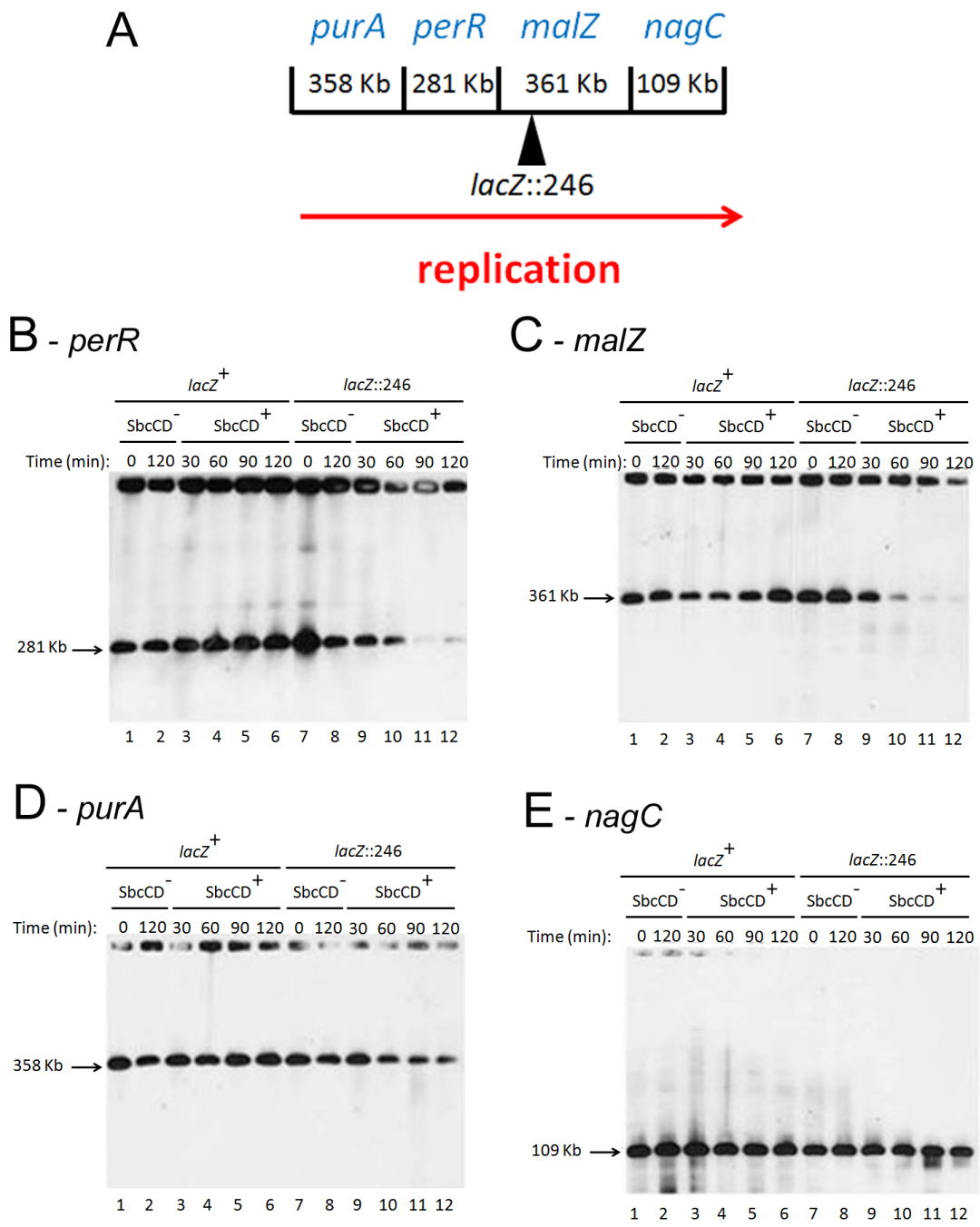

Figure 4. 3 Detection of NotI digested chromosomal DNA using DIG-labelled probes for fragments containing or surrounding *lacZ*

(A) Diagram of NotI fragments and sizes (Kb) containing or surrounding *lacZ*, and the respective probes (shown as black rectangles) used to detect them (diagram not to scale). (B-E) Detection of DNA fragments using *perR*, *malZ*, *purA*, and *nagC* probes, respectively. Strains used were DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), and DL4313 (Δ*ruvAB* Δ*recG lacZ*$^+$).

Figure 4.3 B to E presents the results of the Southern blotting using *perR*, *malZ*, *purA*, and *nagC* probes, respectively. Each of these probes revealed a unique fragment at the expected size. The signal for the linear 361 Kb *malZ* fragment, which contained the palindrome, gradually decreased as exposure to arabinose, and therefore to DSBs, increased from 60 minutes to 120 minutes (Fig. 4.3 C lanes 10-12). This disappearance was not seen when the *lacZ*$^{+}$ variant of this strain expressed SbcCD (grown in the presence of 0.2 % arabinose for 120 minutes; lane 6) or when the *lacZ*::246 variant of the strain did not express SbcCD and therefore was not subject to DSBs (grown in 0.5 % glucose for 120 minutes; Fig. 4.3 C lanes 8). These observations indicate that in the Δ*ruvAB* Δ*recG* mutant, a 361 Kb fragment containing the palindrome disappears from the gel as a result of SbcCD-mediated cleavage at the palindrome. Further probing revealed that this disappearance was not only restricted to the fragment that contained the palindrome. The fragment directly upstream, detected with the probe *perR*, also disappeared from the gel (Fig. 4.3 B lanes 9-12). Interestingly, this disappearance was delayed by 30 minutes compared to the *malZ* fragment. The two adjacent fragments, *purA* and *nagC* (Fig. 4.3 D and E), did not show the same pattern of disappearance, although it should be noted that there may have been a small reduction in the amount of the *purA* fragment following the induction of DSBs (Fig. 4.3 D, lanes 9-12), but this reduction was not as dramatic as for the *malZ* and *perR* fragments. The results presented in this experiment demonstrate that inducing a DSB at a 246 bp interrupted palindrome in the chromosomal *lacZ* gene of a Δ*ruvAB* Δ*recG* mutant causes loss of the linear DNA within a region of 642 Kb surrounding the breakpoint. This observation is consistent with the results obtained when DNA from the same

mutant was run on a pulse-field electrophoresis gel and stained with EtBr, as shown in Figure 4.2 (lanes 12 and 13), where the bulk of the chromosome appeared to be in a linear conformation and the disappearance of a fragment of high molecular weight, presumably either the *malZ* or the *perR* fragment, was noted.

Disappearance of DNA is normally attributed to degradation at the hands of cellular nucleases. If a DSB is formed but not repaired, for example in a Δ*recA* mutant, linear DNA persists in the cell. Cellular nucleases would eventually start degrading this DNA, resulting in the potential loss of an entire chromosome (White et al., 2008). This scenario is not expected to occur in Δ*ruvAB* Δ*recG* mutants as they ought to be able to undergo DSB repair up to the formation of joint molecules. Joint molecules are not a substrate for the action of cellular exonucleases. Assuming that this hypothesis is correct, the disappearance of the fragments shown in Figures 4.2 and 4.3 could be due to the accumulation of DSB repair intermediates, which prevent the DNA from migrating into the gel, as was reported by Wardrope and collaborators (Wardrope et al., 2009). These species should therefore be present in the gel's wells. Nevertheless, it is also possible that joint molecules are not formed stably in a Δ*ruvAB* Δ*recG* mutant as RecG has been implicated in stabilising D-loops that are generated by RecA-mediated strand invasion (Whitby et al., 1993). Absence of this stabilising activity may result in abortion of the repair process and persistence of DNA free ends that become degraded over time, resulting in loss of DNA. Analysing a smaller fragment surrounding the breakpoint may help to distinguish between these two hypotheses as loss of DNA by degradation ought to be detected close to the breakpoint but recombination intermediates may not be, as

these structures may be formed further away from the breakpoint or, if made close to the DSB, may migrate from their point of origin leaving linear DNA behind.

4.4 A more detailed analysis of the DNA surrounding the palindrome; digestion of the chromosome with the rare-cutting endonuclease, I-SceI

Digestion of the chromosome with I-SceI released a 174 Kb fragment containing *lacZ* 104 Kb from the origin proximal restriction site and 70 Kb from the origin distal restriction site. Probes *perR* and *malZ*, both of which bind to this fragment, were used to detect the DNA (Fig. 4.4 B). The linear 174 Kb DNA fragment disappeared upon DSB formation, as seen after the digestion with NotI shown in the previous section (Fig. 4.3). A control DNA sample (*lacZ*::I-SceI), from a strain containing an additional I-SceI restriction site in place of the palindrome, was loaded into the first lane of each gel shown. Using this strain as a marker for cleavage at the palindrome, the results exposed in Figure 4.4 indicates that, after induction of DSBs no significant amount of broken DNA was detected upstream or downstream of the palindrome (B lanes 11-14 and C lanes 11-14). Nevertheless, it is not possible to exclude the possibility that broken DNA is formed and then degraded by the exonuclease RecBCD as has been previously reported (Clark and Chamberlin, 1966; Cleaver and Boyer, 1972; Willetts and Clark, 1969). In order to determine whether the DNA was being degraded or was accumulating in the wells

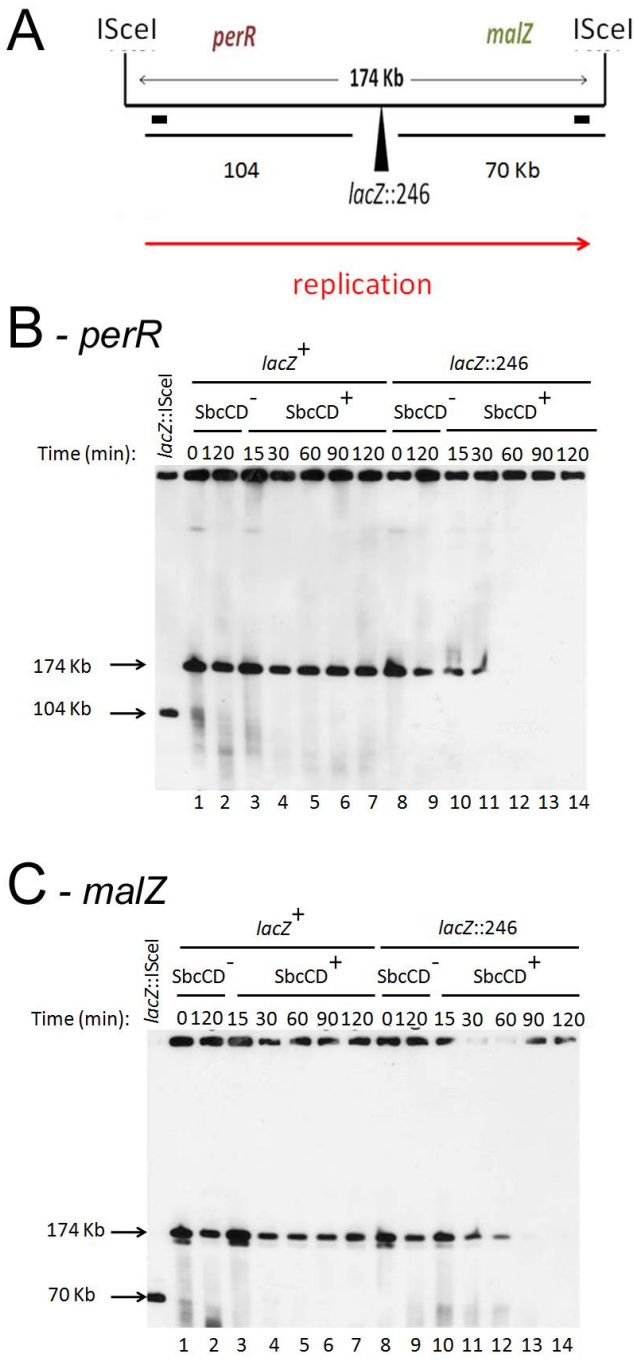

Figure 4. 4 Detection of I-SceI digested chromosomal DNA using DIG-labelled probes for a fragment containing *lacZ*

(A) Diagram of the I-SceI fragment and size (Kb). The location of *lacZ* and the position of *perR* and *malZ* probes (shown as black rectangles) are indicated (diagram not to scale). Detection of DNA using *perR* (B) or *malZ* (C) probes. Strains used were DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), DL4313 (Δ*ruvAB* Δ*recG lacZ*$^+$), and DL2849 (Rec$^+$ *lacZ*::I-SceI).

of the pulsed-field gel, the abundance of fragments surrounding the DSB would need to be quantified in relation to a control fragment located on the opposite side of the chromosome, the abundance of which should not be affected by DSB formation at the palindrome. In order to quantify DNA following Southern blot, ^{32}P-labelled probes need to be used for the detection.

4.5 Quantitative study of recombination intermediates following the induction of a DSB in *lacZ*

In order to improve the sensitivity of the Southern blots and allow for the quantification of the DNA detected, the protocol was modified from using DIG-labelled probes to using ^{32}P-labelled probes. DIG-labelling requires the use of an antibody for the detection of the DIG-labelled UTP in the probe's sequence, whereas ^{32}P-labelled probes can be detected directly without the need for an antibody binding step. This direct approach makes the technique more reliable. In addition, the radioactive counts emitted by ^{32}P-labelled probes can be quantified using a variety of software. For the purpose of this work, the GE healthcare software ImageQuant™ TL was used.

4.5.1 In detail analysis of a 126.8 Kb region surrounding the palindrome; digestion of the chromosome with SalI

A 174 Kb region surrounding the palindrome was identified as a region in which DSB repair intermediates may accumulate following DSB formation. To understand the distribution of these intermediates and determine whether or not this region was being degraded upon DSB formation, SalI digestions of the chromosomes from Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* mutants were carried out. SalI digestion of the palindrome region generated three major fragments of 34.1 Kb, 23.7 Kb and 36.4 Kb that were recognised by probes *yagV*, *lacZ*, and *araJ*, respectively (Fig. 4.5 A). Notably, the 23.7 Kb *lacZ* fragment contained the palindrome. In addition, a fourth probe, *cysN*, was designed to detect a SalI fragment of 31.7 Kb on the opposite side of the chromosome, to be used as a control.

4.5.1.1 Accumulation of branched DNA

For the purpose of quantifying the information obtained from any one lane of the Southern blots so as to detect the accumulation of DNA in the wells (where recombination intermediates are predicted to accumulate), the background signal emitted by the membrane was subtracted from the signals emitted from the DNA in the well and the DNA in the single linear band. Following background subtraction, the proportion of DNA in the well over the DNA in the linear band was established. This proportion was normalised to the proportion of DNA in the well obtained for

the control strain, *lacZ*$^+$ grown in 0.2 % arabinose for 60 minutes, probed with the same probe. This control was chosen over the *lacZ*::246 variant of the strain at either time 0 minute or time 60 minutes in 0.5 % glucose because of the possible leaky expression of *sbcCD* in this background.

Figure 4.5 A shows the SalI restriction map for the region analysed, while panels B and C show, respectively, the Southern blots performed on the DNA of Rec$^+$ strains and the quantification of the DNA accumulated in the wells of the pulsed-field gel as described above. In Rec$^+$ conditions, a very low signal was detected in all wells of the Southern blots (Fig. 4.5, panels B and C). This finding is not surprising as intermediates of DSB repair ought to be very transient in a recombination proficient background, making their detection difficult. Nevertheless, there appeared to be a marginal increase of the *lacZ* fragment in the well containing DNA from a strain in which DSBs were induced (*lacZ*::246 strain grown in 0.2 % arabinose, Fig. 4.5 C). To support this observation, a one-way analysis of variance (ANOVA), comparing the values obtained for the *lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % Glucose, and 60 minutes in 0.2 % Arabinose, showed that for both the yagV and lacZ fragments there was a significant accumulation of DNA in the wells upon induction of DSBs ($P = 0.016$ and $P = 0.014$, respectively; Table 4.1). This result suggests that recombination intermediates are accumulated in both the *yagV* and *lacZ* fragments. It is important to note that this accumulation is more prominent in the *lacZ* fragment and no significant accumulation was detected in either the *araJ* or *cysN* fragments (Table 4.1). These branched DNA species probably do represent DSBR intermediates, presumably on their way to being resolved.

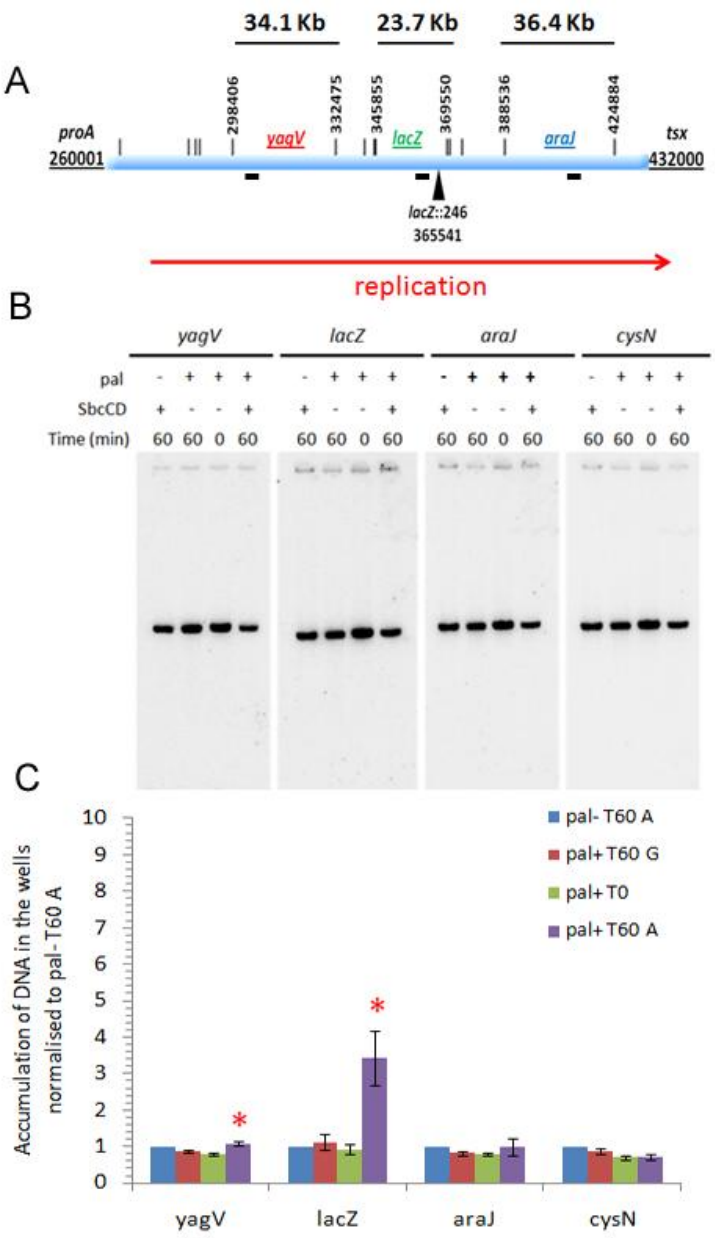

Figure 4. 5 Pulsed-field gel electrophoresis and Southern blot analysis of SalI digested chromosomal DNA of Rec$^+$ strains

(A) SalI digestion map of the E. coli chromosome from coordinates 260,001 (*proA*) to 432,000 (*tsx*). SalI restriction sites are marked by a black vertical line. Restriction sites of interest to this analysis are marked with the coordinates, the probes used to detect the fragments (*yagV*, *lacZ*, and *araJ*) are indicated above the respective fragment and their binding site is represented below, by black rectangles. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4184 (Rec$^+$ *lacZ*::246) and DL4201 (Rec$^+$ *lacZ*$^+$). (C) Quantification of DNA present in the wells of the Southern blots. Red * indicates data

that is significantly different from controls. Error bars are standard error of the mean where n = 3.

When the same analysis was carried out using DNA from Δ*ruvAB* strains, a low signal was detected in the wells of all control strains lacking the palindrome (*lacZ*$^+$) or containing the palindrome (*lacZ*::246) but in which the expression of *sbcCD* was not induced (time point 0 minute and time point 60 minutes in 0.5 % glucose; Fig. 4.6 B, the first three lanes of each gel). In contrast, when DSBs were formed (*lacZ*::246 grown in 0.2 % arabinose), a significant amount of the *lacZ* fragment (up to 70 times more than the control strain *lacZ*$^+$ strain grown in 0.2 % arabinose) accumulated in the well. This dramatic accumulation appeared to be restricted to the 23.7 Kb *lacZ* fragment as the *yagV*, *araJ*, and *cysN* fragments did not appear to accumulate branched DNA upon induction of DSBs (Fig. 4.6 panels B and C). Nevertheless, statistical analysis shown in Table 4.1 shows that there was a small, but highly significant, amount of branched DNA detected in the *yagV* and *araJ* fragments ($P < 0.01$ for both fragments). No significant accumulation of branched DNA was detected in the control *cysN* fragment (Table 4.1). This result strengthens the hypothesis that DSB repair intermediates generated from a site-specific DSB arise locally to the breakpoint. Additionally, it suggests that HJs are an intermediate of the repair process, as it is these species of branched DNA that are expected to accumulate in the absence of RuvAB.

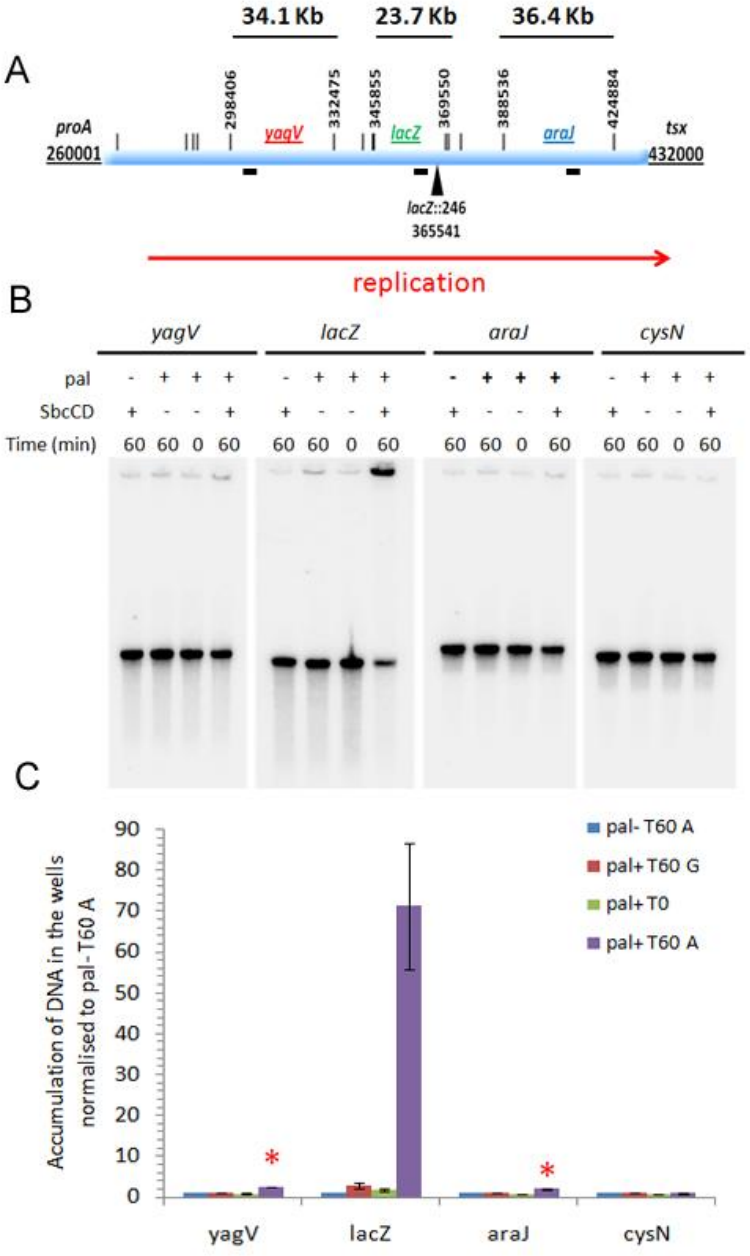

Figure 4. 6 Pulsed-field gel electrophoresis and Southern blot analysis of SalI digested chromosomal DNA of Δ*ruvAB* strains

(A) SalI digestion map of the *E. coli* chromosome from coordinates 260,001 (*proA*) to 432,000 (*tsx*). SalI restriction sites are marked by a black vertical line. Restriction sites of interest to this analysis are marked with the coordinates, the probes used to detect the fragments (*yagV, lacZ,* and *araJ*) are indicated above the respective fragment and their binding site is represented below, by black rectangles. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4243 (Δ*ruvAB lacZ*::246), DL4257 (Δ*ruvAB lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Red * indicates data that is significantly different from controls. Error bars are standard error of the mean where n = 3.

The results obtained with the DNA from the Δ*recG* mutants are shown in Figure 4.7. As seen previously with the DNA of the Rec$^+$ and the Δ*ruvAB* strains, intermediates accumulated in the Δ*recG* mutant were also localised to the region of the chromosome surrounding the breakpoint, as shown by the significant absence of branched DNA in the *cysN* fragment (Fig. 4.7 panels B and C; Table 4.1). Additionally, no branched DNA was detected in the DNA from control strains in which DSBs were not induced. When DSBs were induced, a lot of DNA seemed to be accumulated in the wells of all three SalI fragments surrounding the site of the palindrome, which was strikingly different to the results obtained with DNA from the Rec$^+$ and Δ*ruvAB* strains. Nevertheless this accumulation was more prominent and significant for the *yagV* and *lacZ* fragments (Table 4.1). Fragments beyond *yagV* were not probed for and therefore it is not possible to determine how much further away from the breakpoint these branched DNA species may be found in a Δ*recG* mutant. Another interesting observation is that the accumulation upstream of the breakpoint (in the *yagV* fragment, which is located 33 Kb upstream of the palindrome) appeared to be more prominent than the accumulation downstream (in the *araJ* fragment, which is located 23 Kb downstream of the palindrome). A two-sample T-test comparing the *yagV* and *araJ* values for the *lacZ*::246 strain grown in 0.2 % arabinose shows that a significantly higher amount of branched DNA was located upstream of the breakpoint ($P = 0.013$). These results conclude that branched DNA species are accumulated preferentially upstream of the breakpoint in a Δ*recG* mutant. This data is in accordance with the data obtained from the DNA isolated from the Rec$^+$ strain (Fig. 4.5).

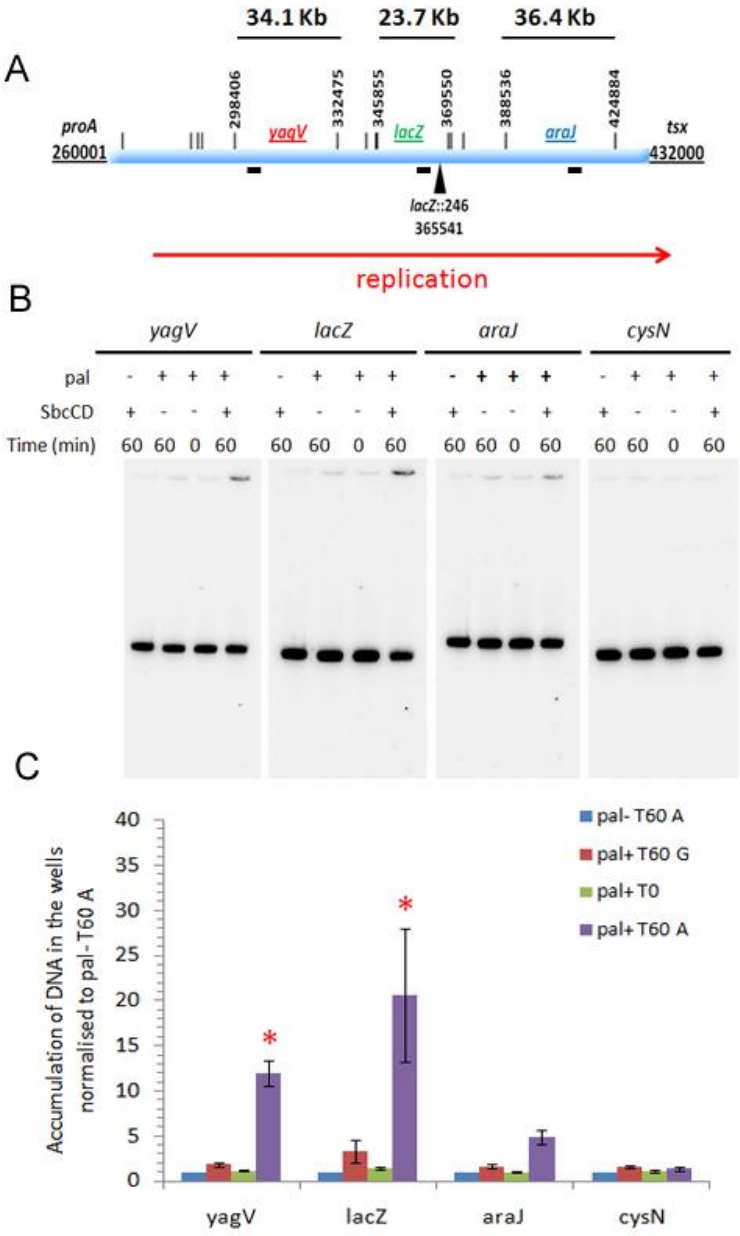

Figure 4. 7 Pulsed-field gel electrophoresis and Southern blot analysis of SalI digested chromosomal DNA of Δ*recG* strains

(A) SalI digestion map of the *E. coli* chromosome from coordinates 260,001 (*proA*) to 432,000 (*tsx*). SalI restriction sites are marked by a black vertical line. Restriction sites of interest to this analysis are marked with the coordinates, the probes used to detect the fragments (*yagV*, *lacZ*, and *araJ*) are indicated above the respective fragment and their binding site is represented below, by black rectangles. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4311 (Δ*recG lacZ*::246), DL4312 (Δ*recG lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Red * indicates data that is significantly different from controls. Error bars are standard error of the mean where n = 3.

As was seen when using DNA from the Δ*recG* mutant, upon DSBs formation DNA from the Δ*ruvAB* Δ*recG* mutant appeared to generate branched DNA across a wide area of the chromosome (Fig. 4.8 panels B and C). Nevertheless, the abundance of these intermediates was not higher than was seen in the single Δ*recG* mutant, suggesting that potentially very little intermediates of repair are generated in a Δ*ruvAB* Δ*recG* mutant. For an easier comparison of the quantifications obtained from the Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG*, data presented in Figures 4.4-4.8 were compiled in Figure 4.9, in which the y-axis of all four bar charts is the same.

4.5.1.2 Degradation of DNA

Inducing DSBs in the Δ*ruvAB* Δ*recG* mutant appeared to result in the accumulation of DNA in the wells of the pulse-field gel and caused the disappearance of the linear *yagV*, *lacZ* and *araJ* fragments. Nevertheless, the accumulation of DNA in the wells did not reflect the disappearance of the DNA that was detected in the gel, suggesting that there was some DNA degradation. To investigate whether DNA is degraded upon DSB formation, the total DNA present on the Southern blots from Figures 4.4 to 4.8 (panels B) was quantified by normalising the sum of the signals emitted from the well and the linear band of a fragment to the sum of the signal emitted from the well and the signal emitted from the linear band of the *cysN* control fragment. The background signal was subtracted from all bands quantified prior to normalisation. The results of this quantification are presented in Figure 4.10.

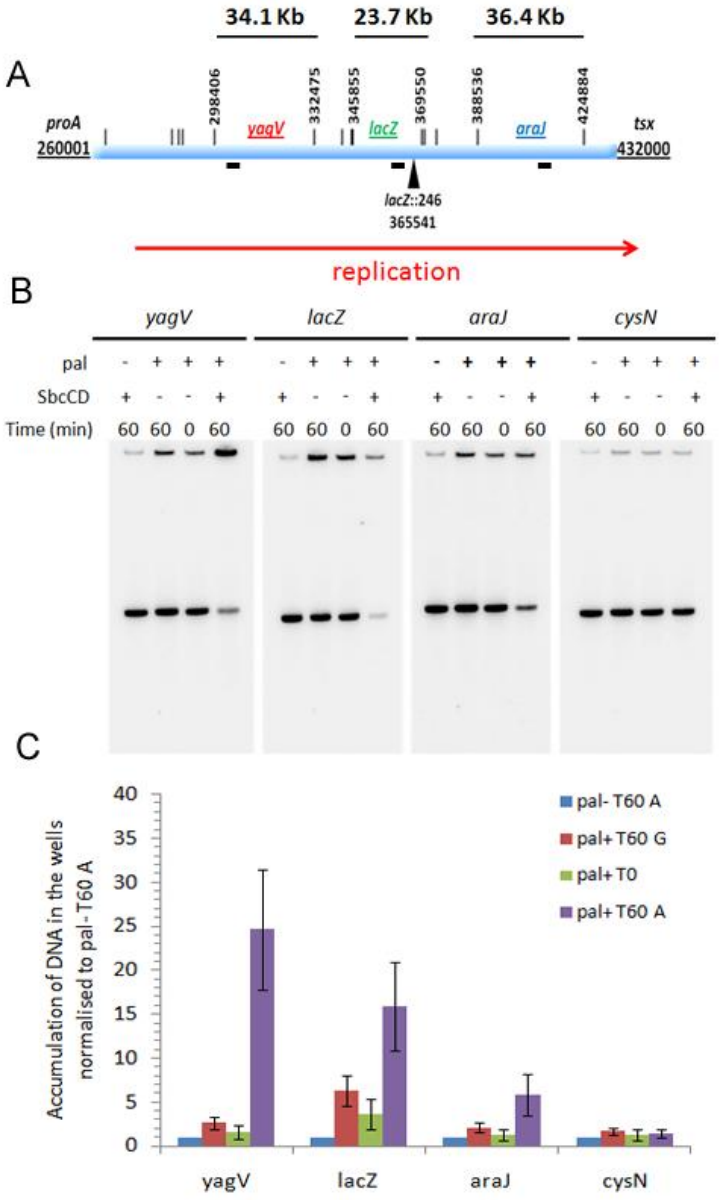

Figure 4. 8 Pulsed-field gel electrophoresis and Southern blot analysis of SalI digested chromosomal DNA of Δ*ruvAB* Δ*recG* strains

(A) SalI digestion map of the *E. coli* chromosome from coordinates 260,001 (*proA*) to 432,000 (*tsx*). SalI restriction sites are marked by a black vertical line. Restriction sites of interest to this analysis are marked with the coordinates, the probes used to detect the fragments (*yagV*, *lacZ*, and *araJ*) are indicated above the respective fragment and their binding site is represented below, by black rectangles. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), DL4313 (Δ*ruvAB* Δ*recG lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Error bars are standard error of the mean where n = 3.

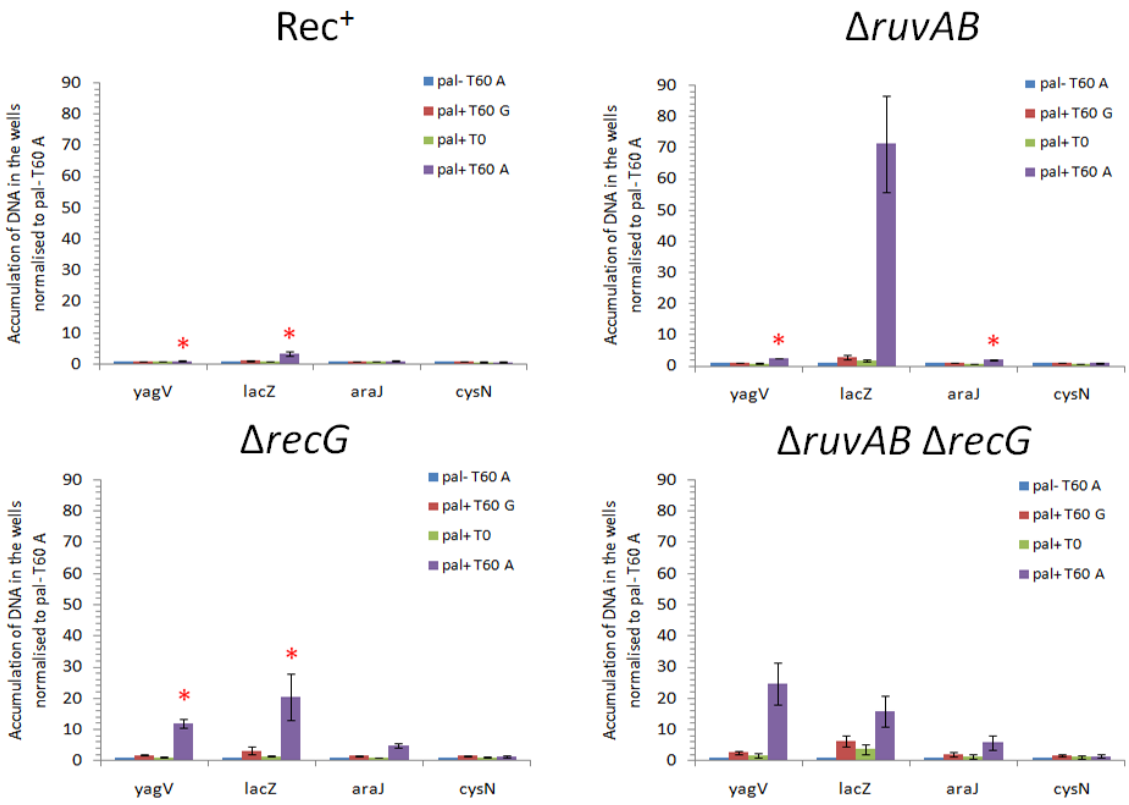

Figure 4. 9 Accumulation of branched DNA in Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains following digestion of the chromosomes with SalI

Bar charts shown in Figures 4.4-4.8 are displayed with the same y-axis values, so as to facilitate the comparison of the results obtained different strains. Red * indicates data that is significantly different from controls.

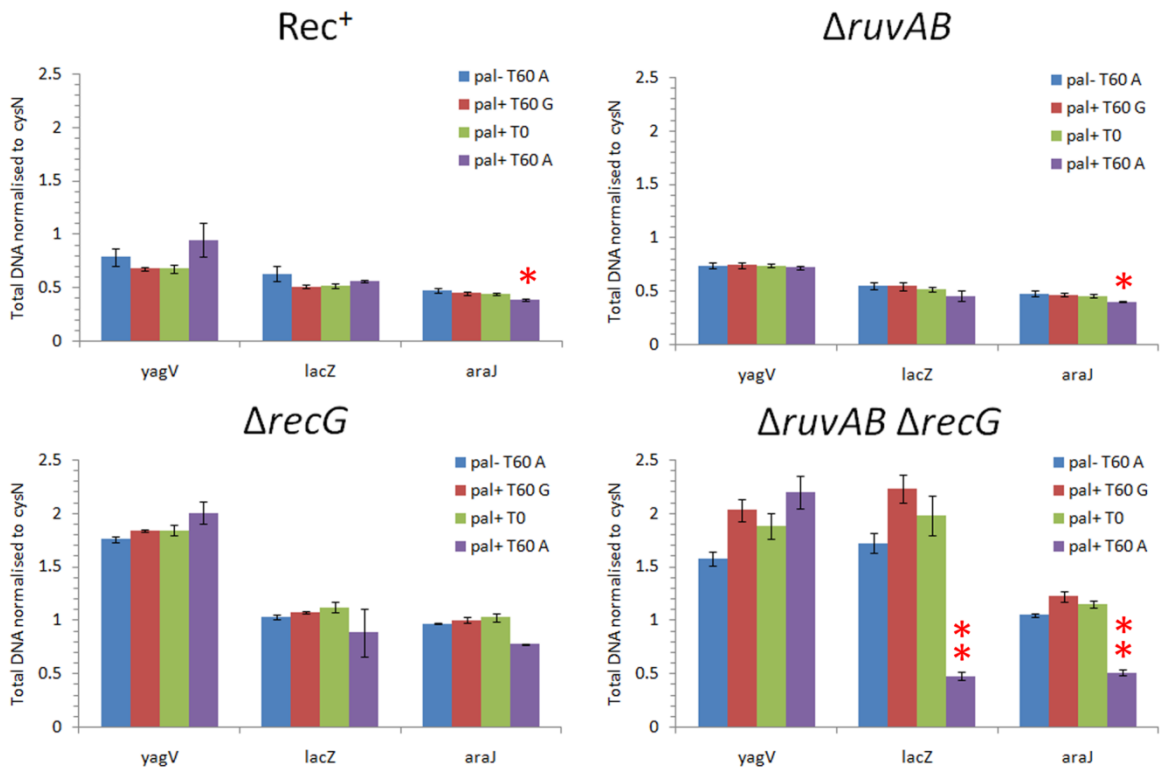

Figure 4. 10 Total amount of DNA in Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains following digestion of the chromosomes with SalI

The total amount of DNA from each sample was quantified and normalised to the total DNA for the *cysN* fragment so as to detect DNA degradation of the region surrounding the breakpoint upon DSB induction. Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4201 (Rec$^+$ *lacZ*$^+$), DL4243 (Δ*ruvAB lacZ*::246), DL4257 (Δ*ruvAB lacZ*$^+$), DL4311 (Δ*recG lacZ*::246), DL4312 (Δ*recG lacZ*$^+$), DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), and DL4313 (Δ*ruvAB* Δ*recG lacZ*$^+$). Red * indicates data that is significantly different from controls, while Red ** indicates a highly significant difference. Error bars represent standard error of the mean where n = 3.

In all the Rec$^+$ strains, there were equal amounts of DNA at the *yagV*, *lacZ*, *araJ*, and *cysN* loci, indicating that the induction of DBSs did not alter the DNA content of the cell in a recombination proficient background. Table 4.2 shows that a small but significant loss of the *araJ* fragment occurred in the Rec$^+$ strain upon DSB induction. A similar pattern was seen in the Δ*ruvAB* mutant, suggesting that despite accumulating large amounts of branched DNA, the cells were still able to maintain a balanced DNA content and no DNA degradation or increase in DNA replication occurred. This observation indicates that stable joint molecules are formed despite the absence of RuvAB. It is interesting to note that the small disappearance of the *araJ* fragment is also significant in the Δ*ruvAB* mutant (Table 4.2). This might suggest that fewer intermediates of repair are accumulated downstream of the breakpoint as seems to be suggested by the accumulation of the branched DNA in the wells (Fig. 4.9). The DNA downstream of, and containing, the breakpoint (the *lacZ* and *araJ* fragments) from the Δ*recG* mutant appeared to be less abundant after induction of DSBs. The statistical analyses presented in Table 4.2 show that this loss is not significant, nevertheless repetition of this analysis may well bring out a significant loss of DNA downstream of the palindrome in the Δ*recG* mutant.

In contrast to the Rec$^+$, Δ*ruvAB*, and Δ*recG* strains, the Δ*ruvAB* Δ*recG* mutant showed a dramatic and highly significant loss of DNA in both the *lacZ* and *araJ* fragments upon DSB induction. A one-way ANOVA, which compared the values obtained for the Δ*ruvAB* Δ*recG lacZ*::246 strain at time point 0 minute, time point 60 minutes in 0.5 % glucose, and time point 60 minutes in 0.2 % arabinose, showed that the loss of DNA for both fragments was highly significant ($P < 0.01$ for both

Table 4. 1 P-values for the accumulation of branched DNA in *lacZ*::246 strains at time point 0 minute, 60 minute in 0.5 % glucose, and 60 minutes in 0.2 % arabinose following digestion of the chromosome with SalI

P-values that are not underlined represent results of a one-way ANOVA, while P-values that are underlined represent the result of a non-parabinosemetric Kruskal-Wallis test. P-values marked by *, represent a significant difference and those marked by **, represent a highly significant difference. P-values in black are not significant.

	P-value for *yagV*	P-value for *lacZ*	P-value for *araJ*	P-value for *cysN*
Rec$^+$	0.016*	0.014*	0.652	0.240
Δ*ruvAB*	0.000**	0.051	0.001**	0.232
Δ*recG*	0.027*	0.027*	0.268	0.268
Δ*ruvAB* Δ*recG*	0.051	0.081	0.130	0.801

Table 4. 2 P-values for the degradation of DNA in *lacZ*::246 strains at time point 0 minute, 60 minute in 0.5 % glucose, and 60 minutes in 0.2 % arabinose following digestion of the chromosome with SalI

P-values that are not underlined represent results of a one-way ANOVA, while P-values that are underlined represent the result of a non-parabinosemetric Kruskal-Wallis test. P-values marked by *, represent a significant difference and those marked by **, represent a highly significant difference. P-values in black are not significant.

	P-value for *yagV*	P-value for *lacZ*	P-value for *araJ*
Rec$^+$	0.113	0.167	0.017*
Δ*ruvAB*	0.754	0.227	0.029*
Δ*recG*	0.229	0.733	0.061
Δ*ruvAB* Δ*recG*	0.329	0.000**	0.000**

fragments). No significant degradation was seen upstream of the breakpoint, in the *yagV* fragment (Table 4.2). In the Δ*ruvAB* Δ*recG* mutant, the remaining DNA at the *lacZ* and *araJ* loci corresponded to half of the DNA content at *cysN*. This result may be an indication that half of the DNA at these loci is being lost, which would correlate with the degradation of the broken chromosome. This would suggest that in this mutant background intermediates of repair are not formed stably and linear DNA persists in the cell. It also suggests that the disappearing fragment detected following digestion of the chromosome with NotI or I-SceI was disappearing due to DNA degradation (Figs. 4.3 and 4.4).

4.5.2 Analysis of the DNA immediately upstream and downstream of the palindrome; digestion of the chromosome with SacI

4.5.2.1 Accumulation of branched DNA

In order to gain more detailed information about the chromosome locus immediately surrounding the breakpoint, a SacI digestion was carried out and analysed as described above for the SalI digests. Two fragments surrounding the breakpoint were released (Fig.4.11 A). The fragment upstream of the DSB was 11.7 Kb, whereas the one downstream, which contained the palindrome within the first 2.1 Kb, was 10.2 Kb in size. These fragments were separated by pulsed-field gel

electrophoresis and probes *lacZ*.prox + 1 and *lacZ*.dist + 1 were used to detect them, respectively. A third fragment of 19.5 Kb on the other side of the chromosome was used as a control and was detected using the *lepA* probe. Quantification of the accumulation of DNA in the wells was carried out as was done for the SalI experiment shown in Section 4.5.1.

Figure 4.11 A shows the SacI restriction map of the region of interest. Figures 4.11 B and C show the Southern blots using DNA from Rec^+ strains and the quantification of the accumulation of DNA in the wells of these Southern blots, respectively. The variances of the values obtained from this experiment were not homogeneous. This meant that the data could not be analysed by a standard ANOVA. Instead, it had to be analysed using non-parametric tests (Table 4.3). A Kruskal-Wallis test, comparing values obtained for the *lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, showed that the accumulation of DNA in the well of the sample that incurred DSBs was not significant for the *lacZ*.proximal + 1 fragment (P = 0.066). Nevertheless, the data do show that intermediates are accumulated in this region of the chromosome, indicating that the statistical test carried out is not appropriate for this analysis. In contrast, the accumulation of DNA in the well of the *lacZ*.distal + 1 fragment following DSB induction was shown to be significanly different from control samples (P = 0.039). A non parametric Mann-Whitney test also showed that there was no significant difference between the accumulation of DNA in the well of the *lacZ*.proximal + 1 fragment compared to the *lacZ*.distal + 1 fragment when DSBs were induced (P = 0.809).

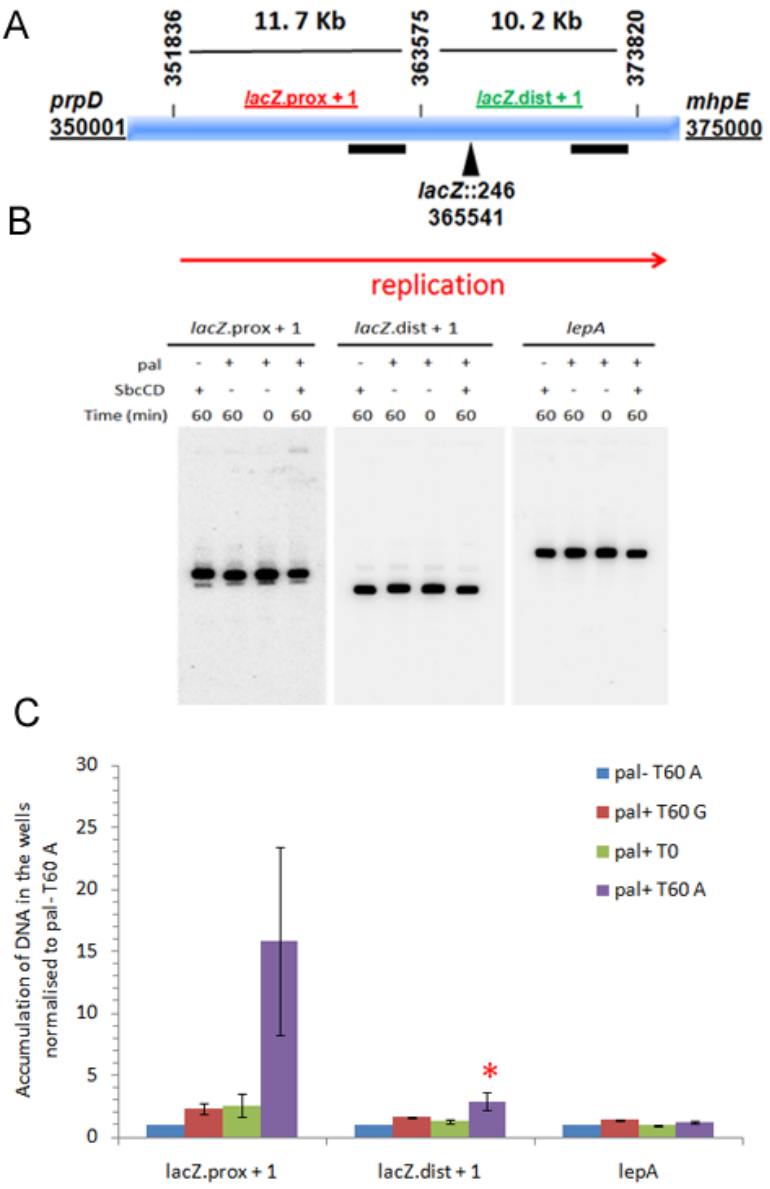

Figure 4. 11 Pulsed-field gel electrophoresis and Southern blot analysis of SacI digested chromosomal DNA of Rec$^+$ strains

(A) SacI digestion map of the *E. coli* chromosome from coordinates 350,001 (*prpD*) to 375,000 (*mhpE*). SacI restriction sites are marked by a grey vertical line and their coordinates are indicated. Probes used to detect the fragments (*lacZ*.prox + 1 and *lacZ*.dist + 1) are indicated above the respective fragment and their binding site is marked by a black rectangle. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4201 (Rec$^+$ *lacZ*$^+$). (C) Quantification of DNA present in the wells of the Southern blots. Red * indicates data that is significantly different from controls. Error bars are standard error of the mean where n = 3.

In the Δ*ruvAB* mutant, a higher accumulation of DNA in the wells of samples incurring DSBs was detected in both fragments analysed (Fig. 4.12). A one-way ANOVA, comparing values obtained for the Δ*ruvAB lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, showed that the accumulation of DNA in the well of the sample that incurred DSBs was not significant for the *lacZ*.proximal + 1 fragment (P = 0.079). To carry out the same comparison for the *lacZ*.distal + 1 fragment, a Kruskal-Wallis non-parabinosemetric Test was carried out. The result of this test showed no significan difference (P = 0.079; Table 4.3). Nonetheless, the quantification clearly shows that a similar amount of recombination intermediates accumulate in both the *lacZ*.proximal + 1 and *lacZ*.distal + 1 fragments when DSBs are induced in the Δ*ruvAB* mutant.

In the Δ*recG* mutant, induction of DSBs also caused an accumulation of DNA in the wells of both fragments analysed (Fig. 4.13). A Kruskal-Wallis non-parametric Test, comparing values obtained for the Δ*recG lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, showed that the accumulation of DNA in the well of the sample that incurred DSBs was not significant for the *lacZ*.proximal + 1 fragment (P = 0.061). To carry out the same comparison for the *lacZ*.distal + 1 fragment, a one-way ANOVA was carried out. The result of this test showed that the accumulation of the DNA in the well of the sample incurring DSBs was highly significant (P < 0.01; Table 4.3). A two-sample T-test, showed there was no significant difference between the amount of DNA accumulated in the well of the *lacZ*.proximal + 1 fragment when compared to the *lacZ*.distal + 1 fragment in both the Δ*ruvAB* and the Δ*recG* mutants (P = 0.417, P = 0.474, respectively). From the quantification presented here, it is possible to

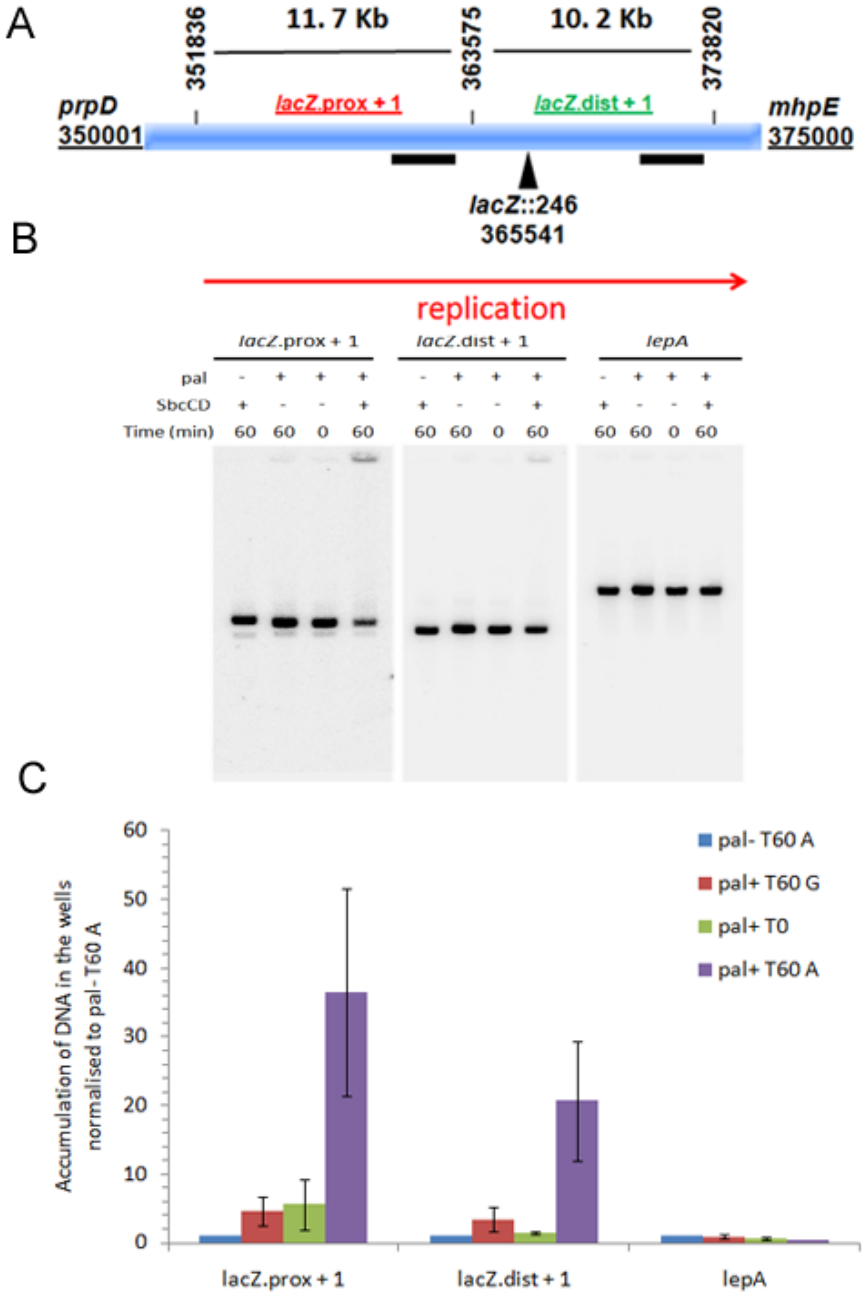

Figure 4. 12 Pulsed-field gel electrophoresis and Southern blot analysis of SacI digested chromosomal DNA of Δ*ruvAB* strains

(A) SacI digestion map of the *E. coli* chromosome from coordinates 350,001 (*prpD*) to 375,000 (*mhpE*). SacI restriction sites are marked by a grey vertical line and their coordinates are indicated. Probes used to detect the fragments (*lacZ*.prox + 1 and *lacZ*.dist + 1) are indicated above the respective fragment and their binding site is marked by a black rectangle. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4243 (Δ*ruvAB lacZ*::246), DL4257 (Δ*ruvAB lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Error bars are standard error of the mean where n = 3.

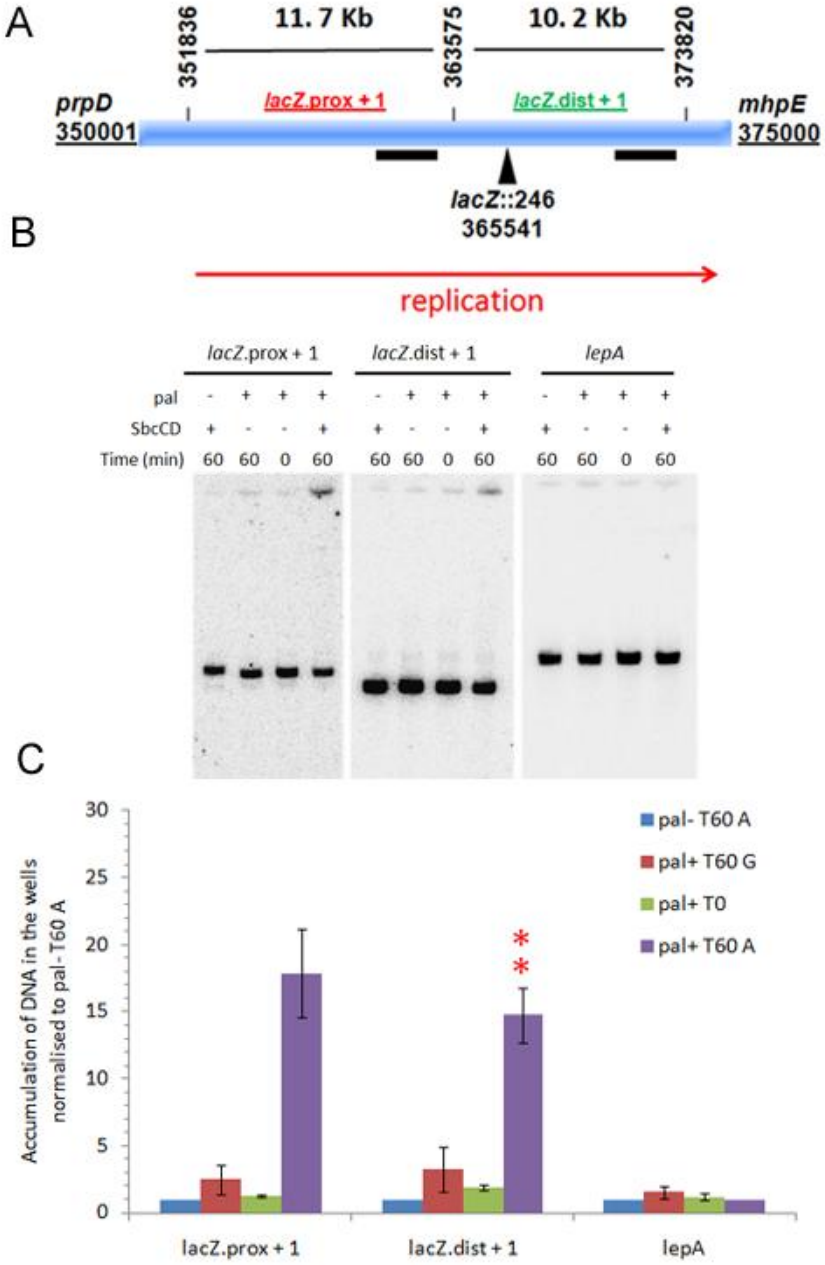

Figure 4. 13 Pulsed-field gel electrophoresis and Southern blot analysis of SacI digested chromosomal DNA of Δ*recG* strains

(A) SacI digestion map of the *E. coli* chromosome from coordinates 350,001 (*prpD*) to 375,000 (*mhpE*). SacI restriction sites are marked by a grey vertical line and their coordinates are indicated. Probes used to detect the fragments (*lacZ*.prox + 1 and *lacZ*.dist + 1) are indicated above the respective fragment and their binding site is marked by a black rectangle. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4311 (Δ*recG lacZ*::246), DL4312 (Δ*recG lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Red ** indicates a highly significant difference from controls. Error bars are standard error of the mean where n = 3.

conclude that a similar amount of DSB repair intermediates are accumulated in the wells of both fragments analysed.

Analysis of the DNA isolated from the Δ*ruvAB* Δ*recG* mutant showed a similar accumulation of recombination intermediates in both the fragments analysed (Fig. 4.14). It is interesting to note that the there appears to be less branched DNA accumulated in the Δ*ruvAB* Δ*recG* mutant compared to the individual Δ*ruvAB* and Δ*recG* mutants. This result is in accordance with the results obtained from the analysis of the SalI digestions, which shows that DNA is degraded in the Δ*ruvAB* Δ*recG* mutant, which would prevent recombination intermediates from forming in this background (Fig. 4.10). For easier comparison of the quantifications obtained from the Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG*, results from strains presented in Figures 4.11-4.14 are compiled in Figure 4.15, in which the y-axis of all four bar charts is the same.

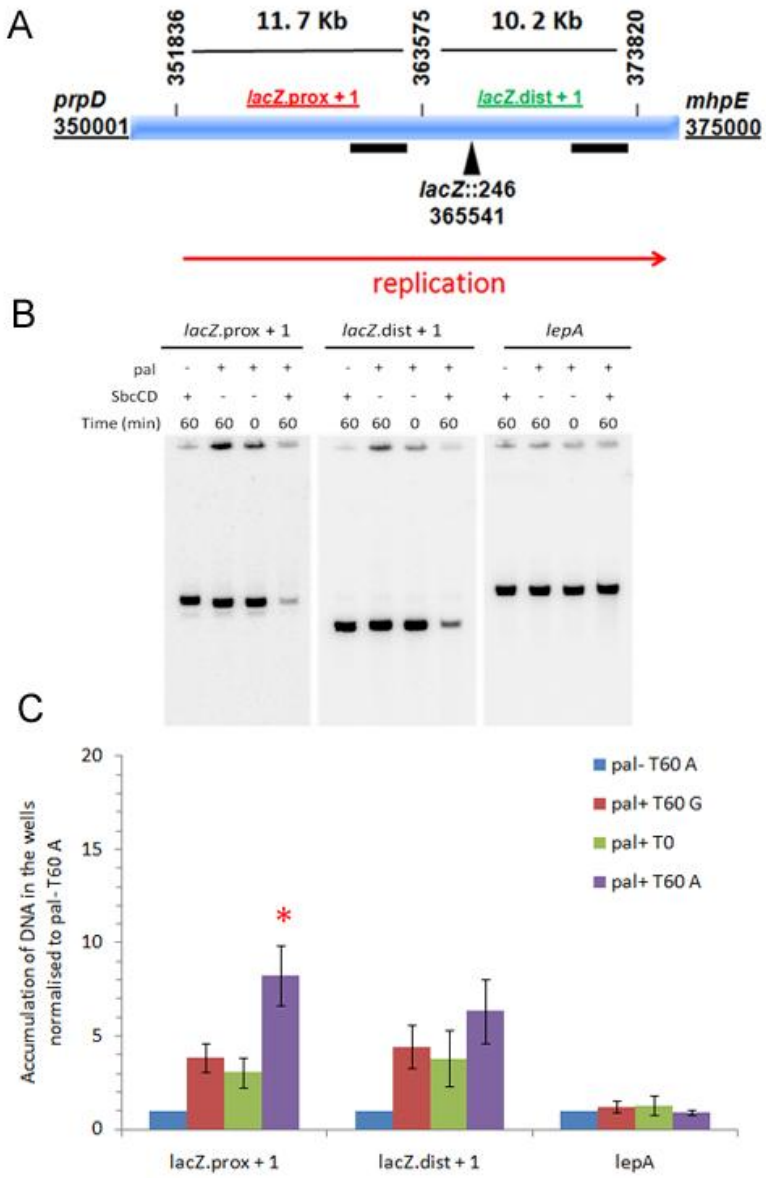

Figure 4. 14 Pulsed-field gel electrophoresis and Southern blot analysis of SacI digested chromosomal DNA of Δ*ruvAB* Δ*recG* strains

(A) SacI digestion map of the *E. coli* chromosome from coordinates 350,001 (*prpD*) to 375,000 (*mhpE*). SacI restriction sites are marked by a grey vertical line and their coordinates are indicated. Probes used to detect the fragments (*lacZ*.prox + 1 and *lacZ*.dist + 1) are indicated above the respective fragment. (B) Southern blots detected with ^{32}P-labelled probes. Strains used were DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), DL4313 (Δ*ruvAB* Δ*recG lacZ*$^{+}$). (C) Quantification of DNA present in the wells of the Southern blots. Red * indicates data that is significantly different from controls. Error bars are standard error of the mean where n = 3.

4.5.2.2 Disappearance of DNA

Once again, digestion of the chromosome of the Δ*ruvAB* Δ*recG* mutant revealed the disappearance of the DNA surrounding the palindrome upon induction of DSBs (Fig. 4.14 B). Figure 4.16 shows the quantification of the total DNA present in each mutants, as calculated for Figure 4.10. All values were normalised to values obtained for the *lepA* fragment. As shown in Figure 4.10, Figure 4.16 shows that there is no disappearance of the DNA surrounding the breakpoint in the Rec$^+$ strain upon DSB formation. A one-way ANOVA, comparing values obtained for the Rec$^+$ *lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, showed that there was no disappearence of DNA in samples that incurred DSBs for both the *lacZ*.proximal + 1 and the *lacZ*.distal + 1 fragment (P = 0.563, P = 0.148, respectively). The same analysis showed that there was slightly less DNA in the *lacZ*.proximal + 1 fragment from the Δ*ruvAB* mutant upon incurring DSBs, but that this was not the case for the *lacZ*.distal + 1 fragment (P = 0.012, P = 0.133, respectively). Nevertheless, the loss of DNA from the *lacZ*.proximal + 1 fragment is marginal. Following the same analysis

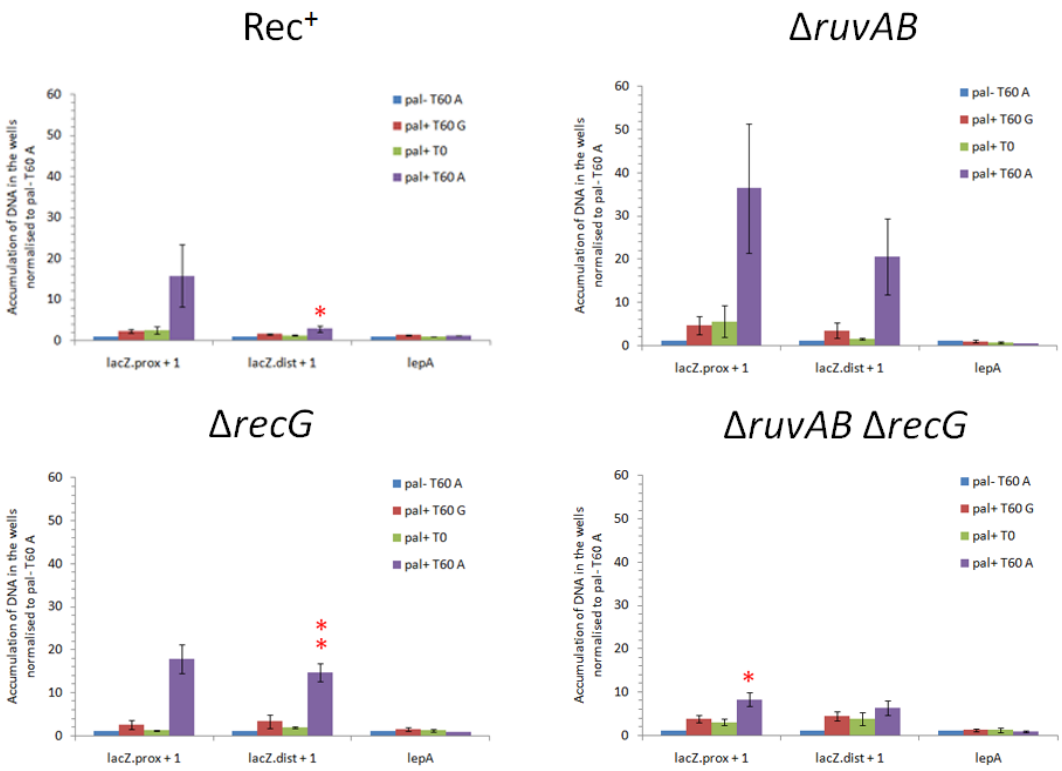

Figure 4. 15 Accumulation of branched DNA in Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains following digestion of the chromosomes with SacI

Bar charts shown in Figures 4.11-4.14 are displayed with the same y-axis values, so as to facilitate the comparison of the different strains. Red * indicates data that is significantly different from controls, while red ** indicates data that is highly significant from controls. Error bars represent standard error of the mean where n = 3.

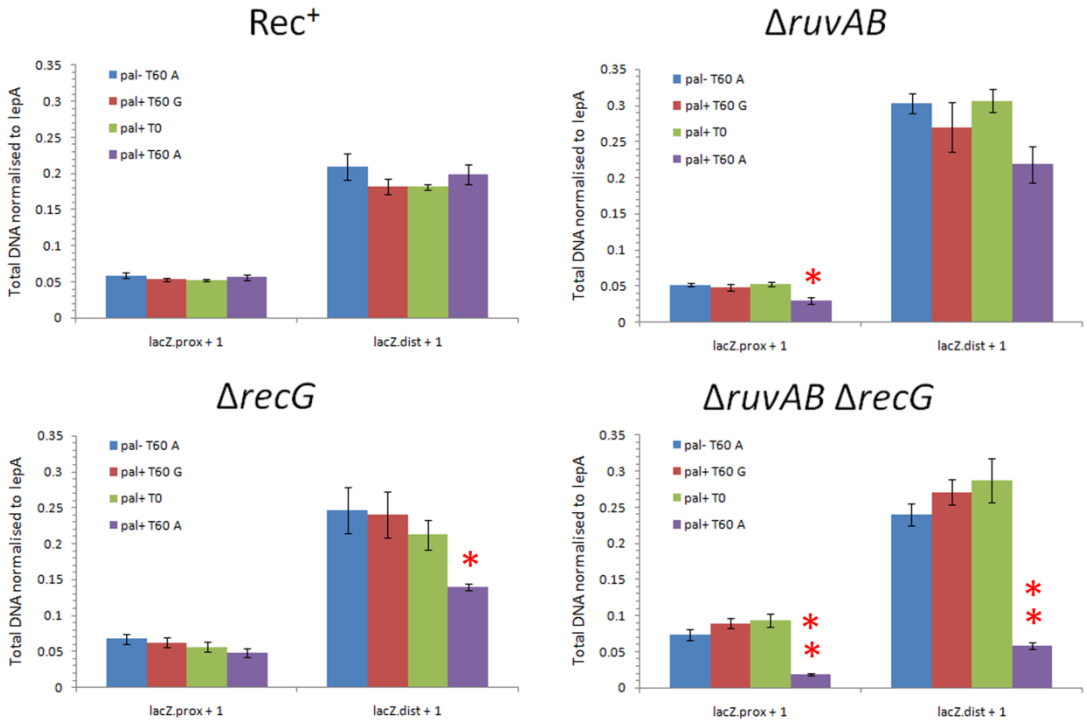

Figure 4. 16 Total amount of DNA in Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains following digestion of the chromosomes with SacI

The total amount of DNA from each sample was quantified and normalised to the total DNA for the *lepA* fragment, so as to detect DNA degradation of the region surrounding the breakpoint upon DSB induction. Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4201 (Rec$^+$ *lacZ*$^+$), DL4243 (Δ*ruvAB lacZ*::246), DL4257 (Δ*ruvAB lacZ*$^+$), DL4311 (Δ*recG lacZ*::246), DL4312 (Δ*recG lacZ*$^+$), DL4260 (Δ*ruvAB* Δ*recG lacZ*::246), and DL4313 (Δ*ruvAB* Δ*recG lacZ*$^+$). Red * indicates data that is significantly different from controls, while a red ** indicates the data are highly significant. Error bars represent standard error of the mean where n = 3.

Table 4. 3 P-values for the accumulation of branched DNA in *lacZ*::246 strains at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose following digestion of the chromosome with SacI

P-values that are not underlined represent results of a one-way ANOVA, while P-values that are underlined represent the result of a non-parabinosemetric Kruskal-Wallis test. P-values marked by *, represent a significant difference and those marked by **, represent a highly significant difference. P-values in black are not significant.

	P-value for *lacZ*.proximal + 1	P-value for *lacZ*.distal + 1	P-value for *lepA*
Rec$^+$	0.066	0.039*	0.051
Δ*ruvAB*	0.079	0.079	0.325
Δ*recG*	0.061	0.002**	0.504
Δ*ruvAB* Δ*recG*	0.036*	0.493	0.770

Table 4. 4 P-values for the degradation of DNA in *lacZ*::246 strains at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose following digestion of the chromosome with SacI

P-values that are not underlined represent results of a one-way ANOVA, while P-values that are underlined represent the result of a non-parabinosemetric Kruskal-Wallis test. P-values marked by a red *, represent a significant difference and those marked by two red *, represent a highly significant difference. P-values in black are not significant.

	P-value for *lacZ*.proximal + 1	P-value for *lacZ*.distal + 1
Rec$^+$	0.563	0.402
Δ*ruvAB*	0.012*	0.133
Δ*recG*	0.381	0.042*
Δ*ruvAB* Δ*recG*	0.000**	0.000**

in the Δ*recG* mutant, the *lacZ*.proximal + 1 fragment did not disappear but there was shown to be less DNA in the *lacZ*.distal + 1 fragment (P = 0.381, P = 0.042, respectively). In contrast, the Δ*ruvAB* Δ*recG* mutant incurred dramatic DNA degradation when DSBs were induced. A one-way ANOVA, comparing values obtained for the Δ*ruvAB* Δ*recG lacZ*::246 strain at time point 0 minute, 60 minutes in 0.5 % glucose, and 60 minutes in 0.2 % arabinose, showed that the loss of DNA was highly significant for both the *lacZ*.proximal + 1 and the *lacZ*.distal + 1 fragments (P < 0.01 for both fragments). This results correlates with the data shown in Figure 4.14, where it was shown that very little branched DNA was accumulated in the Δ*ruvAB* Δ*recG* mutant.

4.6 Discussion

4.6.1 DSB repair intermediates generated from a site-specific DSB are accumulated locally to the breakpoint

In this chapter, pulsed-field gel electrophoreses were used to analyse large chromosomal fragments for the presence of DSB repair intermediates generated from SbcCD-mediated cleavage of a 246 bp interrupted palindrome in the *lacZ* gene of the *E. coli* chromosome. Digestion of the chromosome with NotI released 23 fragments of different sizes, which were visualised by staining the DNA with EtBr (Fig. 4.2 A and B). It was previously shown that NotI digests of DNA from cells unable to resolve recombination intermediates, which incurred multiple, randomly

distributed DSBs, failed to generate a NotI banding pattern on a pulsed-field gel (Wardrope et al., 2009). In the work presented here, in which the DSB was site-specific, a strong banding pattern was obtained in Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* mutants, suggesting that the majority of the DSB repair intermediates accumulate locally to the breakpoint (Fig. 4.2 B). This finding was supported by Southern blotting of the DNA surrounding the breakpoint after a NotI digestion of DNA from the Δ*ruvAB* Δ*recG* mutant, which showed that the induction of DSBs caused the disappearance of the linear DNA in the region surrounding the breakpoint only (Fig. 4.3 B and C lanes 10-12). This trend was also seen following digestion of the chromosome with I-SceI, which released a smaller fragment, of 174 Kb, containing the breakpoint (Fig. 4.4 B and C lanes 12-14). All this work was carried out using DIG-labelled probes. In order to increase the sensitivity of the assay and be able to quantify the data obtained, the Southern blot protocol was changed from using DIG-labelled probes to using ^{32}P-labelled probes. Thereafter, Rec^{+}, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* mutants were analysed following digestion of the chromosomes with either SalI, which released three major fragments containing and surrounding the palindrome, or SacI, which released two smaller fragments containing and surrounding the breakpoint (Figs. 4.5 A and 4.11 A, respectively).

4.6.2 Transient DSB repair intermediates can be detected upstream of the breakpoint in a recombination proficient strain

There are various models for the repair of DSBs that have been generated by the analysis of different model organisms (Paques and Haber, 1999). In the canonical double strand break repair (DSBR) pathway, intermediates of repair are predicted to accumulate equally on both sides of the DSB (Szostak et al., 1983). In other proposed repair pathways, such as break induced replication (BIR), intermediates accumulate preferentially upstream of the breakpoint and the DNA downstream of the breakpoint is degraded. This is followed by new replication of large tracts of the chromosome (Voelkel-Meiman and Roeder, 1990). Another model, synthesis dependent strand annealing (SDSA), does not predict that the resolution of HJs requires the action of a HJ resolvase. Instead, dissolution of the HJs takes place (Chapter 1 Section 1.4) (Kreuzer et al., 1995; Nassif et al., 1994). In this work, the detection of intermediates of repair in a Rec^+ background can help to distinguish between these pathways. Following Southern blotting of SalI digested DNA from a Rec^+ strain, the SalI fragment containing the breakpoint and incurring DSBs and the SalI fragment upstream of the breakpoint and incurring DSBs, accumulated a significant amount of branched DNA (the 34.1 Kb *yagV* fragment and the 23.7 Kb *lacZ* fragment; Fig. 4.5 C). As the palindrome is not centrally located in the *lacZ* SalI fragment, the majority of the DNA analysed was located upstream of the breakpoint. These results may indicate that in a recombination proficient background a repair pathway similar to BIR takes place. This hypothesis is

plausible as the DSB generated at the palindrome is replication dependent, which means that the DNA free end downstream of the DSB may not be far from a replication fork. This arm could be degraded to that replication fork to leave the DNA free end upstream of the DSB to invade and repair the break.

Alternatively, the DNA free end downstream of the breakpoint may not be degraded all the way back to a replication fork but simply degraded beyond the *araJ* fragment, meaning that these branched DNA species are not detected in the assay carried out here. This hypothesis is based on the possibility that the χ array that is located 1.5 Kb downstream of the DSB falls within an Okazaki fragment and is therefore in ssDNA. This would mean that the sequence is not recognised by RecBCD, leaving the DNA free end downstream of the breakpoint to be degraded to the next χ sequences that are available. This is likely as the hairpin formation by the palindrome is thought to occur when the palindrome falls in an Okazaki fragment. As a result, it is also possible that the DNA 1.5 Kb downstream of the palindrome also falls within the same Okazaki fragment.

In order to determine whether there really was a bias in accumulating DSB repair intermediates preferentially on one side of the breakpoint, analysis of restriction fragments that were more equally distributed around the palindrome was required. For this purpose, a SacI digest was carried out (Fig. 4.11 A). Branched DNA was detected in both the *lacZ*.proximal + 1 and the *lacZ*.distal + 1 fragments. Nevertheless a portion of the *lacZ*.distal + 1 fragment (2.1 Kb) is located upstream of the breakpoint and it is therefore difficult to conclude whether the branched DNA species detected are located upstream or downstream of the breakpoint. An experiment in which the rare-cutting I-SceI endonuclease restriction

sites are inserted immediately upstream and downstream of the palindrome to generate two equally sized chromosomal fragments that do not contain the breakpoint would need to be carried out in order to be able to conclude whether intermediates of repair are indeed accumulated preferentially upstream of the DSB or not.

4.6.3 During DSB repair, a recombination proficient strain maintains a balanced DNA content

A repair process such as BIR predicts extensive degradation of the DNA free end downstream of the DSB, as this DNA gets ressected to a pre-existing replication fork (Paques and Haber, 1999). The observation that DNA was being lost in the Δ*ruvAB* Δ*recG* mutant upon DBS induction (Fig. 4.8 B) prompted us to quantify the abundance of DNA surrounding the breakpoint in order to detect whether any DNA degradation was taking place. Following digestion of the DNA from the Rec$^+$ strain with SalI, a very small but significant decrease in the DNA surrounding the breakpoint was detected (Figs. 4.10; Table 4.2). This was not the case following digestion of the chromosome with SacI (Fig. 4.16; Table 4.4). Nevertheless, it is important to emphasize that a portion of the *lacZ*.distal + 1 fragment is located upstream of the DSB and this could generate inaccurate data (Fig. 4.11 A).

BIR in a Rec$^+$ background predicts the establishment of a new replication fork that is generated by RecA-mediated strand invasion of the unbroken chromosome with the DNA free end upstream of the breakpoint. Therefore, new replication generated by BIR in a Rec$^+$ strain could easily mask the degradation of

the downstream DNA free end and might account for the very minor loss of DNA detected in the *araJ* fragment following digestion of the chromosome with SalI (Figs. 4.10; Table 4.2). It would be interesting to carry out this experiment in Δ*priA* or Δ*priB* mutants. In these mutants, the replicative helicase, DnaB, which is required to establish replication from a D-loop, is not loaded onto the replication fork (Sandler and Marians, 2000). Therefore, joint molecules are formed but repair synthesis should not be established. As a result, it may be possible to detect degradation of the DNA free end downstream of the breakpoint if it does occur. In this context, the absence of DNA degradation would indicate that both DNA free ends either invade the unbroken template or are somehow protected from degradation. If lack of degradation is then coupled with the accumulation of branched DNA on both sides of the DSB, this would be indicative of canonical DSBR taking place. If no DNA degradation is detected and branched DNA is accumulated only upstream of the breakpoint then this would be indicative of SDSA taking place. This hypothesis would also predict there to be no requirement for the RuvABC HJ resolvase complex. Indeed, viability tests presented in Chapter 3 suggest that RuvAB may not be required for the repair of acute DSBs (Fig. 3.4 A).

Nevertheless, it is important to mention that Eykelenboom and collaborators did detect recombination on both sides of the breakpoint following the formation of an SbcCD-mediated DSB at the palindrome, which suggests that both DNA free ends are actively involved in repair, indicating that a pathway like the canonical DSBR pathway may be taking place (Eykelenboom et al., 2008). In order to generate more conclusive data, this experiment could be designed so that both fragments

analysed are of equal size and do not contain the palindrome by the insertion of I-SceI restriction sites into the chromosome.

4.6.4 Branched DNA accumulates in the absence of RuvAB within 60 minutes of DSB induction, suggesting that repair of the DSB generates HJs, which need to be processed in order to complete the repair process

In the Δ*ruvAB* mutant, following chromosomal digestion with SalI, a very high proportion, compared to the Rec$^+$ strain, of branched DNA species were detected as a result of DSB formation (Figs. 4.6 panels B and C and 4.9). The detection of branched DNA in a Δ*ruvAB* mutant is strongly indicative of HJs arising during the repair of the DSB. As the Δ*ruvAB* mutant did not lose viability when exposed to DSBs for 60 minutes (Chapter 3; Fig. 3.4 A), this observation strengthens the hypothesis that there may be an alternative mechanism for dealing with persistent HJs that does not rely on their resolution by RuvAB and, presumably, RuvC. Additionally, the repair of the SbcCD-mediated DSBs was shown to be partially reliant on the XerCD site-specific recombination system (Eykelenboom et al., 2008). This system is required to resolve chromosome dimers that arise from crossovers generated by the resolution of two or more HJs (Eykelenboom et al., 2008). This characteristic suggests that in the presence of RuvABC, HJs are resolved and not dissolved, as expected. Testing whether a requirement for the XerCD site-specific recombination system remains in a Δ*ruvAB* mutant might indicate whether

HJs are being resolved or dissolved in the absence of RuvAB. If the requirement for XerCD remains it will suggest that the RuvAB-independent HJ processing pathway also occurs through the resolution (cleavage) of HJs. Alternatively, if this requirement is lost, it will strongly indicate that HJs are dissolved instead of resolved.

Finally, the majority of the branched DNA detected in the Δ*ruvAB* mutant was located in the 23.7 Kb *lacZ* fragment (Fig. 4.6 C). This distribution is not surprising as the strains used contain χ arrays 1.5 Kb either side of the breakpoint. As a result, strand invasion should be initiated around 1.5 Kb from the DSB and all HJs arising from this strand invasion may not branch migrate in the absence of the branch migration complex RuvAB, causing their accumulation close to the breakpoint. In order to support this hypothesis, it would be interesting to carry out the same experiment in strains lacking the χ arrays or in strains retaining RuvAB but lacking RuvC. In either of these situations HJs would be predicted to be located further away from the breakpoint and this should be reflected in the distribution of the branched DNA detected in these mutant backgrounds. If this were to be the case, this would indicate that RuvAB is the major protein involved in the branch migration of HJs.

4.6.5 During DSB repair, a Δ*ruvAB* mutant maintains a balanced DNA content

As was seen in the Rec$^+$ strain following digestion of the chromosome with SalI, a small but significant reduction in the DNA content was detected after induction of DSBs in the downstream *araJ* fragment (Figs. 4.10; Table 4.2). A similar decrease was noted in the *lacZ*.distal + 1 fragment following digestion with SacI (Fig. 4.16). Nevertheless, this could be due to a portion of the lacZ.distal + 1 fragment being partially located upstream of the breakpoint. These results suggests that, despite accumulating a large amount of branched DNA, the joint molecules formed in the Δ*ruvAB* mutant are stable and no DNA free ends persist in the cell. It would be interesting to carry out an experiment similar to the SacI experiment but where I-SceI restriction sites are introduced to generate two small fragments either side of the palindrome that do not contain the breakpoint. This experiment might uncover degradation downstream of the DSB in the Δ*ruvAB* mutant. If this were to occur, it would suggest that repair replication cannot be established efficiently unless HJs are first resolved as these would pose a topological barrier to a travelling replication fork.

4.6.6 Branched DNA accumulated in the absence of RecG and RuvAB RecG is preferentially located upstream of the breakpoint and is distributed across a wide area of the chromosome

Analysis of the branched DNA accumulated in the Δ*recG* and the Δ*ruvAB* Δ*recG* mutants provided more information about the dynamics of DSB repair

initiated at the palindrome and the possible role of the RecG helicase in this process. In the SalI digest of the Δ*recG* mutant incurring DSBs, branched DNA was detected in all three restriction fragments surrounding the breakpoint (Fig. 4.7 panels B and C). A higher amount of branched DNA was located in the *yagV* fragment of the Δ*recG* mutant when compared to the *yagV* fragments of the Rec$^+$ and Δ*ruvAB* strains. This result could be explained by various hypotheses. One explanation is that the Δ*recG* mutant accumulates replication forks, as was recently suggested by Rudolph and collaborators (Rudolph et al., 2009a; Rudolph et al., 2009b). This hypothesis predicts that during repair of damage induced by UV light, two converging replication forks are established following the canonical DSBR pathway. When these forks collide in the absence of RecG, PriA re-primes DNA synthesis. These data, along with data showing that RecG can act on a variety of branched DNA substrates, has lead Rudolph and collaborators to suggest that RecG plays a general role in chromosome maintenance and that over-replication of the chromosome takes place in the absence of RecG and this over-replication is exacerbated when DNA damage is induced (Rudolph et al., 2010). This hypothesis would be confirmed by the detection of y-arcs when DNA from a Δ*recG* mutant incurring DSBs would be analysed by native two-dimensional agarose gel electrophoresis. To fully confirm this hypothesis the directionality of the replication forks would need to be understood as in a situation of over-replication generated at the breakpoint these replication forks would be moving from the DSB towards the origin and terminus.

RecG has also been implicated in resolving HJs through branch migration, presumably by migrating two junctions into each other in a mechanism analogous

to the eukaryotic HJ dissolution pathway mediated by the Sgs1-Top3-Rmi1 complex (Ashton et al., 2011; Lloyd and Sharples, 1993a). This mechanism would predict that HJs would accumulate in a Δ*recG* mutant. Their broad distribution could be a result of branch migration mediated by RuvAB, which is still present in this background. This hypothesis would predict that the distribution of branched DNA in a Δ*recG* mutant is dependent on RuvAB. However, this was not the case. The distribution of branched DNA in the Δ*ruvAB* Δ*recG* mutant was not very different to the distribution of branched DNA detected in the Δ*recG* mutant, suggesting that the branched DNA species accumulating as a result of the absence of RecG are not a substrate for the action of RuvAB and therefore are likely not to be HJs (Fig. 4.9). Furthermore, if HJs were being accumulated in a Δ*recG* mutant, the branched DNA species detected in the *lacZ* fragment of the Δ*ruvAB* Δ*recG* mutant ought to be the sum of the branched DNA species detected in the individual Δ*ruvAB* and Δ*recG* mutants. This, too, was not the case. This result contradicts both the hypotheses that HJs accumulate in the Δ*recG* mutant or that over-replication is established in the Δ*recG* mutant as a result of DSB induction. The profile of the branched DNA detected in the Δ*ruvAB* Δ*recG* mutant suggests that RecG may be acting upstream of RuvAB, as losing the activity of the RuvAB complex in a context in which the activity of the RecG helicase is no longer present does not result in the dramatic accumulation of branched DNA that is seen in the individual Δ*ruvAB* mutant.

The hypothesis that RecG may act upstream of RuvAB has been previously suggested based on *in-vitro* experiments showing that RecG can act very efficiently on three-way DNA junctions (McGlynn and Lloyd, 1999; Whitby and Lloyd, 1995). D-loops, which are formed before a mature joint molecule, which contains HJs, is

generated, is a three-way DNA junction. The role of RecG may be in stabilising these D-loops and allowing for the formation of HJs from them. Indeed, it has been shown *in-vivo* that RecG can act on R-loops, structures very similar to D-loops (Hong et al., 1995). This hypothesis could also explain why branched DNA accumulates preferentially upstream of the breakpoint in the Δ*recG* and, potentially, the Δr*uvAB* Δ*recG* mutants (Figs. 4.7 C and 4.8 C). This could be attributed to the uneven distribution of the χ recombinational hotspot (5'-GCTGGTGG-3') in the chromosome, as this sequence is 50 % more abundant on the leading strand of both replichores (Blattner et al., 1997). This means that in the context of SbcCD-mediated palindrome cleavage, the DNA free end upstream of the breakpoint would contain more correctly oriented χ sequences than the DNA free end downstream of the breakpoint. In mutants where D-loops are unstable, the DNA free end upstream of the breakpoint would attempt to repeatedly strand invade the unbroken chromosome due to the higher abundance of χ. The DNA free end downstream of the breakpoint would be less subject to this event due to the lower abundance of χ and would therefore be subject to higher degradation. As a result, it would be expected that more branched DNA (presumably D-loops) would be detected upstream of the breakpoint. This bias was not seen when SacI digestions of the DNA isolated from the Δ*recG* mutant were analysed (Fig. 4.13 C). Nevertheless, it was still detected in the SacI digestions of the DNA from the Δ*ruvAB* Δ*recG* mutant (Fig. 4.14 C). As mentioned above, results from this experiment should be taken with caution as the *lacZ*.distal + 1 fragment is not a true representation of a DNA fragment downstream of the breakpoint. Nevertheless, if it is found that the Δ*recG* mutant does accumulates branched DNA on both sides of the breakpoint while a

bias for accumulating upstream in a Δ*ruvAB* Δ*recG* mutant is still detected, this might be due to the ability of RuvAB to partially complement for the absence of RecG. If RuvAB could bind unproductively to the D-loop and prevent it from dissociating to generate a new DNA free end, this may generate branched DNA on both sides of the DSB. Indeed the ability of RuvAB to partially complement for the absence of RecG was described in Chapter 3 (Fig. 3.4 A), and previously published literature (Lloyd, 1991).

4.6.7 In the absence of RecG, a small but significant amount of the DNA downstream of the breakpoint is degraded upon DSB formation. This DNA degradation is exacerbated by the additional loss of RuvAB

Following digestion of the chromosome with both SalI and SacI, the Δ*recG* mutant incurred a small loss of DNA downstream of the breakpoint (Figs. 4.10 and 4.16). This result suggests that some DNA free ends may persist in this mutant. This could happen if D-loops were not converted into mature joint molecules. Additionally, the degradation is exacerbated in a Δ*ruvAB* Δ*recG* mutant suggesting that under certain circumstances, endogenous RuvAB can stabilise the D-loops formed, presumably by binding unproductively to them (Figs. 4.10 and 4.16). These results are in accordance with the hypothesis that RecG is required to stabilise D-loops and that RuvAB can bind to D-loops in an unproductive manner, which prevents the D-loops dissociating but does not allow for the completion of repair. It would be interesting to see whether X-spikes are absent from native two-

dimensional agarose gel electrophoreses using DNA from the Δ*recG* and the Δ*ruvAB* Δ*recG* mutants and whether these strains only accumulate Y-arcs. This could be indicative of excessive D-loop formation. Additionally, it would be compelling to determine whether the dismantling of the D-loops is an active or a passive process. UvrD, a protein involved in a variety of DNA repair pathways, has been implicated in stripping RecA polymers from DNA substrates both *in-vitro* and *in-vivo* and may be responsible for the dismantling of persistent and unproductive D-loops generated in the absence of RecG (Centore and Sandler, 2007; Veaute et al., 2005). This hypothesis would predict that a Δ*uvrD* mutation might reduce the loss of viability seen in the Δ*recG* Δ*ruvAB* mutation back to the level of a single Δ*recG* mutation. Another helicase that may carry out this process is PriA (Tanaka et al., 2007). It would be worth investigating whether the *priA300* helicase mutant can rescue the Δ*recG* death phenotype and prevent the degradation of DNA in the 2Δ*ruvAB* Δ*recG* mutant in the SbcCD/palindrome system (Chapter 1; section 1.3.2.2).

4.6.8 Conclusions

The study of the distribution of branched DNA that accumulates in a cell after the induction of a site-specific DSB can shed light on the mechanism of the repair process. Genetic analysis of the Δ*ruvAB* mutant that suggested that the RuvAB protein complex was not required for surviving DSB induction for up to 120 minutes and described in Chapter 3, might mean that the repair of the DSB was occurring by SDSA, a repair pathway in which HJs are dissolved rather than resolved and therefore the activity of a HJ resolvase like RuvABC is not required.

To fully conclude this it is mandatory to identify the proteins required for this RuvAB-independent HJ processing pathway. Meddows and collaborators reached a similar conclusion when describing the repair of a site-specific DSB generated by I-SceI cleavage of the *E. coli* chromosome and suggested that RecG was involved in the dissolution of HJs in the absence of RuvABC (Meddows et al., 2004). In the work presented in this chapter, pulsed-field gel electrophoresis suggested that RecG was not responsible for the dissolution of HJs in the absence of RuvAB. This result illustrates the importance of being able to study both the physical nature of the DNA as well as the viability of different mutant strains.

Conclusions drawn from the analyses of branched DNA by pulsed-field gel electrophoreses alone cannot be conclusive as the detailed three dimensional structures of the DNA cannot be determined using this technique. In order to gain a more complete picture of the repair process, native two-dimensional agarose gel electrophoresis will be used.

Chapter 5

Native two dimensional agarose gel electrophoresis of recombination intermediates generated during DNA double strand break repair in *E. coli*

5.1 Introduction

In Chapter 4, recombination intermediates formed during the repair of an SbcCD-mediated DSB in Rec$^+$, Δ*ruvAB*, Δ*recG*, and Δ*ruvAB* Δ*recG* strains were analysed by pulsed-field gel electrophoresis. It was shown that, in all strains, DSB repair intermediates accumulated in close proximity of the breakpoint, leaving the majority of the chromosome unaffected. In the Rec$^+$ and Δ*ruvAB* strains, the majority of the intermediates were located within 23 Kb surrounding the DSB, while in the Δ*recG* and Δ*ruvAB* Δ*recG* mutants they were located across a larger area of the chromosome and could be detected in a region as big as 59 Kb surrounding the DSB. In addition, the loss of RecG appeared to cause a small amount of degradation of the DNA surrounding the breakpoint when DSBs were induced. This was grossly exacerbated in the absence of both RecG and RuvAB.

Analysing the distribution of intermediates of repair alone does not generate enough information to fully understand the repair process. In order to shed further light onto the mechanism of the repair, the three dimensional structure of these recombination intermediates needs to be understood. For this purpose, it is necessary for them to migrate out of the wells of a gel, yet remain distinguishable from their linear counterpart. This can be achieved by native two dimensional agarose gel electrophoresis (2-D agarose gel). The first part of this chapter describes the optimisation of the Southern blot technique in order to maximise the signal obtained from chromosomal DNA preparations, as this is indispensible for detecting branched DNA by 2-D agarose gel. The second part of this chapter describes results obtained from 2-D agarose gel analysis of DNA extracted from Δ*ruvAB* mutants and further optimisations of the technique using these strains. Finally, 2-D agarose gel analysis of three restriction fragments surrounding the breakpoint was carried out in Rec$^+$, Δ*ruvAB*, and Δ*recG* strains.

5.2 Optimisation of the Southern blotting protocol

5.2.1 Optimising the percentage of agarose in the plugs

The isolation of DNA in agarose plugs is thought to prevent DNA shearing, as the agarose physically protects the DNA. Additionally, branched DNA species are thought to be preserved better when embedded in agarose, as the DNA is not free to move as it would be in a liquid medium (Cromie et al., 2006). This is thought

to limit *in vitro* branch migration of the molecules. Taking these factors into consideration, it was asked whether the percentage of agarose used in the plugs could affect the retention of branched DNA. This question was investigated using pulsed-field gel electrophoresis. SalI digested chromosomal DNA, from a Δ*ruvAB* mutant containing the palindrome (*lacZ*::246) grown in either 0.5 % glucose ($SbcCD^-$) or 0.2 % arabinose ($SbcCD^+$), was prepared in plugs containing different concentrations of agarose (1 %, 0.8 %, 0.6 %, and 0.4 %). The DNA was separated by pulsed-field gel electrophoresis and, following Southern blotting, detected with a probe against the 23.7 Kb *lacZ* fragment containing the breakpoint.

Figure 5.1 A shows the results of the Southern blot and Figure 5.1 B shows the quantification of the radioactive signal emitted from it. As expected DNA was only detected in the wells of samples in which SbcCD was expressed (0.2 % arabinose). Unexpectedly, the amount of DNA retained in the wells increased as the percentage of agarose decreased. In accordance with this finding, an increase in the intensity of the bands of linear DNA from the controls (0.5 % glucose) and the *cysN* fragment was also noticeable as the percentage of agarose decreased. When chromosomal DNA was prepared in agarose plugs, a culture of cells was mixed with liquid agarose and allowed to set. Once set, enzymatic reactions were required to lyse the cells and degrade the proteins, leaving behind the DNA. It was concluded that as the percentage of agarose decreased, the enzymatic reactions required to release the DNA occurred more efficiently, resulting in a greater yield of

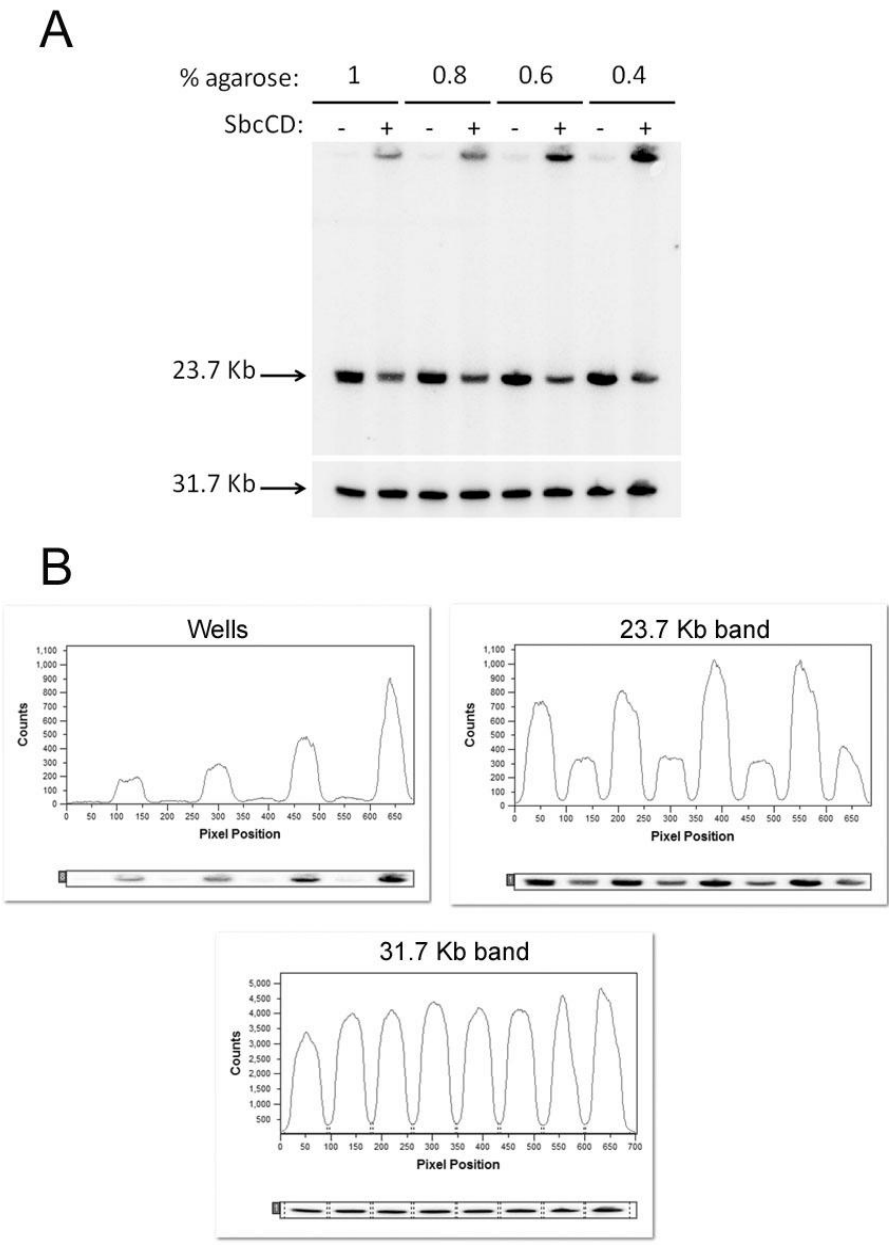

Figure 5. 1 Optimisation of the percentage of agarose used in the plugs

(A) SalI digestions of chromosomal DNA isolated from a Δ*ruvAB* mutant grown either in 0.5 % glucose for 60 minutes (SbcCD$^-$) or 0.2 % arabinose for 60 minutes (SbcCD$^+$) were separated by pulsed-field gel electrophoresis and a Southern blot using a 32P-labelled probe specific for the 23.7Kb *lacZ* fragment was used to detect the DNA. A second probe, *cysN*, was used to detect a 31.7 Kb fragment on the opposite side of the chromosome as a loading control. (B) The DNA in the wells and in the gel was quantified using ImageQuant™ TL. Strain used was DL4243 (Δ*ruv*AB lacZ::246). The experiment was repeated twice and the same trend was found.

genetic material in the plugs. Based on this conclusion, all plugs made hereafter contained 0.4 % agarose.

5.2.2 Optimising the handling of the membrane pre-transfer and the crosslinking of the DNA to the membrane post-transfer

Once DNA has been separated on a gel, it has to be transferred and crosslinked to a membrane in order to ensure that it does not get washed off during the probing stage of the Southern blot. A variety of factors can affect the efficiency of the transfer and crosslinking. Some membranes need to be equilibrated in either water or transfer buffer before the transfer stack is set up. Once the transfer is finished, drying of the membrane as well as varying the intensity of UV light used could increase the efficiency of the crosslinking. These issues were addressed in order to ensure that an optimal protocol was used for detecting recombination intermediates by 2-D agarose gel electrophoresis. NEB 1 Kb ladder was used as a DNA sample because the fragment sizes in this ladder were similar to the sizes of the restriction fragments that were going to be analysed by 2-D gel electrophoresis. The *lacZ* probe was used to detect the DNA, as part of the 1 Kb ladder contains sequence homology to the *E. coli lacZ* gene.

Four samples of the NEB 1 Kb ladder (1 µg each) were separated on a 1 % agarose gel. Two of those DNA samples were transferred to a membrane that had been equilibrated in transfer buffer for 5 minutes. The remaining two samples were transferred onto a dry membrane. Post-transfer, the two membranes were sliced to leave one DNA sample on each slice. For each condition, one slice was allowed to

dry for 1 hour at room temperature prior to crosslinking, while the second was crosslinked while still wet. Crosslinking was achieved by using the autocrosslink function on a Stratalinker, which delivers a UV light dose of 1200 J/m^2 and is generally used for crosslinking DNA to a membrane. Figure 5.2 A shows that, irrespective of whether the membrane was equilibrated pre-transfer, the most DNA was detected on the membranes that were allowed to dry for 1 hour at room temperature before being crosslinked.

In order to address which intensity of UV light resulted in the most efficient crosslinking, six samples of NEB 1 Kb ladder (1 µg each) were separated on a 1 % agarose gel. The DNA was transferred onto a membrane that had been equilibrated in transfer buffer for 5 minutes. Post-transfer, the membrane was allowed to dry at room temperature for 1 hour and then sliced into 6, leaving one DNA sample on each slice. The slices were crosslinked in a Stratalinker using different intensities of UV light (1400, 1200, 1000, 800, 600, and 400 J/m^2). Figure 5.2 B shows that for the type of membrane and size of the DNA fragments used in this experiment, 1200 J/m^2 was not the optimal intensity of UV light. In fact, a dose of UV light as low as 400 J/m^2 fixed more DNA to the membrane than the intensity of UV light that is delivered by the autocrosslink function (1200 J/m^2). Using ImageQuant™ TL from GE Healthcare, the signal emitted from the 3 Kb band was quantified. The background emitted from the membrane was subtracted following quantification, and a UV light dose of 1000 J/m^2 was found to give the highest signal.

For the purpose of analysing intermediates of DSB repair, it was concluded that the optimum agarose concentration for the plugs was 0.4 %. The optimum conditions for handling the membrane pre-transfer and for crosslinking the DNA to

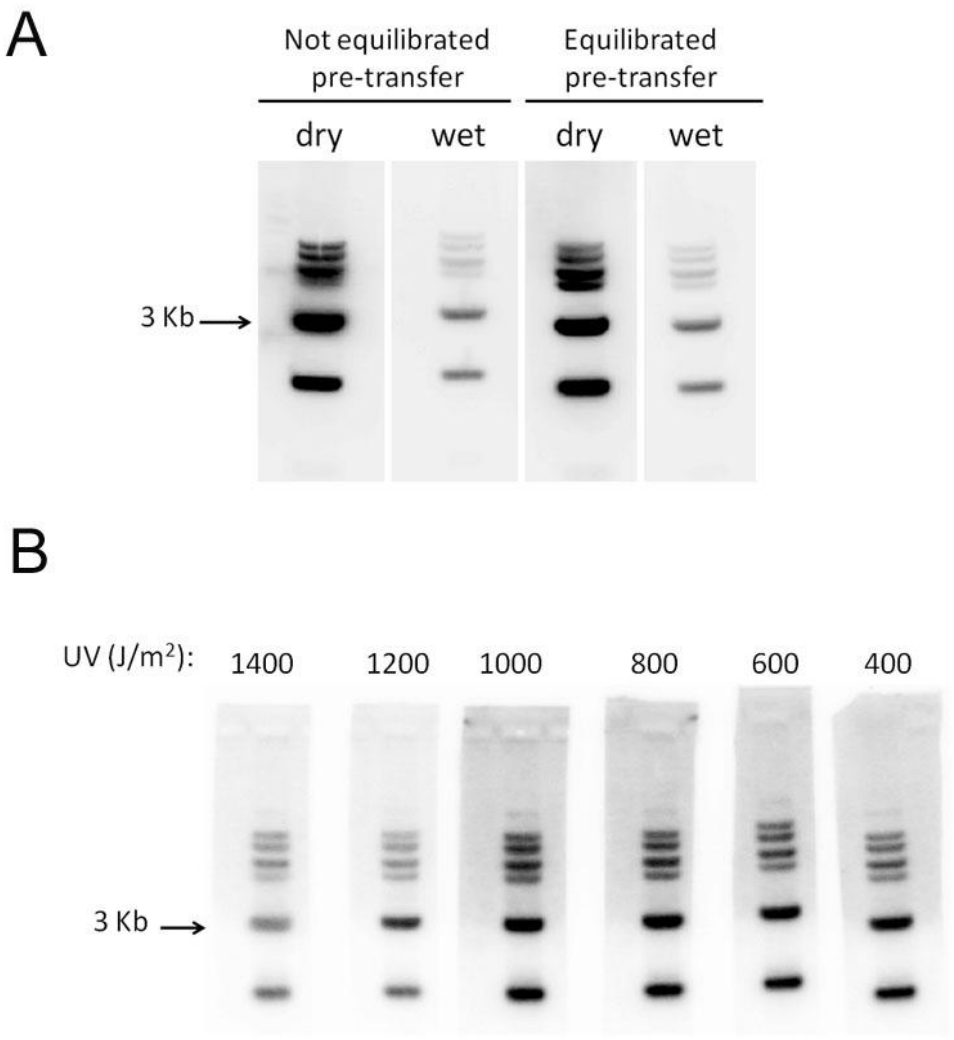

Figure 5. 2 Optimisation of the membrane handling pre-transfer and crosslinking of the DNA post-transfer

(A) 4 samples of NEB 1 Kb ladder (1 µg each) were separated on a 1 % agarose gel. The gel was treated for transfer and 2 membranes, one of which was pre-equilibrated in transfer buffer for 5 minutes, were placed onto the gel, each membrane covering two samples. Post-transfer, the two membranes were sliced in half. One half was taken directly to the Stratalinker and crosslinked using the autocrosslink function (UV light dose of 1200 J/m2). The second half was left to dry at room temperature for 1 hour, after which it was crosslinked in the same manner as the first membrane. (B) 6 samples of NEB 1 KB ladder (1 µg each) were separated on a 1 % agarose gel. The membrane was pre-equilibrated in transfer buffer for 5 minutes before the transfer stack was set up. Post-transfer, the membrane was allowed to dry at room temperature for 1 hour. The dry membrane was sliced into 6 strips, each containing one DNA sample, which were crosslinked using different doses of UV. The experiment was repeated twice and the same trend was found.

the membrane post-transfer were to equilibrate the membrane in transfer buffer for 5 minutes prior to setting up the transfer stack and to allow the membrane to dry at room temperature for 1 hour post-transfer and before being crosslinked in a Stratalinker using a UV light intensity of 1000 J/m^2.

5.3 2-D agarose gel electrophoresis

5.3.1 2-D agarose gels of DNA isolated from Δ*ruvAB* mutants

In order to establish the 2-D agarose gel protocol, Δ*ruvAB* mutants were used, as these strains were found to accumulate the highest amount of branched DNA in close proximity of the breakpoint (Chapter 4; Fig. 4.6). Additionally, the established role of RuvAB in Holliday junction (HJ) processing meant that HJs were the most likely form of branched DNA to be accumulated in a Δ*ruvAB* mutant, making the interpretation of the first 2-D agarose gel easier. Figure 5.3 presents a schematic representation of a 2-D agarose gel highlighting the most common branched species detected and their pattern of migration.

Figure 5.4 A displays the MfeI/SacI restriction map surrounding the palindrome. Figure 5.3 B shows 2-D agarose gel electrophoreses of chromosomal DNA isolated from the Rec$^+$ *lacZ*::246 strain grown in 0.2 % arabinose, the Δ*ruvAB* *lacZ*$^+$ strain grown in 0.2 % arabinose, and the Δ*ruvAB* *lacZ*::246 strain grown either in 0.5 % glucose or 0.2 % arabinose. A probe against the fragment containing the

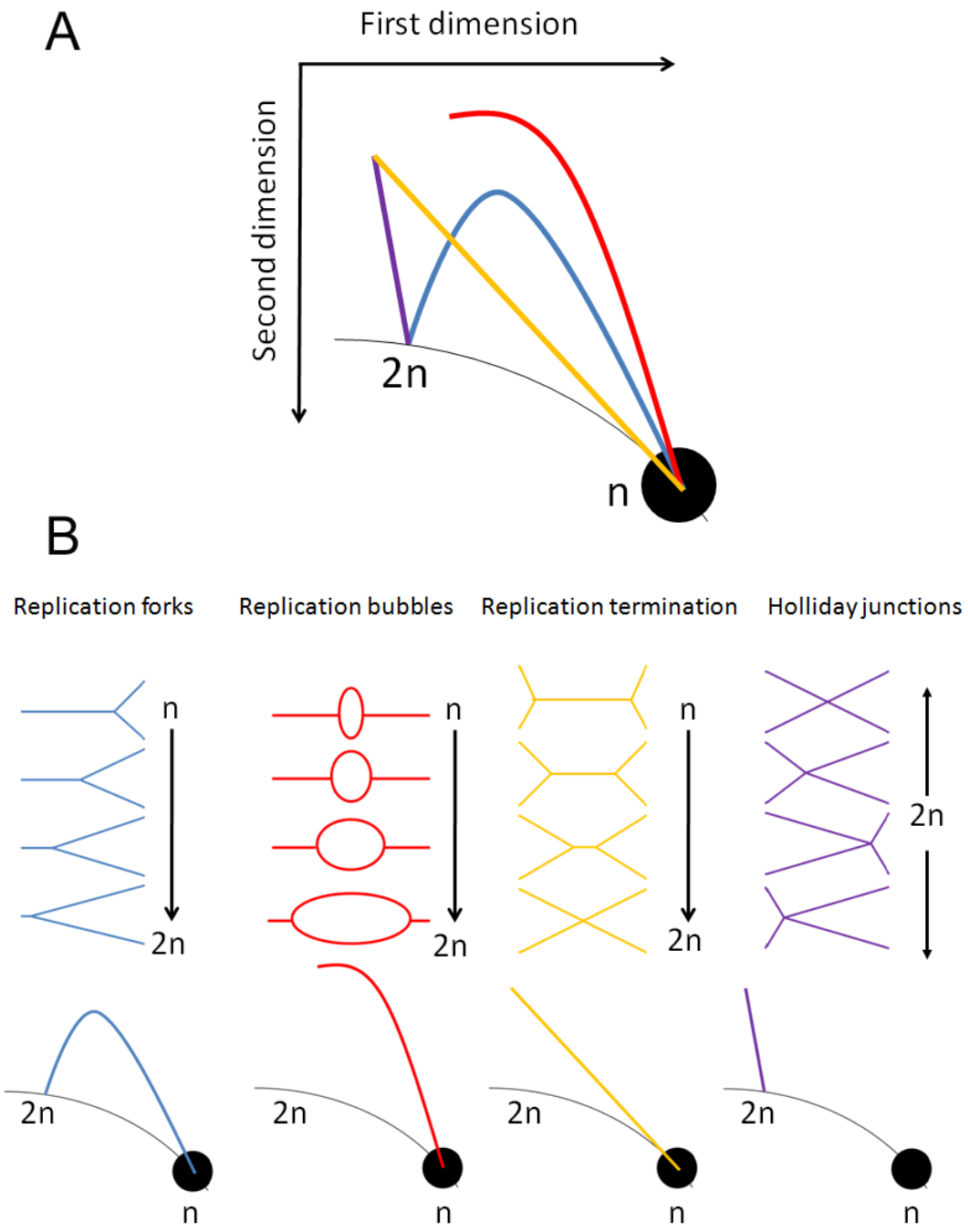

Figure 5. 3 Diagrammatic representation of the migration patterns of different species of branched DNA when separated on a native two dimensional agarose gel

(A) The arc of linear DNA is represented by the thin black line, which runs through n and 2n. n represents linear, un-replicated DNA while 2n represents linear, replicated DNA. (B) Migration patterns of replication forks, replication bubbles, replication termination, and HJs. Figure modified from: (Friedman and Brewer, 1995).

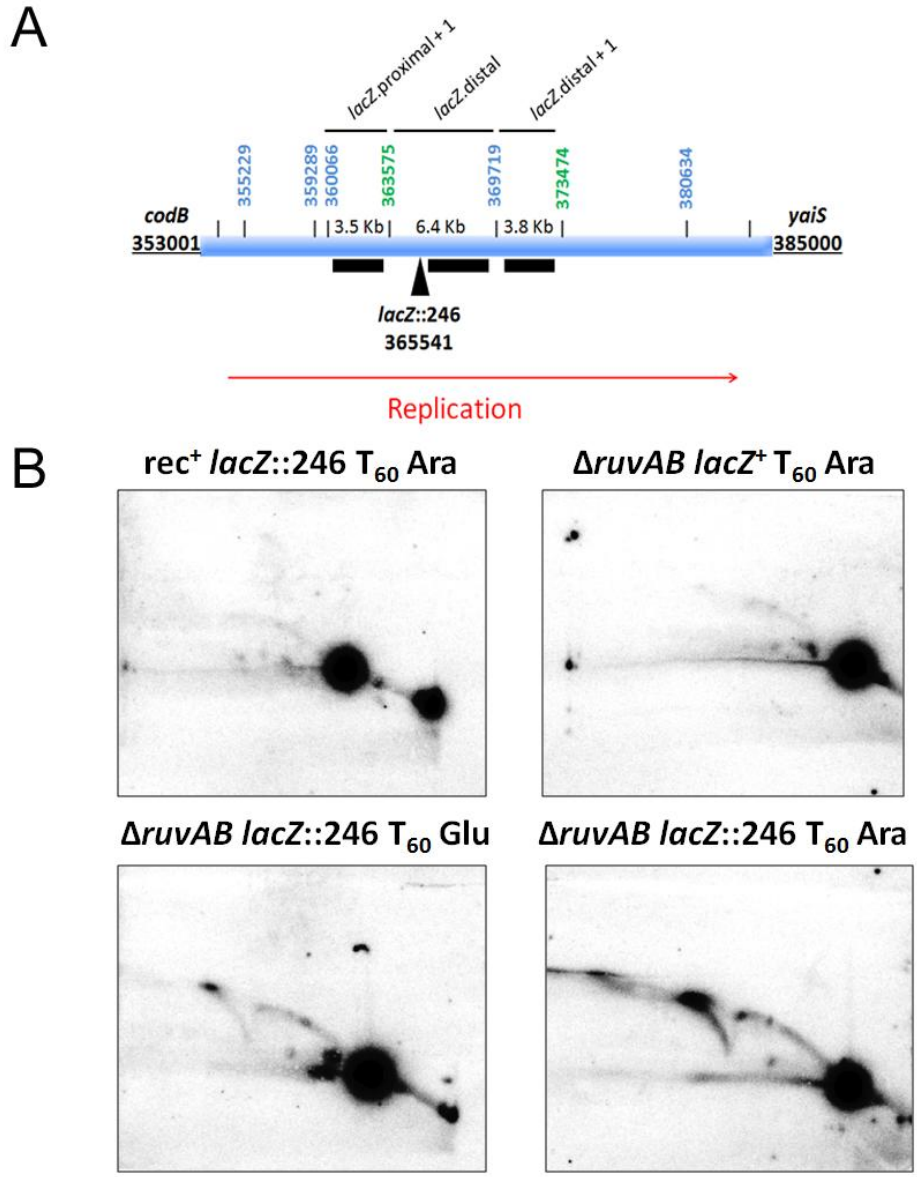

Figure 5. 4 Two dimensional agarose gel electrophoresis of Rec$^+$ and Δ*ruvAB* strains

(A) MfeI/SacI restriction map of the *E. coli* chromosome from *codB* to *yaiS* (coordinates; 353,001 to 358,000). Restriction sites of interest are marked with the respective coordinate. MfeI recognition sites are marked in blue and SacI recognition sites are marked in green. The location of the palindrome is marked by a black triangle and the binding sites for the probes are marked by black rectangles. (B) 2-D agarose gel electrophoresis. DNA was detected using the probe *lacZ*.distal. Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4243 (Δ*ruvAB lacZ*::246), and DL4257 (Δ*ruvAB lacZ*$^+$).

palindrome was used to detect the DNA (*lacZ*.distal). In all four panels of Figure 5.3 B, the linear DNA (6.4 Kb) of the *lacZ*.distal fragment can be seen as an intense spot at the bottom right-hand corner of the autoradiograph. The top two panels, which contain DNA from the Rec$^+$ *lacZ*::246 and the Δ*ruvAB lacZ*$^+$ strains, show very little DNA above the spot of linear DNA indicating that very little branched DNA is accumulated. This is expected, as no DSBs are formed in these strains, in these conditions. This result is consistent with data obtained from the pulsed-field gel analysis. In contrast, the bottom two panels, showing DNA isolated from the Δ*ruvAB lacZ*::246 strain grown either in 0. 5 % glucose or 0.2 % arabinose, both contained branched DNA. These DNA species can be seen above the spot of linear DNA as a Y-arc, which represents intermediates of replication, and an X-spike, which represents intermediates of recombination. It is interesting to note that some intermediates are detected in the Δ*ruvAB* mutant grown in 0.5 % glucose (SbcCD$^-$). This is indicative that the arabinose inducible promoter (P_{araBAD}) that controls the expression of *sbcDC*, is leaky.

5.3.2 2-D agarose gel electrophoresis of Rec$^+$, Δ*ruvAB*, and Δ*recG* strains

Once the separation of the DNA by 2-D agarose gel electrophoresis was optimised, the analysis of three restriction fragments surrounding the breakpoint was carried out in Rec$^+$, Δ*ruvAB*, and Δ*recG* strains in order to confirm the nature of the intermediates generated in these backgrounds during DSB repair. An MfeI/SacI digestion of the chromosome generated three fragments of 3.5 Kb, 6.4 Kb and 3.8

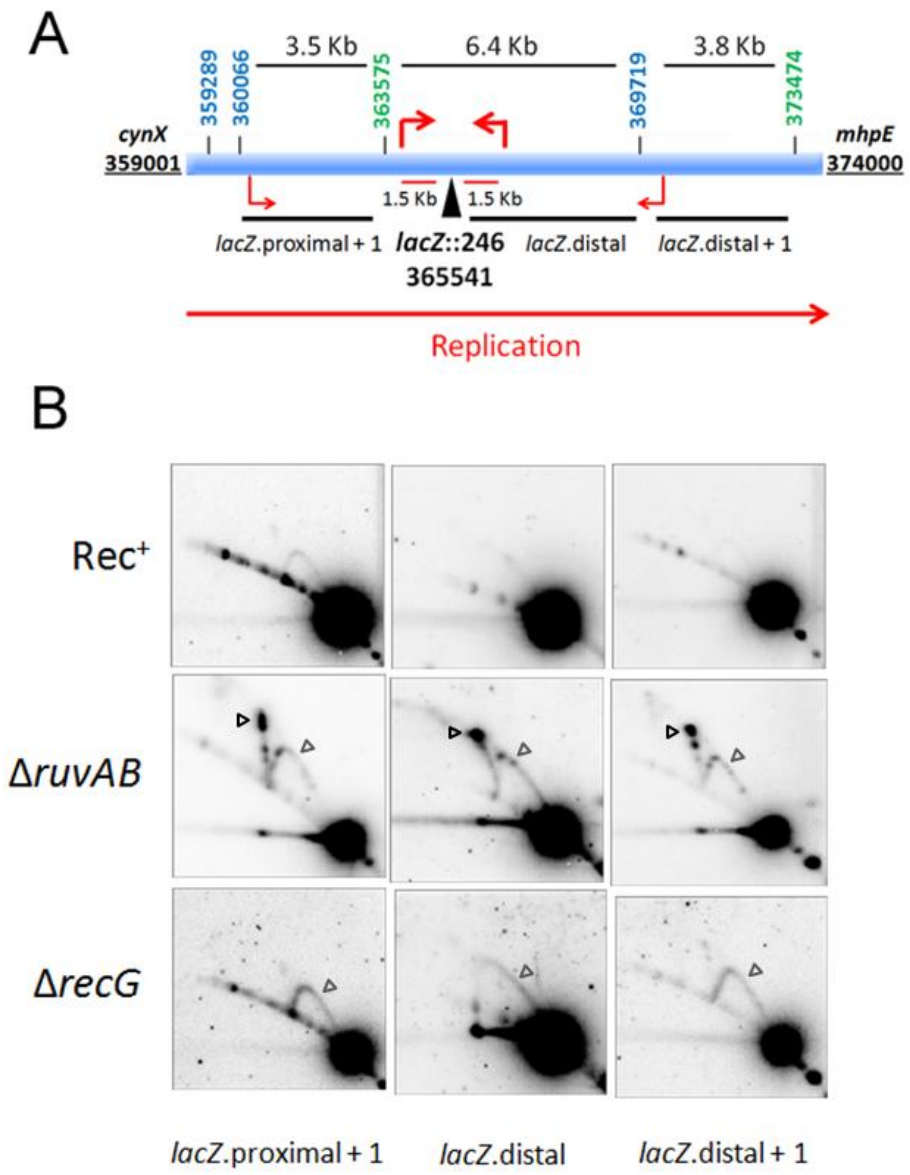

Figure 5. 5 Native 2-D agarose gel electrophoresis of DNA isolated from Rec$^+$, Δ*ruvAB*, and Δ*recG* strains containing the palindrome and in which DSBs were induced for 60 minutes

(A) MfeI/SacI digestion of the locus surrounding the palindrome. The coordinates representing MfeI restriction sites are shown in blue, while coordinates representing SacI restriction sites are shown in green. The palindrome is marked by a black triangle and binding sites of probes by black lines placed below the restriction fragments. X arrays are marked by thick red arrows and single, endogenous χ sites are marked by thin red arrows. (B) Native 2-D agarose gel electrophoreses for Rec$^+$, Δ*ruvAB*, and Δ*recG* strains. X-spikes are indicated by black arrows while grey arrows highlight Y-arcs. Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4243 (Δ*ruvAB lacZ*::246), and DL4311 (Δ*recG lacZ*::246).

Kb, containing or surrounding the palindrome (Fig. 5.5 A). These fragments were recognised using the probes *lacZ*.proximal + 1, *lacZ*.distal, and *lacZ*.distal + 1, respectively. Notably, the 6.4 Kb *lacZ*.distal fragment contained the palindrome within the first 1.966 Kb (Fig. 5.4 A). The χ arrays, which are placed 1.5 Kb from the centre of the palindrome on both the origin proximal and origin distal sides and are present within the *lacZ*.distal fragment, are indicated by thick red arrows in Figure 5.4 A. The χ array on the origin proximal side of the palindrome is located 466 bp from the beginning of the fragment while the χ array on the origin distal side is located just over half-way through the fragment (3.466 Kb from the beginning and 2.934 Kb from the end of the fragment). The smaller red arrows, which are present in the *lacZ*.proximal + 1 and *lacZ*.distal + 1 fragments, indicate the location of a single endogenous χ sequence in the correct orientation for processing a DSB generated at the palindrome. Figure 5.4 B shows the results of the 2-D agarose gel electrophoresis on DNA from the Rec$^+$, Δ*ruvAB*, and Δ*recG* strains containing the palindrome (*lacZ*::246) and grown in 0.2 % arabinose (SbcCD$^+$) for 60 minutes, whereas Figure 5.6 B displays the results of the 2-D agarose gel electrophoresis on DNA from the Rec$^+$, Δ*ruvAB* and Δ*recG* strains in which no DSBs were induced.

5.3.2 Intermediates accumulated in the Rec$^+$ strain

The top three panels of Figure 5.5 B present 2-D agarose gel electrophoreses using DNA isolated from a Rec$^+$ strain in which DSBs were induced and detected using, from left to right, the *lacZ*.proximal + 1, *lacZ*.distal, and *lacZ*.distal + 1 probes.

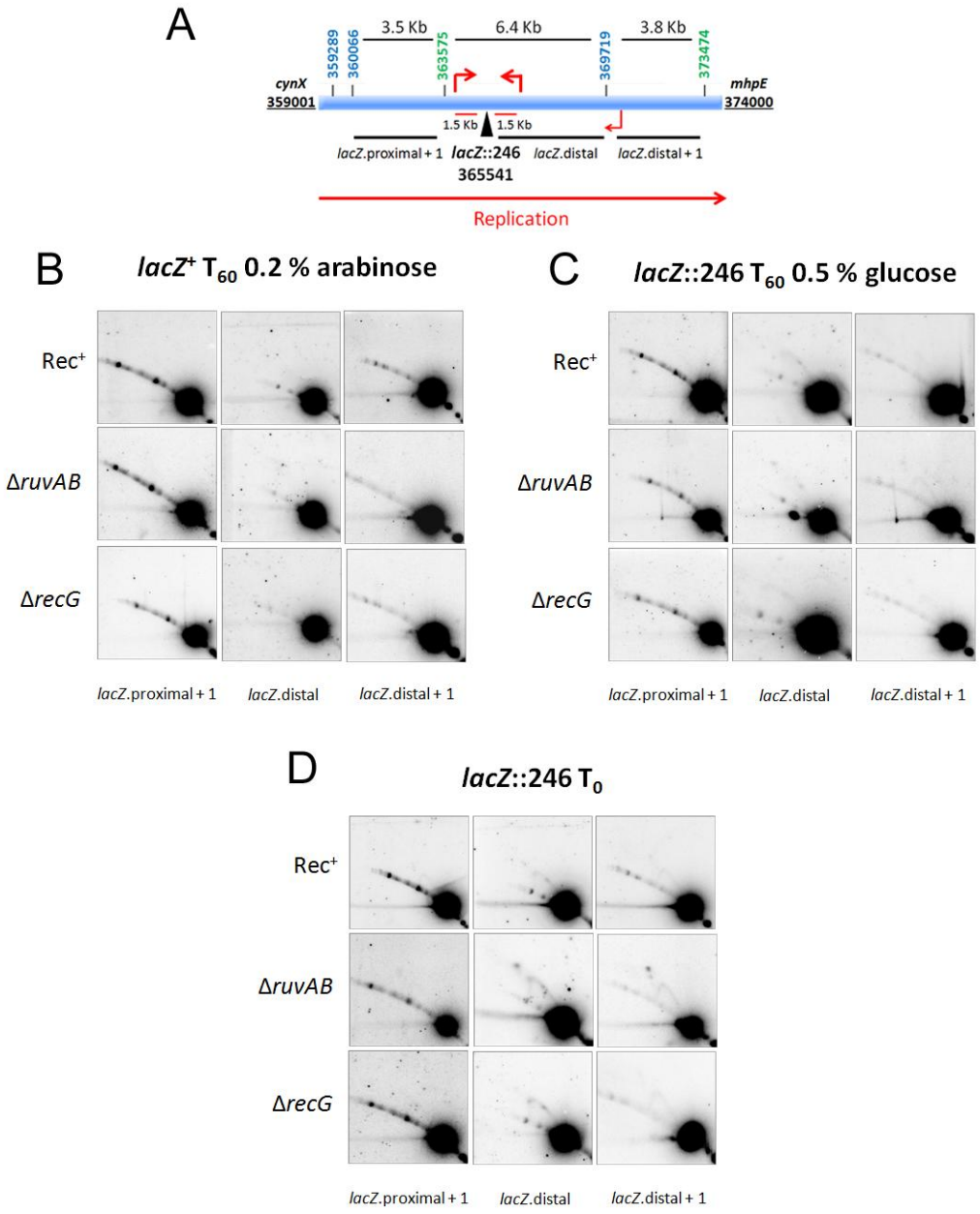

Figure 5. 6 Native 2-D agarose gel electrophoresis of DNA isolated from Rec$^+$, Δ*ruvAB*, and Δ*recG* strains in which no DSBs were induced

(A) MfeI/SacI digestion of the locus surrounding the palindrome as shown in Figure 5.4. (B-D) Native 2-D agarose gel electrophoreses for Rec$^+$, Δ*ruvAB*, and Δ*recG* strains. (B) Strains used were DL4201 (Rec$^+$ *lacZ*$^+$), DL4257 (Δ*ruvAB lacZ*$^+$), and DL4312 (Δ*recG lacZ*$^+$) grown in 0.2 % arabinose for 60 minutes. (C) Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4243 (Δ*ruvAB lacZ*::246), and DL4311 (Δ*recG lacZ*::246) grown in 0.5 % glucose for 60 minutes. (D) Strains used were DL4184 (Rec$^+$ *lacZ*::246), DL4243 (Δ*ruvAB lacZ*::246), and DL4311 (Δ*recG lacZ*::246) at Time point 0 minute.

The majority of the DNA was linear, as shown by the spot in the bottom left-hand corner of each panel. Branched DNA intermediates, located above this spot, were not abundant, which is in accordance with data from Chapter 4 where a very small amount of branched DNA was accumulated in the wells of a pulsed-field gel electrophoresis using DNA isolated from the Rec$^+$ strain upon induction of DSBs (Fig. 4.5). Nevertheless, in Figure 5.5 B, a faint Y-arc was visible in all three panels and these DNA species are likely to represent replication that is initiated during DNA repair as they were not present in controls (*lacZ*$^+$ grown in 0.2 % arabinose for 60 minutes and *lacZ*::246 at time point 0 and time point 60 minutes in 0.5 % glucose; Fig. 5.5 B). It is interesting to note that a shadow of an X-spike, presumably HJs on their way to being resolved, was visible in the *lacZ*.proximal + 1 fragment.

5.3.2.2 Intermediates accumulated in the Δ*ruvAB* mutant

The middle three panels of Figure 5.5 B contain 2-D agarose gel electrophoreses using DNA isolated from the Δ*ruvAB* mutant in which DSBs have been induced and detected using the same probes as were used to detect the DNA from the Rec$^+$ strain. As was seen with the DNA isolated from the Rec$^+$ strain, the linear DNA was the most abundant species of DNA on the autoradiograph of the Δ*ruvAB* mutant. The intermediates of repair detected in the absence of RuvAB were a lot more abundant than in the recombination proficient strain. These intermediates were made up of four-way DNA junctions, which ran along the X-

spike (marked by black arrows), and three-way DNA junctions (marked by grey arrows), which ran along the Y-arc.

It is interesting to note that in fragments *lacZ*.proximal + 1 and *lacZ*.distal + 1, four-way DNA junctions seemed to accumulate at specific locations along the restriction fragment, as shown by more than one spot detected along the X-spike. Similarly, a pattern of spots was also detected along the Y-arcs, this time in all three fragments analysed. The pattern of these spots was consistent in all three repetitions of the experiment. X-structures that branch migrate spontaneously, as might be the case for HJs that accumulate in a Δ*ruvAB* background, may stabilise in regions of the chromosome that are GC-rich. This is because the energy required to melt the three hydrogen bonds that form between guanine and cytosine is higher than the energy required to melt the two hydrogen bonds that form between adenine and thymine. The *lacZ*.proximal + 1 and *lacZ*.distal + 1 fragments have a very similar % GC content (50 and 55, respectively). Therefore, the overall GC content does not account for the pattern of spots detected on the X-spike. In order to determine whether there were clusters of GC-rich areas in the three fragments analysed, stretches of the DNA that contained five or more repeats of either cytosine or guanine, or a mixture of these bases in any order, were located (Appendix). This analysis did not reveal any region of the fragments that was particularly GC-rich. Additionally, the distribution of GC-rich stretches in the *lacZ*.distal fragment was very similar to the distribution of GC-rich stretches in the other two fragments, yet the X-spike in the *lacZ*.distal fragment did not contain more than one spot. These observations suggest that the pattern of spots detected on the X-spikes of the

lacZ.proximal + 1 and *lacZ*.distal + 1 fragments is probably not due to GC-rich stretches stabilising the HJs.

As mentioned above, the X-spike detected in the *lacZ*.distal fragment did not contain more than one spot along its length. Additionally, the shape of this X-spike was different to the shape of the X-spikes detected in the other two fragments. Instead of proceeding in a straight line from the position of the 2n up to the peak of the X-spike, it curved in towards the Y-arc. This migration pattern suggests that in the first dimension, when molecules are separated primarily based on their molecular weight, a proportion of these branched species were smaller than 2n. As a result, these structures may not represent classical HJs.

The spots detected on the Y-arcs of the three fragments analysed also contained multiple spots. The features present in the DNA fragments analysed that are most likely to cause accumulation of three-way DNA junctions (in this case a D-loop) are single, endogenous χ sites and the χ arrays. Following the formation of a DSB at the palindrome, the endogenous χ site in the *lacZ*.proximal + 1 fragment would generate a three-way DNA junction in which the majority of the fragment was in a single copy (not duplicated). This would result in a spot accumulating close to the n of the Y-arc (Fig. 5.7). This spot was not seen. Under the same circumstance, the endogenous χ site in the *lacZ*.distal + 1 fragment would generate a three-way DNA junction in which the majority of the fragment was duplicated. This would result in a spot accumulating close to the 2n of the Y-arc. This spot was also not seen. Following these observations, the endogenous χ sites in the *lacZ*.proximal + 1 and the *lacZ*.distal + 1 fragments are not responsible for generating the spots seen on the Y-arcs of these fragments. The *lacZ*.distal fragment

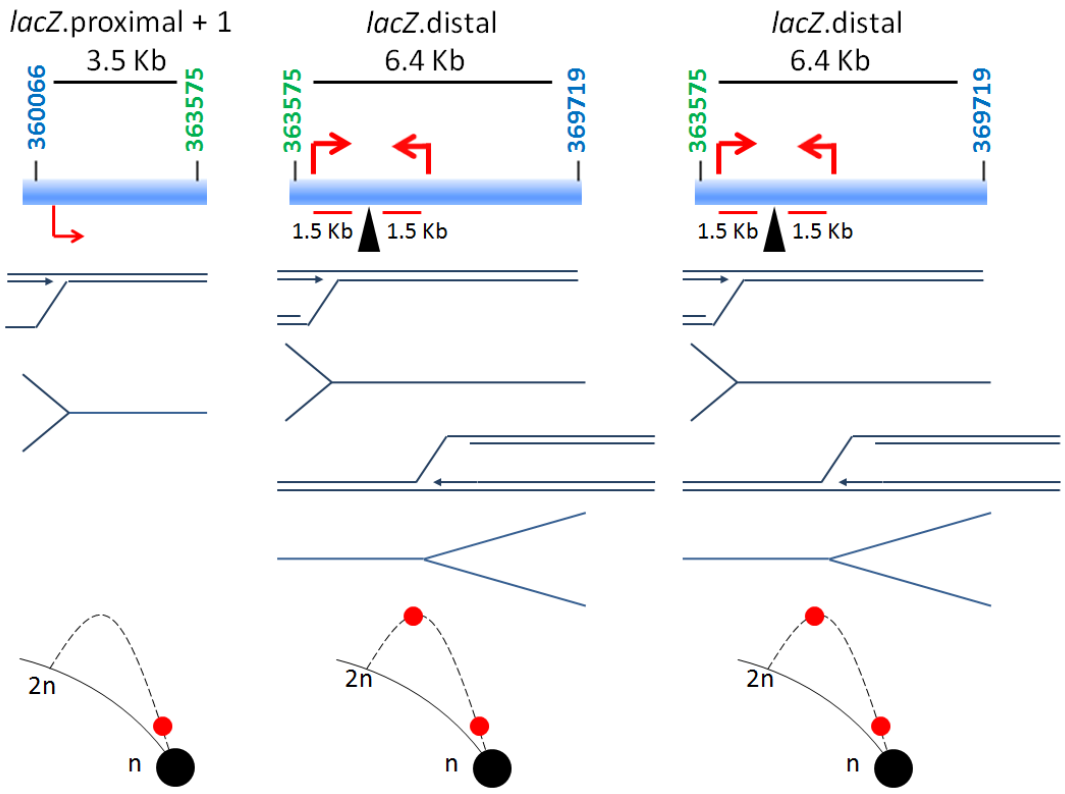

Figure 5. 7 Diagrammatic representation of the expected migration pattern of D-loops generated by the single endogenous χ sites and the χ arrays

Below each restriction fragment are diagrammatic representations of replication forks that result from D-loops, generated at the location of the χ sequences, and the corresponding three-way DNA junction they are predicted to form (shown in thin blue lines). Below these are diagrammatic representations presenting the expected localisation of the D-loops (red spots) on the Y-arcs (dashed black line) of 2-D agarose gel electrophoresis of these branched DNA molecules.

contains two χ arrays that were introduced into the chromosome for the purpose of this work. If these features were causing an accumulation of three-way junctions, spots would be generated close to the n (for the origin proximal χ array) and close to the peak of the Y-arc (for the origin distal χ array), which represents the centre of the fragment. No peak was seen close to the n. However, the signal emitted from the n is so strong that it may be masking a spot generated by the origin proximal χ array. Contrary to this, a spot at the peak of the fragment was detected. This may represent strand invasion stimulated by the origin distal χ array. Nevertheless, it is worth noting that a second spot, which was closer to n and cannot be explained, was also detected. An alternative hypothesis to explain the spots detected on the Y-arcs of the Δ*ruvAB* mutant is that the HJs that are accumulated in these strains generate a topological barrier to replication, which results in regular stalling of the replication fork.

5.3.2.3 Intermediates accumulated in the Δ*recG* mutant

The bottom three panels of Figure 5.5 B show the 2-D agarose gel electrophoreses of the DNA from the Δ*recG* mutant (*lacZ*::246) grown in 0.2 % arabinose for 60 minutes. All control conditions are presented in Figure 5.6 B. From Figure 5.5 B, it is apparent that the only intermediates of DSB repair that were accumulated in the absence of RecG were three-way DNA junctions, although there was a shadow of an X-spike in the *lacZ*.proximal + 1 and the *lacZ*.distal + 1 fragments but the intensity of these was comparable to the X-spikes detected in the

Rec$^+$ strain and were therefore not accumulated due to the absence of RecG. The Y-arcs detected in the Δ*recG* mutant are likely to be either replication forks, which are set up by the repair of the DSB, or D-loops, which are formed by RecA-mediated strand invasion of the unbroken chromosome. An interesting characteristic of these arcs is that they were darker than the Y-arcs detected in the Δ*ruvAB* mutant and did not contain spots. The Y-arc from the *lacZ*.distal fragment, the fragment containing the breakpoint, was also different from the Y-arcs detected in the two adjacent fragments. A signal for branched DNA was detected from the n up to the peak of the Y-arc, but not beyond. This indicates that half, or less, of the fragment was replicated/duplicated or that forks did not accumulate in this portion of the fragment for some other reason.

5.4 Discussion

Analysis of DSB repair intermediates by 2-D agarose gel can shed light on the mechanism of the repair process and uncover the function of different proteins. Analysis of these intermediates in a Rec$^+$ background illustrated the speed of the repair, as very few branched DNA species were detected (Fig. 5.4 B top three panels). Notably there appeared to be a weaker Y-arc in the *lacZ*.distal and *lacZ*.distal + 1 fragments when compared to the *lacZ*.proximal + 1. This is in accordance with data obtained from pulsed-field gel electrophoresis presented in Chapter 4 and may reflect an increased processing of the DNA end downstream of

the palindrome, which results in few recombination intermediates being detected here.

5.4.1 Intermediates accumulated in the Δ*ruvAB* mutant

The most intermediates were detected in the Δ*ruvAB* mutant, in which three-way DNA junctions (replication intermediates) and HJs (recombination intermediates) were detected (Fig. 5.5 B middle three panels black and grey arrows). This result suggests that RuvAB is the major protein complex required to resolve HJs in recombination proficient conditions. This indicates that in a Rec+ background, the majority of DSBs are not repaired by SDSA, as this repair pathway does not require the activity of a HJ resolvase. Additionally, as HJs are accumulated after 60 minutes of exposure to DSBs but the Δ*ruvAB* mutant does not lose viability during this period, this indicates that these HJs must eventually be processed in a RuvAB-independent manner. Finally, these four-way DNA junctions were detected in similar quantities in both the *lacZ*.proximal + 1 and *lacZ*.distal + 1. This result is a strong indication both DNA free ends of the DSB are actively involved in repair. This supports the view that in a Rec+ background, the DSB is repaired by canonical DSBR.

Multiple spots were detected along the X-spikes of the *lacZ*.proximal + 1 and *lacZ*.distal + 1 fragments and sequence analysis of the region did not identify any feature within these fragments that would account for HJs accumulating at specific locations due to the stabilizing of branch migration. It has been reported that the preparation of chromosomal DNA in agarose plugs efficiently blocks branch

migration of branched DNA (Cromie et al., 2006). Nevertheless, it would be interesting to carry out 2-D agarose gel analysis on DNA that has been crosslinked prior to chromosomal plug preparation. This experiment might restrict branch migration even further, confirm whether these spots are a result of spontaneous branch migration or not. If the spots are still detected when the DNA is crosslinked, this result might be an indication that the spots are composed by different branched species, such as single and double HJs. Both single and double HJs have a DNA content of 2n. This means that in the first dimension they would both migrate around the 2n mark. In the second dimension, these two structures might be separated, as this dimension separates branched DNA based on three dimensional shape. A single HJ has a single branch-point, which leaves the arms of the molecule free to invade a large spatial area. Contrary to this, a double HJ has two branch-points. This would tie the arms of the molecule together and prevent them from adopting a large three dimensional structure. Assuming these hypotheses are true, single HJs would migrate closer to the peak of the X-spike while double HJs might be retained closer to the base of the spike and therefore closer to the 2n mark. In fact, according to published data, both branched species are predicted to lie along the X-spike (Bzymek et al.; Bzymek et al., 2010; Cromie et al., 2006; Schwacha and Kleckner, 1994; Schwacha and Kleckner, 1995). To confirm this hypothesis, the individual spots could be isolated following their separation by native 2-D agarose gel electrophoresis and analysed using an electron microscope. Indeed, this technique has been used to identify both species of HJs in recombination intermediates that were isolated from different yeast species (Cromie et al., 2006).

Another interesting point of discussion is the shape of the X-spike detected in the *lacZ*.distal fragment. As mentioned above, the four-way DNA junctions located close to the 2n in the second dimension appear to have a slightly smaller molecular weight than 2n in the first dimension. Branched DNA structures similar to these have been described in *Saccharomyces cerevisiae* and bacteriophage T4 (Doksani et al., 2009; Long and Kreuzer, 2008). These structures represent replication forks (three-way DNA junctions) that are converted into HJs (four-way DNA junctions), typically by replication fork reversal (RFR; Fig. 5.8). RFR has been described to occur in *E. coli* when a replication fork is arrested. In this context, RuvAB has been implicated in catalysing the reversal of the fork (Seigneur et al., 1998). However, a replication fork arrest is not predicted to occur during repair of the DSB formed at the palindrome. Additionally, these putative reversed forks were detected in a Δ*ruvAB* mutant. Taking all this information into account, it does not appear that the X-spike detected in the *lacZ*.distal fragment can be a typical RuvAB-mediated reversed fork.

Another explanation would be that this four-way DNA junction represents a broken DNA fragment onto which RecBCD has loaded, translocated until one of the χ arrays without degrading the DNA free end, nicked at χ and catalysed strand invasion (Fig 5.9; Chapter 1 Section 1.3.1.2) (Taylor and Smith, 1980; Taylor et al., 1985). This D-loop would not be composed of three chromosomal arms, but four, and would be predicted to migrate like a reversed fork on a 2-D agarose gel. In essence, this molecule would be a reversed fork with a nick in it generated by strand invasion (a pseudo-reversed fork). If this hypothesis is true, it may finally shed light on the true nature of RecBCD activity in the cell. This hypothesis predicts that

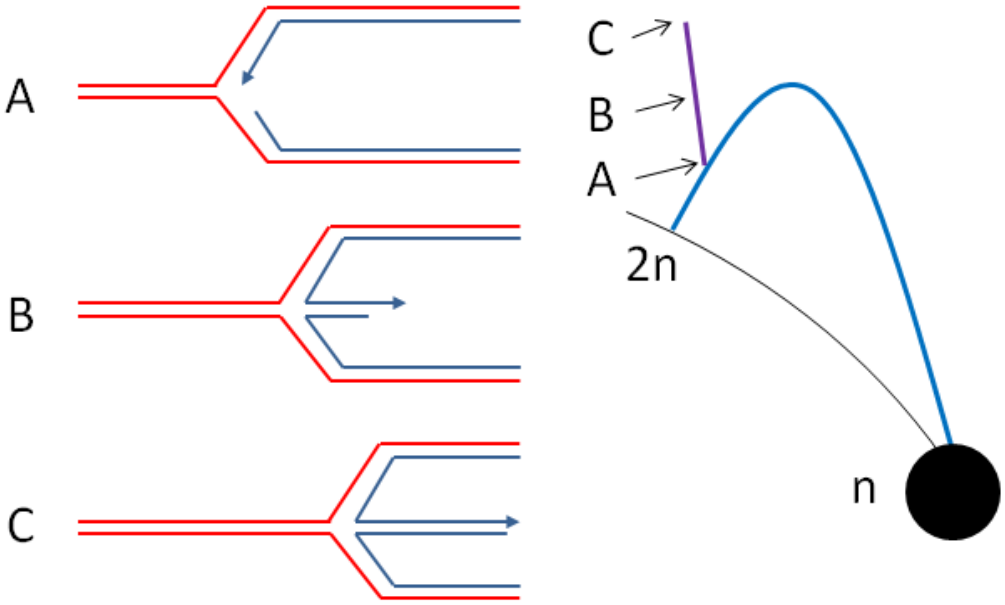

Figure 5. 8 Diagrammatic representation of replication fork reversal and the predicted migration pattern of a reversed fork when separated by native 2-D agarose gel electrophoresis

A, B, and C represent a fork during different stages of reversal. The location of these structures is highlighted on the diagram of the native 2-D agarose gel to the left of the figure. Blue arc represents three-way DNA junctions, while purple arc represents four-way DNA junctions.

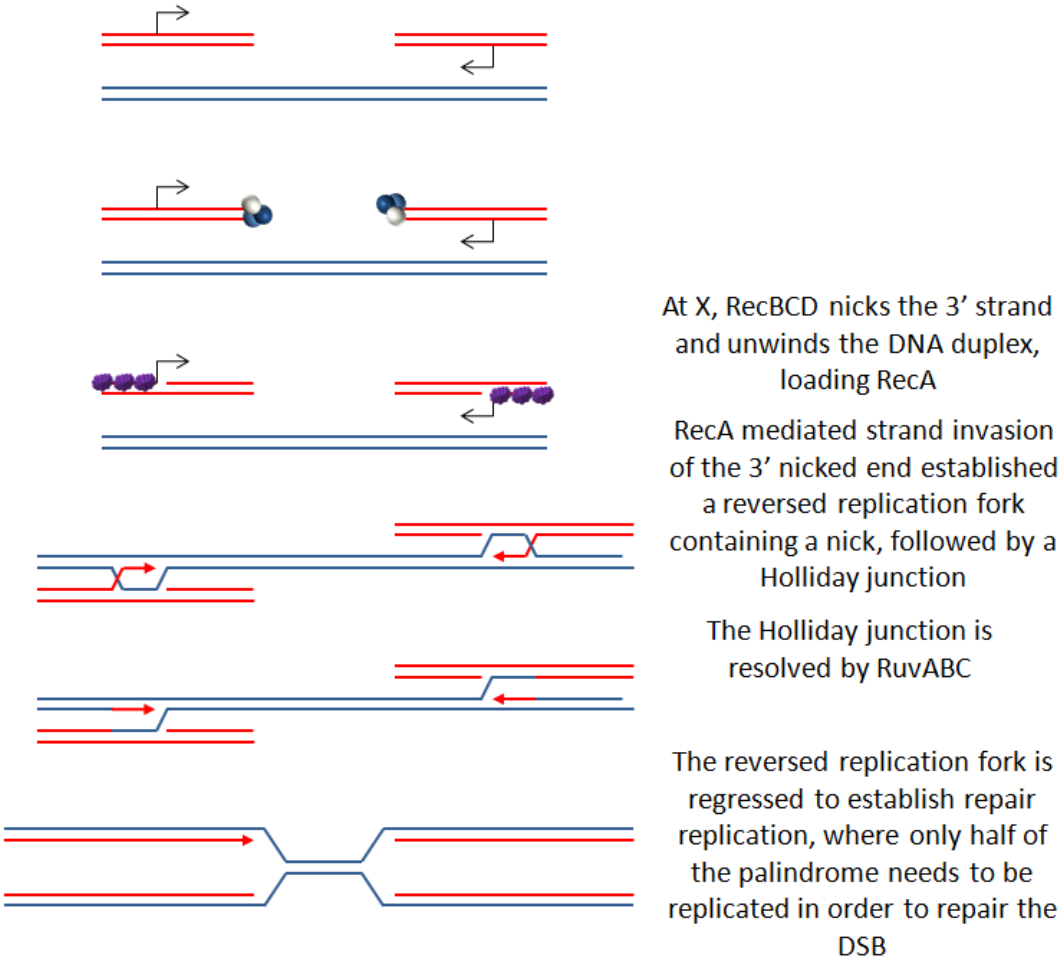

Figure 5. 9 Formation of a Pseudo-reversed fork from a D-loop generated following RecBCD nick at χ model

The DSB, which contains half of the palindrome at the end of the downstream DNA free end, is recognised by RecBCD. RecBCD translocates along the broken DNA strands until a χ sequence is recognised, where it nicks and unwinds the 3' strands. The 3' strands then invade the unbroken chromosome generating four-way DNA D-loops reminiscent of reversed forks. In order to re-establish replication from these D-loops, RuvAB regresses them by branch migration.

one of the chromosome arms would be only 1.5 Kb long, because strand invasion is initiated at the χ array, which would determine the location of the junction for this four-way DNA molecule. The relative size of each arm can be verified by electron microscopy. In addition, altering the location of the χ array should change the length of one of the chromosome arms, which can also be verified using electron microscopy. Another prediction made by this hypothesis is that the end of one of the chromosome arms should contain half of the palindrome, as SbcCD is predicted to cleave close to the centre of the hairpin (Connelly et al., 1998). The presence of this sequence could be verified by Southern blot. Finally, in order to complete repair, this pseudo-reversed fork would need to be processed to re-establish a productive replication fork. This processing could happen either by degradation of the shorter chromosome arm (the reversed arm in conventional reversed forks or the 1.5 Kb arm in the case of the hypothesis described in this work) or by regression of the junction (Long and Kreuzer, 2008). There is no evidence of degradation of one of the arms, as this would generate a smear between the peak of the X-spike and the peak of the Y-arc. Therefore, the pseudo-reversed fork would need to be regressed by branch migration in order to re-establish a replication fork. RuvAB may be required to catalyse this process.

The spots detected on the Y-arc, an arc that represents intermediates of replication, are difficult to explain. The only spot that localised with a DNA sequence that would be expected to generate an accumulation of three-way DNA junctions was the spot at the peak of the Y-arc detected in the *lacZ*.distal fragment. This spot was located close to the origin distal χ array and might have therefore been generated by RecA-mediated strand invasion at this sequence. In order to

verify whether the spot is generated by the χ array, the experiment could be repeated in a Δr*uvAB* strain lacking this sequence, which would predict to cause the disappearance of the spot. Considering that this spot was the only spot that could be explained using the χ array hypothesis, a more plausible hypothesis is that the spots represent replication forks that regularly stall due to the accumulation of unresolved HJs in the absence of RuvAB. If this hypothesis is accurate, it would implicate that HJs need to be resolved prior to the establishment of productive replication. In many models of DSB repair, repair replication happens during synapsis and prior to the resolution of HJs, which takes place in post-synapsis (Chapter 1; Section 1.4). Data presented in this work suggests that these two events, repair replication and HJ resolution, may not be as independent from one another as many models suggest.

5.4.2 Intermediates accumulated in the Δ*recG* mutant

Y-arcs were the only form of branched DNA that accumulated in the absence of RecG. These structures could either be replication forks or D-loops. It is interesting to note that in the *lacZ*.distal fragment only part of the Y-arc was detected. This result is consistent with these branched DNA species being D-loops as the fraction of the arc that was not detected corresponds to the portion of the fragment that is located between the χ arrays. If D-loops were established, presumably at the χ arrays, but never matured to form productive replication forks, the region between the χ arrays would never be replicated and therefore not detected in a 2-D agarose gel. These data are consistent with the data obtained from

pulsed-field gel electrophoresis presented in Chapter 4, which place RecG upstream of RuvAB in the repair pathway and implicate RecG in the maturation of D-loops. It would be interesting to alter the location of the χ arrays within the *lacZ*.distal fragment and observe whether the portion of the fragment that does not generate a y-arc signal changes location accordingly. Another experiment that could confirm this hypothesis and disprove a hypothesis in which over-replication is established in a Δ*recG* mutant would be to determine the directionality of the three-way DNA junctions detected. Finally, the lack of spots detected in the Y-arcs generated by this mutant confirms that there are no DNA sequences that induce the accumulation of three-way DNA junctions at specific locations along the fragments to generate spots. This result strengthens the hypothesis that the spots detected on the Y-arcs of the Δ*ruvAB* mutant are replication forks that stall in the presence of unresolved HJs.

5.4.3 Conclusions

2-D agarose gel analysis of the repair of an SbcCD-mediated, site-specific DSB has shed light on the mechanism of repair. The lack of intermediates detected in a Rec$^+$ strain illustrates the efficiency and speed of this repair process and can explain why a recombination proficient strain does not lose viability in conditions of chronic DSB formation (Chapter 3; Fig. 3.3) (Eykelenboom et al., 2008). Studying the repair process in the Δ*ruvAB* mutant has confirmed that in the absence of this protein complex HJs are accumulated on both sides of the breakpoint, strongly

supporting the view that both DNA free ends invade the unbroken chromosome. Additionally, the results obtained confirm that in a Rec^+ background HJs are resolved by RuvAB, suggesting that the DSB is repaired by the canonical DSBR pathway. Additionally, the detection of HJs within 60 minutes of exposure to DSBs, an exposure time that does not require functional RuvAB for survival, confirms that in the absence of RuvAB, a RuvAB-independent pathway for the resolution/dissolution of HJs becomes available to the cell. Finally, it was interesting to find that HJs were not accumulated in the Δ*recG* mutant, indicating that this helicase is not involved in the resolution of these four-way DNA junctions when RuvAB is present. Three-way DNA junctions were the only form of branched DNA detected but in order to fully understand the role of RecG in DSB repair, it remains important to establish whether these three-way DNA junctions are replication forks or D-loops.

Chapter 6

Conclusions and future work

6.1 Conclusions

The work presented in this thesis aims to identify the HR intermediates that arise during the repair of a site-specific SbcCD-mediated DSB, which is generated at the site of a 246 bp interrupted palindrome in the *E. coli* chromosomal *lacZ* gene, and to determine the function of the RecG recombination helicase in this process. In order to concentrate intermediates close to the breakpoint, strains used for the analysis of the chromosomal DNA by Southern blot were modified to contain three repeated χ recombinational hotspot sequences (χ array) on each side of the palindrome. Based on growth curves and viability assays, it was confirmed that a recombination proficient strain does not lose viability under conditions of chronic DSB formation (Chapter 3). This finding illustrates the efficiency of the repair process. Additionally, it was found that there was no requirement for RuvAB after acute DSB formation (Chapter 3), yet a large amount of branched DNA was

accumulated during this time in the absence of RuvAB (Chapters 4). It was subsequently shown that Holliday junctions (HJs) made up a large proportion of this branched DNA and that these structures were detected either side of the breakpoint (Chapter 5). This result confirms that in a recombination proficient background the repair of this DSB requires a HJ resolvase to resolve HJs located on both sides of the breakpoint. These requirements are indicative of repair occurring via the canonical DSBR pathway (Chapter 1; Fig. 1.5). Additionally, a RuvAB-independent mechanism for HJs resolution/dissolution is available to the cell for survival when RuvAB is not present, as Δ*ruvAB* mutants survived a limited number of DSBs. Finally, data obtained from 2-D agarose gel electrophoreses using DNA from the Δ*ruvAB* mutant suggested that replication, initiated during repair, regularly stalled. It was concluded that un-resolved HJs might generate a topological barrier to efficient replication. Strikingly, this result placed the resolution of HJs upstream of the PriA-mediated fork restart (Chapter 5). It was also shown that HJs were not accumulated in a Δ*recG* mutant incurring DSBs. Instead of HJs, a Δ*recG* mutant accumulated three-way DNA junctions representative of either replication forks or D-loops (Chapter 5). This information, coupled with data obtained from pulsed-field gel electrophoreses of DNA isolated from Δ*recG* and Δ*ruvAB* Δ*recG* mutants, supports the view that the HR intermediates that were accumulated in the absence of RecG were D-loops and suggested that RecG acts upstream of RuvAB and before the formation of HJs (Chapter 4). Nevertheless, this hypothesis remains to be confirmed. All these conclusions are summarised in Figure 6.1.

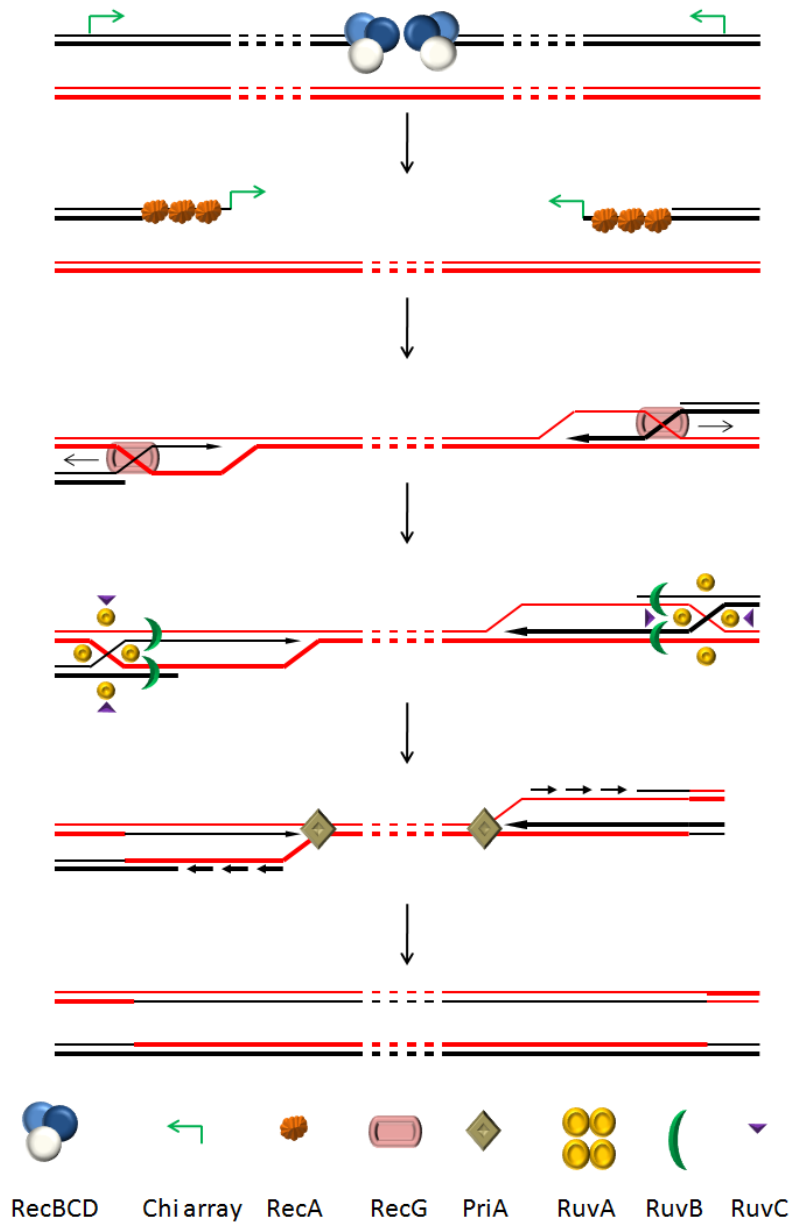

Figure 6. 1 Repair of a two-ended site-specific DSB in the *E. coli lacZ* gene

RecBCD processes the DSB ends to generate 3' ssDNA, which becomes coated with RecA. RecA searches the genome for a homologous sequence and mediates strand-invasion of the unbroken homologue. The products of strand-invasion, D-loops, are then bound by RecG, which catalyses the migration of these three-way DNA junctions to form four-way DNA junctions (HJs). The HJs can then be resolved by RuvABC, and potentially an alternative resolution/dissolution pathway, to allow for the establishment of new replication forks. PriA mediates the assembly of new replisomes and two converging replication forks are established. Replication proceeds to close the gap in the DNA generated by RecBCD-mediated degradation of the DNA DSB and two intact chromosomes are re-established.

Finally, an interesting observation was made when analysing the 2-D agarose gel electrophoresis obtained using DNA isolated from the Δ*ruvAB* mutant. Evidence for branched DNA reminiscent of reversed replication forks was found in the fragment containing the breakpoint (Chapter 5). One suggested explanation for the formation of this structure in the fragment that contained the breakpoint was that it represented a D-loop generated from an event in which RecBCD translocated along the duplex and nicked at χ, without degrading the dsDNA free end, to generate a pseudo-reversed fork, the regression of which was dependent on RuvAB. This hypothesis of RecBCD action was proposed by Taylor and collaborators but concrete evidence to support it has not yet been obtained (Taylor and Smith, 1980; Taylor et al., 1985). The use of the SbcCD/palindrome cleavage system might be the appropriate tool to finally prove or disprove this hypothesis.

If this hypothesis is shown to be correct, it would answer one of the big questions that has emerged from the study of the SbcCD/palindrome cleavage system. The cleavage of the palindrome to generate a DSB is replication-dependent, yet the repair of the DSB to restore the palindrome requires replication. Why is the palindrome not re-cleaved during repair replication? A study that analysed the fate of the 246 bp interrupted palindrome that was introduced into *E. coli* by λ phage lysogenisation, suggested that the mechanism of SbcCD-mediated cleavage of the palindrome generated two asymmetrical DNA free ends. These two ends were processed differently in order for repair of the DSB to take place (Chapter 1; Section 1.5) (Connelly et al., 1998; Cromie et al., 2000). It is plausible that as a result of this hypothesis the DNA downstream of the palindrome contains half of the palindromic sequence at it's extremity, while the palindromic sequence in DNA free

end upstream of the DSB might have been completely lost (Fig. 6.2). Following repair of this DSB by the nick at χ model (Chapter 1; Section 1.3.1.2), two converging replication forks would be established where half of the palindrome has already been replicated. This would mean that during replication associated with repair, only half of the palindrome remains to be synthesized. This would not allow the hairpin structure to form and therefore SbcCD-mediated cleavage of the palindrome would not occur (Fig. 6.3).

6.2 Future work

6.2.1 Synchronisation of replication

All experiments described in this thesis were carried out using asynchronous cell cultures. The disadvantage of this is that the formation of the SbcCD-mediated DSB may not have occurred in all the cells at the same time, which would have reduced the yield of the intermediates of repair. For future work, synchronising the cells prior to any analysis would generate a more accurate picture and allow for easier interpretation of the data. In *E. coli*, there are various techniques to achieve this. One of these techniques relies on the chemical serine hydroxamate, which is an analogue of serine. Addition of this amino acid analogue to the growth media of *E. coli* cultures inhibits growth. Normal growth can be recovered by the addition of serine back to the growth media (Ferullo et al., 2009; Tosa and Pizer, 1971).

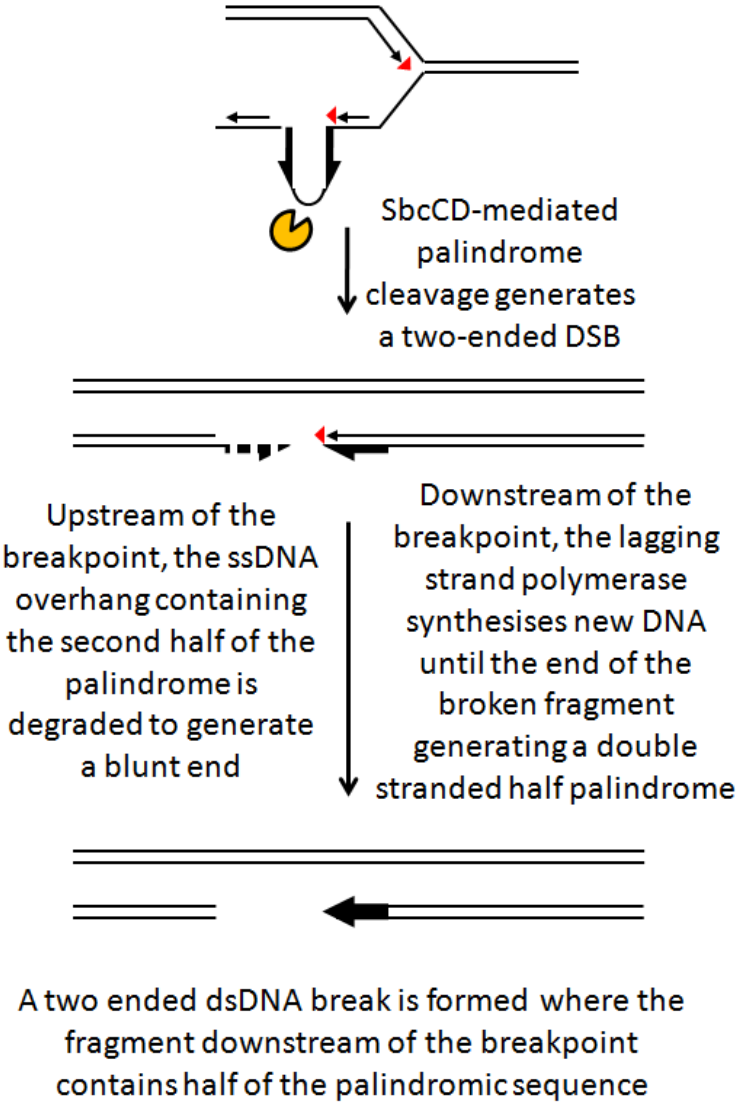

Figure 6. 2 Formation of an asymmetric DSB by SbcCD-mediated palindrome cleavage

During replication the palindrome on the lagging strand template is extruded into a hairpin, which is recognised and cleaved by SbcCD. SbcCD is thought to cleave in the loop region that is located at the tip of the hairpin. The lagging strand polymerase (represented by a red triangle) might carry on synthesising until the cleavage point to generate a blunt end that contains half of the palindrome downstream of the breakpoint. The palindrome on the DNA free end upstream of the DSB cannot be replicated so the region of ssDNA is degraded until a double stranded, blunt end is formed. This generates an asymmetric DSB where the DNA free end upstream contains half of the palindrome.

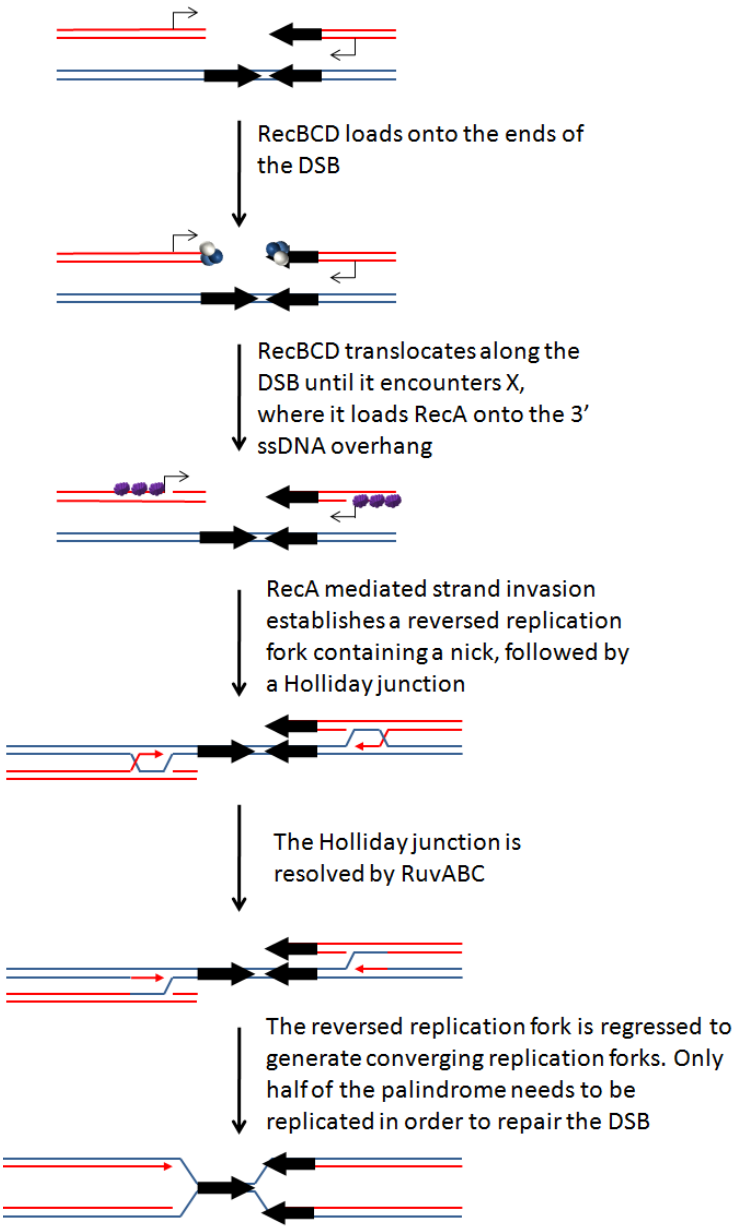

Figure 6. 3 The roles of RecBCD and RuvAB in the repair of an SbCD-mediated DNA double strand break in the *E. coli lacZ* gene

Due to the nature of the SbcCD-mediated DSB, the DNA free end upstream of the breakpoint loses half of the palindrome, while the other half is retained on the DNA free end downstream of the DSB. RecBCD recognises these DNA ends and translocates along them until χ is recognised. Upon χ recognition, the 3' strands are nicked and unwound to allow for the loading of RecA. RecA mediates strand invasion to generate four-way D-loops. These D-loops are then regressed by RuvAB to establish two converging replication forks in which half of the palindromic sequence has already been replicated. During repair replication only half of the palindrome remains to be replicated, meaning that during repair the palindrome does not have the opportunity to form another hairpin for SbcCD to cleave.

6.2.2 Physical analysis of recombination intermediates by pulsed-field gel electrophoresis

Further analysis using pulsed-field gel electrophoresis would be needed to generate a more conclusive picture of exactly where intermediates are generated in Rec+, Δ*ruvAB*, and Δ*recG* strains and how much DNA is degraded in Δ*ruvAB* Δ*recG* mutants. It would be useful to introduce I-SceI restriction sites into the chromosome so as to generate equal sized fragments immediately either side of the palindrome. Additionally, to analyse the extent of degradation surrounding the palindrome, it would be necessary to locate equal sized probes at the same distance either side of the breakpoint. As depicted in Figure 6.2, a portion downstream of the palindrome is predicted to lie in a region of singe-stranded DNA, which gets processed to generate a blunt end for RecBCD to act upon. It is possible that the 1.5 Kb χ array that is located downstream of the palindrome falls into this ssDNA region. This would prevent RecBCD from recognising this χ array. It would be interesting to see whether the degradation on the DNA downstream of the palindrome that was detected in the Δ*ruvAB* Δ*recG* mutant would be reduced in a strain that contains χ arrays that are located further away from the palindrome (3 Kb or more). This experiment will help to ascertain that the repair of the DSB occurs via canonical DSBR. These experiments could then be extended to the study of additional repair and recombination mutants.

6.2.3 Study of the requirement for other repair and recombination proteins

The work in this thesis addresses the requirement for RuvAB and RecG. It would be interesting to extend this analysis to other repair and recombination proteins. Studying the repair of the SbcCD-mediated DSB in cells lacking either RuvC alone or the whole HJs resolvase complex, RuvABC, might highlight a RuvC-independent activity of the RuvAB complex and *vice versa*. Additionally, it would be interesting to study Δ*uvrD* and/or Δ*recQ* mutants, as strains harbouring these mutations have been described to have a small colony phenotype following SbcCD-mediated cleavage of a 246 bp interrupted palindrome (unpublished data).

6.2.4 Two dimensional agarose gel electrophoresis

Native 2-D agarose gels were used in this work to visualise different branched DNA species generated by the repair of a site-specific DSB generated by SbcCD-mediated cleavage of a 246 bp interrupted palindrome located in the chromosomal *lacZ* gene of *E. coli*. These experiments used DNA isolated from asynchronous strains that had been subjected to DSBs for 60 minutes. This provided only a weak snapshot of the intermediates generated during the repair process. Carrying out 2-D agarose gel electrophoresis using DNA isolated following a time-course of DSB induction in synchronised cell cultures might reveal additional branched structures that were not detected in the work presented here. In addition to native 2-D agarose gel electrophoresis, alkaline 2-D agarose gel electrophoresis

would allow for the dissection of the DNA strand make-up of these recombination intermediates. This analysis should allow for the distinction of D-loops from replication forks, as well as reversed replication forks from HJs. The ability to distinguish reversed forks from HJs could be used to prove or disprove whether the four-way DNA junctions detected in the Δ*ruvAB* mutant, in the fragment containing the breakpoint, is a reversed fork.

6.2.5 Analysis of recombination intermediates by electron microscopy

A useful technique used to complement 2-D agarose gel elactrophoresis is the visualisation of DNA using EM, as this technique can confirm the exact structure of intermediates isolated by 2-D agarose gel electrohporesis. For example D-loops could be distinguished from replication forks and double HJs could be distinguished from single HJs. EM would also be a useful tool to investigate the true nature of the reversed forks detected in the *lacZ*.distal fragment of the Δ*ruvAB* mutant, as the length of different chromosome arms can be measured.

References

Al-Deib, A. A., Mahdi, A. A. and Lloyd, R. G. (1996). Modulation of recombination and DNA repair by the RecG and PriA helicases of Escherichia coli K-12. *J Bacteriol* **178**, 6782-9.

Allers, T. and Lichten, M. (2001). Intermediates of yeast meiotic recombination contain heteroduplex DNA. *Mol Cell* **8**, 225-31.

Amundsen, S. K., Taylor, A. F., Chaudhury, A. M. and Smith, G. R. (1986). recD: the gene for an essential third subunit of exonuclease V. *Proc Natl Acad Sci U S A* **83**, 5558-62.

Amundsen, S. K., Taylor, A. F., Reddy, M. and Smith, G. R. (2007). Intersubunit signaling in RecBCD enzyme, a complex protein machine regulated by Chi hot spots. *Genes Dev* **21**, 3296-307.

Anderson, D. G. and Kowalczykowski, S. C. (1997). The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner. *Cell* **90**, 77-86.

Ashton, T. M., Mankouri, H. W., Heidenblut, A., McHugh, P. J. and Hickson, I. D. (2011). Pathways for Holliday Junction processing during homologous recombination in Saccharomyces cerevisiae. *Mol Cell Biol*.

Baharoglu, Z., Bradley, A. S., Le Masson, M., Tsaneva, I. and Michel, B. (2008). ruvA Mutants that resolve Holliday junctions but do not reverse replication forks. *PLoS Genet* **4**, e1000012.

Bar-Ziv, R. and Libchaber, A. (2001). Effects of DNA sequence and structure on binding of RecA to single-stranded DNA. *Proc Natl Acad Sci U S A* **98**, 9068-73.

Barre, F. X., Aroyo, M., Colloms, S. D., Helfrich, A., Cornet, F. and Sherratt, D. J. (2000). FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes Dev* **14**, 2976-88.

Barre, F. X., Soballe, B., Michel, B., Aroyo, M., Robertson, M. and Sherratt, D. (2001). Circles: the replication-recombination-chromosome segregation connection. *Proc Natl Acad Sci U S A* **98**, 8189-95.

Bell, L. and Byers, B. (1983). Separation of branched from linear DNA by two-dimensional gel electrophoresis. *Anal Biochem* **130**, 527-35.

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**, 15513-20.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F. et al. (1997). The complete genome sequence of Escherichia coli K-12. *Science* **277**, 1453-62.

Boehmer, P. E. and Emmerson, P. T. (1992). The RecB subunit of the Escherichia coli RecBCD enzyme couples ATP hydrolysis to DNA unwinding. *J Biol Chem* **267**, 4981-7.

Bowater, R. and Doherty, A. J. (2006). Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet* **2**, e8.

Brewer, B. J. and Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell* **51**, 463-71.

Broker, T. R. and Lehman, I. R. (1971). Branched DNA molecules: intermediates in T4 recombination. *J Mol Biol* **60**, 131-49.

Buss, J. A., Kimura, Y. and Bianco, P. R. (2008). RecG interacts directly with SSB: implications for stalled replication fork regression. *Nucleic Acids Res* **36**, 7029-42.

Bzymek, M., Thayer, N. H., Oh, S. D., Kleckner, N. and Hunter, N. Double Holliday junctions are intermediates of DNA break repair. *Nature* **464**, 937-41.

Bzymek, M., Thayer, N. H., Oh, S. D., Kleckner, N. and Hunter, N. (2010). Double Holliday junctions are intermediates of DNA break repair. *Nature* **464**, 937-941.

Capaldo-Kimball, F. and Barbour, S. D. (1971). Involvement of recombination genes in growth and viability of Escherichia coli K-12. *J Bacteriol* **106**, 204-12.

Centore, R. C. and Sandler, S. J. (2007). UvrD limits the number and intensities of RecA-green fluorescent protein structures in Escherichia coli K-12. *J Bacteriol* **189**, 2915-20.

Chen, Z., Yang, H. and Pavletich, N. P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489-4.

Churchill, J. J. and Kowalczykowski, S. C. (2000). Identification of the RecA protein-loading domain of RecBCD enzyme. *J Mol Biol* **297**, 537-42.

Churchward, G., Belin, D. and Nagamine, Y. (1984). A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene* **31**, 165-71.

Clark, A. J. and Chamberlin, M. (1966). Abnormal metabolic response to ultraviolet light of a recombination deficient mutant of Escherichia coli K12. *J Mol Biol* **19**, 442-54.

Clark, A. J. and Margulies, A. D. (1965). Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12. *Proc Natl Acad Sci U S A* **53**, 451-9.

Cleaver, J. E. and Boyer, H. W. (1972). Solubility and dialysis limits of DNA oligonucleotides. *Biochim Biophys Acta* **262**, 116-24.

Collins, J. (1981). Instability of palindromic DNA in Escherichia coli. *Cold Spring Harb Symp Quant Biol* **45 Pt 1**, 409-16.

Connelly, J. C., Kirkham, L. A. and Leach, D. R. (1998). The SbcCD nuclease of Escherichia coli is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proc Natl Acad Sci U S A* **95**, 7969-74.

Connolly, B., Parsons, C. A., Benson, F. E., Dunderdale, H. J., Sharples, G. J., Lloyd, R. G. and West, S. C. (1991). Resolution of Holliday junctions in

vitro requires the Escherichia coli ruvC gene product. *Proc Natl Acad Sci U S A* **88**, 6063-7.

Cox, M. M. (2007). Motoring along with the bacterial RecA protein. *Nat Rev Mol Cell Biol* **8**, 127-38.

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**, 1163-74.

Cromie, G. A., Hyppa, R. W., Taylor, A. F., Zakharyevich, K., Hunter, N. and Smith, G. R. (2006). Single Holliday junctions are intermediates of meiotic recombination. *Cell* **127**, 1167-78.

Cromie, G. A., Millar, C. B., Schmidt, K. H. and Leach, D. R. (2000). Palindromes as substrates for multiple pathways of recombination in Escherichia coli. *Genetics* **154**, 513-22.

Davis, L. and Smith, G. R. (2001). Meiotic recombination and chromosome segregation in Schizosaccharomyces pombe. *Proc Natl Acad Sci U S A* **98**, 8395-402.

Dayani, Y., Simchen, G. and Lichten, M. (2011). Meiotic recombination intermediates are resolved with minimal crossover formation during return-to-growth, an analogue of the mitotic cell cycle. *PLoS Genet* **7**, e1002083.

de Boer, J. and Hoeijmakers, J. H. (2000). Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**, 453-60.

de Massy, B., Studier, F. W., Dorgai, L., Appelbaum, E. and Weisberg, R. A. (1984). Enzymes and sites of genetic recombination: studies with gene-3 endonuclease of phage T7 and with site-affinity mutants of phage lambda. *Cold Spring Harb Symp Quant Biol* **49**, 715-26.

Deans, A. J. and West, S. C. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**, 467-80.

Demple, B. and Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annu Rev Biochem* **63**, 915-48.

Dickman, M. J., Ingleston, S. M., Sedelnikova, S. E., Rafferty, J. B., Lloyd, R. G., Grasby, J. A. and Hornby, D. P. (2002). The RuvABC resolvasome. *Eur J Biochem* **269**, 5492-501.

Dillingham, M. S., Spies, M. and Kowalczykowski, S. C. (2003). RecBCD enzyme is a bipolar DNA helicase. *Nature* **423**, 893-7.

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**, 87-96.

Doksani, Y., Bermejo, R., Fiorani, S., Haber, J. E. and Foiani, M. (2009). Replicon dynamics, dormant origin firing, and terminal fork integrity after double-strand break formation. *Cell* **137**, 247-58.

Duckett, D. R., Murchie, A. I., Diekmann, S., von Kitzing, E., Kemper, B. and Lilley, D. M. (1988). The structure of the Holliday junction, and its resolution. *Cell* **55**, 79-89.

Dunderdale, H. J., Benson, F. E., Parsons, C. A., Sharples, G. J., Lloyd, R. G. and West, S. C. (1991). Formation and resolution of recombination intermediates by E. coli RecA and RuvC proteins. *Nature* **354**, 506-10.

Dutra, B. E., Sutera, V. A., Jr. and Lovett, S. T. (2007). RecA-independent recombination is efficient but limited by exonucleases. *Proc Natl Acad Sci U S A* **104**, 216-21.

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**, 677-97.

Esposito, M. S. (1978). Evidence that spontaneous mitotic recombination occurs at the two-strand stage. *Proc Natl Acad Sci U S A* **75**, 4436-40.

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. *Mol Cell* **29**, 644-51.

Farmer, P. B., Singh, R., Kaur, B., Sram, R. J., Binkova, B., Kalina, I., Popov, T. A., Garte, S., Taioli, E., Gabelova, A. et al. (2003). Molecular epidemiology studies of carcinogenic environmental pollutants. Effects of polycyclic aromatic hydrocarbons (PAHs) in environmental pollution on exogenous and oxidative DNA damage. *Mutat Res* **544**, 397-402.

Ferguson, D. O. and Holloman, W. K. (1996). Recombinational repair of gaps in DNA is asymmetric in Ustilago maydis and can be explained by a migrating D-loop model. *Proc Natl Acad Sci U S A* **93**, 5419-24.

Ferullo, D. J., Cooper, D. L., Moore, H. R. and Lovett, S. T. (2009). Cell cycle synchronization of Escherichia coli using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods* **48**, 8-13.

Formosa, T. and Alberts, B. M. (1986). DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* **47**, 793-806.

Fricke, W. M. and Brill, S. J. (2003). Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. *Genes Dev* **17**, 1768-78.

Friedberg, E. C. (2003). DNA damage and repair. *Nature* **421**, 436-40.

Friedman, K. L. and Brewer, B. J. (1995). Analysis of replication intermediates by two-dimensional agarose gel electrophoresis. *Methods Enzymol* **262**, 613-27.

Fukuoh, A., Iwasaki, H., Ishioka, K. and Shinagawa, H. (1997). ATP-dependent resolution of R-loops at the ColE1 replication origin by Escherichia coli RecG protein, a Holliday junction-specific helicase. *EMBO J* **16**, 203-9.

Gabbai, C. B. and Marians, K. J. (2010). Recruitment to stalled replication forks of the PriA DNA helicase and replisome-loading activities is essential for survival. *DNA Repair (Amst)* **9**, 202-9.

Ganesan, S. and Smith, G. R. (1993). Strand-specific binding to duplex DNA ends by the subunits of the Escherichia coli RecBCD enzyme. *J Mol Biol* **229**, 67-78.

Gaskell, L. J., Osman, F., Gilbert, R. J. and Whitby, M. C. (2007). Mus81 cleavage of Holliday junctions: a failsafe for processing meiotic recombination intermediates? *EMBO J* **26**, 1891-901.

Gentry, D. R. and Burgess, R. R. (1989). rpoZ, encoding the omega subunit of Escherichia coli RNA polymerase, is in the same operon as spoT. *J Bacteriol* **171**, 1271-7.

Geuting, V., Kobbe, D., Hartung, F., Durr, J., Focke, M. and Puchta, H. (2009). Two distinct MUS81-EME1 complexes from Arabidopsis process Holliday junctions. *Plant Physiol* **150**, 1062-71.

Gorbalenya, A. E. and Koonin, E. V. (1993). Helicase: amino acid sequence comparisons and structure-function relationships *Curr Opin Struct Biol* **3**, 419-429.

Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S. and Glickman, M. S. (2011). Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Mol Microbiol* **79**, 316-30.

Hastings, P. J. (1988). Recombination in the eukaryotic nucleus. *Bioessays* **9**, 61-4.

Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genetical Research* **5**, 282-304.

Holliday, R. (1966). Studies on mitotic gene conversion in Ustilago. *Genet Res* **8**, 323-37.

Holliday, R. (1974). Molecular aspects of genetic exchange and gene conversion. *Genetics* **78**, 273-87.

Hong, X., Cadwell, G. W. and Kogoma, T. (1995). Escherichia coli RecG and RecA proteins in R-loop formation. *EMBO J* **14**, 2385-92.

Hsu, P. L. and Landy, A. (1984). Resolution of synthetic att-site Holliday structures by the integrase protein of bacteriophage lambda. *Nature* **311**, 721-6.

Huberman, J. A., Spotila, L. D., Nawotka, K. A., el-Assouli, S. M. and Davis, L. R. (1987). The in vivo replication origin of the yeast 2 microns plasmid. *Cell* **51**, 473-81.

Ip, S. C., Rass, U., Blanco, M. G., Flynn, H. R., Skehel, J. M. and West, S. C. (2008). Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**, 357-61.

Ivancic-Bace, I., Vlasic, I., Salaj-Smic, E. and Brcic-Kostic, K. (2006). Genetic evidence for the requirement of RecA loading activity in SOS induction after UV irradiation in Escherichia coli. *J Bacteriol* **188**, 5024-32.

Jessop, L. and Lichten, M. (2008). Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis. *Mol Cell* **31**, 313-23.

Kalman, M., Murphy, H. and Cashel, M. (1992). The nucleotide sequence of recG, the distal spo operon gene in Escherichia coli K-12. *Gene* **110**, 95-9.

Karow, J. K., Constantinou, A., Li, J. L., West, S. C. and Hickson, I. D. (2000). The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A* **97**, 6504-8.

Kemper, B., Jensch, F., von Depka-Prondzynski, M., Fritz, H. J., Borgmeyer, U. and Mizuuchi, K. (1984). Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA. *Cold Spring Harb Symp Quant Biol* **49**, 815-25.

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. *Microbiology* **147**, 3241-7.

Klapstein, K., Chou, T. and Bruinsma, R. (2004). Physics of RecA-mediated homologous recognition. *Biophys J* **87**, 1466-77.

Kobayashi, H., Simmons, L. A., Yuan, D. S., Broughton, W. J. and Walker, G. C. (2008). Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of Sinorhizobium meliloti. *Mol Microbiol* **67**, 350-63.

Kobayashi, I. and Ikeda, H. (1983). Double Holliday structure: a possible in vivo intermediate form of general recombination in Escherichia coli. *Mol Gen Genet* **191**, 213-20.

Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev* **61**, 212-38.

Kogoma, T., Cadwell, G. W., Barnard, K. G. and Asai, T. (1996). The DNA replication priming protein, PriA, is required for homologous recombination and double-strand break repair. *J Bacteriol* **178**, 1258-64.

Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. and Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635-47.

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**, 401-65.

Kreuzer, K. N., Saunders, M., Weislo, L. J. and Kreuzer, H. W. (1995). Recombination-dependent DNA replication stimulated by double-strand breaks in bacteriophage T4. *J Bacteriol* **177**, 6844-53.

Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A* **98**, 8241-6.

Lam, S. T., Stahl, M. M., McMilin, K. D. and Stahl, F. W. (1974). Rec-mediated recombinational hot spot activity in bacteriophage lambda. II. A mutation which causes hot spot activity. *Genetics* **77**, 425-33.

Leach, D. R. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **16**, 893-900.

Lennox, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**, 190-206.

Lewis, S. M. and Cote, A. G. (2006). Palindromes and genomic stress fractures: bracing and repairing the damage. *DNA Repair (Amst)* **5**, 1146-60.

Lilley, D. M. (1981). In vivo consequences of plasmid topology. *Nature* **292**, 380-2.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709-15.

Link, A. J., Phillips, 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**, 6228-37.

Lisby, M., Rothstein, R. and Mortensen, U. H. (2001). Rad52 forms DNA repair and recombination centers during S phase. *Proc Natl Acad Sci U S A* **98**, 8276-82.

Liu, J., Nurse, P. and Marians, K. J. (1996). The ordered assembly of the phiX174-type primosome. III. PriB facilitates complex formation between PriA and DnaT. *J Biol Chem* **271**, 15656-61.

Liu, J., Wu, T. C. and Lichten, M. (1995). The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. *EMBO J* **14**, 4599-608.

Lloyd, R. G. (1991). Conjugational recombination in resolvase-deficient ruvC mutants of Escherichia coli K-12 depends on recG. *J Bacteriol* **173**, 5414-8.

Lloyd, R. G., Benson, F. E. and Shurvinton, C. E. (1984). Effect of ruv mutations on recombination and DNA repair in Escherichia coli K12. *Mol Gen Genet* **194**, 303-9.

Lloyd, R. G. and Buckman, C. (1991). Genetic analysis of the recG locus of Escherichia coli K-12 and of its role in recombination and DNA repair. *J Bacteriol* **173**, 1004-11.

Lloyd, R. G. and Sharples, G. J. (1991). Molecular organization and nucleotide sequence of the recG locus of Escherichia coli K-12. *J Bacteriol* **173**, 6837-43.

Lloyd, R. G. and Sharples, G. J. (1993a). Dissociation of synthetic Holliday junctions by E. coli RecG protein. *EMBO J* **12**, 17-22.

Lloyd, R. G. and Sharples, G. J. (1993b). Processing of recombination intermediates by the RecG and RuvAB proteins of Escherichia coli. *Nucleic Acids Res* **21**, 1719-25.

Lobachev, K. S., Gordenin, D. A. and Resnick, M. A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183-93.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C.A., Krieger, M., Scott, M.P., Zipursky, S.L., Darnell, J. . (2004). Molecular Biology of the cell: WH Freeman: New York, NY.

Long, D. T. and Kreuzer, K. N. (2008). Regression supports two mechanisms of fork processing in phage T4. *Proc Natl Acad Sci U S A* **105**, 6852-7.

Lusetti, S. L. and Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem* **71**, 71-100.

Lutz, R. and Bujard, H. (1997). Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**, 1203-10.

Mandal, T. N., Mahdi, A. A., Sharples, G. J. and Lloyd, R. G. (1993). Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of ruvA, ruvB, and ruvC mutations. *J Bacteriol* **175**, 4325-34.

Marians, K. J. (2000). PriA-directed replication fork restart in Escherichia coli. *Trends Biochem Sci* **25**, 185-9.

Maryon, E. and Carroll, D. (1991). Characterization of recombination intermediates from DNA injected into Xenopus laevis oocytes: evidence for a

nonconservative mechanism of homologous recombination. *Mol Cell Biol* **11**, 3278-87.

McGill, C., Shafer, B. and Strathern, J. (1989). Coconversion of flanking sequences with homothallic switching. *Cell* **57**, 459-67.

McGlynn, P., Al-Deib, A. A., Liu, J., Marians, K. J. and Lloyd, R. G. (1997). The DNA replication protein PriA and the recombination protein RecG bind D-loops. *J Mol Biol* **270**, 212-21.

McGlynn, P. and Lloyd, R. G. (1999). RecG helicase activity at three- and four-strand DNA structures. *Nucleic Acids Res* **27**, 3049-56.

McKinnon, P. J. and Caldecott, K. W. (2007). DNA strand break repair and human genetic disease. *Annu Rev Genomics Hum Genet* **8**, 37-55.

Meddows, T. R., Savory, A. P. and Lloyd, R. G. (2004). RecG helicase promotes DNA double-strand break repair. *Mol Microbiol* **52**, 119-32.

Merlin, C., McAteer, S. and Masters, M. (2002). Tools for characterization of Escherichia coli genes of unknown function. *J Bacteriol* **184**, 4573-81.

Meselson, M. S. and Radding, C. M. (1975). A general model for genetic recombination. *Proc Natl Acad Sci U S A* **72**, 358-61.

Michel, B., Ehrlich, S. D. and Uzest, M. (1997). DNA double-strand breaks caused by replication arrest. *EMBO J* **16**, 430-8.

Mosig, G. (1987). The essential role of recombination in phage T4 growth. *Annu Rev Genet* **21**, 347-71.

Mueller, J. E., Clyman, J., Huang, Y. J., Parker, M. M. and Belfort, M. (1996). Intron mobility in phage T4 occurs in the context of recombination-dependent DNA replication by way of multiple pathways. *Genes Dev* **10**, 351-64.

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**, 1613-25.

Ng, J. Y. and Marians, K. J. (1996a). The ordered assembly of the phiX174-type primosome. I. Isolation and identification of intermediate protein-DNA complexes. *J Biol Chem* **271**, 15642-8.

Ng, J. Y. and Marians, K. J. (1996b). The ordered assembly of the phiX174-type primosome. II. Preservation of primosome composition from assembly through replication. *J Biol Chem* **271**, 15649-55.

Nishinaka, T., Ito, Y., Yokoyama, S. and Shibata, T. (1997). An extended DNA structure through deoxyribose-base stacking induced by RecA protein. *Proc Natl Acad Sci U S A* **94**, 6623-8.

Norbury, C. J. and Hickson, I. D. (2001). Cellular responses to DNA damage. *Annu Rev Pharmacol Toxicol* **41**, 367-401.

Oh, S. D., Lao, J. P., Taylor, A. F., Smith, G. R. and Hunter, N. (2008). RecQ helicase, Sgs1, and XPF family endonuclease, Mus81-Mms4, resolve aberrant joint molecules during meiotic recombination. *Mol Cell* **31**, 324-36.

Ohsato, T., Muta, T., Fukuoh, A., Shinagawa, H., Hamasaki, N. and Kang, D. (1999). R-Loop in the replication origin of human mitochondrial DNA is resolved by RecG, a Holliday junction-specific helicase. *Biochem Biophys Res Commun* **255**, 1-5.

Oppenheim, A. (1981). Separation of closed circular DNA from linear DNA by electrophoresis in two dimensions in agarose gels. *Nucleic Acids Res* **9**, 6805-12.

Osman, F. and Subramani, S. (1998). Double-strand break-induced recombination in eukaryotes. *Prog Nucleic Acid Res Mol Biol* **58**, 263-99.

Otsuji, N., Iyehara, H. and Hideshima, Y. (1974). Isolation and characterization of an Escherichia coli ruv mutant which forms nonseptate filaments after low doses of ultraviolet light irradiation. *J Bacteriol* **117**, 337-44.

Panyutin, I. G. and Hsieh, P. (1994). The kinetics of spontaneous DNA branch migration. *Proc Natl Acad Sci U S A* **91**, 2021-5.

Paques, F. and Haber, J. E. (1999). Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **63**, 349-404.

Paques, F., Leung, W. Y. and Haber, J. E. (1998). Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol Cell Biol* **18**, 2045-54.

Pegg, A. E. (2000). Repair of O(6)-alkylguanine by alkyltransferases. *Mutat Res* **462**, 83-100.

Pegg, A. E. and Byers, T. L. (1992). Repair of DNA containing O6-alkylguanine. *FASEB J* **6**, 2302-10.

Pellman, D. (2007). Cell biology: aneuploidy and cancer. *Nature* **446**, 38-9.

Pommier, Y., Redon, C., Rao, V. A., Seiler, J. A., Sordet, O., Takemura, H., Antony, S., Meng, L., Liao, Z., Kohlhagen, G. et al. (2003). Repair of and checkpoint response to topoisomerase I-mediated DNA damage. *Mutat Res* **532**, 173-203.

Ponticelli, A. S., Schultz, D. W., Taylor, A. F. and Smith, G. R. (1985). Chi-dependent DNA strand cleavage by RecBC enzyme. *Cell* **41**, 145-51.

Pritchard, R. H. (1955). The linear arrangement of a series of alleles of *Aspergillus nidulans*. *Heredity* **9**, 343-371.

Register, J. C., 3rd 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**, 12308-12.

Resnick, M. A. and Martin, P. (1976). The repair of double-strand breaks in the nuclear DNA of Saccharomyces cerevisiae and its genetic control. *Mol Gen Genet* **143**, 119-29.

Roe, S. M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I. R. and Pearl, L. H. (1998). Crystal structure of an octameric RuvA-Holliday junction complex. *Mol Cell* **2**, 361-72.

Rothstein, R., Michel, B. and Gangloff, S. (2000). Replication fork pausing and recombination or "gimme a break". *Genes Dev* **14**, 1-10.

Rudolph, C. J., Upton, A. L., Briggs, G. S. and Lloyd, R. G. (2010). Is RecG a general guardian of the bacterial genome? *DNA Repair (Amst)* **9**, 210-23.

Rudolph, C. J., Upton, A. L., Harris, L. and Lloyd, R. G. (2009a). Pathological replication in cells lacking RecG DNA translocase. *Mol Microbiol* **73**, 352-66.

Rudolph, C. J., Upton, A. L. and Lloyd, R. G. (2009b). Replication fork collisions cause pathological chromosomal amplification in cells lacking RecG DNA translocase. *Mol Microbiol* **74**, 940-55.

Sagi, D., Tlusty, T. and Stavans, J. (2006). High fidelity of RecA-catalyzed recombination: a watchdog of genetic diversity. *Nucleic Acids Res* **34**, 5021-31.

Sancar, A., Stachelek, C., Konigsberg, W. and Rupp, W. D. (1980). Sequences of the recA gene and protein. *Proc Natl Acad Sci U S A* **77**, 2611-5.

Sandler, S. J. (1996). Overlapping functions for recF and priA in cell viability and UV-inducible SOS expression are distinguished by dnaC809 in Escherichia coli K-12. *Mol Microbiol* **19**, 871-80.

Sandler, S. J. and Marians, K. J. (2000). Role of PriA in replication fork reactivation in Escherichia coli. *J Bacteriol* **182**, 9-13.

Sandler, S. J., Marians, K. J., Zavitz, K. H., Coutu, J., Parent, M. A. and Clark, A. J. (1999). dnaC mutations suppress defects in DNA replication- and recombination-associated functions in priB and priC double mutants in Escherichia coli K-12. *Mol Microbiol* **34**, 91-101.

Sandler, S. J., Samra, H. S. and Clark, A. J. (1996). Differential suppression of priA2::kan phenotypes in Escherichia coli K-12 by mutations in priA, lexA, and dnaC. *Genetics* **143**, 5-13.

Sasaki, K., Ose, T., Okamoto, N., Maenaka, K., Tanaka, T., Masai, H., Saito, M., Shirai, T. and Kohda, D. (2007). Structural basis of the 3'-end recognition of a leading strand in stalled replication forks by PriA. *EMBO J* **26**, 2584-93.

Sattin, B. D. and Goh, M. C. (2004). Direct observation of the assembly of RecA/DNA complexes by atomic force microscopy. *Biophys J* **87**, 3430-6.

Schwacha, A. and Kleckner, N. (1994). Identification of joint molecules that form frequently between homologs but rarely between sister chromatids during yeast meiosis. *Cell* **76**, 51-63.

Schwacha, A. and Kleckner, N. (1995). Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* **83**, 783-91.

Seigneur, M., Bidnenko, V., Ehrlich, S. D. and Michel, B. (1998). RuvAB acts at arrested replication forks. *Cell* **95**, 419-30.

Sharples, G. J., Benson, F. E., Illing, G. T. and Lloyd, R. G. (1990). Molecular and functional analysis of the ruv region of Escherichia coli K-12 reveals three genes involved in DNA repair and recombination. *Mol Gen Genet* **221**, 219-26.

Sharples, G. J., Chan, S. N., Mahdi, A. A., Whitby, M. C. and Lloyd, R. G. (1994). Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. *EMBO J* **13**, 6133-42.

Sharples, G. J. and Lloyd, R. G. (1991). Resolution of Holliday junctions in Escherichia coli: identification of the ruvC gene product as a 19-kilodalton protein. *J Bacteriol* **173**, 7711-5.

Shiba, T., Iwasaki, H., Nakata, A. and Shinagawa, H. (1991). SOS-inducible DNA repair proteins, RuvA and RuvB, of Escherichia coli: functional interactions between RuvA and RuvB for ATP hydrolysis and renaturation of the cruciform structure in supercoiled DNA. *Proc Natl Acad Sci U S A* **88**, 8445-9.

Silberman, R. and Kupiec, M. (1994). Plasmid-mediated induction of recombination in yeast. *Genetics* **137**, 41-8.

Singleton, M. R., Dillingham, M. S., Gaudier, M., Kowalczykowski, S. C. and Wigley, D. B. (2004). Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**, 187-93.

Smith, G. R., Kunes, S. M., Schultz, D. W., Taylor, A. and Triman, K. L. (1981). Structure of chi hotspots of generalized recombination. *Cell* **24**, 429-36.

Spies, M., Amitani, I., Baskin, R. J. and Kowalczykowski, S. C. (2007). RecBCD enzyme switches lead motor subunits in response to chi recognition. *Cell* **131**, 694-705.

Stephanou, N. C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S. and Glickman, M. S. (2007). Mycobacterial nonhomologous end joining mediates mutagenic repair of chromosomal double-strand DNA breaks. *J Bacteriol* **189**, 5237-46.

Storm, P. K., Hoekstra, W. P., de Haan, P. G. and Verhoef, C. (1971). Genetic recombination in Escherichia coli. IV. Isolation and characterization of recombination-deficiency mutants of Escherichia coli K12. *Mutat Res* **13**, 9-17.

Strathern, J. N., Klar, A. J., Hicks, J. B., Abraham, J. A., Ivy, J. M., Nasmyth, K. A. and McGill, C. (1982). Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. *Cell* **31**, 183-92.

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.

Takata, M., Sasaki, M. S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A. and Takeda, S. (1998). Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* **17**, 5497-508.

Tanaka, T. and Masai, H. (2006). Stabilization of a stalled replication fork by concerted actions of two helicases. *J Biol Chem* **281**, 3484-93.

Tanaka, T., Mizukoshi, T., Sasaki, K., Kohda, D. and Masai, H. (2007). Escherichia coli PriA protein, two modes of DNA binding and activation of ATP hydrolysis. *J Biol Chem* **282**, 19917-27.

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**, 115-29.

Taylor, A. and Smith, G. R. (1980). Unwinding and rewinding of DNA by the RecBC enzyme. *Cell* **22**, 447-57.

Taylor, A. F., Schultz, D. W., Ponticelli, A. S. and Smith, G. R. (1985). RecBC enzyme nicking at Chi sites during DNA unwinding: location and orientation-dependence of the cutting. *Cell* **41**, 153-63.

Taylor, A. F. and Smith, G. R. (2003). RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* **423**, 889-93.

Tosa, T. and Pizer, L. I. (1971). Effect of serine hydroxamate on the growth of Escherichia coli. *J Bacteriol* **106**, 966-71.

Tracy, R. B. and Kowalczykowski, S. C. (1996). In vitro selection of preferred DNA pairing sequences by the Escherichia coli RecA protein. *Genes Dev* **10**, 1890-903.

Truglio, J. J., Croteau, D. L., Van Houten, B. and Kisker, C. (2006). Prokaryotic nucleotide excision repair: the UvrABC system. *Chem Rev* **106**, 233-52.

van Gent, D. C., Hoeijmakers, J. H. and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. *Nat Rev Genet* **2**, 196-206.

Veaute, X., Delmas, S., 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**, 180-9.

Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. and Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75-84.

Vilenchik, M. M. and Knudson, A. G. (2003). Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proc Natl Acad Sci U S A* **100**, 12871-6.

Voelkel-Meiman, K. and Roeder, G. S. (1990). Gene conversion tracts stimulated by HOT1-promoted transcription are long and continuous. *Genetics* **126**, 851-67.

Wallace, S. S. (1998). Enzymatic processing of radiation-induced free radical damage in DNA. *Radiat Res* **150**, S60-79.

Wardrope, L., Okely, E. and Leach, D. (2009). Resolution of joint molecules by RuvABC and RecG following cleavage of the Escherichia coli chromosome by EcoKI. *PLoS ONE* **4**, e6542.

West, S. C. (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu Rev Genet* **31**, 213-44.

Weterings, E. and Chen, D. J. (2008). The endless tale of non-homologous end-joining. *Cell Res* **18**, 114-24.

Whitby, M. C. and Lloyd, R. G. (1995). Branch migration of three-strand recombination intermediates by RecG, a possible pathway for securing exchanges initiated by 3'-tailed duplex DNA. *EMBO J* **14**, 3302-10.

Whitby, M. C., Ryder, L. and Lloyd, R. G. (1993). Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* **75**, 341-50.

White, M. A., Eykelenboom, J. K., Lopez-Vernaza, M. A., Wilson, E. and Leach, D. R. (2008). Non-random segregation of sister chromosomes in Escherichia coli. *Nature* **455**, 1248-50.

Willetts, N. S. and Clark, A. J. (1969). Characteristics of some multiply recombination-deficient strains of Escherichia coli. *J Bacteriol* **100**, 231-9.

Xu, L. and Marians, K. J. (2000). Purification and characterization of DnaC810, a primosomal protein capable of bypassing PriA function. *J Biol Chem* **275**, 8196-205.

Yancey-Wrona, J. E. and Camerini-Otero, R. D. (1995). The search for DNA homology does not limit stable homologous pairing promoted by RecA protein. *Curr Biol* **5**, 1149-58.

Yu, M., Souaya, J. and Julin, D. A. (1998a). The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from Escherichia coli. *Proc Natl Acad Sci U S A* **95**, 981-6.

Yu, M., Souaya, J. and Julin, D. A. (1998b). Identification of the nuclease active site in the multifunctional RecBCD enzyme by creation of a chimeric enzyme. *J Mol Biol* **283**, 797-808.

Yu, X. and Egelman, E. H. (1990). Image analysis reveals that Escherichia coli RecA protein consists of two domains. *Biophys J* **57**, 555-66.

Zou, H. and Rothstein, R. (1997). Holliday junctions accumulate in replication mutants via a RecA homolog-independent mechanism. *Cell* **90**, 87-96.

Appendix

Sequences shown in bold **black** and underlined, represent single, endogenous χ sites in the correct orientation for processing a DSB at the palindrome.

The location of *lacZ*χ- (See Table 2.) is marked by the bold grey

The 44 bp sequences shown in bold pale blue are the χ arrays. The underlined portions of these sequences are the 8 pb χ sites.

The 246 bp interrupted palindrome is shown in bold green.

Regions of the chromosome made up of five or more repeats of either cytosine or guanine are highlighted.

lacZ.proximal + 1 (% GC = 50)

caattgccggtattgtgggcgatggtgtgtgggttaggtctgggcggcgcgtttccgctctgtttgct

gctggcgctcgatcactctgtgcaaccggctattgctggcaa**gctggtgg**cgtttatgcagggaatcg

gttttatcatcgccgggcttgccccgtggttttctggcgtgctgcgtagtatcagcggcaattacctg

atggactgggcatttcatgcgctgtgcgtcgttgggctgatgatcataaccctgcgttttgcaccagt

acgttttccgcagctgtgggtcaaagaggcatgatgcgacgcttgttcctgcgctttgttcatgccgg

atgcggctaatgtagatcgctgaacttgtaggcctgataagcgcagcgtatcaggcaatttttataat

ttaaactgacgattcaactttataatctttgaaataatagtgcttatcccggtcgtttatttcgcgaa

taacccgacaaggaacgccagccgccacgacgtttggtggaatgtcttttgtgacgatactacccgcg

ccaataacagaattatccccgatggtgacgcctggattaataaccacatgacttccgatccagacgtt

attgccaatcgttatcggaaaagagtacatctcgccgttttttctcaattcatggtgtacagggtgtc

ccgtaacggaaagagtaacgttgggtgcaatcagtacgttatcaccgattgttaccgtgtagtcatcg

acaatggttaaattgaaatttgcataaaaattgcggcctatatggatgttggaaccgtaagagaaata

gacaggcggttctacccaggcgttttcccctaccgtggcaaacatttctttaatcaggctttctcttt

tttcaacttctgatggatgcgagtgattaaactcatacattaacgttttcccacgaagtctttttttcc

ggtaagccttcgcacatatcggtaaatagcttgcctgctcttattctttcggtcattggcatgtcatt

cacctgacgacgcagcagggaaagcgggccggggccgctaagcgtgaacacggaaattaaggtgaagc

ccagcgccaccagacccagcaccagataagcgcccctggaaaccgatgctttcatacatattgcccgcc

agtacagacataaaaatcatcgccagttgcttaaagaagcagaaacagaccagataaatcgtcgctga

aaaacgcacttcaaactggctggtaatatatttaaagcagcccaccagcaggaacggtacttcaaaca

tatgcagcgttttcagaataaccacttccagcgctgaggtggcgaacgatgagccaataatacgtaca

gacataatagtgccagccagcagcagggcgttttcccaccgatgcgattaatgatcagtggcgcaaa

gaacataatcgaggcgttaagtaattcgcccattgtcgttacgtagccaaatacccgcgtaccctgtt

caccggtagcaaagaacgaagtaaagaaattagcaaactgttggtcaaaaacatcgtaggtgcaggaa

acgccaataacatacagtgacaaaaaccacagttttggctgtctgaacagttccagtgccagcttaag

gctaaatgccgaatggttggcacctaccgcattggcaaccgtggcagaagagggcgcatccgttttgg

cgaaaaagagtaaaacggcgaggatgagtgcacagccagagcccagccagaaaacaaactgattattg

atggtgaacatgatgccgacaatcgaggcacacagcgcccagccaacacagccaaacatccgcgcgcg

accaaattcgaaattactgcgacggctgactttctcaataaatgcctctactgctggcgcaccggcgt

taaaacaaaagcctagataaataccaccaacaatcgatcctactaaaatgttgtattgtaacagtggc

ccgaagataaaaataaagaacggcgcaaacatcactaacatgccggtaataatccacagcaggtattt

gcgcagcccgagtttgtcagaaagcagaccaaacagcggttggaataatagcgagaacagagaaatag

cggcaaaaataatacccgtatcacttttgctgatatggttgatgtcatgtagccaaatcgggaaaaac

gggaagtaggctcccatgataaaaaagtaaaagaaaaagaataaaccgaacatccaaaagtttgtgtt

ttttaaatagtacataatggatttccttacgcgaaatacgggcagacatggcctgcccggttattatt

atttttgacaccagaccaactggtaatggtagcgaccggcgctcagctggaattccgccgatactgac

gggctccaggagtcgtcgccaccaatccccatatggaaaccgtcgatattcagccatgtgccttcttc

cgcgtgcagcagatggcgatggctggtttccatcagttgctgttgactgtagcggctgatgttgaact

ggaagtcgccgcgccactggtgtgggccataattcaattcgcgcgtcccgcagcgcagaccgttttcg

ctcgggaagacgtacggggtatacatgtctgacaatggcagatcccagcggtcaaaacaggcggcagt

aaggcggtcgggatagttttcttgcggcccctaatccgagccagtttacccgctctgctacctgcgcca

gctggcagttcaggccaatccgcgccggatgcggtgtatcgctcgccacttcaacatcaacggtaatc

gccatttgaccactaccatcaatccggtaggttttccggctgataaataaggttttcccctgatgctg

ccacgcgtgagcggtcgtaatcagcaccgcatcagcaagtgtatctgccgtgcactgcaacaacgctg

cttcggcctggtaatggcccgccgccttccagcgttcgacccaggcgttagggtcaatgcgggtcgct

tcacttacgccaatgtcgttatccagcggtgcacgggtgaactgatcgcgcagcggcgtcagcagttg

ttttttatcgccaatccacatctgtgaaagaaagcctgactggcggttaaattgccaacgcttattac

ccagctcgatgcaaaaatccatttcactgatggtcagatgcgggatggcgtgggacgcggcggggagc

gtcacactgaggttttccgccagacgccactgctgccaggcgctgatgtgcccggcttctgaccatgc

ggtcgcgttcggttgcactacgcgtactgtgagccagagttgcccggcgctctccggctgcggtagtt

caggcagttcaatcaactgtttaccttgtggagcgacatccagaggcacttcaccgcttgccagcggc

ttaccatccagcgccaccatccagtgcaggagctc

lacZ.distal (% GC = 55)

gagctcgttatcgctatgacggaacaggtattcgctggtcacttcgatggtttgcccggataaacgga
actggaaaaactgctgctggtgttttgcttccgtcagcgctggatgcggcgtgcggtcggcaaagacc
agaccgttcatacagaactggcgatcgttcggcgtatcgccaaaatcaccgccgtaagccgaccacgg
gttgccgttttcatcatatttaatcagcgactgatccacccagtcccagacgaagccgccctgtaaac
ggggatactgacgaaacgcctgccagtatttagcgaaaccgccaagactgttacccatcgcgtgggcg
tattcgcaaaggatcagcgggcgcgtctctccaggtagcgaaagccatttttgatggaccatttcgg
cacagccgggaagggctggtcttcatccacgcgcgcgtacatcggggcaaataatatcggtggccgtgg
tgtcggctccgccgccttcatactgcaccgggcgggaaggatcgacagatttgatccagcgatacagc
gcgtcgtgaGCTGGTGGTCGATGCTGAGCTGGTGGACACGCGCTGGCTGGTGGttagcgccgtggcct
gattcattccccagcgaccagatgatcacactcgggtgattacgatcgcgctgcaccattcgcgttac
gcgttcgctcatcgccggtagccagcgcggatcatcggtcagacgattcattggcaccatgccgtggg
tttcaatattggcttcatccaccacatacaggccgtagcggtcgcacagcgtgtaccacagcggatgg
ttcggataatgcgaacagcgcacggcgttaaagttgttctgcttcatcagcaggatatcctgcaccat
cgtctgctcatccatgacctgaccatgcagaggatgatgctcgtgacggttaacgcctcgaatcagca
acggcttgccgttcagcagcagcagaccattttcaatccgcacctcgcggaaaccgacatcgcaggct
tctgcttcaatcagcgtgccgtcggcggtgtgcagttcaaccaccgcacgatagagattcgggatttc
ggcgctccacagtttcgggttttcgacgttcagacgtagtgtgacgcgatcggcataaccaccacgct
catcgataatttcaccgccgaaaggcgcggtgccgctggcgacctgcgtttcaccctgccataaagaa
actgttacccgtaggtagtcacgcaactcgccgcacatctgaacttcagcctccagtacagcgcggct
gaaatcatcattaaagcgagtggcaacatggaaatcgctgatttgtgtagtcggtttatgcagcaacg
agacgtcacggaaaatgccgctcatccgccacatatcctgatcttccagataactgccgtcactccag
cgcagcaccatcaccgcgaggcggttttctccggcgcgtaaaaatgcgctcaggtcaaattcagacgg
caaacgactgtcctggccgtaaccgacccagcgcccgttgcaccacagatgaaacgccgagttaacgc
catcaaaaataattcgcgtctggccttcctgtagccagctttcatcaacattaaatgtgagcgagtaa
caacccgtcggattctccgtgggaacaaacggcggattgaccgtaatgggataggtcacgttggtgta
gatgggcgcatcgtaaccgtgcatctgccagtttgaggggacgacgacagtatcggcctcaggaagat

cgcactccagccagctttccggcaccgcttctggtgccggaaaccaggcaaagcgccattcgccattc

aggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagg

gggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacga

cg**GAATTCATTTCAGCATTTATTGGTTGTATGAGAGTAGATAGAAAAAGACAACTCTGGCTTGAAGCT**

ATCAAAAAACTAAGTAGTGATGAAAACTTTTCAAATATGGAACTCTTCATATTTAGAAATGAGGCTGA

TGAGTTCCATATTTGAAAAGTTTTCATCACTACTTAGTTTTTTGATAGCTTCAAGCCAGAGTTGTCTT

TTTCTATCTACTCTCATACAACCAATAAATGCTGAAATGAATTCcatggtcatagctgtttcctgtgt

gaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggt

gcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacct

gtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgccagg

gtggtttttctttcaccagtgagacgggcaacagctgattgcccttcaccgcctggccctgagagag

ttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtggttaacggcg

ggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagc

ccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtggg

aacgatgccctcattcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttccc

gttccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgcc

gagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccac

gcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaa

gaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatccagcggatag

ttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgcc

gcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcga

caatttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgccc

gccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccacttttcccg

cgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacaccggcat

actctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttccgggcgc

tatcatgccataccgcgaaaggttttgcgccattcgatggtgtcaacgtaaatgcatgccgcttcgcc

ttccggccaccagaatagcctgcgattcaaccccttcttcgatctgttttgctacccgttg**CCACCAG**

CCGCCATGTGA**CCACCAGC**GAGTCTGCGC**CCACCAGC**tagcgccggaagatgcttttccgctgcctgt

tcaatggtcattgcgctcgccatatacaccagattcagacagccaatcacccgttgttcactgcgcag

cggtacggcgatagaggcgatcttctcctcctgatcccagccgcggtagttctgtccgtaaccctctt

tgcgcgcgcgcgccagaatggcttccagctttaacggttcccgtgccagttgatagtcatcaccgggg

cgggaggctaacatttcgattaattccttgcggtcttgttccgggcaaaaggccagccaggtcaggcc

cgaggcggttttcagaagcggcaaacgtcgcccgaccattgcccggtgaaaggataagcggctgaaac

ggtgagtggtttcgcgtaccaccattgcatcaacatccagcgtggacacatctgtcggccataccact

tcgcgcaacagatcgcccagcagtggggccgccagtgcagaaatccactgttcgtcacgaaatccttc

gcttaattgccgcactttgatggtcagtcgaaaactatcatcggaggggctacggcggacatatccct

cttcctgcagcgtctccagcagtcgccgcacagtggtgcgatgcaggccgctgagttccgccagcagc

ccgacgctggcaccgccatcaagtttatttaacatatttaataacattagaccgcgggttaagccgcg

cacggttttgtattccgtctgctcattgttctgcatattaattgacatttctatagttaaaacaacgt

ggtgcacctggtgcacattcgggcatgttttgattgtagccgaaaacacccttcctatactgagcgca

caataaaaaatcatttacatgtttttaacaaaataagttgcgctgtactgtgcgcgcaacgacatttt

gtccgagtcgtgaggtactgaaatggcaatacaacaccctgacatccagcctgctgttaaccatagcg

ttcaggtggcgatcgctggtgccggcccggttgggctgatgatggcgaactatctcggccagatgggc

attgacgtgctggtggtggagaaactcgataagttgatcgactacccgcgtgcgattggtattgatga

cgaggcgctgcgcaccatgcagtcggtcggcctggtcgatgatgttctgccgcacactacgccgtggc

acgcgatgcgttttctcaccccgaaaggccgctgttttgctgatattcagccaatgaccgatgaattt

ggctggccgcgccgtaacgcctttattcagccgcaggtcgatgcggtgatgctggaaggggtgtcgcg

ttttccgaatgtgcgctgcttgttttcccgcgagctggaggccttcagtcagcaagatgacgaagtga

ccttgcacctgaaaacggcagaagggcagcgggaaatagtcaaagcccagtggctggtagcctgtgac

ggtggagcaagttttgtccgtcgcactctgaatgtgccgtttgaaggtaaaactgcgccaaatcagtg

gattgtggtagatatcgccaacgatccgttaagtacgccgcatatctatttgtgttgcgatccggtgc

gcccgtatgtttctgccgcgctgcctcatgcggtacgtcgctttgaatttatggtgatgccgggagaa

accgaagagcagctgcgtgagccgcaaaatatgcgcaagctgttaagcaaagtgctgcctaatccgga

caatgttgaattgattcgccagcgtgtctacacccacaacgcgcgactggcgcaacgtttccgtattg

atcgcgtactgctggcgggcgatgccgcgcacatcatgccggtatggcaggggcagggctataacagt

ggtatgcgcgacgcctttaacctcgcatggaaactggcgttggttatccaggggaaagcccgcgatgc

gctgctcgatacctatcaacaagaacgtcgcgatcacgccaaagcgatgattgacctgtccgtgacgg

cgggcaacgtgctggctccgccgaaacgctggcagggtacgttacgtgacggcgtttcctggctgttg

aattatctgccgccagtaaaacgctacttcctcgaaatgcgcttcaagccgatgccgcaatattacgg

cggtgcgctgatgcgtgagggcgaagcgaagcactctccggtcggcaagatgtttattcagccgaaag

tcacgctggaaaacggcgacgtgacgctgctcgataacgcgatcggcgcgaacttcgcggtaattggc

tggggatgcaatccactgtgggggatgagcgacgagcaaatccagcagtggcgcgcgttgggcacacg

cttcattcaggtggtgccggaagtgcaaattcataccgcacaggataaccacgacggcgtactacgcg

tgggcgatacgcaaggtcgcctgcgtagctggttcgcgcaacacaatgcttcgctggtggtgatgcgc

ccggatcgctttgttgccgccaccgccattccgcaaaccctgggcaagaccctgaataaactggcgtc

ggtgatgacgctgacccgcccctgatgccgacgtttctgtcgaaaaggtagcctgatatgcacgcttat

cttcactgtctttcccactcgccgctggtggggtatgtcgacccggcgcaagaggtgctcgatgaggt

caatggcgtgattgccagcgcccgcgagcgtattgcggcattctcccctgaactggtggtgctgtttg

cgccagatcactacaacggctttttctatgacgtgatgccaccgttctgtttaggcgttggagcgacg

g**caattg**

lacZ.distal + 1 (% GC = 55)

caattggtgatttcggcagtgcggcaggagagctgcccgtgcctgtggagctggcggaggcctgtgcg

catgccgtcatgaagagcgggatcgatcttgccgtttcttactgtatgcaggtggaccacgggttcgc

ccagccgctggagttcctgctcggtgggctggataaggtgccagttctgcctgtgttcatcaacggtg

tcgccacgccgctgcccggtttccagcgtacccgcatgttgggtgaagccattggacgtttcaccagc

actctcaataaacgcgtgctgttcctgggttccggtgggctttcccatcagccgccggtgcccgaact

ggcgaaagccgatgcccatatgcgcgaccgtctgttggggagcgggaaagatttacccgccagtgagc

gcgaattgcgtcagcaacgggtgattagcgccgctgagaagtttgttgaggatcagagaacgctgcat

ccgctcaacccgatttgggataaccagttcatgactttgctggagcagggacgcatacaggaactgga

tgccgtcagtaacgaagagctttccgccattgccggaaagtcgacacatgaaatcaaaacctgggtcg

ccgctttgccgctatttctgcgtttggcaactggcgtagcgaagggcgttattaccgcccaatcccg

gagtggattgccggatttggctcgttaagcgccagaacagagaactgaatatgcaggagaagatgatg

agttatcagccacaaaccgaagccg**ccaccagc**cgttttctgaatgtagaagaagcgggtaaaacgct

gcgcatccattttaatgactgcggacaaggcgacgaaaccgttgtcctgctgcatggttccggcccgg

gtgctactggctgggcgaacttcagccgcaatatcgatccgctggtagaggcgggctatcgggtgatc

ctgctggattgtccgggttggggcaagagcgattcggtcgttaatagtggttcgcgatcggatcttaa

tgcacgaatcctgaaaagcgtggtggatcaactggatatcgccaaaatccacctgctgggcaactcga

tgggggggccatagttctgtggcgttcacccttaaatggccggagcgcgtcggcaaactggtgctgatg

ggcggcggtacgggcggcatgagtttgtttacgccgatgccaaccgaaggtattaagcgactgaatca

gctttatcgtcagccgactatcgaaaacctgaagctgatgatggatatcttcgtttttgataccagcg

atttgaccgacgcccctgtttgaagcgcgcctgaataatatgctgtcgcgccgcgatcacctggaaaac

ttcgttaagagcctggaagctaatccgaaacagttcccggattttggcccacgtctggcggaaatcaa

agcgcaaaccctgattgtctgggggcgcaacgaccgctttgtgccgatggatgcgggtctgcgtctgc

tgtccggcattgccggttctgaactgcatatcttccgcgactgtggtcactgggcgcagtgggaacat

gccgacgctttcaatcaactggtgctgaatttcctcgcacgcccttaaggaatggtcatgacgaagca

tactcttgagcaactggcggcggatttacgccgcgccgcagagcagggcgaagcgattgcaccgctgc

gcgatctgattggtatcgataacgctgaagcggcttacgccattcagcacataaatgtgcaacatgac

gttgcgcaggggcgtcgcgtggtagggcgtaaagtgggcctgacacatccgaaagtgcaacaacaact

gggcgttgatcaaccggattttgggacgttatttgccgacatgtgttatggcgataacgaaatcattc

ctttttcccgtgttctgcaaccccgcattgaagcggagatcgcactggtgttgaaccgcgatttgccc

gcaaccgatatcaccttcgacgaattgtataacgccattgaatgggtacttccggcgctggaagtggt

ggggagccgcattcgcgactggtcgattcagtttgtcgataccgtggcagataacgcctcctgtgggg

tgtatgtcatcggcggtccggcgcaacgtccggcggggttagacctgaaaaactgcgccatgaagatg

acgcgtaataacgaagaggtttctagcgggcgcggcagcgaatgcctgggacatccgcttaatgcggc

cgtctggctggcacgcaaaatggccagtctgggtgaaccgctgcgcaccggagatatcattcttaccg

gggcattaggtccgatggtggcggtgaatgcgggcgatcgttttgaagcccatattgaaggcataggt

tcagttgctgcgacattttcaagcgcagccccaaaaggaagtctgtcatgagtaagcgtaaagtcgcc

attatcggttctggcaacattggtaccgatctgatgattaaaattttgcgtcacggtcagcatctgga

gatggcggtgatggttggcattgatcctcagtccgacggtctggcgcgcgccagacgtatgggcgtcg

ccaccacccatgaaggggtgatcggactgatgaacatgcctgaatttgctgatatcgacattgtattt

gatgcgaccagcgccggtgctcatgtgaaaaacgatgccgctttacgcgaagcgaaaccggatattcg

cttaattgacctgacgcctgctgccatcggcccttactgcgtgccggtggttaacctcgaggcgaacg

tcgatcaactgaacgtcaacatggtcacctgcggcggccaggccaccattccaatggtggcggcagtt

tcacgcgtggcgcgtgttcattacgccgaaattatcgcttctatcgccagtaaatctgccggacctgg

cacgcgtgccaatatcgatgaatttacggaaaccacttcccgagccattgaagtggtgggcggcgcgg

caaaagggaaggcgattattgtgcttaacccagcagagccaccgttgatgatgcgtgacacggtgtat

gtattgagcgacgaagcttcacaagatgatatcgaagcctcaatcaatgaaatggctgaggcggtgca

ggcttacgtaccgggttatcgcctgaaacagcgcgtgcagtttgaagttatcccgcaggataaaccgg

tcaatttaccgggcgtggggcaattctccggactgaaaacagcggtctggctggaagtcgaaggcgca

gcgcattatctgcctgcctatgcgggcaacctcgacattatgacttccagtgcgctggcgacagcgga

aaaaatggcccagtcactggcgcgcaaggcaggagaagcggcatgaacggtaaaaaactttatatctc

ggacgtcacattgcgtgacggtatgcacgccattcgtcatcagtattcgctggaaaacgttcgccaga

ttgccaaagcactggacgatgcccgcgtggattcgattgaagtggcccacggcgacggtttgcaaggt

tccagctttaactatggtttcggcgcacatagcgaccttgaatggattgaagcggcggcggatgtggt

gaagcacgccaaaatcgcgacgttgttgctgccaggaatcggcactattcacgatctgaaaaatgcct

ggcaggctggcgcgcgggtggttcgtgtggcaacgcactgtaccgaagctgatgtttccgcccagcat

attcagtatgcccgc**gagctc**