Recombination at a 246 bp interrupted palindrome in *Escherichia coli*

John K. Eykelenboom

Thesis presented for the degree of Doctor of Philosophy

Institute of Cell Biology

University of Edinburgh

August 2006

Declaration

I hereby declare that this thesis was composed by me, and the research presented is my own, except where otherwise stated.

John K. Eykelenboom

August 2006

Acknowledgements

I would like to thank David Leach for giving me the opportunity to work in his laboratory and for his patient support and guidance during the course of my Ph.D. Also, of course, thank you to all the past and present members of the lab that have helped out during my time here. A special big thank you to Elise Darmon for her critical reading of the first draft of this thesis.

Thank you to Bénédicte Michel and Vladimir Bidnenko in Paris who helped me get to grips with Pulsed-field gel electrophoresis and provided several mutant strains and plasmids used in this study. Thank you also to Garry Blakely for providing the SOS reporter plasmid pGB150 used in this study.

A special mention for all the friends I have made in Edinburgh whose company has made the time pass by very quickly and given me many happy memories during my postgraduate studies. Especially the 'lads' John Blackwood, Phil Jordan and Florin Dale who helped me settle in Edinburgh at the beginning and were always available to share a cheeky pint and watch the football.

A huge thank you to Jenn, whose love and support has helped me through the 'last-few' experiments and the 'writing-up' stages of my PhD.

I would like to dedicate this thesis to my parents, who have encouraged and helped me throughout my education in Liverpool, Nottingham and Edinburgh. I am eternally grateful. Thank you also to my sisters, Sarah, Louise and Emily, my Nan, and the rest of my family including my Gran and Grandad who I am sure are looking down on me.

Finally thank you to the MRC whose scholarship allowed me to study for my PhD at the University of Edinburgh.

Abstract

Long DNA palindromes are sites of genome instability (deletions, gene amplification and translocations) in eukaryotic cells. In both prokaryotic and eukaryotic cells they are sites of DNA breakage that can be repaired by homologous recombination. Genetic evidence suggests that in *E. coli* breakage is mediated by the SbcCD complex (Rad50/Mre11) that can cleave DNA-hairpin structures. Here an arabinose inducible-*sbcDC* (P_{BAD}-*sbcDC*) system has been constructed and used to obtain genetic and *in vivo* physical evidence of DNA double-strand breaks (DSBs) at a 246 bp interrupted palindrome. These breaks are shown to have two ends suggesting that they occur behind the replication fork probably on the lagging strand. A two-ended break is not compatible with the palindrome blocking DNA replication and cleavage causing the fork to collapse. Repair of the breaks requires homologous recombination facilitated by the RecA and RecBCD proteins as well as the Holliday junction resolvases RuvABC and RecG. Contrary to previous reports this study rules out a significant role of the RecFOR proteins in the repair of breaks and implicates a role for the PriA protein – probably to establish replication forks at the break site. Removing the XerCD/*dif* system (chromosome dimer resolution) has a minor negative effect on cell viability suggesting that chromosome crossover events occur at a detectable (but unquantified) frequency. There is also a requirement for cleavable LexA protein demonstrating that the SOS response is required for efficient repair.

Common abbreviations

ATP	adenosine triphosphate
BFB	bridge-fusion-break
BIR	break-induced replication
bp	base-pairs
CHO	Chinese hampster ovarian
cm	centimetre
°C	degrees Celcius
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
DIG	digoxygenin
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DSB	double-strand break
dsDNA	double-stranded DNA
dTTP	deoxythymidine triphosphate
g	gram
GFP	green-fluorescent protein
kb	kilobase pair
l	litre
M	molar
µg	microgram
µl	microlitres
µm	micrometre
µM	micromolar
mg	milligram
ml	millilitre
mM	millimolar
MRN	Mre11/Rad50/Nbs1 (*S. pombe* homologue of SbcCD)
mRNA	messenger ribonucleic acid
MRX	Mre11/Rad50/Xrs2 (*S. cerevisiae* homologue of SbcCD)
ng	nanogram
nm	nanometre
nt	nucleotide
OD	optical density
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
pfu	plaque forming units
pH	power of hydrogen
PMGR	plasmid mediated gene replacement
rpm	revolutions per minute
SSB	single-strand binding protein
ssDNA	single-stranded DNA
SSG	single-strand gap

T_m	melting temperature
UV	ultra-violet light
v/v	volume per unit volume
w/v	weight per unit volume

Table of Contents

List of figures

List of tables

Chapter 1: Introduction

The gram-negative bacterium *Escherichia coli* has proved a useful model organism for studying genetics of many different cellular processes. Homologous recombination has been intensively investigated in *E. coli* and, through the use of genetic analyses, physical analyses of DNA and biochemical studies of purified proteins, the pathways involved in this process are now well defined. Many of the gene products involved in recombination appear to have homologues in higher organisms, suggesting that the basic mechanisms of recombination are conserved through evolution. In addition to studies in recombination, *E. coli* has also been used to investigate the effect of DNA palindromes on the viability of an organism. Palindromic sequences are sites of genetic instability. In *E. coli,* they are targets of SbcCD cleavage and can interfere with DNA synthesis. In eukaryotes, they can lead to gene amplification and chromosome translocation: processes that in higher eukaryotes can contribute to onset of disease.

The introduction of this thesis focuses on processes involved in recombination in *E. coli* and on mechanisms of palindrome-mediated inviability and instability. Some of the processes of palindrome instability in *E. coli* are applicable to higher eukaryotes, so palindromes in these organisms are also discussed.

1.1 Homologous recombination in Escherichia coli

Since early descriptions of homologous recombination in *E. coli* and its involvement in gene transfer and inheritance, it has become clear that many of the genes involved in these processes are also involved in repairing damaged DNA (Kuzminov, 1999).

E. coli cells lacking the recombination genes are not only defective in genomic integration of linear DNA molecules during processes of gene transfer like conjugation or P1 transduction (Lloyd and Buckman, 1995) but they are also more sensitive to DNA damaging agents such as UV rays, X-rays or γ-rays (Lloyd et al., 1984), as well as chemical agents such as mitomycin C and cisplatin (Bhattacharya and Beck, 2002; Friedberg, 1995; Keller et al., 2001). In addition, endogenous enzymatic processes that can cause DNA damage also exist (Cromie and Leach, 2001; Cromie et al., 2000; Leach et al., 1997) and DNA breaks can be caused by replication fork collapse (Michel et al., 1997). Clearly, a major role of homologous recombination in *E. coli* is to repair DNA damage, with high fidelity, thus allowing faithful transmission of genetic material through generations.

Homologous recombination for repair of DNA breaks can be described in a three-step process: presynapsis, synapsis and postsynapsis (see figure 1.1) (Clark, 1971). Each of the steps can be summarised as follows: (1) Presynapsis: a free DNA end is processed to give a recombinogenic, single-stranded, 3' DNA end coated with the RecA protein. (2) Synapsis: the RecA-coated single-stranded 3' DNA end (ssDNA) molecule invades and pairs with a double-stranded homologous sequence with ejection of the identical strand. (3) Postsynapsis: the junctions of this joint molecule are moved away from the initial invasion site, generating a Holliday junction. DNA synthesis from the 3' end of the invading molecule, using the intact / undamaged invaded molecule as a template, completes the high-fidelity repair of the broken molecule. Finally, the joined molecules are separated by Holliday junction resolution to give intact, recombinant DNA molecules. In *E. coli,* the proteins

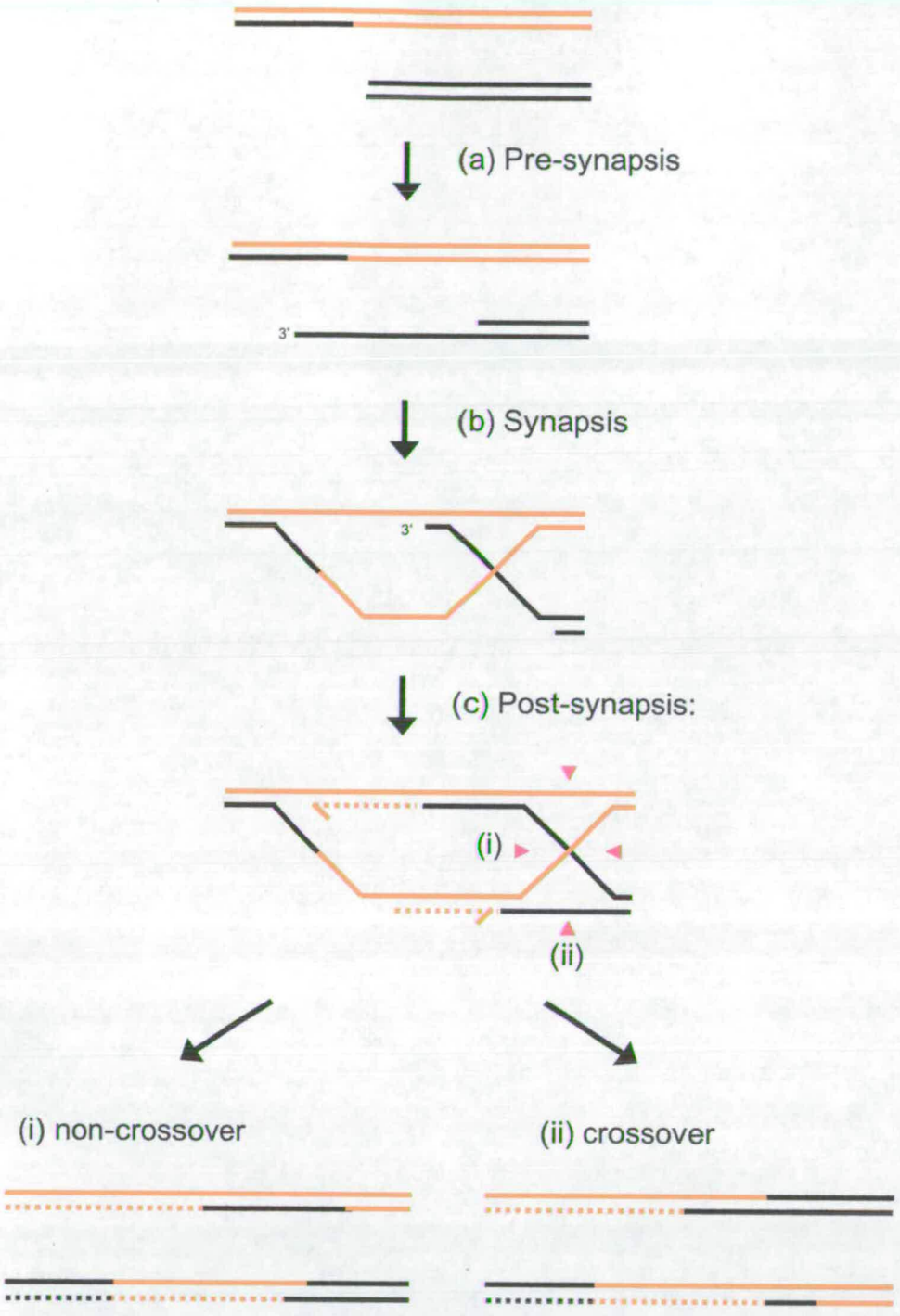


Figure 1.1: The three stages of homologous recombination. Following the collapse of a replication fork that generates a one-ended DNA DSB the processes of (a) Pre-synapsis (b) synapsis and (c) post-synapsis ensures the repair of the DNA without loss of genetic information. Depending on the DNA strands cleaved at post-synapsis either non-crossover or crossover recombinant products are generated.

involved in these three stages of homologous recombination have well defined roles, which will be discussed in more detail.

Recombination in *E. coli* has important roles in ensuring complete chromosomal replication. Replication forks often encounter obstacles (including DNA damage) that impede their progress and several components of the recombination machinery are important for reestablishment of broken or interrupted replication forks. These processes will not be discussed in detail here but have been the subject of several reviews (Courcelle et al., 2004; McGlynn, 2004; Michel et al., 2004).

Finally, homologous recombination is a very important process for generating genome variation and repairing DNA damage in eukaryotes. The basic processes and some of the proteins involved are similar to those in *E. coli* and will not be discussed here. A review on this subject can be found by Cromie and collaborators (Cromie et al., 2001).

1.1.1 Pre-synapsis

Pre-synapsis involves the preparation of damaged DNA for invasion into an intact homologous DNA molecule. Two distinct protein groups appear to act at this stage to process different types of damaged DNA and facilitate loading of the RecA protein filament. Broadly speaking, the RecBCD complex has a role in the repair of DNA double-strand breaks (DSBs) and the RecF, O and R proteins are required for the repair of DNA single-strand gaps (SSGs).

1.1.1.1 RecBCD and processing of DNA DSBs

The RecBCD complex plays a very important role in processing the ends of a DNA DSB and stimulating recombination by loading RecA onto ssDNA. It acts as a hetero-trimer that rapidly unwinds and degrades blunt-ended double-strand DNA in a highly regulated manner (Boehmer and Emmerson, 1991; Taylor and Smith, 1995a). The RecB and RecD subunits provide 3' to 5' and 5' to 3' helicase activities, respectively (Dillingham et al., 2003; Taylor and Smith, 2003). RecB also provides both 3' – 5' and 5' – 3' nuclease activities (Yu et al., 1998). The RecC subunit recognises a DNA sequence, known as *chi,* that is important in the regulation of RecBCD activities (Handa et al., 1997). Recently the crystal structure of RecBCD was solved, revealing more information on how the subunits work together to perform their function (Singleton et al., 2004). A diagram summarising RecBCD function is shown in figure 1.2.

The RecBCD complex recognises and binds blunt ended DNA (Taylor and Smith, 1995a). It then translocates along and unwinds double-stranded DNA using its helicase domains (Bianco et al., 2001; Dillingham et al., 2003; Taylor and Smith, 2003), while the nuclease domain of the RecB subunit progressively cleaves the single-strand DNA ends with an initial cleavage-bias to the 3' end (Dixon and Kowalczykowski, 1993). The cleavage-bias probably reflects the rigid positioning of the single nuclease domain within the complex that allows easier access by the 3' ssDNA end (Singleton et al., 2004).

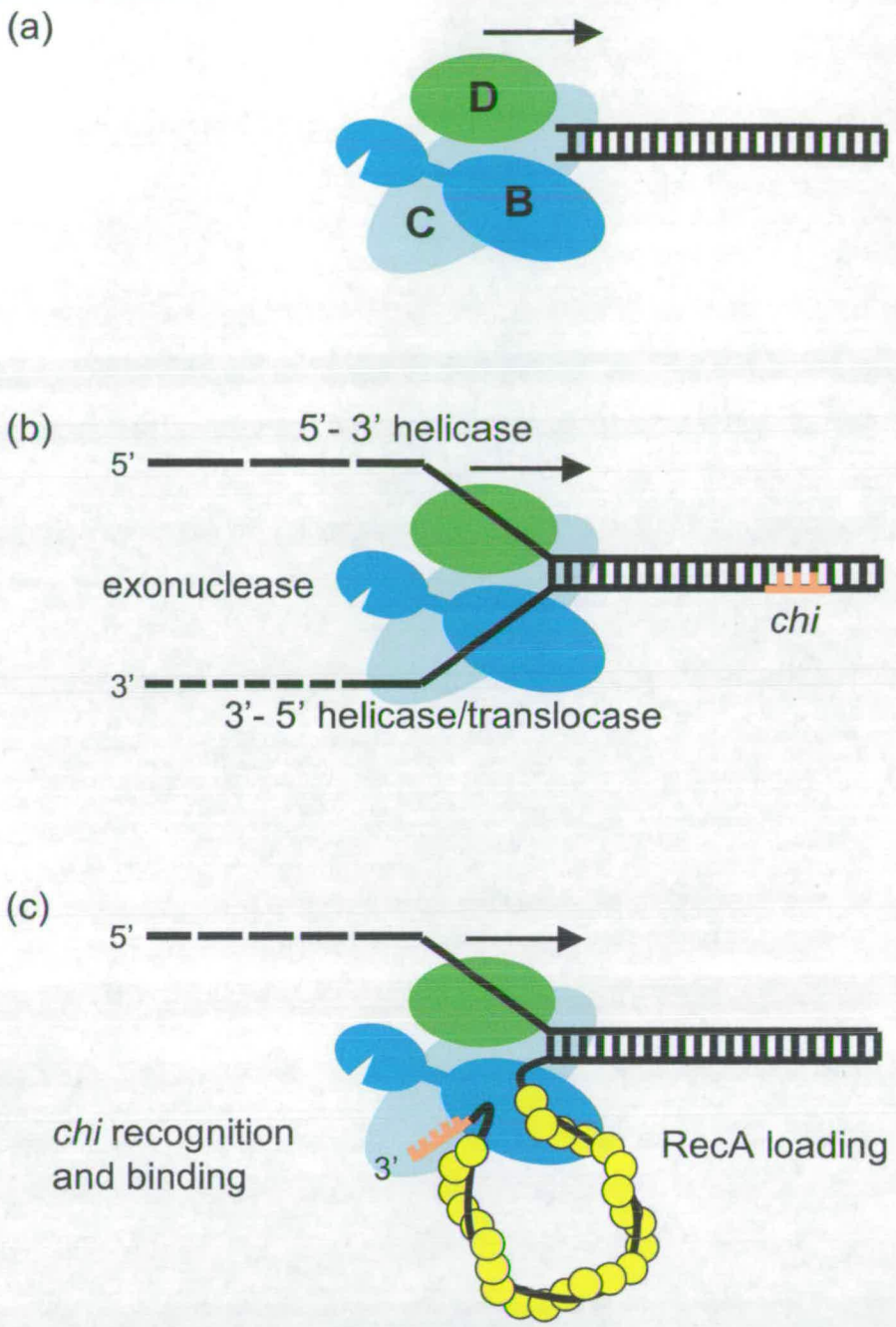


Figure 1.2: The activity of the RecBCD complex during presynapsis at a broken DNA end. (a) RecBCD binds and begins unwinding a blunt-ended broken DNA molecule, (b) the RecBCD nuclease site favours cleavage of the 3' end, (c) recognition of a *chi* sequence in the 3' end leads to an activity change that favours 5' end cleavage and the loading of RecA onto the generated 3' tail. Based on model proposed by Singleton and collaborators (Singleton, Dillingham et al. 2004).

Upon encountering a *chi* sequence the activities of RecBCD are altered: (1) the rate of translocation is reduced, (2) the activity of the nuclease domain is attenuated and the strand bias is reversed and (3) the RecA-loading activity is stimulated (Anderson and Kowalczykowski, 1997a; Dixon and Kowalczykowski, 1993; Spies et al., 2003). The *chi* sequence reads 5' GCTGGTGG 3' and was discovered as a genetic initiator of recombination (see Stahl, 2005), a property later demonstrated in biochemical studies (Dixon and Kowalczykowski, 1991). The *chi* sequence is recognised in its ssDNA form by the translocating RecBCD complex (Bianco and Kowalczykowski, 1997; Taylor and Smith, 1995b). The reduced rate of translocation of the RecBCD complex provoked by the *chi* sequence is reversible and is not caused by ejection of the RecD motor subunit, as previously thought, but probably by a conformational change in the complex (Dixon et al., 1994; Handa et al., 2005). The effects of *chi* upon the nuclease activity are probably propagated through the binding of the 3' ssDNA end containing the *chi* sequence to the RecBCD complex, preventing it from entering the nuclease site and allowing greater access of the 5' ssDNA end (Singleton et al., 2004). The *chi* sequence also stimulates RecBCD-mediated RecA filament formation on the 3' ssDNA overhang. RecBCD allows RecA to overcome the competition with single-strand binding protein (SSB) for binding to ssDNA (Anderson and Kowalczykowski, 1997b; Churchill et al., 1999; Churchill and Kowalczykowski, 2000).

1.1.1.2 RecFOR and processing of DNA single-strand gaps (SSGs) and DSBs

The *recF* gene product was originally shown to be required for the suppression of UV sensitivity / recombination deficiency of *recB⁻ recC⁻ sbcB⁻ sbcC⁻* strains (Horii and Clark, 1973; Kushner et al., 1971; Lloyd and Buckman, 1985). In fact, cells lacking RecF are themselves sensitive to UV irradiation, and those lacking RecF and RecB/C are more UV sensitive than either of the single mutants, suggesting that the two gene products act in different pathways for repair of UV-damage (Horii and Clark, 1973). RecF, RecO and RecR were shown to be required for the repair of ssDNA gaps that arise following UV irradiation (Wang and Smith, 1984). Indirect genetic evidence led to the proposal that RecF, RecO and RecR are involved in loading RecA onto ssDNA gaps. The mutant *recA803* was shown to be a suppressor of *recF⁻* UV sensitivity (Volkert and Hartke, 1984; Wang and Smith, 1986). RecA803 is a mutant variant that is more efficient at displacing SSB protein from ssDNA than wild-type RecA (Madiraju et al., 1992; Madiraju et al., 1988).

Despite strong genetic implications for a role of RecF, RecO and RecR in loading RecA protein onto ssDNA gaps, characterising the activities of these proteins biochemically has been difficult. The proteins do not cooperate in a single complex like the RecBCD complex but have a number of interactions that lead to the generation of RecA-coated ssDNA. *In vitro,* a combination of the RecO and RecR proteins are sufficient to overcome SSB competition and stimulate RecA-filament formation and homologous strand invasion (Bork et al., 2001; Umezu et al., 1993; Umezu and Kolodner, 1994). The RecF protein can bind both ssDNA or dsDNA

(Griffin and Kolodner, 1990; Madiraju and Clark, 1992; Webb et al., 1997), but preferentially binds the 5' paired end of a ssDNA-dsDNA junction (Hegde et al., 1996; Morimatsu and Kowalczykowski, 2003). In fact, a combination of both RecF and RecR prevents the extension of a RecA filament into dsDNA regions and can stimulate RecA loading by RecOR onto gapped-DNA substrates coated in SSB protein. This indicates that RecF(R) targets RecOR activity to gapped DNA junctions (Morimatsu and Kowalczykowski, 2003) (see figure 1.3).

The model described above for RecFOR-stimulated loading of RecA at single-strand gaps can also be applied to the repair of DNA double-strand breaks, for example in *recB⁻ recC⁻ sbcB⁻ sbcC⁻* strains (Horii and Clark, 1973; Kushner et al., 1971; Lloyd and Buckman, 1985). The *recQ* and *recJ* genes were also found to be involved in the RecF pathway of suppression of *recB⁻ recC⁻ sbcB⁻ sbcC⁻* strains (Lovett and Clark, 1984; Nakayama et al., 1985). By the concerted activities of the RecQ helicase (Umezu et al., 1990) and the RecJ 5' - 3' exonuclease (Lovett and Kolodner, 1989), a blunt DNA end could be converted to a DNA end with a 3' single-strand overhang providing a dsDNA-ssDNA junction with a 5' paired end for RecF(R)/RecOR stimulated loading of RecA.

1.1.1.3 Interchangeable components of the pre-synapsis machinery

Although the RecBCD complex and the RecF, RecO and RecR proteins process damaged DNA in two discrete pathways, there is evidence that functional components of one presynapsis machine can substitute for dysfunctional subunits of the other to form a hybrid pathway for presynapsis.

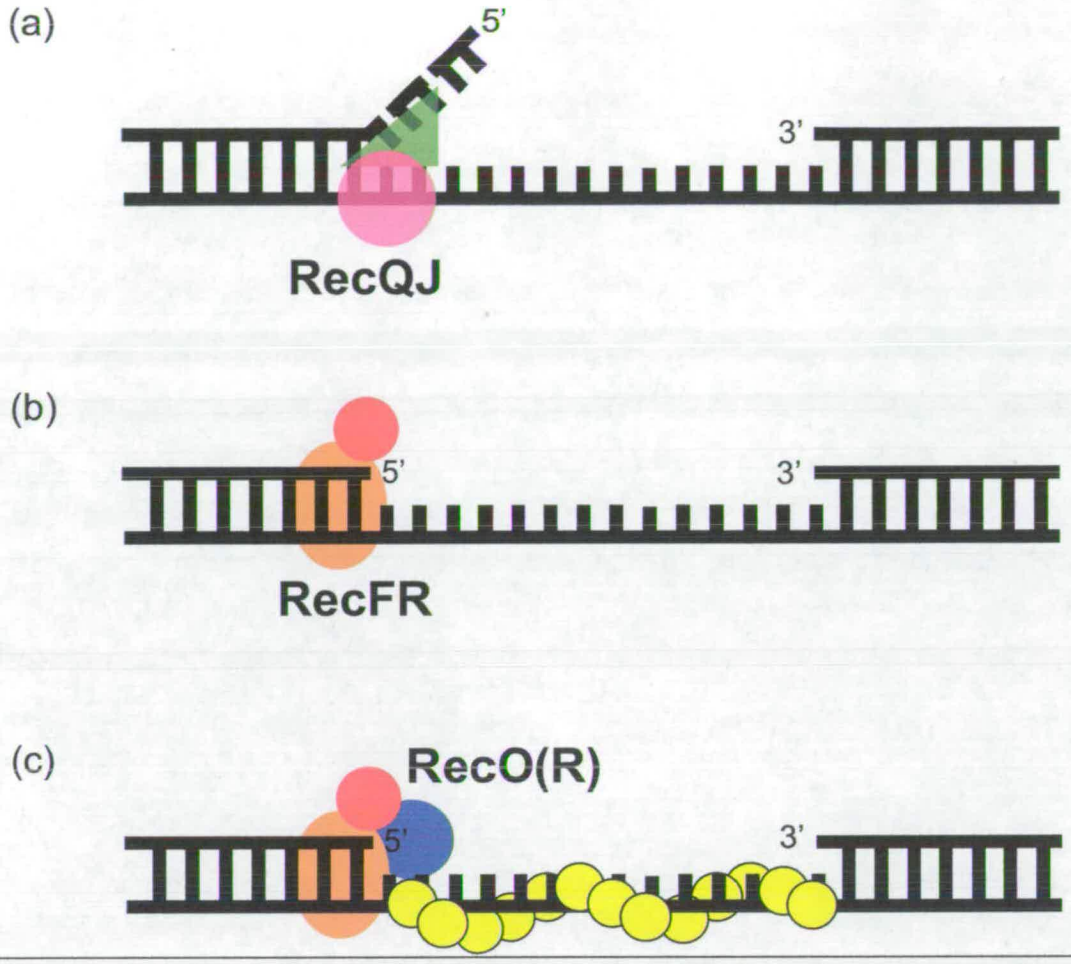


Figure 1.3: The activities of RecF, RecO and RecR during presynapsis at a single-strand gap (a) RecQ and RecJ cooperatively extend the region of ssDNA, (b) RecFR binds the 5' junction between the ssDNA and dsDNA, (c) RecOR is employed to the gap and allows efficient 5' to 3' loading of RecA. Based on model proposed by Morimatsu and Kowalczykowski (Morimatsu and Kowalczykowski 2003).

The mutant allele *recB1080* produces a protein without nuclease activity and is unable to load RecA onto ssDNA *in vitro* (Anderson et al., 1999; Wang et al., 2000; Yu et al., 1998). Genetic analysis of strains containing this mutation revealed that they are recombination proficient only in the presence of functional RecFOR proteins for loading of RecA, and the nuclease activity of RecJ (Amundsen et al., 2000; Ivancic-Bace et al., 2003; Jockovich and Myers, 2001).

1.1.2 Synapsis

Synapsis is the defining stage of recombination where homologous DNA strand invasion occurs. In *E. coli,* synapsis is facilitated in the presence of a RecA protein filament. Strains without functional RecA are compromised for all types of homologous recombination and are severely sensitive to DNA damage (Clark and Margulies, 1965; Howard-Flanders and Theriot, 1966; Kapp and Smith, 1970; Krasin and Hutchinson, 1977; Smith and Meun, 1970; Wang and Smith, 1983).

The mechanism by which RecA catalyses synapsis is complicated and will not be discussed here in detail though RecA function has been the subject of a number of reviews (Kowalczykowski et al., 1994; Lusetti and Cox, 2002). Presynapsis generates a recombinogenic DNA molecule: a ssDNA molecule coated with a right-handed helical filament of RecA. The RecA-filament stimulates alignment and pairing of ssDNA with an intact, homologous, complementary DNA molecule (figure 1.1). This is coupled with the ejection of the identical strand from the intact dsDNA to generate an unpaired D-loop.

1.1.3 Post-synapsis

Post-synapsis includes all the stages of recombination that follow RecA-mediated strand invasion. Firstly, RecA catalyses pairing of the D-loop with the unpaired damaged strand generating a four-way DNA molecule called a Holliday junction, which is productively moved away from the site of damage (figures 1.1 and 1.4). DNA synthesis across the damaged region or re-establishment of a replication fork from the invading molecule completes the high-fidelity repair of a damaged region. Finally, resolution of Holliday junctions by cleavage of symmetrically opposed DNA strands, and rejoining in the opposite configuration separates the joined molecules. In *E. coli* this process can generate crossover products or non-crossover products depending on the DNA strands cleaved (see figure 1.1).

1.1.3.1 Holliday Junction Processing by RuvABC

The RuvABC complex is required for the migration and resolution of Holliday junctions formed early during post-synapsis. Strains that are deficient for any of the *ruvA*, *ruvB* or *ruvC* genes are sensitive to DNA damage (for example caused by UV light) (Otsuji et al., 1974). Biochemical analysis of the different proteins has revealed the role of RuvABC in recombination (West, 1997).

Purified RuvC can resolve RecA-catalysed joint DNA molecules to produce recombinant DNA of either crossover or non-crossover conformations (Dunderdale et al., 1991) (figure 1.4). RuvC exists as a dimer that introduces symmetrical nicks into the DNA backbone of a Holliday junction at a consensus sequence 5'- $^{A}/_{T}$ TT↓ $^{G}/_{C}$ -3' (Bennett et al., 1993; Iwasaki et al., 1991; Shah et al., 1994; Shah et al., 1997).

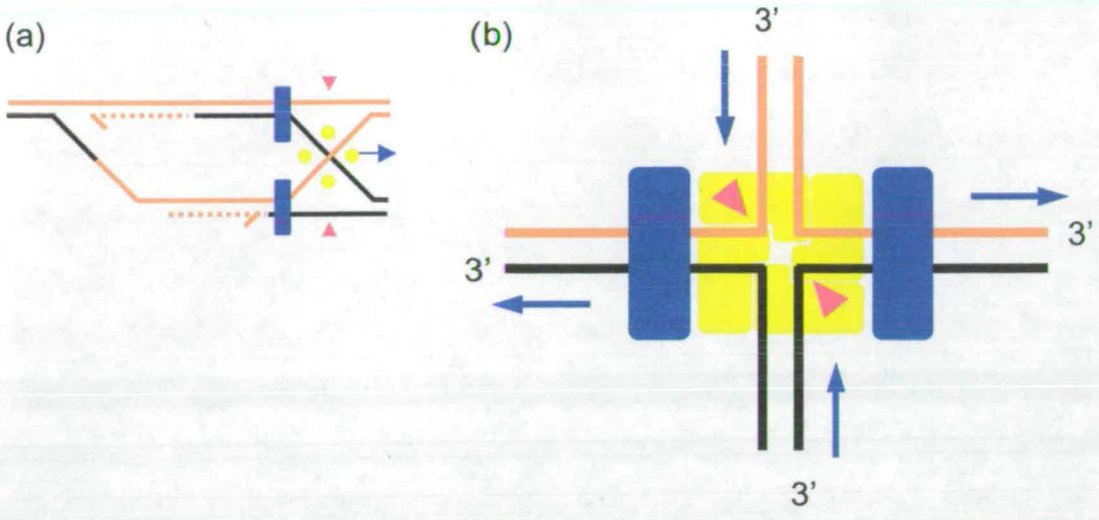


Figure 1.4: The interaction of RuvABC with a Holliday junction. The yellow square represents RuvA, the blue rectangle the RuvB hexamers and the pink arrows RuvC dimer directed cleavage. In (a) the blue arrow shows the direction of movement of the Holliday junction, and (b) the blue arrows show the direction the DNA is pulled across the RuvA tetramer by the RuvC hexamers. This diagram represents the square planar conformation of the DNA in a RuvAB(C) bound Holliday Junction.

Purification and analysis of the RuvA and RuvB proteins revealed that together they efficiently migrate Holliday junctions in the presence of ATP (Parsons et al., 1992; Shiba et al., 1991; Tsaneva et al., 1992). A tetramer of RuvA (with four-fold rotational symmetry) recognises and binds a Holliday junction, holding it in a square planar conformation (Hargreaves et al., 1999; Rafferty et al., 1996), while two hexameric rings of RuvB stimulate Holliday junction branch migration, hydrolysing ATP (Parsons and West, 1993; Stasiak et al., 1994) (figure 1.4). The RuvAB complex recruits the RuvC dimer at a Holliday junction (Whitby et al., 1996) and migration by RuvAB to a suitable RuvC cleavage site stimulates Holliday junction resolution (Eggleston and West, 2000; van Gool et al., 1998; Zerbib et al., 1998). The orientation of RuvAB at a Holliday junction specifies strand cleavage; RuvC cleaves the two strands passing in the 3' direction through the RuvB hexameric rings towards the Holliday junction (van Gool et al., 1999) (figure 1.4). As a consequence of productive loading of RuvB, so that a Holliday junction is moved away from damage the polarity of cleavage has important implications in the generation of predictable recombinant products (crossover or non-crossover; figure 1.1) following repair of different types of DNA DSBs (Cromie and Leach, 2000).

1.1.3.2 The RecG helicase

E. coli strains without functional RecG protein are defective in recombination and moderately sensitive to UV light (Lloyd and Buckman, 1991). These defects are greatly increased in *recG*$^-$ *ruvA*$^-$/*ruvB*$^-$/*ruvC*$^-$ double mutants (also to a greater extent than defects/sensitivity of a *ruv*$^-$ single mutant) (Lloyd, 1991). In addition, it was recently shown that *recG*$^-$ *ruv*$^-$ double mutants are extremely defective in repair of a

DNA DSB generated at a unique chromosomal cleavage site wereas single mutants are only modestly defective (Meddows et al., 2004). These studies suggest that RecG might be or direct an alternative Holliday junction resolvase.

Biochemical studies demonstrate that RecG is a monomeric DNA-dependent ATPase that can dissociate synthetic Holliday junctions by helicase activity (like the RuvAB complex), but no nuclease activity (analogous to RuvC) was detected (Lloyd and Sharples, 1993; McGlynn et al., 2000). RecG can also act at three-way DNA junctions, such as those catalysed by RecA during synapsis, an activity that is anti-recombinogenic, reversing RecA-stimulated reactions (Whitby and Lloyd, 1995b; Whitby et al., 1993). Since RecG has been shown to act at three-way and four-way DNA junctions, models have been proposed where RecG is involved in processing of stalled or collapsed replication forks and the predicted role in resolution of recombination intermediates remains unclear (Briggs et al., 2004; McGlynn and Lloyd, 1999; McGlynn and Lloyd, 2000; McGlynn and Lloyd, 2002; McGlynn et al., 2001).

1.1.3.3 PriA establishes replication forks following synapsis

E. coli strains lacking functional PriA are sick, sensitive to DNA-damaging agents and defective in homologous recombination (Kogoma et al., 1996; Lee and Kornberg, 1991; Sandler et al., 1996). These phenotypes reflect the important role of the PriA protein in ensuring complete DNA replication and in homologous recombination.

Chromosome replication in *E. coli* is initiated at a single origin known as *oriC* and requires many proteins that are involved directly in DNA synthesis, or indirectly as initiation factors that lead to appropriate loading of the relevant protein machines. This process is a stepwise process that begins with the recognition and binding of *oriC* by the DnaA protein and includes the loading of the primosome complex that provides RNA primers for lagging strand DNA synthesis (Marians, 1992). Studies in *E. coli* suggest that, under normal growth conditions, up to 18% of replication forks arrest and require the activity of the DnaC protein to complete replication (Maisnier-Patin et al., 2001). DnaC is one of the initiation factors of replication and biochemical analysis reveals that it is required for loading the primosome complex (specifically the DnaB helicase) onto DNA (Arai and Kornberg, 1981; Kaguni and Kornberg, 1984). Interestingly, certain mutant versions of the *dnaC* gene suppress the UV sensitivity and defect in homologous recombination of a *priA*$^-$ mutant suggesting that PriA interacts with DnaC (Sandler et al., 1996). However, it had been previously demonstrated that PriA is not required for the assembly of the primosome at *oriC* (Kaguni and Kornberg, 1984).

Biochemical analysis showed that PriA is required for loading the primosome complex onto the ssDNA of ΦX174 via an interaction with DnaC, allowing DNA Polymerase III replication. PriA initiates this reaction by recognition and binding of a hairpin structure in the ssDNA substrate, a role analogous to DnaA binding of *oriC* (Arai and Kornberg, 1981). Later, it was shown that PriA could direct primosome loading at a D-loop structure (generated by RecA-mediated strand invasion), linking the PriA protein with a specific role in recombination (Liu and Marians, 1999;

McGlynn et al., 1997). In this way, PriA can direct leading and lagging strand DNA synthesis following RecA-mediated strand invasion (Xu and Marians, 2003).

By the same mechanism it would be possible for PriA to underpin DNA replication. Following replication fork arrest, the recombination machinery and PriA can re-establish a replication fork with both leading and lagging strand synthesis (Michel et al., 2004).

1.1.4 Resolution of chromosome dimers following crossover product formation

Holliday junction resolution of recombination intermediates into crossover products has important implications for the circular chromosome of *E. coli.* The final crossover product is a circular chromosome dimer, which cannot be segregated to daughters at cell division. To separate chromosome dimers, *E. coli* employs the XerCD site-specific recombinase, which targets an additional crossover recombination reaction to a 32 bp site (termed *dif*) found near the terminus of the chromosome. This separates the dimeric chromosomes into two monomers that can be segregated into daughter cells (Sherratt et al., 1995).

1.2 The SOS response to DNA damage

In *E. coli,* exposure to DNA-damaging agents such as UV light or mitomycin C is followed by the increased expression of a set of genes known as the "SOS genes" (Courcelle et al., 2001; Kenyon and Walker, 1980; Walker G.C., 2000). This response increases the chance of survival of cells following DNA-damage by helping

to remove and repair damaged DNA, or to replicate beyond the damage. Up-regulated genes include some of the recombination genes including *recA* and *ruvAB*.

1.2.1 Propagation of the SOS response

The SOS response and increased expression of the SOS genes are elicited by dissociation of the LexA repressor protein from the SOS promoters. Under normal conditions, LexA binds and excludes RNA polymerase from the promoters of the SOS genes (Brent and Ptashne, 1981; Little et al., 1981). However this protein has an autocatalytic cleavage property, which is stimulated following DNA damage, leading to dissociation from its DNA binding sites and allowing transcription of the previously repressed genes (Little, 1984; Little, 1993). Interestingly, the LexA 'recognition' of presence of DNA damage is via an interaction with ssDNA-RecA filaments. *In vitro* experiments show that the presence of RecA bound to ssDNA greatly enhances expression from SOS-promoters due to an increase in LexA cleavage (Anderson and Kowalczykowski, 1998; Little, 1984; Little et al., 1981).

1.2.2 SOS polymerases and translesion synthesis

The SOS response in the cell, as well as increasing expression of some of the recombination genes, presumably to allow repair of the DNA damage, increases the expression of the polymerases PolII, PolIV and PolV.

PolII (encoded by *polB / dinA*) levels are increased during early stages of the SOS response and this error-free polymerase may act in the same pathway as PriA and therefore may be important in replication restart (Pham et al., 2001; Rangarajan et al., 1999). PolIV is encoded by the *dinB* gene and is an error-prone polymerase

implicated in mutagenesis and it is unclear how this helps a cell during an SOS response (Wagner et al., 1999).

PolV is an error-prone polymerase encoded by the *umuDC* genes. The protein complex is activated by autocatalytic cleavage of the UmuD protein, a reaction that is stimulated by ssDNA-RecA (Shinagawa et al., 1988). In brief, the active PolV complex is a heterotrimer (two subunits of activated UmuD and one of UmuC) that acts with the PolIII processivity complex and can effectively synthesise beyond DNA lesions in the presence of a RecA filament. Disassembly of the RecA filament coincides with dissociation of the PolV complex from the DNA. Thus, PolV is targeted to damaged DNA templates (Pham et al., 2001).

1.3 The SbcCD complex

In *E. coli,* the *sbcC* (and *sbcD*) genes were first described when mutations inactivating the gene products, in association with the *sbcB*$^-$ mutation were found to suppress the DNA-damage sensitivity of *recB*$^-$ strains (Lloyd and Buckman, 1985). Leach and Stahl reported that *recB*$^-$ *sbcB*$^-$ *sbcC*$^-$ strains could also support the replication of long palindromic sequences (Leach and Stahl, 1983). The original observation of suppression of the *recB*$^-$ phenotype by *sbcDC*$^-$ is still not understood and the natural role of SbcCD in *E. coli* remains a mystery. Biochemical and structural studies have given a better understanding of SbcCD and genetic and biochemical studies of eukaryotic homologues have revealed a number of possible roles for this complex.

1.3.1 SbcCD has a variety of nuclease activities *in vitro*

Biochemical studies have shown that, in the presence of Mn^{2+} ions, SbcCD has ATP-dependent exonuclease and ATP-independent endonuclease activities (Connelly et al., 1997; Connelly and Leach, 1996). Further studies have demonstrated that SbcCD can cleave the loop of a closed hairpin DNA molecule (Connelly et al., 1998). This activity is consistent with genetic phenotypes of palindrome inviability in *E. coli* (Cromie et al., 2000; Leach et al., 1997; Leach and Stahl, 1983). More recently, it was shown that SbcCD can remove a tightly bound protein from the end of a double-stranded DNA molecule by generation of a DNA double-strand break adjacent to the protein. It was proposed that SbcCD might recognise and cleave blocked double-strand DNA ends, allowing access to damaged DNA ends by other repair proteins (Connelly et al., 2003).

1.3.2 Eukaryotic SbcCD homologues

The SbcCD complex shows protein sequence identity and structural similarities with the Rad50 (SbcC) and Mre11 (SbcD) complex in eukaryotes (humans and yeast) suggesting that SbcCD has an important, conserved role *in vivo* (Connelly and Leach, 2002; de Jager et al., 2004). Similar to SbcCD *in vitro* characteristics, eukaryotic Mre11/Rad50 is required *in vivo* (1) for removing proteins covalently bound to DNA ends (Connelly and Leach, 2004) and (2) in processing of hairpin-capped DNA ends (Lobachev et al., 2002). Other roles of the Mre11/Rad50 complex in eukaryotes are numerous and may also yield clues as to the roles of SbcCD in *E. coli*. Mre11/Rad50 complexes are involved in DNA double-strand break repair, generation of DNA double-strand breaks during meiosis, maintenance of telomere

length, and regulation of the cell cycle, roles which will not be discussed here but have been reviewed by Connelly and Leach (Connelly and Leach, 2002).

1.4 Palindromic DNA

In the Oxford English dictionary, a palindrome is defined as a word or phrase that reads the same backwards as forwards (for example *rotator* or *able was I ere I saw elba*). Likewise, in molecular biology a DNA palindrome is a sequence of DNA that is identical when read 5' to 3' on either strand. Palindromes are also inverted repeats and have two-fold rotational symmetry. This gives each single-strand of a palindromic DNA molecule the ability to base pair with itself and to form a secondary structure (figure 1.5).

A palindromic DNA sequence in double-stranded DNA has the potential to form a cruciform structure (a four-way DNA molecule like a Holliday junction) as the two DNA strands become dissociated from each other and form intramolecular base pairs (figure 1.5 (b)). Cruciform formation can become energetically more favourable than regular Watson-Crick base-pairing with an increase in negative supercoiling. Cruciform structures can be detected by altered migration of DNA in an agarose gel, cleavage by structure-specific nucleases, or by electron microscopy (Courey and Wang, 1988; Leach, 1994; Mizuuchi et al., 1982).

There is a distinction between perfect and interrupted palindromes. An interrupted palindrome consists of a pair of inverted repeats separated by a short non-symmetrical spacer. Interrupted palindromes share some but not all properties of perfect palindromes. A spacer at the centre of a palindrome reduces its ability to

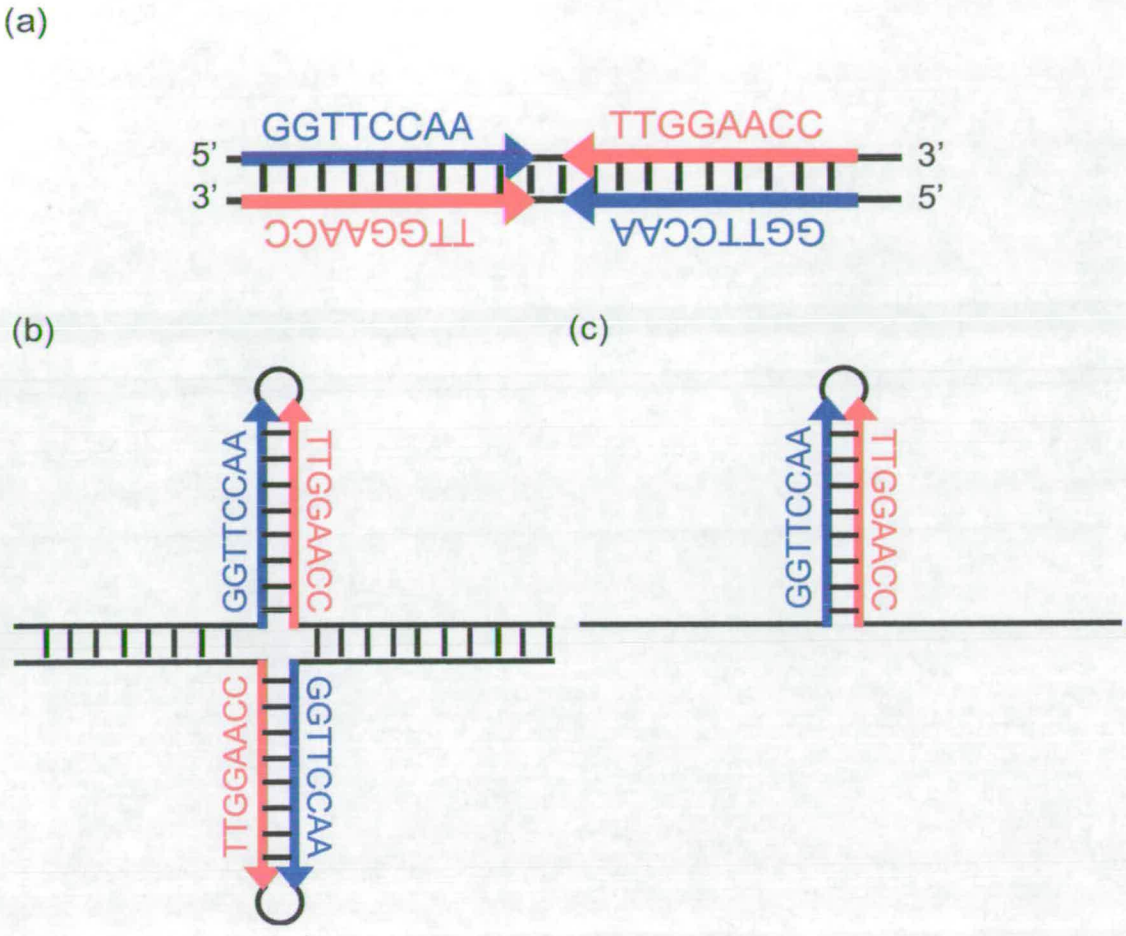


Figure 1.5: Palindromic sequences and secondary structures. (a) A hypothetical palindromic sequence forming (b) a cruciform, or (c) a hairpin secondary structure (adapted from Leach 1994).

form a cruciform structure *in vitro* in negatively supercoiled DNA but does not prevent the formation of a DNA hairpin in single-stranded DNA (see figure 1.5). The interrupted palindrome used in this study measures 246 bp and contains two 111 bp inverted repeats separated by a 24 bp non-symmetrical central spacer (Pinder et al., 1998).

1.5 *Palindromes in* Escherichia coli.

In *E. coli,* the introduction of long perfect, and some interrupted, palindromes of above 200 bp is associated with host inviability or palindrome instability. These observations appear to be caused by the formation of secondary structures that are: (1) targeted by nucleases to form DNA double-strand breaks (DSBs), (2) able to interfere with DNA synthesis, or (3) able to stimulate template strand slippage during replication (Bzymek and Lovett, 2001; Cromie et al., 2000; Gibson et al., 1992; Leach, 1994; Leach et al., 1997; Leach and Stahl, 1983; Lindsey and Leach, 1989; Pinder et al., 1998).

1.5.1 Palindrome inviability in episomal DNA

The introduction of long palindromic DNA sequences into *E. coli* on plasmids or bacteriophage lambda was often associated with "episome death" (inviability). In 1983, it was shown that inviability was eliminated in *recBC⁻ sbcB⁻* mutants (Leach and Stahl, 1983). It was later discovered that the strain also had a mutation in *sbcC* and that the *sbcC* and *sbcD* genes were responsible for conferring inviability (Chalker et al., 1988; Gibson et al., 1992). The *sbcC* and *sbcD* genes are present on a single operon (Naom et al., 1989) and encode a protein complex which biochemical studies have shown to have ATP-dependent double-strand DNA

exonuclease, ATP-independent single-strand DNA endonuclease and hairpin nuclease activity (Connelly et al., 1997; Connelly et al., 1998; Connelly and Leach, 1996). Despite episomal inviability observed in previous studies, *sbcDC*$^{+}$ cells containing a chromosomal 246 bp interrupted palindrome (harboured in a lambda prophage) were viable. However, this viability requires functional copies of the recombination genes implicated in DNA DSB repair (Cromie et al., 2000; Leach et al., 1997).

The biochemical properties of SbcCD and the involvement of SbcCD in palindrome inviability imply that hairpin or cruciform structures probably form during the processing of palindromic DNA. Presumably, the extent of inviability is dependent upon the propensity for a sequence to form a secondary structure (that can be attacked by SbcCD) and the ability for the cell to repair damage induced by that structure.

1.5.2 Palindrome instability

Although the inviability phenotype associated with palindromic DNA is related to the activity of the SbcCD complex, long palindromic sequences remain unstable and susceptible to deletion even in strains lacking functional SbcCD. Early experiments showed that a phage lambda containing a 3.2 kb palindrome could be successfully plated on a *recBC*$^{-}$ *sbcB*$^{-}$ *sbcC*$^{-}$ *E. coli* strain (Leach and Stahl, 1983). However, this palindromic sequence was very unstable and analysis of the lambda DNA from these experiments revealed that the palindromic DNA had undergone deletions to an average length of 680 bp.

The instability of palindromic sequences in *E. coli* has been the subject of several studies. An early study looking at plasmid-borne palindrome deletions demonstrated that deletion events could occur by replication slippage at a secondary structure between short direct repeats (one within the palindromic sequence and one outside; see figure 1.6 pathway ii) (Weston-Hafer and Berg, 1991).

Work in this laboratory followed up this idea by looking at deletion patterns of long DNA palindromes in *sbcDC*⁻ strains, in the *E. coli* chromosome, in bacteriophage lambda and in a high-copy-number plasmid. A number of palindrome deletions were observed and characterised. Two modes of stabilizing deletions resulted from this work; both were consistent with replication slippage at direct-repeats proposed by previous studies. The deletions could be explained by disassociation of newly replicated DNA at sequences within the palindrome with subsequent reannealing and replication (1) from identical sequences further in the palindrome, generating a non-symmetrical deletion at the centre of the palindrome (see figure 1.6 pathway i) or (2) with an identical sequence outside the palindrome, deleting most of the palindromic sequence (see figure 1.6 pathway ii). Considering the direction of replication with respect to the palindrome and assuming that the direct repeat closest to the centre of the palindrome was the point at which replication was interrupted and the DNA became disassociated, the deletion events observed should have a specific strand bias. All of the deletion events could be explained by secondary structure formation in the lagging-strand template (Pinder et al., 1998). A similar conclusion was also obtained in work by Trinh and Sinden (Trinh and Sinden, 1991).

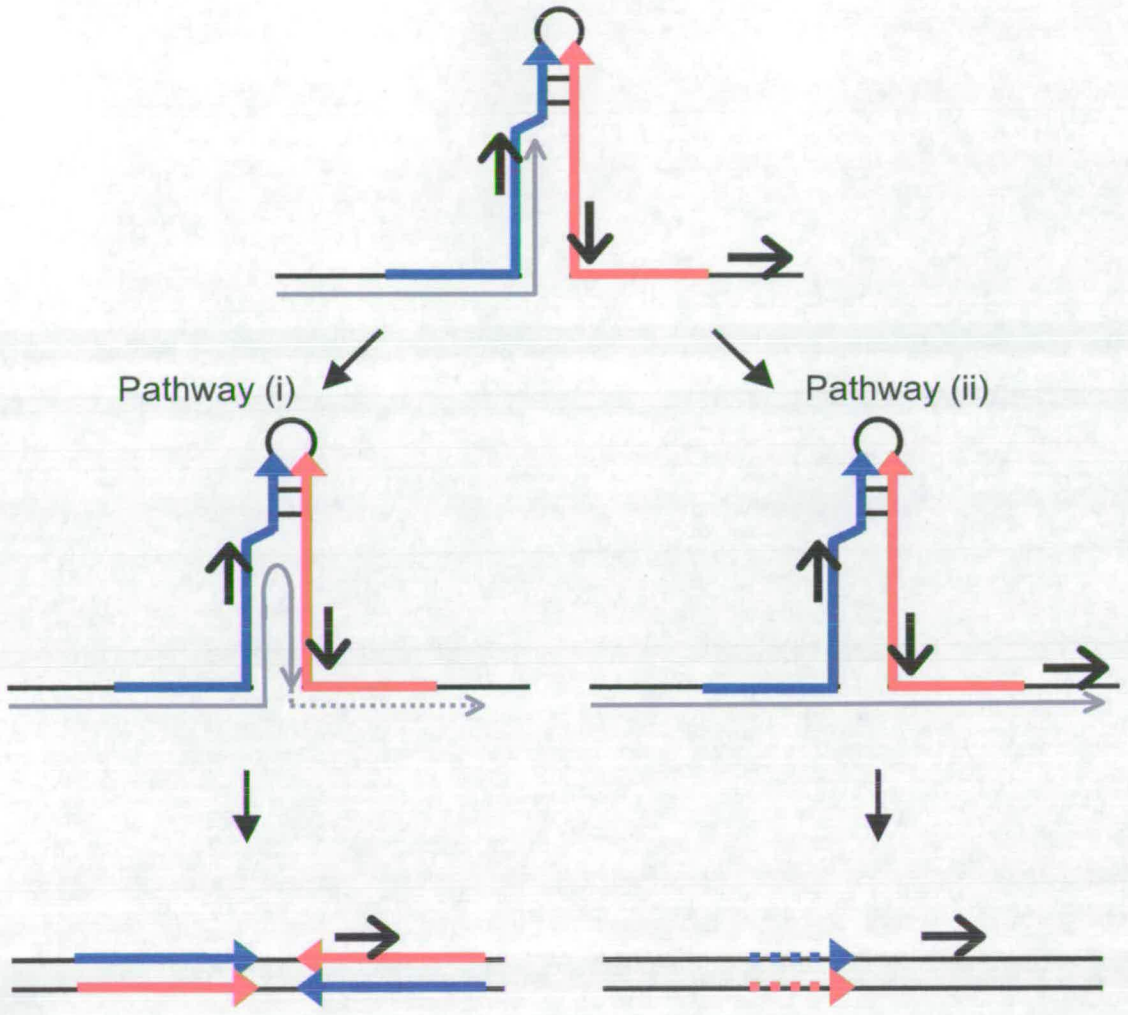


Figure 1.6: Replication slippage at palindromic hairpin structures. Modes of palindrome instability caused by hairpin stimulated replication slippage and mis-allignment at direct repeats (represented by black arrows). (i) The pathway for a central deletion and formation of a shorter interrupted palindrome, and (ii) deletion of one of the inverted repeats leaving no palindromic sequence. (Diagram adapted from Pinder, Blake et al. 1998).

In a study of plasmids containing palindromic sequences, Bzymek and Lovett observed similar SbcCD-independent palindrome deletion facilitated by replication slippage between two large direct-repeats flanking the palindrome. These events were also consistent with replication slippage at the palindromic sequence on the lagging strand template (Bzymek and Lovett, 2001). Albertini and collaborators described a similar, frequent deletion event in the *E. coli* chromosome that occurs between natural, direct-repeats flanking a palindromic sequence (Albertini et al., 1982).

Bzymek and Lovett also described an SbcCD-dependent palindrome deletion pathway that occurs in a construct where two long direct repeats flank a perfect palindrome. These deletions show no bias to the leading or lagging strand and fit a model of replication-independent cruciform extrusion followed by SbcCD-cleavage. Repair follows an uncharacterised RecA-independent pathway with a precise single-strand annealing mechanism between the long direct-repeats flanking the palindrome (Bzymek and Lovett, 2001). However, this class of events occurs with very specific requirements of long direct-repeats flanking a perfect palindrome and may not be a general mode of palindrome instability.

1.6 Palindromes in eukaryotic genomes

Despite a boom in genome sequencing projects during the last decade, the distribution of palindromes in eukaryotic organisms is largely unknown. At various stages during genome sequencing, DNA sequences native to the organism in question are fragmented and cloned into plasmids that are introduced into strains of

E. coli for propagation and long-term storage. Any palindromes in these DNA fragments would be subject to the palindrome-constraints of *E. coli* (described above) and would potentially be inviable or unstable. This problem was specifically demonstrated in a study of the human intronic sequence at the NF1 locus that has multiple conflicting entries with varying levels of palindromy in GenBank (Lewis et al., 2005). In the future, improved genome sequencing methods will have to be employed to find out the true palindromic content of the genomes of higher organisms.

Nonetheless, some eukaryotic genomes (or some regions) have been fully sequenced including the problematic regions. Lisnic and collaborators looked at palindromic sequences in the genome of *Saccharomyces cerevisiae*. Similarly, LeBlanc and collaborators looked at palindromic sequences in chromosomes III and X of *Caenorhabditis elegans*. In yeast, there was an under-representation of palindromes smaller than 10 bp (compared to those found in randomly generated genomes) probably because a large proportion would exist in coding regions and could have detrimental mutagenic effects, or affect mRNA metabolism. In both yeast and *C. elegans*, intermediate and large palindromes are over-represented but exist almost entirely outside coding regions and are probably involved in regulation of gene expression (LeBlanc et al., 2000; Lisnic et al., 2005). Palindromes above a certain threshold may be subject to similar negative selective pressures in higher organisms as they are in *E. coli*.

1.7 Palindrome-mediated genomic instability

In *E. coli*, palindromic sequences appear to stimulate recombination, and can present potential obstacles during DNA synthesis. As a result, they are sites of genetic instability. In the last ten years, work has been carried out on palindromes in eukaryotes indicating similar characteristics. In addition, palindromes in eukaryotes can initiate gene amplification and chromosomal translocation leading to genomic instability.

1.7.1 Palindromes as recombination hotspots in yeast

In *E. coli*, *sbcDC*$^{+}$ strains containing a long interrupted palindrome are only viable if they are recombination proficient, suggesting that palindromes can be sites of DNA DSBs and can stimulate recombination. Palindromes have also been shown to be recombination hotspots in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Farah et al., 2005; Farah et al., 2002; Lobachev et al., 2002; Nasar et al., 2000). The eukaryotic SbcCD homolog, MRN seems to have a role in the formation of a DSB at a palindrome during meiosis in *S. pombe* (Farah et al., 2005) and MRX in the processing of a DSB at a palindrome during mitosis in *S. cerevisiae* (Lobachev et al., 2002). Furthermore, stimulation of recombination in *S. pombe* can lead to the elimination of a palindrome by gene conversion (Farah et al., 2005). This process might contribute to the elimination of potentially unstable sequences in yeast (and other eukaryotes).

1.7.2 Gene amplification at palindromes

Palindromic sequences are frequently the mediators of gene amplification in eukaryotes (Butler et al., 1995; Tanaka et al., 2005; Tanaka et al., 2002). In some

organisms such as the ciliated protozoan *Tetrahymena thermophila*, gene amplification is an important part of the life cycle (Butler et al., 1995), in other organisms this process can be detrimental and can lead to progression of cancer by amplification of oncogenes (Lengauer et al., 1998) or the development of drug resistance in advanced tumours (Stark et al., 1989).

1.7.2.1 Palindrome-mediated gene amplification in *Tetrahymena thermophila*

As mentioned above, during macronuclear development in *T. thermophila* an 84 bp DNA palindrome plays an important role in the amplification of the ribosomal RNA (rDNA) gene (see figure 1.7). The DNA fragment containing the rDNA gene (and the 84 bp palindrome) is excised from its chromosomal location in the micronucleus and a telomere is synthesised at the 3' end of this molecule (Yao, 2002). The 84 bp palindrome is near to the 5' end of the molecule and by an unknown mechanism intramolecular pairing generates a hairpin end and ligase seals the nick to generate a large DNA hairpin molecule. Bi-directional replication of this giant hairpin molecule forms a 20 kb palindrome with two functional telomeres and two copies of the rDNA gene (Butler et al., 1995). Subsequent amplification of this palindrome to approximately 10,000 copies per cell completes the amplification process.

1.7.2.2 Palindrome-mediated gene amplification in extrachromosomal DNA

Similar mechanisms that lead to the generation of large palindromes in *Tetrahymena* also exist in other organisms such as the eukaryote *S. cerevisiae* and the prokaryote *Streptomyces lividans* (Butler et al., 1996; Qin and Cohen, 2000). Large palindrome

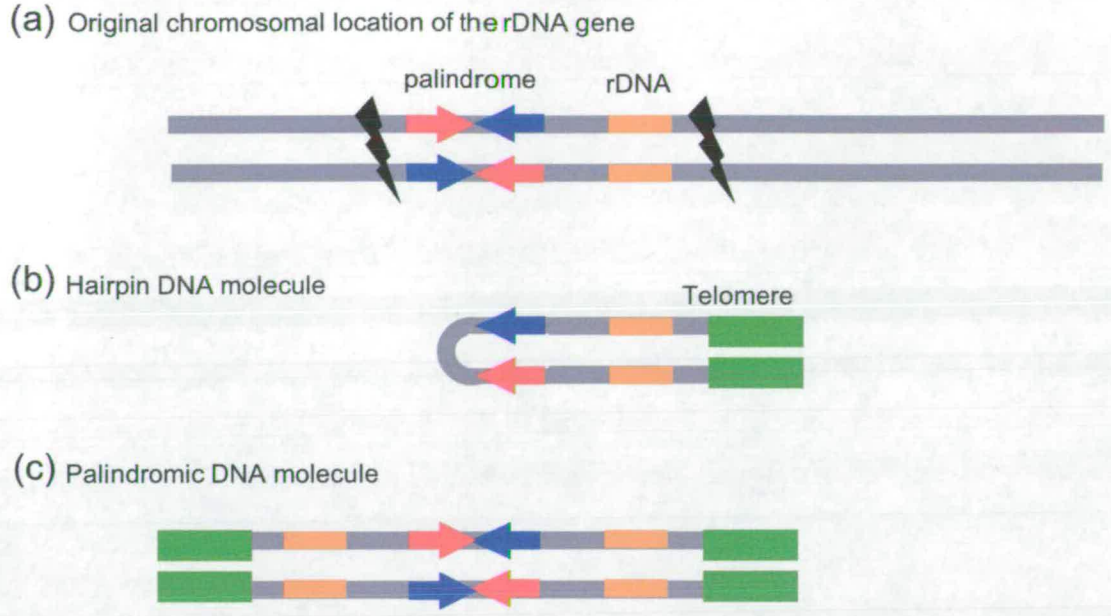


Figure 1.7: rDNA gene amplification at a palindrome in *Tetrahymena thermophila*. (a) The rDNA gene is excised from its chromosomal position in the micro-nucleus, (b) a telomere is synthesised onto one end and DNA processing of the other end generates ssDNA containing a palindrome that forms intra-strand base pairing, (c) bidirectional replication leads to the generation of a large palindrome. This palindromic molecule is amplified to about 10,000 copies per cell to complete the process.

formation has even been observed in *E. coli* (Lin et al., 2001). Butler and collaborators found that induction of a DNA break, in an extra-chromosomal DNA construct, near an 84 bp palindrome (or a 38 bp imperfect palindrome) would lead to the formation of large DNA palindromes in *S. cerevisiae* (Butler et al., 1996).

Likewise, work carried out by Qin and Cohen in the prokaryote *Streptomyces lividans* (a gram positive bacterium with linear chromosomes) demonstrated that large palindromic plasmids were obtained following transformation with a plasmid with only one functional telomere. Analysis of these palindromic structures revealed that the initiating palindromic sequences were between 7 and 19 bp long and could be several kilobases away from the broken DNA end. They suggest that, following a DNA break, the resultant end is processed to generate a 3' ssDNA end and allows short palindromic sequences to form intrastrand base pairs (see figure 1.8). DNA synthesis from the 3' paired-end of the short palindrome (following exonuclease degradation of the unpaired 3' strand) and ligation leads to the generation of a large hairpin structure that becomes a large palindrome following bi-directional replication (Qin and Cohen, 2000). In fact, work by Lin and collaborators isolated large circular palindromic plasmids (head-to-head dimers) following transformation of *E. coli* strains with a linearised plasmid containing palindromic sequences at each end (Lin et al., 2001). These palindromic plasmids were formed following replication of a linear molecule capped at both ends with a hairpin structure.

1.7.2.3 Break-induced replication (BIR) in yeast

The studies discussed above demonstrate that extra-chromosomal DNA molecules can generate large palindromes following a DNA DSB near a short palindrome

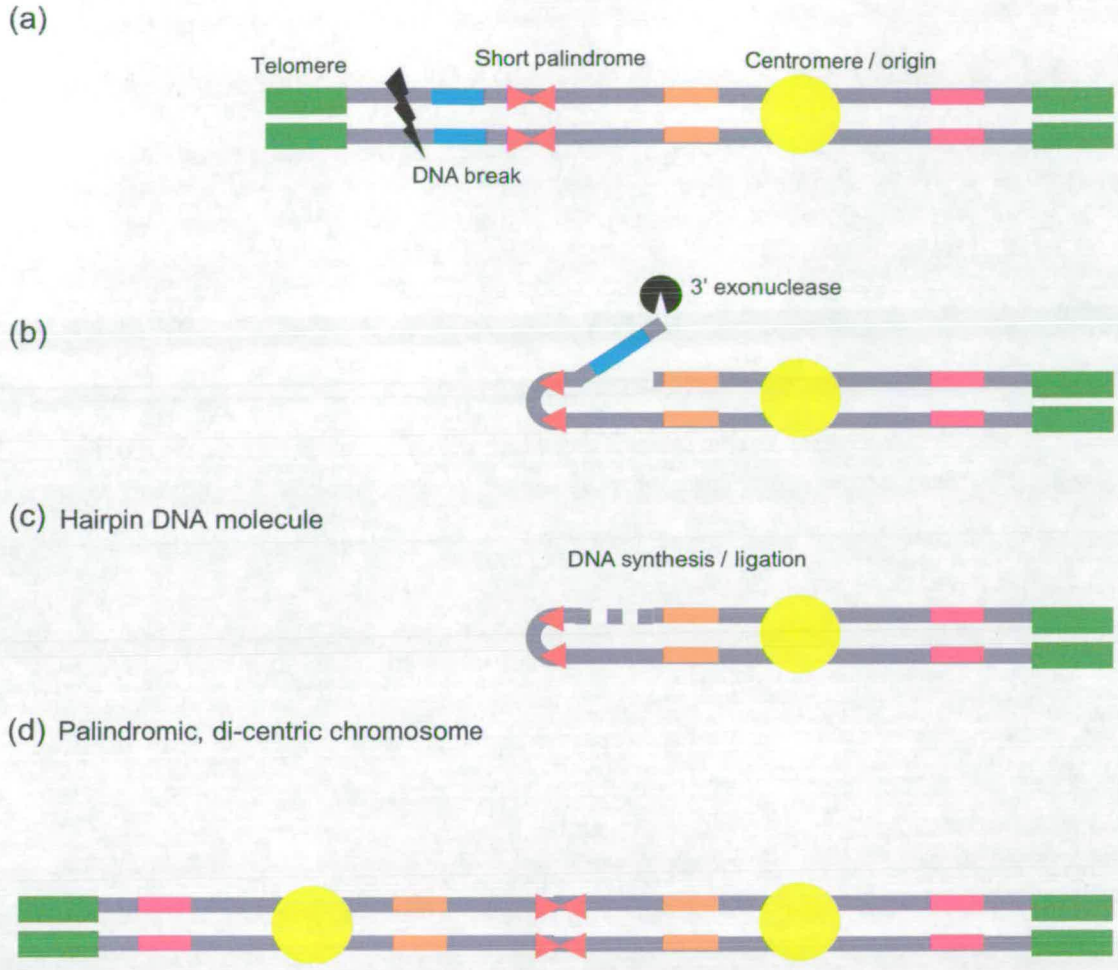


Figure 1.8: Palindrome formation from episomal DNA in *Streptomyces lividans* (Qin and Cohen 2000). (a) The chromosome is broken and processed to leave a 3' ssDNA overhang, (b) a palindromic sequence (as small as 7 bp) forms intrastrand base-pairs and the unpaired 3' overhang is degraded (c) DNA synthesis initiated from the 3' end of the palindrome followed by ligation generates a giant hairpin molecule, (d) bi-directional replication forms a large palindromic molecule with some gene amplification as well as loss of some genetic information. In this diagram the DNA palindrome is shown with two origins of replication but this could also be a dicentric molecule in eukaryotes.

(Butler et al., 1996; Lin et al., 2001; Qin and Cohen, 2000). In these situations large palindromic molecules were formed at the expense of loss of non-essential DNA from the small extra-chromosomal constructs. Presumably this would act as a selective pressure against this occurring in chromosomes.

Rattray and collaborators constructed a reporter cassette for studying palindrome-formation in one of the chromosomes of *Saccharomyces cerevisiae* (Rattray et al., 2005). They demonstrated a mode of gene amplification and large palindrome formation in the region of a HO-induced DNA DSB that does not lead to loss of chromosomal DNA. They propose a model of break-induced replication (BIR) that is instigated when a DNA sequence near the break invades a homologous sequence in an inverted position towards the telomere of the same chromosome (figure 1.9). This invasion is followed by the capture of the non-invading DNA end and subsequent replication from this strand (figure 1.9 (b) and (c)). Because the event is within the same chromosome, the broken ends are still present and hairpin formation at a short palindrome (of 4 – 6 bp), near the end of one of the molecules could join the two broken molecules. Replication of the other end followed by Holliday junction migration in the same direction would complete the formation of a large palindrome (inside the chromosome), with gene amplification, without the loss of genetic information (figure 1.9 (d)). It is possible that large palindromes could lead to further genome instability by driving further BIR cycles and subsequent gene amplification. Similarly, a study by Lobachev and collaborators observed gene amplification following a break near a palindromic sequence (Lobachev et al., 2002).

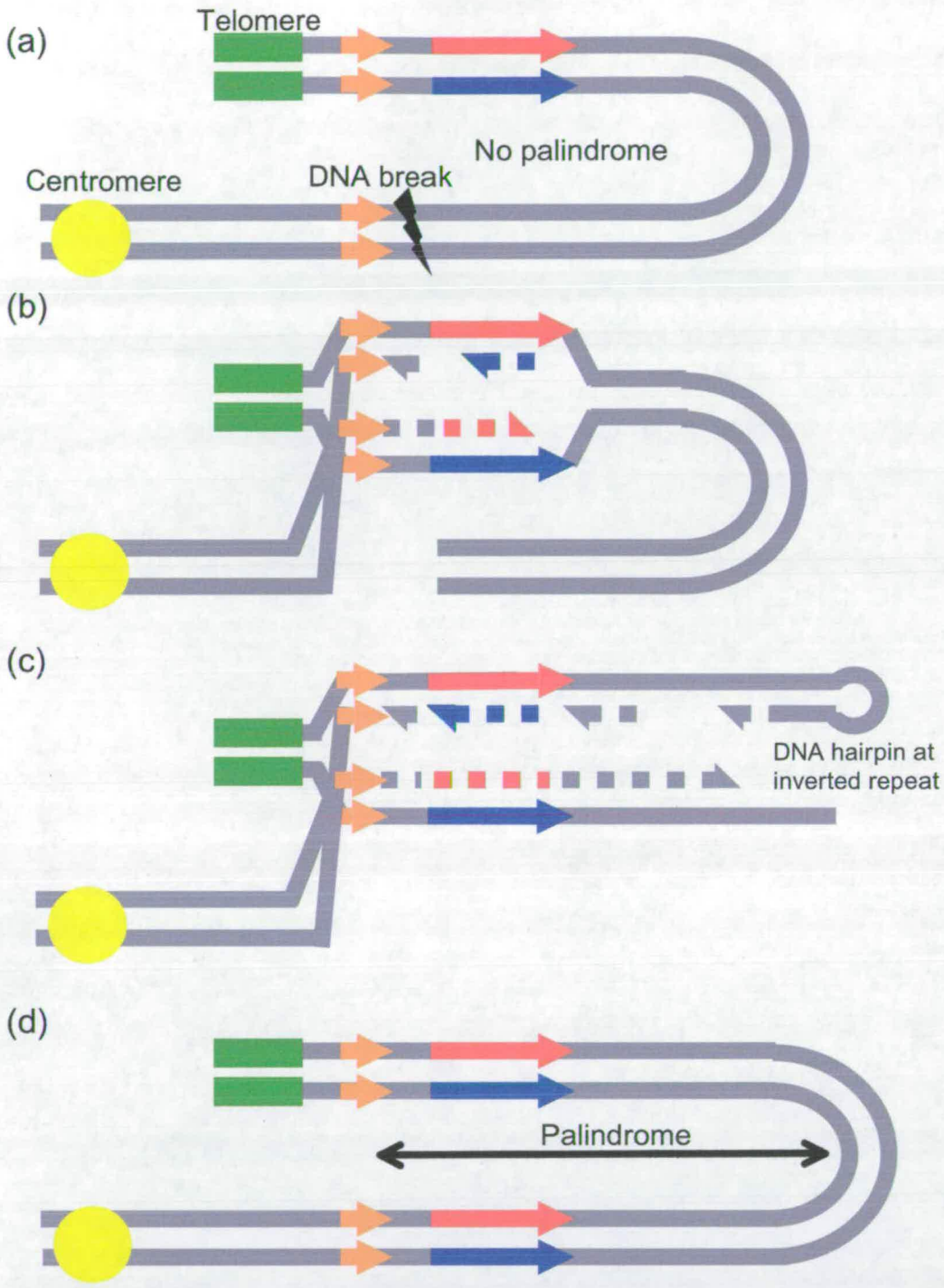


Figure 1.9: Palindrome formation by break-induced replication in *Saccharomyces cerevisiae* (Rattray, Shafer et al. 2005). (a) a DNA break occurs near a dispersed repeat (orange arrow), (b) repair follows intra-molecule invasion of the other indirect repeat and replication is initiated, (c) short inverted repeats lead to hairpin end formation and the Holliday junction is migrated towards this hairpin, (d) the repair process and generation of a large palindrome is complete by sealing of the replicated DNA ends opposite the short inverted repeat

1.7.2.4 Palindrome-mediated gene amplification in higher eukaryotes

Tanaka and collaborators studied gene amplification by large palindrome formation in Chinese Hamster Ovarian (CHO) cells. A break caused by the I-*Sce*I enzyme near a short palindrome was shown to frequently stimulate large palindrome formation and amplification of a nearby drug resistance gene (Tanaka et al., 2002). When the selection pressure for drug resistance was increased, further intra-chromosomal gene amplification (up to 20 copies of the gene) was observed.

The bridge-fusion-break (BFB) cycle offers an explanation for this high level of amplification (see figure 1.10). Formation of a dicentric chromosome can lead to a DNA DSB at cell division as centromeres are pulled to opposing poles. This can lead to further gene amplification and formation of another dicentric chromosome if newly replicated, broken sister homologues are joined by non-homologous end joining.

It is probable that the models for gene amplification seen in the study of CHO cells could be applied to gene amplification during development of human cancers where immortal cells might be selected for their fast growth following oncogene amplification, or increased drug resistance. A combination of BIR and BFB cycles may contribute to the process of gene amplification.

1.7.2.5 Palindromes in human cancers

Tanaka and collaborators screened human cancer cell lines for palindromic DNA and revealed that cells originating from a colorectal cancer cell line and a breast cancer

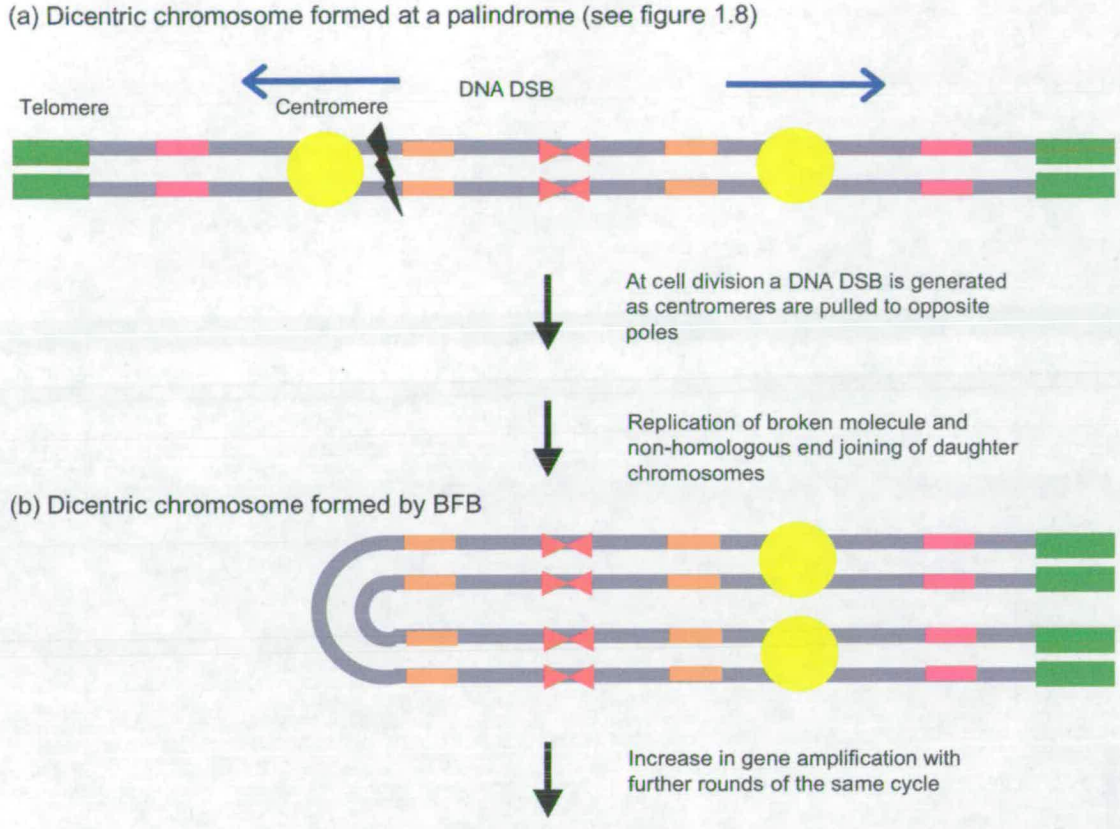


Figure 1.10: Gene amplification by the break-fusion-bridge cycle. (a) Following large dicentric palindrome formation another chromosomal break is generated as centromeres are pulled to opposite poles, replication of the broken chromosome generates two broken sister homologs, which can be joined together by the non-homologous end-joining pathway, (b) another dicentric chromosome with further gene amplification is generated, and the cycle can continue.

cell line contained more palindromic DNA than normal, non-cancerous cell lines. Interestingly, these palindromes are clustered in similar chromosomal regions between the two cell lines. In addition, some of these palindromes were associated with gene amplification. The amplified regions were different between the two cancer types, which may reflect different selection pressures on tumour cells at different locations. DNA was also examined from five independent human medulloblastomas. Here, palindromic DNA was associated with different chromosomal regions (compared to colorectal cancer or breast cancer cell lines) and some of these regions were associated with gene amplification (Tanaka et al., 2005). This study did not distinguish between palindromes formed (1) as a result of a cell becoming cancerous or (2) as germline rearrangements that led to development of the cancer. It remains unclear if palindromic sequences contribute to a predisposition for cancer development, arise as tumour cells evolve, or if cancer develops by a combination of the two (Tanaka et al., 2005).

Tanaka and collaborators suggest that palindrome formation might be associated with fragile chromosomal regions that are susceptible to break (Tanaka et al., 2005). Ciullo and collaborators have reported a case of gene amplification and palindrome formation at a chromosome fragile site. This appears to be a stable gene amplification consistent with the BIR mode of amplification (Ciullo et al., 2002).

1.7.3 Mechanisms for stabilizing palindromic sequences

Palindromic sequences can lead to erroneous gene amplification and subsequent genome instability and potentially contribute to cancer progression. Therefore, cells

have mechanisms that help prevent formation of palindromes or that recognise and stabilize palindromes.

1.7.3.1 MRX (SbcCD homologue) prevents formation of palindromes and leads to elimination of palindromes in yeast

Lobachev and collaborators looked at recombination during mitosis at an interrupted palindrome in *S. cerevisiae* (Lobachev et al., 2002). They found that the palindrome was highly unstable, and a hotspot for DNA DSBs. The observed recombination was dependent on the MRX complex and Sae2p. These proteins were shown to have a role in processing of hairpin-capped DNA ends after the formation of a DSB. In the absence of any single component of the MRX complex or Sae2p, the hairpin-capped ends stimulate inverted duplication (Lobachev et al., 2002). Similarly, Rattray and collaborators also observed a high rate of break-induced, large palindrome formation in *S. cerevisiae* cells defective for Sae2p or the MRX complex (Rattray et al., 2005).

Both groups also report palindrome instability (Lobachev et al., 2002) or palindrome-mediated instability of chromosomes (Rattray et al., 2005) in strains lacking Sae2p. In *S. cerevisiae,* Sae2p may be required to stimulate MRX hairpin cleavage activity and prevent palindrome-mediated duplication and large palindrome formation. If Sae2p/MRX fails to prevent palindrome formation it may have a role in stabilizing palindromes by a similar mode of hairpin cleavage.

Studies in *S. pombe* by Farah and collaborators show that palindromic sequences are targets for MRN-cleavage during meiosis and repair of the resultant DSBs by gene conversion is shown to eliminate palindromic sequences (Farah et al., 2005).

1.7.3.2 Stabilizing palindrome rearrangements by a centre-break mechanism

Collick and collaborators investigated the stability of two palindromic sequences in transgenic mice (Collick et al., 1996). They demonstrated that these palindromic sequences had substantial somatic instability with high rates of central or complete palindrome deletions. Similarly, Lewis and collaborators demonstrated that, in transgenic mice, palindromes are hotspots for homologous recombination (Lewis et al., 1999). Additionally, they observed frequent stabilizing, central-deletion rearrangements of these palindromes in the germ cell line. Cunningham and collaborators established somatic cell lines from transgenic mice with palindromes (Cunningham et al., 2003). A cell line containing a palindrome had a mutation rate at least 25 times greater than a cell line containing a palindrome that had undergone a stabilizing mutation. The majority of stabilizing mutations were due to deletion of the central region of the palindrome and could be explained by a model in which a palindrome is broken at the centre (following cruciform extrusion) and the broken ends are rejoined imprecisely (maybe repair by non-homologous end joining). The observations in germ lines and somatic cells are very similar and it is possible that the same mechanism of stabilisation may act in both somatic and germ cell lines.

1.7.4 Palindromes as sites of translocation

In higher eukaryotes palindromes are subject to stabilizing, central-deletion mutations. It has been suggested that this takes place by a centre-break mechanism whereby a cruciform structure is cleaved and then rejoined by an imprecise repair-system such as in non-homologous end joining. This repair could lead to

chromosomal translocation if two breaks occurred in a cell, simultaneously and were rejoined to the incorrect partner.

In humans, the chromosomal region 22q11.2 has been shown to be involved in two recurrent constitutional translocations t(11;22) and t(17;22), and a third non-recurrent translocation t(4;22). It is also involved in a balanced ependymoma-associated translocation t(1;22). In each of these translocations, the break-point in chromosome 22 has been traced to the centre of an almost perfect palindrome (Edelmann et al., 2001; Gotter et al., 2004; Kurahashi and Emanuel, 2001; Kurahashi et al., 2003; Nimmakayalu et al., 2003). Additionally, Gotter and collaborators have described the break-point locations of the other chromosomes involved in these translocations. They also lie in regions of large inverted repeats with varying degrees of mismatch. It is apparent that regions of recurrent translocation have palindromic sequences resembling the break-point region of chromosome 22. In fact, there is a correlation between the similarity of the physical properties of the breakpoint sequences (in terms of GC-content, melting temperature and the propensity for the sequence to form secondary structures) and the frequency of a translocation. For example, the t(4;22) translocation is non-recurrent and the palindrome at the 4q35.1 break-point site has a large central-spacer region of 547 nt, which probably reduces the effectiveness of the palindromic sequence to form a secondary structure (and be a target for breakage). The palindrome-associated translocations in the 22q11.2 region are associated with palindromes at the reciprocal loci, and the starting point for these translocations could be a mechanism that directs cleavage at palindromes in order to stabilize them.

1.8 The scope of this thesis

This thesis is concerned with recombination stimulated by a 246 bp interrupted palindrome in the *E. coli* chromosome. Chapter three describes the construction of the main tool of this work, a strain carrying the *sbcDC* operon under the control of an arabinose inducible promoter (P_{BAD}-*sbcDC*). Lysogenisation experiments, using a bacteriophage lambda containing a 246 bp palindrome and the P_{BAD}-*sbcDC* strain grown under $SbcCD^+$ and $SbcCD^-$ conditions, reveal some of the genetic requirements for palindrome viability. Chapter four describes further genetic analyses in a lambda-free system using the P_{BAD}-*sbcDC* strain containing a 246 bp palindrome inserted at the *lacZ* locus. Palindrome viability in an $SbcCD^+$ strain requires genes from the RecBCD recombination pathway (but not the RecFOR pathway), the PriA protein for re-establishment of replication, genes involved in SOS-induction (RecA and LexA) and finally, there is some requirement for XerCD-mediated chromosome dimer resolution. Chapter five describes the physical analysis of the DNA surrounding the palindromic sequence and provides evidence that two-ended DNA double-strand breaks are generated following expression of SbcCD. Finally, chapter six describes the construction of a set of recombination cassettes for measuring rates of recombination either side of a palindrome in both *sbcDC*$^+$ and Δ*sbcDC* strains. These experiments demonstrate that a palindrome is a hotspot for recombination in an *sbcDC*$^+$ strain. In addition, a modest stimulation of recombination is observed at a palindrome in a Δ*sbcDC* strain suggesting that recombination can be stimulated by SbcCD-independent pathway.

Chapter 2: Materials and methods

2.1 Materials

2.1.1 Stock solutions

50 mg ml^{-1} Chloramphenicol (Calbiochem)

Made up to 50 mg ml^{-1} in 100% ethanol and stored at -20 °C. It was used at a concentration of 50 μg ml^{-1}.

10 mg ml^{-1} Gentamicin (Sigma)

Made up to 10 mg ml^{-1} in sterile Milli-Q water and stored at -20 °C in 1 ml aliquots. It was used at a concentration of 10 μg ml^{-1}.

50 mg ml^{-1} Kanamycin (Calbiochem)

Made up to 50 mg ml^{-1} in sterile Milli-Q water and stored at -20 °C in 1 ml aliquots. It was used at a concentration of 50 μg ml^{-1}.

10 mg ml^{-1} Spectinomycin (Sigma)

Made up to 10 mg ml^{-1} in sterile Milli-Q water and stored at -20 °C in 1 ml aliquots. It was used at a concentration of 50 μg ml^{-1}.

15 mg ml^{-1} Tetracyclin (Calbiochem)

Made up to 15 mg ml^{-1} in 50% ethanol and stored at -20 °C. It was used at a concentration of 15 μg ml^{-1}.

100mg ml^{-1} Zeocin (Invitrogen)

Obtained as a 100 mg ml^{-1} solution and stored at -20 °C in 0.2 ml aliquots in the dark. It was used at a concentration of 35 μg ml^{-1} using low-salt (5 g l^{-1} NaCl) media (see low-salt LB agar 2.1.2).

0.1 M or 0.5 M $CaCl_2$

Made up fresh to 0.1 M or 0.5 M in Milli-Q water and sterilised using a 0.45 μm syringe filter.

1 M Sodium Citrate ($Na_3C_6H_5O_7 \bullet 2H_20$)

Made up to 1 M using Milli-Q water and autoclaved. Stored at room temperature.

10 mM $FeCl_3$

Made up to 10 mM using Milli-Q water and autoclaved. Stored at room temperature.

1 M $MgSO_4$

Made up to 1 M using Milli-Q water and autoclaved. Stored at room temperature.

5 mg ml^{-1} Vitamin B1

Made up to 5 mg ml^{-1} using Milli-Q water and sterilised using a 0.45 μm syringe filter. Stored at 4°C.

3 mg ml^{-1} Proline

Made up to 3 mg ml^{-1} using Milli-Q water and sterilised using a 0.45 μm syringe filter. Stored at 4°C.

2.5 mg ml^{-1} Leucine

Made up to 2.5 mg ml^{-1} using Milli-Q water and sterilised using a 0.45 μm syringe filter. Stored at 4°C.

20% (w/v) Glucose

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

20% (w/v) Arabinose

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

20% (w/v) Maltose

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

20% (w/v) Sucrose

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

1 M Tris

Made up to 1 M solution using Milli-Q water and adjusted to desired pH using concentrated hydrochloric acid and autoclaved. Stored at room temperature.

0.5 M EDTA (pH 8)

Made up to 0.5 M solution using Milli-Q water and adjusted to pH 8.0 using NaOH and then autoclaved. Stored at room temperature.

5 M NaCl

Made up to 5 M solution in Milli-Q water and then autoclaved. Stored at room temperature.

1 mg ml^{-1} Ethidium bromide

Made up to 1 mg ml^{-1} using Milli-Q water and stored at room temperature in the dark. Diluted to a 0.5 μg ml^{-1} working concentration with Milli-Q water.

80% (v/v) Glycerol

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

---

*Autoclaving is 115°C for 20 minutes for sterilising sugar solutions (compared to the usual 121°C for 15 minutes).

500 mg ml^{-1} Isopropyl-β-D-Thiogalactopyranoside (IPTG) (Melford Labs.)

Made up to 500 mg ml^{-1} in sterile Milli-Q water and stored at -20°C in 1 ml aliquots. Used at a concentration of 500 μg ml^{-1}.

50 mM Phenylmethanesulphonyl fluoride (PMSF)

Made up to 50 mM in isopropanol and stored at -20°C. Used at a concentration of 1 mM.

2.1.2 Media and Buffers

LB agar

10 g Bacto-tryptone (Difco), 5 g yeast extract (Difco), 10 g NaCl and 15 g Bacto-agar (Difco) made up to 1 litre with distilled water and adjusted to pH 7.5 using NaOH and autoclaved.

Low-salt LB agar

10 g Bacto-tryptone (Difco), 5 g yeast extract (Difco), 0.5 g / 5 g NaCl and 15 g Bacto-agar (Difco) made up to 1 litre with distilled water and adjusted to pH 7.5 using NaOH and autoclaved.

LC agar

10 g tryptone, 5 g yeast extract, 5 g NaCl and 10 g Difco-agar made up to 1 litre with distilled water and adjusted to pH 7.2 with NaOH and autoclaved.

LC top agar

10 g tryptone, 5 g yeast extract, 5 g NaCl and 7 g Difco-agar made up to 1 litre with distilled water and adjusted to pH 7.2 with NaOH and autoclaved.

BBL agar

10 g Trypticase-peptone, 5 g NaCl, 10 g agar (Difco) made up to 1 litre with distilled water and autoclaved.

BBL top agar

10 g Trypticase-peptone, 5 g NaCl, 6.5 g agar (Difco) made up to 1 litre with distilled water and autoclaved.

Minimal agar

20 g Oxoid agar No. 3 made up to 1 litre with distilled water and autoclaved. Immediately before use 25% (v/v) spitzizen salts added.

L broth

10 g bacto-tryptone (Difco), 5 g yeast extract (Difco) and 10 g NaCl made up to 1 litre with distilled water and autoclaved.

Low-salt L broth

10 g bacto-tryptone (Difco), 5 g yeast extract (Difco) and 0.5 g / 5 g NaCl made up to 1 litre with distilled water and autoclaved.

Minimal liquid medium

Using sterile stock solutions and sterile water a solution of 1 x M9 salts, 0.2 % (v/v) glycerol, 0.002 % (w/v) $MgSO_4$, 0.0001 % (w/v) $CaCl_2$ and 0.5 % (w/v) Casamino acids was made up and used immediately.

Spizizen salts

10 g $(NH_4)_2SO_4$, 70 g K_2HPO_4, 30 g KH_2PO_4, 5 g $Na_3C_6H_5O_7$ and 1 g $MgSO_4$ made up to 1 litre with distilled water and autoclaved.

Phage buffer

3 g KH_2PO_4, 7 g Na_2HPO_4, 5 g NaCl, 1 mM $MgSO_4$, 1 mM $CaCl_2$ and 1% (w/v) gelatine made up to 1 litre with distilled water and autoclaved.

4 x M9 salts

28 g Na_2HPO_4, 12 g KH_2PO_4, 2 g NaCl, 4 g NH_4Cl made up to 1 litre with distilled water and autoclaved.

TM buffer

Made up to 10 mM Tris (using 1 M Tris pH 7.5) and 10mM $MgSO_4$ (using 1 M $MgSO_4$) using Milli-Q water. Sterilised using a 0.45 μm syringe filter. Stored at room temperature.

50 x Tris-acetate (TAE)

242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA (pH 8.0) made up to 1 litre with Milli-Q water. Stored at room temperature. Diluted to 1 x working concentration using Milli-Q water.

5 x Tris-borate (TBE)

55 g Tris base, 27.6 g boric acid, 10 ml 0.5 M EDTA (pH 8.0) made up to 1 litre with Milli-Q water. Stored at room temperature. Diluted to 0.5 x working concentration using Milli-Q water.

TE buffer

Made up to 10 mM Tris (using 1 M Tris pH 7.6) and 1 mM EDTA (using 0.5 M EDTA pH 8) in Milli-Q water. Stored at room temperature

TEE buffer

Made up to 10 mM Tris (using 1 M Tris pH 9), 100 mM EDTA and 100 mM EGTA using Milli-Q water. Stored at room temperature.

Lysozyme solution

Made up to 0.05% (w/v) sarcosyl and 5 mg ml^{-1} lysozyme in TEE buffer. Made up fresh immediately before use.

Proteinase K solution

Made up to 1 mg ml^{-1} proteinase K and 1% (w/v) SDS in TEE buffer. Made up fresh immediately before use.

PMSF solution

Made up to 1 mM PMSF in TE buffer. Made up fresh immediately before use.

Depurination solution

Made up to 250 mM HCl in Milli-Q water. Stored at room temperature.

Denaturation solution

Made up to 0.5 M NaOH and 1.5 M NaCl (using 5 M NaCl solution) in Milli-Q water. Stored at room temperature.

Neutralisation solution

Made up to 0.5 M Tris (using 1 M Tris pH 7.5) and 1.5 M NaCl (using 5 M NaCl) in Milli-Q water. Stored at room temperature.

20 x SSC

Made up to 3 M NaCl and 300 mM sodium citrate in Milli-Q water and adjusted to pH 7.0 using concentrated NaOH and autoclaved. Stored at room temperature. Diluted appropriately in Milli-Q water.

Low stringency buffer

Made up to 2 x SSC and 0.1% (w/v) SDS in Milli-Q water. Stored at room temperature.

High stringency buffer

Made up to 0.1 x SSC and 0.1% (w/v) SDS in Milli-Q water. Stored at room temperature.

Maleic acid buffer

Made up to 0.1 M Maleic acid and 0.15 M NaCl in Milli-Q water and adjusted to pH 7.5 using concentrated NaOH and autoclaved. Stored at room temperature.

Washing buffer

Made up to 0.3% (w/v) Tween 20 in maleic acid buffer. Stored at room temperature.

Detection buffer

Made up to 0.1 M Tris, 0.1 M NaCl and adjusted to pH 9.5 with concentrated NaOH. Stored at room temperature.

Stripping buffer

Made up to 0.2 M NaOH and 0.1% (w/v) SDS in Milli-Q water.

100x Deoxynucleotide triphosphate (dNTPs) mix

Individual 100 mM dATP, dCTP, dGTP and dTTP (Promega; catalogue number U1330) were mixed in equal volumes to give a concentration of 25 mM of each. Stored at -20°C in 50 µl aliquots.

2.2 General Methods

2.2.1 Bacterial methods

2.2.1.1 Overnight cultures

Bacterial cultures were made by inoculating 5 ml L broth (with relevant additives) with a single isolated colony. They were incubated at 37°C (unless otherwise stated) with shaking over night.

2.2.1.2 Storage of bacteria

0.75 ml of an overnight culture was mixed with 0.75 ml 80% (v/v) glycerol in a 1.5 ml eppendorf tube and stored at -80°C.

2.2.1.3 Serial dilutions

Bacterial cultures were diluted using 1 x M9 salts.

2.2.1.4 UV irradiation

In order to test whether a strain was UV sensitive following introduction of a *rec*$^-$ or *ruv*$^-$ mutation overnight cultures of the parental strain (*rec*$^+$) and the *rec*$^-$ strain were grown up and were serially diluted and plated on LB agar plates. These plates were exposed to different amounts of UV light (ranging between 0 and 14 milli Joules) using a UV Stratalinker. The plates were incubated at 37°C overnight in the dark. UV sensitive strains were less viable when exposed to UV light.

2.2.2 Phage P1 transduction methods

P1 is a general transducing phage that requires an *E. coli* host for its propagation and can be used to transfer a chromosomal marker gene from one strain of *E. coli* to another. These methods take advantage of the tendency for a proportion (about 3 %) of phage P1 to package host *E. coli* chromosomal DNA (about 100 kb) into the

phage head during its lytic growth cycle. If another strain of *E. coli* is then infected by such a phage particle containing *E. coli* DNA with the marker gene then this region can replace the homologous chromosomal region by RecA-dependent recombination (Neidhert F.C.).

2.2.2.1 Preparing a P1 lysate

An overnight culture of the strain of *E. coli* containing the marker gene of interest was prepared (as described in 2.2.1) and diluted 10-fold in 5 ml of L broth containing 2.5 mM $CaCl_2$. This was grown at 37°C with shaking for 2 hours. 0.2 ml was mixed with 0.1 ml of a P1 lysate of approximately 10^6 pfu ml^{-1} and incubated for 30 minutes at 37°C with gentle shaking to allow phage adsorption. 2.5 ml of melted LC top agar containing 5 mM $CaCl_2$ (cooled to approximately 45°C) was added to the cell-phage mix and poured onto a LC agar plate containing 5 mM $CaCl_2$. The plates were incubated upright overnight at 37°C. The phage were collected from the plate by removing the top agar in 5 ml of phage buffer and vortexing with 100 µl chloroform. They were left at 4°C for 30 minutes, the agar was removed by centrifugation (5,000 rpm for 10 minutes at room temperature) and the supernatant transferred to a fresh bottle containing 100 µl of chloroform. P1 lysates were stored at 4°C in the dark.

2.2.2.2 Titration of a P1 lysate

An overnight culture of *E. coli* was prepared and 200 µl was mixed with 2.5 ml of LC top agar and poured onto a LC agar plate (containing 2.5 mM of $CaCl_2$). The lysate was diluted in phage buffer (10^{-1} to 10^{-7}) and 10 µl spots of the phage were plated onto the lawn of cells. The plates were incubated overnight at 37°C and the plaque forming units (pfu) were counted.

2.2.2.3 P1 transduction for novel strain construction

An overnight culture of the receiving strain of *E. coli* was prepared in L broth plus 2.5 mM $CaCl_2$. 1 ml aliquots were spun down by centrifugation (top speed in a micro centrifuge for 1 minute) and re-suspended in 100 µl fresh L broth plus 2.5 mM $CaCl_2$. This was mixed with 100 µl of P1 lysate and incubated at 37°C with gentle shaking for 30 minutes. 800 µl of L broth plus 2mM sodium citrate was added and this was incubated at 37°C with gentle shaking for 1 hour. 200 µl of the cells were spread onto LB agar plates supplemented with the appropriate selective agent and incubated at 37°C overnight. To purify away from residual phage particles successful transductants were streaked onto fresh selective LB agar plates twice.

2.2.3 DNA techniques

2.2.3.1 Preparing total-cell DNA for PCR (Basic method)

A single isolated colony of the desired strain was picked from a LB agar plate and suspended in 30 µl of sterile water in a 0.5 ml eppendorf tube. The cell suspension was heated at 95°C for 10 minutes in a thermocycler to lyse the cells and release the DNA. The cell debris was then spun down for 3 minutes at full speed in a micro centrifuge. 2 µl of supernatant was used in a 25 µl PCR reaction.

2.2.3.2 Preparing genomic DNA (Kit method)

For applications requiring pure genomic DNA the Wizard® Genomic DNA purification kit was used (Promega; catalogue number A1120). The manufacturers instructions where followed and DNA was rehydrated overnight at 4°C.

2.2.3.3 Preparing plasmid DNA (Kit method)

For applications requiring pure plasmid DNA the QIAprep® Spin Miniprep Kit was used (QIAGEN; catalogue number 27104). For high-copy number plasmids the manufacturers instructions were followed. For low-copy number plasmids DNA was isolated from a 15 ml overnight culture and eluted in 30 µl of elution buffer.

2.2.3.4 Checking strains using the Polymerase Chain Reaction (PCR)

The thermostable DNA polymerase from *Thermus aquaticus* (*Taq*) was used to check strains for presence of particular DNA constructs (Promega; catalogue number M1661).

A typical PCR reaction using *Taq* polymerase shown below:

0.5 – 2 µl	DNA Template*
0.25 µl	100 x dNTP mix (final concentration of 0.25 mM of each NTP)
2.5 µl	10 x Mg^{2+}-free Reaction buffer
2.5 µl	10 x $MgCl_2$ solution (final concentration of 2.5 mM)
1 µl	Forward primer (final concentration of 0.2 mM)
1 µl	Reverse primer (final concentration of 0.2 mM)
0.15 µl	*Taq* polymerase (0.75 units total)
15.6 – 17.1 µl	Sterile water (to a total volume of 25 µl)

* 2 µl boiled-cell/ 0.5 µl genomic (100 ng μl^{-1}) / 0.5 µl plasmid (30 ng μl^{-1})

Temperature cycling was carried out using Applied Biosytems GeneAmp PCR System 2400 or Hybaid PCR Express thermo-cyclers. A typical PCR programme was as follows (were T_m = the melting temperature of the primer):

1		95°C	5 minutes	x 1
2	i	95°C	30 seconds	
	ii	T_m - 5 °C	15 seconds	
	iii	72°C	1 minute per kilobase of DNA	x 30
3		72°C	10 minutes	x 1
4		4°C	∞	

2.2.3.5 PCR using a proof-reading polymerase

The polymerase with proof-reading properties from *Pyrococcus furiosus* (*Pfu*) was used for applications that require cloning of a PCR fragment (Promega; catalogue number M7741).

A typical PCR reaction using *Pfu* polymerase:

0.5 µl	DNA template**
0.5 µl	100 x dNTP mix (final concentration of 0.25 mM of each NTP)
5 µl	10 x Reaction buffer with $MgSO_4$
2 µl	Forward primer (final concentration of 0.2 mM)
2 µl	Reverse primer (final concentration of 0.2 mM)
0.3 µl	*Pfu* polymerase (1 unit total)
39.7 µl	Sterile water (to a total volume of 50 µl)

** 0.5 µl genomic (100 ng μl^{-1}) / 0.5 µl plasmid (30 ng μl^{-1})

Temperature cycling was carried out using Applied Biosytems GeneAmp PCR System 2400 or Hybaid PCR Express thermo-cyclers. A typical PCR programme was as follows (where T_m = the melting temperature of the primer):

1		95°C	5 minutes	x 1
2	i	95°C	30 seconds	
	ii	T_m - 5 °C	15 seconds	
	iii	72°C	2 minute per kilobase of DNA	x 30
3		72°C	10 minutes	x 1
4		4°C	∞	

2.2.3.6 DNA restriction for cloning

Restriction digests of PCR products or purified plasmids were carried out using enzymes and buffers received from New England Biolabs (NEB) or Roche. Reaction conditions vary for different enzymes but in general were set up as follows:

30 µl	Plasmid DNA (30 ng μl^{-1}) / PCR product (25 - 50 ng μl^{-1})
10 µl	10 x Reaction buffer
1 µl	100 x Bovine Serum Albumin (BSA) (only NEB enzymes)
2.5 µl	Restriction enzyme 1 (obtained at 10 – 50 units μl^{-1})
2.5 µl	Restriction enzyme 2
39 - 45 µl	Sterile water

Reactions were left at the recommended incubation temperature for 4 hours.

2.2.3.7 Removing 5' phosphate groups from linear DNA

Calf intestinal alkaline phosphatase (CIP) (New England Biolabs; catalogue number M0290S) was used to remove the 5' phosphates from the ends of linearised DNA vectors prior to using in a ligation. This reduces the number of clones that contained re-ligated original plasmid. Immediately after digestion 1 µl of the enzyme (10 units μl^{-1}) was added to the reaction, and incubated at 37°C for 1 hour.

2.2.3.8 Purification of linear DNA

Linear DNA molecules for cloning (or PCR products for sequencing) were separated from protein and salts using the QIAquick® PCR Purification Kit (QIAGEN; catalogue number 28104). The QIAquick® Gel Extraction Kit was used when DNA products needed to be separated (QIAGEN; catalogue number 28704). In either case the manufacturers instructions were followed.

2.2.3.9 Ligation of DNA molecules

Ligations were carried out using a Quick Ligation™ Kit (New England Biolabs; catalogue number M2200S). Purified, digested DNA was observed on an agarose gel and quantified. Ligation reactions were carried out according to the manufacturers instructions.

2.2.3.10 Transformation of chemically competent *E. coli* cells

An overnight culture of the desired strain was diluted 1 ml in 50 ml in L broth and grown at 37°C for 2 hours with shaking. 1 ml aliquots of cells were spun down by centrifugation (full speed on a micro centrifuge for 1 minute) and gently re-suspended in 0.5 ml of fresh, ice-cold 0.1 M $CaCl_2$ and left on ice for 30 minutes. The cells were spun down as before and gently re-suspended in 0.1 ml of fresh, ice-cold 0.1 M $CaCl_2$. Approximately 20 ng of pure plasmid DNA (about 0.5 µl) or 25

ng (about 10 µl) of ligation mix was added to the cell suspension and left on ice for 30 minutes. The cells were heat-shocked at 37°C for 5 minutes, and placed on ice for 5 minutes. 1 ml of L broth was added and incubated with gentle shaking at 30°C / 37°C (depending on the plasmid) for 2 hours. 200 µl of the cell suspension was spread onto LB agar plates supplemented with appropriate antibiotics and incubated overnight at the appropriate growth temperature.

2.2.3.11 Agarose Gel Electrophoresis

DNA fragments obtained following a PCR reaction or restriction digest where separated and visualised in 1% (w/v) agarose gels. Gels were prepared by melting the appropriate amount of agarose (Invitrogen; catalogue number 15510-027) in 1 x TAE (see 2.1.1). Samples were made up to 20% (v/v) glycerol. Gels were ran between 90 and 130 volts for up to 1 hour. Gels were stained in 0.5 µg ml^{-1} ethidium bromide solution (see 2.1.1) with gentle agitation for 20 minutes and visualised under a UV lamp. DNA ladders were used to estimate sizes and quantities of linear bands (New England Biolabs; catalogue numbers N3200S, N3232S and/or N3231S).

2.2.3.12 Sequencing of PCR products

Sequencing was carried out on purified PCR products (see 2.2.3.8) using BigDye® Terminator v3.1 Cycle-Sequencing Kit according to manufacturers instructions (Applied Biosystems; catalogue number 4337455). Sequencing reactions were processed and analysed by the SBS Sequencing Service, Ashworth Laboratories, University of Edinburgh using an ABI PRISM® 3100-Avant Genetic Analyzer.

2.2.4 Specialised PCR

2.2.4.1 Site-Directed Mutagenesis (SDM)

Site-directed mutagenesis is a PCR based technique used to introduce specific point mutations to a plasmid. The QuickChange® Site-Directed Mutagenesis Kit (obtained from Stratagene; catalogue number 200518) provided instructions for oligo design and PCR reaction conditions. Plasmid template was grown up and isolated from the *dam*$^+$ *E. coli* strain XL1-blue. The proof-reading *Pfu* DNA polymerase was used to amplify the plasmid. Methylated template DNA was digested at 37°C for 4 hours with 20 units of *Dpn*I (obtained from New England Biolabs; catalogue number R0176S). XL1-blue cells were transformed with the nicked, circular PCR products according to the method in 2.2.3.10.

2.2.4.2 Cross-over PCR

This PCR technique allows precise hybridisation of DNA molecules without the use of restriction endonucleases or DNA ligase (described by Horton et al., 1989) Separate PCR reactions are used to create the initial separate DNA fragments. The oligos for amplification are designed so that these DNA molecules have 24 bp of homology to each other at the ends corresponding to the joining point (internal oligos have 12 bp overhang). Using DNA from the previous reactions as template and the external most oligos a second PCR reaction (the 'cross-over' stage) results in annealing of the complementary regions of the PCR products and self-priming to give a novel, joined DNA product. The reactions were set up using *Pfu* polymerase according to section 2.2.3.5. Between PCR stages the PCR products were purified from excess oligo and dNTPs according to section 2.2.3.8. For cross-over stages

equimolar amounts of DNA template were added to the reaction (approximately 20 ng of DNA).

2.2.5 Bacteriophage lambda techniques

2.2.5.1 Preparing a fresh lambda lysate

An overnight culture of the appropriate *E. coli* strain was diluted 1 ml in 10 ml L broth containing 0.2 % (w/v) maltose and 5 mM $MgSO_4$ and incubated at 37°C with shaking for 2 hours. An equal volume of TM buffer was added and 250 µl of this was added to 250 µl of the old bacteriophage lambda lysate (various dilutions were used between 10^{-2} and 10^{-5} in phage buffer). 2 ml of BBL top agar was added and the mixture poured onto a BBL agar plate. This was incubated overnight at 37°C. From these plates a single lambda plaque was picked and suspended in 1 ml TM buffer and 0.5 ml, 0.3 ml and 0.1 ml added to separate 250 µl aliquots of fresh *E. coli* cells (prepared as above). These mixtures were incubated for 30 minutes at 37°C with gentle shaking. 2 ml of BBL top agar was added and the mixture poured onto BBL agar plates containing 10 µg ml^{-1} vitamin B1, 4 µM $FeCl_3$, 0.8 mM $CaCl_2$ and 0.3% (w/v) glucose. These plates were incubated overnight at 37°C. The phage from the plate that had confluent lysis was collected by removing the top agar in 5 ml of phage buffer and vortexing with 100 µl of chloroform. It was left at 4°C for 30 minutes, the agar was removed by centrifugation (5,000 rpm for 10 minutes at room temperature) and the supernatant transferred to a fresh bottle containing 100 µl of chloroform. Lambda phage were stored in the dark at 4°C.

2.2.5.2 Titration of a lambda lysate

An overnight culture of the appropriate *E. coli* strain was prepared in 5 ml L broth containing 2% (w/w) maltose and 5 mM Mg_2SO_4. 0.25 ml was mixed with 2.5 ml of

BBL top agar and poured onto a BBL agar plate. Serial dilutions of the lambda lysate (10^{-1} – 10^{-8}) were prepared using phage buffer and 10 µl spots were plated onto the BBL agar plate. Plates were incubated overnight at 37°C and plaque forming units (pfu) were counted.

2.2.5.3 Purifying lambda DNA from a lysate

DNA was isolated from lambda lysates (as prepared in 2.2.5.1) using QIAGEN® Lambda mini kit (QIAGEN; catalogue number 12523) according to manufacturers instructions.

2.2.5.4 Quantitative lysogenisation using lambda

Overnight cultures of the strains were prepared and diluted 1:10 in 5 ml L broth containing 2% (w/w) maltose and 5 mM Mg_2SO_4 and grown at 37°C with shaking to an OD_{650} of 0.9 (approximately 4 x 10^8 cells ml^{-1}). This was diluted 1:2 in TM buffer. Bacteriophage lambda lysates were diluted to 2 x 10^9 pfu ml^{-1} in phage buffer. 0.15 ml of diluted *E. coli* cells were mixed with 0.15 ml of diluted phage and incubated at 30°C for 1 hour. Infected cells were diluted in M9 salts and plated on L agar plates containing the appropriate antibiotic. For purposes of calculating lysogenisation frequency the cells diluted in TM buffer were further diluted in M9 salts and plated onto L agar plates.

2.2.6 Pulsed-Field Gel Electrophoresis (PFGE)

Conventional agarose gel electrophoresis methods can resolve DNA molecules up to 20 kilo-base pairs (kb). PFGE is a specialised agarose gel electrophoresis that is used to separate DNA molecules of large molecular weights (between 10 kb and 1.5 mega-bases pairs).

2.2.6.1 Preparing *E. coli* DNA for PFGE

Traditional methods of preparing DNA from *E. coli* cultures inevitably results in random shearing of large DNA molecules, to fragments less than 100 kb - this is undesirable if you want to separate and analyse larger molecules. Methods for preparing DNA are adapted from techniques used by Michel and collaborators (Michel et al., 1997). All buffers and solutions used are listed in section 2.1.2.

Overnight cultures were diluted to an optical density at 650 nm (OD_{650}) of 0.01 in 10 ml of appropriate media and supplements and incubated at the appropriate temperature with shaking for between 1 and 2 hours. The OD_{650} of the cultures were measured and the cells were spun down by centrifugation 10 minutes at 3500 rpm at 4°C. The cells were suspended in an appropriate amount of TEE buffer to give a final OD_{650} of 0.9 and 350 µl was mixed with melted 2% (w/v) low melting point (LMP) agarose (Gibco Brl; catalogue number 15517-014) in TEE buffer and allowed to solidify in plug molds. Plugs were incubated at 37°C in 10ml lysozyme solution with gentle shaking for 2 hours and then at 55°C in 5 ml proteinase K solution overnight. All remaining washes were carried out at room temperature and with gentle shaking. The plugs were rinsed with 3 x 10 ml of TE buffer for 3 x 1 hours and rinsed with 2 x 10 ml PMSF solution for 2 x 1 hours and rinsed with 2 x 10 ml TE buffer for 2 x 30 minutes. Plugs were stored in fresh TE buffer at 4°C.

2.2.6.2 Restriction digestion of agarose embedded DNA

Restriction enzymes were obtained from New England Biolabs. Methods were based on instructions provided by the manufacturer; agarose plugs were equilibrated in 1 ml of 1 x reaction buffer for 1 hour at room temperature, this was replaced with 0.3

ml 1 x reaction buffer containing 30 – 50 units of restriction enzyme and incubated for 4.5 hours at the appropriate temperature

2.2.6.3 Running a pulsed-field gel of digested DNA

Plugs were loaded on a 1% (w/v) high-strength agarose gel (AquaPor™ ES; Fisher catalogue number ELR-300-040F) in 0.5 x TBE. The gel was ran in 0.5 x TBE with CHEF-DR™ II PFGE equipment set to the following parameters: switch time 5 to 30 seconds; 160 Volts; included angle 120°; total run time 17 hours at 8°C.

2.2.7 Southern blotting

2.2.7.1 DNA transfer by capillary method

All stages were carried out at room temperature. All solutions are listed in section 2.1.2.

Prior to transfer the DNA gel was submerged first in depurination solution for 10 minutes, then denaturation solution for 2 x 15 minutes, then neutralisation solution for 2 x 15 minutes, finally the gel was equilibrated in 20 x SSC solution for 10 minutes. Between solution changes the gel was rinsed briefly in Milli-Q water.

The DNA was transferred to a positively charged nylon membrane (Roche; catalogue number 1417240) by capillary transfer. A piece of Whatmann 3MM paper was soaked in 20 x SSC and placed on top of a bridge with the edges dipping into a shallow reservoir of 20 x SSC. The agarose gel was placed on top of this (wells facing down) and the membrane placed on top of the gel. A dry sheet of Whatmann 3MM paper, a stack of approximately 7 cm of tissue paper, a glass plate and a 500 g

weight were then placed on top of the membrane to complete the blot. Transfer was allowed to occur overnight. The DNA was fixed to the membrane using the auto-cross linking setting on a UV Stratalinker. The membrane was briefly rinsed in Milli-Q water and allowed to air dry. The membrane was stored between two sheets of Whatmann 3MM paper in a sealed bag at 4°C.

2.2.7.2 Making a digoxygenin (DIG) labelled probe

Probes were made using the PCR DIG Probe synthesis kit (obtained from Roche; catalogue number 1636090). Oligos to amplify the probes were designed and ordered as described in section 2.4.1. PCR labelling reactions were set up according to manufacturers instructions. Probes were stored at -20°C.

2.2.7.3 Hybridisation using a digoxygenin labelled probe to *E. coli* DNA

Hybridisation was carried out using a PCR DIG labelled probe and DIG Easy Hyb hybridisation solution (obtained from Roche; catalogue number 11796895001), according to manufacturers instructions for target DNA with greater than 50% GC content and 100% homology to the probe. All steps were carried out using a hybridisation tube. Membranes were air dried and stored at 4°C in a sealed bag.

2.2.7.4 Detection of a probe bound to target DNA by chemiluminescence

The membrane was initially blocked using Blocking Reagent (obtained from Roche; catalogue number 1096176) according to manufacturers instructions. Detection of probe-target hybrids was carried out using anti-DIG-alkaline phosphatase antibody (obtained from Roche; catalogue number 1093274) and chemiluminescent alkaline phosphatase substrate (CSPD) (obtained from Roche; catalogue number 11655884) according to the manufacturers instructions, and using CRONEX® medical X-ray

film (obtained from Sterling Diagnostic Imaging). Exposure times varied according to the strength of the signal but were typically between 15 and 60 minutes.

2.2.7.5 Stripping a membrane of a probe (for re-probing)

The membrane was rinsed thoroughly in Milli-Q water and then washed twice for 15 minutes at 37°C in Stripping buffer. The membrane was then rinsed briefly in 2 x SSC solution. This was carried out just prior to reprobing with another probe. To reprobe the membrane the normal procedures are followed starting with the prehybridisation step (see section 2.2.7.3).

2.2.8 Plasmid Mediated Gene Replacement (PMGR)

Precise chromosomal alterations (deletions, insertions or point mutations) can be engineered into any recombination proficient *E. coli* chromosome using the pKO3 "gene replacement" system designed by Link and collaborators (Link et al., 1997). pKO3 contains a temperature-sensitive origin of replication, the *cat* gene that confers chloramphenicol resistance, and the *Bacillus subtilis sacB* gene that is lethal to *E. coli* if expressed in the presence of sucrose. A derivative of this vector, pTOF24 (Merlin et al., 2002) also contains a cloning region spanning the *aph* gene that confers kanamycin resistance. It can be appreciated that a plasmid containing chromosomal homology can integrate into the chromosome by homologous recombination, in the case of pKO3 (and derivatives) this event can be selected for by growing cells in media containing chloramphenicol at a temperature that does not allow plasmid replication (non-permissive). Likewise an integrated plasmid can excise by homologous recombination, this event can be selected for by growing in media supplemented with sucrose, furthermore cells that have lost the plasmid can be

selected by growing cells at a non-permissive temperature and then selecting colonies that are chloramphenicol sensitive.

For creating any given change a PCR product was generated (by cross-over PCR; section 2.2.4.2.) creating the desired altered locus surrounded by approximately 400 base-pairs of upstream and downstream homology. Oligos were designed to incorporate specific restriction enzyme recognition sites at the ends of the PCR fragments and they were cloned into the pTOF24 plasmid by digestion and ligation in place of the *aph* gene. Recipient cells took up the plasmid by transformation (see section 2.2.3.10) and were grown up on LB agar plates containing 50 μg ml^{-1} chloramphenicol at 30°C (permissive temperature). Chromosomal integrants were selected by streaking a single transformant onto a LB agar plate containing 50 μg ml^{-1} chloramphenicol and incubating overnight at 42°C integrants (large colonies) were purified by restreaking and growing again at 42°C. An overnight culture grown at 30°C (without selection) was prepared using a single isolated colony and 200 μl of a 10^{-6} dilution was plated on an LB agar plate supplemented with 5% (w/v) sucrose to select for plasmid excision. Loss of the plasmid was confirmed by testing colonies for chloramphenicol sensitivity. If integration and excision occur at the different homologous regions (flanking the mutation) then the mutation should be retained in the chromosome. Potential clones were screened for the desired change either phenotypically, by boiled-cell PCR, restriction digest or cycle-sequencing (or a combination of the four).

2.2.9 Fluctuation analysis

In 1943 Luria and Delbrück published work describing the rate at which *E. coli* became resistant to phage T1 infection (Stent, 1971). They showed that a mutation conferring resistance would arise spontaneously and did not arise as a result of exposure to the phage. This conclusion was reached due to a fluctuating distribution of numbers of mutants arising in different populations (variation between repeated experiments). Based on the conclusions from this work a mutation rate can be calculated and standard errors calculated by combining data from multiple experiments.

In this study fluctuation analysis was used to measure the rate of recombination of a zeocin sensitive cassette to give zeocin resistance gene in various strains. Cultures were grown from single isolated colonies in 5 ml of low-salt L broth overnight at 37°C with shaking. To establish the number of cells in the culture 100 µl of a 10^{-6} dilution was spread onto a LB agar plate, to establish the number of recombinants 100 µl of a 10^{-2} or 10^{-3} dilution was plated onto low-salt LB agar plates containing 35 µg ml^{-1} zeocin. The numbers were put into fluctuation analysis tables that accordingly calculated the median frequency of mutation and from this a rate of recombination was calculated (mutations $cell^{-1}$ $generation^{-1}$) with 95% confidence intervals (Spell and Jinks-Robertson, 2004).

2.2.10 Microscopy

2.2.10.1 Preparing samples

Slides were prepared with a thin layer of 1 % (w/v) agarose melted in Milli-Q water. After the agarose had set a 4 µl sample of cells was placed on the surface and a coverslip was placed over this.

2.2.10.2 Obtaining images

A Zeiss Axiovert 200 microscope with a Semrock dichroic GFP filter set was used to obtain both bright-field and fluorescent images using Metamorph v 6.3r2.

2.2.10.3 Analysing images

Images were manipulated and analysed using Metamorph v 6.3r2.

2.3 Strains

2.3.1 Bacterial strains

The *Escherichia coli* strains used are listed in table 2.1.

2.3.2 Bacteriophage lambda strains

The bacteriophage lambda strains used are listed in table 2.2.

2.4 Oligonucleotides and Plasmids

2.4.1 Oligonucleotides

Oligos were designed using the internet-based tool Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with MG1655 DNA sequences obtained from the *E. coli* genome database (http://genolist.pasteur.fr/Colibri/). They are listed in table 2.3. Oligos were ordered

through MWG Biotech (https://ecom.mwgdna.com/register/index.tcl) where they were synthesised, HPSF purified and lyophilized. They were dissolved in sterile water to give a 100 mM stock solution and stored at -20°C. Working solutions were made up to 5 mM in sterile water and stored at -20°C.

2.4.2 Plasmids

The plasmids used are listed in table 2.4.

Table 2.1: *E. coli* strains.

Strain	Background	Relevant/New Genotype	Source or construction
AB1157 derivatives			
AB1157		*his4 proA2 argE3 thr1 leuB6 ara14 galK2 lacYi xyl5 mtl1 str^R*	
DS984		*recF lacZ ΔM15 lacI^q xerCY17 (mini Mu lacCm^R in xerC)*	D. Sherratt
JJC450		*recF400::Tn5 (Kan^R)*	B. Michel
JJC777		*recB268::Tn10/ containing pDW52 (= pBR322-recBCD+)*	B. Michel
JJC1398		*priA2::Kan^R hsdR thr+ pro+/ containing pAMpriA+*	Grompone *et al* (2003)
N1642		*malE::Tn10 lexA3*	R. Lloyd
N2313		*sbcC201 proC29 phoR79::Tn10*	Lloyd and Buckman (1985)
N2362		*recB21*	Lloyd and Buckman (1985)
N2365		*sbcC7623 recB21*	Lloyd and Buckman (1985)
N2679		*sbcC201*	Naom et al (1989)
N2691		*recA269::Tn10*	Lloyd and Buckman (1985)
N2693		*sbcC201 recA269::Tn10*	Lloyd and Buckman (1985)
N3793		*ΔrecG263::Kan^R*	R. Lloyd
DL1092		*recF332::Tn3*	Cromie et al (2000)
DL1093		*sbcC201 recF332::Tn3*	Cromie et al (2000)
DL1108		*recO1504::Tn5*	Cromie et al (2000)
DL1109		*sbcC201 recO1504::Tn5*	Cromie et al (2000)
DL1110		*recR252::Tn10kan*	Cromie et al (2000)
DL1111		*sbcC201 recR252::Tn10kan*	Cromie et al

			(2000)
Other strains:			
MG1655		*F⁻ lambda⁻ ilvG⁻ rfb-50 rph-1*	
BW27784	BW25113	*DE(araFGH)* (ΔP_{araE} P_{CP18}-*araE)*	Khlebnikov *et al* (2001)
DB1318		*recD1014 hsdR2 zgq202::Tn10 recA::CmR*	D. Botstein
GR47		*lexA3 difΔ6::Kan*^{R}	Reccia *et al* (1999)
DL1645	BW27784	*sbcC201 proC29 phoR79::Tn10*	P1 from N2313
DL1651	DL1645	*proC*$^+$ *sbcC201 phoR*$^+$	P1 from N2679
DL1751	MG1655	*lacZ*$^+$ *chi site at 2230bp removed (ACCACCAGC to ACCACTAGT)*	PMGR using pDL2309
DL1777	DL1751	*lacI*^{q}	PMGR using pDL1828
DL1780	BW27784	ΔP_{sbcDC} P_{BAD}-*sbcDC*	PMGR using pDL1779
DL1803	DL1777	ΔP_{sbcDC} P_{BAD}-*sbcDC*	PMGR using pDL1779
DL1979	DL1803	*lacZ::246bp palindrome*	PMGR using pDL2774
DL2005	DL1979	*cynX::Gm*^{R}	PMGR using pDL2812
DL2006	DL1780	*lacZ::246bp palindrome cynX::Gm*^{R}	P1 from DL2005
DL2075	DL2006	*recA::Cm*^{R}	P1 from DB1318
DL2076	DL2006	*recB268::Tn10*	P1 from JJC777
DL2077	DL2006	*recF400::Tn5*	P1 from JJC450
DL2078	DL2077	*recB268::Tn10*	P1 from JJC777
DL2151	DL1777	*ΔsbcDC*	PMGR using pTOFsbcDC
DL2155	DL2006	*ΔrecG::Kan*^{R}	P1 from N3793
DL2237	DL2006	*xerC Y17::Cm*^{R}	P1 from DS984
DL2241	DL2006	*difΔ6::Kan*^{R}	P1 from GR47
DL2294	DL2006	*priA2::Kan*^{R}*/ containing pAMpriA+*	P1 from JJC1398

DL2409	BW27784	Δ*sbcDC*	PMGR using pTOFsbcDC
DL2473	BW27784	ΔP_{recA} P_{BAD}-*recA*	PMGR using pDL2378
DL2518	DL2409	ΔP_{recA} P_{BAD}-*recA*	PMGR using pDL2378
DL2522	MG1655	*lacZ::*I-*SceI*$_{cs}$	PMGR using pDL2521
DL2523	DL2522	Δ*araBAD::*P_{BAD}-I-*SceI*	PMGR using pDL2655
DL2537	DL2473	*lacZ::246bp palindrome cynX::Gm*^{R}	P1 from DL2005
DL2538	DL2518	*lacZ::246bp palindrome cynX::Gm*^{R}	P1 from DL2005
DL2551	DL1777	*cynX::Gm*^{R}	PMGR using pDL2812
DL2573	DL1780	*lacZ*$^+$ *cynX::Gm*^{R}	P1 from DL2551
DL2589	DL2573	*containing pAMpriA+ (from JJC1398)*	
DL2605	DL2573	*recA::Cm*^{R}	P1 from DB1318
DL2606	DL2573	*recB268::Tn10*	P1 from JJC777
DL2608	DL2573	*recF400::Tn5*	P1 from JJC450
DL2610	DL2573	Δ*recG263::Kan*^{R}	P1 from N3793
DL2611	DL2573	*xerC Y17::Cm*^{R}	P1 from DS984
DL2612	DL2573	*difΔ6::Kan*^{R}	P1 from GR47
DL2630	DL2589	*priA2::Kan*^{R}	P1 from JJC1398
DL2643	DL2523	*recB268::Tn10*	P1 from JJC777
DL2726	DL2409	*lacZ::246bp palindrome cynX::Gm*^{R}	P1 from DL2005
DL2727	DL2726	*recB268::Tn10*	P1 from JJC777
DL2735	DL2608	*recB268::Tn10*	P1 from JJC777
DL2775	DL2573	*tsx::*I-*SceI*$_{cs}$	PMGR using pDL2755
DL2776	DL2006	*tsx::*I-*SceI*$_{cs}$	PMGR using pDL2755
DL2777	DL2726	*tsx::*I-*SceI*$_{cs}$	PMGR using pDL2755

DL2792	DL2775	*proA::*I-*Sce*I$_{cs}$	PMGR using pDL2736
DL2793	DL2776	*proA::*I-*Sce*I$_{cs}$	PMGR using pDL2736
DL2794	DL2777	*proA::*I-*Sce*I$_{cs}$	PMGR using pDL2736
DL2797	DL2792	*recB268::Tn10*	P1 from JJC777
DL2798	DL2793	*recB268::Tn10*	P1 from JJC777
DL2799	DL2794	*recB268::Tn10*	P1 from JJC777
DL2800	DL2573	Δ*ruvAB*	PMGR using pDL2757
DL2801	DL2006	Δ*ruvAB*	PMGR using pDL2757
DL2829	DL2573	*malE::Tn10 lexA3*	P1 from N1642
DL2830	DL2006	*malE::Tn10 lexA3*	P1 from N1642
DL2849	DL2792	*lacZ::*I-*Sce*I$_{cs}$	PMGR using pDL2521
DL2852	DL2151	*lacZ::'zeo chi*$^{+}$ *cass'*	PMGR using pDL2851
DL2859	DL1777	*lacZ::246bp palindrome cynX::Gm*^{R}	P1 from DL2005
DL2874	DL2859	Δ*sbcDC*	PMGR using pTOFsbcDC
DL2883	DL2859	*cynX::'zeo chi*$^{+}$ *cass'*	PMGR using pDL1998
DL2884	DL2859	*cynX::'zeo chi*^{0} *cass'*	PMGR using pDL1999
DL2886	DL1777	*cynX::'zeo chi*$^{+}$ *cass'*	PMGR using pDL1998
DL2887	DL1777	*cynX::'zeo chi*^{0} *cass'*	PMGR using pDL1999
DL2888	DL2151	*cynX::'zeo chi*$^{+}$ *cass'*	PMGR using pDL1998
DL2889	DL2151	*cynX::'zeo chi*^{0} *cass'*	PMGR using pDL1999
DL2899	DL2874	*lacZ::(246bp palindrome), 'zeo chi*$^{+}$ *cass'*	PMGR using pDL2851
DL2901	DL2874	*cynX::'zeo chi*$^{+}$ *cass'*	PMGR using

			pDL1998
DL2902	DL2874	*cynX::'zeo chi^0 cass'*	PMGR using pDL1999
DL2963	DL1777	*mhpC::'zeo chi$^+$ cass'*	PMGR using pDL2941
DL2964	DL1777	*mhpC::'zeo chi^0 cass'*	PMGR using pDL2942
DL2965	DL2151	*mhpC::'zeo chi$^+$ cass'*	PMGR using pDL2941
DL2966	DL2151	*mhpC::'zeo chi^0 cass'*	PMGR using pDL2942
DL2967	DL2859	*mhpC::'zeo chi$^+$ cass'*	PMGR using pDL2941
DL2968	DL2859	*mhpC::'zeo chi^0 cass'*	PMGR using pDL2942
DL2969	DL2874	*mhpC::'zeo chi$^+$ cass'*	PMGR using pDL2941
DL2970	DL2874	*mhpC::'zeo chi^0 cass'*	PMGR using pDL2942

Table 2.2: Bacteriophage Lambda Strains.

Strain	Description	Reference or source
DRL152	*Δ(red gam) cI857 chi*$^{+}$	Cromie *et al* 2000
DRL154	*pal571 Δ(red gam) cI857 chi*$^{+}$	Cromie *et al* 2000
SKK43	*pal246 red3 gam1342::Kan*^{R} *imm21*	Kulkarni, 1990
DRL243	SKK43 with deletion of *pal246* and 46 bp of the surrounding region	Pinder *et al*, 1998
DRL246	λ TXF97 (*imm21 nin5*) with *zeo*^{R} marker cloned in (to the left of *att*)	Cromie *et al.* 2000
DRL282	DRL246 with the 246bp palindrome (from SKK43) cloned in (to the left of *att*)	Cromie *et al.* 2000

Table 2.3. Oligonucleotides

Name	Sequence (5' to 3')	Summary of use
pKO.F	AGG GCA GGG TCG TTA AAT AGC	PCR across and sequencing through pTOF24 clone site
pKO.R2	AGG GAA GAA AGC GAA AGG AG	PCR across and sequencing through pTOF24 clone site
Ex-test-F	TTA TGC TTC CGG CTC GTA TG	PCR across and sequencing through lacZ 'clone site'
Ex-test-R	GGC GAT TAA GTT GGG TAA CG	PCR across and sequencing through lacZ 'clone site'
DIGyaiOF	CAT GAA GAG CGT CAG TCT GC	for making PCR DIG labelled Southern probes
DIGyaiOR	TTG CCA GAA CGA CAG AAC TC	for making PCR DIG labelled Southern probes
DIGcodBF	TGG TCG CAA AGT GCT CAT AG	for making PCR DIG labelled Southern probes
DIGcodBR	AAA AGG GGG AAT TTC GTG TC	for making PCR DIG labelled Southern probes
zeocass_F	AAA AAT CTA GAG ATA AAA CCG ATT CCC TGC AC	for amplifying 2x zeo cassette
zeocass_R	AAA AAA GAT CTG ATC CCC GGG AAT TCA GAC	for amplifying 2x zeo cassette
zeogap_F	AAA AAA GGA TCC GAT AAA ACC GAT TCC CTG CA	for amplifying 2x zeo cassette
zeogap_R	AAA AAA GCG GCC GCG ATC CCC GGG AAT TCA GA	for amplifying 2x zeo cassette
ygaD_F	AAA AAG TCG ACC GAG TCT TGT ACC GGT GGT T	crossover-PCR to make ygaD-paraBAD-recA contruct
ygaD_R	AGT ATG AAA AGT TCA AGT GTT TTG TAG AAA TTG TTG C	crossover-PCR to make ygaD-paraBAD-recA contruct
paraBrecA_F	CAA AAC ACT TGA ACT TTT CAT ACT CCC ACC ATT C	crossover-PCR to make ygaD-paraBAD-recA contruct
paraBrecA_R	GTC GAT AGC CAT CGT TTC ACT CCA TCC AAA AA	crossover-PCR to make ygaD-paraBAD-recA contruct
recA_F	TGG AGT GAA ACG ATG GCT ATC GAC GAA AAC AAA C	crossover-PCR to make ygaD-paraBAD-recA contruct
recA_R	AAA AAC GAT CGC ACG GAG TCA ACG ACG ATA A	crossover-PCR to make ygaD-paraBAD-recA contruct
PhoB.F	AAA AAC TGC AGC GAA AAC	crossover-PCR to make phoB-

	GGC TTG GTG ATA	paraBAD-sbcDC contruct
PhoB.R2	CTT CTC TGA ATG GCT GCG CCA CGG AAA TCA ATA ACC T	crossover-PCR to make phoB-paraBAD-sbcDC contruct
Pbad.F2	CCG TGG CGC AGC CAT TCA GAG AAG AAA CCA ATT GTC C	crossover-PCR to make phoB-paraBAD-sbcDC contruct
Pbad.R2	GTG AAG GAT GCG CAT CGT TTC ACT CCA TCC AAA AAA AC	crossover-PCR to make phoB-paraBAD-sbcDC contruct
sbcD.F2	GAG TGA AAC GAT GCG CAT CCT TCA CAC CTC AGA C	crossover-PCR to make phoB-paraBAD-sbcDC contruct
sbcD.R2	AAA AAG TCG ACC ATC GGC ATA GTG TTG TTG G	crossover-PCR to make phoB-paraBAD-sbcDC contruct
LacZ-chi-OF	AAA AAG TCG ACT CGC CAG TTC TGT ATG AAC G	crossover-PCR for knock-out of lacZ chi site (introduce a SpeI site)
LacZ-chi-IR	ATC CAT TTC ACT AGT GGT CAG ATG CGG GAT G	crossover-PCR for knock-out of lacZ chi site (introduce a SpeI site)
LacZ-chi-IF	CTG ACC ACT AGT GAA ATG GAT TTT TGC ATC G	crossover-PCR for knock-out of lacZ chi site (introduce a SpeI site)
LacZ-chi-OR	AAA AAC TGC AGT TTA CCC GCT CTG CTA CCT G	crossover-PCR for knock-out of lacZ chi site (introduce a SpeI site)
TetO-CF1	AAA AAG CTA GCA AAT ATC TGC CGA CCA AAC C	crossover PCR to make tetO homology arms with clone site at centre
TetO-CR1	CTC GAG AAG GAT CCA ATC TAG ATT TAA TCA CCG AAG GCA TCA C	crossover PCR to make tetO homology arms with clone site at centre
TetO-CF2	TCT AGA TTG GAT CCT TCT CGA GTA TCA AAC ACT CGC CTG GTG	crossover PCR to make tetO homology arms with clone site at centre
TetO-CR2	AAA AAC TGC AGC CCA GAC CTA ACC CAC ACA C	crossover PCR to make tetO homology arms with clone site at centre
Bsd-CR2	AAA AAC TGC AGT CAT CAG CTT CAG GTT TTC G	crossover PCR to make mhpCB homology arms with clone site at centre
MhpC_Rchi+	ACC AGC AGA TCT AAA AAT CTA GAT TTA ATG ACT GCG	crossover PCR to make mhpCB homology arms with clone site

	GAC AAG G	at centre
MhpB_Fchi+	ATC TAG ATT TTT AGA TCT GCT GGT GGT AAT AAC GCC CTT CGC TAC G	crossover PCR to make mhpCB homology arms with clone site at centre
Bsd-CF1-2	AAA AAG CGG CCG CAT GTT GGG TGA AGC CAT TG	crossover PCR to make mhpCB homology arms with clone site at centre
lacZ_F1	AAA AAC TGC AGA ACG TCG TGA CTG GGA AAA C	crossover PCR to make lacZ homology arms with clone site at centre
lacZ_R1	CGC GGC CGC CCC AAA GGA TCC GTT GCA CCA CAG ATG AAA CG	crossover PCR to make lacZ homology arms with clone site at centre
lacZ_F2	AAC GGA TCC TTT GGG GCG GCC GCG TCG TTT GCC GTC TGA ATT T	crossover PCR to make lacZ homology arms with clone site at centre
lacZR2	AAA AAG TCG ACA CCA CCG CAC GAT AGA GAT T	crossover PCR to make lacZ homology arms with clone site at centre
Gent_F	CCG CTC GAG ACC CAG TTG ACA TAA GCC TGT T	amplify gentamicin gene (from pLAU44)
Gent_R	CGC GGA TTC TTA GGT GGC GGT ACT TGG GT	amplify gentamicin gene (from pLAU44)
ruvA-KO-F1	AAA AAC TGC AGG ATC CCG ACG TGA TTA CTC C	check for pDL2757 mediated ruvAB deletion
ruvB-KO-R2	AAA AAG TCG ACT GAC GAT TGG TGT AGC GAT G	check for pDL2757 mediated ruvAB deletion
DIGmalZ_F	GCG TGA CTG GGA TGA ACC	for making PCR DIG labelled Southern probes
DIGmalZ_R	CGG ATC GTT TTT GCC ATC	for making PCR DIG labelled Southern probes
DIGperR_F	ATA GCG CTG GCT GCA TCT	for making PCR DIG labelled Southern probes
DIGperR_R	GCT GGG GGA AAA CCA AAT	for making PCR DIG labelled Southern probes
sdm_kochi_f	CTA GAT TTT TAG ATC TAC TAG TGG TAA TAA CGC CCT TCG C	Site-directed mutagenesis of mhpBC homology arms to make chiO version
sdm_kochi_r	GCG AAG GGC GTT ATT ACC ACT AGT AGA TCT AAA AAT CTA G	Site-directed mutagenesis of mhpBC homology arms to make chiO version

Table 2.4 Plasmids

Plasmid	Brief description	Stored in:	Source or Reference
pAMpriA$^+$		JJC1398	Grompone *et al* (2003)
pGB150	pACYC184 p_{sfiA}-*gfp*	JM109	G. Blakely -unpublished
pLAU44	Contains Gm^R gene		Lau, Fillipe *et al* (2003)
pLacD1	pTOF24 + *lacZ* fragment with *Mfe*I site	XL1 blue	This lab – John Blackwood
pDL2757	pTOF24 + Δ*ruvAB* k.o. fragment	XL1 blue	This lab - Ewa Okely
pTOFsbcDC	pTOF24 +Δ*sbcDC* fragment	XL1 blue	This lab - Elise Darmon
pDL1998	pTOF24 + *cynX*::'*zeo chi*$^+$ cass' fragment	XL1 blue	This lab - John Blackwood
pDL1999	pTOF24 + *cynX*::'*zeo chi*^{0} cass' fragment	XL1 blue	This lab - John Blackwood
pDL2736	pTOF24 + *proA*::I-*Sce*I$_{cs}$ fragment	XL1 blue	This lab - John Blackwood
pDL2755	pTOF24 + *tsx*::I-*Sce*I$_{cs}$ fragment	XL1 blue	This lab - John Blackwood
pDL1828	pTOF24 + *lacI*^{q} mutation fragment	XL1 blue	This lab - John Blackwood
pDL2521	pLacD1 + I-*Sce*I$_{cs}$ fragment	XL1 blue	This lab - Katie Tully
pDL2655	pTOF24 + P$_{BAD}$-I-*Sce*I fragment	XL1 blue	This lab - Martin White/John Eykelenboom
pDL1779	pTOF24 + *phoB*, P$_{BAD}$-*sbcD* fragment	MG1655	
pDL2069	pTOF24 + *mhpC chi*$^+$ fragment	XL1 blue	
pDL2282	pTOF24 + *cynX* fragment	XL1 blue	
pDL2309	pTOF24+ *lacZ chi* knock out fragment	XL1 blue	
pDL2378	pTOF24 + P$_{BAD}$-*recA* fragment	XL1 blue	
pDL2444	pTOF24 + *lacZ* fragment	XL1 blue	
pDL2774	pLacD1 + *pal246* fragment	DL733	
pDL2812	pDL2282+ *Gm*^{R} fragment	XL1 blue	

pDL2851	pDL2444 + *lacZ*::'*zeo chi*$^{+}$ cass' fragment	XL1 blue
pDL2940	pDL2069 sdm to remove *chi* sequence - replace with *Spe*I	XL1 blue
pDL2941	pDL2069 + *mhpC*::'*zeo chi*$^{+}$ cass' fragment	XL1 blue
pDL2942	pDL2940 + *mhpC*::'*zeo chi*^{0} cass' fragment	XL1 blue

Chapter 3: Lysogenisation with lambda containing a palindrome in a P_{BAD}-*sbcDC* strain.

3.1 Introduction

As discussed in chapter one, the SbcCD nuclease can cleave DNA-hairpin structures (as can be formed by palindromic DNA sequences) *in vitro*, without the need of a 3' or 5' end (Connelly et al., 1998). Genetic evidence suggests that SbcCD cleaves DNA hairpin structures *in vivo*. *E. coli* strains with functional SbcCD and deficient in DNA double-strand break repair (*recA*$^-$, *recB*$^-$) cannot harbour large palindromic sequences in their chromosomes (Cromie et al., 2000; Leach et al., 1997). However, so far there is no *in vivo* physical evidence that SbcCD activity at a palindromic sequence results in a DNA double-strand break. In order to study the effect of SbcCD under conditions where the presence of the complex might be lethal the *sbcDC* operon was placed under the control of a repressible promoter, allowing strains to be constructed under SbcCD$^-$ conditions and then monitored under SbcCD$^+$ conditions.

A large body of research has been dedicated to understanding transcriptional control from the *araBAD* promoter because of its potential for tightly controlling the expression of genes (on bacterial chromosomes or from expression vectors) (Guzman et al., 1995). The *araBAD* genes encode proteins responsible for the metabolism of the sugar arabinose. Under no-arabinose conditions, a dimer of AraC binds the I_1 site of the *araBAD* promoter and an upstream O_2 regulatory region causing a kink in

the DNA and preventing transcription of the *araBAD* genes (Kolodrubetz and Schleif, 1981a; Schleif, 2003). Introduction of arabinose in the medium induces a conformation change in the AraC dimer so that it binds the I_1 and I_2 sites of the promoter which upregulate expression. Two arabinose import systems have been described in *E. coli*, encoded by *araFGH* and *araE* (Kolodrubetz and Schleif, 1981a). Expression of these genes is upregulated by the arabinose-bound AraC dimer (Kolodrubetz and Schleif, 1981b). Expression from the *araBAD* promoter in *E. coli* follows an all-or-nothing pattern propagated by a positive feedback effect of arabinose uptake.

This chapter describes the construction of a strain of *E. coli* where the native *sbcDC* promoter has been replaced by the *araBAD* promoter (P_{BAD}). This strain has been used for lysogenisation experiments adapted from previous work by Cromie and collaborators (Cromie et al., 2000) to re-examine the genetic requirements for lysogenisation with a phage lambda containing a 246 bp interrupted palindrome.

3.2 P_{BAD}-sbcDC – an arabinose inducible system

Plasmid mediated gene replacement (PMGR) (Link et al., 1997) was used to replace the native *sbcDC* promoter with the arabinose inducible *araBAD* promoter (P_{BAD}) resulting in an arabinose inducible *sbcDC* operon (figure 3.1 (e)).

3.2.1 Construction of pDL1779

A 460 bp region of the *phoB* gene (upstream of the *sbcDC* operon), the first 511 bp of the *sbcD* gene and a 349 bp region containing the *araBAD* promoter /regulatory region were PCR amplified using the primer pairs PhoB.F (with *Pst*I site at 5' end)

and PhoB.R2; sbcD.F2 and sbcD.R2 (with *Sal*I site at 5' end); Pbad.F2 and Pbad.R2 respectively and genomic DNA isolated from *E. coli* MG1655 as template. Cross-over PCR was used to fuse the three PCR products using the external primers PhoB.F and sbcD.R2. The novel DNA fragment that has a deletion of the *sbcDC* promoter and an insertion of the *araBAD* promoter, was cloned into pTOF24 using *Pst*I and *Sal*I. The resulting pDL1779 plasmid was verified by sequencing using the oligonucleotides pKO.F and pKO.R2.

3.2.2 A strain for homogeneous expression of P_{BAD}-*sbcDC*

These experiments need homogeneous expression of *sbcDC*. However, wild-type cells display heterogeneous expression from the *araBAD* promoter, due to the nature of the *ara* genes and regulatory regions.

Khlebnikov and collaborators have created *E. coli* strains that allow homogeneous expression from P_{BAD} (Khlebnikov et al., 2001; Khlebnikov et al., 2000). They deleted the *araFGH* operon and placed the *araE* gene under the control of a constitutive promoter, removing the positive feedback loop associated with arabinose uptake. Measurement of the fluorescence intensity across a population of cells containing a plasmid with the *gfpuv* gene under the control of P_{BAD} showed that the expression pattern from P_{BAD} was more heterogeneous when *araE* was under the control of its wild-type promoter than under the control of the constitutive promoter.

The BW27784 strain (obtained from Khlebnikov *et al*) contains the mutations permitting homogeneous expression from the P_{BAD} promoter.

pDL1779 was used to replace the *sbcDC* promoter with P_{BAD} in BW27784. The promoter replacement was selected by PCR using the primer pair PBAD.F_test and SbcD.R_test (clones with the desired chromosomal change give a ~1 kb product). This strain was named DL1780.

3.2.3 Testing the P_{BAD}-*sbcDC* construct

It was noticed in 1983 by Leach and Stahl (Leach and Stahl, 1983) that *red⁻ gam⁻* lambda phages containing long palindromes were viable in recombination proficient *recBC⁻ sbcB⁻* strains of *E. coli*. It was later shown by Lloyd and Buckman (Lloyd and Buckman, 1985) that these suppressor strains carry mutations not only in *sbcB* but also in another gene that became known as *sbcC*. Finally it was demonstrated that palindrome inviability is actually due to the *sbcC* gene (Chalker et al., 1988).

The P_{BAD}-*sbcDC* strain (DL1780) was prepared for plating of lambda phage in the presence of 0.2% (w/v) arabinose (P_{BAD} induced) or 0.5% (w/v) glucose (P_{BAD} repressed). 10 µl dilutions of the phage lambda DRL152 (control; Δ(*red-gam*) *cI857 chi*$^+$) and DRL154 (containing the palindrome; *pal571* Δ(*red-gam*) *cI857 chi*$^+$) were spotted on lawns of cells. *sbcDC*$^+$ and *sbcC201*$^-$ strains were used as controls. Figure 3.1 shows that the P_{BAD}-*sbcDC* strain grown in media supplemented with glucose behaved like the *sbcC⁻* strain (the phage containing the palindrome was viable) and the P_{BAD}-*sbcDC* strain grown in media supplemented with arabinose behaved like the *sbcC*$^+$ strain (the phage containing the palindrome showed reduced viability).

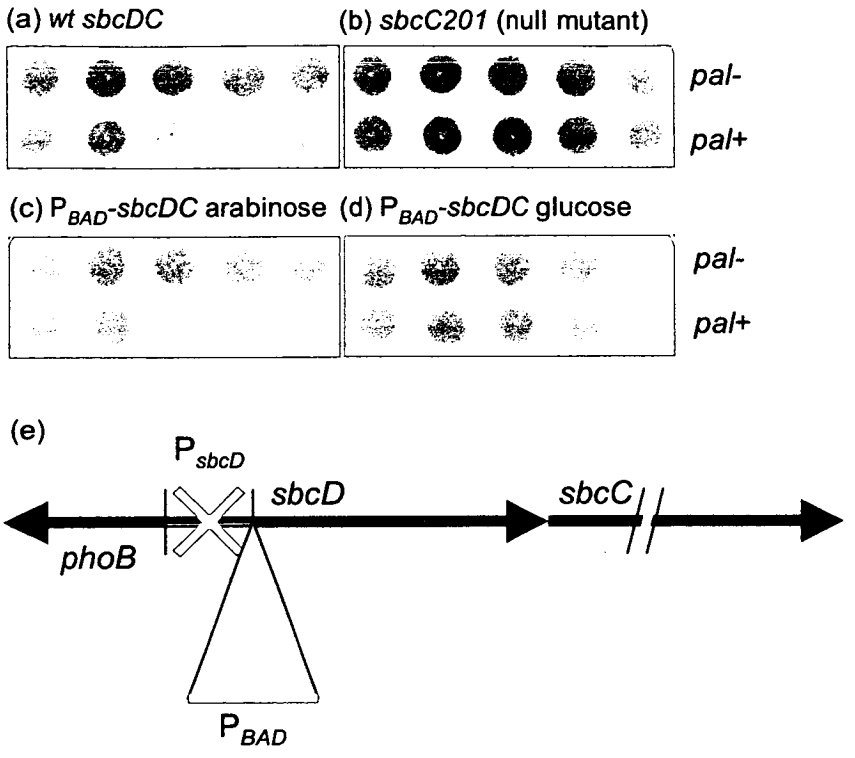


Figure 3.1: Lambda plating test for *sbcC*$^{+/-}$ using the P$_{BAD}$-*sbcDC* strain. Strains used were (a) BW27784, (b) DL1651, (c) DL1780 with arabinose and (d) DL1780 with glucose. (e) cartoon depicting the replacement of P$_{sbcDC}$ with P$_{BAD}$.

3.3 Lysogenisation of P_{BAD}-sbcDC recombination mutants

The lambda phage SKK43 was isolated as a palindrome-containing phage that could form stable lysogens in *sbcDC*$^+$ recombination-proficient strains of *E. coli* (Leach et al., 1997). The palindrome was shown to be a 246 bp interrupted palindrome (containing a 26 bp asymmetrical centre). This phage and an isogenic palindrome-deleted phage were used to lysogenise various recombination-deficient mutants in *sbcDC*$^+$ and *sbcDC*$^-$ genetic backgrounds (Leach et al., 1997). While the non-palindrome phage could lysogenise all genetic backgrounds equally well, there was a 10^4 - 10^5-fold reduction in lysogenisation frequency of strains that were *sbcDC*$^+$ and *recA*$^-$, *recB*$^-$ or *recC*$^-$ using the palindrome phage. A different pair of isogenic phage were created containing a zeocin resistance marker. The first phage contains the 246 bp interrupted palindrome (with 3 internal mismatches introduced during the cloning process) and the second is isogenic but without the palindrome. These phages were used to study lysogenisation frequency of a larger set of recombination-deficient mutants in *sbcDC*$^+$ and *sbcDC*$^-$ backgrounds (Cromie et al., 2000). The results with regard to *recA*$^-$ and *recB*$^-$ mutants reflected those previously observed. In addition it was reported that lysogenisation using the palindrome-containing phage was severely reduced in *sbcDC*$^+$ strains with mutations in *recF*, *recO* or *recR*.

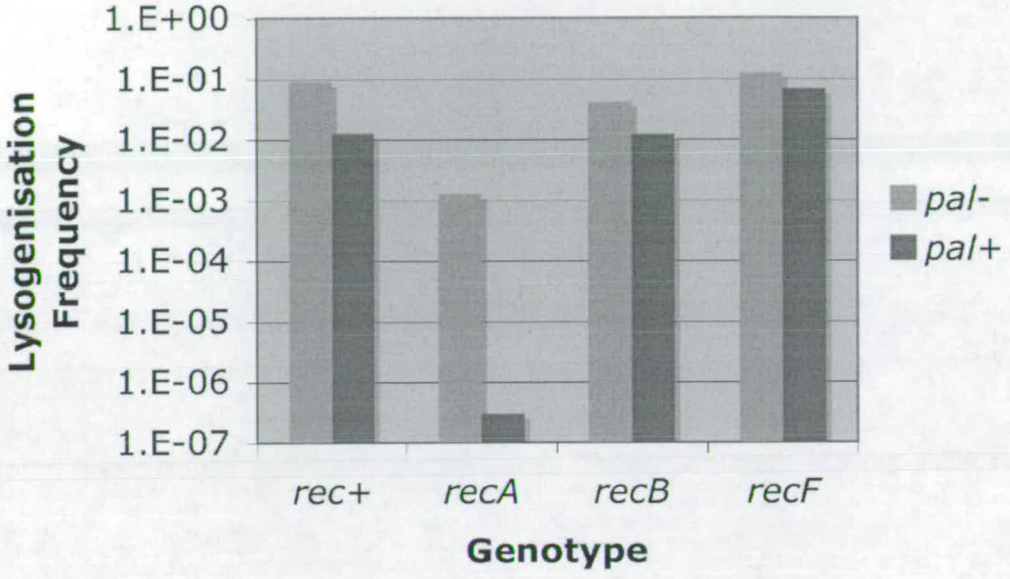
In order to further test the ability to control the expression of *sbcDC* from the P_{BAD} promoter to allow conditional palindrome viability, the *recA*$^-$, *recB*$^-$ and *recF*$^-$ mutations were transferred to the P_{BAD}-*sbcDC* strain by P1 transduction. The frequencies of formation of viable lysogens (referred to from here as lysogenisation

frequency) with the zeocin resistant phage DRL246 (without a palindrome) and DRL282 (containing the palindrome) (Cromie et al., 2000) were calculated for each of these mutants grown under SbcCD$^+$ (arabinose) or SbcCD$^-$ (glucose) conditions (figure 3.2). Neither phage gave significantly higher lysogenisation frequency than the other in any of the strains grown under SbcCD$^-$ conditions (figure 3.2 (b)). However, when grown under SbcCD$^+$ conditions there was a 5 to 10-fold drop in lysogenisation frequency of *rec*$^+$, *recB*$^-$ or *recF*$^-$ strains with the palindrome containing phage (DRL282). Only the *recA*$^-$ mutant grown under SbcCD$^+$ conditions showed the same scale of reduction of about 10^4-fold in lysogenisation frequency with a phage containing a palindrome as in the previous studies.

Lysogenisation experiments using SKK43 and DRL243 (lambda that confer kanamycin resistance) in the P_{BAD}-*sbcDC* strain *rec*$^+$, *recB*$^-$ and *recA*$^-$ versions were carried out and provide further evidence that a *recB*$^-$ strain expressing SbcCD could be lysogenised with a phage containing a palindrome (see figure 3.3).

In order to test if RecF-mediated recombination was substituting for the RecBCD complex and allowing lysogenisation with a phage containing a palindrome in a *recB*$^-$ strain, a *recB*$^-$ *recF*$^-$ double mutant was made by P1 transduction. Under SbcCD$^+$ conditions this mutant had a very low lysogenisation frequency with the phage containing the palindrome (about 10^3 –fold drop compared to lysogenisation with the phage without a palindrome; see figure 3.4). This suggests that a lysogen containing a palindrome in an SbcCD$^+$ strain requires at least one of the RecBCD or RecFOR recombination pathways for survival. This is in

(a)



(b)

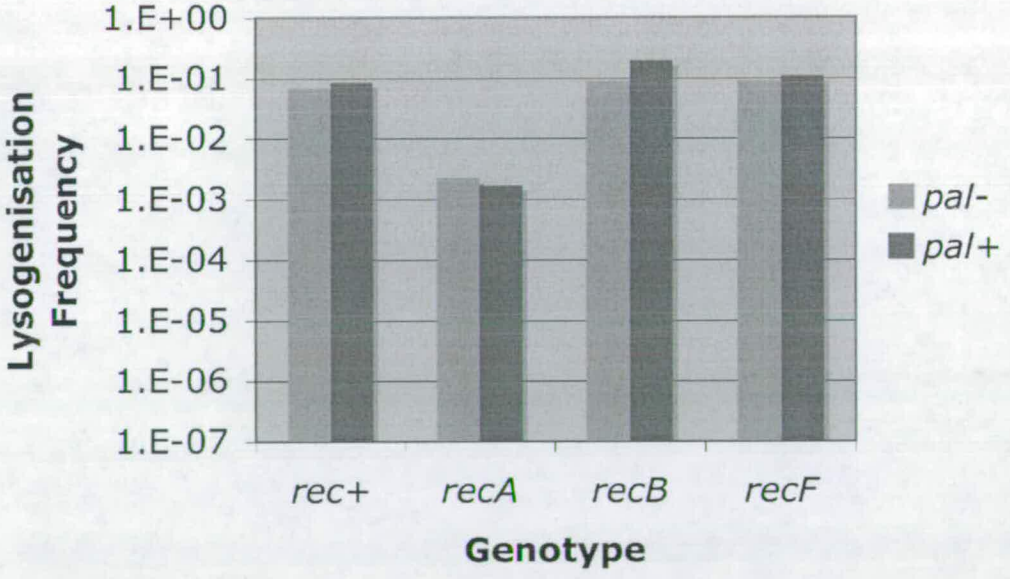
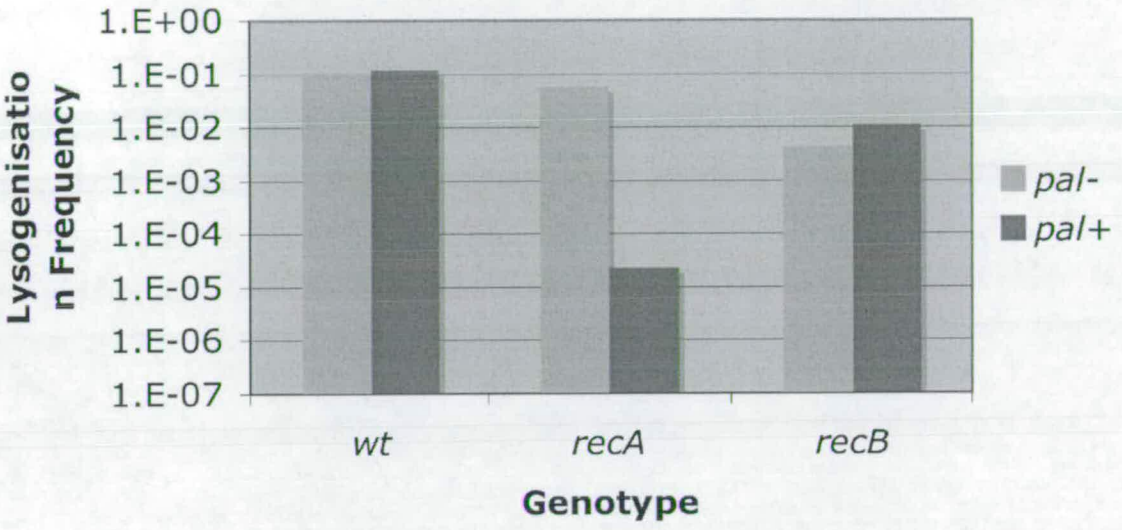


Figure 3.2: Effects of *recA*⁻, *recB*⁻ or *recF*⁻ mutations on lysogenisation frequencies of P_{BAD}-*sbcDC* strains with lambda DRL246 (*pal*⁻) and DRL282 (*pal*⁺). Strains were grown in media supplemented with (a) arabinose and (b) glucose. Strains used were DL2573 (*wt*), DL2605 (*recA*⁻), DL2606 (*recB*⁻) and DL2608 (*recF*⁻). Values are the average of two separate experiments.

(a)



(b)

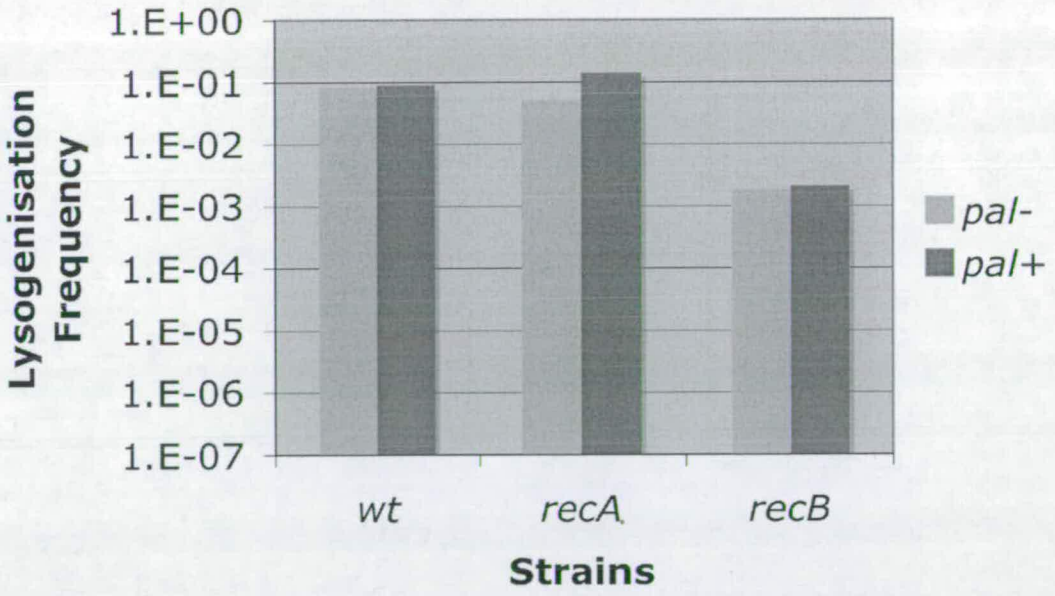


Figure 3.3: Effects of *recA*⁻ and *recB*⁻ mutations on lysogenisation frequencies of P_{BAD}-*sbcDC* strains with lambda DRL243 (*pal*⁻) and SKK43 (*pal*⁺). Strains were grown in media supplemented with (a) arabinose and (b) glucose. Strains used were DL2573 (*wt*), DL2605 (*recA*⁻), DL2606 (*recB*⁻). Values are the average of two separate experiments.

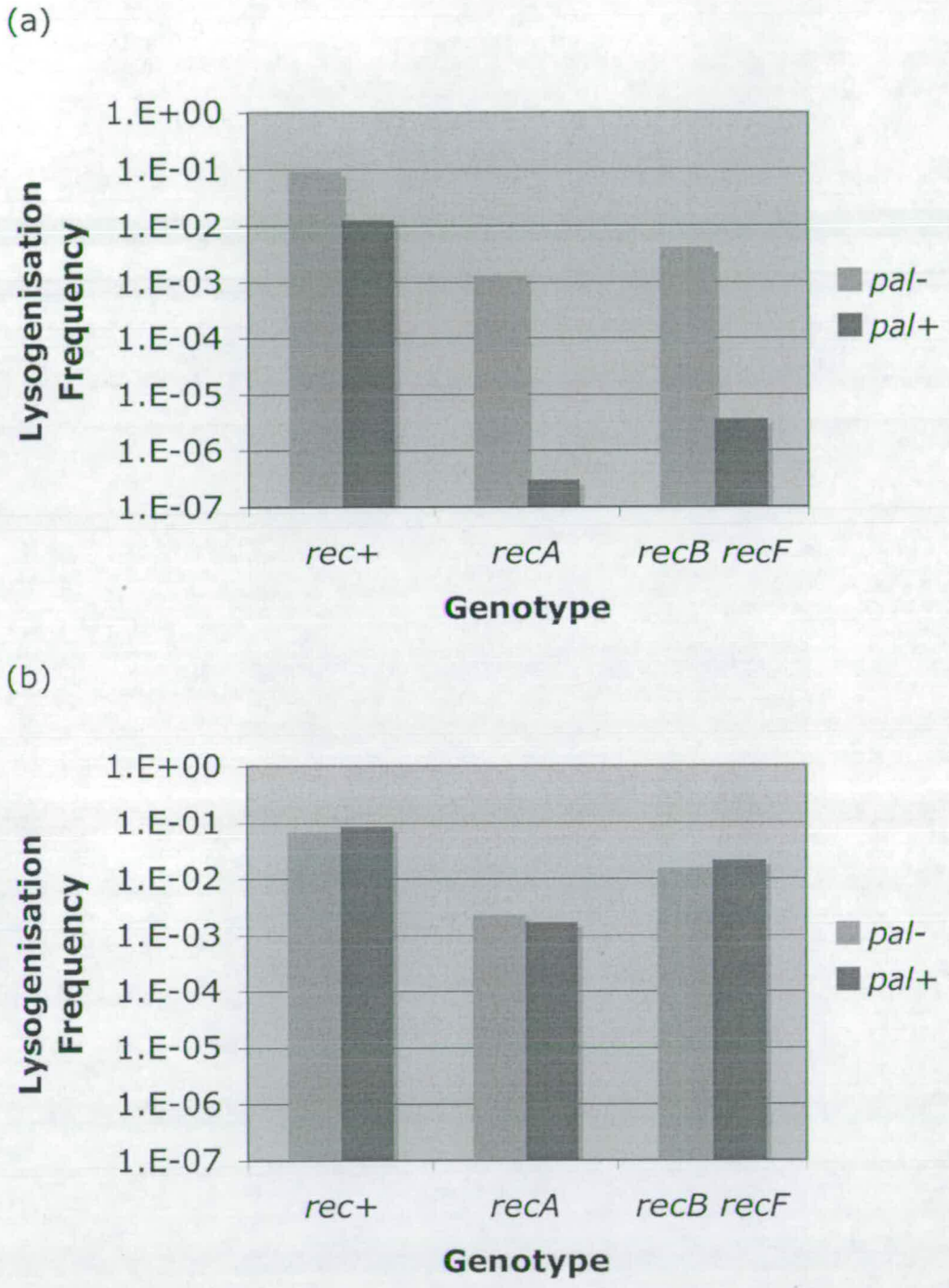


Figure 3.4: Effects of both *recB*⁻ *recF*⁻ mutations on lysogenisation frequencies of P_{BAD}-*sbcDC* strains with lambda DRL246 (*pal*⁻) and DRL282 (*pal*⁺). Strains were grown in media supplemented with (a) arabinose and (b) glucose. Strains used were DL2573 (*wt*), DL2605 (*recA*⁻) and DL2735 (*recB*⁻ *recF*⁻). Values for *rec*+ strains are taken from figure 3.2, other values are the average of two separate experiments.

contrast to the results from the previous work that suggested that RecBCD and RecFOR pathways were both required for survival (Cromie et al., 2000).

An advantage of using an inducible-*sbcDC* system is that lysogens can be created in SbcCD$^-$ conditions and then shifted to SbcCD$^+$ conditions to observe the effects of SbcCD. Colonies lysogenised (using lambda DRL246 and DRL282) under SbcCD$^-$ conditions were picked and grown overnight in LB broth containing 0.5 % (w/v) glucose and then spread onto L agar plates supplemented with either 0.5% (w/v) glucose or 0.2% (w/v) arabinose (see figure 3.5). When grown under SbcCD$^+$ or SbcCD$^-$ conditions *rec*$^+$, *recA*$^-$, *recF*$^-$ lysogens containing DRL246 (without the palindrome) grew equally well. When grown under SbcCD$^+$ conditions (compared to SbcCD$^-$ conditions) the *recB*$^-$ and *recB*$^-$ *recF*$^-$ lysogens containing DRL246 were generally slightly sicker (see figure 3.5). In the case of lysogens containing lambda DRL282, *recA*$^-$, *recB*$^-$ and *recB*$^-$ *recF*$^-$ strains were inviable when plated on LB containing arabinose, although they grew normally on plates containing glucose. *rec*$^+$ and *recF*$^-$ strains containing lambda DRL282 grew equally on plates containing arabinose or glucose.

These initial lysogenisation experiments using the P_{BAD}-*sbcDC* strains have suggested differences with published literature (Cromie et al., 2000). In these experiments there appears to be no requirement for RecFOR-mediated recombination in a strain containing a palindrome grown under SbcCD$^+$ conditions. In addition a *recB*$^-$ strain can be lysogenised with a phage containing a palindrome but does not form a stable healthy lysogen.

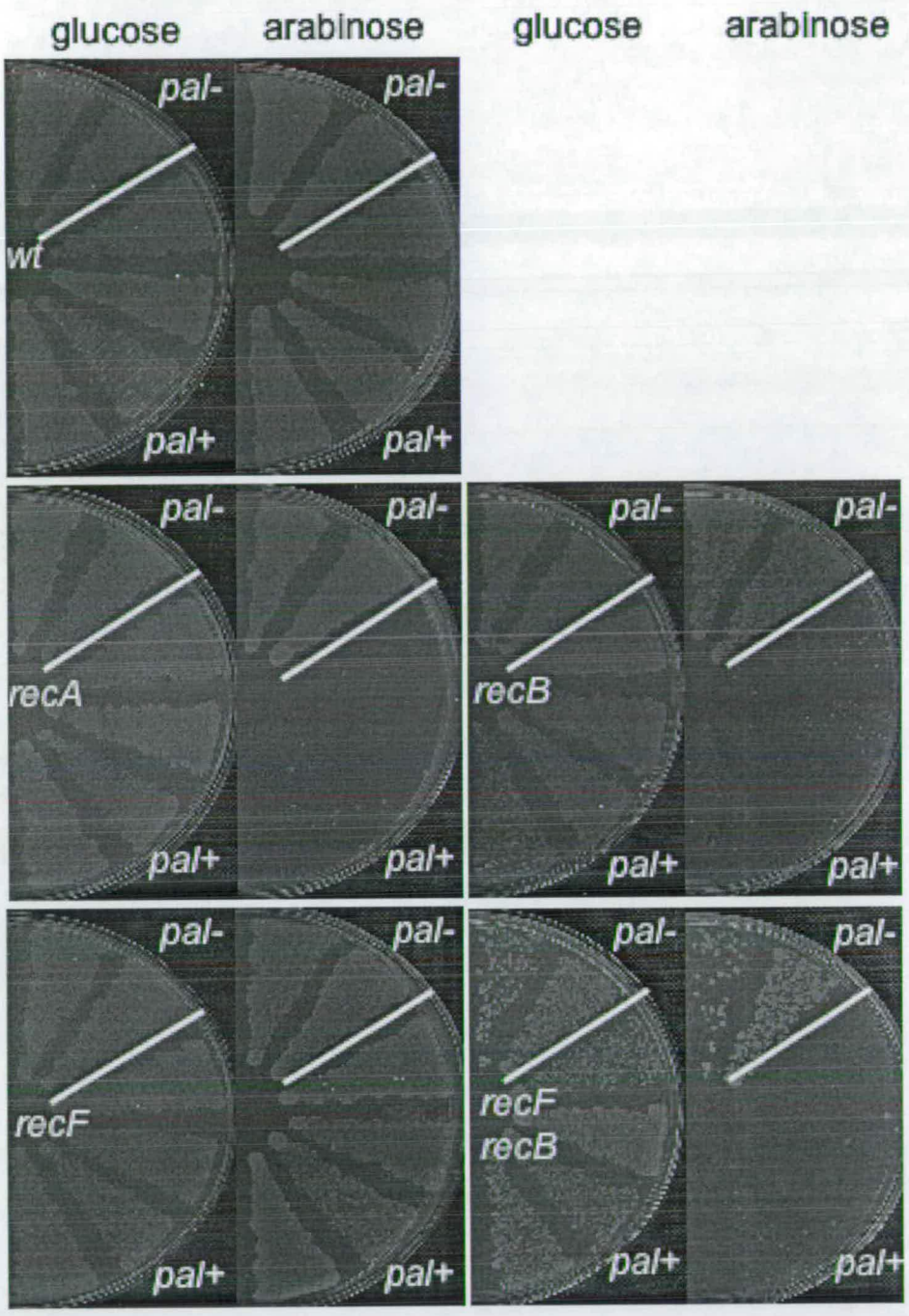


Figure 3.5: Effect of shifting stable P_{BAD}-*sbcDC* lysogens made under SbcCD$^-$ conditions to SbcCD$^+$ conditions. Strains used are DRL246 or DRL282 lysogens of DL2573 (*rec*$^+$), DL2075 (*recA*$^-$), DL2076 (*recB*$^-$), DL2077 (*recF*$^-$) and DL2078 (*recF*$^-$ *recB*$^-$). Each patch represents a separate isolated lysogen.

3.4 Lysogenisation of AB1157 sbcC$^{+/-}$ strains

In order to investigate differences between this work and published work (Cromie et al., 2000) lysogenisation experiments were repeated with the lambda DRL246 and DRL282 (and where possible SKK43 and DRL243) using both *sbcC*$^{+}$ and *sbcC*$^{-}$ mutants in an AB1157 background, and in conjunction with *recA*$^{-}$, *recB*$^{-}$, *recF*$^{-}$, *recO*$^{-}$ and *recR*$^{-}$ mutations (see figures 3.6 and 3.7). The *recA*$^{-}$ and *recB*$^{-}$ strains gave results consistent with previous studies - with either of the phage containing a palindrome in an *sbcDC*$^{+}$ background a large decrease in lysogenisation was observed. In contrast to previous studies the *rec*$^{+}$ *sbcDC*$^{+}$ strain showed a slight reduction in lysogensisation with the zeocin phage containing the palindrome, as did the *sbcDC*$^{+}$ strains carrying *recF*$^{-}$, *recO*$^{-}$ or *recR*$^{-}$ mutations (figure 3.6). This reduction was not seen with the kanamycin phage (figure 3.7). The lysogenisation frequencies obtained with these AB1157 strains were consistent with the viabilities of lysogens derived from the P_{BAD}-*sbcDC* strain (see figure 3.5).

3.5 Effect of zeocin concentration on lysogenisation frequency

Experiments carried out in this laboratory (John Blackwood – unpublished data) (see also chapter 6) using the zeocin resistance gene suggested that the gene had variable effectiveness depending on the strain genetic background and on the concentration of the antibiotic in the growth medium. We decided to examine if different antibiotic concentrations in the growth medium would give varying lysogenisation frequencies.

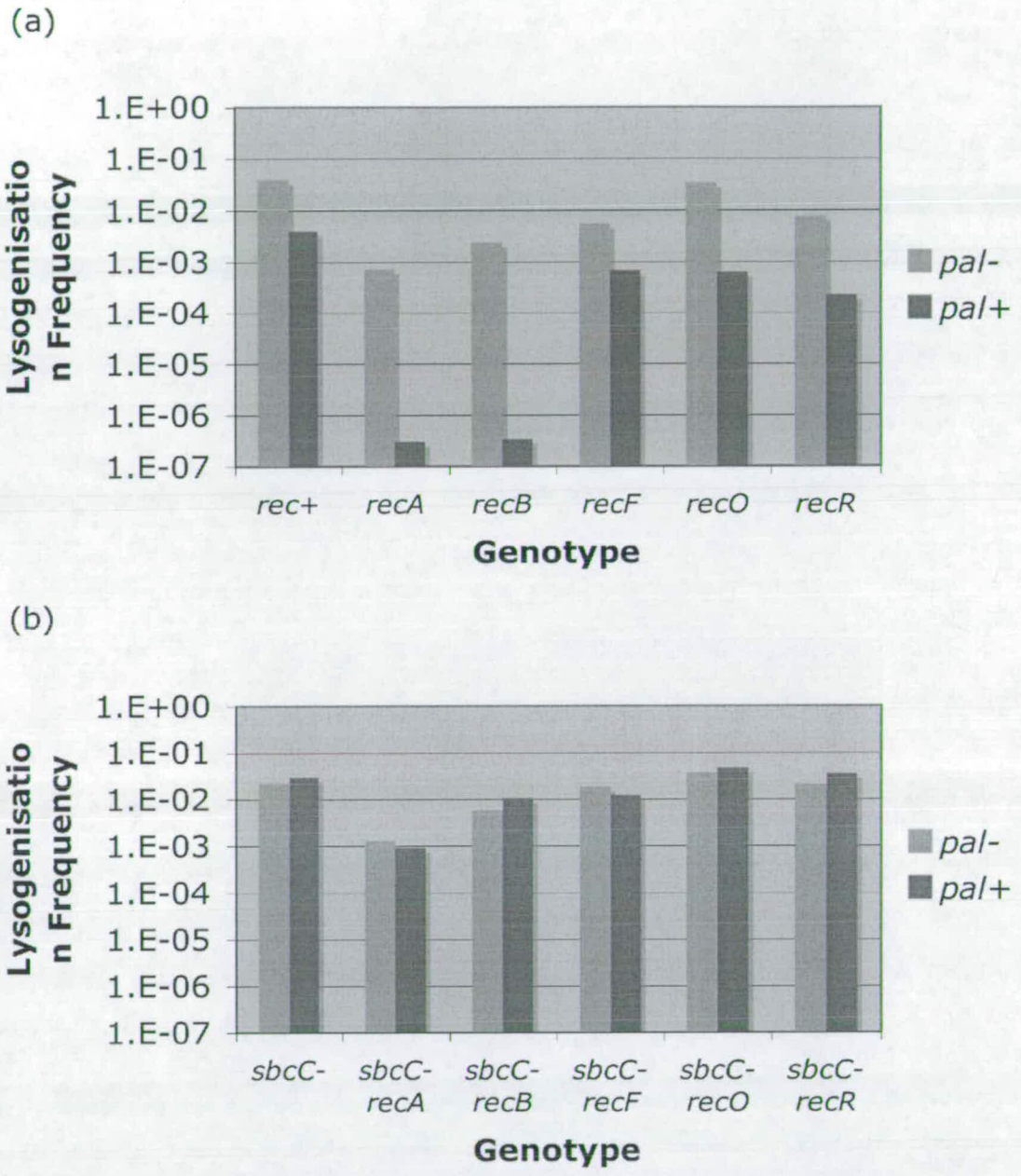


Figure 3.6: Effects of recombination mutations on lysogenisation frequencies of AB1157 background strains with lambda DRL246 (*pal*⁻) and DRL282 (*pal*⁺). Strains were either (a) *wt* or (b) *sbcC*⁻ background. Strains used were AB1157 (*wt*), N2691 (*recA*) and N2362 (*recB*⁻), DL1092 (*recF*⁻), DL1108 (*recO*⁻), DL1110 (*recR*⁻), N2679 (*sbcC*⁻), N2693 (*sbcC*⁻ *recA*⁻), N2365 (*sbcC*⁻ *recB*⁻), DL1093 (*sbcC*⁻ *recF*⁻), DL1109 (*sbcC*⁻ *recO*⁻), DL1111 (*sbcC*⁻ *recR*⁻).

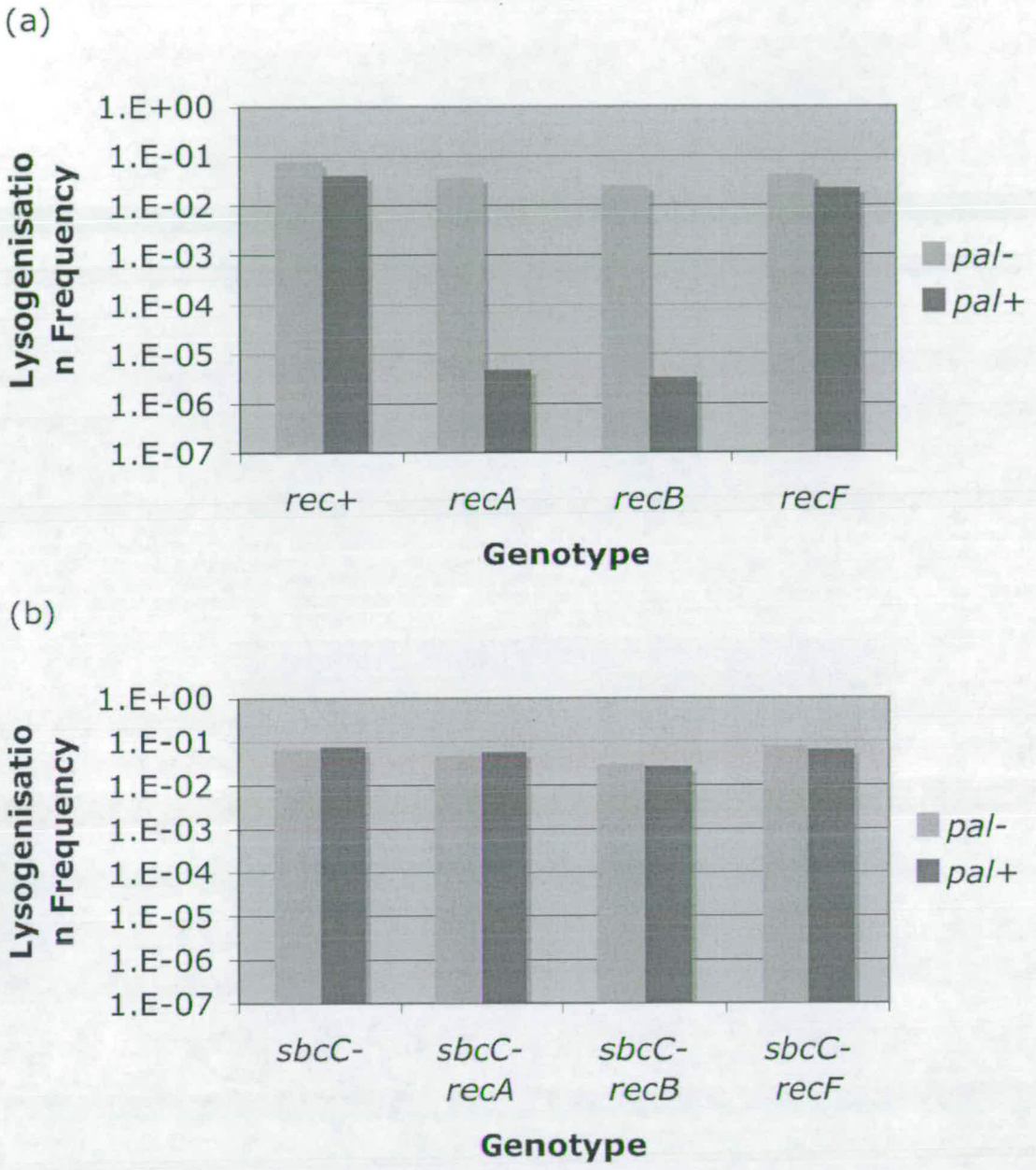


Figure 3.7: Effects of recombination mutations on lysogenisation frequencies of AB1157 background strains with lambda DRL243 (*pal*$^-$) and SKK43 (*pal*$^+$). Strains were (a) *wt* or (b) *sbcC*$^-$ background. Strains used were AB1157 (*wt*), N2691 (*recA*$^-$) and N2362 (*recB*$^-$), DL1092 (*recF*$^-$), N2679 (*sbcC*$^-$), N2693 (*sbcC*$^-$ *recA*$^-$), N2365 (*sbcC*$^-$ *recB*$^-$), DL1093 (*sbcC*$^-$ *recF*$^-$).

Using the recombination proficient AB1157 strains (either *sbcC*$^+$ or *sbcC*$^-$) used by Cromie and collaborators (Cromie et al., 2000), lysogenisation experiments using lambda DRL246 and DRL282 were repeated with different concentrations of zeocin in the medium. Increasing the concentration of zeocin by four-fold reduced by 500-fold the lysogenisation with the phage containing a palindrome in an *sbcDC*$^+$ strain (see figure 3.8). Lysogenisation with the control phage in an *sbcDC*$^+$ strain was not significantly affected by a similar change in zeocin concentration (a 3.5-fold reduction). Finally, lysogenisation with either of the phage in the *sbcC*$^-$ strain was approximately equal under all zeocin concentrations investigated (using DRL246 a 2.2-fold reduction and DRL282 a 3.7-fold reduction with a four-fold increase in zeocin concentration).

3.5 Discussion

These experiments and previous lysogenisation experiments have shed light on the *E. coli* genetic requirements for harbouring palindromic sequences. These results confirm previous models where cleavage of a DNA-hairpin by SbcCD leads to a DNA double-strand break. An SbcCD$^+$ strain carrying a 246 bp palindrome requires the RecBCD (and RecA) recombination pathway (Cromie et al., 2000; Leach et al., 1997). However these experiments do not find a requirement for the RecFOR recombination pathway as observed by Cromie and collaborators (Cromie et al., 2000).

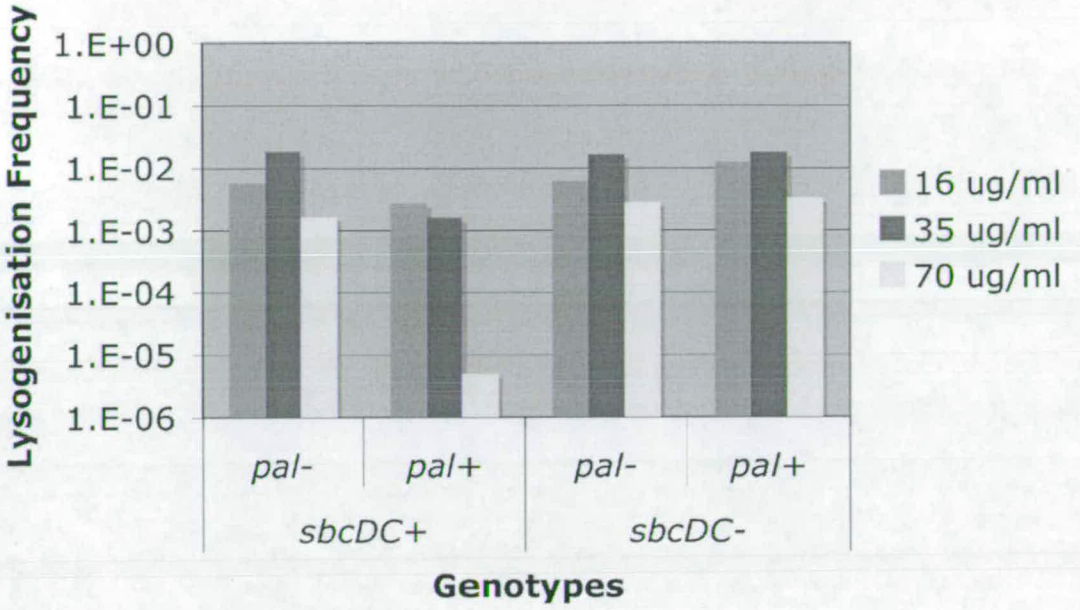


Figure 3.8: Effect of zeocin concentration on lysogenisation frequency of *rec*$^+$ AB1157 strains using lambda DRL246 (*pal*$^-$) and DRL282 (*pal*$^+$). Strains used were AB1157 (*sbcDC*$^+$) and DL515 (*sbcDC*$^-$).

These experiments have highlighted some pit-falls of this experimental system that may have given misleading results. In particular, lysogenisation efficiency with a palindrome containing phage in an *sbcCD*$^+$ host can drop dramatically when the zeocin antibiotic concentration is increased.

3.5.1 DNA double-strand break repair at a palindromic sequence.

The viability of lysogens derived from the P_{BAD}-*sbcDC* strains under SbcCD$^-$ conditions when grown in SbcCD$^+$ conditions support the hypothesis that a palindromic DNA sequence can be attacked by the SbcCD nuclease to give a DNA double-strand break that can be repaired by the RecBCD recombination pathway (but not the RecFOR recombination pathway). The lysogenisation experiments using the AB1157 background also gave results consistent with this model.

From these data, we cannot conclude whether DNA breaks are two-ended or one-ended (like a replication fork collapse). From current knowledge, both scenarios would require RecBCD-mediated recombination for repair.

3.5.2 Lysogenisation of *recB*$^-$ mutants.

In contrast to previous studies it was possible to lysogenise an SbcCD$^+$, *recB*$^-$ strain with a phage containing a palindrome. The success of this lysogenisation may reflect the difference between the genetic backgrounds, or the level of expression of SbcCD used in this and previous studies.

Interestingly, in the P_{BAD}-*sbcDC* background under SbcCD$^+$ conditions a *recF*$^-$ *recB*$^-$ double mutant had a very low lysogenisation frequency with the phage containing a palindrome. This suggests that initial survival of an SbcCD$^+$ strain containing the *recB*$^-$ mutation after lysogenisation with a phage containing a palindrome relies on the RecF protein. How RecF promotes survival is unclear. It may have a role in a cryptic repair pathway that requires RecF and proteins expressed from the lambda genome prior to repression. The notion of a short-lived positive effect of RecF is supported by the inviability of a stable *recB*$^-$ lysogen containing a palindrome when grown under SbcCD$^+$ conditions.

3.5.3 Limitations of the lysogenisation assay

In experiments carried out by Leach and collaborators (Leach et al., 1997) a limit was imposed on the range of mutants which could be tested due to overlapping antibiotic resistances between the lambda (kanamycin resistant) and the strains being tested. Cromie and collaborators (Cromie et al., 2000) created a strain of lambda (zeocin resistant) that contained the same 246 bp interrupted palindrome (with three internal mismatches) and was compatible with a wider range of mutants.

Zeocin is a copper-chelated glycopeptide that was isolated from *Streptomyces verticillus* and is structurally similar to bleomycin and phleomycin-type antibiotics. This family of antibiotics intercalate with DNA, and following removal of the copper compound can cleave the DNA leading to cell death (Reference: Invitrogen). A zeocin resistance gene, *Sh ble* (*Streptoalloteichus hindustanus*), encodes a 14 kiloDalton protein that works by binding the glycopeptide and inhibiting its DNA cleavage activity (Reference: Invitrogen). It is possible that the mode of action of

this drug has added another level of complexity for interpretation of the results discussed in this chapter.

Here, experiments were carried out to explore the effect of zeocin concentration on lysogenisation frequency using the recombination proficient AB1157 background *sbcDC*$^{+/-}$ strains. It was shown that increasing zeocin concentration resulted in reduced lysogenisation frequencies with the phage containing a palindrome in an *sbcDC*$^{+}$ strain. If the effective potency of zeocin varies between different batches, or is very sensitive to variations in media content then misleading lysogenisation frequencies may be calculated.

Chapter 4: A lambda-free system for study of palindrome viability in *E. coli*

4.1 Introduction

The lysogenisation experiments of previous studies (described in chapter three) (Cromie et al., 2000; Leach et al., 1997), and those carried out in chapter three have been useful in studies of palindrome inviability in *E. coli*. However, these methods have the disadvantage of introducing a lambda prophage into the *E. coli* chromosome, adding a level of complexity when interpreting results. For example, SOS induction in *E. coli*, as a result of DNA damage, will provoke lambda prophage induction and host death (Hendrix, 1983). Furthermore, with respect to the lysogenisation experiments carried out in different recombination mutants, the level and timing of SOS induction is dependent on the genetic background; for example, following UV irradiation the SOS response is severely reduced in a *recA*$^-$ mutant and is delayed but increased in a *recF*$^-$ mutant (Hegde et al., 1995; Whitby and Lloyd, 1995a). The survival of a particular mutant therefore depends on (1) its ability to repair SbcCD-induced damage at a palindromic sequence and (2) the likelihood of lambda induction. In a hypothetical situation, a particular mutant may be proficient in repairing DNA damage but have an increased SOS response that induces lambda. It is therefore difficult to determine the exact origin of the strain lethality.

This chapter describes the construction of a strain with a 246 bp interrupted palindrome at the *lacZ* locus in the *E. coli* chromosome (*lacZ::pal246*) and not in a lambda prophage. A P_{BAD}-*sbcDC* version of this strain has been used to re-examine

the genetic requirements for repair of SbcCD-induced damage at a palindromic sequence.

4.2 A new system for study of recombination events at a 246bp interrupted palindrome

4.2.1 Alteration of the *lacZ* gene region

For the purposes of the experiments in this thesis and other studies in the laboratory, two changes were made to the *lacZ* region of an MG1655 strain using PMGR; (1) the introduction of the $lacI^q$ mutation to the P_{lacI} to increase the strength of repression of expression from P_{lacZ}, (2) removal of a *chi* sequence in *lacZ* (for purposes explained in chapter six).

pDL1828 was designed to introduce the $lacI^q$ mutation, a one base-pair change to P_{lacI}, that increases the expression of the *lacI* gene and therefore confers stronger repression of the *lacZ* gene (Calos, 1978). It was constructed by cloning in pTOF24 an 800 bp crossover PCR product with homology to the *lacI* gene and the upstream *mhpR* gene, and containing the $lacI^q$ mutation at the centre (carried out by John Blackwood).

pDL2309 was designed to remove an undesirable *chi* sequence from the *lacZ* gene. It was constructed by cloning into pTOF24 an 800 bp crossover PCR product. This was generated using the primer pairs LacZ-chi-OF (has a *Sal*I restriction site at the 5' end), LacZ-chi-IR and LacZ-chi-IF, LacZ-chi-OR (has a *Pst*I restriction site at the 5' end) with genomic DNA template from the MG1655 strain. The novel DNA

fragment is homologous to a region of *lacZ* but with a *chi* sequence removed by introduction of two point mutations generating a unique *Spe*I restriction site. Potential clones were checked by *Spe*I restriction of PCR products generated using pKO-F and pKO-R2 primers. Correct clones were confirmed by sequencing using pKO-F and pKO-R2 primers.

Sequential PMGR with these plasmids removed the *chi* sequence from the *lacZ* gene (strain denoted DL1751) and introduced the *lacI*^{q} mutation (strain denoted DL1777). Subsequent PMGR with pDL1779 (see chapter 3) was used to replace the *sbcDC* promoter with the arabinose inducible promoter P_{BAD}. The final strain was named DL1803.

4.2.2 Introduction of the 246 bp palindrome at *lacZ*

As part of a laboratory-wide collaboration, the *lacZ* gene (and surrounding region) was chosen to accommodate a variety of genetic cassettes for genetic and physical analysis of recombination induced by trinucleotide repeats, unique restriction sites or palindromic sequences (John Blackwood, Rabaab Zahra and Martin White). This ensured the construction of a library of strains that could be used in a variety of experiments. In this study, a palindrome was introduced into the *lacZ* gene by PMGR using pDL2774 (derived from pLacD1 - see below).

The pLacD1 plasmid is a pTOF24 derivative (Link et al., 1997; Merlin et al., 2002) that was constructed using site-directed mutagenesis and basic cloning strategies. pLacD1 contains an 800 bp crossover PCR product homologous to the *lacZ* gene (and region) and contains some point mutations near the ATG start codon that

generate an *Mfe*I site for cloning, without altering *lacZ* expression (carried out by John Blackwood).

A pLacD1 derivate, denoted pDL2774, was constructed by cloning the 246 bp interrupted palindrome, liberated by *EcoR*I digestion of lambda SKK43 genomic DNA, into the *Mfe*I site of pLacD1. PMGR using pDL2774 was used to introduce the palindrome into DL1803 (see previous section) in the presence of glucose (represses expression of SbcCD) to give a strain named DL1979. The presence of the palindrome was confirmed by PCR across the *Mfe*I restriction site using the oligonucleotides Ex-testF and Ex-testR (see figure 4.1).

4.2.3 Moving the palindrome between strains

Because the introduction of the 246 bp palindrome using PMGR with pDL2774 was difficult, a gentamicin resistant gene was introduced at *cynX* approximately 5 kb away from *lacZ* in order to transfer the palindrome between strains by P1 transduction.

The gentamicin resistance gene was introduced in *cynX* using the plasmid pDL2812. Firstly, pDL2282 was constructed by cloning into pTOF24 an 800 bp crossover PCR product, with homology to part of the *cynX* gene. This novel DNA fragment has unique *Bam*HI and *Xho*I sites at its centre and was generated using the primer pairs TetO-CF1 (*Pst*I site at 5' end), TetO-CR1 and TetO-CF2, TetO-CR2 (*Nhe*I site at 5' end) with MG1655 genomic DNA as template.

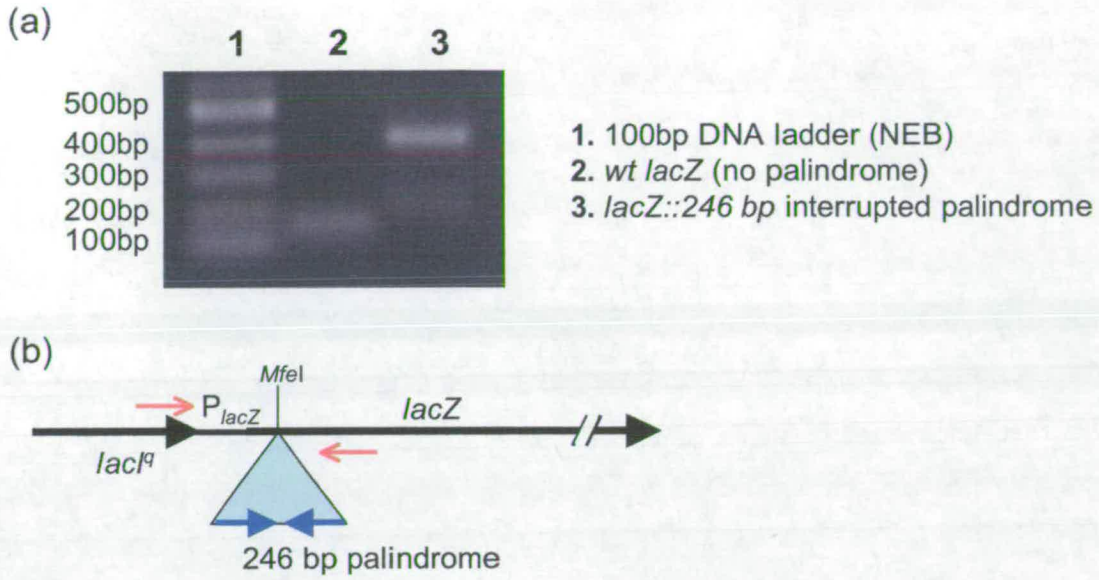


Figure 4.1: Insertion of a 246 bp interrupted palindrome at the *lacZ* locus. (a) PCR products from Ex-testF and Ex-testR PCR on a strain without the palindrome (DL2573) and a strain with the 246bp palindrome (DL2006) (b) a cartoon showing the important features of the *lacZ* region including the Ex-testF and Ex-testR binding sites (red arrows).

Secondly, pDL2812 was constructed by cloning into pDL2282 a DNA fragment, of approximately 750 bp, containing the gentamicin resistance gene. This was generated by PCR using the primer pair Gent_F (contains *Xho*I site) and Gent_R (contains *Bam*HI site) and using pLAU44 as template (Lau et al., 2003)).

PMGR with pDL2812 was used to introduce the gentamicin resistance gene into the chromosome of DL1979 (P_{BAD}-*sbcDC*, *lacZ::pal246*) to give the strain named DL2005. Similarly, it was introduced into DL1777, a version of DL1979 without a palindrome at *lacZ*, to give the strain named DL2551.

For these experiments homogeneous expression of SbcCD was required, so the *lacZ*$^+$, *cynX::Gm*^{R} and the *lacZ::pal246, cynX::Gm*^{R} regions (of DL2551 and DL2005, respectively) were introduced by P1 transduction into the strain DL1780 (P_{BAD}-*sbcDC*, Δ*lacZ4787*) to give strains named DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*). The presence of the palindrome was confirmed by PCR using Ex-test_F and Ex-test_R (for an example see figure 4.1).

4.3 Palindrome viability in an SbcCD$^+$ strain is dependent on the RecA and RecBCD recombination pathway

In the presence of glucose, *recA*$^-$, *recB*$^-$ or *recF*$^-$ mutations were introduced using P1 transduction to DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*). In addition, *recF*$^-$ *recB*$^-$ double mutants were made by P1 transduction of the *recB*$^-$ mutation into the *recF*$^-$ strains. Overnight cultures of these strains were grown in L broth supplemented with

0.5 % (v/v) glucose and dilutions were plated onto LB agar plates containing 0.5 % (w/v) glucose or 0.2 % (w/v) arabinose (see figure 4.2).

No difference in viability was observed when a *recA*⁻ strain without the palindrome was plated on plates containing glucose or arabinose, however a *recA*⁻ strain containing a palindrome was less viable on plates containing arabinose compared to glucose. This is consistent with previous studies suggesting that cleavage of a palindromic sequence by SbcCD requires one or more functions of RecA for repair and survival. A *recB*⁻ mutant containing the palindrome had decreased viability when grown on plates containing arabinose compared to glucose suggesting a requirement of RecBCD recombination for repair of an SbcCD-induced break. The *recF*⁻ mutant containing a palindrome was equally viable on plates containing glucose or arabinose, confirming the results of the lysogenisation experiments in chapter three and in contrast to a previous published study (Cromie et al., 2000). A *recF*⁻ *recB*⁻ double mutant containing a palindrome had the same viability as the isogenic *recB*⁻ single mutant when grown on plates containing glucose or arabinose. From these data, it is unlikely that the RecFOR pathway is required for cell survival following an SbcCD-induced break at a palindrome. Whereas RecBCD and RecA pathways seem essential in this process.

Interestingly, the *recB*⁻ (and *recB*⁻ *recF*⁻) strains that did not contain the palindrome showed decreased viability when plated on medium containing arabinose. This is reminiscent of the discovery of *sbcC*⁻ as a co-suppressor with *sbcB*⁻ of sensitivity to DNA damage (by UV irradiation) of *recB*⁻ mutants (Lloyd and Buckman, 1985).

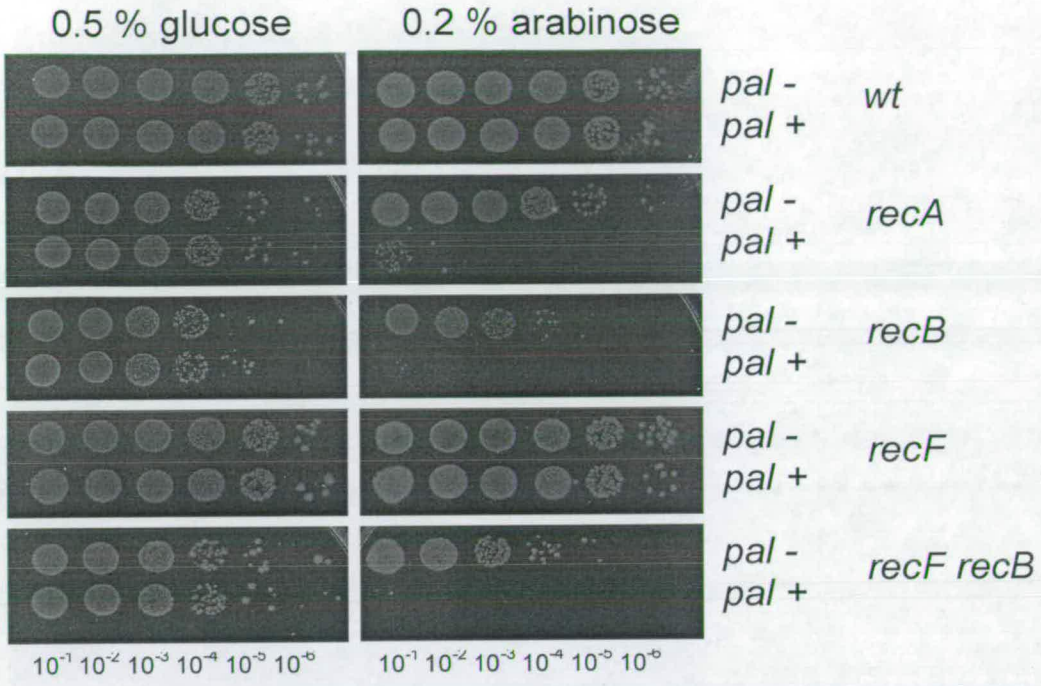


Figure 4.2: Effect of *recA*⁻, *recB*⁻, *recF*⁻ or *recB*⁻ *recF*⁻ mutations on the viability of P_{BAD}-*sbcDC* strains with or without a palindrome. Strains used were DL2573 (*lacZ*⁺ *rec*⁺), DL2006 (*lacZ::pal246 rec*⁺), DL2605 (*lacZ*⁺ *recA*⁻), DL2075 (*lacZ::pal246 recA*⁻) DL2606 (*lacZ*⁺ *recB*⁻), DL2076 (*lacZ::pal246 recB*⁻), DL2608 (*lacZ*⁺*recF*⁻), DL2077 (*lacZ::pal246 recF*⁻), DL2735(*lacZ*⁺ *recF*⁻ *recB*⁻) and DL2078 (*lacZ::pal246 recF*⁻ *recB*⁻).

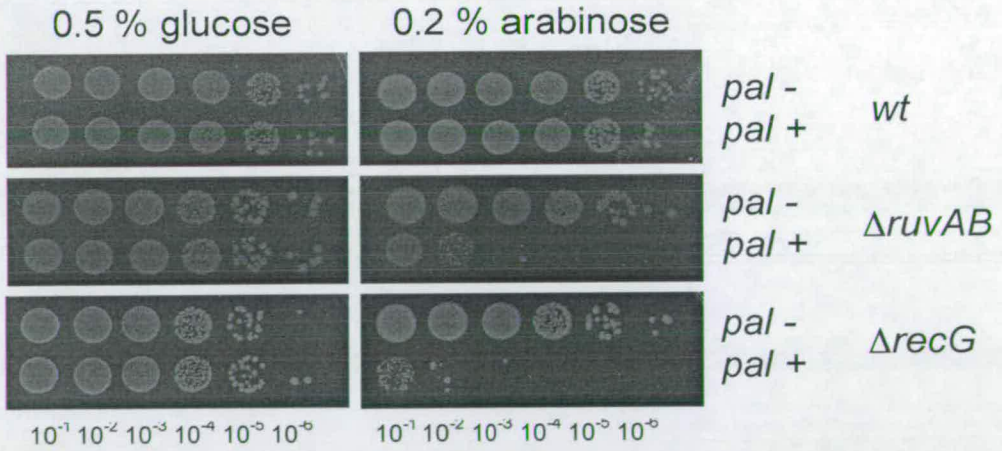


Figure 4.3: Effect of Δ*ruvAB* or Δ*recG* mutations on the viability of P_{BAD}-*sbcDC* strains with or without a palindrome. Strains used were DL2573 (*lacZ*⁺ *rec*⁺), DL2006 (*lacZ::pal246 rec*⁺), DL2800 (*lacZ*⁺ Δ*ruvAB*), DL2801 (*lacZ::pal246* Δ*ruvAB*), DL2610 (*lacZ*⁺ Δ*recG*) and DL2155 (*lacZ::pal246* Δ*recG*).

4.4 Palindrome viability in an SbcCD$^+$ strain requires both the RuvABC complex and RecG

RuvABC migrates and resolves Holliday junctions following RecA-mediated strand invasion. RecG is a helicase that seems to play a role in resolution of joint DNA molecules, such as Holliday junctions. Cromie and collaborators reported that palindrome viability in *sbcDC*$^+$ strains requires functional RuvABC and RecG (Cromie et al., 2000).

PMGR using pDL2757 (constructed by Ewa Okely; a pTOF24 derivative with a novel crossover PCR product of approximately 800 bp homology to the *znuB* and *yebB* (genes surrounding the *ruvAB* operon) and with an almost total deletion of *ruvAB*) in the presence of glucose was used to delete the *ruvAB* genes in DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*). It had previously been reported that a 246 bp interrupted palindrome stimulates recombination in *sbcC*$^-$ strains and that these strains require RuvABC for Holliday junction resolution (Cromie et al., 2000). However, no difficulties were encountered when introducing the Δ*ruvAB* deletion into an SbcCD$^-$ strain containing a palindrome. Overnight cultures of these strains were grown in L broth supplemented with glucose and dilutions were plated onto LB agar plates containing arabinose or glucose (see figure 4.3).

The Δ*ruvAB* strain containing the palindrome grew well on plates containing glucose confirming that there is little or no requirement for RuvABC in SbcCD$^-$ cells. In

contrast, this strain grew poorly on plates containing arabinose demonstrating a requirement of RuvABC for repair of an SbcCD-induced break.

In the presence of glucose, a Δ*recG* mutation was moved using P1 transduction into DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*). Overnight cultures of these strains were grown in L broth supplemented with glucose and dilutions were plated onto LB agar plates containing arabinose or glucose (see figure 4.3).

The Δ*recG* strain containing the palindrome showed decreased viability when plated on medium containing arabinose, this is consistent with previous work (Cromie et al., 2000) and suggests RecG has an important role in repair of an SbcCD-induced break.

4.5 Palindrome viability in an SbcCD$^+$ strain requires PriA

The PriA protein is required for loading the replication machinery at a D-loop formed following RecA-mediated strand invasion. PriA activities allow the loading of the DnaB helicase and processive replication on both leading and lagging strands, following recombination. *priA*$^-$ mutants are sick and can quickly accumulate suppressor mutations. Cromie and collaborators reported that palindrome viability in *sbcDC*$^+$ strains does not require functional PriA protein (Cromie et al., 2000).

In the presence of glucose and IPTG (see below), a *priA*$^-$ mutation was introduced using P1 transduction into DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*) strains carrying pAM*priA*$^+$. pAM*priA*$^+$ is a plasmid with an IPTG-dependent origin of replication, it has a functional copy of the *priA* gene (and regulatory region) and was

present to prevent the occurrence of *priA*⁻ suppressor mutations. Overnight cultures of these strains were grown in LB broth supplemented with glucose but not IPTG (to cure the cells of pAM*priA*⁺). In order to encourage slow growth of these strains to reduce the occurrence of suppressor mutations, dilutions were plated on minimal medium plates containing arabinose or glucose (see figure 4.4).

The *priA*⁻ strain containing a palindrome was less viable on media containing arabinose. This suggests that PriA is required for repair of an SbcCD-induced break in contrast to the results of Cromie and collaborators (Cromie et al., 2000).

4.6 Palindrome viability in an SbcCD⁺ strain requires the XerCD complex and dif sequence

Resolution of a Holliday junction can give a crossover product (and therefore a chromosome dimer). The XerCD complex directs a recombination event at the *dif* sequence to generate monomeric chromosomes that can be segregated to different daughters at cell division. Cromie and collaborators reported that palindrome viability in *sbcDC*⁺ strains requires functional XerCD and the *dif* sequence, suggesting that resolution of recombination products following repair of SbcCD-induced damage is biased towards giving crossover products (Cromie and Leach, 2000).

Using P1 transduction in the presence of glucose, the *xerC*⁻ or *dif*Δ mutations were introduced to DL2573 and DL2006. Overnight cultures of these strains were grown

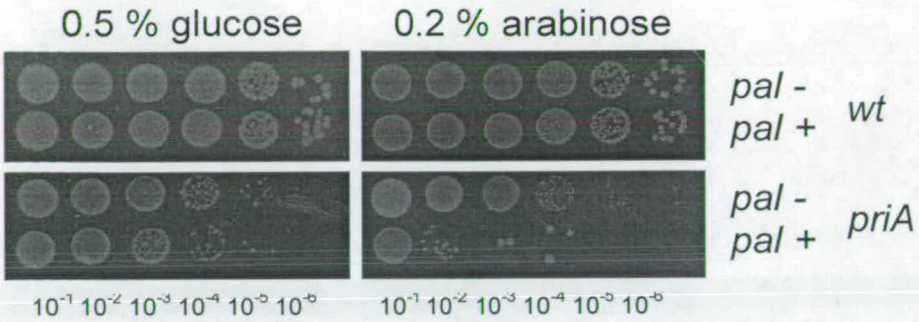


Figure 4.4: Effect of the *priA*⁻ mutation on the viability of P_{BAD}-*sbcDC* strains with or without a palindrome. Strains used were DL2573 (*lacZ*⁺ *rec*⁺), DL2006 (*lacZ::pal246 rec*⁺), DL2630 (*lacZ*⁺ *priA*⁻) and DL2294 (*lacZ::pal246 priA*⁻).

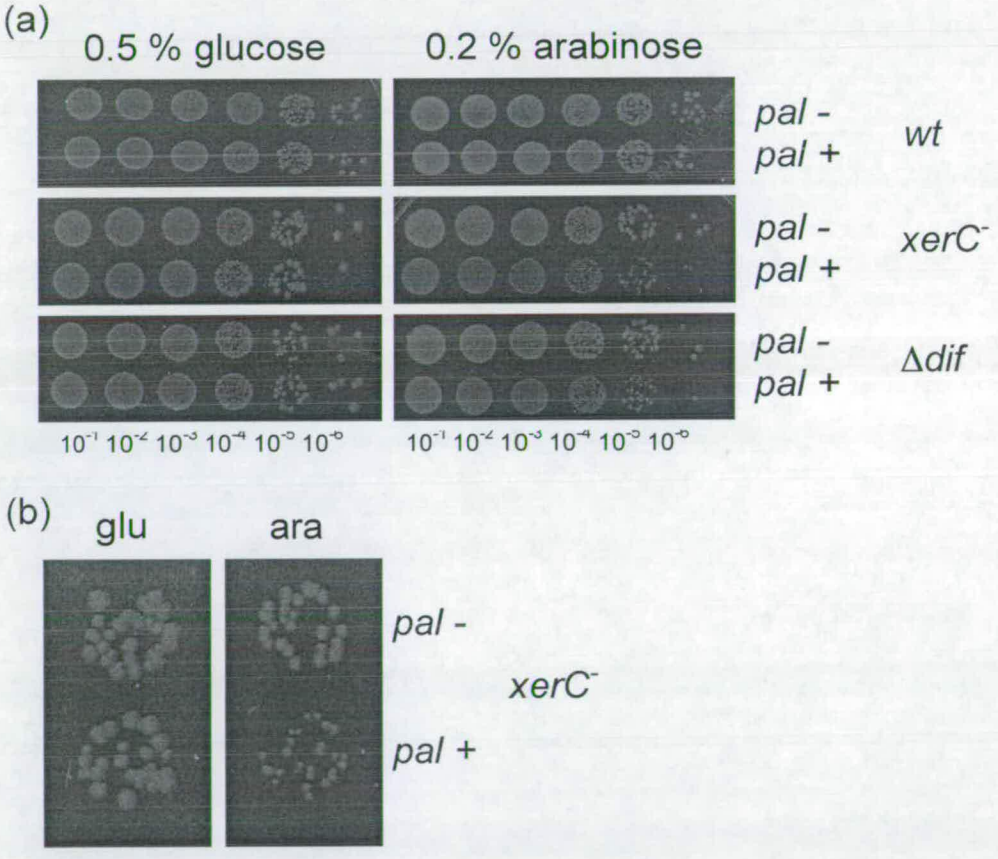


Figure 4.5: (a) **Effect of *xerC*⁻ or Δ*dif* mutations on the viability of P_{BAD}-*sbcDC* strains with or without a palindrome** (b) Close-up of the colonies of the *xerC*⁻ strains, 10^{-5} dilutions. Strains used were DL2573 (*lacZ*⁺ *rec*⁺), DL2006 (*lacZ::pal246 rec*⁺), DL2611 (*lacZ*⁺ *xerC*⁻), DL2237 (*lacZ::pal246 xerC*⁻), DL2612 (*lacZ*⁺ Δ*dif*) and DL2241 (*lacZ::pal246* Δ*dif*).

in L broth supplemented with glucose and dilutions were plated onto LB agar plates containing glucose or arabinose (see figure 4.5).

A *xerC*⁻ or *dif*Δ mutant containing a palindrome had an equal number of viable colonies on glucose or arabinose plates. Upon closer inspection, the colonies of *xerC*⁻ or *dif*Δ strains containing a palindrome, grown on plates containing arabinose were smaller than either the colonies of the same strains grown on plates containing glucose or the colonies of the *xerC*⁻ or *dif*Δ strains without a palindrome, grown on plates containing arabinose. These results suggest that there is a requirement for the XerCD complex and the *dif* sequence for full viability following an SbcCD-induced break at a palindrome. These results do not fully reproduce the results obtained from lysogenisation experiments by Cromie and Leach (Cromie and Leach, 2000). Their results suggested that repair of an SbcCD-induced break resulted in a strong bias towards resolution to a crossover product. The experiments here suggest that an undefined proportion of repair events result in crossover products.

4.7 Palindrome inviability of rec⁻ strains is not an artefact of over-expression of SbcCD

Palindrome viability experiments described so far in this chapter have been done with mutant strains following expression of *sbcDC* under the control of the P_{BAD} promoter. In this context, SbcCD is probably expressed above normal levels (Elise Darmon – unpublished data). To observe if DNA breaks are occurring in strains containing a palindrome with normal levels of SbcCD the *recA* gene was put under the control of the P_{BAD} promoter in wild-type *sbcDC*⁺ *lacZ::pal246* and Δ*sbcDC*

lacZ::pal246 strains. If breaks occur with normal levels of SbcCD then growth should be impaired in the presence of glucose when expression of the *recA* gene is repressed.

4.7.1 Construction of pDL2378

The pTOF24 derivative plasmid pDL2378 was constructed in order to replace the natural *recA* promoter and regulatory region, determined by Weisemann and Weinstock (Weisemann and Weinstock, 1991), with the *araBAD* promoter.

A 418 bp region of the *ygaD* gene (upstream of *recA*), the first 441 bp of the *recA* gene and a 338 bp region containing the *araBAD* promoter/regulatory sequence were PCR amplified using the primer pairs ygaD.F (a *Sal*I site at the 5' end), ygaD.R and RecA.F, RecA.R (a *Pvu*I site at the 5' end) and ParaBrecA.F, ParaBrecA.R with MG1655 genomic DNA as template. Crossover PCR, with the external primers ygaD.F and RecA.R, was used to fuse the three PCR products. This novel DNA fragment with a deletion of the *recA* promoter and insertion of the *araBAD* promoter was cloned into pTOF24 between *Pvu*I and *Sal*I restriction sites. The pDL2378 sequence was verified by sequencing with the oligonucleotides pKO.F and pKO.R2.

PMGR using pDL2378 in the presence of arabinose was used to replace the natural *recA* promoter with the P_{BAD} promoter in wild-type *sbcDC*$^+$ and Δ*sbcDC* strains in the BW27784 background (allowing homogeneous expression from P_{BAD}). P1 transduction in the presence of arabinose was used to introduce the 246 bp interrupted palindrome into these strains.

4.7.2 Depletion of RecA in an *sbcDC*$^+$ strain containing a palindrome leads to inviability

Overnight cultures of the above strains were grown in L broth supplemented with 0.2 % (v/v) arabinose and plated onto LB medium containing 0.2 % arabinose or 0.5 % glucose (figure 4.6).

The wild-type *sbcDC*$^+$ strain has reduced viability when grown on plates containing glucose compared to arabinose, or the Δ*sbcDC* strain grown on plates containing glucose. The Δ*sbcDC* strain grew equally well on plates containing arabinose or glucose. This suggests that a wild type level of SbcCD can cause a DNA DSB at a palindrome that requires RecA for repair. This confirms the results obtained with viability assays using the P_{BAD}-*sbcDC* strains.

4.8 The SOS response is activated following SbcCD-induced damage at a 246 bp palindrome

4.8.1 Palindrome viability in an SbcCD$^+$ strain requires a cleavable LexA protein

In *E. coli*, extensive DNA damage, for example following UV irradiation or treatment with mitomycin C, (Whitby and Lloyd, 1995a) leads to an SOS response that involves increased expression of a set of genes that aid DNA repair and increase the chance of cell survival. This response is elicited by self-cleavage of the LexA repressor protein which is stimulated by a RecA-ssDNA filament. In *E. coli*, an un-cleavable mutant of the LexA protein, encoded by *lexA3*, will remain bound to

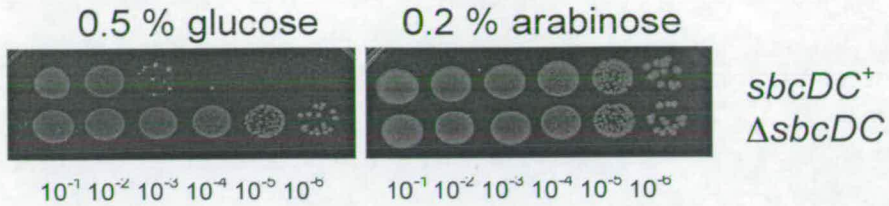


Figure 4.6: Effect of Δ*sbcDC* mutation on the viability of P_{BAD}-*recA* strains containing a palindrome. Strains used were DL2537 (*lacZ::pal246 sbcDC*$^+$) and DL2538 (*lacZ::pal246* Δ*sbcDC*).

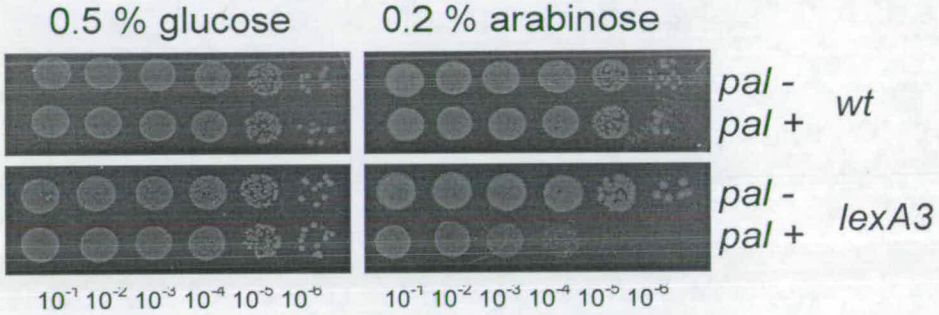


Figure 4.7: Effect of the *lexA*3 mutation on the viability of P_{BAD}-*sbcDC* strains with or without a palindrome. Strains used were DL2573 (*lacZ*$^+$ *rec*$^+$), DL2006 (*lacZ::pal246 rec*$^+$), DL2829 (*lacZ*$^+$ *lexA3*) and DL2830 (*lacZ::pal246 lexA3*).

repressor sequences following DNA damage, suppressing the SOS response (Little et al., 1980).

In the presence of glucose the *lexA3* mutation was introduced using P1 transduction into DL2573 (*lacZ*$^+$) and DL2006 (*lacZ::pal246*). Overnight cultures were grown in L broth supplemented with glucose and dilutions were plated onto LB medium containing arabinose or glucose (see figure 4.7).

A strain with the *lexA3* mutation and containing a palindrome had reduced viability on plates containing arabinose. This suggests that the cleavage of LexA has an important role for cell survival following an SbcCD-induced DNA DSB at a palindromic sequence.

4.8.2 An increased SOS response is observed in an *sbcDC*$^+$ strain containing a palindrome

PMGR using pTOFsbcDC (constructed by Elise Darmon) was used to delete the *sbcDC* operon from DL1777 (*lacZ*$^+$) and DL2859 (*lacZ::pal246*) to give strains named DL2151 and DL2874, respectively. The SOS reporter plasmid pGB150 (a gift from Garry Blakely, The University of Edinburgh) was introduced into each of these strains by transformation. pGB150 is a low-copy number plasmid that contains the *gfp* gene under the control of the SOS-inducible *sfiA* promoter. Under normal growth conditions LexA is bound to the *sfiA* promoter and cells are not fluorescent.

Overnight cultures of these strains were grown in LB broth supplemented with chloramphenicol (for pGB150 selection). The cells were then diluted 100-fold in

fresh medium and grown for four additional hours. Multiple pairs of brightfield and fluorescence images were collected and individual cell lengths and average cell fluorescence were measured for about 1000 cells for each strain (see figure 4.8).

The population of cells with wild-type *sbcDC*$^+$ and with *lacZ;:pal246*, displays higher incidence of individual cells with elevated fluorescence. This indicates that the SOS response follows an SbcCD-induced DNA DSB at a palindrome. This reflects the requirement for a cleavable LexA protein for viability after an SbcCD-induced DNA DSB at a palindromic sequence.

4.9 Discussion

4.9.1 RecA has multiple roles for viability of strains containing a palindrome

This study demonstrates the importance of RecA for the survival of cells following an SbcCD-induced DNA DSB at a palindromic sequence. Firstly, RecA is required to facilitate DNA strand exchange between the broken DNA end and an intact homologous DNA molecule, thus initiating homologous recombination and repair (see below and figure 4.9). Without this function a cell could not repair a break. Secondly, in *E. coli*, RecA is required for induction of the SOS response. Following DNA damage, RecA is loaded onto ssDNA and forms a nucleoprotein filament that stimulates self-cleavage of the LexA repressor followed by up-regulation of the SOS genes. This study shows that there is a requirement for a cleavable LexA protein for survival of cells following an SbcCD-induced DNA DSB at a palindromic sequence, and that there is an increase in expression from SOS promoters following damage,

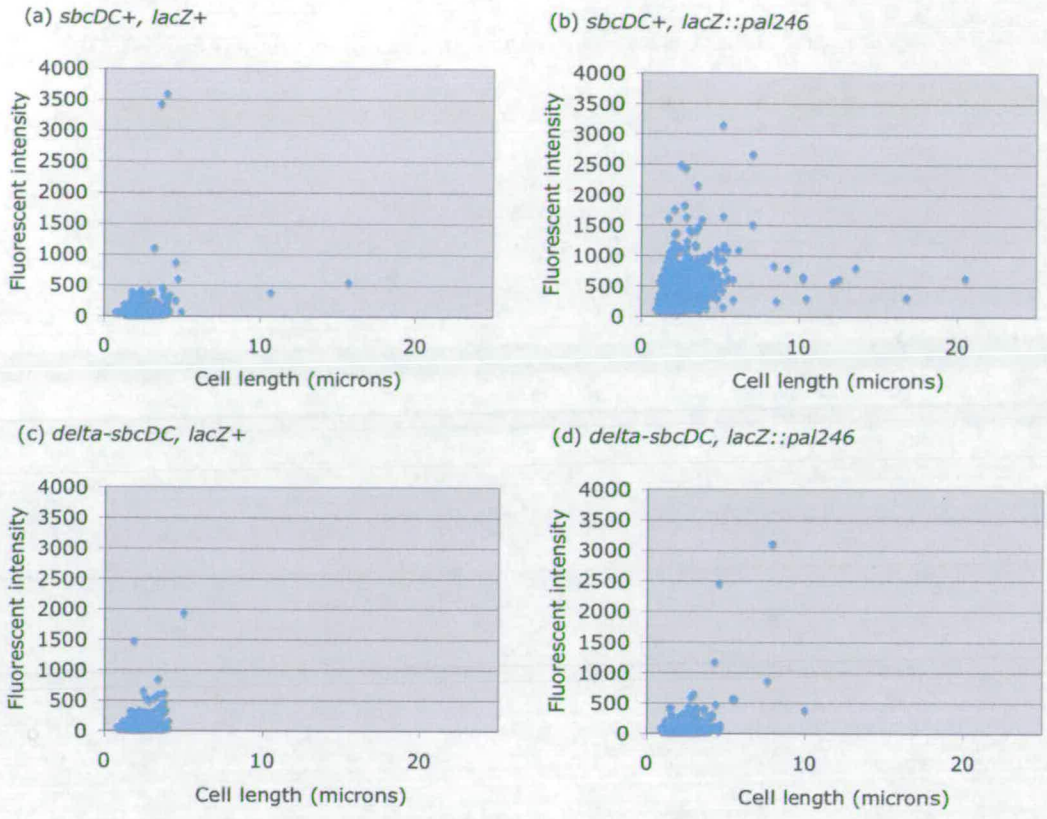


Figure 4.8: Effect of a 246 bp interrupted palindrome and SbcCD on SOS induction. Scatter plots of cell length against fluorescent intensity in various strains. Fluorescence from the green fluorescent protein, whose gene is present on a low copy number plasmid (pGB150) and under the control of the SOS-inducible *sfiA* promoter provides an indication of the level of SOS induction individual cells. Strains used were (a) DL1777 (*sbcDC*$^+$ *lacZ*$^+$), (b) DL2859 (*sbcDC*$^+$ *lacZ::pal246*), (c) DL2151 (Δ*sbcDC lacZ*$^+$) and (d) DL2874 (Δ*sbcDC lacZ::pal246*).

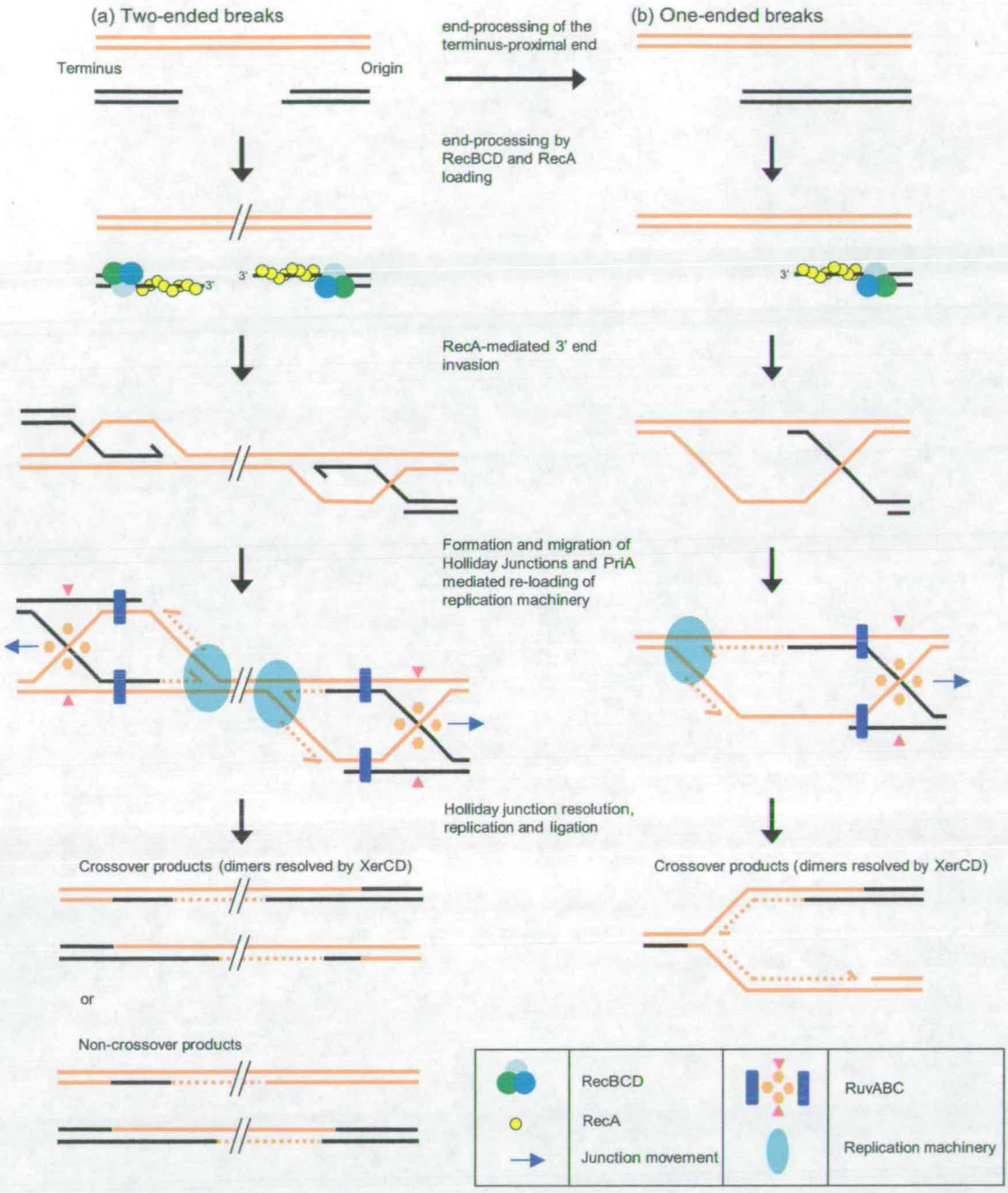


Figure 4.9: Schematic drawing showing repair of DNA double-strand breaks following SbcCD cleavage of a palindromic sequence. Repair followed by a (a) two-ended DNA DSB or (b) one-ended DNA DSB at a palindromic sequence. Resolution of Holliday junctions by RuvC (indicated by pink arrow heads) or an alternative RecG mechanism results in crossover or non-crossover products,. The origin and terminus sides are indicated in the first diagram only but remain the same throughout.

presumably stimulated by loading of RecA protein onto ssDNA. RecA and RuvAB are upregulated following cleavage of LexA and this facilitates the repair of DNA damage. This study does not rule out the possibility that other SOS-upregulated genes are involved in the survival of cells following SbcCD-induced DNA damage (for example the genes involved in trans-lesion synthesis; see figure 4.10).

4.9.2 Repair of an SbcCD-induced break requires RecBCD recombination

Data in this chapter support results in chapter three and a model where SbcCD causes a DNA double-strand break at a palindromic sequence. Palindrome viability in an $SbcCD^+$ strain is dependent on a number of the recombination proteins (a model for repair is suggested; figure 4.9). The presynapsis stage of recombination appears to be entirely dependent on the RecBCD complex and not the RecFOR proteins as previously suggested by Cromie and collaborators (Cromie et al., 2000). This is consistent with either one-ended DNA breaks or DNA breaks with two equivalent ends. However, these data do not rule out the possibility that a break could occur in a non-replicating chromosome (although this is less likely, see discussion in chapter five). Data here implicate roles for the RuvABC complex and the RecG protein; presumably for resolution of Holliday junctions. Palindrome viability requires functional PriA protein probably to load the replication machinery at D-loop junctions following RecA-mediated strand invasion. The requirement for PriA contradicts studies by Cromie and collaborators (Cromie et al., 2000), but the results in this study are probably more reliable as the method for introduction of the *priA*

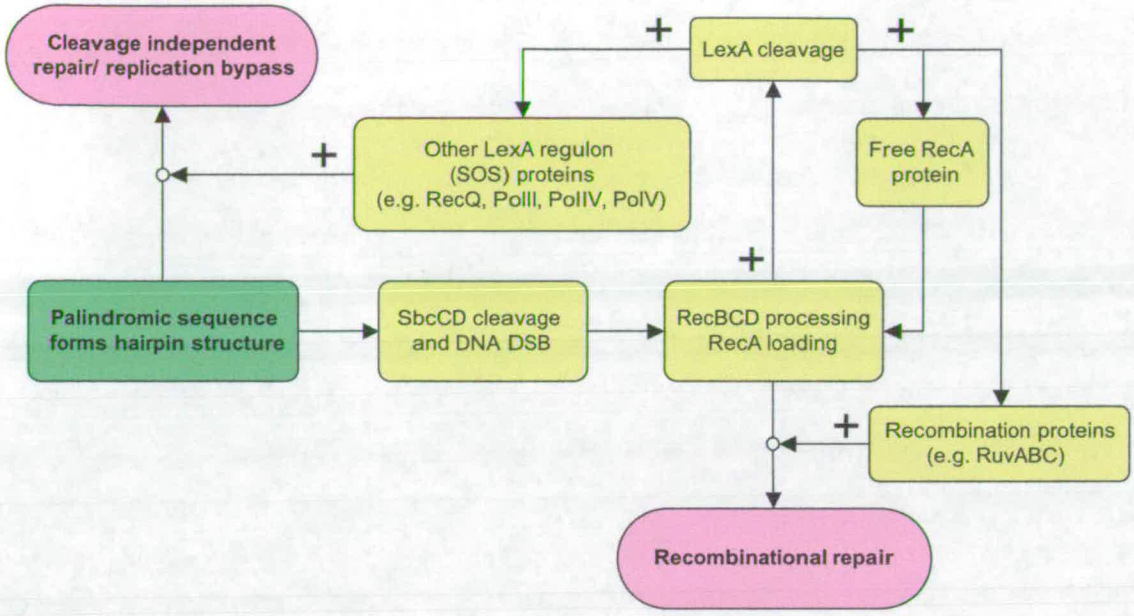


Figure 4.10: Possible contributions of the SOS response for repair or avoidance of SbcCD-induced breaks at a palindromic sequence.

mutation in these experiments was less likely to have resulted in suppressor mutations giving false-negative results.

A previous study by Cromie and collaborators (Cromie et al., 2000) found a requirement of *ruvABC* in an *sbcDC*$^-$ strain containing a palindrome. Upon the introduction of a *recF*$^-$ (or *recO/R*$^-$) mutation to this strain there was no longer a requirement for *ruvABC* for palindrome viability. This suggested that a palindrome could stimulate RecFOR-stimulated recombination in the absence of SbcCD. In this study, in an SbcCD$^-$ strain containing a palindrome, it was possible to delete the *ruvAB* genes. Although this does not rule out the possibility that a palindrome can stimulate RecFOR recombination in an SbcCD$^-$ strain it does suggest that this stimulation is either weak or that there is an alternative resolution pathway.

4.9.3 Repair of SbcCD-induced damage is not associated solely with crossover

A previous study in *E. coli* suggested that SbcCD-induced damage is repaired to generate crossover products and therefore chromosome dimers (Cromie and Leach, 2000). Re-examination of *xerC*$^-$ / *dif*Δ mutants in this study implicates a minor role for XerCD / *dif*. Crossover events occur at a detectable level but these data are not sufficient to calculate a ratio of crossover to non-crossover events.

4.9.4 SbcCD has a palindrome-independent negative effect on *recB*$^-$ strains

In these experiments, a *recB*$^-$ strain without a palindrome showed decreased viability on plates containing arabinose. These results reflect a previous study (Lloyd and

Buckman, 1985) and suggest that in the absence of RecB protein, SbcCD (albeit in these experiments SbcCD protein is at higher than normal levels) has a detrimental effect on cell viability. The *sbcC* (and *sbcD*) genes were first described as suppressor mutants necessary for suppression of *recB*$^-$ *sbcB*$^-$ UV sensitivity, and *recB*$^-$ *sbcC*$^+$ strains were reported as showing poor growth (Lloyd and Buckman, 1985). The actual role of the SbcCD complex in *E. coli* remains a mystery (in spite of eukaryotic homologues having defined roles). From these studies it is not clear what the detrimental effect of SbcCD is, or what its natural cellular function is (or whether these are related). The P_{BAD}-*sbcDC* strains may prove to be useful tools for studying this.

Chapter 5: *in vivo* physical evidence of two-ended DNA double-strand breaks

5.1 Introduction

The genetic evidence described in previous *E. coli* studies (Cromie et al., 2000; Leach et al., 1997) and this thesis strongly suggest that SbcCD can cleave a hairpin structure formed by a palindromic sequence generating a DNA DSB *in vivo* (discussed in chapter four). However, so far no physical evidence of this phenomenon has been presented. A number of laboratories have previously used refined agarose gel electrophoresis techniques and Southern blotting to detect DNA DSBs in the *E. coli* chromosome (Bidnenko et al., 2002; Meddows et al., 2004). Bidnenko and collaborators used pulsed-field gel electrophoresis (outlined in chapter two) to demonstrate that replication fork arrest at prematurely positioned *Ter* sequences led to fork collapse and to a DNA break. Detection of these breaks by this method was only possible in cells lacking RecBCD, presumably because the broken DNA molecules are rare and susceptible to the nuclease activities of the RecBCD complex. Meddows and collaborators used standard agarose gel electrophoresis to observe chromosomal DNA DSBs at a unique 18 bp DNA restriction site (I-*Sce*I$_{cs}$) following expression of the I-*Sce*I endonuclease. Broken molecules were detected even in *recBCD*$^{+}$ strains probably because they occur at a very high frequency and/or they are less susceptible to the nuclease activities of the RecBCD complex.

This chapter describes pulsed-field gel electrophoresis and Southern blotting experiments to visualise SbcCD induced DNA DSBs at a palindromic sequence *in*

vivo. This chapter also describes the construction of a control strain similar to that used by Meddows and collaborators (Meddows et al., 2004) known to give two-ended DNA double-strand breaks at a specific locus.

5.2 Construction of a P_{BAD}-*I-SceI* lacZ::*I-SceI*$_{cs}$ control strain

The I-*Sce*I endonuclease is found within a mitochondrial gene intron in *S. cerevisiae* (Jacquier and Dujon, 1985) and recognises and cleaves a specific 18 bp DNA sequence (I-*Sce*I$_{cs}$) that does not exist naturally in the *E. coli* chromosome. Meddows and collaborators engineered the I-*Sce*I$_{cs}$ into various loci of the *E. coli* chromosome and observed DNA double-strand breaks following conditional expression of the I-*Sce*I gene (Meddows et al., 2004).

Construction of pDL2655 was carried out as part of a collaborative study in this laboratory (John Blackwood, Katy Tully, Rachel Weisz and Martin White). It is a derivative of pKO3 designed to replace the *araBAD* locus with the P_{BAD}-I-*Sce*I construct. It has 400 bp homology to a region upstream of the *araB* gene (including the P_{BAD} promoter) and 400bp downstream of the *araB* gene (within *araD*). The I-*Sce*I gene is located between homology arms (in a position under the control of the arabinose promoter). The P_{BAD}-I-*Sce*I construct was inserted into the chromosome of an MG1655 background strain with *lacZ::*I-*Sce*I$_{cs}$ (constructed by Katy Tully). Finally, the *recB*$^-$ mutation was introduced into this strain by P1 transduction in the presence of glucose. This strain was shown to be inviable on L agar plates supplemented with arabinose (Martin White, unpublished data), consistent with previous work by Meddows and collaborators (Meddows et al., 2004).

5.3 Initial PFGE analysis reveals SbcCD palindrome-induced DNA DSBs

The P_{BAD}-*sbcDC lacZ*$^{+}$ *recB*$^{-}$ and P_{BAD}-*sbcDC lacZ::pal246 recB*$^{-}$ strains (described in chapter three, DL2606 and DL2076 respectively), the P_{BAD}-I-*Sce*I *recB*$^{-}$ strain (described above, DL2643) and a Δ*sbcDC lacZ::pal246 recB*$^{-}$ strain (DL2727, constructed using pTOFsbcDC and by P1 transduction) were grown in LB broth supplemented with arabinose or glucose. Samples were taken periodically during exponential growth and agarose plugs containing DNA were prepared for each. The chromosomal DNA was digested with the *Not*I restriction enzyme and fragments separated on a pulsed-field gel. *Not*I has 23 recognition sites in the *E. coli* chromosome (Colibri website; http://genolist.pasteur.fr/Colibri/) and the *lacZ* locus lies in a 361 kb fragment with *Not*I cleavage sites 59 kb towards the origin and 302 kb towards the terminus (see figure 5.1 (a)).

After separation on the PFGE-apparatus, the gel was soaked in ethidium bromide solution and the DNA visualised under a UV lamp. No differences were observed between the 361 kb band of different samples and no novel bands were observed at 59 kb or 302 kb following expression of SbcCD or I-*Sce*I (see figure 5.1(b)).

Using a capillary transfer method, the DNA was transferred to a nylon membrane and fragments either side of *lacZ* were visualised using 1 kb DIG-labelled DNA probes generated by PCR using primer pairs binding in the *codB* gene (DIGcodBF and DIGcodBR; 10 kb away from *lacZ*) or the *yaiO* gene (DIGyaiOF and DIGyaiOR; 15 kb away from *lacZ*) (see figure 5.2).

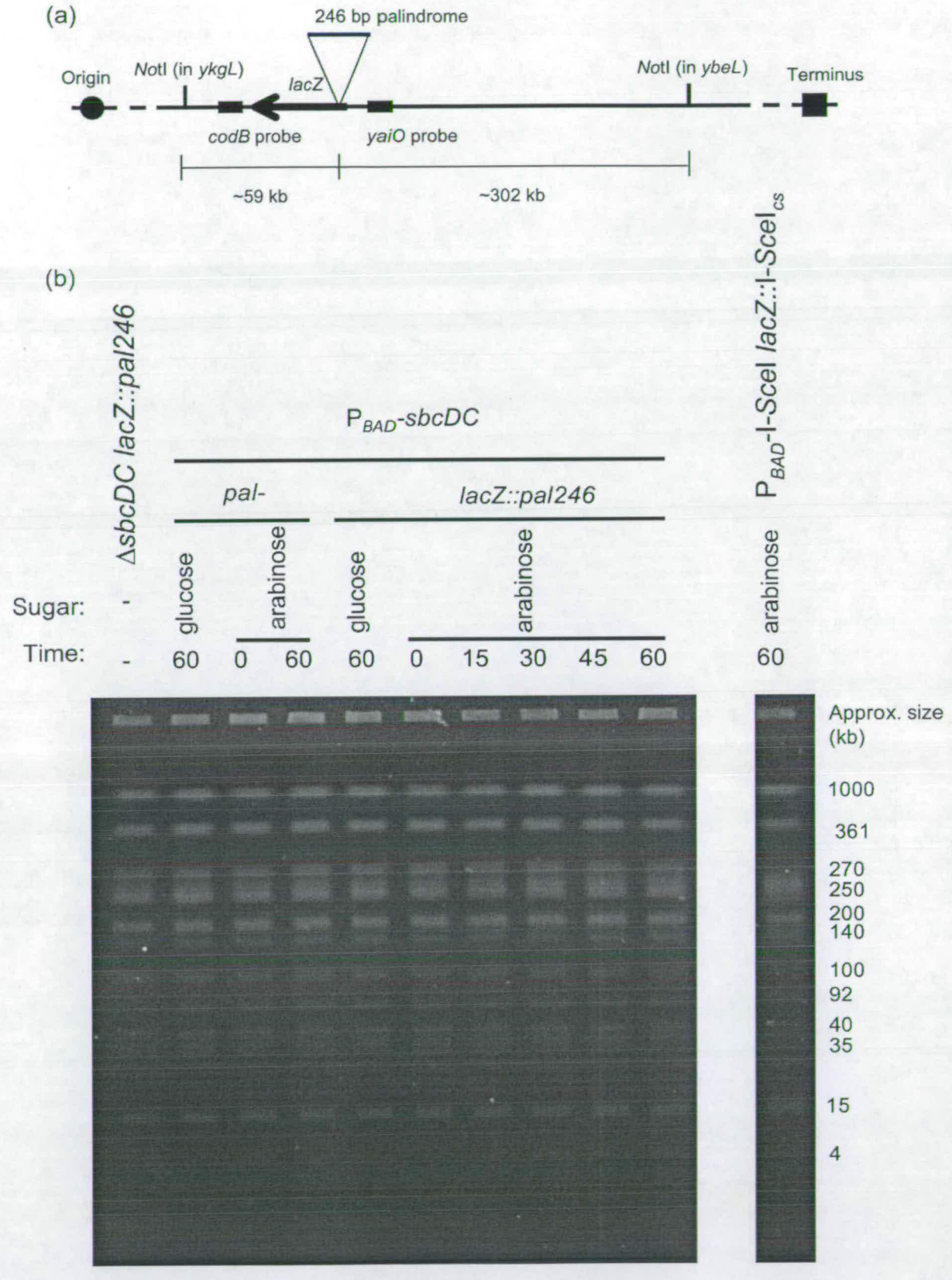


Figure 5.1: Pulsed-field gel electrophoresis of *Not*I digested chromosomal DNA of P_{BAD}-*sbcDC recB*⁻ strains with or without a palindrome (a) Cartoon showing the relative positions of important sites for *Not*I PFGE analysis (not to scale), (b) Ethidium bromide stained gel of DNA from *recB*⁻ strains. Strains used are DL2727 (Δ*sbcDC lacZ::pal246 recB*⁻), DL2606 (P_{BAD}-*sbcDC lacZ*⁺ *recB*⁻), DL2076 (P_{BAD}-*sbcDC lacZ::pal246 recB*⁻) and DL2643 (P_{BAD}-I-*Sce*I *lacZ::*I-*Sce*I$_{cs}$ *recB*⁻).

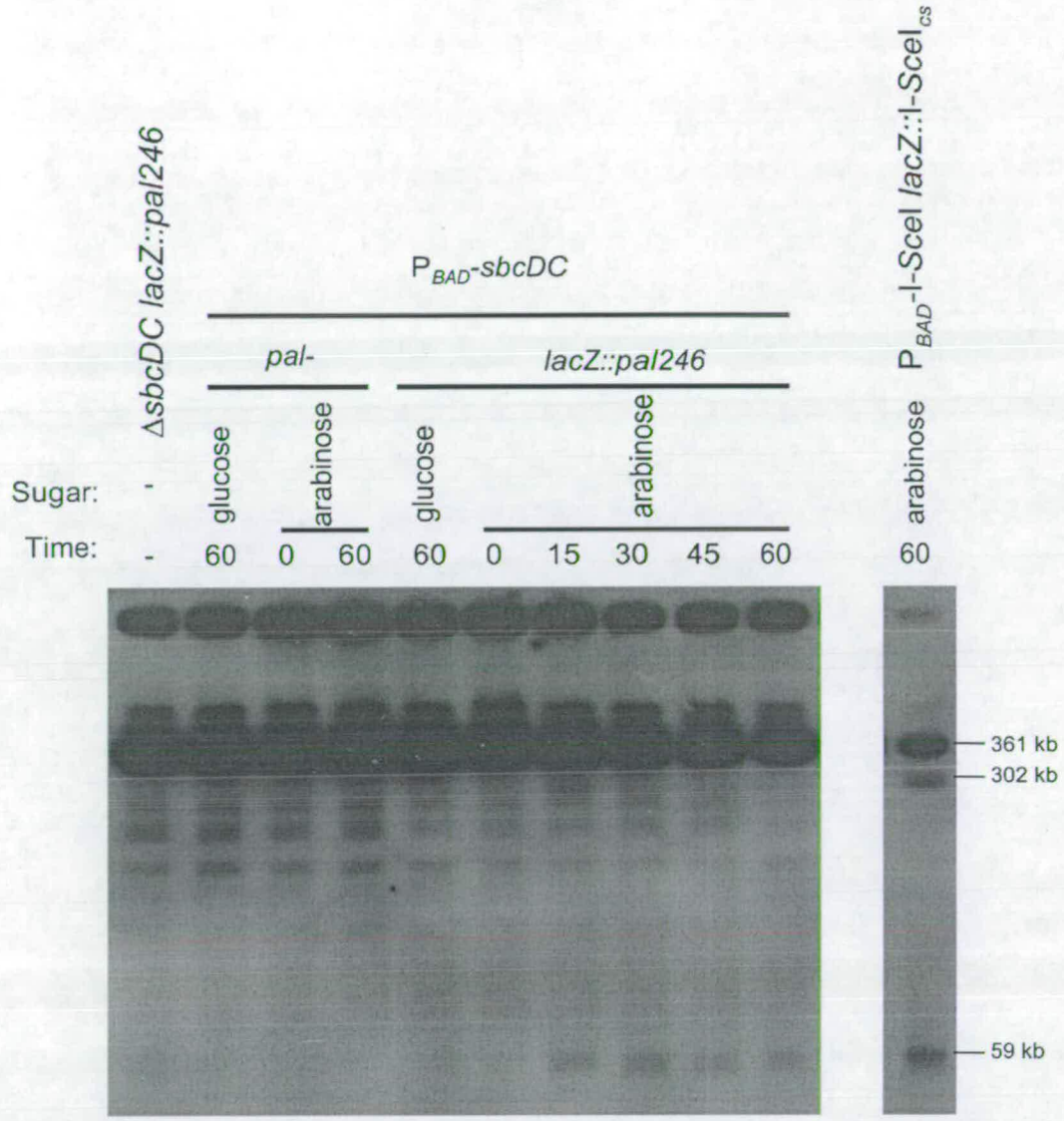


Figure 5.2: Southern blot of *Not*I digested chromosomal DNA of P$_{BAD}$-*sbcDC recB*$^-$ strains with or without a palindrome using *codB* and *yaiO* probes. Strains used were DL2727 (Δ*sbcDC lacZ::pal246 recB*$^-$), DL2606 (P$_{BAD}$-*sbcDC lacZ*$^+$ *recB*$^-$), DL2076 (P$_{BAD}$-*sbcDC lacZ::pal246 recB*$^-$) and DL2643 (P$_{BAD}$-I-*Sce*I *lacZ::*I-*Sce*I$_{cs}$ *recB*$^-$). The positions of the important *Not*I sites and the probe binding sites are indicated in figure 5.1 (a).

The Southern blot shown in figure 5.2 reveals that the *yaiO* and *codB* DIG-labelled probes bound 59 kb and 302 kb DNA fragments created by the I-*Sce*I cleavage of *lacZ*::I-*Sce*I_{cs} following growth in LB broth containing arabinose. Similarly a novel 59 kb band was observed following SbcCD expression in a strain containing *lacZ::pal246*. Due to problems of high background and an abundance of the intact 361 kb fragment it was difficult to determine whether the corresponding novel 302 kb fragment was also present in the DNA of the *lacZ::pal246* strain following SbcCD expression. The novel 59 kb fragment was not detected in the DNA of any of the palindrome-free (*lacZ*$^+$) samples. Interestingly, there was a weak band of 59 kb detected in the DNA of the P_{BAD}-*sbcDC lacZ::pal246* strain when grown in LB broth supplemented with glucose and the Δ*sbcDC lacZ::pal246* strain, indicating that there may be an SbcCD-independent breakage pathway.

5.4 Construction of strains with additional I-SceI_{cs} for in vitro restriction

Digestion with the *Not*I enzyme was not ideal for this experimental system and no other restriction enzyme was found with a better distribution of recognition sites in the *E. coli* chromosome. As part of a collaborative project within this laboratory a strategy was adopted to introduce two recognition sequences for the I-*Sce*I enzyme either side of the *lacZ* locus (John Blackwood). Two derivatives of pKO3 were constructed by crossover PCR and standard cloning methods for the introduction of the 18 bp I-*Sce*I recognition sequence at the *proA* gene (~104 kb towards the origin) (pDL2236) and the *tsx* gene (~70 kb toward the terminus) (pDL2755) (see figure 5.3

(a); carried out by John Blackwood). Using PMGR the I-*Sce*I_{cs} restriction sites were introduced into the chromosome of P_{BAD}-*sbcDC lacZ*$^+$, the P_{BAD}-*sbcDC lacZ::pal246* and the Δ*sbcDC, lacZ::pal246* strains.

5.5 PFGE analysis reveals SbcCD causes DNA DSBs that have two ends

In the presence of glucose the *recB*$^-$ mutation was introduced using P1 transduction into the above strains. Cultures were grown in LB broth supplemented with arabinose or glucose (as appropriate). Samples were taken periodically during exponential growth and agarose plugs containing DNA were prepared for each. The chromosomal DNA was digested with the I-*Sce*I enzyme and fragments separated on a pulsed-field gel.

The gel was initially soaked in ethidium bromide solution and the DNA visualised under a UV lamp (data not shown), this revealed the 174 kb and 4,465 kb chromosomal fragments but was not sensitive enough to reveal any novel bands. The DNA was transferred by a capillary method to a positively charged nylon membrane and probed using the DIG-labelled *codB* and *yaiO* probes (see figure 5.3 (a)).

As shown in figure 5.3 (b) probes bound to a high level of unbroken 174 kb fragment in all of the samples. Following expression of SbcCD, novel 104 kb and 70 kb DNA fragments were visible in the *lacZ::pal246* strain. Neither of these novel bands was detected in the DNA of any of the palindrome-free (*lacZ*$^+$) samples, the P_{BAD}-*sbcDC*

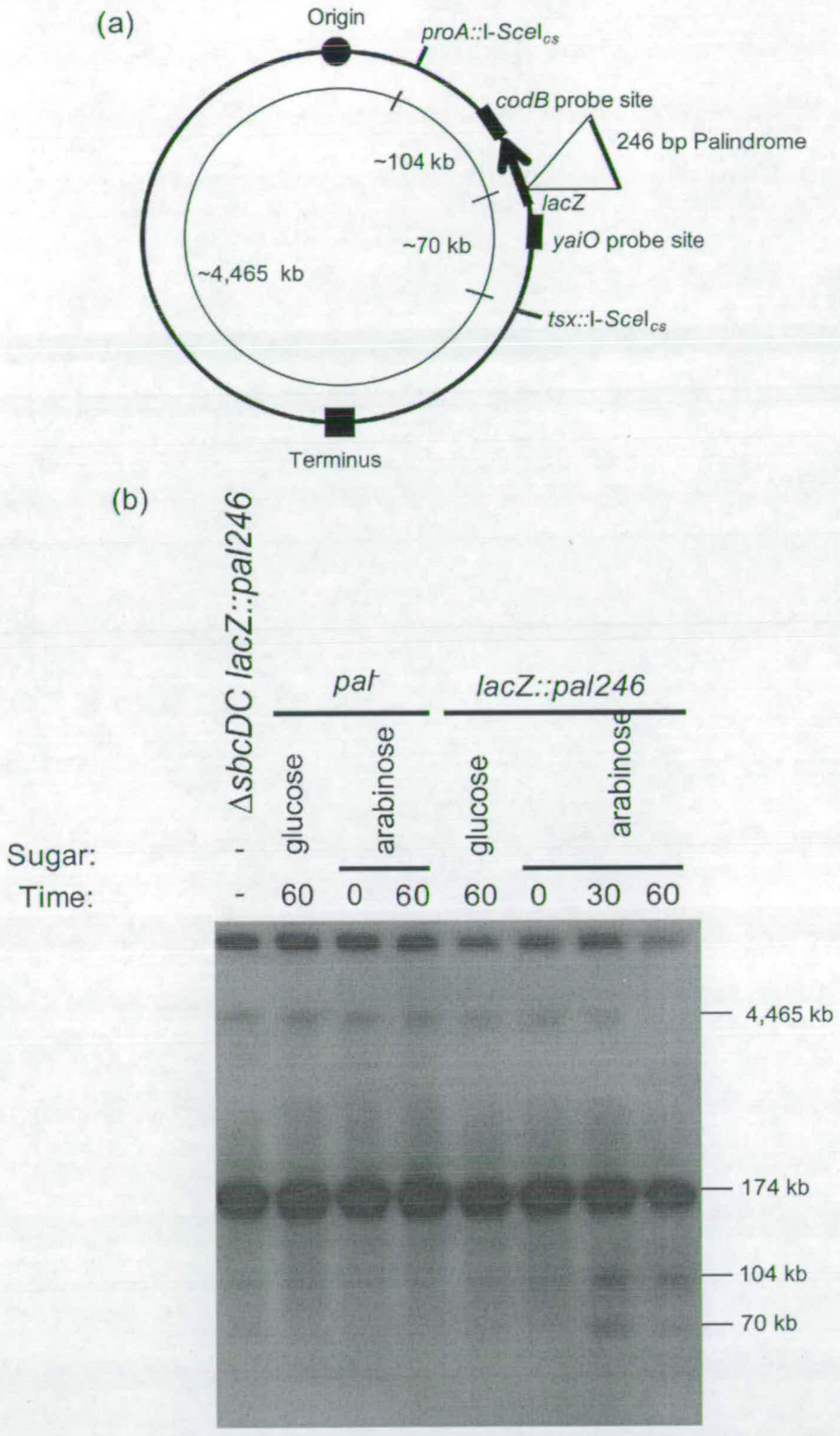


Figure 5.3: Southern blot of I-*Sce*I digested chromosomal DNA of P$_{BAD}$-*sbcDC recB*$^{-}$ strains with or without a palindrome using *codB* and *yaiO* probes (a) Cartoon showing the relative positions of important sites for I-*Sce*I PFGE analysis (not to scale) (b) Southern blot of I-*Sce*I digested DNA from various *recB*$^{-}$ strains. Strains used were DL2797 (P$_{BAD}$-*sbcDC lacZ*$^{+}$ *recB*$^{-}$), DL2798 (P$_{BAD}$-*sbcDC lacZ::pal246 recB*$^{-}$) and DL2799 (Δ*sbcDC lacZ::pal246 recB*$^{-}$).

lacZ::pal246 grown in LB broth supplemented with glucose, or the Δ*sbcDC lacZ::pal246* strain.

The appearance of the two novel bands in this gel demonstrates that SbcCD cleavage of a palindromic sequence generates a two-ended DNA DSB. This could be explained by a break occurring behind the replication fork or in a non-replicating chromosome but not by cleavage leading to replication fork collapse.

Noticeably, broken bands detected following expression of SbcCD in the *lacZ::pal246* strain become less intense with time, especially the 70 kb *yaiO* band. Broken DNA ends are probably being degraded beyond the probe binding sites.

5.6 The two ends of the DNA DSB are degraded in the same manner by a RecB-independent mechanism

In order to investigate whether the degradation of the two DNA ends is equal another gel was run using DNA samples from the P_{BAD}-*sbcDC lacZ*$^+$ and P_{BAD}-*sbcDC lacZ::pal246* strains, with subsequent transfer to a positively charged nylon membrane. Two new 1 kb DIG-labelled DNA probes were generated by PCR, binding the *perR* gene (using primer pair DIGperR_F and DIGperR_R) or the *malZ* gene (using primer pair DIGmalZ_F and DIGmalZ_R; see figure 5.4 (a)). The *malZ* and *perR* probes hybridise to sequences approximately 56 kb and 97 kb respectively from the site of the palindrome (but between the two I-*Sce*I$_{cs}$ at *tsx* and *proA*). For this particular experiment, the probes were used sequentially; the membrane was stripped after detection with the first probe.

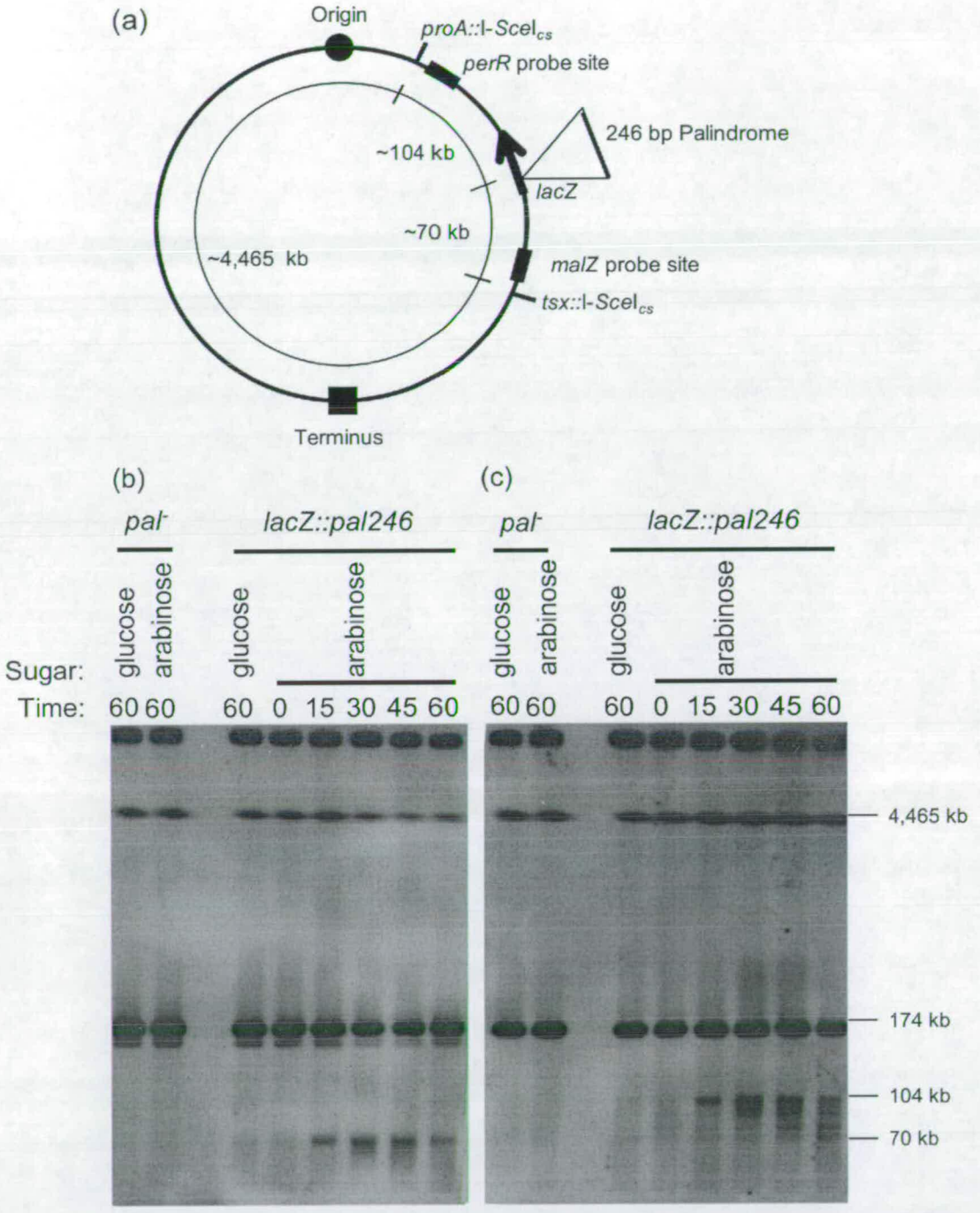


Figure 5.4: Southern blot of I-*Sce*I digested chromosomal DNA of P$_{BAD}$-*sbcDC recB*$^-$ strains with or without a palindrome using *perR* and *malZ* probes. (a) Cartoon showing the relative positions of important sites for I-*Sce*I PFGE analysis (not to scale). Southern blot using (b) *malZ*, or (c) *perR* DIG-labelled probes, and I-*Sce*I digested DNA from various *recB*$^-$ strains. Strains used were DL2797 (P$_{BAD}$-*sbcDC lacZ*$^+$ *recB*$^-$) and DL2798 (P$_{BAD}$-*sbcDC lacZ::pal246 recB*$^-$).

The gel in figure 5.4 (b) and (c) confirms the observation of the previous experiment that a two-ended DNA DSB appears following cleavage at the palindrome by SbcCD. In addition, it shows that both ends of the DNA break are subject to RecB-independent degradation. However, it is not clear if degradation occurs because of cell death or as an attempt by the cell to repair the break in the absence of RecBCD.

When the membrane was probed with the *malZ* probe (figure 5.4 (b)) a DNA fragment slightly smaller than 174 kb was observed. It is not clear what this band is but it is not specific to either $SbcCD^+$ strains or strains containing a palindrome.

5.7 Discussion

Physical methods have previously shown that a palindrome is a target of DNA DSBs in yeast (both Mre11/Rad50 dependent and independent breaks) (Farah et al., 2005; Lobachev et al., 2002). The work in this chapter shows for the first time that SbcCD cleaves a palindromic sequence leaving a two-ended DNA DSB. In light of the data in this chapter, and the fact that the *E. coli* chromosome has a single origin of replication, a model is supported where a hairpin structure is cleaved once the replication fork passes beyond the palindromic sequence (figure 5.5 (a)). These data could also be interpreted as cleavage of a non-replicating chromosome following cruciform extrusion (see figure 5.5 (c)). Previous studies that showed replication-dependent palindrome inviability in phage lambda support the first hypothesis (Leach and Lindsey, 1986). In addition, the palindrome used in these experiments has a 26 bp asymmetric centre and is more likely to form a hairpin structure during replication than a cruciform structure in a non-replicating chromosome (Leach,

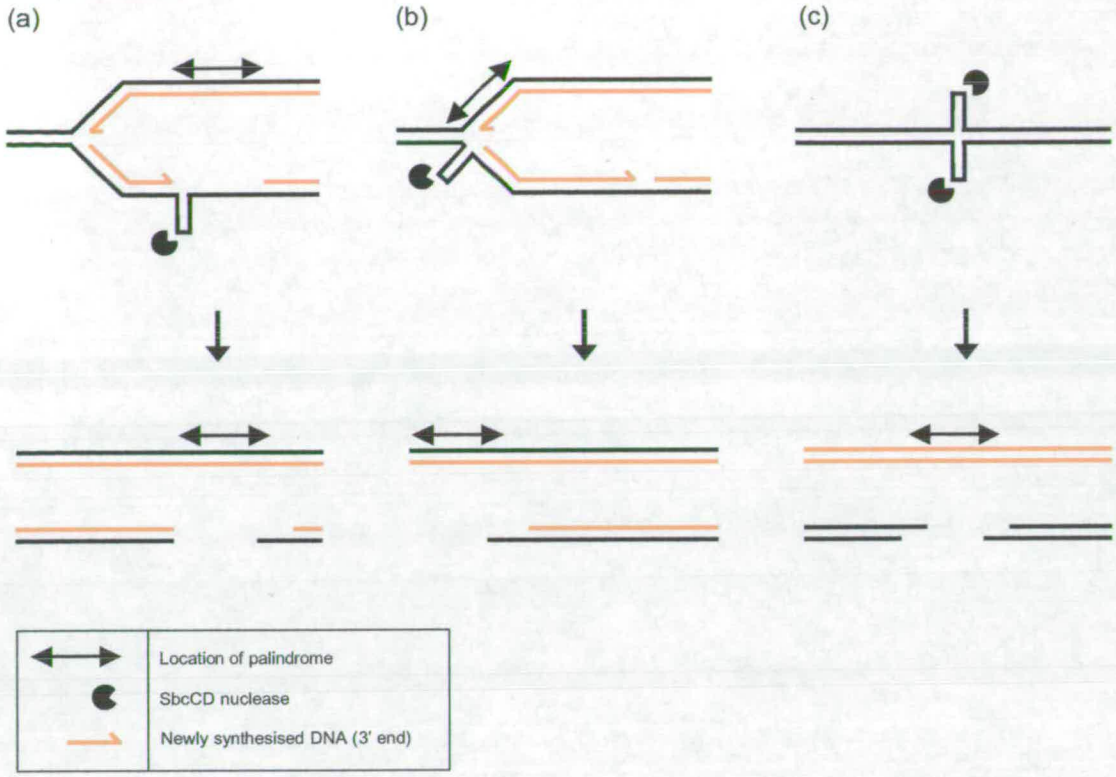


Figure 5.5 Possible mechanisms of SbcCD cleavage of palindromic sequences. (a) Cleavage in the lagging-strand template giving a two-ended DNA DSB, (b) cleavage causing replication fork collapse giving a single DNA end, (c) cleavage of a non-replicating chromosome following cruciform extrusion giving a 2-ended DNA DSB. At least one of (a) and (c) must occur, and the event at (b) might occur in a subset of breaks. In all cases the DNA ends produced are open to nuclease attack in a RecBCD-independent manner.

1994). It may be possible to distinguish between the two scenarios by observing the effects of SbcCD expression in cells that are not actively replicating their DNA.

Although, it is shown that DNA DSBs occur following SbcCD cleavage of a palindromic sequence the results here do not distinguish between DNA DSBs on the leading or lagging strand templates. However, a previous study showed that large palindromes are subject to stabilizing, central deletions in *sbcDC*⁻ strains of *E. coli* following replication slippage on the lagging strand template (Pinder et al., 1998). It would therefore seem likely that the DNA breaks observed in these experiments occur following formation of secondary structures in the lagging strand template during replication. This could be a consequence of discontinuous replication of the lagging strand in which single-stranded DNA regions are longer lived.

Observation of ethidium bromide stained gels revealed intact (not broken) chromosomal DNA suggesting that breaks are not occurring in the majority of chromosomes. In fact, assuming breaks occur only (or more commonly) in the lagging strand template then a maximum ratio of 1:1 broken to intact fragments would be expected. In these experiments, the breaks appear to be occurring at a lower frequency. However, the Southern blotting detection methods employed here are not readily quantifiable. Detection with ^{32}P labelled probes with a phosphoimager may provide an accurate measure of the frequency of chromosomal breaks. In addition, it may be possible to quantify the ratio of the two DNA ends in a population of cells with DSBs; a ratio of 1(*perR*):1(*malZ*) would be expected if all

breaks are two-ended, and a deviation from a 1:1 ratio in favour of the *perR* probe would be expected if a proportion of breaks are one-ended (see figure 5.5 (b)).

The results in this chapter also show that the DNA ends resulting from SbcCD cleavage of a palindromic sequence are degraded over time in a RecBCD-independent manner. This observation may be because SbcCD$^+$ *lacZ::pal246 recB*$^-$ cells are inviable, and DNA degradation is a consequence of cell death. Alternatively, a RecBCD-independent process that cannot facilitate DNA DSB repair may degrade the DNA ends following a break.

Chapter 6: Measuring recombination caused by a palindrome in *E. coli*

6.1 Introduction

The work in chapter five demonstrates that SbcCD-dependent cleavage of a palindromic sequence *in vivo* generates a two-ended DNA double-strand break. Some of the genetic requirements for palindrome viability have been described in chapters three and four, and repair of breaks relies on RecA and the RecBCD recombination pathway. However, it is unclear if one or both DNA ends are substrates for the recombination machinery. In addition, a previous report suggested that a palindrome can stimulate RecFOR-mediated recombination in the absence of SbcCD (Cromie et al., 2000) but the work described in chapter four did not provide support for this idea. This chapter describes the construction of a set of recombination reporter cassettes to measure rates of recombination at different loci near a 246 bp interrupted palindrome. These cassettes were used to measure rates of recombination in both *sbcDC*$^{+}$ and Δ*sbcDC* strains at loci upstream, downstream and immediately adjacent to the N-terminus of *lacZ* or *lacZ::pal246*. The contribution of RecBCD was assessed by measuring recombination rates in cassettes with or without a *chi* sequence. In *E. coli* the octameric *chi* sequence stimulates the RecA-loading activity of RecBCD and therefore indirectly stimulates homologous strand invasion (see section 1.1.1.1).

6.2 Construction of a 'zeocin sensitive' recombination reporter cassette

6.2.1 The zeocin recombination reporter cassette

As part of a laboratory wide collaboration a recombination reporter cassette was designed based on the zeocin resistance gene (collaboration with John Blackwood). The cassette consists of two incomplete direct-repeats of the zeocin resistance gene separated by three stop codons (figure 6.1 (a)). The first copy is under the control of an *E. coli* constitutive promoter and is missing the 3' part of the gene (the last 100 bases). The second copy is missing the 5' part of the gene (first six bases including the ATG start codon). Importantly, this cassette does not confer zeocin resistance.

6.2.2 A set of strains for measuring recombination

In order to measure the recombination rates of the DNA ends following SbcCD cleavage at a palindromic sequence, the zeocin recombination cassette (described above) was integrated into the *E. coli* chromosome 5 kb either side of the *lacZ* locus (at the *mhpBC* and *cynX* loci), in *sbcDC*$^+$ *lacZ*$^+$ and *sbcDC*$^+$ *lacZ::pal246* strains, (figure 6.1 (b)). *mhpBC* and *cynX* are non-essential genes that have no reported or predicted roles in homologous recombination.

In order to assess if a palindrome can stimulate SbcCD-independent recombination the zeocin recombination cassettes were integrated into the *mhpBC* and *cynX* loci of Δ*sbcDC lacZ*$^+$ and Δ*sbcDC lacZ::pal246*. In addition, the *zeo*^{S} cassette was

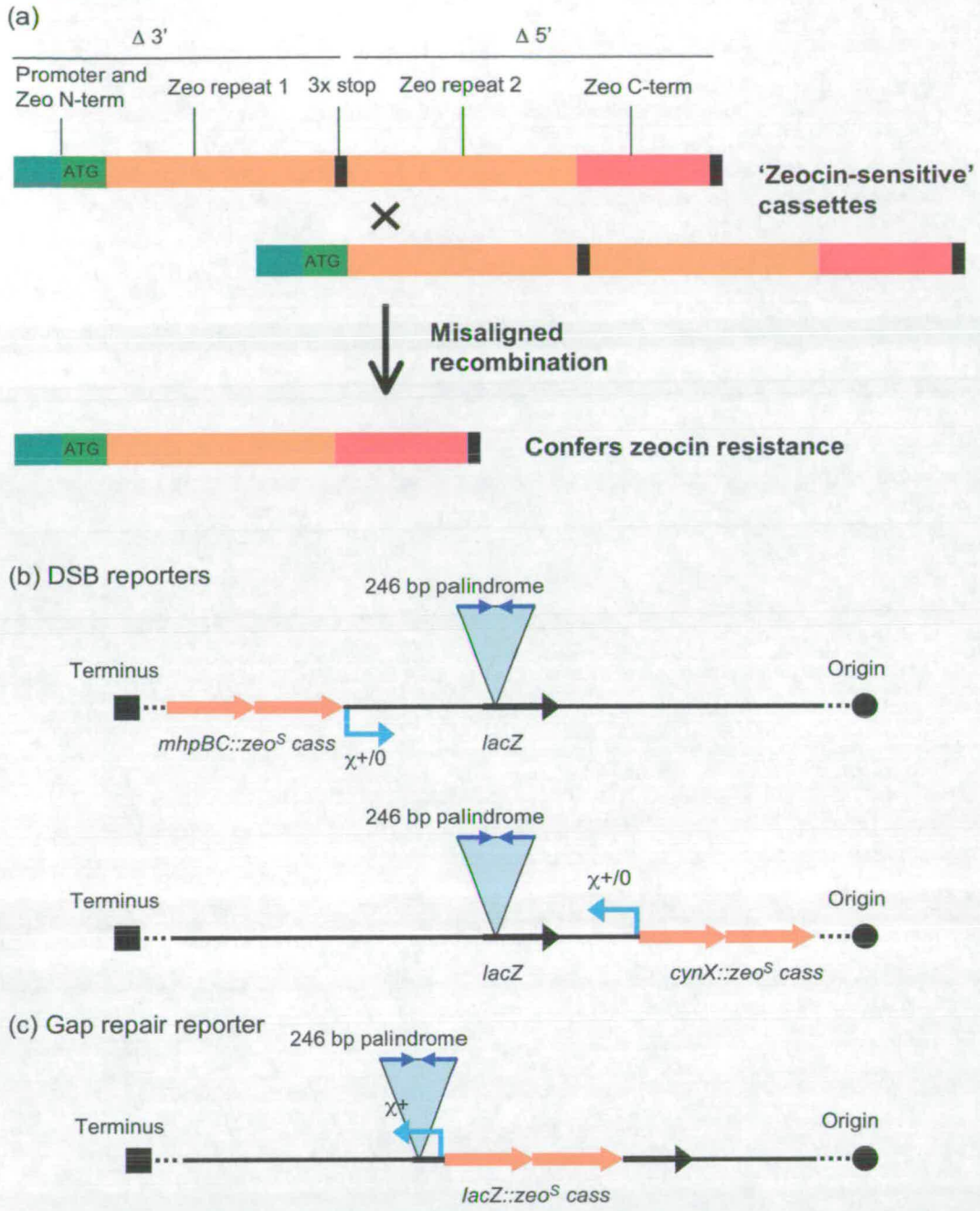


Figure 6.1: The zeocin recombination cassette and sites of integration. Cartoons representing (a) the 'zeocin-sensitive' cassette and the recombination event that confers zeocin resistance, (b) and (c) the relative positions of the cassettes with respect to the origin, the terminus, the palindrome and the positions of the chi sites.

integrated 400bp into the *lacZ* gene of Δ*sbcDC lacZ*$^{+}$ and Δ*sbcDC lacZ::pal246* strains (figure 6.1 (c)).

6.2.2.1 Construction of pDL1998 and pDL1999

To facilitate integration of *zeo*^{S} cassettes for measurement of *chi*-dependent and *chi*-independent recombination at the *cynX* locus, two pTOF24-derivative plasmids, pDL1998 and pDL1999, were constructed by John Blackwood. Both of these plasmids contain the *zeo*^{S} cassette located between two 400 bp arms of *cynX* homology. The plasmid pDL1998 has a *chi* sequence at the N-terminal end of the first zeocin-repeat, pDL1999 does not carry this *chi* sequence.

6.2.2.2 Construction of pDL2941 and pDL2942

To facilitate integration of *zeo*^{S} cassettes for measurement of *chi*-dependent and *chi*-independent recombination at the *mhpBC* locus, two pTOF24-derivative plasmids pDL2941 and pDL2942 were constructed. Both of these plasmids contain the *zeo*^{S} cassette located between two 400 bp arms of *mhpBC* homology. The plasmid pDL2941 has a *chi* sequence near the 3' end of the second zeocin-repeat, pDL2942 does not carry this *chi* sequence.

Initially, the plasmid pDL2069 was constructed by cloning into pTOF24, between *Pst*I and *Not*I, an 800 bp crossover PCR product generated using the primer pairs Bsd-CR2 (*Pst*I site at 5' end), MhpC_Rchi+ and MhpB_Fchi+, Bsd-CF1-2 (*Not*I site at 5' end) with *E. coli* MG1655 genomic DNA as template. This novel fragment is homologous to regions of the *mhpBC* operon and has an *Xba*I *Bgl*II cloning site and a *chi* sequence located at its centre. Correct clones were confirmed by sequencing with the primer pair pKOF and pKOR2.

A *chi*^{0} version of pDL2069 was created by site-directed mutagenesis using the primer pair sdm_kochi_f and sdm_kochi_r. Two point mutations were introduced to the *chi* sequence, generating a *Spe*I restriction site. Correct clones were confirmed by restriction, PCR and sequencing with the primer pair pKOF and pKOR2. This new plasmid was named pDL2940.

Finally, pDL2941 and pDL2942 were constructed from pDL2069 and pDL2940, respectively. The *zeo*^{S} cassette was PCR amplified using the primer pair zeocass_F (*Xba*I site at 5' end) and zeocass_R (*Bgl*II site at 5' end) with pDL1999 as template. This PCR product was cloned into each of the plasmids pDL2069 and pDL2940, using *Xba*I and *Bgl*II restriction sites. The presence of the *zeo*^{S} cassette was confirmed by PCR with the primer pair zeocass_F and zeocass_R.

6.2.2.3 Construction of pDL2851

To facilitate integration of *zeo*^{S} cassettes for measurement of recombination 400 bp downstream of the *lacZ* N-terminus, a pTOF24-derivative plasmid, pDL2851 was constructed. This plasmid contains the *zeo*^{S} cassette located between two 400 bp arms of *lacZ* homology, with a *chi* sequence at the N-terminus of the first zeocin repeat.

Initially, the plasmid pDL2444 was constructed by cloning into pTOF24, between *Pst*I and *Sal*I, an 800 bp crossover PCR product generated using the primer pairs LacZ-F1 (*Pst*I site at 5' end), LacZ-R1 and LacZ-F2, LacZ-R2 (*Sal*I site at 5' end) with MG1655 genomic DNA as template. This novel fragment is homologous to

regions of the *lacZ* gene and has a *Not*I *Bam*HI cloning site, and a *chi* sequence located at its centre. A correct clone was confirmed by sequencing using the primer pair pKOF and pKOR2.

Finally, pDL2851 was constructed by cloning into pDL2444 between *Not*I and *Bam*HI, the zeo^S cassette amplified by PCR using the primer pair zeogap_F (*Bam*HI site at 5' end) and zeogap_R (*Not*I site at 5' end) with pDL1999 as template. The presence of the zeo^S cassette was confirmed by PCR with zeogap_F and zeogap_R.

6.2.2.4 PMGR to introduce the zeo^S cassettes

For these experiments the DL1777 strain (and derivatives) were used. DL1777 is an MG1655 derivative with an altered *lacZ* region (containing the $lacI^q$ mutation and with a *chi* sequence removed from the *lacZ* gene; described in section 4.2).

PMGR using pDL1998, pDL1999, pDL2941 or pDL2942, was carried out to introduce the zeo^S chi^+/chi^0 cassettes into the *mhpBC* and *cynX* loci of $sbcDC^+$ and Δ*sbcDC* strains with or without a palindrome ($lacZ^+$ or *lacZ::pal246*). PMGR using pDL2851 was used to introduce the zeo^S chi^+ cassette into the *lacZ* locus of Δ*sbcDC* $lacZ^+$ and Δ*sbcDC lacZ::pal246* strains.

6.3 Evidence for SbcCD-independent recombination at a palindromic sequence

In a previous study by Cromie and collaborators palindrome inviability of an $sbcC^-$ $ruvA^-$ strain was alleviated following the introduction of a $recF^-$, $recO^-$ or $recR^-$ mutation (Cromie et al., 2000). A model was proposed in which a palindrome

stimulates RecFOR recombination at a ssDNA gap formed when a hairpin structure blocks lagging-strand DNA synthesis. Similar experiments shown in chapter four found no requirement of RuvABC in an SbcCD$^-$ strain containing a palindrome. To explore this idea further the *lacZ zeo*^{S} recombination cassette was designed to be integrated as near to the palindromic sequence as possible, where a ssDNA gap would be predicted if the model described above were true.

Fluctuation analyses (see section 2.2.9) were used to measure the recombination rates at each of the *cynX* / *mhpBC* / *lacZ zeo*^{S} cassettes in Δ*sbcDC* strains, with or without a palindrome (figure 6.2 and 6.3).

6.3.1 Recombination rates at *cynX* and *mhpBC*

The recombination rates (expressed as recombinants per cell per generation) of the *zeo*^{S} cassettes at *cynX* and *mhpBC* in Δ*sbcDC* strains without a palindrome were all approximately the same, between 1.8 x 10^{-6} and 2.1 x 10^{-6} (see figure 6.2). The presence of the *chi* sequence had no significant effect on recombination rates in these strains. The recombination rates in the equivalent strains containing the palindrome were slightly increased. At the *cynX* locus, the recombination rate was up 1.5-fold (to 2.6 x 10^{-6}) with the *chi*^{0} cassette and up 2.5-fold (to 4.5 x 10^{-6}) with the *chi*$^+$ cassette. At the *mhpBC* locus, the recombination rate was not significantly different with the *chi*^{0} cassette and up 1.6-fold (3.5 x 10^{-6}) with the *chi*$^+$ cassette (figure 6.2).

These results suggest that there is a small but significant level of SbcCD-independent recombination stimulated by the palindrome and at least part is stimulated by RecBCD. It would appear that there is more stimulation at the *cynX* locus, this could

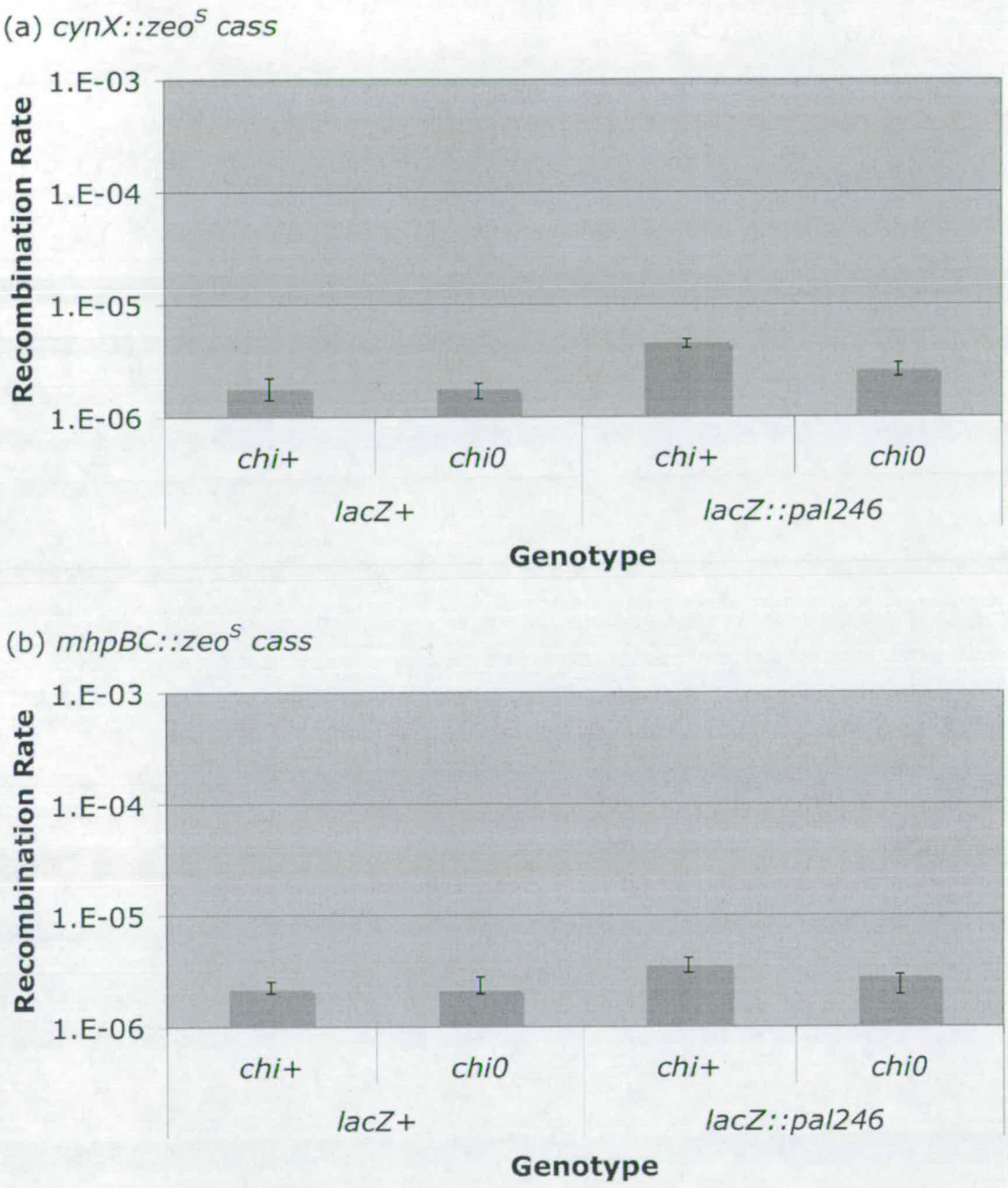


Figure 6.2: Recombination rates of the *mhpBC* and *cynX* zeocin cassette in Δ*sbcDC* strains with or without a palindrome. Rates measured at (a) the *cynX* locus and (b) the *mhpBC* locus. Calculated as a mean derived from the median frequency of 36 measurements with 95% confidence intervals indicated. Strains used were DL2888 (*lacZ*$^{+}$ *cynX::zeo chi*$^{+}$), DL2889 (*lacZ*$^{+}$ *cynX::zeo chi*^{0}), DL2901 (*lacZ::pal246 cynX::zeo chi*$^{+}$), DL2902 (*lacZ::pal246 cynX::zeo chi*^{0}), DL2965 (*lacZ*$^{+}$ *mhpBC::zeo chi*$^{+}$), DL2966 (*lacZ*$^{+}$ *mhpBC::zeo chi*^{0}), DL2969 (*lacZ::pal246 mhpBC::zeo chi*$^{+}$) and DL2970 (*lacZ::pal246 mhpBC::zeo chi*^{0}).

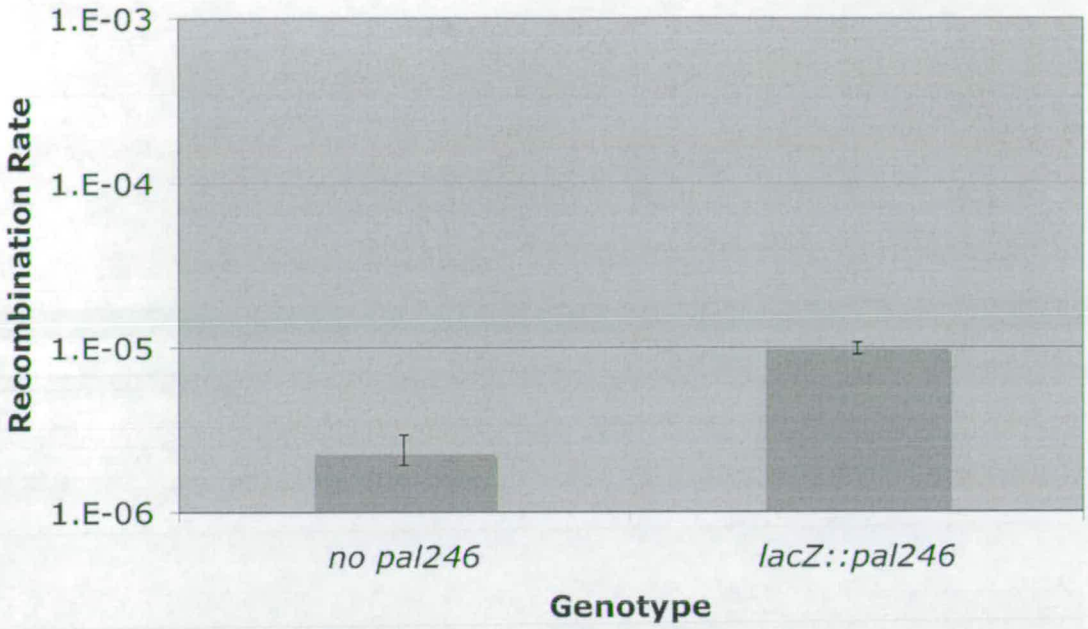


Figure 6.3: Recombination rates of the *lacZ* zeocin cassette in Δ*sbcDC* strains with or without a palindrome. Calculated as a mean derived from the median frequency of 36 measurements with 95% confidence intervals indicated. Strains used were DL2852 (*lacZ::zeo chi*$^{+}$) and DL2899 (*lacZ::pal246 zeo chi*$^{+}$).

be because the origin proximal side (with respect to the palindrome) is more likely to recombine, or alternatively could reflect a difference between the orientations of the two cassettes with respect to the palindrome.

6.3.2 Recombination rates at *lacZ*

The recombination rate of the reporter cassette at the *lacZ* locus (400bp origin proximal of the palindrome) in a Δ*sbcDC* strain without a palindrome was 2.2×10^{-6} (about the same as the recombination rates seen with the *mhpBC* or *cynX zeo*^{S} cassettes in equivalent strains) (figure 6.3). By comparison the mean recombination rate in the equivalent strain containing the palindrome was increased 4.3-fold (to 9.7×10^{-6}). This is a higher rate and a bigger difference than seen between the equivalent *chi*$^{+}$ cassettes at the *cynX* locus. It may be because the cassette lies closer to the palindrome in a region that could form a ssDNA gap. However, it remains to be determined whether the recombination observed at this cassette is RecBCD or RecFOR dependent.

6.4 RecBCD mediated recombination of the zeo^{S} cassette is increased in an sbcDC^{+} strain containing a palindrome

In chapter five, physical evidence was presented demonstrating that SbcCD cleavage of a palindromic sequence gives a two-ended DNA DSB. In order to find out if both the DNA ends generated are recombinogenic, fluctuation analyses were used to measure the recombination rates at each of the *cynX* or *mhpBC zeo*^{S} cassettes in *sbcDC*$^{+}$ strains with or without a palindrome (figure 6.4).

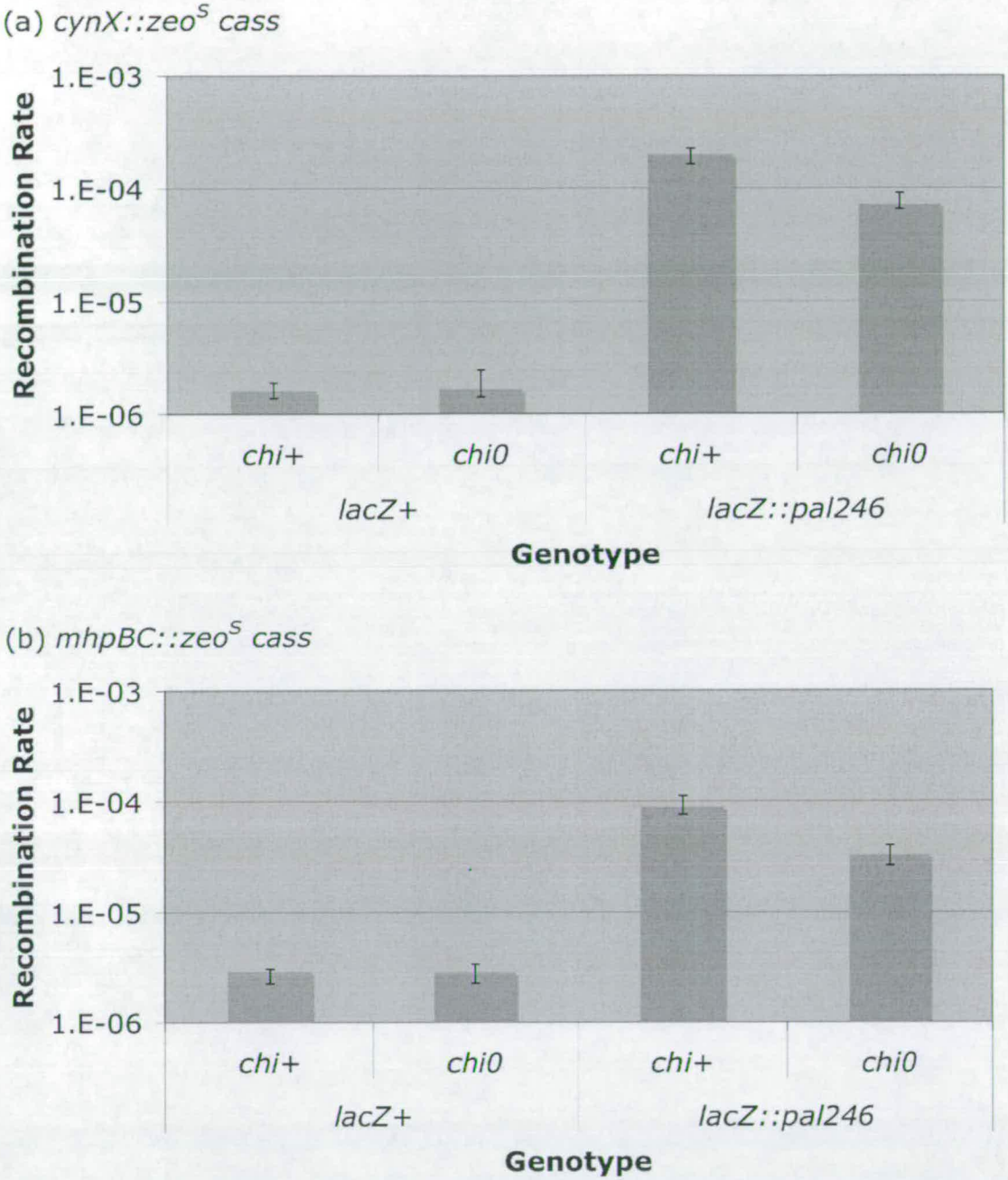


Figure 6.4: Recombination rates of the *mhpBC* and *cynX* zeocin cassettes in *sbcDC*$^+$ strains with or without a palindrome. Rates measured at (a) the *cynX* locus and (b) the *mhpBC* locus. Calculated as a mean derived from the median frequency of 36 measurements with 95% confidence intervals indicated. Strains used were DL2886 (*lacZ*$^+$ *cynX::zeo chi*$^+$), DL2887 (*lacZ*$^+$ *cynX::zeo chi*^{0}), DL2883 (*lacZ::pal246 cynX::zeo chi*$^+$), DL2884 (*lacZ::pal246 cynX::zeo chi*^{0}), DL2963 (*lacZ*$^+$ *mhpBC::zeo chi*$^+$), DL2964 (*lacZ*$^+$ *mhpBC::zeo chi*^{0}), DL2967 (*lacZ::pal246 mhpBC::zeo chi*$^+$) and DL2968 (*lacZ::pal246 mhpBC::zeo chi*^{0}).

The recombination rates of the *zeo*^{S} cassettes at *mhpBC* and *cynX* in *sbcDC*$^{+}$ strains without a palindrome were all approximately the same, between 1.6 x 10^{-6} and 2.9 x 10^{-6}. The presence of the *chi* sequence had no significant effect on recombination rates in these strains. The recombination rates in the equivalent strains containing the palindrome were greatly elevated. At the *cynX* locus, the recombination rate was increased 44-fold (to 7.2 x 10^{-5}) with the *chi*^{0} cassette and 124-fold (to 2.0 x 10^{-4}) with the *chi*$^{+}$ cassette. Similarly, at the *mhpBC* locus, the recombination rate was increased 10-fold (to 3.3 x 10^{-5}) with the *chi*^{0} cassette and 30-fold (to 8.9 x 10^{-5}) with the *chi*$^{+}$ cassette.

The difference between the recombination rates in the *sbcDC*$^{+}$ *lacZ::pal246* and Δ*sbcDC lacZ::pal246* strains confirms that the majority of the palindrome stimulation is SbcCD-dependent. In the strains containing a palindrome, at the *cynX* locus, the recombination rate of the *chi*$^{+}$ cassette was increased 44-fold with the presence of SbcCD, the *chi*^{0} cassette was increased 28-fold. Similarly, at the *mhpBC* locus, the recombination rate of the *chi*$^{+}$ cassette was increased 25-fold with the presence of SbcCD, the *chi*^{0} cassette was increased 11-fold (figure 6.2 and 6.4).

These data (and data presented in section 6.3) indicate that an SbcCD-induced DNA DSB at a palindromic sequence stimulates recombination on both sides of the break. The recombination rate increased by about three-fold with the addition of a *chi* sequence between the break site and the recombination cassette suggesting that RecBCD processing is responsible for at least part of the recombination observed at

both DNA ends. The difference observed between recombination rates at the different loci might reflect a higher propensity for the origin proximal DNA end to recombine, or alternatively, could reflect a difference between the cassette orientations with respect to the palindrome (one cassette faces inwards the other outwards; see figure 6.1 (b)).

6.5 Discussion

6.5.1 SbcCD cleaves a palindromic sequence to give two recombinogenic DNA ends

Physical observations in chapter five revealed that SbcCD-induced breaks at a palindrome are (or can be) two-ended events. The work in chapter six demonstrates that SbcCD-induced DNA double-strand breaks stimulate recombination of both origin-proximal and origin-distal DNA ends. The stimulation is increased in the presence of a *chi* sequence suggesting that at least a proportion of recombination is RecBCD-dependent. The experiments described in this chapter do not assess the role of the RecFOR proteins in recombination following SbcCD cleavage of a palindromic sequence. Although genetic evidence presented in chapters three and four show that these proteins are not required for the repair, a minor role in recombination cannot be ruled out.

The strains containing the palindrome gave relatively higher rates of recombination at the *cynX zeo*^{S} cassettes (origin proximal with respect to the palindrome). This could reflect the difference in orientation (with respect to the palindrome) of the cassettes at the two different loci with the cassette pointing away from the

palindrome being a better substrate for recombination. Alternatively, an origin-proximal end could have a higher recombination rate than an origin-distal end following an SbcCD-induced break. This could be because the DNA ends are not equivalent, and a competitive end-processing pathway that does not stimulate recombination acts efficiently at the origin-distal end. Alternatively, the DNA ends may be equivalent, and equally good substrates for RecBCD, but a proportion of breaks in a population of cells are one-ended with the presence of the origin-proximal end only (figure 6.5).

6.5.2 A palindrome stimulates a low-level of SbcCD-independent recombination

6.5.2.1 RecBCD-dependent recombination

The rates of recombination measured at the *cynX* and *mhpBC* loci in Δ*sbcDC* strains suggest there is a low, but significant stimulation of recombination induced by a palindrome at *lacZ*. Stimulation of recombination at the *cynX zeo*^{S} cassette (and to a lesser extent the *mhpBC* cassette) is increased by the addition of a *chi* sequence and therefore a proportion of the recombination events are probably RecBCD-dependent. A number of models could explain this: one compatible with stimulation of both *mhpBC* and *cynX zeo*^{S} cassettes and two compatible with stimulation of recombination at the *cynX zeo*^{S} cassettes only (see below). The higher stimulation of recombination at the *cynX* locus (compared to the *mhpBC* locus) suggested that a combination of the different pathways probably contribute to its recombination. Alternatively, this could reflect the difference in orientation (with respect to the palindrome) of the cassettes at the two different loci.

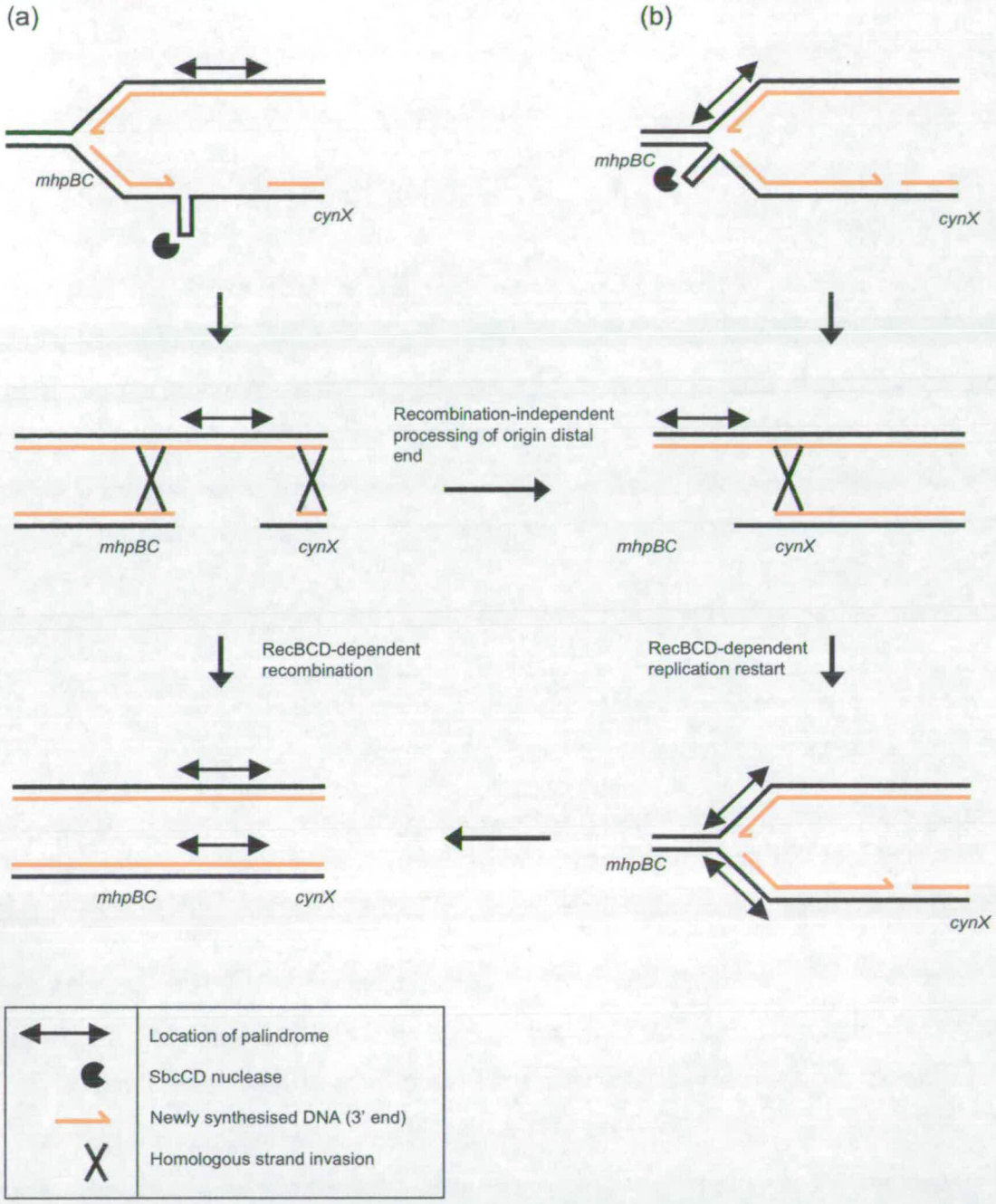


Figure 6.5: Schematic drawing showing possible RecBCD repair mechanisms following SbcCD cleavage of a palindromic sequence. Following cleavage (a) in the lagging strand template or (b) at the replication fork. It is also possible that a break occurs in a non-replicating chromosome (see figure 5.5 (c)). This figure indicates the relative positions of the *cynX* and *mhpBC* loci; the chromosomal positions of the zeocin recombination reporter cassettes.

First of all an SbcCD-independent palindrome cleavage pathway generating a two-ended DNA DSB could stimulate RecBCD recombination of both DNA ends (figure 6.5). This is consistent with *chi* stimulated recombination of both *mhpBC* and *cynX* zeo^S cassettes in Δ*sbcDC* strains. The low-level of recombination of the cassettes in Δ*sbcDC* strains compared to the equivalent $sbcDC^+$ strains is probably because SbcCD-independent cleavage of a palindrome is very infrequent. Physical analyses in chapter five were unable to detect SbcCD-independent DNA DSBs at a palindromic sequence.

A second possible RecBCD-dependent, SbcCD-independent recombination pathway stimulated by a palindrome would be if a hairpin secondary structure on the lagging strand leads to incomplete replication and generation of a nick in the DNA phosphate backbone (figure 6.6 (a)). A subsequent round of replication would lead to replication fork collapse. RecBCD-mediated recombination at the broken end could reload the replication machinery and could lead to rearrangement of the *cynX* zeo^S cassettes (but not the *mhpBC* zeo^S cassettes; figure 6.7).

A third possible RecBCD-dependent, SbcCD-independent recombination pathway stimulated by a palindrome would be if a palindromic hairpin structure impedes the progress of a replication fork (Lindsey and Leach, 1989) and stimulates replication fork reversal. RecBCD processing and recombination of the reversed DNA end can lead to initiation of a new replication fork allowing another opportunity for a replication fork to bypass the DNA-block (Seigneur et al., 1998). In this hypothesis,

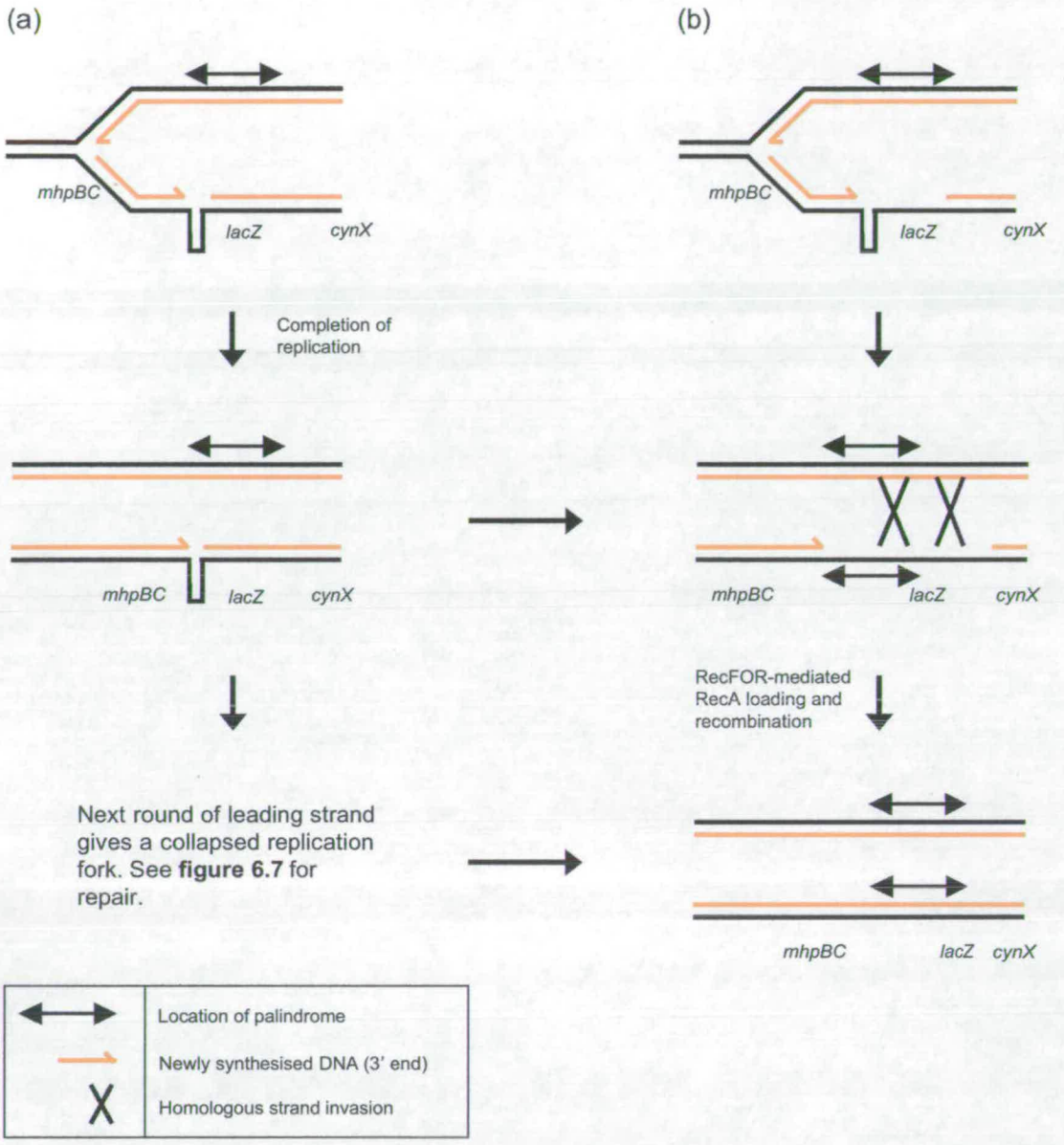


Figure 6.6: Schematic drawing of a possible SbcCD-independent recombination pathway at a palindromic sequence. Possibly stimulated by DNA hairpin in the lagging strand template (a) to form a DNA nick (see also figure 6.7) or (b) to form a ssDNA gap. This figure also shows the relative positions of the *cynX* and *mhpBC* loci, the chromosomal positions of the zeocin recombination reporter cassettes.

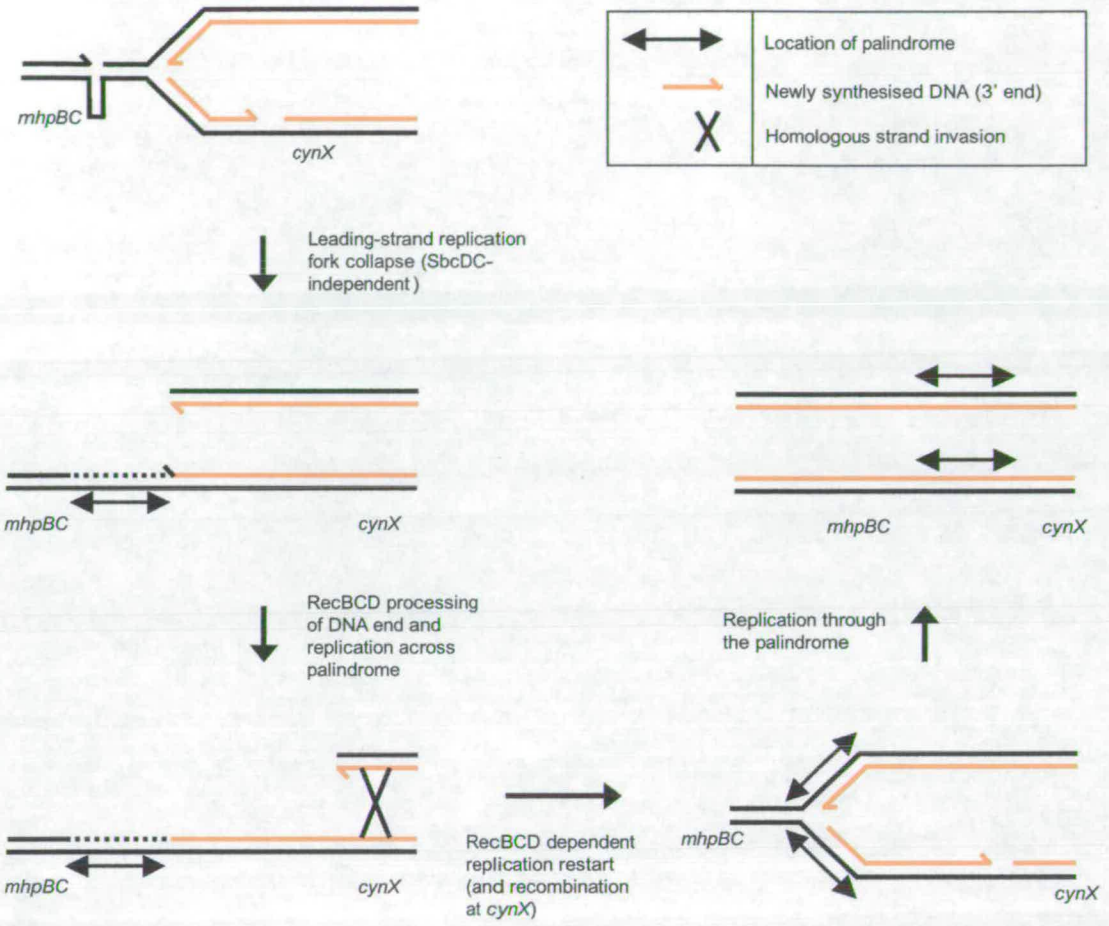


Figure 6.7: Schematic drawing of a possible SbcCD-independent RecBCD recombination pathway at a palindromic sequence (follows figure 6.6). This shows RecBCD-dependent recombination following polymerisation across the palindromic sequence. This figure also shows the relative positions of the *cynX* and *mhpBC* loci; the chromosomal positions of the zeocin recombination reporter cassettes.

RecBCD-mediated recombination could stimulate rearrangement of *cynX zeo*^{S} cassettes (but not *mhpBC zeo*^{S} cassettes; see figure 6.8).

6.5.2.2 RecFOR-dependent recombination

The rates of recombination measured at the *lacZ zeo*^{S} *chi*$^+$ cassette in Δ*sbcDC* strains suggest there is a low, but significant stimulation of SbcCD-independent recombination induced by a palindrome at *lacZ*. The experiments here only describe recombination rates of recombination in *chi*$^+$ cassettes and therefore the dependence on RecBCD of this recombination remains unknown. In fact the stimulation of recombination of the *lacZ zeo*^{S} cassette could be by any of the RecBCD-dependent pathways described above. However, SbcCD-independent stimulation of recombination by the palindrome at the *lacZ zeo*^{S} cassette is higher than in an equivalent strain with the *cynX zeo*^{S} *chi*$^+$ cassette. This could be because it is closer to the site of the palindrome or an alternative recombination pathway may be acting at the *lacZ zeo*^{S} cassette.

Recombination at the *lacZ zeo*^{S} cassette might be stimulated by a palindrome hairpin block of lagging strand DNA synthesis that causes a ssDNA gap, which is a target for RecFOR recombination (figure 6.6). This additional pathway may not stimulate recombination of the equivalent *cynX zeo*^{S} cassette because it is too far away from the palindrome (5 kb) to be affected by RecFOR recombination.

6.5.2.3 Limitations of these assays

The assays described in this chapter using Δ*sbcDC* strains have given an idea of possible recombination pathways that are occurring at a palindrome in the *E. coli* chromosome and are largely based on small but significant differences of rates of

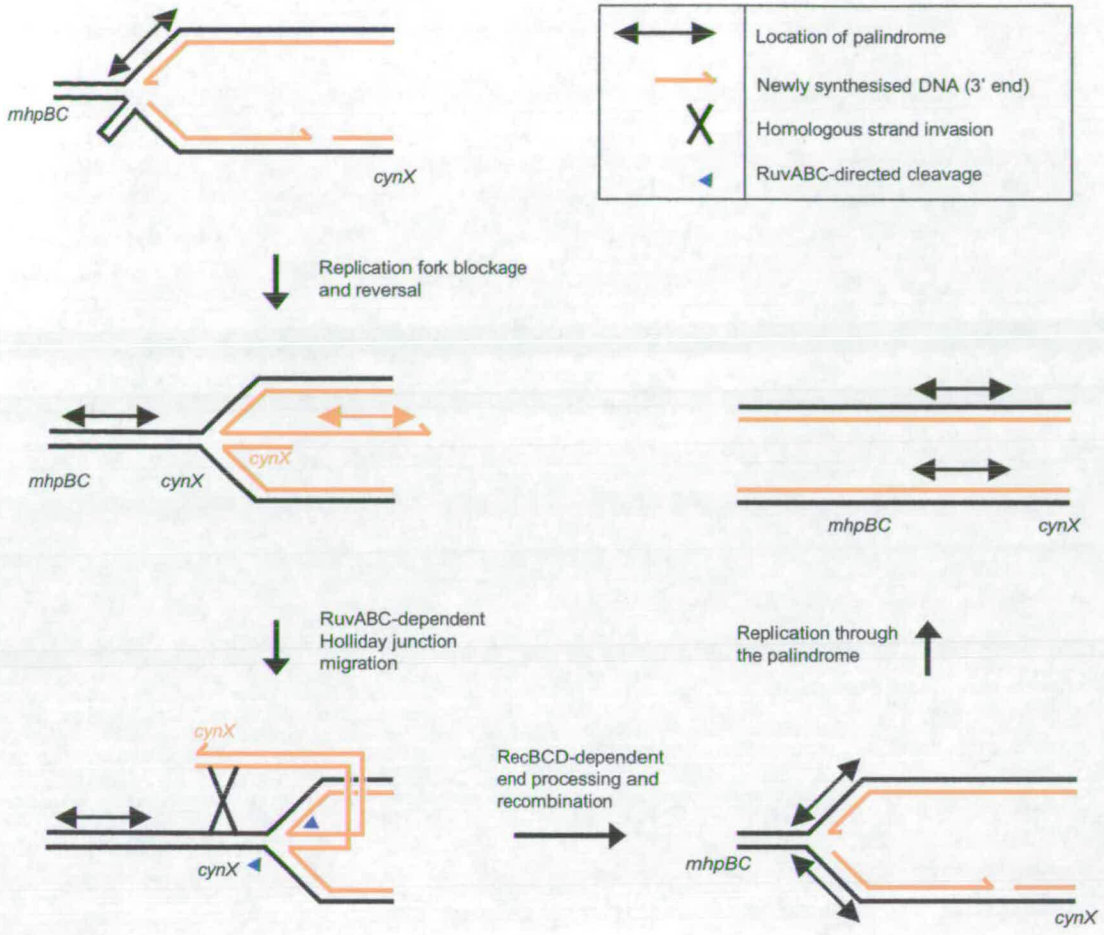


Figure 6.8: Schematic drawing of a possible replication fork reversal recombination pathway at a palindromic sequence. A palindrome stimulated RecBCD RuvABC-dependent recombination event allows a second chance of replication past the palindrome. This figure also shows the relative positions of the *cynX* and *mhpBC* loci; the chromosomal positions of the zeocin recombination reporter cassettes.

recombination in *rec*$^{+}$ strains. To distinguish between the different pathways (for example the RecBCD-dependent pathways), recombination rates in specific mutants will have to be assessed. This should be possible since, from the genetic data presented in chapter three, palindrome-viability is possible in most SbcCD^{-}, *rec*$^{-}$ mutants.

Chapter 7: Conclusions and future work

7.1 Conclusions

The work in this thesis aims to give a better understanding of recombination at a palindromic sequence in *E. coli*. For this purpose the natural chromosomal *sbcDC* promoter was replaced by an arabinose-inducible promoter (P_{BAD}) in an *E. coli* strain displaying homogeneous expression from P_{BAD}. Additionally, a 246 bp interrupted palindrome was integrated into the chromosomal *lacZ* locus of this strain. Pulsed-field gel electrophoresis of the chromosomal DNA of derivatives of this strain demonstrates that the SbcCD complex can cleave the palindromic sequence and generate a DNA DSB. Previous studies in yeast have shown by physical analysis that a palindrome is a target of the MRN complex (SbcCD homologue) generating DNA DSBs (Farah et al., 2005). However, the work here, in the *E. coli* chromosome, provides evidence that the DSB can consist of two DNA ends. This suggests that breaks can occur following cleavage of either (1) a cruciform structure in a non-replicating chromosome, or (2) a hairpin structure in the lagging strand template, behind the replication fork.

This thesis defines some of the genetic requirements of $SbcCD^+$ and $SbcCD^-$ cells carrying a 246 bp interrupted palindrome using viability assays in the P_{BAD}-*sbcDC* strains. Previous published work, reporting the genetic requirements of *E. coli* for lysogenisation, were done with a bacteriophage lambda containing the 246 bp interrupted palindrome (Cromie et al., 2000; Leach et al., 1997). The work here

reflects some, but not all of these results. In this study, an SbcCD$^+$ strain containing a palindrome requires the RecA protein for SOS induction and recombination. The RecB protein that makes up part of the RecBCD complex is also required, presumably for processing broken DNA ends following SbcCD cleavage of a palindromic sequence. In contrast to a previous study (Cromie et al., 2000), in this experimental system, the RecFOR proteins were not required for repair of SbcCD induced DSBs. Furthermore, RuvAB was not required by an SbcCD$^-$ strain containing a palindrome, suggesting that the RecFOR recombination pathway is not a major pathway in the absence of SbcCD or there is an alternative pathway for resolution. Moreover, it was found that the PriA protein is required in an SbcCD$^+$ strain containing a palindrome, probably for establishing replication following a DSB. The RuvAB proteins, from the RuvABC complex, and the RecG protein are required in an SbcCD$^+$ strain containing a palindrome, presumably for the resolution of Holliday junctions formed following recombination. Finally, in contrast to a previous study (Cromie et al., 2000), in these experiments, the XerC protein and *dif* sequences only have a minor role in survival of an SbcCD$^+$ strain containing a palindrome suggesting that the bias of genetic crossing over following a DNA DSB may not be as strong as previously suggested (Cromie and Leach, 2000).

Further genetic analyses were carried out in *sbcDC*$^+$ and Δ*sbcDC* strains containing a palindrome. For this purpose, recombination reporter cassettes were integrated at three different loci around the palindrome. Rates of recombination measured at different loci confirm that both sides of a palindrome can instigate recombination following SbcCD cleavage. Recombination is increased when a *chi* sequence is

added to the reporter cassette, indicating the presence of RecBCD-dependent recombination. The recombination rates measured in Δ*sbcDC* strains suggest that a palindrome can stimulate SbcCD-independent recombination. The exact mechanisms for this stimulation are yet to be elucidated.

Figure 7.1 is a flow diagram that summarises the possible pathways by which a palindrome can be processed in *E. coli*. Two SbcCD-dependent palindrome cleavage events are shown with the more probable event of cleavage behind a replication fork shown with a bold line. In addition, SbcCD-independent events are shown that may lead to recombination or may be bypassed by some other mechanism, for example by simply removing the secondary structure. Events such as replication slippage and palindrome deletion are not considered in this diagram.

In *E. coli,* recombination following an SbcCD-induced DNA break leads to generation of the original palindrome (follow the heavy black line in figure 7.1). In contrast, in *S. pombe* a meiotic, MRN-induced DNA break frequently leads to palindrome deletion by gene conversion (Farah et al., 2005). The recombination mechanisms involved in both cases may be related and it is possible that MRN-independent pathways of recombination are similar between *E. coli* and higher organisms too.

Finally, the P_{BAD}-*sbcDC recB*$^-$ control strain without a palindrome is less viable in the presence of SbcCD than in its absence. This is reminiscent of the difference in

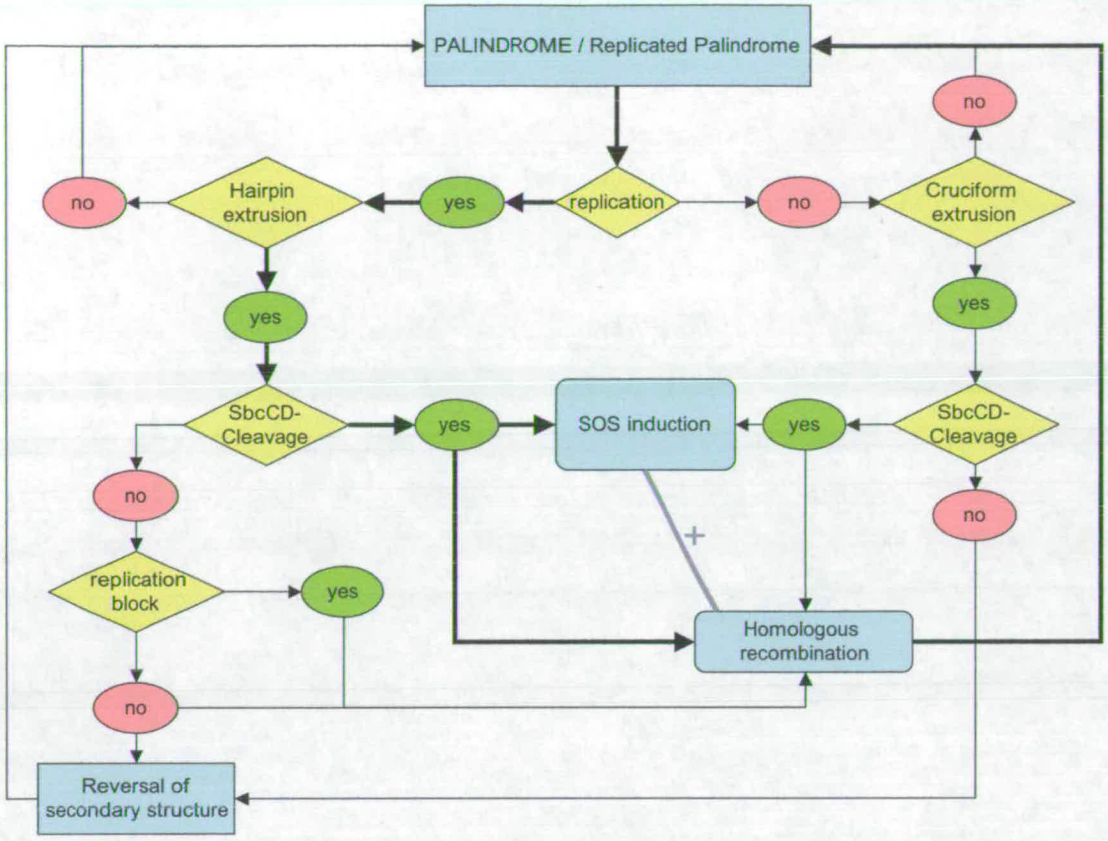


Figure 7.1: SbcCD-dependent and independent pathways of palindrome processing. All possible pathways suggested in this thesis are included. The SbcCD-dependent pathway indicated by the heavy black arrows is arguably the major pathway for recombination at a palindrome. The grey line represents the stimulatory feedback between SOS induction and homologous recombination proteins.

fitness observed between *recB⁻ sbcB⁻ sbcC⁻* and *recB⁻ sbcB⁻ sbcC⁺* strains (Lloyd and Buckman, 1985). There has yet been no explanation for this observation.

7.2 Future work

7.2.1 Genetic requirements for palindrome viability

The work in this thesis has demonstrated that homologous recombination and SOS induction are required for palindrome viability in SbcCD$^+$ strains. The SOS response is induced following SbcCD-cleavage of a palindromic sequence and there is a requirement for some of the SOS genes such as *recA* and *ruvAB*. However, it is not clear if these are the only SOS proteins that are required for survival, many others could be tested using a similar approach. Alternatively, the only important SOS proteins required for repair of SbcCD-induced damage could be RecA (and / or RuvAB). In this case over-expression of the RecA protein (or RuvAB or both) might rescue the viability defect of the uncleavable LexA3 mutant.

7.2.2 Different recombination pathways at a palindrome

The experiments in this thesis mainly describe the SbcCD-dependent recombination pathways using the P_{BAD}-*sbcDC lacZ::pal246* strain. However, the experiments using the zeocin recombination cassettes show evidence of a low level of recombination in Δ*sbcDC* strains as previously published (Cromie et al., 2000). The zeocin recombination cassettes could prove a valuable tool for defining subtle palindrome stimulated recombination pathways. For example, measuring the recombination rates in various recombination mutant strains might provide evidence

for or against some of the hypothesised SbcCD-independent recombination pathways suggested in chapter six. The first set of proteins to test would be the RecFOR proteins, which have been implicated in recombination at a palindromic sequence in a previous study (Cromie et al., 2000).

In addition, the same study suggests that the RecQ helicase plays a role in palindrome viability (Cromie et al., 2000). There was a requirement for functional RecQ for plating of a phage containing a 571 bp palindrome on an *sbcC*$^-$ strain. Furthermore, RecQ was needed for lysogenisation of an *sbcC*$^-$ *recA*$^-$ strain with a phage containing a 246 bp palindrome. The authors suggested that this protein provides a recombination-independent pathway for replication beyond a palindrome. It will be interesting to examine the requirements for RecQ using the lambda-free, P_{BAD}-*sbcDC* system. Additionally, the recombination rates could be measured using the zeocin recombination cassettes in Δ*sbcDC* strains containing a palindrome following the introduction of a *recQ*$^-$ mutation. If RecQ is providing an alternative and competitive, recombination-independent pathway, then the rate of recombination may increase in a *recQ*$^-$ mutant.

7.2.3 Further physical analysis

Pulsed-field gel electrophoresis experiments described here demonstrate that SbcCD-cleavage of a palindromic sequence can result in a two-ended DNA DSB. However, there are a number of important factors still unknown about the breaks (1) at what frequency are breaks occurring? (2) Are all breaks two-ended or are a proportion of them one-ended? (3) Are breaks generated only in replicating chromosomes, and if so, (4) are breaks occurring on the lagging strand template? By observing the

chromosomal DNA of various mutant strains, or by using different DNA separation and probing techniques it may be possible to address these points.

7.2.3.1 Quantitative analysis of DNA DSBs

An altered Southern blotting method may allow quantification of (1) the frequency of DNA breaks with respect to unbroken DNA or (2) the ratio of the DNA-ends in the breaks following SbcCD-cleavage at a palindrome. By using ^{32}P labelled DNA probes relative signals from the unbroken and broken DNA fragments can be measured using a storage phosphor screen.

7.2.3.2 Analysis of DNA replication mutants

A number of temperature dependent replication mutants have been described in *E. coli* that affect genes required for establishment of replication forks at *oriC,* like a mutation in the *dnaC* gene (Maisnier-Patin et al., 2001). The DnaC protein is essential for loading the DnaB helicase, which is required for lagging strand DNA synthesis from the origin (Funnell et al., 1987). In a *dnaC*(ts) mutant a shift to the non-permissive temperature allows the progression of established replication forks (assuming replication remains uninterrupted) but new rounds of replication cannot begin (Maisnier-Patin et al., 2001). By shifting a P_{BAD}-*sbcDC recB*$^-$ *dnaC*(ts) strain containing a palindrome to a non-permissive temperature (under SbcCD$^-$ conditions) and allowing initiated replication forks to run to completion before shifting to SbcCD$^+$ conditions, it may be possible to see if DNA DSBs can occur in non-replicating chromosomes or if breaks only occur during replication.

7.2.3.3 Analysis of DNA DSBs by 2D-denaturing gel electrophoresis

It may be possible to use physical methods to investigate whether SbcCD-cleavage of a palindromic sequence occurs during replication in the leading or lagging strand

template. By separating digested DNA fragments through two dimensions (neutral and alkaline gels) and probing for specific DNA strands it should be possible to reveal on which strand breaks occur.

7.2.4 Behaviour of perfect palindromes in the *E. coli* chromosome

All the work in this thesis is based on recombination at a 246 bp interrupted palindrome. However, using the systems described in this thesis it may be possible to integrate a perfect palindrome into the *E. coli* chromosome at *lacZ*. Carrying out similar experiments using strains containing a perfect palindrome may reveal differences and similarities between the two types of palindrome.

7.2.5 Investigating the role of SbcCD in *E. coli*

The actual role of SbcCD in *E. coli* still remains a mystery and although it has palindrome-cleavage properties *in vivo* (as shown in this thesis), this function has only been demonstrated following the insertion of an unnatural, large palindrome into the *E. coli* chromosome. The roles of SbcCD-homologues in eukaryotes have been well described (Connelly and Leach, 2002) and could be conserved in *E. coli*. In this study it was confirmed that a *recB*$^-$ P_{BAD}-*sbcDC* strain was sicker in the presence of SbcCD. Further investigation of this phenotype might help to reveal the actual role of SbcCD in *E. coli*. For example, random mutagenesis of the *recB*$^-$ P_{BAD}-*sbcDC* strain grown under SbcCD$^-$ conditions might yield variants that grow faster under SbcCD$^+$ conditions. Suppressors variants might reveal clues as to the natural role of SbcCD.

References

Albertini, A. M., Hofer, M., Calos, M. P. and Miller, J. H. (1982). On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* **29**, 319-28.

Amundsen, S. K., Taylor, A. F. and Smith, G. R. (2000). The RecD subunit of the Escherichia coli RecBCD enzyme inhibits RecA loading, homologous recombination, and DNA repair. *Proc Natl Acad Sci U S A* **97**, 7399-404.

Anderson, D. G., Churchill, J. J. and Kowalczykowski, S. C. (1999). A single mutation, RecB(D1080A) eliminates RecA protein loading but not Chi recognition by RecBCD enzyme. *J Biol Chem* **274**, 27139-44.

Anderson, D. G. and Kowalczykowski, S. C. (1997a). The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes Dev* **11**, 571-81.

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

Anderson, D. G. and Kowalczykowski, S. C. (1998). Reconstitution of an SOS response pathway: derepression of transcription in response to DNA breaks. *Cell* **95**, 975-9.

Arai, K. and Kornberg, A. (1981). Unique primed start of phage phi X174 DNA replication and mobility of the primosome in a direction opposite chain synthesis. *Proc Natl Acad Sci U S A* **78**, 69-73.

Bennett, R. J., Dunderdale, H. J. and West, S. C. (1993). Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. *Cell* **74**, 1021-31.

Bhattacharya, R. and Beck, D. J. (2002). Survival and SOS induction in cisplatin-treated Escherichia coli deficient in Pol II, RecBCD and RecFOR functions. *DNA Repair (Amst)* **1**, 955-66.

Bianco, P. R., Brewer, L. R., Corzett, M., Balhorn, R., Yeh, Y., Kowalczykowski, S. C. and Baskin, R. J. (2001). Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* **409**, 374-8.

Bianco, P. R. and Kowalczykowski, S. C. (1997). The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single strand

of DNA containing the sequence 5'-GCTGGTGG-3'. *Proc Natl Acad Sci U S A* **94**, 6706-11.

Bidnenko, V., Ehrlich, S. D. and Michel, B. (2002). Replication fork collapse at replication terminator sequences. *Embo J* **21**, 3898-907.

Boehmer, P. E. and Emmerson, P. T. (1991). Escherichia coli RecBCD enzyme: inducible overproduction and reconstitution of the ATP-dependent deoxyribonuclease from purified subunits. *Gene* **102**, 1-6.

Bork, J. M., Cox, M. M. and Inman, R. B. (2001). The RecOR proteins modulate RecA protein function at 5' ends of single-stranded DNA. *Embo J* **20**, 7313-22.

Brent, R. and Ptashne, M. (1981). Mechanism of action of the lexA gene product. *Proc Natl Acad Sci U S A* **78**, 4204-8.

Briggs, G. S., Mahdi, A. A., Weller, G. R., Wen, Q. and Lloyd, R. G. (2004). Interplay between DNA replication, recombination and repair based on the structure of RecG helicase. *Philos Trans R Soc Lond B Biol Sci* **359**, 49-59.

Butler, D. K., Yasuda, L. E. and Yao, M. C. (1995). An intramolecular recombination mechanism for the formation of the rRNA gene palindrome of Tetrahymena thermophila. *Mol Cell Biol* **15**, 7117-26.

Butler, D. K., Yasuda, L. E. and Yao, M. C. (1996). Induction of large DNA palindrome formation in yeast: implications for gene amplification and genome stability in eukaryotes. *Cell* **87**, 1115-22.

Bzymek, M. and Lovett, S. T. (2001). Evidence for two mechanisms of palindrome-stimulated deletion in Escherichia coli: single-strand annealing and replication slipped mispairing. *Genetics* **158**, 527-40.

Calos, M. P. (1978). DNA sequence for a low-level promoter of the lac repressor gene and an 'up' promoter mutation. *Nature* **274**, 762-5.

Chalker, A. F., Leach, D. R. and Lloyd, R. G. (1988). Escherichia coli sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* **71**, 201-5.

Churchill, J. J., Anderson, D. G. and Kowalczykowski, S. C. (1999). The RecBC enzyme loads RecA protein onto ssDNA asymmetrically and independently of chi, resulting in constitutive recombination activation. *Genes Dev* **13**, 901-11.

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

Ciullo, M., Debily, M. A., Rozier, L., Autiero, M., Billault, A., Mayau, V., El Marhomy, S., Guardiola, J., Bernheim, A., Coullin, P. et al. (2002). Initiation of the breakage-fusion-bridge mechanism through common fragile site activation in

human breast cancer cells: the model of PIP gene duplication from a break at FRA7I. *Hum Mol Genet* **11**, 2887-94.

Clark, A. J. (1971). Toward a metabolic interpretation of genetic recombination of E. coli and its phages. *Annu Rev Microbiol* **25**, 437-64.

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.

Collick, A., Drew, J., Penberth, J., Bois, P., Luckett, J., Scaerou, F., Jeffreys, A. and Reik, W. (1996). Instability of long inverted repeats within mouse transgenes. *Embo J* **15**, 1163-71.

Connelly, J. C., de Leau, E. S. and Leach, D. R. (2003). Nucleolytic processing of a protein-bound DNA end by the E. coli SbcCD (MR) complex. *DNA Repair (Amst)* **2**, 795-807.

Connelly, J. C., de Leau, E. S., Okely, E. A. and Leach, D. R. (1997). Overexpression, purification, and characterization of the SbcCD protein from Escherichia coli. *J Biol Chem* **272**, 19819-26.

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.

Connelly, J. C. and Leach, D. R. (1996). The sbcC and sbcD genes of Escherichia coli encode a nuclease involved in palindrome inviability and genetic recombination. *Genes Cells* **1**, 285-91.

Connelly, J. C. and Leach, D. R. (2002). Tethering on the brink: the evolutionarily conserved Mre11-Rad50 complex. *Trends Biochem Sci* **27**, 410-8.

Connelly, J. C. and Leach, D. R. (2004). Repair of DNA covalently linked to protein. *Mol Cell* **13**, 307-16.

Courcelle, J., Belle, J. J. and Courcelle, C. T. (2004). When replication travels on damaged templates: bumps and blocks in the road. *Res Microbiol* **155**, 231-7.

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. and Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli. *Genetics* **158**, 41-64.

Courey, A. J. and Wang, J. C. (1988). Influence of DNA sequence and supercoiling on the process of cruciform formation. *J Mol Biol* **202**, 35-43.

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. and Leach, D. R. (2000). Control of crossing over. *Mol Cell* **6**, 815-26.

Cromie, G. A. and Leach, D. R. (2001). Recombinational repair of chromosomal DNA double-strand breaks generated by a restriction endonuclease. *Mol Microbiol* **41**, 873-83.

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.

Cunningham, L. A., Cote, A. G., Cam-Ozdemir, C. and Lewis, S. M. (2003). Rapid, stabilizing palindrome rearrangements in somatic cells by the center-break mechanism. *Mol Cell Biol* **23**, 8740-50.

de Jager, M., Trujillo, K. M., Sung, P., Hopfner, K. P., Carney, J. P., Tainer, J. A., Connelly, J. C., Leach, D. R., Kanaar, R. and Wyman, C. (2004). Differential arrangements of conserved building blocks among homologs of the Rad50/Mre11 DNA repair protein complex. *J Mol Biol* **339**, 937-49.

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

Dixon, D. A., Churchill, J. J. and Kowalczykowski, S. C. (1994). Reversible inactivation of the Escherichia coli RecBCD enzyme by the recombination hotspot chi in vitro: evidence for functional inactivation or loss of the RecD subunit. *Proc Natl Acad Sci U S A* **91**, 2980-4.

Dixon, D. A. and Kowalczykowski, S. C. (1991). Homologous pairing in vitro stimulated by the recombination hotspot, Chi. *Cell* **66**, 361-71.

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.

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.

Edelmann, L., Spiteri, E., Koren, K., Pulijaal, V., Bialer, M. G., Shanske, A., Goldberg, R. and Morrow, B. E. (2001). AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am J Hum Genet* **68**, 1-13.

Eggleston, A. K. and West, S. C. (2000). Cleavage of holliday junctions by the Escherichia coli RuvABC complex. *J Biol Chem* **275**, 26467-76.

Farah, J. A., Cromie, G., Steiner, W. W. and Smith, G. R. (2005). A novel recombination pathway initiated by the Mre11/Rad50/Nbs1 complex eliminates palindromes during meiosis in Schizosaccharomyces pombe. *Genetics* **169**, 1261-74.

Farah, J. A., Hartsuiker, E., Mizuno, K., Ohta, K. and Smith, G. R. (2002). A 160-bp palindrome is a Rad50.Rad32-dependent mitotic recombination hotspot in Schizosaccharomyces pombe. *Genetics* **161**, 461-8.

Friedberg, E. C., Walker, G.C., Siede, W. (1995). DNA Repair and Mutagenesis. Washington D.C., ASM Press.

Funnell, B. E., Baker, T. A. and Kornberg, A. (1987). In vitro assembly of a prepriming complex at the origin of the Escherichia coli chromosome. *J Biol Chem* **262**, 10327-34.

Gibson, F. P., Leach, D. R. and Lloyd, R. G. (1992). Identification of sbcD mutations as cosuppressors of recBC that allow propagation of DNA palindromes in Escherichia coli K-12. *J Bacteriol* **174**, 1222-8.

Gotter, A. L., Shaikh, T. H., Budarf, M. L., Rhodes, C. H. and Emanuel, B. S. (2004). A palindrome-mediated mechanism distinguishes translocations involving LCR-B of chromosome 22q11.2. *Hum Mol Genet* **13**, 103-15.

Griffin, T. J. t. and Kolodner, R. D. (1990). Purification and preliminary characterization of the Escherichia coli K-12 recF protein. *J Bacteriol* **172**, 6291-9.

Guzman, L. M., Belin, D., Carson, M. J. and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121-30.

Handa, N., Bianco, P. R., Baskin, R. J. and Kowalczykowski, S. C. (2005). Direct visualization of RecBCD movement reveals cotranslocation of the RecD motor after chi recognition. *Mol Cell* **17**, 745-50.

Handa, N., Ohashi, S., Kusano, K. and Kobayashi, I. (1997). Chi-star, a chi-related 11-mer sequence partially active in an E. coli recC1004 strain. *Genes Cells* **2**, 525-36.

Hargreaves, D., Rafferty, J. B., Sedelnikova, S. E., Lloyd, R. G. and Rice, D. W. (1999). Crystallization of Escherichia coli RuvA complexed with a synthetic Holliday junction. *Acta Crystallogr D Biol Crystallogr* **55**, 263-5.

Hegde, S., Sandler, S. J., Clark, A. J. and Madiraju, M. V. (1995). recO and recR mutations delay induction of the SOS response in Escherichia coli. *Mol Gen Genet* **246**, 254-8.

Hegde, S. P., Rajagopalan, M. and Madiraju, M. V. (1996). Preferential binding of Escherichia coli RecF protein to gapped DNA in the presence of adenosine (gamma-thio) triphosphate. *J Bacteriol* **178**, 184-90.

Hendrix, R. W., Roberts, J.W., Stahl, F. W., Weisberg, R.A. (1983). Lambda II: Cold Spring Harbor Monograph series.

Horii, Z. and Clark, A. J. (1973). Genetic analysis of the recF pathway to genetic recombination in Escherichia coli K12: isolation and characterization of mutants. *J Mol Biol* **80**, 327-44.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-8.

Howard-Flanders, P. and Theriot, L. (1966). Mutants of Escherichia coli K-12 defective in DNA repair and in genetic recombination. *Genetics* **53**, 1137-50.

Ivancic-Bace, I., Peharec, P., Moslavac, S., Skrobot, N., Salaj-Smic, E. and Brcic-Kostic, K. (2003). RecFOR function is required for DNA repair and recombination in a RecA loading-deficient recB mutant of Escherichia coli. *Genetics* **163**, 485-94.

Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A. and Shinagawa, H. (1991). Escherichia coli RuvC protein is an endonuclease that resolves the Holliday structure. *Embo J* **10**, 4381-9.

Jacquier, A. and Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* **41**, 383-94.

Jockovich, M. E. and Myers, R. S. (2001). Nuclease activity is essential for RecBCD recombination in Escherichia coli. *Mol Microbiol* **41**, 949-62.

Kaguni, J. M. and Kornberg, A. (1984). Replication initiated at the origin (oriC) of the E. coli chromosome reconstituted with purified enzymes. *Cell* **38**, 183-90.

Kapp, D. S. and Smith, K. C. (1970). Repair of radiation-induced damage in Escherichia coli. II. Effect of rec and uvr mutations on radiosensitivity, and repair of x-ray-induced single-strand breaks in deoxyribonucleic acid. *J Bacteriol* **103**, 49-54.

Keller, K. L., Overbeck-Carrick, T. L. and Beck, D. J. (2001). Survival and induction of SOS in Escherichia coli treated with cisplatin, UV-irradiation, or mitomycin C are dependent on the function of the RecBC and RecFOR pathways of homologous recombination. *Mutat Res* **486**, 21-9.

Kenyon, C. J. and Walker, G. C. (1980). DNA-damaging agents stimulate gene expression at specific loci in Escherichia coli. *Proc Natl Acad Sci U S A* **77**, 2819-23.

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.

Khlebnikov, A., Risa, O., Skaug, T., Carrier, T. A. and Keasling, J. D. (2000). Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *J Bacteriol* **182**, 7029-34.

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.

Kolodrubetz, D. and Schleif, R. (1981a). L-arabinose transport systems in Escherichia coli K-12. *J Bacteriol* **148**, 472-9.

Kolodrubetz, D. and Schleif, R. (1981b). Regulation of the L-arabinose transport operons in Escherichia coli. *J Mol Biol* **151**, 215-27.

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.

Krasin, F. and Hutchinson, F. (1977). Repair of DNA double-strand breaks in Escherichia coli, which requires recA function and the presence of a duplicate genome. *J Mol Biol* **116**, 81-98.

Kurahashi, H. and Emanuel, B. S. (2001). Long AT-rich palindromes and the constitutional t(11;22) breakpoint. *Hum Mol Genet* **10**, 2605-17.

Kurahashi, H., Shaikh, T., Takata, M., Toda, T. and Emanuel, B. S. (2003). The constitutional t(17;22): another translocation mediated by palindromic AT-rich repeats. *Am J Hum Genet* **72**, 733-8.

Kushner, S. R., Nagaishi, H., Templin, A. and Clark, A. J. (1971). Genetic recombination in Escherichia coli: the role of exonuclease I. *Proc Natl Acad Sci U S A* **68**, 824-7.

Kuzminov, A. (1999). Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol Mol Biol Rev* **63**, 751-813, table of contents.

Lau, I. F., Filipe, S. R., Soballe, B., Okstad, O. A., Barre, F. X. and Sherratt, D. J. (2003). Spatial and temporal organization of replicating Escherichia coli chromosomes. *Mol Microbiol* **49**, 731-43.

Leach, D. and Lindsey, J. (1986). In vivo loss of supercoiled DNA carrying a palindromic sequence. *Mol Gen Genet* **204**, 322-7.

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

Leach, D. R., Okely, E. A. and Pinder, D. J. (1997). Repair by recombination of DNA containing a palindromic sequence. *Mol Microbiol* **26**, 597-606.

Leach, D. R. and Stahl, F. W. (1983). Viability of lambda phages carrying a perfect palindrome in the absence of recombination nucleases. *Nature* **305**, 448-51.

LeBlanc, M. D., Aspeslagh, G., Buggia, N. P. and Dyer, B. D. (2000). An annotated catalog of inverted repeats of Caenorhabditis elegans chromosomes III and X, with observations concerning odd/even biases and conserved motifs. *Genome Res* **10**, 1381-92.

Lee, E. H. and Kornberg, A. (1991). Replication deficiencies in priA mutants of Escherichia coli lacking the primosomal replication n' protein. *Proc Natl Acad Sci U S A* **88**, 3029-32.

Lengauer, C., Kinzler, K. W. and Vogelstein, B. (1998). Genetic instabilities in human cancers. *Nature* **396**, 643-9.

Lewis, S., Akgun, E. and Jasin, M. (1999). Palindromic DNA and genome stability. Further studies. *Ann N Y Acad Sci* **870**, 45-57.

Lewis, S. M., Chen, S., Strathern, J. N. and Rattray, A. J. (2005). New approaches to the analysis of palindromic sequences from the human genome: evolution and polymorphism of an intronic site at the NF1 locus. *Nucleic Acids Res* **33**, e186.

Lin, C. T., Lin, W. H., Lyu, Y. L. and Whang-Peng, J. (2001). Inverted repeats as genetic elements for promoting DNA inverted duplication: implications in gene amplification. *Nucleic Acids Res* **29**, 3529-38.

Lindsey, J. C. and Leach, D. R. (1989). Slow replication of palindrome-containing DNA. *J Mol Biol* **206**, 779-82.

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.

Lisnic, B., Svetec, I. K., Saric, H., Nikolic, I. and Zgaga, Z. (2005). Palindrome content of the yeast Saccharomyces cerevisiae genome. *Curr Genet* **47**, 289-97.

Little, J. W. (1984). Autodigestion of lexA and phage lambda repressors. *Proc Natl Acad Sci U S A* **81**, 1375-9.

Little, J. W. (1993). LexA cleavage and other self-processing reactions. *J Bacteriol* **175**, 4943-50.

Little, J. W., Edmiston, S. H., Pacelli, L. Z. and Mount, D. W. (1980). Cleavage of the Escherichia coli lexA protein by the recA protease. *Proc Natl Acad Sci U S A* **77**, 3225-9.

Little, J. W., Mount, D. W. and Yanisch-Perron, C. R. (1981). Purified lexA protein is a repressor of the recA and lexA genes. *Proc Natl Acad Sci U S A* **78**, 4199-203.

Liu, J. and Marians, K. J. (1999). PriA-directed assembly of a primosome on D loop DNA. *J Biol Chem* **274**, 25033-41.

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. (1985). Identification and genetic analysis of sbcC mutations in commonly used recBC sbcB strains of Escherichia coli K-12. *J Bacteriol* **164**, 836-44.

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 Buckman, C. (1995). Conjugational recombination in Escherichia coli: genetic analysis of recombinant formation in Hfr x F- crosses. *Genetics* **139**, 1123-48.

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

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.

Lovett, S. T. and Clark, A. J. (1984). Genetic analysis of the recJ gene of Escherichia coli K-12. *J Bacteriol* **157**, 190-6.

Lovett, S. T. and Kolodner, R. D. (1989). Identification and purification of a single-stranded-DNA-specific exonuclease encoded by the recJ gene of Escherichia coli. *Proc Natl Acad Sci U S A* **86**, 2627-31.

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.

Madiraju, M. V. and Clark, A. J. (1992). Evidence for ATP binding and double-stranded DNA binding by Escherichia coli RecF protein. *J Bacteriol* **174**, 7705-10.

Madiraju, M. V., Lavery, P. E., Kowalczykowski, S. C. and Clark, A. J. (1992). Enzymatic properties of the RecA803 protein, a partial suppressor of recF mutations. *Biochemistry* **31**, 10529-35.

Madiraju, M. V., Templin, A. and Clark, A. J. (1988). Properties of a mutant recA-encoded protein reveal a possible role for Escherichia coli recF-encoded protein in genetic recombination. *Proc Natl Acad Sci U S A* **85**, 6592-6.

Maisnier-Patin, S., Nordstrom, K. and Dasgupta, S. (2001). Replication arrests during a single round of replication of the Escherichia coli chromosome in the absence of DnaC activity. *Mol Microbiol* **42**, 1371-82.

Marians, K. J. (1992). Prokaryotic DNA replication. *Annu Rev Biochem* **61**, 673-719.

McGlynn, P. (2004). Links between DNA replication and recombination in prokaryotes. *Curr Opin Genet Dev* **14**, 107-12.

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.

McGlynn, P. and Lloyd, R. G. (2000). Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell* **101**, 35-45.

McGlynn, P. and Lloyd, R. G. (2002). Genome stability and the processing of damaged replication forks by RecG. *Trends Genet* **18**, 413-9.

McGlynn, P., Lloyd, R. G. and Marians, K. J. (2001). Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. *Proc Natl Acad Sci U S A* **98**, 8235-40.

McGlynn, P., Mahdi, A. A. and Lloyd, R. G. (2000). Characterisation of the catalytically active form of RecG helicase. *Nucleic Acids Res* **28**, 2324-32.

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.

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

Michel, B., Grompone, G., Flores, M. J. and Bidnenko, V. (2004). Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A* **101**, 12783-8.

Mizuuchi, K., Mizuuchi, M. and Gellert, M. (1982). Cruciform structures in palindromic DNA are favored by DNA supercoiling. *J Mol Biol* **156**, 229-43.

Morimatsu, K. and Kowalczykowski, S. C. (2003). RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell* **11**, 1337-47.

Nakayama, K., Irino, N. and Nakayama, H. (1985). The recQ gene of Escherichia coli K12: molecular cloning and isolation of insertion mutants. *Mol Gen Genet* **200**, 266-71.

Naom, I. S., Morton, S. J., Leach, D. R. and Lloyd, R. G. (1989). Molecular organization of sbcC, a gene that affects genetic recombination and the viability of DNA palindromes in Escherichia coli K-12. *Nucleic Acids Res* **17**, 8033-45.

Nasar, F., Jankowski, C. and Nag, D. K. (2000). Long palindromic sequences induce double-strand breaks during meiosis in yeast. *Mol Cell Biol* **20**, 3449-58.

Neidhert F.C., C. R., Ingraham J.L., Lin E.C.C., Low K.B., Magasanik B., Reznikoff W.S., Riley M., Schaechter M., Umburger H.W. (ed). Escherichia coli and Salmonella: ASM Press, Washington DC, USA.

Nimmakayalu, M. A., Gotter, A. L., Shaikh, T. H. and Emanuel, B. S. (2003). A novel sequence-based approach to localize translocation breakpoints identifies the molecular basis of a t(4;22). *Hum Mol Genet* **12**, 2817-25.

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.

Parsons, C. A., Tsaneva, I., Lloyd, R. G. and West, S. C. (1992). Interaction of Escherichia coli RuvA and RuvB proteins with synthetic Holliday junctions. *Proc Natl Acad Sci U S A* **89**, 5452-6.

Parsons, C. A. and West, S. C. (1993). Formation of a RuvAB-Holliday junction complex in vitro. *J Mol Biol* **232**, 397-405.

Pham, P., Rangarajan, S., Woodgate, R. and Goodman, M. F. (2001). Roles of DNA polymerases V and II in SOS-induced error-prone and error-free repair in Escherichia coli. *Proc Natl Acad Sci U S A* **98**, 8350-4.

Pinder, D. J., Blake, C. E., Lindsey, J. C. and Leach, D. R. (1998). Replication strand preference for deletions associated with DNA palindromes. *Mol Microbiol* **28**, 719-27.

Qin, Z. and Cohen, S. N. (2000). Long palindromes formed in Streptomyces by nonrecombinational intra-strand annealing. *Genes Dev* **14**, 1789-96.

Rafferty, J. B., Sedelnikova, S. E., Hargreaves, D., Artymiuk, P. J., Baker, P. J., Sharples, G. J., Mahdi, A. A., Lloyd, R. G. and Rice, D. W. (1996). Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science* **274**, 415-21.

Rangarajan, S., Woodgate, R. and Goodman, M. F. (1999). A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated Escherichia coli. *Proc Natl Acad Sci U S A* **96**, 9224-9.

Rattray, A. J., Shafer, B. K., Neelam, B. and Strathern, J. N. (2005). A mechanism of palindromic gene amplification in Saccharomyces cerevisiae. *Genes Dev* **19**, 1390-9.

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.

Schleif, R. (2003). AraC protein: a love-hate relationship. *Bioessays* **25**, 274-82.

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

Shah, R., Bennett, R. J. and West, S. C. (1994). Genetic recombination in E. coli: RuvC protein cleaves Holliday junctions at resolution hotspots in vitro. *Cell* **79**, 853-64.

Shah, R., Cosstick, R. and West, S. C. (1997). The RuvC protein dimer resolves Holliday junctions by a dual incision mechanism that involves base-specific contacts. *Embo J* **16**, 1464-72.

Sherratt, D. J., Arciszewska, L. K., Blakely, G., Colloms, S., Grant, K., Leslie, N. and McCulloch, R. (1995). Site-specific recombination and circular chromosome segregation. *Philos Trans R Soc Lond B Biol Sci* **347**, 37-42.

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.

Shinagawa, H., Iwasaki, H., Kato, T. and Nakata, A. (1988). RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. *Proc Natl Acad Sci U S A* **85**, 1806-10.

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, K. C. and Meun, D. H. (1970). Repair of radiation-induced damage in Escherichia coli. I. Effect of rec mutations on post-replication repair of damage due to ultraviolet radiation. *J Mol Biol* **51**, 459-72.

Spell, R. M. and Jinks-Robertson, S. (2004). Determination of mitotic recombination rates by fluctuation analysis in Saccharomyces cerevisiae. *Methods Mol Biol* **262**, 3-12.

Spies, M., Bianco, P. R., Dillingham, M. S., Handa, N., Baskin, R. J. and Kowalczykowski, S. C. (2003). A molecular throttle: the recombination hotspot chi controls DNA translocation by the RecBCD helicase. *Cell* **114**, 647-54.

Stahl, F. W. (2005). Chi: a little sequence controls a big enzyme. *Genetics* **170**, 487-93.

Stark, G. R., Debatisse, M., Giulotto, E. and Wahl, G. M. (1989). Recent progress in understanding mechanisms of mammalian DNA amplification. *Cell* **57**, 901-8.

Stasiak, A., Tsaneva, I. R., West, S. C., Benson, C. J., Yu, X. and Egelman, E. H. (1994). The Escherichia coli RuvB branch migration protein forms double hexameric rings around DNA. *Proc Natl Acad Sci U S A* **91**, 7618-22.

Stent, G. S. (1971). Molecular Genetics (An Introductory Narrative). San Francisco: W.H. Freeman and Company.

Tanaka, H., Bergstrom, D. A., Yao, M. C. and Tapscott, S. J. (2005). Widespread and nonrandom distribution of DNA palindromes in cancer cells provides a structural platform for subsequent gene amplification. *Nat Genet* **37**, 320-7.

Tanaka, H., Tapscott, S. J., Trask, B. J. and Yao, M. C. (2002). Short inverted repeats initiate gene amplification through the formation of a large DNA palindrome in mammalian cells. *Proc Natl Acad Sci U S A* **99**, 8772-7.

Taylor, A. F. and Smith, G. R. (1995a). Monomeric RecBCD enzyme binds and unwinds DNA. *J Biol Chem* **270**, 24451-8.

Taylor, A. F. and Smith, G. R. (1995b). Strand specificity of nicking of DNA at Chi sites by RecBCD enzyme. Modulation by ATP and magnesium levels. *J Biol Chem* **270**, 24459-67.

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.

Trinh, T. Q. and Sinden, R. R. (1991). Preferential DNA secondary structure mutagenesis in the lagging strand of replication in E. coli. *Nature* **352**, 544-7.

Tsaneva, I. R., Muller, B. and West, S. C. (1992). ATP-dependent branch migration of Holliday junctions promoted by the RuvA and RuvB proteins of E. coli. *Cell* **69**, 1171-80.

Umezu, K., Chi, N. W. and Kolodner, R. D. (1993). Biochemical interaction of the Escherichia coli RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *Proc Natl Acad Sci U S A* **90**, 3875-9.

Umezu, K. and Kolodner, R. D. (1994). Protein interactions in genetic recombination in Escherichia coli. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J Biol Chem* **269**, 30005-13.

Umezu, K., Nakayama, K. and Nakayama, H. (1990). Escherichia coli RecQ protein is a DNA helicase. *Proc Natl Acad Sci U S A* **87**, 5363-7.

van Gool, A. J., Hajibagheri, N. M., Stasiak, A. and West, S. C. (1999). Assembly of the Escherichia coli RuvABC resolvasome directs the orientation of holliday junction resolution. *Genes Dev* **13**, 1861-70.

van Gool, A. J., Shah, R., Mezard, C. and West, S. C. (1998). Functional interactions between the holliday junction resolvase and the branch migration motor of Escherichia coli. *Embo J* **17**, 1838-45.

Volkert, M. R. and Hartke, M. A. (1984). Suppression of Escherichia coli recF mutations by recA-linked srfA mutations. *J Bacteriol* **157**, 498-506.

Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P. and Nohmi, T. (1999). The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell* **4**, 281-6.

Walker G.C., S., B.T. and Sutton, M. (2000). The SOS Response to DNA damage. In *Bacterial Stress Responses*, (ed. G. a. H.-A. Storz, R), pp. 131 - 144. Washington D.C.: ASM Press.

Wang, J., Chen, R. and Julin, D. A. (2000). A single nuclease active site of the Escherichia coli RecBCD enzyme catalyzes single-stranded DNA degradation in both directions. *J Biol Chem* **275**, 507-13.

Wang, T. C. and Smith, K. C. (1983). Mechanisms for recF-dependent and recB-dependent pathways of postreplication repair in UV-irradiated Escherichia coli uvrB. *J Bacteriol* **156**, 1093-8.

Wang, T. C. and Smith, K. C. (1986). recA (Srf) suppression of recF deficiency in the postreplication repair of UV-irradiated Escherichia coli K-12. *J Bacteriol* **168**, 940-6.

Wang, T. V. and Smith, K. C. (1984). recF-dependent and recF recB-independent DNA gap-filling repair processes transfer dimer-containing parental strands to daughter strands in Escherichia coli K-12 uvrB. *J Bacteriol* **158**, 727-9.

Webb, B. L., Cox, M. M. and Inman, R. B. (1997). Recombinational DNA repair: the RecF and RecR proteins limit the extension of RecA filaments beyond single-strand DNA gaps. *Cell* **91**, 347-56.

Weisemann, J. M. and Weinstock, G. M. (1991). The promoter of the recA gene of Escherichia coli. *Biochimie* **73**, 457-70.

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

Weston-Hafer, K. and Berg, D. E. (1991). Deletions in plasmid pBR322: replication slippage involving leading and lagging strands. *Genetics* **127**, 649-55.

Whitby, M. C., Bolt, E. L., Chan, S. N. and Lloyd, R. G. (1996). Interactions between RuvA and RuvC at Holliday junctions: inhibition of junction cleavage and formation of a RuvA-RuvC-DNA complex. *J Mol Biol* **264**, 878-90.

Whitby, M. C. and Lloyd, R. G. (1995a). Altered SOS induction associated with mutations in recF, recO and recR. *Mol Gen Genet* **246**, 174-9.

Whitby, M. C. and Lloyd, R. G. (1995b). 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.

Xu, L. and Marians, K. J. (2003). PriA mediates DNA replication pathway choice at recombination intermediates. *Mol Cell* **11**, 817-26.

Yao, M. C., Duharcourt, S. and Chalker, D.L. (2002). Mobile DNA II. Washington D. C.: ASM Press.

Yu, M., Souaya, J. and Julin, D. A. (1998). 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.

Zerbib, D., Mezard, C., George, H. and West, S. C. (1998). Coordinated actions of RuvABC in Holliday junction processing. *J Mol Biol* **281**, 621-30.