

# **Towards Construction and Validation of an Ends-In Recombination System in *Escherichia coli***

A Thesis Submitted to the College of Graduate Studies and Research in Partial Fulfillment of the  
Requirement for the Degree of Master of Science  
in the Department of Microbiology and Immunology  
University of Saskatchewan

By

**Kunal Baxi, M.Sc**

© Copyright Kunal Baxi, April 2011, All Rights Reserved.

## **PERMISSION TO USE**

In presenting this thesis in partial fulfillment of the requirements for a Master degree from the University of Saskatchewan, I agree that the libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or make other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Microbiology and Immunology  
Health Science Building, 107 Wiggins Road  
University of Saskatchewan  
Saskatoon SK  
S7N 5E5 Canada

Dean  
College of Graduate Studies and Research  
University of Saskatchewan  
107 Administration Place  
Saskatoon, Saskatchewan  
S7N 5A2 Canada

## ABSTRACT

Homologous recombination is the primary DNA repair pathway in bacteria and it is immensely important in repairing DNA double strand breaks. Components of the homologous recombination pathway have been well conserved throughout evolution as an essential part of cell survival. Homologous recombination plays an important role in cellular processes like DNA repair as well as exchange of genetic information through chromosomal crossover.

During homologous recombination, DNA strand exchange leads to formation of a heteroduplex joint between the invading and displaced DNA strands. This heteroduplex joint is called a Holliday Junction. Resolution of the Holliday Junction proceeds via one of two pathways. In the presence of RuvC and/or RecG, Holliday Junction resolution proceeds via a “cut and paste” pathway where the invading DNA strand replaces a region of homologous DNA on the target DNA. In the absence of RuvC and RecG, Holliday Junction resolution takes place via a “copy and paste” pathway during which DNA synthesis needs to be primed at Holliday Junction intermediates formed during strand invasion.

In an effort to separate this myriad of different requirements, I have attempted to develop a novel “ends-in” recombination assay system using *E. coli* as a model organism. This ends-in system would allow recombinant molecule formation by DNA synthesis of approximately 200 to 2000 bp size interval between the two converging ends of an invading linear dsDNA substrate oriented just like the greek letter  $\Omega$ , but with the arms pointing inwards. In this study, a number of linear dsDNA assay templates were constructed and analyzed. All the constructs had two “arms” of homology to the chromosome pointing inwards i.e. in the ends-in orientation. Using this ends-in system, it was demonstrated that the presence of *chi* (*Crossover Hotspot Initiator*) sites was an important requirement for ends-in recombination in wild type *E. coli* cells. Our studies also showed that ends-in homologous recombination did not occur if *chi* sites were placed at or very near to the ends of the incoming linear dsDNA molecule, suggesting that the *chi* site recognition is efficient only if the incoming dsDNA has *chi* sites internal to the ends. Moreover, it was shown that neither RuvC nor RecG were required for successful recombinant product formation using the *ends-in* assay. This finding reinforces previous observations that suggest the idea that Holliday Junctions can be resolved independent of both RuvC and RecG.

## **ACKNOWLEDGEMENTS**

I have a lot of people that I would like to thank for all the guidance, support and encouragement I have received throughout this Masters' project. To begin, I thank my supervisor Dr. Harold Bull for providing me with this research opportunity. His infinite patience and support and encouragement were of great help in the completion of this research.

I would also like to thank my advisory committee consisting of Dr. Wei Xiao, Dr. Sidney Hayes and Dr. Peter Bretscher for their support, useful insights and helpful suggestions throughout my research.

I am thankful to Dr. Peter Howard for letting me use some of his lab facilities for my research from time to time.

I am grateful to Dr. Landon Pastushok and Timothy Strozen for their feedback and suggestions throughout my research.

I would like to thank Dr. Wei Xiao and the Department of Microbiology and Immunology for supporting me financially during the final months of my Masters' program.

My thanks also go out to all the faculty and staff of the Department of Microbiology and Immunology. I will never forget the assistance I received and the friendships I made while I was a part of this wonderful institution.

Finally, I would like to thank my parents and family members for their undying support, love, and encouragement at all times. Thank you for letting me pursue my dreams and supporting me in my educational endeavors.

# TABLE OF CONTENTS

<b>PERMISSION TO USE</b> .....	<b>i</b>
<b>ABSTRACT</b> .....	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>iii</b>
<b>TABLE OF CONTENTS</b> .....	<b>iv</b>
<b>LIST OF TABLES</b> .....	<b>vii</b>
<b>LIST OF FIGURES</b> .....	<b>viii</b>

## CHAPTER ONE: INTRODUCTION

<b>1.1. DNA damage</b> .....	<b>1</b>
1.1.1. Endogenous DNA damage .....	1
1.1.2. DNA damage caused by environmental DNA damaging agents .....	2
1.1.2.1. DNA damage caused by chemical agents .....	2
1.1.2.2. DNA damage caused by physical agents .....	2
<b>1.2. Mechanisms of DNA damage and repair</b> .....	<b>3</b>
1.2.1. Damage reversal and one strand repair .....	3
1.2.2. Two strand repair .....	4
1.2.3. Daughter strand gap repair (RecFOR pathway) .....	6
1.2.4. Double strand break repair (RecBCD pathway) .....	9
1.2.5. Resolution of recombination intermediates .....	11
1.2.6. Replicative resolution & replication restart .....	12
<b>1.3. Assays to study DNA repair/recombination</b> .....	<b>14</b>
1.3.1. P1 transduction assays .....	14
1.3.2. Hfr crosses .....	15
1.3.3. $\lambda$ x $\lambda$ crosses .....	16
<b>1.4. Objectives of this study</b> .....	<b>17</b>

## CHAPTER TWO: MATERIALS AND METHODS

<b>2.1. <i>E. coli</i> genetics</b> .....	<b>18</b>
2.1.1. <i>E. coli</i> strains and culture .....	18
2.1.2. Plasmid DNA isolation .....	18
2.1.2.1 Plasmid isolation by miniprep method .....	18
2.1.2.2 Plasmid isolation by midiprep method .....	19
2.1.3. UV sensitivity tests .....	20
<b>2.2. Molecular biology techniques</b> .....	<b>20</b>
2.2.1. Bacterial culture and storage .....	20
2.2.2. Preparation of competent cells .....	21
2.2.3. <i>E. coli</i> transformation .....	21
2.2.4. Agarose gel electrophoresis and DNA fragment isolation .....	21
2.2.5. PCR amplification .....	22
2.2.6. Linear DNA electroporation for qualitative analysis .....	24
2.2.7. Linear DNA electroporation for quantitative analysis .....	24
2.2.8. DNA sequencing .....	24

## CHAPTER THREE: RESULTS

<b>3.1. <i>In-vivo</i> cleavage system using I-SceI linearization</b> .....	<b>25</b>
3.1.1. Verification of <i>in-vivo</i> cleavage system plasmid .....	25
3.1.2. Testing of the <i>in-vivo</i> cleavage system .....	26
<b>3.2. <i>In-vitro</i> cleavage system using <i>lacZ</i> plasmid system</b> .....	<b>29</b>
3.2.1. Construction of <i>lacZ</i> plasmid system .....	29
3.2.2. Testing of the <i>lacZ</i> plasmid system .....	30
<b>3.3. Promoter capture (GFP) plasmid system</b> .....	<b>34</b>
3.3.1. Construction of the promoter capture plasmid system .....	34
3.3.2. Testing of the promoter capture plasmid system .....	37

3.3.3. Modification to the promoter capture plasmid system	
– addition of tandem <i>chi</i> sites .....	37
3.3.4. Testing of the modified promoter capture plasmid system	
with tandem <i>chi</i> sites .....	38
3.3.5. Modification to the promoter capture plasmid system	
– addition of the <i>sacB</i> gene .....	46
3.3.6. Testing of the modified promoter capture plasmid system	
with tandem <i>chi</i> sites and <i>sacB</i> .....	49

## **CHAPTER FOUR: DISCUSSION**

4.1. Ends-in recombination – an overview .....	67
4.2. Location-specific recognition of <i>chi</i> sites .....	70
4.3. Role of RecG and RuvC during ends-in recombination .....	71

<b>Future directions</b> .....	<b>73</b>
--------------------------------	-----------

<b>REFERENCES</b> .....	<b>76</b>
-------------------------	-----------

## List of Tables Used In This Study

Table 2 – 1: List of primers used in this study .....	23
Table 3 – 1: Number of transformants obtained for each relevant strain genotype using linear <i>gfp</i> assay plasmid for transformation .....	42
Table 3 – 2: Results of Transformation Round 1 .....	57
Table 3 – 3: Results of Transformation Round 2 .....	57
Table 3 – 4: Results of Transformation Round 3 .....	58



## List of Figures Used In This Study

Figure 1 – 1 (a): Two-strand damage arising from a single strand lesion .....	5
Figure 1 – 1 (b): Two-strand damage arising from a single strand gap .....	5
Figure 1 – 2 (a): Daughter strand gap repair .....	8
Figure 1 – 2 (b): Translesion DNA synthesis .....	8
Figure 1 – 3: Double strand break repair mediated by RecBCD .....	10
Figure 3 – 1: Diagrammatic representation of the detailed gene layout for the <i>in vivo</i> I-SceI linearization assay plasmid system .....	25
Figure 3 – 2: Measurement of plasmid survival in cells grown in medium containing ampicillin supplemented with glucose or arabinose .....	28
Figure 3 – 3: Diagrammatic illustration of the <i>lacZ</i> plasmid layout .....	29
Figure 3 – 4: Analysis of plasmid DNA extracted from <i>recD<sup>-</sup> lacZ</i> assay plasmid transformants .....	33
Figure 3 – 5: Construction of the <i>gfp</i> plasmid system .....	35
Figure 3 – 6: Diagrammatic representation of the linear promoter capture plasmid system with tandem <i>chi</i> sites at the ends of the regions of homology .....	38
Figure 3 – 7: Restriction map showing <i>PstI</i> restriction sites on the recombined pGFP assay plasmid .....	39
Figure 3 – 8: Restriction map showing a) <i>SacI</i> and b) <i>XcmI</i> & <i>StuI</i> sites on the recombined pGFP assay plasmid .....	40
Figure 3 – 9: Restriction map showing a) <i>XcmI</i> and b) <i>XcmI</i> & <i>SphI</i> sites on the recombined pGFP assay plasmid .....	41
Figure 3 – 10: Restriction map showing a) <i>XcmI</i> & <i>SphI</i> and b) <i>ScaI</i> sites on the recombined pGFP assay plasmid .....	43
Figure 3 – 11: Analysis of pGFP plasmid DNA isolated from transformants .....	45
Figure 3 – 12: Construction of the <i>gfp</i> assay plasmid system with the addition of a <i>sacB</i> gene cassette .....	48

Figure 3 – 13: Plasmid DNA isolated from 18 wild type transformants and digested with <i>ScaI</i> .....	49
Figure 3 – 14: Screening wild type transformant plasmid DNA with <i>XcmI</i> and <i>SphI</i> digestion .....	50
Figure 3 – 15: Screening <i>ruvC</i> <sup>-</sup> transformant plasmid DNA with <i>XcmI</i> and <i>SphI</i> digestion .....	51
Figure 3 – 16: Diagrammatic illustration of the pGFP assay plasmid (a) before and (b) after addition of the CFA termination codon with the ribosome binding site .....	53
Figure 3 – 17: Restriction digest of transformants with <i>XbaI</i> & <i>XcmI</i> to verify insertion of the correct <i>gfp</i> gene with the ribosome binding site .....	53
Figure 3 – 18: Restriction digest of plasmid DNA isolated from wild type transformants ..	54
Figure 3 – 19: Restriction digest of plasmid DNA isolated from <i>recA</i> <sup>-</sup> transformants .....	55
Figure 3 – 20: <i>SspI</i> restriction digest of plasmid DNA isolated from wild type and <i>recA</i> <sup>-</sup> transformants .....	56
Figure 3 – 21: <i>SspI</i> restriction digest of plasmid DNA isolated from wild type transformants .....	59
Figure 3 – 22: <i>SspI</i> restriction digest of plasmid DNA isolated from <i>recA</i> <sup>-</sup> Transformants .....	60
Figure 3 – 23: <i>SspI</i> restriction digest of plasmid DNA isolated from <i>ruvC</i> <sup>-</sup> <i>recG</i> <sup>-</sup> transformants .....	61
Figure 3 – 24: <i>SspI</i> restriction digest of plasmid DNA isolated from <i>ruvC</i> <sup>-</sup> transformants .....	62
Figure 3 – 25: <i>SspI</i> restriction digest of plasmid DNA isolated from <i>recG</i> <sup>-</sup> transformants .....	63
Figure 3 – 26: <i>SspI</i> restriction digest of plasmid DNA isolated from <i>recD</i> <sup>-</sup> transformants .....	64
Figure 3 – 27: Confirmatory digest of plasmid DNA isolated from <i>ruvC</i> <sup>-</sup> <i>recG</i> <sup>-</sup> transformant using <i>SspI</i> .....	65
Figure 3 – 28: Plate showing UV sensitivity of different assay strains tested .....	65

# CHAPTER ONE

## INTRODUCTION

### **1.1 DNA damage**

DNA is the genetic information repository for living organisms. It is therefore important that the information encoded by DNA be extremely stable as any change in this information would be deleterious to the cell. Under physiological conditions, DNA is subjected to physical and chemical modifications (Sakai et al., 2006; Duncan & Miller, 1980; Steenken, 1989). These modifications can lead to a change in the genetic information encoded by DNA. When the change is copied to daughter cells it is called mutation. DNA damage that arises as a part of the cell's normal metabolic activities, DNA replication or cell division is termed as endogenous DNA damage (Lindahl, 1993). DNA damage arising from external chemical or physical agents is termed exogenous damage (Friedberg, 2003).

#### **1.1.1 Endogenous DNA damage**

DNA damage can arise spontaneously due to normal metabolic activities in the cell. This includes damage arising from reactive oxygen species (ROS) (Sakai et al., 2006) or oxidative deamination (Duncan & Miller, 1980).

A common type of endogenous DNA damage is the generation of abasic sites by enzymatic removal or hydrolysis of modified bases that have undergone chemical modifications such as alkylations or aminations (Hevroni & Livneh, 1988). When a growing replication fork hits an abasic site, it can either cause strand breakage (Lindahl, 1993) or it can cause mutations via misincorporation of DNA bases by the intrinsically error-prone Y family of DNA damage-bypass polymerases (Takeshita & Eisenberg, 1994) (Ohmori et al., 2001).

DNA is damaged by superoxide radicals formed by the autoxidation of different reduced electron carriers and redox enzymes of the electron transport chain (Imlay & Fridovich, 1991). Certain hydroxyl groups (OH<sup>•</sup>) damage DNA by chemically modifying the bases to generate OH adduct radicals. For example, adenine gets converted to two OH adducts namely: C4-OH and C8-OH adducts (Steenken, 1989). Such modified DNA bases lead to misincorporation of

nucleotides during DNA replication. The OH<sup>-</sup> groups react with the sugar moiety of DNA to give rise to either sugar modifications or DNA breaks (Teoule, 1987).

### **1.1.2. DNA damage caused by environmental DNA damaging agents**

#### **1.1.2.1. DNA damage caused by chemical agents**

DNA can be damaged by chemical agents that lead to formation of DNA-alkyl adducts. These alkylating agents work by binding to nucleophilic sites on DNA. One such alkylating agent is MMS (methyl methanesulfonate) that alkylates DNA at N<sup>7</sup>-guanine and N<sup>3</sup>-adenine positions (Lindahl et al., 1988). Such lesions are an impediment for replication fork progression and may lead to replication fork collapse (Tercero & Diffley, 2001). Another alkylating agent is N-methyl-N-nitro-N'-nitrosourea (MNNG), a derivative of nitrosoguanidine that forms O-alkyl lesions such as O<sup>6</sup>-MeG and O<sup>4</sup>-MeT. These modified bases can pair with both deoxythymine and deoxyguanine, thereby leading to mismatches which give rise to transversion mutations (Preston et al., 1986).

Damage to DNA called crosslinking arises between opposite (interstrand) or same (intrastrand) DNA strands. Chemicals like nitrogen mustard can crosslink with DNA at N<sup>7</sup> position of guanine on opposite strands forming interstrand crosslinks (Kohn et al., 1965; Kohn et al., 1966). Cisplatin, on the other hand, causes formation of both interstrand and intrastrand DNA crosslinks, besides forming DNA-protein crosslinks (Macquet et al., 1980; Hannan et al., 1984).

#### **1.1.2.2. DNA damage caused by physical agents**

One of the commonest physical agents to cause exogenous DNA damage is UV light. Upon exposure to UV light, adjacent pyrimidine molecules become covalently linked at the C<sup>6</sup> and C<sup>4</sup> positions leading to formation of a pyrimidine-pyrimidine dimer. UV light can also bring about other DNA base damage as well as DNA strand breaks but these are not a major consequence of DNA damage by UV light (Friedberg, 2003).

Unlike UV light, x-rays and gamma rays are potential sources of ionizing radiation. Ionizing radiation also causes direct damage to DNA. One such direct damage to DNA is the formation of double strand breaks (DSBs) (Bresler et al., 1984). Ionizing radiation also brings

about formation of hydroxyl radicals that damage DNA and causes formation of single strand breaks (SSBs). Formation of interstrand cross links is an additional consequence of ionizing radiation (Friedberg et al., 1995).

## **1.2. Mechanisms of DNA damage and repair**

### **1.2.1. Damage reversal and one strand repair**

Bacterial DNA is subjected to constant stress from endogenous and exogenous agents. These agents may be physical or chemical, and give rise to a variety of lesions. It is imperative for the cell to repair these lesions in order to ensure faithful replication of its genetic material. Depending on the type of lesions, the cell has numerous repair pathways at its disposal to repair these lesions. For example, the enzyme photolyase, using near-UV light, is capable of repairing UV induced pyrimidine dimers in the cell (Sancar, 1994b). For repairing lesions formed due to hydroxyl radicals like *O*<sup>6</sup>-methylguanine, *E. coli* uses a methyltransferase protein called Ada (**A**Daptive response to **A**lkylating agents) that transfers a methyl group from the modified base *O*<sup>6</sup>-methylguanine to itself (Shevell et al., 1990). It has been found that this Ada protein is converted into a transcriptional regulator for methylation-resistance genes due to the transfer of a methyl group from a DNA methylphosphotriester to its own Cys69 residue (Nakamura et al., 1988).

Besides DNA damage arising due to lesions or modified bases, there exist other ways in which the duplex DNA strand might get damaged. For example, misincorporation of nucleotides in the nascent DNA strands during DNA replication might give rise to mismatches. These mismatches are repaired by the removal of a flanking stretch of DNA bases including the mismatch. The resulting single stranded gap in the DNA is then repaired by using the intact single stranded DNA as a complimentary template strand. This repair system, termed as methyl directed mismatch repair (MMS), usually excises several hundred bases at once from a single strand of DNA to correct mismatches arising during DNA replication (Modrich, 1994). On the other hand, the base excision repair (BER) system is used to repair DNA damage that results from chemical modifications of bases. The BER system utilizes an enzyme called a DNA glycosylase to remove a modified base, thereby producing an apurinic or apyrimidinic (AP) site by breaking the phosphodiester bond at the 5' end of the modified base. The single base gap is then re-synthesized by DNA polymerase and the nick is sealed by DNA ligase (Dianov et al.,

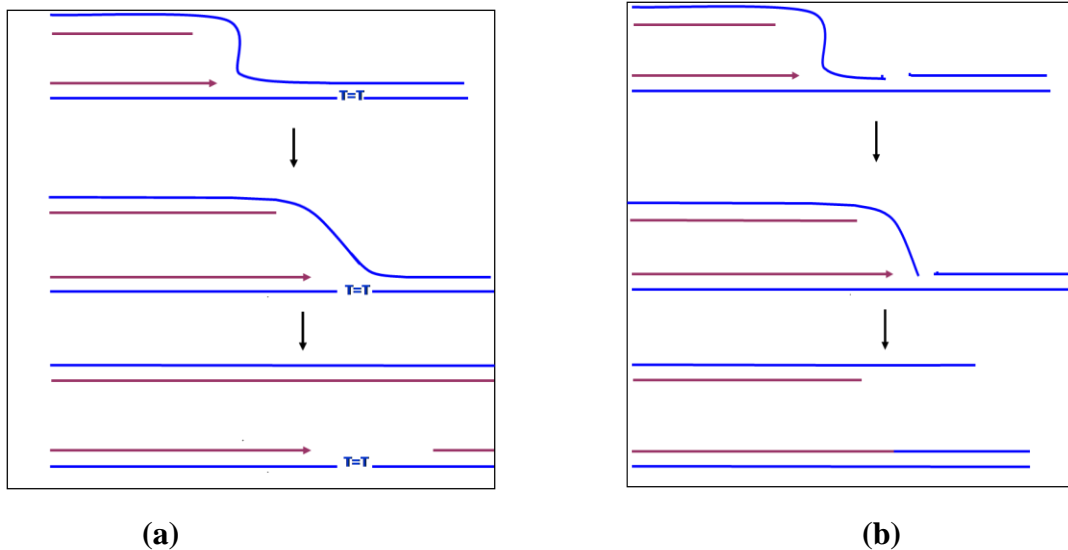
1992). Another repair mechanism called the nucleotide excision repair (NER) system is brought into play when there is a DNA-distorting lesion such as a pyrimidine dimer that needs to be repaired. For nucleotide excision repair, an enzyme called excinuclease (consisting of the UvrA, UvrB and UvrC proteins) recognizes and binds to the site of the lesion and then hydrolyzes two phosphodiester bonds, one 7 nucleotides 5' and the other 3 or 4 nucleotides 3' of the lesion, following which the UvrD helicase removes the nicked DNA segment (Sancar & Rupp, 1983) (Yeung et al., 1983). The ensuing gap is filled in by DNA polymerase and the nick is then sealed by DNA ligase (Sancar, 1994a) (Selby & Sancar, 1994).

### **1.2.2. Two strand repair**

Although the majority of DNA damage affects only one strand, there are occasions wherein both DNA strands are damaged at the same points opposite each other. While this type of DNA damage is less frequent, it poses great risk to the cell if not repaired, since they represent a serious impediment to the replication fork. While studying DNA recombination occurring at sites of DNA replication, it was discovered that replication of DNA in regions having one-strand lesions leads to a higher frequency of recombination at these sites, suggesting that DNA replication generates two-strand lesions when it runs into a single strand gap. (Lin & Howard-Flanders, 1976) This is illustrated in figure 1-1 (b) below.

There are two distinct ways in which a two-strand lesion might occur. One way that two-strand lesions arise is when the replication fork hits a noncoding lesion. The presence of the lesion blocks the replicative DNA polymerases from proceeding past it. DNA replication can then resume downstream of the lesion but it would leave behind a single stranded gap in the nascent DNA strand that is termed as a daughter strand gap (Rupp & Howard-Flanders, 1968; Wilkins & Howard-Flanders, 1968). If a subsequent replication fork attempts to replicate through this gapped template it is thought to cause a replication fork collapse, as illustrated in Figure 1 – 1 (b). This replication fork collapse can be detected as replication induced chromosomal fragmentation (Bonura & Smith, 1975; Wang & Smith, 1986). Another way that two-strand lesions arise is through exposure to ionizing radiation like X-rays and gamma rays. Ionizing radiation generates free radicals that damage molecules in their vicinity, which accounts for multiple instances of damage in a molecule (Ward, 1988). Another two-strand lesion termed as crosslink is observed in DNA treated with psoralen and UV-light or mitomycin C (Cimino et

al., 1985; Tomasz et al., 1987). Hence, two-strand lesions arise either directly due to physical or chemical agents or due to DNA replication on a DNA template that has a one-strand lesion.



**Figure 1-1 (a): Two-strand damage arising from a single strand lesion:** As the replication fork travels from left to right in the above illustration, a single strand lesion in one of the parental strands leads to formation of a single stranded gapped region in the daughter strand. If left unrepaired, this gap can then get converted to a double stranded break when a subsequent replication fork reaches this gap, as illustrated in figure 1-1 (b). Adapted from Kuzminov (1999).  
**Figure 1-1 (b): Two-strand damage arising from a single strand gap:** As the replication fork travels from left to right in the above illustration, a single strand interruption in one of the parental strands leads to disintegration of the replication fork and detachment of a double stranded end from the replication fork. Adapted from Kuzminov (1999).

Although two-strand DNA lesions are less frequent than one strand lesions, it is important for the cell to repair these lesions. It has been observed for mutants defective in certain types of DNA repair that even a single two-strand lesion would be fatal (Howard-Flanders et al., 1969; Sinden & Cole, 1978). From an evolutionary perspective, it has been speculated that the significant recombinational repair capacity of *E. coli* cells probably evolved for repairing frequent dsDNA damage in the wild (Kuzminov, 1999).

A majority of two-strand DNA breaks in growing *E. coli* cells arise due to attempted DNA replication on template DNA that has one-strand lesions. Hence, it would only seem logical for the cell to stop DNA replication upon sensing DNA damage. Eukaryotic cells employ DNA damage checkpoints to sense DNA damage and stop DNA replication if necessary until the damage is repaired (Kastan & Kuerbitz, 1993; Maity et al., 1994). Prokaryotes like *E. coli*, on the other hand, rely on a concerted action by repair proteins that are specifically activated following

DNA damage. One of the proteins (SulA) that is activated following DNA damage specifically inhibits cell division until the DNA damage has been repaired (Johnson, 1977) (George et al., 1975) (Huisman et al., 1984). The concerted activation of repair proteins greatly enhances the repair capacity of the cell by several orders of magnitude and is termed as the SOS response (Radman, 1975).

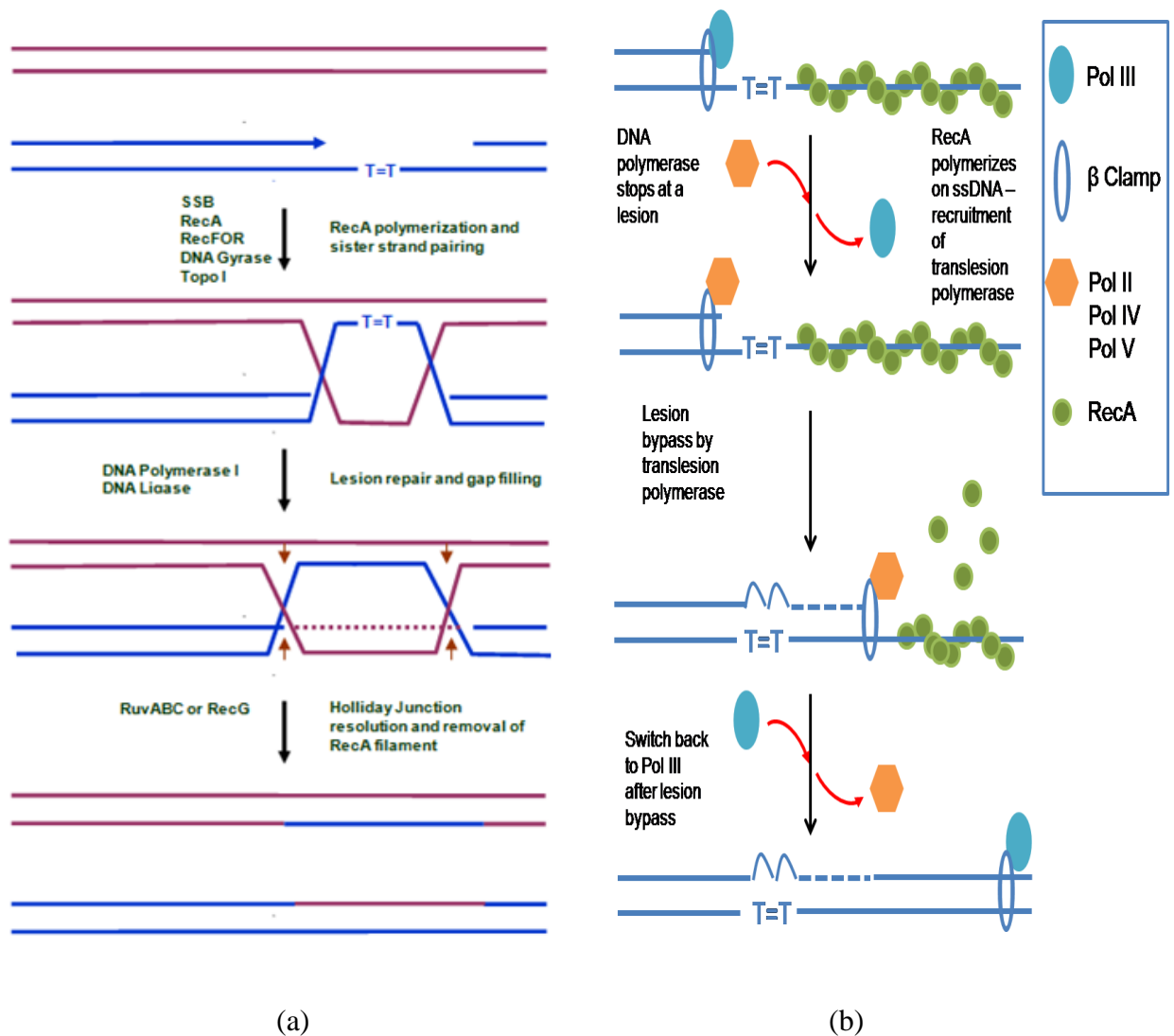
### **1.2.3. Daughter strand gap repair (RecFOR pathway)**

Early studies on the formation of daughter strand gaps and their repair was initiated after it was found that when excision repair-deficient cells are irradiated with low doses of UV light, the rate of DNA synthesis remained unchanged (Rupp & Howard-Flanders, 1968) although the newly synthesized DNA in these cells had a lower molecular weight, indicating single-strand interruptions. It was later found that these single-stranded regions are formed when the replisome meets a non-coding lesion like a pyrimidine dimer which blocks the progress of the replisome (Bonner et al., 1992), as shown in figure 1-1 (a). When the replication fork encounters such non-coding lesions, the replication fork stalls and replication is restarted downstream of the lesion (Rupp & Howard-Flanders, 1968). Studies using transformation of phage M13 circular ssDNA have shown that a single lesion in ssDNA decreases the transformation efficiency to less than 0.5% of lesion-free control circular ssDNA (Lawrence et al., 1993; Horsfall & Lawrence, 1994). A blocked replisome is likely to restart 100 (Wang & Chen, 1992) to 800 (Iyer & Rupp, 1971) nucleotides downstream of the blocking lesion, but the region of DNA from the lesion to the replisome restart is left single stranded. It would seem reasonable to assume that this single stranded region could be repaired by using translesion synthesis to replicate DNA but this repair would come at the cost of fidelity. However, it was found that the gap filling was accomplished by formation of hybrid DNA strands wherein nascent DNA strands were linked to template DNA strands (Rupp et al., 1971). The numbers of such strand exchanges were roughly equal to the number of lesions on the template DNA indicating that daughter strand gap repair is brought about by DNA strand exchange arising due to recombinational repair. This conclusion was in line with the findings that *recA* mutants, deficient in homologous recombination, are also deficient in daughter strand gap repair (Howard-Flanders et al., 1969; Radman et al., 1970; Smith & Meun, 1970; Ganesan, 1974). Daughter strand gap repair repairs the single stranded gaps but



the lesions still persist, and these lesions are usually repaired by excision repair. In excision repair mutants, these lesions persist in the DNA for many more generations.

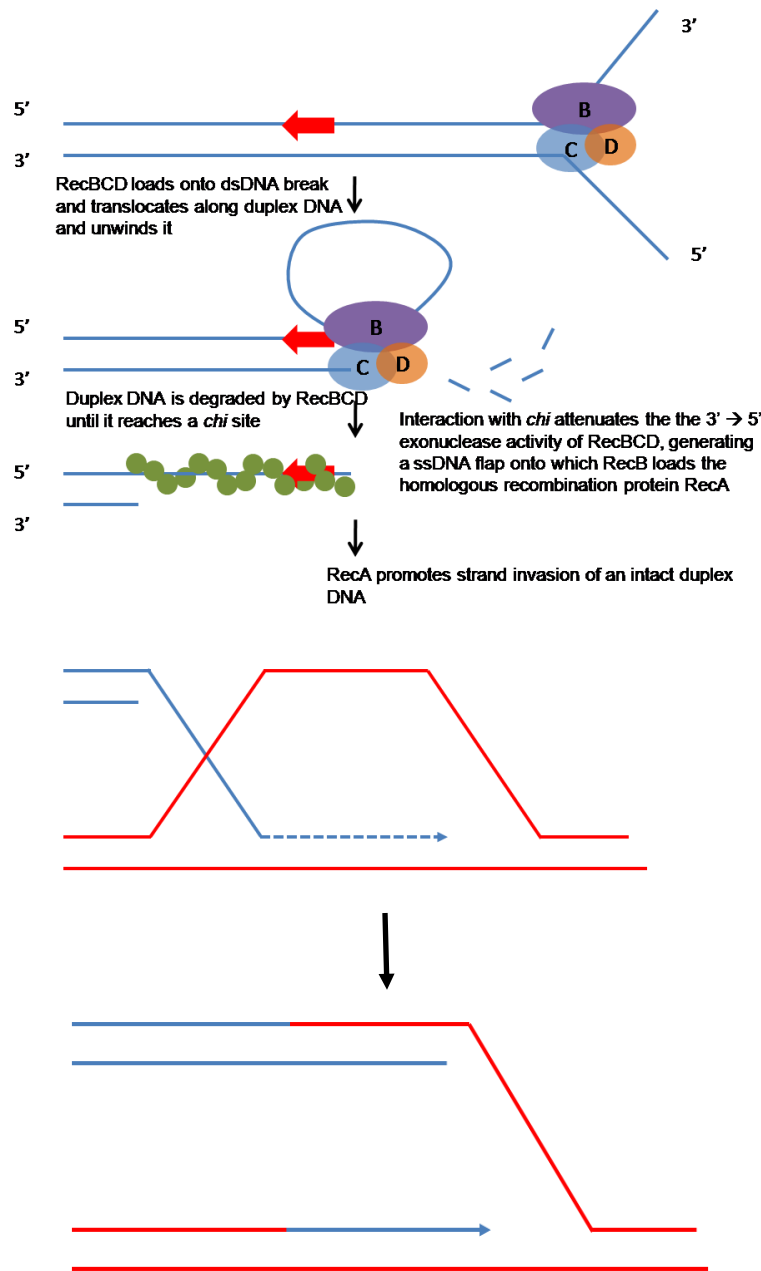
In *E. coli*, daughter strand gaps are repaired by the RecFOR pathway, which is named after the three primary proteins that specifically participate in this mode of repair (Horii & Clark, 1973; Ganesan & Seawell, 1975; Rothman et al., 1975). The sequence of events leading to single strand gap repair is postulated to be as follows (Clark & Sandler, 1994): In vivo, generation of single strand gaps in DNA leads to SSB polymerization on the single stranded gap region. The recFR complex is postulated to recognize ssDNA-dsDNA junctions following which RecOR are guided to the SSB coated DNA. The presence of RecO enables RecA to polymerize onto SSB coated DNA. Electron microscope studies using purified *Thermus thermophilus* RecF pathway proteins have shown that it is RecO that catalyzes displacement of SSB from ssDNA (Inoue et al., 2008). Pull-down assays using purified *Thermus thermophils* proteins have shown that after displacement, SSB still remains attached to RecO, which in turn is attached to ssDNA and that RecR catalyzes the ssDNA-dependent ATPase activity of RecA (Inoue et al., 2008). Following RecA polymerization, the ssDNA is paired with its complementary ssDNA from an intact dsDNA by displacing one of the double strands, thereby creating a D-loop (displacement loop). DNA polymerase is then loaded onto the gapped DNA substrate and the gap is filled (Mosig, 1987). The final step is the removal of Holliday junctions by either RuvABC or RecG (Kowalczykowski, 2000). This process is illustrated in Figure 1-2 (a). In the absence of an intact homologous duplex, the cell resorts to translesion DNA synthesis, which repairs the ssDNA gap, albeit at the cost of possible mutagenesis (illustrated in Figure 1-2 (b)).



**Figure 1-2: (a) Daughter strand gap repair:** To initiate repair, the ssDNA at the gapped region invades an intact duplex DNA mediated by RecA, following which a single intact strand of DNA invades a gapped ssDNA mediated by RecA following which single stranded gaps left in daughter strand DNA are repaired by homologous recombination by means of the RecFOR pathway as illustrated above. Adapted from Kuzminov (1999). **(b) Translesion DNA synthesis:** When the main replicative polymerase DNA Pol III stalls upon encountering a lesion on template ssDNA, certain specialized polymerases termed translesion polymerases are recruited to sites of DNA lesions. The main replicative polymerase Pol III is displaced from the replication machinery, and the translesion polymerase resumes DNA replication until the lesion is bypassed, at which point DNA Pol III displaces the translesion polymerase and takes over DNA replication. Adapted from (Woodgate, 2001).

#### **1.2.4. Double strand break repair (RecBCD pathway)**

While repair of single strand gaps requires the RecFOR pathway, repair of dsDNA breaks on the other hand is dependent on the RecBCD pathway. RecBCD is a heterotrimeric enzyme that possesses DNA helicase and exonuclease activity. The RecBCD pathway is specific for dsDNA breaks and acts at free dsDNA ends that arise either as a consequence of replication of DNA containing single stranded gaps or directly by physical agents like gamma rays (Kuzminov, 1999). Since this process of double strand break repair is catalyzed by the heterotrimeric complex of RecB, RecC and RecD jointly known as the RecBCD complex, this pathway is termed as RecBCD pathway (Stahl & Stahl, 1977). Both the RecB and RecD subunits are helicases, but they unwind dsDNA with different polarities. While RecB possesses a 3' to 5' helicase activity, RecD possesses a 5' to 3' helicase activity (Taylor & Smith, 2003). RecBCD binds to blunt or nearly blunt dsDNA ends. Once it recognizes a free dsDNA end, RecBCD loads onto it and starts to unwind the DNA helix, with the RecD helicase translocating faster compared to the slower RecB helicase. This difference in DNA translocation speeds leads to formation of a 3' ssDNA loop ahead of the RecB subunit. Until this enzyme complex hits the recombination hotspot *chi*, both the DNA strands are degraded uniformly. Upon encountering the recombinational hotspot *chi* (5'-GCTGGTGG-3'), the RecBCD enzyme pauses, following which there is a *chi*-mediated conformational change in the enzyme that leads to attenuation of the 3' to 5' nuclease activity but the 5' to 3' nuclease activity is not affected (Taylor & Smith, 2003). This attenuation gives rise to a 3' end ssDNA flap onto which RecA loads, facilitated by RecB (Braedt & Smith, 1989) (Amundsen et al., 2007), following which RecA initiates homologous pairing between the ssDNA and an intact homologous duplex DNA.



**Figure 1-3:** Double strand break repair mediated by RecBCD: Double stranded ends are first acted upon by RecBCD with the RecB and RecD helicases unwinding the duplex DNA until the RecBCD complex reaches a *chi* site. Upon *chi* site interaction, the 3' to 5' nuclease activity is attenuated, and the 5' end of DNA is degraded. This leads to formation of a 3' ended flap onto which RecC loads the RecA recombinase. RecA primes the 3' flap of ssDNA for strand invasion followed by DNA replication restart. The Holliday Junctions that are formed are then resolved by the RuvABC resolvase (or, alternately, by RecG helicase). Adapted from Kuzminov (1999).

### **1.2.5. Resolution of recombination intermediates**

Once homologous pairing has been established, the cell has two different means at its disposal to resolve the Holliday Junctions formed by homologous pairing. One method utilizes the RuvABC resolvase complex and the other utilizes the RecG helicase.

The RuvABC resolvase is an enzyme complex made up of 3 subunits – RuvA, RuvB and RuvC. RuvA is a 22kDa protein that exists as a tetramer in solution (Tsaneva et al., 1992). Its binding affinity is greater for Holliday Junctions but it also binds to both ssDNA and dsDNA (Shiba et al., 1991). Once it binds to a Holliday Junction, RuvA is thought to target RuvB to the Holliday Junction. RuvB then forms hexameric rings along the DNA (Dickman et al., 2002) and together, the RuvA-RuvB complex “threads” the duplex DNA between them using energy derived from ATP hydrolysis. RuvC is a sequence specific endonuclease that symmetrically cleaves duplex DNA at Holliday Junctions in a sequence specific manner, acting at the consensus sequence 5'(A/T)TT(G/C) (Shida et al., 1996).

Early studies found that *ruvAB E. coli* cells were only moderately defective in DNA repair and homologous recombination thereby suggesting that other protein(s) might be providing Holliday Junction resolution in the absence of RuvAB. Later on, it was found that a *recG* mutation combined with a *ruv* mutation nearly eliminates the DNA repair and conjugational recombination activity of the cell (Lloyd et al., 1984; Lloyd, 1991). RecG is a helicase that has activities very similar to RuvAB, but with a few differences. Unlike RuvAB, RecG is unable to drive branch migration for DNA sequences having a heterologous region of 30 or more basepairs (Lloyd & Sharples, 1993b; Whitby & Lloyd, 1998). RecG can also dissociate DNA strands brought together by RecA-mediated strand pairing, even if the DNA strands are covered by a RecA nucleofilament (Whitby et al., 1993). In order to find out if RecG substitutes for RuvABC for resolving Holliday Junctions, conjugational recombination and P1 transduction tests were carried out using *ruvA recG*, *ruvB recG* and *ruvC recG* strains and in each of these cases it was found that UV resistance and conjugational and P1 recombination were greatly depressed, similar to a *recA*<sup>-</sup> strain (Lloyd, 1991). This led researchers to believe that RecG has functional overlap with RuvABC, but it acts by a different route and that it does not substitute for RuvAB in Holliday Junction branch migration in facilitating RuvC dependent resolution. It was strongly suggested that RecG and RuvABC provide two alternative and partially overlapping pathways for resolution of Holliday Junctions. However, unlike the RuvABC pathway, RecG

does not possess any nuclease activity (Lloyd & Sharples, 1993a), leading to the hypothesis that an unknown protein provides the nuclease activity that is necessary for resolution of Holliday Junctions, in a manner similar to the RuvC resolvase activity. A genetic screening undertaken to identify mutations that are incompatible with a *ruvC* deletion identified three genes (*dam*, *polA* and *uvrD*), none of which encode a protein with any nuclease activity (Zhang et al.). This study suggested that the RecG pathway of branch migration is very ineffective, and is probably a secondary function of RecG. It was suggested that the primary function of RecG was to suppress “promiscuous” initiation of recombination at D-loops and branched DNA structures. It is interesting to note that a suppressor mutation of *ruvC* (Mandal et al., 1993) called *rus-1* was identified and mapped to the *rus* open reading frame. It was shown that the *rus-1* suppressor mutation is a 1.4 kb insertion upstream of the *rus* open reading frame which activates its transcription (Sharples et al., 1994), and that the Rus protein binds to and symmetrically cleaves four-way DNA junctions, just like those generated by RecG and RuvAB (Bolt & Lloyd, 2002).

#### **1.2.6. Replicative resolution & replication restart**

Models proposed for the resolution of recombination intermediates are defined based on the involvement of DNA replication and fall into two categories – (a) Break-join recombination (RuvABC and/or RecG dependent) and (b) Break-copy recombination (RuvABC and RecG independent). Although the RecBCD pathway of recombination in *Escherichia coli* is known to produce recombinant molecules through the break-join mechanism, there has not been extensive scientific evidence to demonstrate the existence of the break-copy mechanism. The most prevalent scientific research utilized bacteriophage lambda DNA substrates as a means to study the bacterial recombination system. Such studies have proven the existence of the break-join mechanism for recombination in *E. coli* (Meselson & Weigle, 1961; Meselson, 1964; McMilin & Russo, 1972). This evidence supported the hypothesis that break-join was the major route for RecBCD recombination, although the break-copy mechanism was not excluded (Thaler & Stahl, 1988; West, 1992; Kowalczykowski et al., 1994). However, recently, there has been evidence to show that RecBCD mediated recombination can occur by a break-copy mechanism and that this mechanism is responsible for approximately half of the products arising out of RecBCD mediated recombination, and that these mechanisms are mutually exclusive (Motamedi et al., 1999). Various studies have tried to provide evidence for the break-copy mechanism of RecBCD

mediated recombination, but it has been difficult to physically connect the two processes in the same DNA molecule using current assay systems.

Using classical genetics, it has been shown that there are multiple pathways to restart replication forks that stall or collapse in *E. coli* (Sandler, 2000). The proteins that constitute these pathways were first identified as being essential components of the phage  $\phi$ x174 replication machinery (Lee et al., 1990). In *E. coli*, three pathways are responsible for restarting stalled and/or collapsed replication forks. These pathways are named after the proteins that take part in each pathway and are called the PriA-PriB pathway, the PriA-PriC pathway and the Rep-PriC pathway (Sandler, 2000). PriA is an ATPase and 3'  $\rightarrow$  5' helicase that has been shown to be absolutely required for replication of  $\phi$ x174 as well as ColE1 type plasmids (Marians, 1999). PriB and PriC are accessory primosomal proteins that act together with PriA to repair collapsed replication forks (Zavitz et al., 1991). Rep is an ATP-dependent DNA helicase with a 3'  $\rightarrow$  5' polarity of unwinding that has been shown to have a role in replication restart along with PriC. The PriA-PriB, PriA-PriC and Rep-PriC proteins act by modulating DNA structure at collapsed replication forks and catalyze loading of the DnaB helicase in a DnaC (helicase loader)-dependent manner (Zavitz et al., 1991). It has been shown that these pathways often overlap in function and therefore provide a robust means of replication restart. Inactivation of any single pathway by itself does not result in a great loss of viability, although inactivation of the PriA protein results in UV sensitivity, defects in DNA recombination and rich media sensitivity (Marians, 1999). Loss of PriA also induces the SOS response (Masai et al., 1994), presumably due to inactivation of the PriA-PriB and the PriA-PriC pathways. It is unclear what types of substrates are processed by these replication restart pathways, although biochemical studies have shown that both the PriA-PriC and Rep-PriC pathways process replication forks that have stalled due to a lesion on the leading strand by unwinding nascent lagging strand DNA and creating a ssDNA region onto which the DnaB helicase can be loaded by the DnaC helicase loader (Heller & Marians, 2005a) (Heller & Marians, 2005b). It has been suggested that the PriA-PriB pathway serves to restart replication forks that have collapsed at sites of double stranded DNA breaks, although this has yet to be demonstrated conclusively (Harinarayanan & Gowrishankar, 2004). A lot of questions remain unanswered and there are certainly gaps in the current knowledgebase about how replication restart is carried out. It is unclear how certain suppressor mutations in

DnaC alleviate phenotypes associated with defects in key replication restart proteins and what kind of replication restart substrates are handled by each pathway (Boonsombat et al., 2006); (Sandler et al., 1999). The extent of functional overlap of these pathways and how they are utilized to provide a robust means of replication restart is the subject of ongoing research.

### **1.3. Assays to study DNA repair/recombination**

The process of homologous recombination is ubiquitous in nature. Homologous recombination plays a key role in DNA repair and chromosomal crossing over and it has been extensively studied in both *Escherichia coli* and *Saccharomyces cerevisiae*. In *E. coli*, homologous recombination has been associated with the repair of collapsed replication forks, with different assay systems being utilized to study DNA recombination coupled to replication fork restart. Some of these assays make use of lambda by lambda crosses, whereas others utilize Hfr mating, or P1 transduction. It is important to note, though, that most of these techniques utilize an ends-out orientation of linear DNA very much like the Greek letter omega  $\Omega$  and that these techniques mostly give rise to gene replacement events.

#### **1.3.1. P1 transduction assays**

Early studies revealed the usefulness of bacteriophage P1 as a vector for transfer of DNA from one host to the other (Lennox, 1955). Horizontal transfer of genetic information such as P1 transduction relies on ends-out targeting of host DNA by the incoming duplex DNA and occurs through RecA-mediated homologous recombination. It was noticed early on that generation of transductants during P1 transduction was marker dependent and varied significantly depending on the region of chromosome that was being transduced (Newman & Masters, 1980). For example, markers closer to the origin of replication transduced at a higher frequency than markers that were away from the origin (Masters, 1977). During P1 transduction, the incoming dsDNA is integrated into the chromosome of the host *E. coli* cell using homologous recombination. This process is dependent on RecA and the RecBCD pathway of homologous recombination. It is now known that in P1 transduction, the dsDNA ends of the incoming linear DNA is acted upon by one or more nucleases, which prepare the DNA for loading of the RecA recombinase, which will then facilitate strand displacement and homologous pairing of the ssDNA strand, followed by resolution of the DNA crossover (Centore et al., 2008).



Early on, Stahl made the observation that *chi* sites enhance (Dower & Stahl, 1981) P1 mediated transduction and conjugation. Soon, it was observed that defects in genes encoding certain recombination or replication restart proteins such as PriB and RuvC led to changes in the frequency at which P1 transductants were obtained, and this observation enabled the use of P1 transduction as an assay to dissect the role of certain proteins in replication restart and recombination (Sandler et al., 1996) (McCool & Sandler, 2001). However, the genetic requirements for certain processes like recombination coupled replication restart cannot be elucidated by P1 transduction assays. For example, the genetic requirements for generation of transductants during P1 transduction are different in the presence or absence of RuvC and RecG. In the presence of RuvC and/or RecG, transductants arise via a “cut and paste” reaction where incoming foreign DNA replaces a region of homologous DNA in the host chromosome target. In the absence of RuvC and RecG, transductants arise via a “copy and paste” reaction during which DNA synthesis needs to be primed at Holliday Junction intermediates formed during strand invasion (Cromie et al., 2001). It is difficult to differentiate between transductants that arise via a “cut and paste” mechanism from those that arise via a “copy and paste” mechanism. Moreover, generation of transductants in the absence of RuvC and RecG might require DNA synthesis across the entirety of the chromosome and this could potentially perturb successful recombination product formation (Smith, 1989). It is also unclear whether P1 transduction relies on replication fork assembly at recombination intermediates such as D-loop substrates, or whether the presence of the recombination intermediates themselves block replication progression, thereby creating a requirement for replication fork restart and subsequent replication fork assembly (McCool & Sandler, 2001).

### **1.3.2. Hfr crosses**

Another assay to detect defects in homologous recombination relies on DNA recombination occurring during bacterial conjugation (Lederberg & Tatum, 1946), when the DNA from one donor cell is transferred to a recipient cell using cell to cell contact. During conjugation, extrachromosomal DNA from a donor strain undergoes homologous recombination with the host recipient chromosome once it enters the recipient cell. Defects in genes encoding homologous recombination proteins impairs the frequency of recombination between the donor DNA and the recipient thereby leading to a reduction in the number of conjugants. In fact,

genetic studies of bacterial conjugation led to the isolation of the first recombination-deficient mutants (Clark & Margulies, 1965). Since the rate of DNA transfer from the donor to the recipient cell is constant, it was also a very useful tool for generating a linkage map of different genes in *E. coli* (Bachmann, 1990). Similar to P1 transduction, integration of foreign DNA that is transferred via conjugation relies on homologous recombination. Therefore, defects in homologous recombination proteins can be assayed as a measure of the frequency of conjugants obtained (Birge & Low, 1974) (Lloyd, 1991). As with P1 transduction, homologous recombination between incoming foreign DNA and the chromosome occurs through an ends-out orientation which needs to be resolved by either a “cut and paste” mechanism, or through a “copy and paste” mechanism, with the latter presumably requiring DNA synthesis across the entirety of the chromosome thereby complicating successful recombinant product formation.

### **1.3.3. $\lambda$ x $\lambda$ crosses**

Early studies on recombination in *E. coli* also utilized recombination between two duplex DNA molecules of bacteriophage lambda, also termed as lambda by lambda crosses. The use of phage lambda as a tool to study bacterial recombination has been in use since the early 1960s (Kellenberger et al., 1961; Meselson & Weigle, 1961). Using lambda recombination, it was proved that DNA recombination involves breakage and reunion of DNA molecules (Meselson & Weigle, 1961). Normally, bacteriophage lambda possesses genes encoding its own recombination proteins (Red and Gam), but in order to study RecBCD mediated recombination, lambda strains deficient for Red and Gam are used. During phage lambda infection, packaging of viral DNA is initiated from the *cos* site (cohesive end site), wherein the lambda dsDNA is packaged as a linear concatemer with *cos* sites on either ends. If a Red<sup>-</sup> Gam<sup>-</sup> phage lambda strain is used to infect a *E. coli* cell, formation of the linear concatemers takes place by host RecBCD and RecA mediated recombination where the packaging machinery binds to the *cos* site on the left end of the lambda DNA, and RecBCD mediated processing of the other end leads to RecA-dependent crossover with another duplex lambda DNA molecule. Using sensitive methods for labeling DNA (physical and genetic), it is possible to detect such recombination events accurately; wherein the host RecBCD and RecA mediated recombination system is used.

#### **1.4. Objectives of this study**

The objective of this study was to create and validate an ends-in recombination assay for studying RecBCD dependent recombination in *E. coli* which will be useful in dissecting Holliday Junction resolution by allowing processing via the Ruv pathway or via the DNA synthesis pathway. As stated previously, conventional assay systems utilize an “ends out” conformation of recombining DNA substrates, which places a constraint on successful recombinant product formation and recovery. The specific characteristic of the ends-in recombination assay that differentiates it from existing assays (such as ends-out assays) is that it allows us to ask whether successful recombinant product formation was obtained via the RuvC pathway (DNA synthesis independent) or via a replicative (DNA synthesis dependent) route. During ends-out recombination, resolution of Holliday Junctions in cells deficient in RuvC resolvase and RecG helicase would be predicted to occur via a DNA synthesis route that would proceed around the whole of the chromosome. This has the potential to confound analysis of the requirements for successful recombinant product formation during homologous recombination coupled DNA synthesis. The ends-in recombination assay proposed in this study circumvents this problem by allowing the amount of DNA needed to be synthesized following ends-in recombination to be very small (researcher defined – ~200 bp to 2000 bp). Recombinant product formation at D-loop intermediates is likely to require a multitude of replication restart proteins. Utilizing an ends-out assay could complicate analysis of the requirements of these replication restart proteins, since the ends-out DNA substrate might itself block replication fork progression, thereby necessitating replication restart. This pitfall can be eliminated using the ends-in assay system proposed here. We anticipate that the ends-in assay would therefore have a significant advantage over conventional ends-out assays, in that the ends-in assay would facilitate dissection of genetic requirements specific for replicative resolution of recombination intermediates.

# CHAPTER TWO

## MATERIALS AND METHODS

### **2.1 *E. coli* Genetics**

#### **2.1.1 *E. coli* strains and cell culture**

All *E. coli* strains used in this study are isogenic derivatives of *E. coli* MG1655. Depending on the *E. coli* strain, the cells were inoculated at 30°C or 37°C in Luria Broth (LB). LB contains 0.5% w/v yeast extract, 0.5% w/v sodium chloride, 1% tryptone and the final pH was adjusted to 7.5 with 1N NaOH. For making LB plates, 1.5% agar was added to the LB prior to autoclaving. Filter sterilized antibiotic stock solution was added to the medium (50 µg/ml kanamycin, 25 µg/ml chloramphenicol, 10 µg/ml tetracycline and 100 µg/ml ampicillin) if required after autoclaving and cooling.

For short term storage, *E. coli* strains were streaked onto LB plates as required and then stored at 4°C. For long term storage, *E. coli* strains were grown in L-Broth overnight at 37°C (or 30°C for certain strains). 1.0 ml of the culture was added to 0.5 ml sterile 50% glycerol and then stored at -80°C.

For allowing recovery of cells following electroporation, the cells were re-suspended in SOC recovery medium (Fiedler & Wirth, 1988) at 37°C for 1 hour. SOC recovery medium contains 2% Tryptone, 0.5% Yeast Extract, 0.058% NaCl, 0.018% KCl, 0.20% MgCl<sub>2</sub>·6H<sub>2</sub>O, 0.12% MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.36% glucose and the final volume was made up to 1 liter using ddH<sub>2</sub>O.

#### **2.1.2. Plasmid DNA isolation**

##### **2.1.2. 1. Plasmid DNA isolation by miniprep method**

Plasmid isolation was performed according to the miniprep protocol using the EZNA plasmid miniprep kit (Omega Bio-Tek) as per manufacturer's instructions. Briefly, 10 ml LB (supplemented with desired antibiotic) was inoculated with a single colony and then grown overnight for 14 to 16 hours at 37°C (or 30°C for certain strains) and then 1.5 ml of the overnight culture was centrifuged at 13,000 RPM for 5 minutes in an eppendorf tube to harvest cells.

Following centrifugation, the cell pellet was re-suspended in 100 µl re-suspension buffer with RNase I added to it (solution I) and then 200 µl of lysis buffer (solution II) was added to it. The eppendorf tube was then inverted five or six times to facilitate adequate cell lysis following which 350 µl of neutralization buffer (solution III) was added to it and then the eppendorf tube was again inverted five or six times and then centrifuged at 13,000 rpm for 5 minutes. After centrifugation, the supernatant containing plasmid DNA was promptly transferred to a membrane DNA column and this column was placed in a 2 ml collection tube and centrifuged at 13,000 rpm for 30 seconds to allow the supernatant to pass through the membrane column. 500 µl of wash buffer was then added into the membrane column and the column was centrifuged at 13,000 rpm for 3 minutes to dry the column matrix. The column matrix was then placed in a 1.5 ml eppendorf tube and then 50 µl of elution buffer was added to it. The column was allowed to stand at room temperature for 2 minutes following which it was centrifuged at 13,000 rpm for 2 minutes to elute the plasmid DNA into the 1.5 ml centrifuge tube.

#### **2.1.2. 2. Plasmid DNA isolation by midiprep method**

Plasmid isolation was performed according to the midiprep protocol using a Qiagen plasmid midiprep kit as described. 10 ml L-Broth (supplemented with desired antibiotic) was inoculated with a single colony and then grown overnight for 14 to 16 hours at 37°C (or 30°C for certain strains). After 14 to 16 hours, 200 µl of this overnight culture was inoculated into 25 ml LB (supplemented with desired antibiotic) and then incubated for 14 to 16 hours at 37°C in a shaking water bath. After 14 to 16 hours, the 25 ml of the overnight culture was decanted into a 30 ml polypropylene centrifuge tube and then centrifuged at 7,000 RPM for 15 minutes at 4°C in a Beckman Coulter JA centrifuge to harvest cells. Following centrifugation, the cells were re-suspended in 4 ml re-suspension buffer (solution P1) with RNase I and lyseblue buffer added to it and then 4 ml of lysis buffer (solution P2) was added to it. The centrifuge tube was then inverted five or six times to facilitate adequate cell lysis following and allowed to stand at room temperature for 5 minutes. 4 ml of chilled neutralization buffer (solution P3) was added to the centrifuge tube and then the centrifuge tube was again inverted five or six times and then kept on ice for 15 minutes. The centrifuge tube was then centrifuged at 13,000 rpm for 30 minutes. A single Qiagen Tip 100 was placed in a 40 ml waste collection tube and equilibrated with 10 ml of buffer QBT by letting the buffer flow out from the tip due to gravity flow. Immediately after

centrifugation, the supernatant containing plasmid DNA was promptly transferred to the equilibrated Qiagen Tip 100 and the supernatant was allowed to flow out. Thereafter, the Qiagen Tip 100 was washed twice with 10 ml of buffer QC. 5 ml of buffer QF was added to the Qiagen Tip 100 to elute the plasmid DNA. Following elution, 3.5 ml of room temperature isopropanol was added to the eluted DNA solution. Thereafter, the solution containing isopropanol and eluted DNA was aliquoted into 6 eppendorf tubes (approx. 1.4 ml into each eppendorf tube) and the 6 tubes were centrifuged at 13,000 RPM for 30 minutes at room temperature. Following centrifugation, the supernatant was decanted and 600 µl of cold 70% ethanol was added to each eppendorf tube. The tubes were then centrifuged for another 30 minutes at 13,000 RPM at room temperature. The ethanol supernatant was then decanted and the tubes were dried at room temperature for 30 or 40 minutes. The DNA was then resuspended in approx. 20 µl of TE buffer per tube (120 ul TE buffer in total), and samples were pooled.

### **2.1.3. UV sensitivity tests**

For UV sensitivity tests, wild type *E. coli* MG1655 and isogenic mutants were streaked onto a LB plate (using a wild type strain as control) and portions of the plate perpendicular to the streaked cells were variously irradiated with 25, 50, 75 and 100 µJ of UV radiation using a BioRad UV crosslinker apparatus. The plate was then incubated overnight at 37°C (or 30°C for certain strains) and then the ability of the mutants to show visible growth in UV irradiated sections was compared to wild type MG1655 *E. coli* to score relative sensitivity or resistance.

## **2.2 Molecular Biology Techniques**

### **2.2.1. Bacterial culture and storage**

For preparation and storage of bacterial strains, 10 ml LB was inoculated with the desired strain and incubated in a shaking water bath at 37°C (or 30°C for certain strains) for 14 to 16 hours and then 1 ml of this overnight inoculum was mixed with 500 µl 50% glycerol in a sterile cryovial and then this cryovial was frozen at -80°C. When needed, a small amount of this stock solution was streaked onto an LB plate and this plate was incubated overnight at 37°C (or 30°C for certain strains).

### **2.2.2 Preparation of electro-competent cells**

To make electro-competent cells, 10 ml of L-Broth (plus required antibiotic for plasmid containing strains) was inoculated with the desired *E. coli* strain and was grown overnight at 37° C (or 30° C in the case of certain strains) in a shaking water bath for 14 to 16 hours. 5 ml of this culture was subcultured into 1000 ml LB. When the cells reached a mid-logarithmic phase of growth with an approximate O.D. of 0.6 – 0.8, they were pelleted by centrifugation at 6,000 rpm for 15 minutes. The cells were washed with 1 liter chilled distilled water, then pelleted by centrifugation at 6,000 rpm for 15 minutes. The cells were again washed with 500 ml chilled 10% glycerol and then pelleted by centrifugation at 6,000 rpm for 15 minutes. The cells were subsequently washed with 50 ml chilled 10% glycerol and pelleted by centrifugation at 6,000 rpm for 15 minutes. Finally, the cells were suspended in 3 ml final volume of 10% glycerol and aliquoted into 60 µl stocks until needed at -80°C.

### **2.2.3. *E. coli* transformation**

*E. coli* cells were transformed using the standard electroporation method as described by (Taketo, 1988). 60 µl of electro-competent cells was thawed on ice and appropriate amount of plasmid DNA was added to it. The cells were then pipetted into a chilled electroporation cuvette and the cuvette was placed inside a Bio-Rad Gene Pulser. The cells were shocked with 1.70 mV of electric voltage following which the cells were immediately resuspended in SOC recovery medium and then incubated in a shaking water bath at 37°C (or 30°C for certain strains) for 1 hour after which the cells were plated onto LB plates supplemented with antibiotics.

### **2.2.4. Agarose gel electrophoresis and DNA fragment isolation**

For analysis of plasmid DNA, a 0.8% agarose gel in 1X TAE buffer (40 mM Tris-acetate, 2 mM Na<sub>2</sub>EDTA) was used. Gels were run in 1 x TAE buffer and stained in 0.5 µg/ml ethidium bromide (EtBr). DNA was visualized under long-wave UV light.

To extract the desired DNA fragment after restriction digestion from agarose gels, the sample was electrophoresed through 0.8% agarose gel and stained with EtBr. The band of interest was identified using a long-wave UV-illuminator and cut out of the gel. The band was then purified using an EZNA gel purification kit (Omega Bio-Tek) as follows. Initially, the gel slice was suspended in 300 µl of solution I per 100mg of gel in an eppendorf tube and then

incubated in a water bath at 60°C until the gel slice dissolved completely. A DNA binding column was placed in a collection tube and the buffer containing the dissolved gel was applied to the DNA column and then centrifuged at 13,000 RPM for 30 seconds. Thereafter, the eluate was discarded and the DNA column was washed with 500 µl wash buffer containing ethanol by centrifugation at 13,000 RPM for 1 minute. The column was again centrifuged for 2 minutes at 13,000 RPM to dry the column matrix and then the DNA was eluted into an eppendorf tube with 50 µl elution buffer by centrifugation for 2 minutes at 13,000 RPM.

For estimating the concentration of eluted DNA, 2 µl of the eluted DNA was run on a 0.8% agarose gel. The concentration of the eluted DNA was estimated by comparing the intensity of the eluted DNA to the intensity of DNA bands of known concentration in a NEB 1 kb DNA ladder.

#### **2.2.5. PCR amplification**

Genomic DNA amplification was performed by first liberating the genomic DNA from *E. coli* MG1655 by re-suspending a colony of wild type MG1655 *E. coli* in 50 to 100 µl distilled water and then boiling this cell suspension for 5 minutes. For PCR amplification, 2 µl of this suspension was used as template DNA in a 100 µl reaction. Table 2 – 1 lists all the primers used in this study.

All primers were suspended in buffer to a final concentration of 1 µg/ul and a 10x working primer stock was made up with buffer to a final concentration of 0.033 µg/ul.

For PCR reactions that did not require high fidelity, GoTaq PCR mastermix from Promega was used. The PCR mixture consisted of 2 µl of template DNA, 50 µl of Promega GoTaq Green Master Mix, 10 µl of left primer (40 pmol), 10 µl of right primer (40 pmol) and 28 µl of sterile distilled water.

All PCR reactions that did not require high fidelity were set up as follows: Step 1 – 95°C for 2 minutes; Step 2 – 95°C for 30 seconds; Step 3 – 56°C for 30 seconds; Step 4 – 72°C for 2 minutes; repeat steps 2 to 4 thirty times; Step 5 – 72°C for 10 minutes; Step 6 – Hold at 4°C.

For PCR reactions requiring high fidelity, AccuPrime Pfx from Invitrogen was used. High fidelity PCR reactions consisted of: - 2 µl of template DNA (approx. 200 ng); 10 µl of 10x AccuPrime Pfx Reaction Buffer; 15 µl of left primer (10x); 15 µl of right primer (10x); 2 µl of AccuPrime Pfx DNA Polymerase; 56 µl of sterile distilled water.



For high fidelity PCR reactions, the following reaction steps were followed: Step 1 – 95°C for 2 minutes; Step 2 – 95°C for 30 seconds; Step 3 – 56°C for 30 seconds; Step 4 – 68°C for 1 minute per kb; repeat steps 2 to 4 thirty times; Step 5 – 68°C for 10 minutes; Step 6 – Hold at 4°C.

**Table 2 – 1:** List of primers used in this study

<b>Primer Name</b>	<b>Primer Sequence (5' to 3')</b>
<i>att-left</i>	TTGCTCCGGGCTATGAAATA
<i>att-right</i>	ATGTAGCGATGAGTTTCGCC
<i>lacZ-left-in</i>	GAAAGCTTCACCGGCGTTAAAACAAAAG
<i>lacZ-left-out</i>	TCGGATCCCACGGCCACCGATATTATTT
<i>lacZ-right-in</i>	GAGGATCCTATTGGCTTCATCCACCACA
<i>lacZ-right-out</i>	TCCGGCCGTCTGACCAGACACCCATCAA
<i>linear-lacZ-left</i>	TATTGGCTTCATCCACCACA
<i>linear-lacZ-right</i>	CCGATATTATTTGCCCGATG
<i>L4</i>	GCCCTGTAAACGGGGATACT
<i>R5</i>	TTGAAAATGGTCTGCTGCTG
<i>cfa-left-in</i>	AATCGGCCTGACAATTTTTGC
<i>cfa-left-out</i>	GCTCTAGATATCGATATAGATCAGTGTGATTCCG
<i>cfa-right-in</i>	GCGATATCACCGACCAGTGATGGAG
<i>cfa-right-out</i>	GACCCGGGCCTGCCAGAGCTGAATATCAC
<i>cfa-right-out-2</i>	GCATCCCGGGGTAGCCTTTAGCGAGCCACT
<i>cfa-left-2chi</i>	GGTGGTCGGGTGGTTCGATCGATATAGATCAGTGTGAT TCG
<i>cfa-right-2chi</i>	CCACCAGCCCACCAGCATCACCGACCAGTGATGGAG
<i>L1</i>	AATACGCAAACCGCCTCTC
<i>R1</i>	TGGGACAACCTCCAGTGAAAA
<i>sacB-EcoRV-left</i>	TCGAGATATCCTGCAGCCAAGCTAGACCTA
<i>sacB-EcoRV-right</i>	TCGAGATATCCTGGGACGGAAGTCGCTGTC
<i>gfp-left-BsmI</i>	ATGCAGGAATGCCTAACACACAGGAAACAGCTATGA
<i>gfp-right</i>	TCAGAGGTTTTACCGTCATC

### **2.2.6. Linear DNA electroporation for qualitative analysis**

For a qualitative analysis of each ends-in assay plasmid construct, linear DNA was electroporated into recipient electrocompetent cells and plasmid DNA isolated from transformants was screened for the correct recombinant clone.

### **2.2.7. Linear DNA electroporation for quantitative analysis**

For quantitative analysis of transformation efficiency among batches of electrocompetent *E. coli* strains (i.e. among strains mutated in key recombination genes), linear DNA was first mixed with a defined amount of covalently closed circular DNA to prepare a master mix of DNA. Then, this DNA master mix was electroporated into recipient electrocompetent cells and plasmid DNA isolated from transformants was screened for the correct recombinant clone.

### **2.2.8. DNA sequencing**

All DNA sequencing was carried out at the Plant Biotechnology Institute – National Research Council of Canada using an ABI 3730x DNA sequencer.

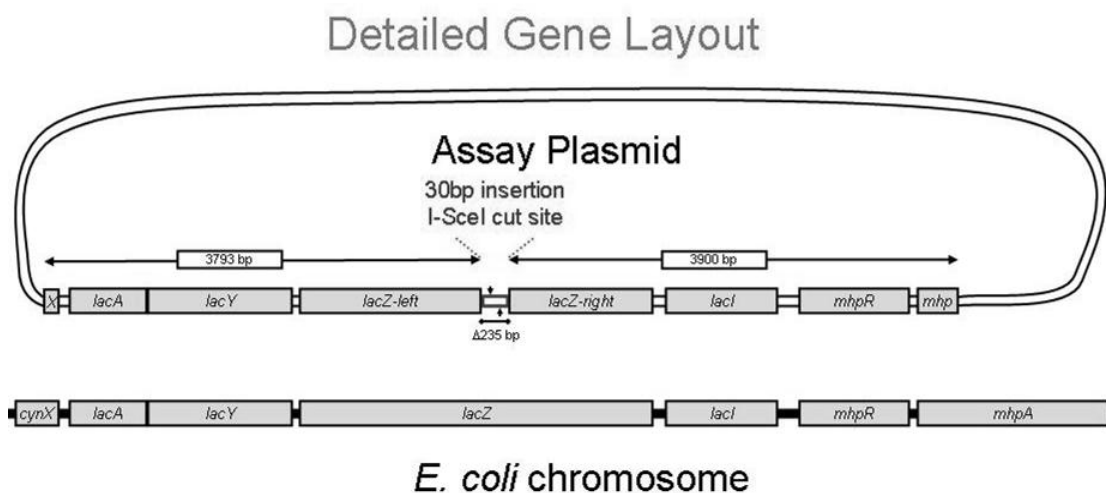
# CHAPTER THREE

## RESULTS

### 3.1. In-vivo cleavage system using I-SceI linearization

#### 3.1.1. Verification of *in-vivo* cleavage system plasmid

The plasmid used for the *in vivo* linearization recombination assay was a modified (*rop*<sup>+</sup>) pBR322 constructed by a previous student Sarah Helgeson. This plasmid possessed ~4,000 bp of homology to the *lacZ* region on either side of an I-SceI recognition site, as illustrated in figure 3 – 1. The orientation of the two regions was such that during strand invasion mediated by *E. coli*'s homologous recombination proteins, the two free ends flanking the double strand break would be facing each other (similar to the greek letter “omega”, but with the “arms” pointing inwards), as shown in figure 3-1.



**Figure 3-1:** Diagrammatic representation of the detailed gene layout for the *in vivo* I-SceI linearization assay plasmid system. (Genes drawn to scale)

The *in-vivo* cleavage assay required an assay strain that carries an inducible I-SceI gene to perform *in vivo* cleavage. A test strain was constructed in the following manner: - a *leu-63::Tn10* allele was moved into wild type MG1655 by P1 transduction, selecting for transductants on LB plates containing tetracycline. Thereafter, a  $\Delta araBAD$  mutation was moved

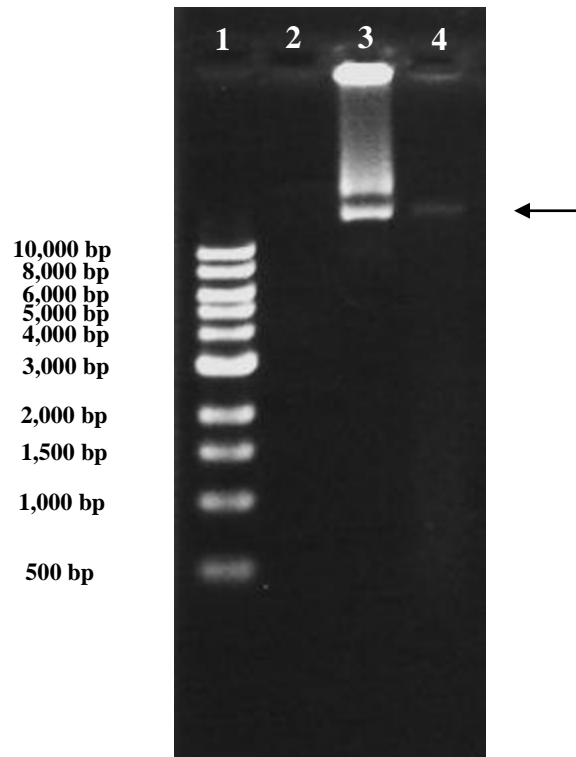
into this *leu<sup>-</sup>* strain by P1 transduction selecting for *leu<sup>+</sup>* transductants on minimal plates containing glucose. The transductants were separately screened for sensitivity to tetracycline, while also screening on minimal media containing arabinose for those transductants that had lost the ability to grow on arabinose i.e. had become  $\Delta araBAD$ . Those transductants that had lost the ability to grow on arabinose and which were also sensitive to tetracycline were then used for further steps (approximately 50% of *leu<sup>+</sup>* transformants were  $\Delta araBAD$ ). An  $\Delta araBAD567$   $\Delta attB::P_{BAD}I$ -SceI gene cassette carrying an arabinose inducible I-SceI expression system (Gumbiner-Russo et al., 2001) was then moved into this strain using a *malE::Tn10* linked marker by P1 transduction. Tetracycline resistant transductants were isolated and screened for the  $\Delta attB::P_{BAD}I$ -SceI gene cassette by a PCR based screening using *att* specific primers *att-left* and *att-right* (3 out of 12 transductants tested were found to contain the  $\Delta attB::P_{BAD}I$ -SceI cassette). A *recA::dTn10* (CAM) (Kleckner et al., 1991) mutation was moved into this strain via P1 transduction and transductants were selected on LB plates containing chloramphenicol.

### **3.1.2. Testing of the *in vivo* cleavage system**

Initial attempts to transform the assay plasmid into strains containing the  $\Delta attB::P_{BAD}I$ -SceI gene cassette using a  $CaCl_2$  protocol for generating competent cells were unsuccessful. Control transformations done using pBR322 were successful, giving an indication that the size of the assay plasmid (about 11.5 kb) may necessitate the use of electroporation for transformation. Using electroporation, transformants were obtained for the assay plasmid (2 transformants from an electroporation with 2  $\mu$ l of standard plasmid miniprep). *recA::dTn10*(CAM) was then transduced into a test strain containing the  $\Delta attB::P_{BAD}I$ -SceI gene cassette along with the assay plasmid. As a preliminary test of I-SceI mediated plasmid linearization, plasmid survival was measured in *recA* cells that were induced for I-SceI expression by growing on LB + ampicillin + arabinose plates using LB + ampicillin + glucose plates as a parallel control. It was presumed that in cells lacking RecA, once linearized, the assay plasmid would not be able to recircularize via homologous recombination with the chromosomal target; hence those cells should show a reduced viability on arabinose + ampicillin media since plasmid-free cells will be unable to form colonies in the presence of ampicillin. However, there was equal growth in both arabinose (plus ampicillin) and glucose (plus ampicillin) plates, implying that either the I-SceI was not being expressed, or that the *recA* cells were somehow remaining recombination proficient, or that the

levels of I-SceI were simply not sufficient to linearize all the plasmid molecules in each cell of the assay strain.

In order to verify the absence of RecA, the cells used for the above experiment were subjected to a UV sensitivity test. The results showed that the cells were UV sensitive and hence indeed *recA*<sup>-</sup>. Furthermore, plasmid preps were performed on RecA-deficient cells grown in the presence of ampicillin with either glucose or arabinose to determine if there was any reduction of plasmid quantity due to I-SceI linearization. A 0.8% agarose gel run with plasmid samples (Figure 3-2) from both the cultures demonstrated that there was a significant reduction in the yield of plasmid DNA in cells grown in arabinose (lane 4). However, covalently closed circular DNA was detectable in the plasmid preparation of *recA*<sup>-</sup> assay plasmid containing cells after growth on LB + ampicillin + arabinose (lane 4). Note that the plasmid DNA in lanes 3 and 4 was apparently migrating higher than the expected 11.5 kb position, likely indicating the plasmid DNA was in the form of a dimer. This result indicated that the induced I-SceI expression system was not sufficiently robust to linearize all the plasmids present in the cell, thus explaining the observed insensitivity of the *recA*<sup>-</sup> plasmid containing cells to growth on LB + ampicillin + arabinose plates. Together, the plate count and gel electrophoresis results showed that there was likely expression of I-SceI, but apparently not enough to linearize ALL the plasmid molecules in each cell, which would explain a reduction in plasmid copy number in cells grown in the presence of arabinose. Unfortunately this background level of uncut plasmid molecules would confound analysis of potential recombinants generated following I-SceI induction.



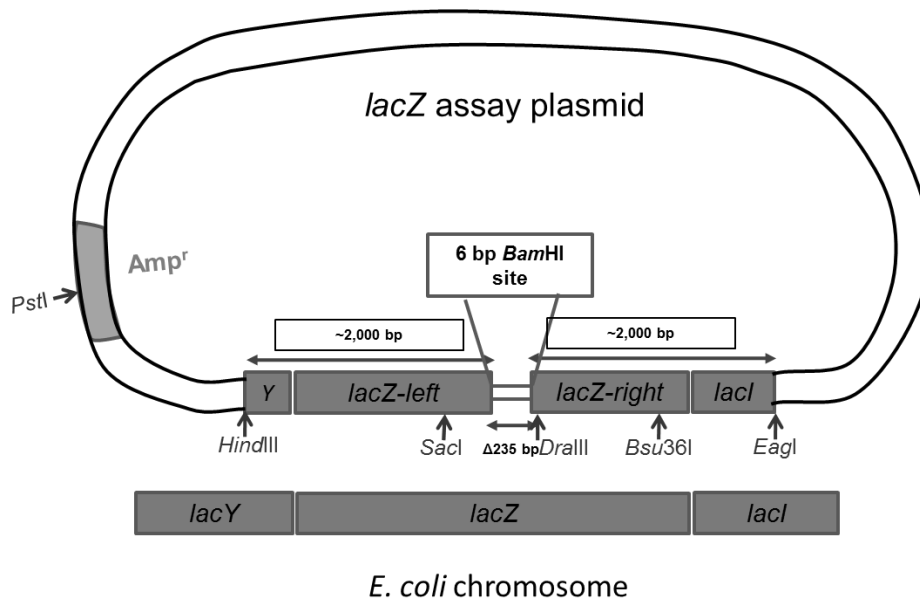
**Figure 3 – 2:** Measurement of plasmid survival in cells grown in medium containing ampicillin supplemented with glucose or arabinose. Lane 1 – 1 kb DNA ladder; Lane 2 – blank; Lane 3 – Miniprep of cells grown overnight in glucose supplemented medium; Lane 4 – Miniprep of cells grown overnight in arabinose supplemented medium.

It was evident from these tests that an *in vivo* assay relying on I-SceI expression would not work, since the amount of I-SceI expression was not enough to cleave all the copies of the assay plasmid in the cell. Since analysis of recombinational repaired plasmids in this system requires purification of plasmids from I-SceI induced cells followed by transformation into a *lacZ* reporter strain to score efficiency of rescuing the chromosomal *lacZ*<sup>+</sup> region, it was concluded that the level of non-cleaved plasmid would be too high for efficient screening of potential recombinants. In addition, the size of the assay plasmid (about 11kb) made it difficult to electroporate into *E. coli* recipient cells, thus it was decided to modify the system to work around these pitfalls.

## 3.2. In-vitro cleavage using lacZ plasmid system

### 3.2.1. Construction of lacZ assay system

To address the shortcomings in the previous assay plasmid, a new assay plasmid was designed to study ends-in double strand break repair, as illustrated in figure 3 – 3. The main drawback of using the *in vivo* linearized assay plasmid was the incomplete linearization of the plasmid by I-SceI expression. Hence, the *lacZ* assay plasmid design was modified to be linearized *in vitro*. In addition, the total plasmid size was reduced by shortening the regions of homology from ~4 kb on either side to ~2 kb on either side of the linearization site. *In vitro* linearization was designed by inclusion of a unique *Bam*HI restriction site between the regions of homology to the *lacZ* region.



**Figure 3 – 3:** Diagrammatic illustration of the *lacZ* plasmid layout with the cloned regions of homology flanking the *Bam*HI site including restriction sites that were used. (Genes are drawn to scale)

The construction of the *lacZ* assay plasmid was a three step process. Initially, the left fragment of homology, termed as the *lacZ*-left fragment, was PCR amplified (using primers *lacZ*-left-in and *lacZ*-left-out containing a *Hind*III and *Bam*HI site, respectively) and digested with *Hind*III and *Bam*HI. Plasmid pBR322 was digested with *Hind*III and *Bam*HI and the desired 4015 bp fragment was gel purified. The recovered fragments were then run on a 0.8% agarose

gel to determine their concentration, following which they were ligated together, transformed into competent *E. coli* cells and the resulting intermediate plasmid was purified. The plasmid DNA was sequenced and a clone that had the correct sequence of *lacZ*-left was selected for subsequent cloning.

The *lacZ*-right fragment was PCR amplified (using primers *lacZ*-right-in and *lacZ*-right-out containing a *Bam*HI and an *Eag*I site, respectively). Initial attempts to clone the *lacZ*-right fragment to the assay plasmid pBR322 containing the *lacZ*-left fragment proved unsuccessful. It was suspected this was likely due to the fact that there was just a 2 bp extension outside of the restriction sites on the primers and these 2 bp overhangs were not enough to enable complete digestion of the PCR fragment by *Bam*HI or *Eag*I, which was surprising given NEB's claim of >90% cutting efficiencies with as little as two base pairs of flanking DNA. It was possible that the flanking sequence context influences this efficiency. In an attempt to work around this roadblock, the undigested right fragment was ligated into the pGEM-T easy TA cloning vector (Holton & Graham, 1991) which allows PCR products with A-overhangs to be cloned into a ddT-tailed vector. This *lacZ*-right fragment was then excised from the pGEM-T easy vector using *Bam*HI and *Eag*I restriction enzyme digestion, gel purified and then ligated into gel purified *Bam*HI-*Eag*I digested pBR322 with the *lacZ*-left fragment already cloned into it. In order to ensure that the final plasmid had the correct left and right fragments, plasmid clones were sequenced using primers outside the cloned left and right regions of homology and selected a clone that had the correct sequence of both *lacZ*-left and *lacZ*-right fragments.

It is important to note that the plasmid construct described above contained a native, properly oriented *chi* site on the left arm of homology to the *lacZ* gene, although at this stage it was not known if this might influence plasmid re-circularization.

### **3.2.2. Testing of the *lacZ* plasmid system**

In order to test the *lacZ* plasmid assay, the assay plasmid was first linearized with *Bam*HI and then mixed with a pre-determined concentration of pACYC184 circular plasmid DNA as a transformation control. This mix of linear and circular DNA was electroporated into competent cells for wild type, *recA*<sup>-</sup> and *recD*<sup>-</sup> genotypes. A low background level of transformants were observed in a *recA*<sup>-</sup> strain, and it was hard to eliminate this background level of contamination even in spite of using gel purified linear DNA.



It was initially observed that the *Bam*HI site would be non-homologous to the target DNA and that might lower the frequency of homologous recombination dependent marker rescue, hence, it was decided that the plasmids would be linearized with *Sac*I and *Bsu*36I (which have recognition sites in the left and right regions of homology respectively). This would exclude the *Bam*HI site and would provide a completely homologous substrate for ends-in recombination. Consequently, the assay plasmid was digested first with *Sac*I, following which the reaction mixture was purified using a Qiagen PCR cleanup kit. The column purified *Sac*I-digested DNA was then digested with *Bsu*36I (as illustrated in Figure 3 – 3). The double digested plasmid DNA was then gel purified out of a 0.8% agarose gel using the gel purification spin column and then eluted in buffer. The concentration of this DNA was determined by running it on a 0.8% agarose gel with standards of known concentration and this DNA was used for electroporation of wild type and *recA*<sup>-</sup> strains.

Upon electroporation, the *recA*<sup>-</sup> negative control strain yielded a few transformants, which might have arose due to a small percentage of contaminating plasmid molecules that evaded linearization by *Sac*I and *Bsu*36I.

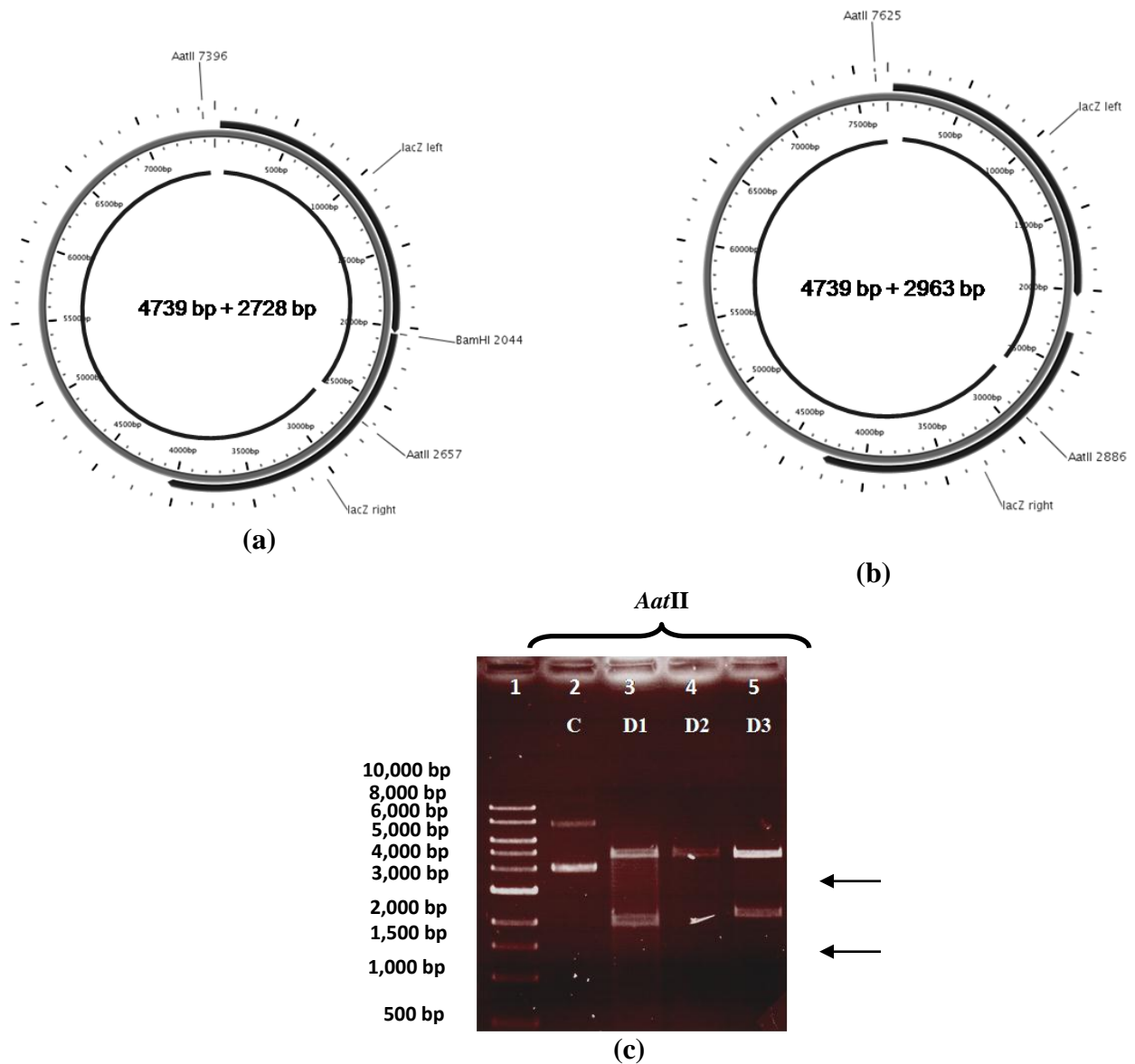
The plasmids recovered from these ampicillin resistant transformants were analyzed to find out if they had recombined with the chromosomal target as predicted. These plasmids were digested separately with *Bam*HI and *Pst*I to determine if they had lost the *Bam*HI site (which they shouldn't have if they recombined with the chromosomal target site). *Pst*I was chosen because this recognition site was within the beta lactamase gene encoding ampicillin resistance and linearization would allow the size of the recombinant plasmids to be estimated. The recovered plasmids were found to retain an intact *Bam*HI site, indicating that the recovered plasmids were not recombinants but were uncut contaminating parental plasmid DNA. Subsequent rounds of electroporation using *Sac*I and *Dra*III digested linear plasmid DNA revealed that most of the recovered plasmid DNA had an intact *Bam*HI site, suggesting that the putative linear DNA was being contaminated by circular undigested DNA.

As a possible alternative to restriction digestion for generating linear DNA molecules for electroporation, PCR primers were designed (forward primer *linear-lacZ-left* and reverse primer *linear-lacZ-right*) flanking the *Bam*HI site which would amplify the whole *lacZ* assay plasmid as a linear molecule. Linear *lacZ* plasmid was PCR amplified using *Bam*HI digested template *lacZ* plasmid as template. This linear *lacZ* assay plasmid was gel purified from a 0.8% agarose gel and

used for subsequent electroporation reactions. Wild type MG1655 *E. coli* competent cells were then electroporated with the following substrates – (a) *Bam*HI digested, gel purified *lacZ* plasmid and (b) PCR generated spin column purified linear *lacZ* plasmid.

In spite of using the linear substrates obtained by *Bam*HI digestion and PCR, background transformants were observed in a *recA*<sup>-</sup> negative control strain, and the number of transformants obtained for PCR generated linear molecules was four to five fold higher compared to those obtained when using *Bam*HI digested linear plasmids. Also, restriction analysis showed that most of those transformants had an intact *Bam*HI site which led me to believe that most of these transformants had been transformed with circular covalently closed DNA used as a template for PCR along with the linear DNA.

The *lacZ* assay plasmid was digested using *Sac*I and *Dra*III restriction enzymes, following which the linear DNA was purified using a spin column purification column and then this linear DNA was used for electroporation. Post electroporation, plasmid DNA from three representative *recD*<sup>-</sup> transformants were analyzed using *Aat*II restriction digestion to find out if the plasmids have recombined or not. Figure 3 – 4 (a) and 3 – 4 (b) illustrate the native assay plasmid and the recombined assay plasmid respectively with the restriction sites for *Aat*II marked. *Aat*II cuts once in the right region of homology and once in the plasmid backbone upstream of the *Hind*III site. The restriction pattern with *Aat*II (Figure 3 – 4 (c)) show that all the samples showed a common larger band around ~4,700 bp and a truncated smaller band of ~2,000 bp (in lanes 3 and 5), suggesting that all the three samples had lost some region of homology on either side of the *Bam*HI site. For sample D2 in lane 4, no smaller band was seen, suggesting extensive resection between the two restriction sites, maybe even to the extent of eliminating one restriction site altogether.



**Figure 3 – 4:** Analysis of plasmid DNA extracted from *recD* *lacZ* assay plasmid transformants. (a) and (b) Illustrations showing *AatII* restriction sites for non-recombined and recombined plasmids respectively. (c) Restriction digest of plasmid DNA extracted from *recD* *lacZ* assay plasmid transformants using *AatII*. Lane 1 – 1 kb DNA ladder; Lane 2 – undigested sample 2; Lane 3 – Sample 1 digested with *AatII*; Lane 4 – Sample 2 digested with *AatII*; Lane 5 – Sample 3 digested with *AatII*. (Expected band size for recombined plasmid – 4739 bp and 2963 bp; Expected band size for non-recombined plasmid – 4739 bp and 2728 bp).

Although the double digested *lacZ* plasmid seemed to be generating transformants, the plasmid backbone was a pBR322 plasmid backbone with a medium-low copy number. This made plasmid purification difficult, and generating enough plasmid DNA to linearize and

transform was laborious, especially since linear DNA transformation requires comparatively much more DNA than circular DNA (Conley & Saunders, 1984). Also, this assay relied on isolating plasmid DNA from recombinant transformants and then screening for *lacZ*<sup>+</sup> plasmid transformants from amongst the transformants obtained, which was itself a cumbersome process when done on a large scale. Hence, a novel plasmid substrate was designed that would enable a one-step verification of recombination events.

### **3.3. Promoter capture (GFP) plasmid system**

#### **3.3.1. Construction of the promoter capture plasmid system**

To address some of the shortcomings of the previous assay plasmid such as low copy number and to ease the screening process, a novel “promoter capture” assay plasmid was constructed. This assay plasmid was designed such that the gene encoding the green fluorescent protein (*gfp*) would be placed under the control of the cyclopropane fatty acid synthase (*cfa*) promoter in successful recombinants. The *cfa* promoter flanking regions were cloned as two separate fragments separated by an *EcoRV* site as shown in figure 3 – 5 (a). Since these two *cfa* fragments lack a central 254 bp piece of DNA encoding  $\sigma^s$  and  $\sigma^{70}$  promoters, no *gfp* expression was observed. This assay was designed as a one-step assay that would report successful recombinants as GFP<sup>+</sup> cells generated by marker rescue of the *cfa* promoter region from the chromosome. The *cfa* promoter was chosen for this assay because it contained a properly oriented *chi* site which was anticipated to facilitate RecBCD mediated homologous recombination of the plasmid substrate.

The construction of the promoter capture assay plasmid was carried out in the following manner. First, the CFA-left fragment was PCR amplified using primers *cfa-left-in* and *cfa-left-out*. The *cfa-left-out* primer had an inner *EcoRV* site flanked by an outer *XbaI* site. This 1,008 bp DNA fragment was digested with *HincII* (present downstream of *cfa-left-in*) and *XbaI* respectively, and the resulting 845 bp fragment was gel purified. Plasmid pGFPuv was digested with *PvuII* (compatible with *HincII*) and *XbaI* and gel purified. The concentration of the fragments was estimated by staining intensity relative to a known concentration of molecular



**Figure 3 – 5:** Construction of the *gfp* plasmid system (a) Diagrammatic illustration of the primary pGFP assay plasmid constructed with regions of homology flanking a missing 254 bp *cfa* promoter. (b) Gel showing restriction digestion of transformants to screen for the correct clone having *cfa*-left fragment inserted in the pGFPuv backbone using 1) *Nco*I and *Pst*I digestion and 2) *Nco*I and *Xcm*I digestion. Lane 1 – 1 kb DNA ladder; Lane 2 to 6 – Samples 1 to 5 digested with *Nco*I and *Pst*I; Lane 7 to 11 – Samples 1 to 5 digested with *Nco*I and *Xcm*I. (Expected band sizes for *Nco*I and *Pst*I digested correct clone – 3342 bp and 633 bp; Expected band sizes for *Nco*I and *Xcm*I digested correct clone – 2962 bp and 1013 bp). (c) Gel showing restriction digestion of transformants to screen for the correct clone having *cfa*-right fragment inserted in the pGFPuv backbone using *Eco*RV and *Dra*III digestion. Lane 1 – 1 kb DNA ladder, Lane 1 to 9 – samples 1 to 9 digested with *Eco*RV and *Dra*III, Lane 10 – Undigested sample 1 as control. (Expected band sizes for *Eco*RV and *Dra*III digested correct clone – 4205 bp and 940 bp).

Transformants samples S1, S4 and S5 with the correct restriction pattern (lanes 2, 5, 6, 7, 10, and 11 in figure 3 – 5 (b)) were sequenced across the left region of homology. A clone with the correct DNA sequence was digested with *Eco*RV and *Xma*I and the resultant double digested plasmid was gel purified. The 1,140 bp CFA-right fragment was PCR amplified with *cfa-right-in* and *cfa-right-out* primers. The resulting PCR generated fragment was digested with *Eco*RV and *Xma*I respectively and spin-column purified. The digested vector and insert DNA were run on a 0.8% agarose gel with molecular weight standards of known intensity to verify their concentration and subsequently ligated together.

Following ligation and transformation, plasmid DNA was isolated from transformants and screened for the presence of the CFA-right fragment by restriction digestion using *Eco*RV and *Dra*III enzymes, as shown in figure 3 – 5 (c). Those transformants that had the correct restriction pattern of 4205 bp and 940 bp (all the samples from S1 to S9 in lanes 1 to 9 turned out to be the correct ones) were sequenced across the right region of homology and a clone with the correct DNA sequence was used for further experiments.

The promoter capture plasmid was constructed such that ligation of the CFA-left and CFA-right fragments re-created an *Eco*RV site between the two fragments, with half the *Eco*RV site on the left fragment and half of the *Eco*RV site on the right fragment. This also ensured that after *Eco*RV digestion, the fragments would be totally homologous to the chromosomal target site.

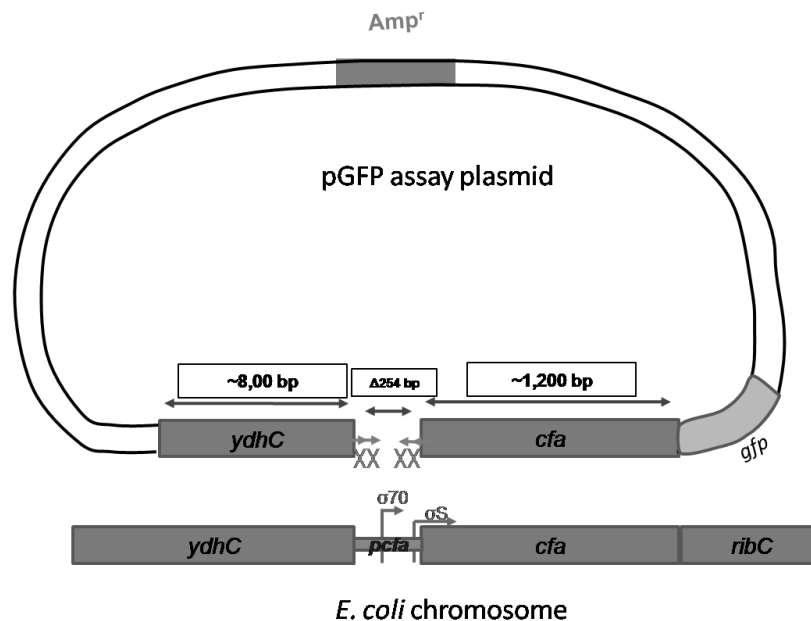
### **3.3.2. Testing of the promoter capture assay**

The *gfp* plasmid was linearized by restriction digestion with *EcoRV* between the left and right regions of homology, and then this linear plasmid was electroporated into wild type, *recA*<sup>-</sup>, *recD*<sup>-</sup>, *recA*<sup>-</sup> *recD*<sup>-</sup> and *ruvC*<sup>-</sup> competent cells of *E. coli*. This electroporation was carried out as a proof-of-concept experiment rather than to generate quantifiable data, hence it was done without a control circular plasmid. The primary goal was to determine if the assay plasmid would express fluorescence following re-circularization. Unfortunately, the results of the electroporation showed that none of the transformants in any of the strains gave rise to glowing colonies. The number of transformants arising in *recA*<sup>-</sup> MG1655 was about 5% (11 colonies) of the number of transformants arising in wild type MG1655 strain (247 colonies). The number of colonies in a *recD*<sup>-</sup> MG1655 strain was higher than wild type (335 colonies), and the *recA*<sup>-</sup> *recD*<sup>-</sup> and *ruvC*<sup>-</sup> strains showed 6 and 21 colonies respectively. Although the number of transformants does not necessarily reflect the level of competence of each strain, the fact that none of the transformants expressed any fluorescence suggested that the assay system would need to be modified to enable expression of *gfp*. Since the primary goal of designing the new assay plasmid was to facilitate screening of transformants to identify those with the correct recombined assay plasmid, these transformants were not analyzed. Instead, more emphasis was placed on modifying the assay plasmid in the next sub-section of this thesis.

### **3.3.3. Modification to the promoter capture plasmid system – addition of tandem *chi* sites**

In order to determine the reason for the absence of *gfp* expression in the transformants, the plasmid sequence was subjected to an ORF Finder analysis at NCBI's website. The analysis revealed that the gene encoding *cfa* downstream of the double stranded break was lacking a termination codon that resulted in the potential production of a translational fusion of *gfp* to the *cfa* gene. Hence, a new reverse primer was designed to amplify the right fragment of homology (*cfa-right-out-2*) with an *XmaI* site. The right fragment of homology was then amplified using the primers *cfa-right-in* and *cfa-right-out-2* such that it now included the termination codon. The plasmid containing the *cfa* left fragment of homology was then digested with *XbaI* and *XmaI* and then the *XbaI* + *XmaI* digested *cfa* right fragment with the termination codon was ligated to it. Following ligation, the right fragment that was inserted between the *XbaI* and *XmaI* sites was sequenced to make sure that there were no base pair changes in the DNA sequence.

Moreover, since previous experience with the *lacZ* assay plasmid had revealed that one *chi* site was not wholly sufficient for efficient plasmid re-circularization, it was suspected that the single *chi* site present on the left *cfa* region of homology on the *gfp* plasmid would not be adequate for plasmid re-circularization. Hence, to add *chi* sites to the ends of the CFA-left and CFA-right regions of homology, primers specific to the two CFA regions of homology were constructed with added tandem *chi* sites on both ends (*cfa-left-2chi* and *cfa-right-2chi*; these primers were cartridge purified for high fidelity and accuracy). PCR amplification of the promoter capture plasmid substrate by Pfx High Fidelity Polymerase (Invitrogen) using the primers with tandem *chi* sites produced a linear dsDNA molecule with tandem *chi* sites at the ends of the regions of homology, as illustrated in figure 3 – 6.



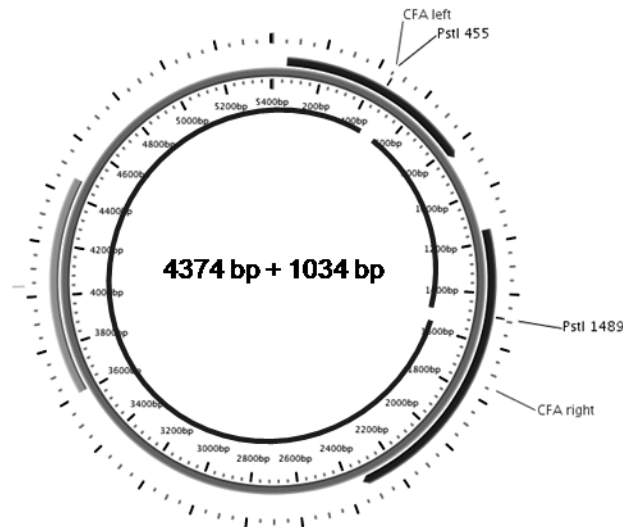
**Figure 3 – 6:** Diagrammatic representation of the linear promoter capture plasmid system with tandem *chi* sites at the ends of the regions of homology

### **3.3.4. Testing of the modified promoter capture plasmid system with tandem *chi* sites**

Once linear plasmid molecules with tandem *chi* sites on either ends were amplified, this linear molecule was electroporated into different MG1655 recipient strains. Following electroporation, 7 transformants were obtained in a *recD*<sup>-</sup> strain; one transformant was obtained in a *recA*<sup>-</sup> strain and one transformant in the wild type MG1655 strain. After extracting plasmid DNA from each of these nine transformants, *Pst*I restriction digestion was used to analyze the extracted plasmid DNA. *Pst*I cuts once on each arm of homology, and the correct plasmid clone



was expected to generate two fragments, with the bigger fragment around 4374 bp and the smaller one being around 1034 bp (as illustrated in figure 3 – 7).

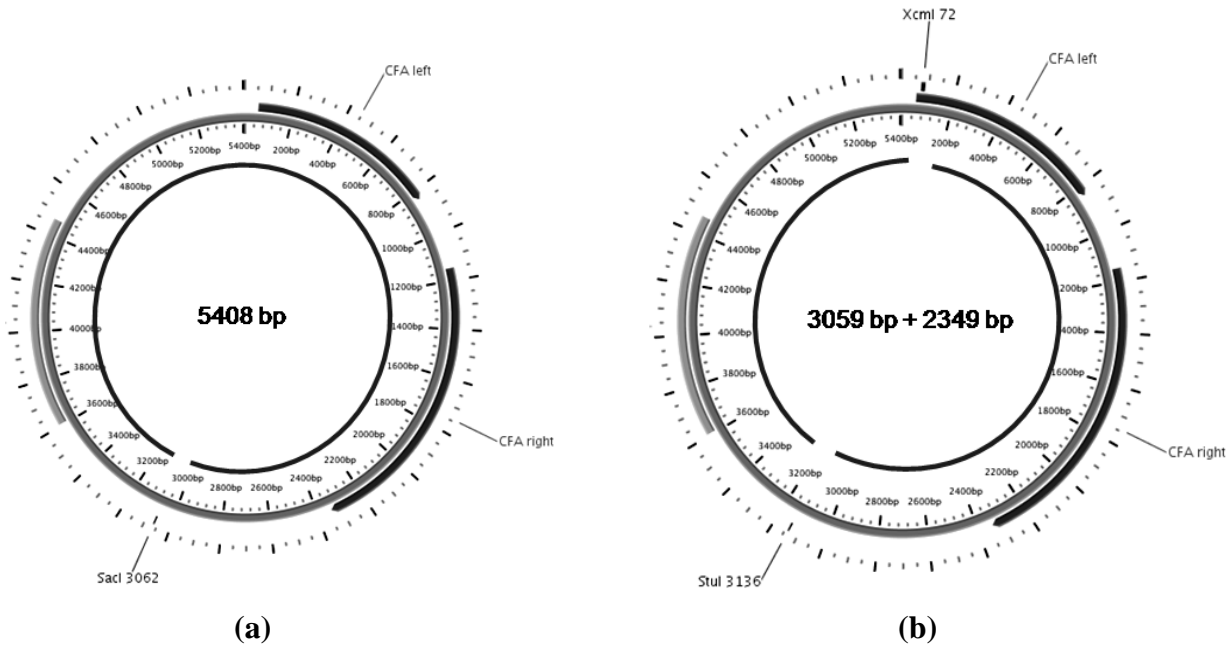


**Figure 3 – 7:** Restriction map showing *PstI* restriction sites on the recombinant pGFP assay plasmid.

Restriction analysis revealed that none of the nine transformants analyzed generated the expected band pattern. The band sizes seen for all the samples corresponded to neither the starting assay plasmid DNA nor the expected recombinant plasmid DNA. It was quite likely that all the samples had lost both the *PstI* cut sites and therefore the bands seen represented relaxed and supercoiled forms of plasmid multimers and monomers.

In an effort to further characterize the plasmid DNA extracted from these transformants, five representative plasmid DNA samples (one from wild type, one from *recA*<sup>-</sup> and three from *recD*<sup>-</sup> (samples 1, 2 and 3)) were subjected to the following restriction digestion a) *SacI* and b) *XcmI* and *StuI*. *SacI* has a recognition site downstream of the right arm of homology (figure 3 – 8 (a)) and since it cut the plasmid backbone only once, it was chosen to give an indication about the size of the plasmid. *XcmI*, on the other hand, had a recognition site near the 5' end of the left region of homology, whereas *StuI* had one recognition site in the center of the right region of homology and the other recognition site was located 59 bp downstream of the *gfp* termination codon (figure 3 – 8 (b)). The *StuI* site in the middle of the right region of homology was blocked by Dcm methylation, and hence it cut the plasmid only once (downstream of the *gfp* gene).

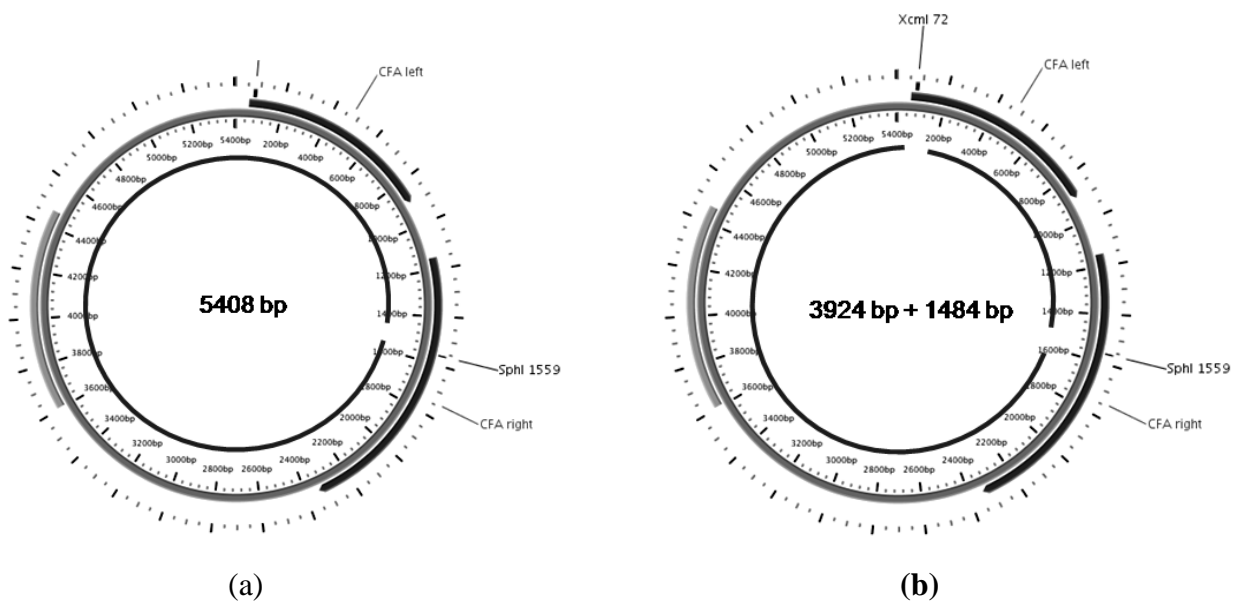
The single digests with *SacI* revealed that most of these plasmids had been reduced in size, with the greatest reduction being evident in wild type as well as the third *recD*<sup>-</sup> sample. The plasmid recovered from the *recA*<sup>-</sup> transformant showed a band pattern similar to the expected band pattern for the non-recombined assay plasmid. Curiously, one of the plasmid samples recombined from the *recD*<sup>-</sup> showed a band pattern that suggested that the size of the plasmid had increased and was now much larger than the expected plasmid size.



**Figure 3 – 8:** (a) Restriction map showing the *SacI* restriction site on the recombined pGFP assay plasmid. (b) Restriction map showing the *XcmI* and *StuI* restriction sites on the recombined pGFP assay plasmid.

Since it was obvious that both wild type and *recD*<sup>-</sup> sample 3 had lost a major portion of DNA and were not even as big as the starting plasmid, another set of restriction digestion was performed on the remaining transformants to find out more about these plasmids. These three plasmid samples (one from the *recA*<sup>-</sup> strain and two from the *recD*<sup>-</sup> strain (D1 and D2)) were digested with a) *SphI* and b) *XcmI* and *SphI*. The recognition site of *SphI* was present on the right fragment of homology, about 410 bp downstream of the ds break site. The *XcmI* site, on the other hand, was present on the left fragment of homology, about 20 bp from its 5' end. The restriction sites are diagrammatically elucidated in figure 3 – 9 (a) and (b). A diagnostic gel was run with these digested samples, revealing that the size of the plasmid DNA obtained for the *recA*<sup>-</sup> sample

was equal to the starting plasmid. Hence it was hypothesized that this transformant arose by transformation of an undigested circular DNA molecule. For sample 1 isolated from the *recD*<sup>-</sup> strain, there was a small but noticeable reduction in plasmid size, as evident from the restriction digests. For sample 2 isolated from the *recD*<sup>-</sup> strain, surprisingly, the plasmid size was shown to have increased even more than expected, and was close to 7 kb. Also, these three plasmid samples had retained most of the plasmid backbone starting from the *XcmI* site to the *SphI* site, since they all generated a larger ~3,900 bp band seen for *XcmI* + *SphI* double digested samples. The gel analysis strongly suggests that for these samples, most of the re-arrangement was at the site of the dsDNA break.



**Figure 3 – 9:** (a) Restriction map showing the *XcmI* restriction site on the recombinant pGFP assay plasmid. (b) Restriction map showing the *XcmI* and *SphI* restriction site on the recombinant pGFP assay plasmid.

In order to get further information about the fate of the linear plasmid molecules following electroporation, wild type, *recA*<sup>-</sup>, *recD*<sup>-</sup> and *recA*<sup>-</sup> *recD*<sup>-</sup> strains of MG1655 *E. coli* were electroporated with linear *gfp* assay plasmid that had been amplified using primers containing tandem *chi* sites (*cfa-left-2chi* and *cfa-right-2chi*). This time, however, a higher concentration of plasmid DNA was used in order to obtain a higher yield of transformants. The results of electroporation are shown in Table 3 – 1:

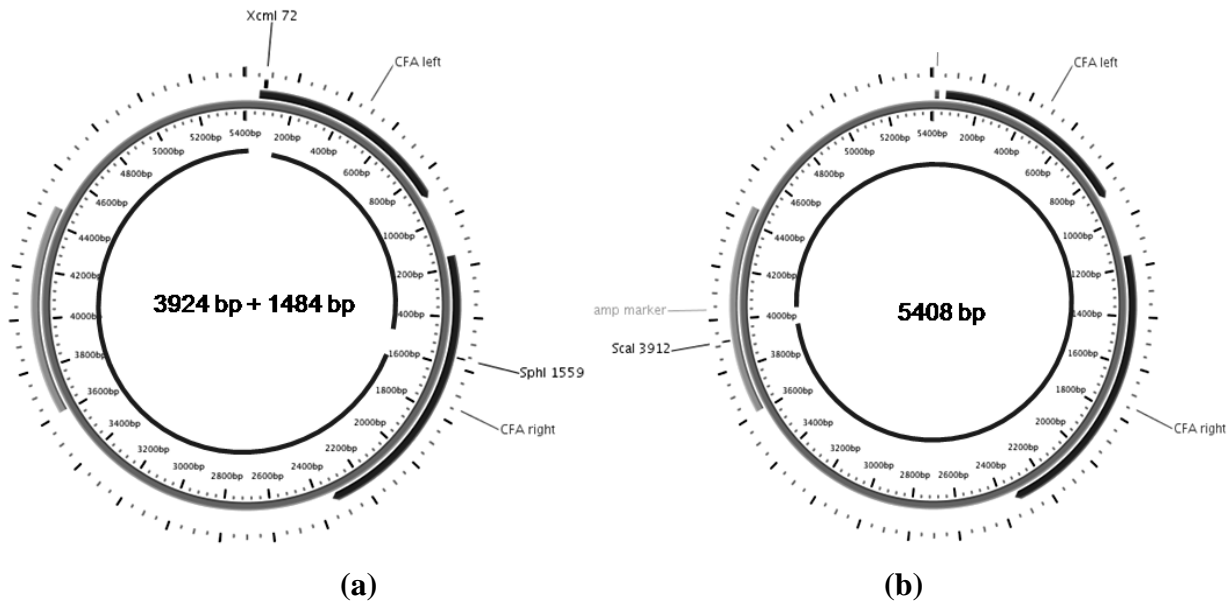
**Table 3 – 1:** Number of transformants obtained for each relevant strain genotype using linear *gfp* assay plasmid for transformation

Strain	Number of Transformants
Wild Type	2
<i>recA</i> <sup>-</sup>	8
<i>recA</i> <sup>-</sup> <i>recD</i> <sup>-</sup>	2
<i>recD</i> <sup>-</sup>	18

The above 30 transformants were inoculated in LB containing ampicillin and then plasmid DNA was isolated from each transformant. The isolated plasmid DNA was digested with *XcmI* and *SphI* and the restriction digestion profile was studied (figure 3 – 10). The plasmids isolated from wild type strains were, not surprisingly, only cut once by these two enzymes, indicating that they had lost one of the restriction sites. For the *recA*<sup>-</sup> transformants, there were two types of restriction patterns seen. One restriction pattern corresponded to the restriction pattern that would be seen if the starting plasmid were to be digested by *XcmI* and *SphI*. The other restriction pattern was just like that for plasmids isolated from wild type transformants, wherein the plasmid had lost either one or the other restriction site, possibly by resection of that particular arm of homology.

For the *recD*<sup>-</sup> transformants, there was no clear restriction pattern observed, and the sizes of the DNA fragments obtained ranged from two to six fragments, implying that there was more than one plasmid clone in each cell. The loss of exonuclease activity in *recD*<sup>-</sup> cells might allow plasmid molecules to persist long enough to undergo re-arrangement and re-circularize, possibly in a RecA-dependent manner. This would very well be in accordance with the observations by (Biek & Cohen, 1986) that inter-plasmid recombination in RecD-deficient cells is six-fold higher than RecD proficient cells. If this was the case, then such multiple bands would not be seen in a restriction digest of plasmid samples extracted from a *recA*<sup>-</sup> *recD*<sup>-</sup> strain. Indeed, this was observed for plasmid samples isolated from a *recA*<sup>-</sup> *recD*<sup>-</sup> double mutant strain. However, the total size of the plasmid samples obtained from a *recA*<sup>-</sup> *recD*<sup>-</sup> strain was definitely less than the starting plasmid size, suggesting some loss and/or rearrangement of DNA, presumably between the *XcmI* and *SphI* sites.

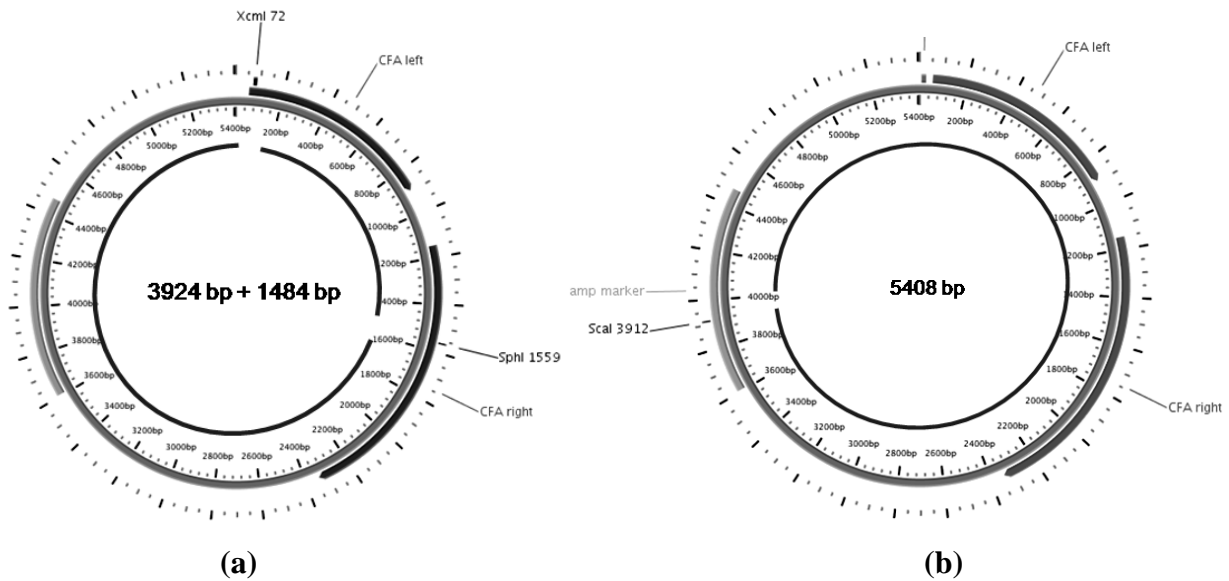
The plasmid samples were then linearized using *ScaI* restriction digestion, which would cut the plasmid backbone inside the ampicillin gene. This would give a nearly-accurate estimate of the size of the plasmid DNA. The results of *ScaI* digestion revealed that the transformant plasmid samples had varying sizes, and almost all of them had lost DNA instead of gaining the desired piece of chromosomal DNA. It was also interesting that most of the plasmid DNA sizes ranged from 2 to 3 kb, which was much less than the expected size of recombined plasmids.



**Figure 3 – 10:** (a) Restriction map showing the *XcmI* & *SphI* restriction site on the recombined pGFP assay plasmid. (b) Restriction map showing the *ScaI* restriction sites on the recombined pGFP assay plasmid.

One thing that was common for the analyzed plasmid DNA samples was the fact that they were all truncated and re-circularized apparently in a RecA-independent manner. Searching for possible mechanisms of RecA-independent plasmid re-circularization led to a paper (Centore et al., 2008) highlighting the existence of a RecA-independent plasmid re-circularization mechanism that relied on exonuclease III, encoded by *xthA*, along with exonuclease IX (encoded by *xni*). Hence, in order to test this hypothesis, the assay plasmid was electroporated into *xthA*<sup>-</sup> cells, *xni*<sup>-</sup> cells and *sbcB*<sup>-</sup> cells as well (since *sbcB* (exonuclease I) has been implicated in plasmid re-circularization along with *xthA* and *xni*). Plasmid DNA was isolated from transformants arising in each of these strains and analyzed via restriction digestion using *XcmI* (which would cut 815 bp upstream of the *EcoRV* site on the left region of homology) and *SphI* (which would

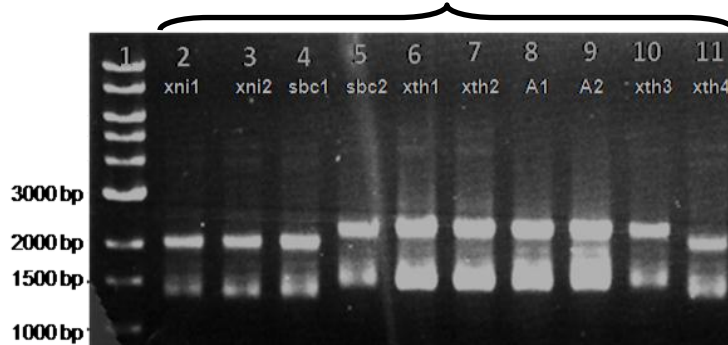
cut 415 bp downstream of the *EcoRV* site on the right region of homology). The restriction digestion revealed that all the plasmid samples had lost both restriction sites, as evidenced by DNA bands corresponding to uncut supercoiled DNA (figure 3 – 11 (c)). All the samples analyzed had some residual undigested DNA. However, a clear picture emerged when the samples were digested with *ScaI*, which cuts the plasmid backbone within the gene encoding ampicillin resistance (figure 3 – 11 (b)). Once linearized with *ScaI*, all these samples showed about the same size, around 2 to 2.5 kb (figure 3 – 11 (d)), which suggested that the plasmids were all being truncated in size even in these strains, and the absence of either *xthA*, *sbcB* or *xni* genes did not have any effect on reducing or eliminating the plasmid resection that was seen earlier.



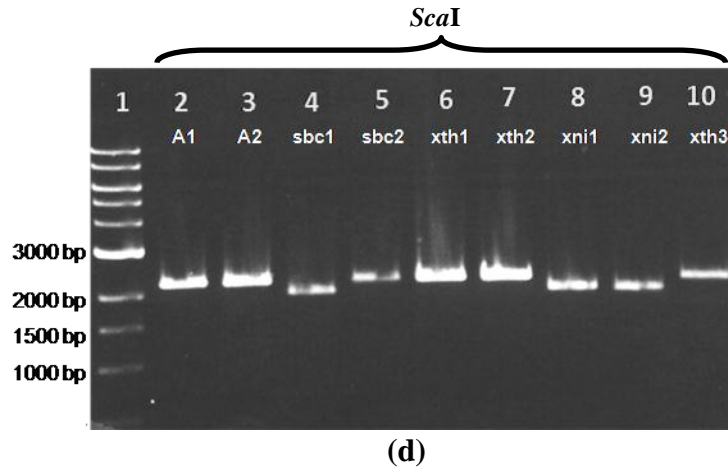
(a)

(b)

*XcmI* and *SphI*



(c)



**Figure 3 – 11:** Analysis of pGFP plasmid DNA isolated from transformants (a) Restriction map showing the *XcmI* and *SphI* restriction sites on the recombined pGFP assay plasmid. (b) Restriction map showing the *ScaI* restriction site on the recombined pGFP assay plasmid. (c) Analytical gel of samples digested with *XcmI* and *SphI*. Lane 1– 1 kb DNA ladder; Lane 2 and 3 – *xni*<sup>-</sup> transformant plasmid DNA samples digested with *SphI* and *XcmI*; Lane 4 and 5 – *sbcB*<sup>-</sup> transformant plasmid DNA samples digested with *SphI* and *XcmI*; Lane 6 and 7 – *xthA1* transformant plasmid DNA samples digested with *SphI* and *XcmI*; Lane 8 and 9 – *recA*<sup>-</sup> transformant plasmid DNA samples digested with *SphI* and *XcmI*; Lane 10 and 11 –  $\Delta$ *xthA* transformant plasmid DNA samples digested with *SphI* and *XcmI*; Lane 12 – Undigested control (Expected size of the correct recombinant clone digested with *SphI* and *XcmI* – 3924 bp and 1484 bp). (d) Analytical gel of samples digested with *ScaI*. Lane 1 – 1 kb DNA ladder; Lane 2 and 3 – *recA*<sup>-</sup> transformant plasmid DNA samples digested with *ScaI*; Lane 4 and 5 –  $\Delta$ *sbcB* transformant plasmid DNA samples digested with *ScaI*; Lane 6 and 7 – *xthA1* transformant plasmid DNA samples digested with *ScaI*; Lane 8 and 9 – *xni*<sup>-</sup> transformant plasmid DNA samples digested with *ScaI*. ; Lane 10 –  $\Delta$ *xthA* transformant plasmid DNA samples digested with *ScaI*. (Expected size of the correct recombinant clone digested with *ScaI* – 5408 bp; Expected size of the wild type plasmid digested with *ScaI* – 5154 bp).

Once it was proven that the recovered plasmids were being resected relative to the transformed substrate, the next step was to somehow devise a strategy to prevent the DNA from being resected. Although the linear dsDNA already possessed tandem *chi* sites (5' GCTGGTGG 3') on both ends, it was hypothesized that perhaps the tandem *chi* sites were not effective at the termini of the linear dsDNA plasmid, and that they would need to be internal to the ends for RecBCD to recognize it. An extensive literature search revealed that this conclusion was also supported by previous findings by (Kulkarni & Julin, 2004) that show that the RecBCD dsDNA exonuclease activity is not affected by *chi* sites present at the termini of the dsDNA and that even a 3 bp of DNA flanking the *chi* site can increase *chi* site recognition by 2.2 fold.

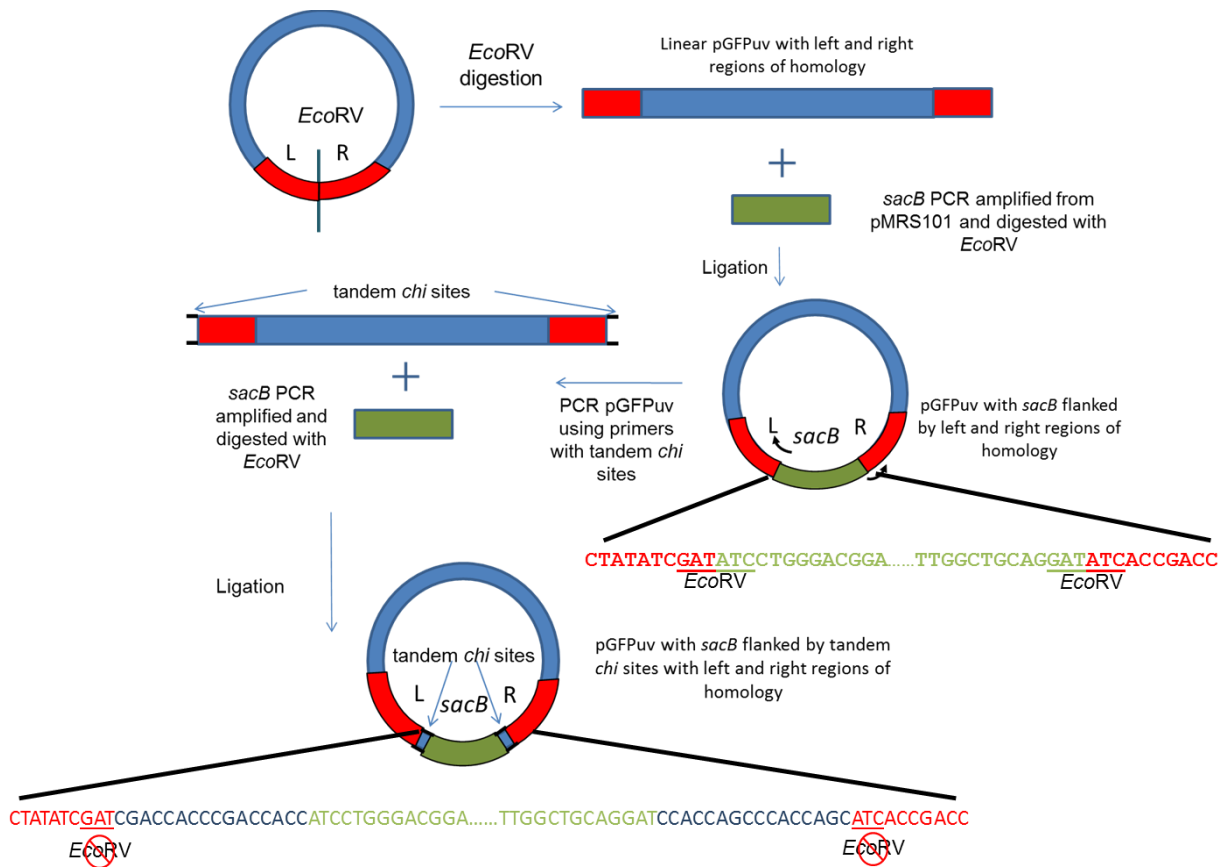
### **3.3.5. Modification to the promoter capture plasmid system – addition of the *sacB* gene**

In order to test whether the plasmid substrate would recombine if the *chi* sites were internal to the ends, the assay plasmid was modified by adding a piece of “stuffer” *sacB* DNA to the linear dsDNA molecule. This was done for two reasons – a) to insert a piece of “stuffer” DNA so as to improve *chi* site recognition and b) to provide counterselectivity against sucrose so that background transformants arising due to transformation of circular parental plasmid DNA could be avoided.

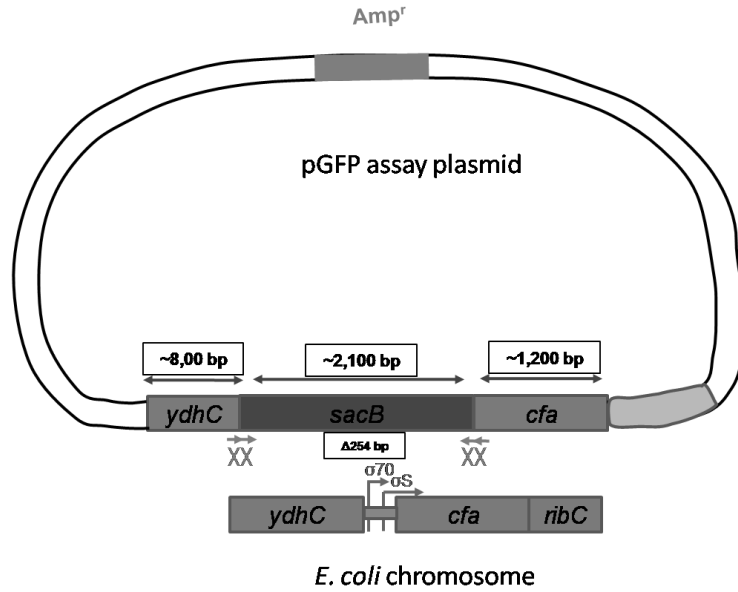
To clone the *sacB* gene into the *gfp* assay plasmid, the *sacB* gene was PCR amplified using primers *sacB-EcoRV-left* and 5' *sacB-EcoRV-right* from plasmid pMRS101. This PCR amplified fragment was then digested with *EcoRV* and ligated to the *gfp* assay plasmid likewise digested at the *EcoRV* site located between the two regions of homology. Following this, XL-1Blue cells were transformed with the ligation mixture and transformants were selected on LB plates supplemented with ampicillin. The resulting Amp<sup>r</sup> transformants were screened for sucrose sensitivity on LB plates containing ampicillin and 10% sucrose. One such sucrose sensitive transformant was inoculated into LB supplemented with ampicillin and allowed to grow overnight. Plasmid extraction was performed the following day using Qiagen's miniprep kit according to the manufacturer's instructions. This plasmid sample was then used as a PCR template to amplify the pGFP backbone along with the left and right regions of homology minus the *sacB* gene using primers *cfa-left-2chi* and *cfa-right-2chi*. The reaction was carried out using Invitrogen's Pfx High Fidelity DNA polymerase, which amplifies DNA fragments without adding “A” overhangs to the amplified product. The *sacB* gene was similarly PCR amplified using primers *sacB-EcoRV-left* and *sacB-EcoRV-right*. Again, this PCR reaction was carried out using Invitrogen's Pfx High Fidelity DNA polymerase. Following PCR, the *sacB* fragment was restriction digested with *EcoRV* and blunt-ligated to linear pGFP having the two arms of homology along with the tandem *chi* sites at the end. The desired ligation product will have destroyed the *EcoRV* restriction sites as the linear pGFP with added Chi sites lacks the corresponding half of the *EcoRV* recognition sites. Following ligation, XL-1Blue cells were transformed with the ligation mixture and transformants were selected on LB plates supplemented with ampicillin. Amp<sup>r</sup> transformants were screened for sucrose sensitivity on LB plates containing ampicillin and 10% sucrose. Since it was highly likely that some of the parental pGFP plasmid with the two arms of homology and *sacB* that was used as a template to generate a



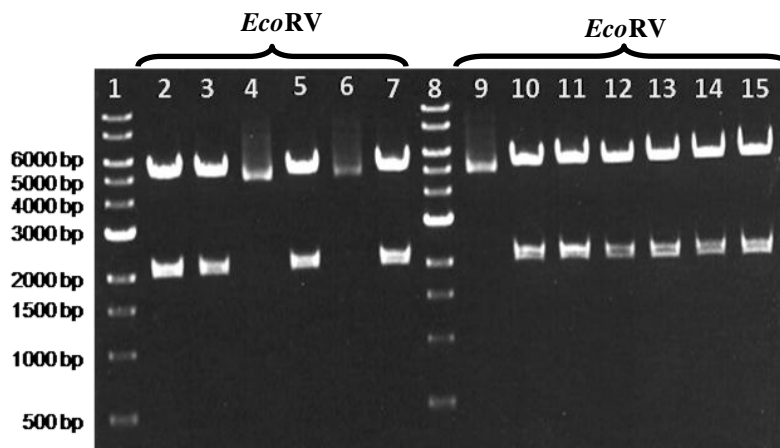
linear plasmid molecule with *chi* sites might have contaminated the ligation reaction, the sucrose resistant clones were then digested with *EcoRV* to screen for the contaminants. Since the *sacB* gene was cloned between the *EcoRV* sites in the parental plasmid, transformants arising due to contamination by the parental pGFP plasmid would give rise to two bands after *EcoRV* digestion. Those clones that would not be digested by *EcoRV* would be the correct clones. The schematic illustration of the whole cloning procedure is shown in figure 3 – 12 (a). As shown in figure 3 – 12 (c), 3 out of 13 transformants are not cut by *EcoRV* (lanes 4, 6, and 9) and therefore these were purified for further experiments. A schematic illustration of the final plasmid layout is given in figure 3 – 12 (b).



(a)



(b)

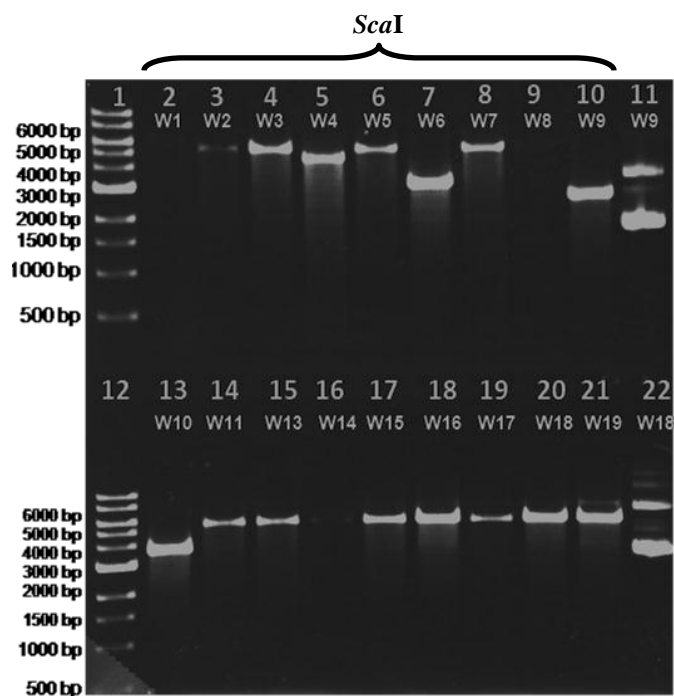


(c)

**Figure 3 – 12:** Construction of the *gfp* assay plasmid system with the addition of a *sacB* gene cassette (a) Schematic illustration showing the cloning scheme for the modified *gfp* assay plasmid with tandem *chi* sites and *sacB*. (b) Diagrammatic illustration of the modified promoter capture assay plasmid with the addition of tandem *chi* sites and *sacB* (c) Restriction digests using *EcoRV* to screen for presence of *sacB*. Lane 1 and 8 – 1 kb DNA ladder; Lane 2 to 7 and 9 to 15 – Clones digested with *EcoRV* to detect the presence of *EcoRV* sites. (Expected size – 5145 bp and 2140 bp)

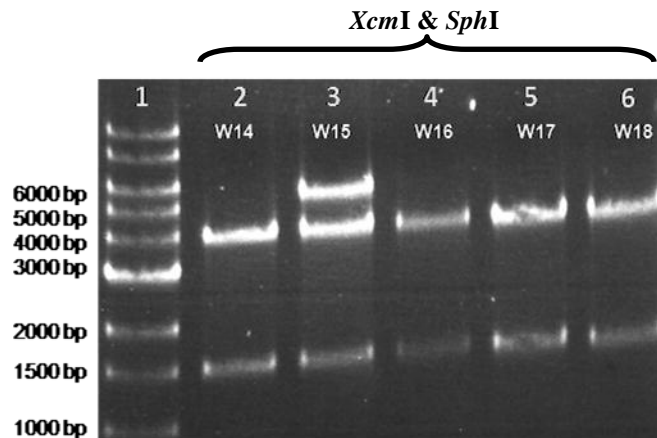
### 3.3.6. Testing of the modified promoter capture plasmid system with tandem *chi* sites and *sacB*

The three plasmid clones that had *sacB* inserted between the two regions of homology with tandem *chi* sites at the end (figure 3 – 12 lanes 4, 6 and 9) were pooled together and digested with *PsiI* which cuts in the *sacB* gene. The linear DNA fragment was extracted from a 0.8% agarose gel and electroporated into wild type MG1655 cells. Cells were allowed to recover for 1 hour in SOC recovery medium and then plated onto LB plates supplemented with ampicillin to select for Amp<sup>r</sup> transformants. None of the Amp<sup>r</sup> colonies displayed fluorescence. 18 colonies from the wild type transformants were selected and plasmid DNA was extracted from them. All 18 plasmid samples were linearized using *ScaI* (that cuts within the gene encoding ampicillin resistance) and the size of each plasmid was examined following electrophoresis on a 0.8% agarose gel (figure 3 – 13).



**Figure 3 – 13:** Plasmid DNA isolated from 18 wild type transformants and digested with *ScaI*. Lane 1 and 12 – 1 kb DNA ladder; Lanes 2 to 10 – Samples 1 to 9 digested with *ScaI*; Lane 11 – Undigested sample 9; Lanes 13 to 21 – Samples 10 to 18 digested with *ScaI*; Lane 22 – Undigested sample 18. (Expected size for the correct recombinant clone – 5408 bp)

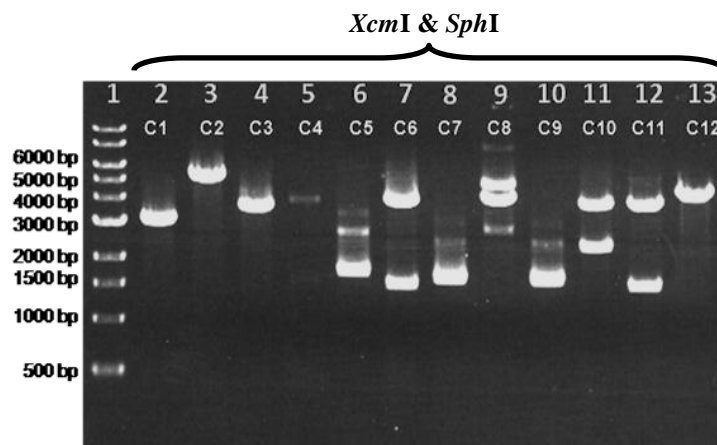
As a confirmatory test, the five plasmid samples that generated the expected size of around 5500 bp following *ScaI* digestion (i.e. samples 14 to 18 in lanes 17 to 21) were digested with *XcmI* and *SphI* and run on a 0.8% agarose gel. As can be seen in figure 3 – 14, with the possible exception of sample 15 (lane 3) which shows three bands, (likely due to incomplete digestion) the samples digested with *XcmI* and *SphI* showed a restriction pattern of two bands, corresponding to the expected 3924 bp and 1484 bp size bands. This result implied that ends-in recombination does occur using this assay plasmid. To confirm that the plasmids had picked up the *cfa* promoter region by homologous recombination, these 5 samples were sent for DNA sequencing. The sequencing results across the arms of homology confirmed that these plasmid samples had all gained the missing 254 bp *cfa* promoter, and the predicted marker rescue had occurred in 5 out of 18 transformants analyzed.



**Figure 3 – 14:** Screening wild type transformant plasmid DNA with *XcmI* and *SphI*. Lane 1 – 1 kb DNA ladder; Lanes 2 to 5 – DNA samples from wild type transformants 14 to 18 digested with *XcmI* and *SphI*. (Expected size – 3924 bp and 1484 bp)

Once it was demonstrated that the assay was working for a wild type strain, the next step was to find out if true recombinants could be obtained in a recipient strain lacking the Holliday Junction resolvase RuvC. Hence, the linearized assay plasmid was electroporated into RuvC-deficient cells and plasmid DNA was isolated from Amp<sup>r</sup> transformants and subjected to restriction digestion analysis using *SphI* and *XcmI*. Similar to the transformants in the wild type strain, no fluorescence was detected in any of the transformants in the *ruvC* strain, although three out of twelve transformants had plasmid DNA that had the expected restriction digestion pattern of ~3900 bp and ~1500 bp (lanes 5, 7 and 12 corresponding to samples C4, C6 and C11 in figure

3 – 15). These samples were sent for DNA sequencing, and the sequencing results proved that the plasmid samples had indeed gained the missing fragment of the *cfa* promoter. This result implies that productive homologous recombination using this assay system can occur in the absence of a Holliday junction resolvase.



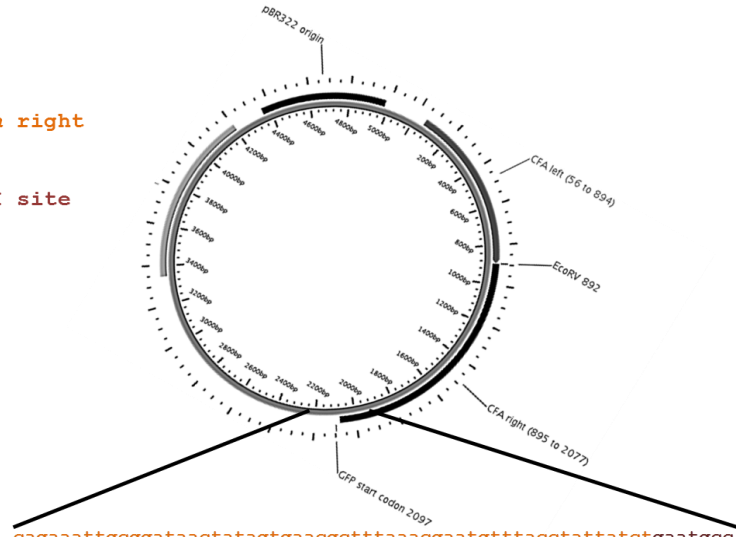
**Figure 3 – 15:** Screening *ruvC* transformant plasmid DNA with *XcmI* and *SphI* digestion. Lane 1 – 1 kb DNA ladder; Lane 2 to 13 – *ruvC* transformant plasmid DNA samples 1 to 12 digested with *XcmI* & *SphI* (Expected size for the correct recombined clone – 3924 bp and 1484 bp)

While it was encouraging to finally extract recombinant plasmids that had successfully undergone homologous recombination with the *E. coli* chromosome, it was also puzzling as to why these transformants clones did not glow. Finally, careful analysis revealed that the *cfa* regions of homology had been cloned such that the ribosome binding site of the *gfp* gene was deleted, and which was probably why the transformants did not express *gfp* and fluoresce. To correct this, the *gfp* gene was amplified from the pGFPuv vector with primers *gfp-left-BsmI* (with a *BsmI* site) and *gfp-right* and then double-digested with *BsmI* and *EagI*, and then ligated to the assay plasmid that was previously digested with *BsmI* and *EagI*. This step also introduced a termination codon TAA right after the *BsmI* site, thereby ensuring that the *gfp* gene was not translationally fused to the *cfa* region of homology. This modified construct resulted in an 89 bp insertion containing the ribosome binding site for *gfp* flanked by the TAA termination codon for *cfa* and AUG codon for *gfp*. The original and modified plasmid constructs are diagrammatically illustrated in figure 3 – 16 (a) and (b). XL-1 Blue cells were transformed with this ligation mixture and plasmid DNA was isolated from the 14 transformants obtained and digested with *XbaI* and *XcmI*. Five clones had the correct restriction digestion pattern of ~4000 bp and ~3000

bp and hence presumably contained the correct version of *gfp* with the ribosome binding site (lanes 4, 5, 7, 11 and 12 in figure 3 – 17) and out of those five, one was selected for further studies. The presence of *sacB* in these plasmids was verified by screening transformants for sucrose sensitivity on LB plates containing ampicillin and 10% sucrose.

**Legend**

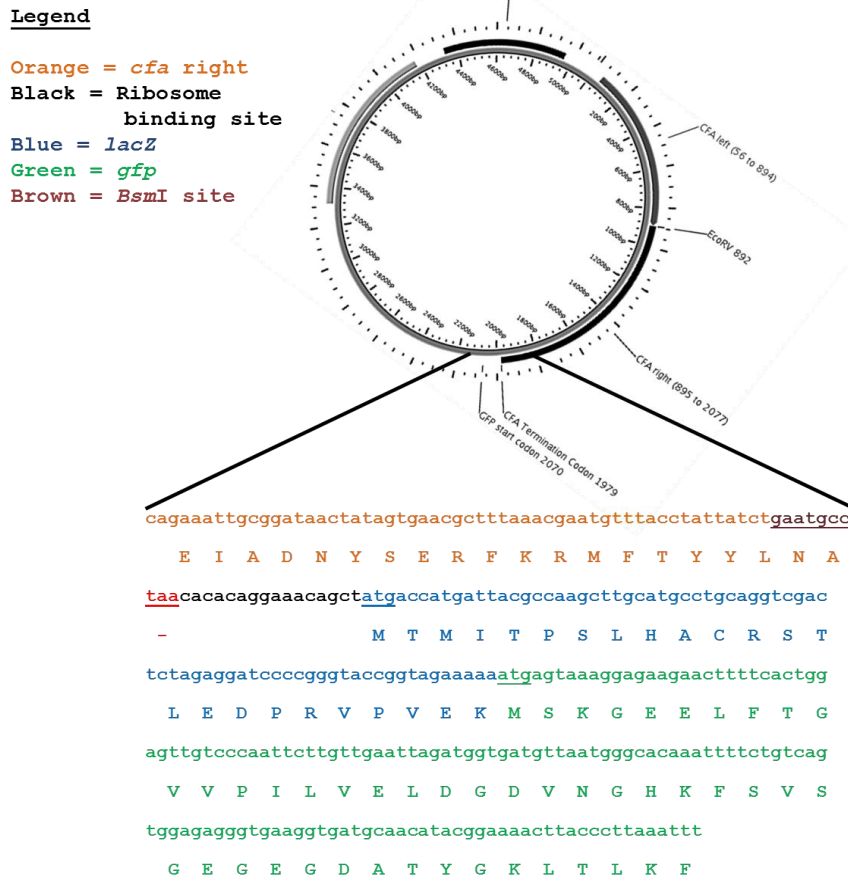
- Orange = *cfa* right
- Blue = *lacZ*
- Green = *gfp*
- Brown = *BsmI* site



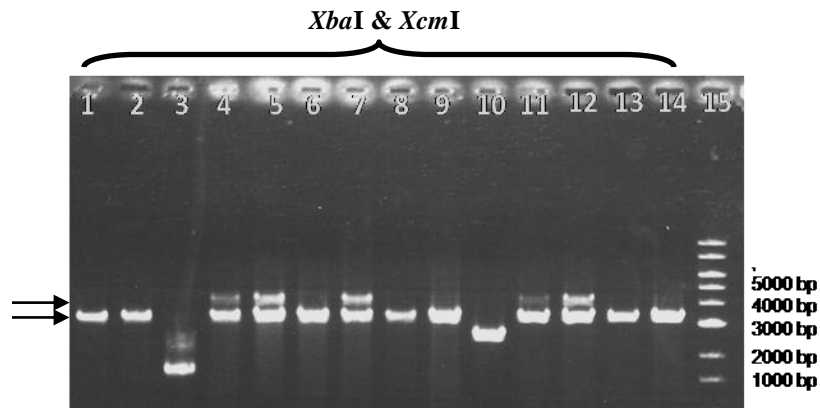
```

cagaaattgCGgataactatagTgaacgctttaaacgaatgTttacotattatctgaatgcc
  E I A D N Y S E R F K R M F T Y Y L N A
tgtgcaggtgctttccgcgccgtgatattcagctctggcaggcggctegtgttctcaccg
  C A G A F R A R D I Q L W Q A V V F S R
ggtgttgaaaacggccttcogagtgctcgcTaaaggctaccggtagaaaaaat
  G V E N G L R V A R - M
gagtaaaggagaagaacttttcaactggagttgtoccaattctgttgtaattagatggtga
  S K G E E L F T G V V P I L V E L D G D
tgtaaattgggcacaaattttctgtcagtgagagggTgaaggtgatgcaacatacgaaa
  V N G H K F S V S G E G E G D A T Y G K
acttacccttaaatttatttgcact
  L T L K F I C T
  
```

(a)



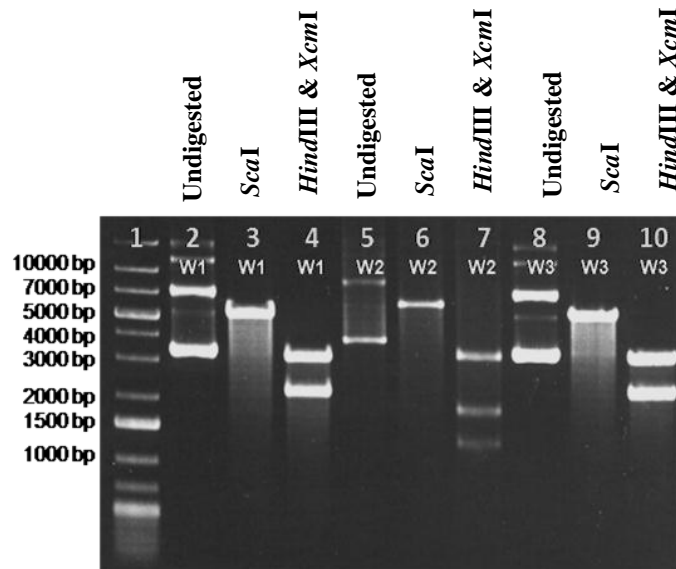
**Figure 3 – 16:** Diagrammatic illustration of the pGFP assay plasmid (a) before and (b) after addition of the CFA termination codon with the ribosome binding site.



**Figure 3 – 17:** Restriction digest of transformants with *XbaI* & *XcmI* to verify insertion of the correct *gfp* gene with the ribosome binding site. Lane 15 – 1 kb DNA ladder; Lanes 1 to 14 – plasmid DNA isolated from transformants after restriction digestion with *XbaI* and *XcmI*. (Expected band size for the correct version of *gfp* with the ribosome site – 3977 bp and 3158 bp).

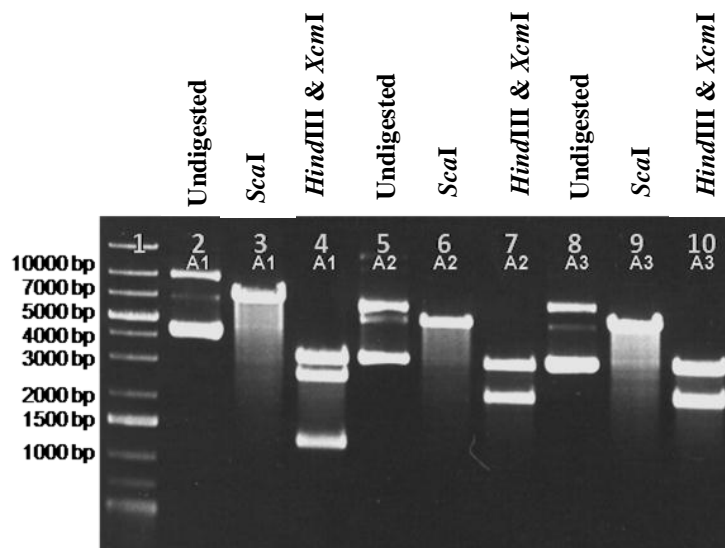
In order to determine whether inclusion of the ribosome binding site allows transformants to express *gfp*, the assay plasmid substrate was linearized with *Sna*BI (which cuts the assay plasmid once near the center of the *sacB* gene) and then electroporated it into wild type and *recA*<sup>-</sup> electrocompetent cells. Amp<sup>r</sup> colonies were selected on LB plates containing ampicillin. Surprisingly, fluorescent colonies were observed for BOTH wild type and *recA*<sup>-</sup> transformants, suggesting that perhaps the *gfp* reporter expression was not limited to those plasmids which had successfully undergone homologous recombination with the chromosomal *cfa* promoter region. Plasmid DNA was extracted from three glowing transformants each from wild type and *recA*<sup>-</sup> strains and digested with a) *Sca*I and b) *Hind*III and *Xcm*I. The restriction digested plasmid DNA was then run on a 0.8% agarose gel for analysis.

Figure 3 – 18 shows the results of restriction digestion for the three isolates from wild type cells. Based on the restriction pattern seen, it appeared that samples 1 and 3 (in lanes 3, 4, 9 and 10) were true recombinants which had regained the missing *cfa* fragment as they showed the expected band pattern (~5000 bp for *Sca*I and ~3900 bp and ~1500 bp for *Hind*III and *Xcm*I).



**Figure 3 – 18:** Restriction digest of plasmid DNA isolated from wild type transformants. Lane 1 – 1 kb DNA ladder; Lanes 2, 5 and 8 – Wild type glowing transformant plasmid DNA samples (undigested control); Lanes 3, 6 and 9 – Wild type glowing transformant plasmid DNA samples digested with *Sca*I; Lanes 4, 7 and 10 – Wild type glowing transformant plasmid DNA samples digested with *Hind*III and *Xcm*I. (Expected size for the correct recombined clone – 5408 bp with *Sca*I and 3924 bp and 1484 bp with *Hind*III & *Xcm*I)

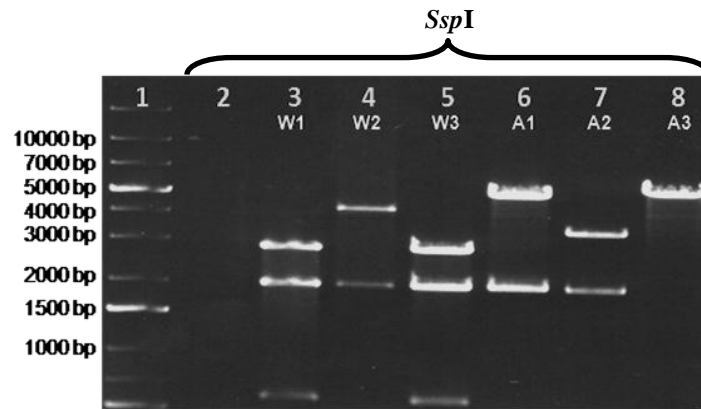




**Figure 3 – 19:** Restriction digest of plasmid DNA isolated from *recA*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2, 5 and 8 – *recA*<sup>-</sup> glowing transformant plasmid DNA samples (undigested control); Lanes 3, 6 and 9 – *recA*<sup>-</sup> glowing transformant plasmid DNA samples digested with *ScaI*; Lanes 4, 7 and 10 – *recA*<sup>-</sup> glowing transformant plasmid DNA samples digested with *HindIII* and *XcmI*.

Figure 3 – 19 shows the results of the *HindIII* and *XcmI* digestion for *recA*<sup>-</sup> transformants. The restriction pattern for plasmid DNA isolated from *recA*<sup>-</sup> transformants was very similar to the plasmid DNA samples isolated from wild type transformants (both for *ScaI* digested and *HindIII* and *XcmI* digested samples in figure 3 – 18). Samples 2 and 3 from *recA*<sup>-</sup> transformants produced identical restriction patterns to those seen with samples 1 and 3 from wild type transformants. Given the extremely low probability of homologous recombination occurring in the absence of RecA, wild type samples 1 and 3 and *recA*<sup>-</sup> samples 2 and 3 were tested further by digestion with *SspI*. The advantage of using *SspI* was that it had a recognition site in the missing region of *cfa* promoter, as well as two additional sites on the assay plasmid. A true recombinant would be cleaved into bands of 2802 bp, 2008 bp and 598 bp. Figure 3 – 20 revealed that the two plasmid samples from wild type transformants did indeed have the correct region of *cfa* inserted between the two regions of homology since they gave the expected sized bands at around ~2800 bp, ~2000 bp and ~600 bp – lanes 3 and 5 following *SspI* digestion. As predicted, none of the plasmids from the *recA*<sup>-</sup> transformants had picked up the *cfa* region as none of them gave the correct sized bands (lanes 6, 7 and 8 in figure 3 – 20). However, due to the fact that fluorescent colonies were observed that contained plasmids that were not true recombinants, it was

concluded that not all transformants that fluoresce have the correct recombinant plasmid with the missing *cfa* promoter.



**Figure 3 – 20:** *SspI* restriction digest of plasmid DNA isolated from wild type and *recA*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 3 to 5 – Wild type glowing transformant plasmid DNA samples digested with *SspI*; Lanes 6 to 8 – *recA*<sup>-</sup> glowing transformant plasmid DNA samples digested with *SspI*.

Since it was ascertained that *SspI* digestion was the most reliable method to analyze plasmid DNA from transformants, *SnaBI* linearized assay plasmid was sequentially transformed into isogenic strains that were either wild type or deleted for one or more recombination genes and scored for those that had gained the missing *cfa* promoter region. As a transformation control, a pre-determined amount of circular plasmid DNA (pCDF Duet 1 carrying streptomycin resistance) was mixed into the batch of linear *SnaBI* digested plasmid DNA and then this mastermix was used for transformation. Results of these transformations have been summarized in table 3 – 2, table 3 – 3 and table 3 – 4 respectively. The tables also include a ratio of the number of transformants (glowing, non-glowing and intermediate glowing) arising from the linear assay plasmid to the number of transformants arising from a covalently closed circular plasmid (control).

**Table 3 – 2: Results of Transformation Round 1**

Genotype	Glowing	Intermediate	Non-glowing	Streptomycin resistant	Ratio
					Linear plasmid/circular plasmid
Wild type	71	25	427	3	174.33
<i>recA</i> <sup>-</sup>	40	22	303	21	17.38
<i>recD</i> <sup>-</sup>	TNTC*	TNTC*	TNTC*	12	ND <sup>†</sup>
<i>recA</i> <sup>-</sup> <i>recD</i> <sup>-</sup>	59	33	0	0	ND <sup>†</sup>
<i>ruvC</i> <sup>-</sup>	10	23	372	0	ND <sup>†</sup>
<i>recG</i> <sup>-</sup>	40	174	629	24	35.12
<i>ruvC</i> <sup>-</sup> <i>recG</i> <sup>-</sup>	0	0	3	2	ND <sup>†</sup>

\* = Too Numerous To Count

† = Not Determined

**Table 3 – 3: Results of Transformation Round 2**

Genotype	Glowing	Intermediate	Non-glowing	Streptomycin resistant	Ratio
					Linear plasmid/circular plasmid
Wild type	81	18	474	2	286.5
<i>recA</i> <sup>-</sup>	40	32	488	8	70
<i>recD</i> <sup>-</sup>	TNTC*	TNTC*	TNTC*	2	ND <sup>†</sup>
<i>recA</i> <sup>-</sup> <i>recD</i> <sup>-</sup>	0	0	0	0	0
<i>ruvC</i> <sup>-</sup>	2	24	420	3	148.67
<i>recG</i> <sup>-</sup>	200	96	324	20	31
<i>ruvC</i> <sup>-</sup> <i>recG</i> <sup>-</sup>	0	0	13	0	ND <sup>†</sup>

\* = Too Numerous To Count

† = Not Determined

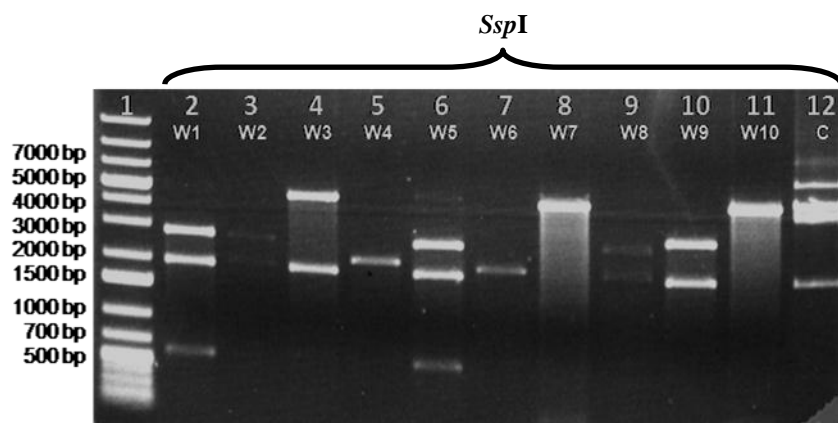
**Table 3 – 4: Results of Transformation Round 3**

Genotype	Glowing	Intermediate	Non-glowing	Streptomycin resistant	Ratio
					Linear plasmid/circular plasmid
Wild type	58	31	431	11	47.27
<i>recA</i> <sup>-</sup>	21	59	501	4	145.25
<i>recD</i> <sup>-</sup>	TNTC*	TNTC*	TNTC*	6	ND <sup>†</sup>
<i>recA</i> <sup>-</sup> <i>recD</i> <sup>-</sup>	0	0	0	0	0
<i>ruvC</i> <sup>-</sup>	12	27	500	2	269.5
<i>recG</i> <sup>-</sup>	180	112	640	2	466
<i>ruvC</i> <sup>-</sup> <i>recG</i> <sup>-</sup>	0	0	11	0	ND <sup>†</sup>

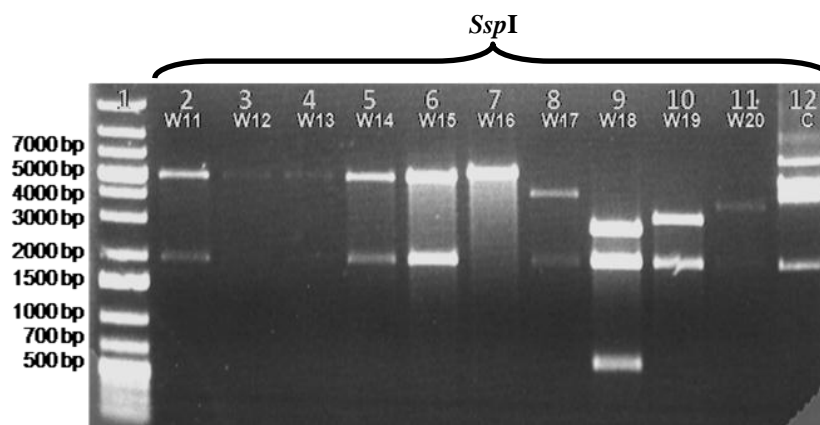
\* = Too Numerous To Count

<sup>†</sup> = Not Determined

Initially, ten transformants from a wild type strain that were non-glowing, five that were glowing but were not very highly expressed for *gfp* and five that were fully expressed for *gfp* and were glowing brightly were selected and subjected to plasmid extraction followed by restriction analysis of the extracted plasmids using *SspI* digestion. Figure 3 – 21 shows that two plasmid clones from non-glowing transformants (lanes 2 and 6 in figure 3 – 21 (a)), and one plasmid clone from a glowing transformant (lane 9 in figure 3 – 21 (b)) generated the expected banding pattern of 2802 bp, 2008 bp and 598 bp indicating they were true recombinant plasmids which regained the missing *cfp* promoter. As expected, a *recA*<sup>-</sup> strain did not give rise to any transformants which contained the correct recombinant plasmid clone (figure 3 – 22 (a) and (b)). Three of the plasmids analyzed from *recA*<sup>-</sup> transformants had a restriction pattern corresponding to the native assay plasmid (figure 3 – 22 (a) lane 8 and (b) lanes 8 and 9). The remaining plasmids were truncated in size.

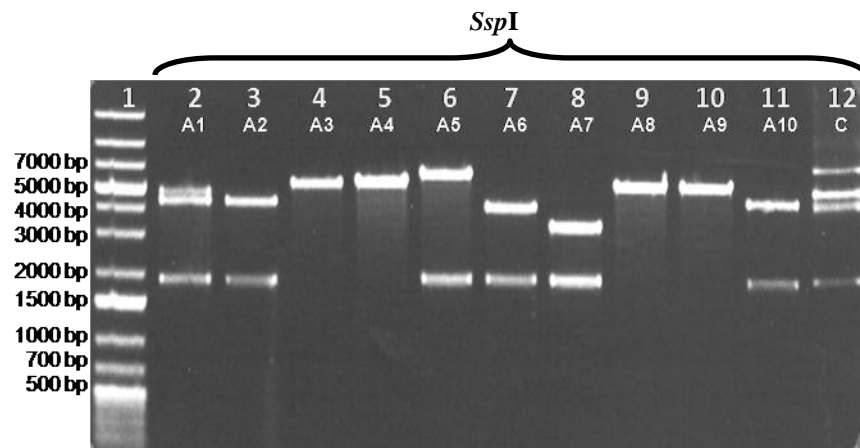


(a)

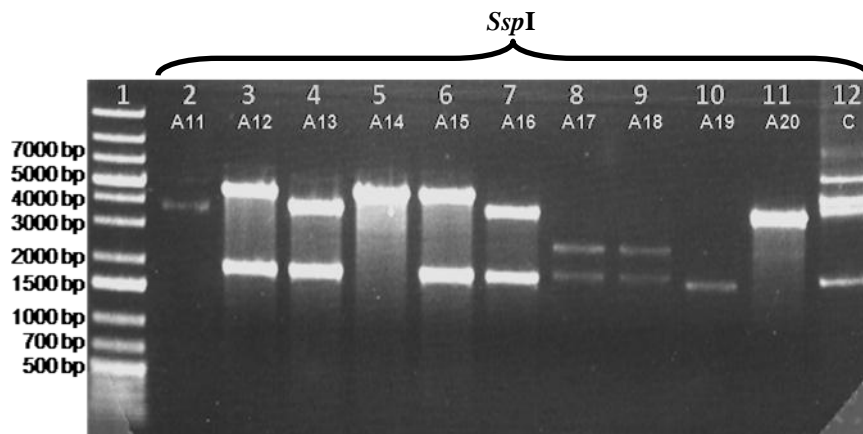


(b)

**Figure 3 – 21:** *SspI* restriction digest of plasmid DNA isolated from wild type transformants. **(a)** *SspI* restriction digest of plasmid DNA isolated from wild type non-glowing transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – Wild type non glowing transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp). **(b)** *SspI* restriction digest of plasmid DNA isolated from glowing wild type transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 6 – Wild type intermediate glowing transformant plasmid DNA samples digested with *SspI*; Lanes 7 to 11 – Wild type glowing transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).



(a)

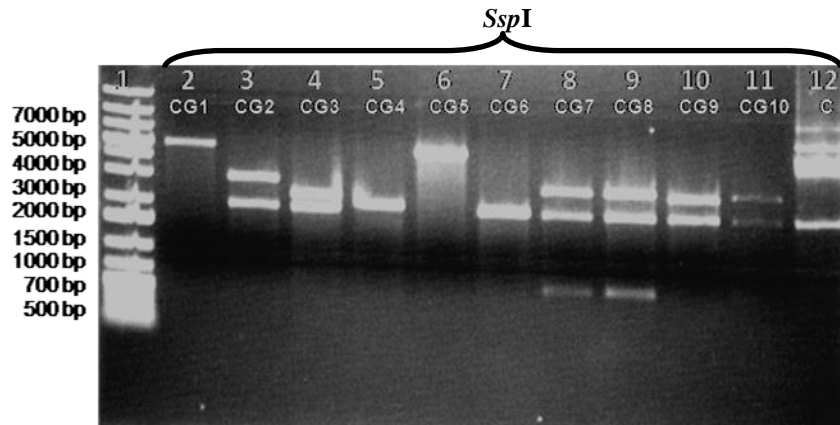


(b)

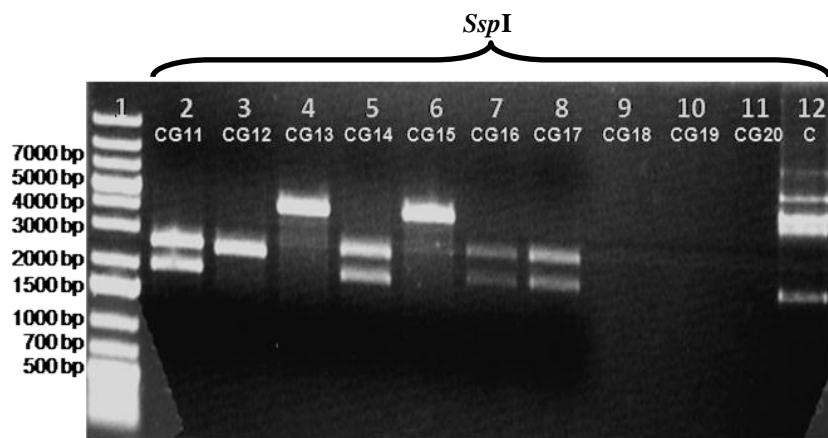
**Figure 3 – 22:** *SspI* restriction digest of plasmid DNA isolated from *recA*<sup>-</sup> transformants. **(a)** *SspI* restriction digest of plasmid DNA isolated from non-glowing *recA*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *recA*<sup>-</sup> non glowing transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp). **(b)** *SspI* restriction digest of plasmid DNA isolated from glowing *recA*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 6 – *recA*<sup>-</sup> intermediate glowing transformant plasmid DNA samples digested with *SspI*; Lanes 7 to 11 – *recA*<sup>-</sup> glowing transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).

For a *ruvC*<sup>-</sup> *recG*<sup>-</sup> cell, 17 transformants were obtained, none of which were expressing *gfp* but two clones were found to have the correct insertion of the missing *cfa* fragment (lanes 8 and 9 in figure 3 – 23 (a)), and these results are of special significance, as explained below in the discussion section. Unfortunately, for a *ruvC*<sup>-</sup> background, none of the transformants analyzed had the correct plasmid clone (figure 3 – 24 (a) and (b)); even though previous results had shown that ends-in recombination does allow the assay plasmid to “pick up” the missing 254 bp piece of DNA from the chromosome in the absence of RuvC. However, this might just be due to the possibility that the batch of transformants selected for restriction digestion analysis did not have the right plasmid in them.

It was evident now that screening transformants for *gfp* expression would not be feasible since even transformants that do not express *gfp* were found to possess the expected recombinated plasmid clone. Hence random transformants were selected for the *recG*<sup>-</sup> strain as well as for the *ruvC*<sup>-</sup> strain and then plasmid DNA was isolated from them and digested with *Ssp*I.

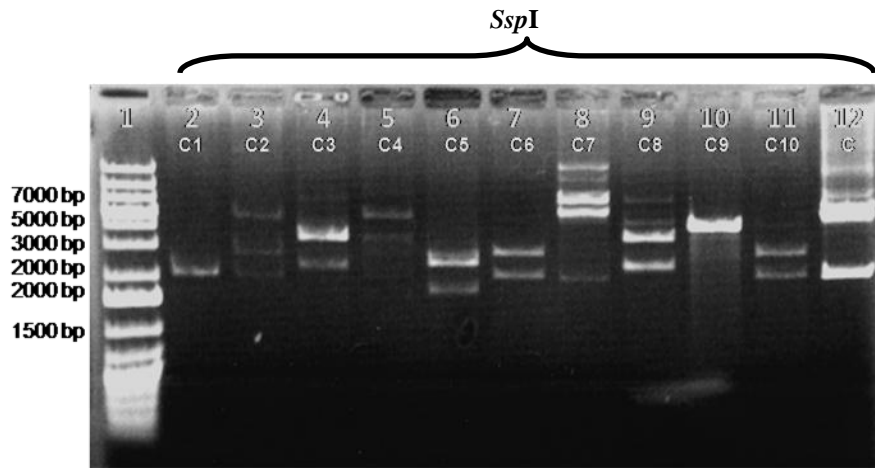


(a)

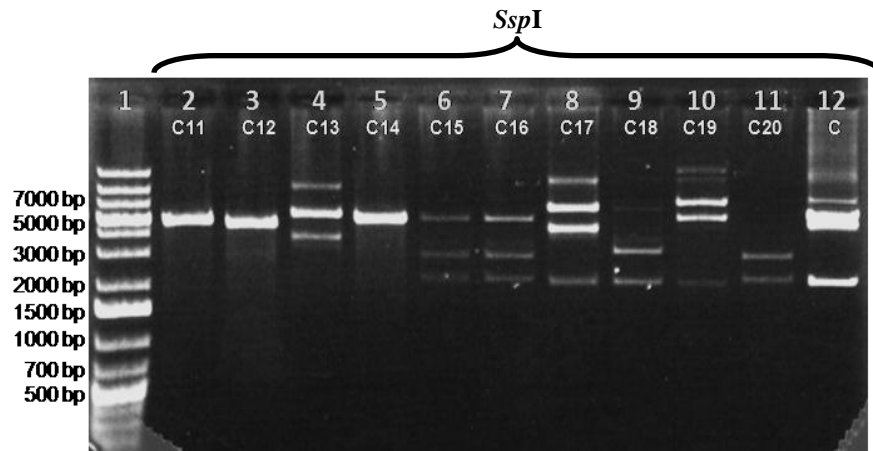


(b)

**Figure 3 – 23:** (a) *SspI* restriction digest of plasmid DNA isolated from *ruvC<sup>-</sup> recG<sup>-</sup>* transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *ruvC<sup>-</sup> recG<sup>-</sup>* transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp). (b) *SspI* restriction digest of plasmid DNA isolated from *ruvC<sup>-</sup> recG<sup>-</sup>* transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 8 – *ruvC<sup>-</sup> recG<sup>-</sup>* transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).



(a)



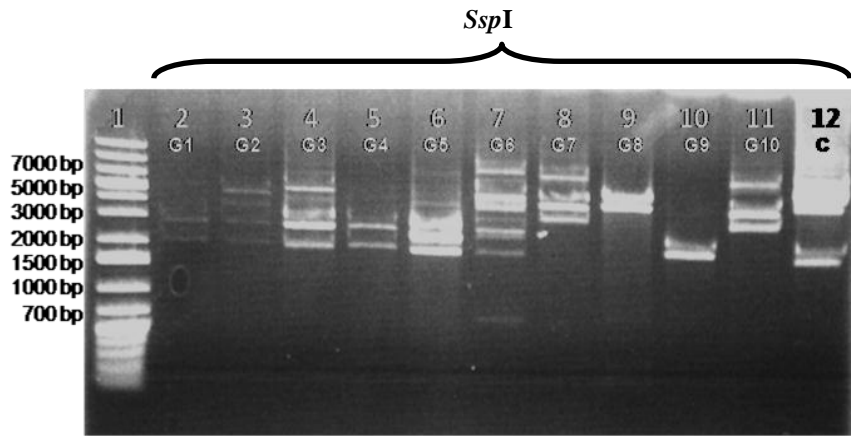
(b)

**Figure 3 – 24:** (a) *SspI* restriction digest of plasmid DNA isolated from *ruvC<sup>-</sup>* transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *ruvC<sup>-</sup>* transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp). (b) *SspI* restriction digest of plasmid DNA

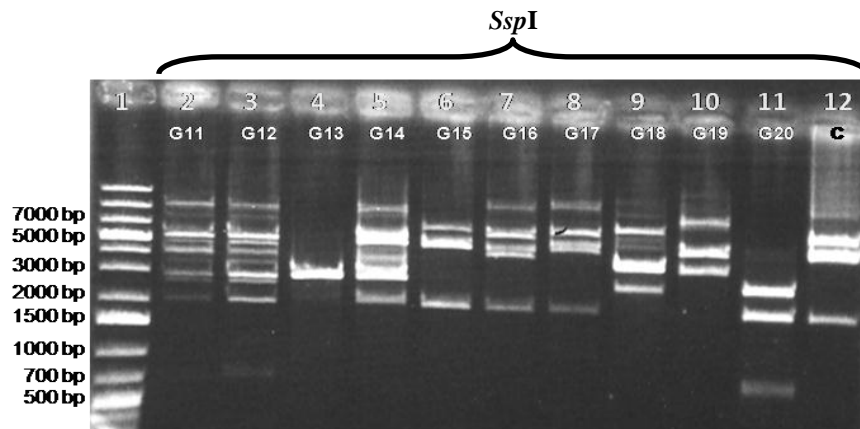


isolated from *ruvC* transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *ruvC* transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).

For a *recG*<sup>-</sup> strain, three clones seemed to have the correct restriction digestion pattern, one of which was isolated from a non-glowing transformant (lane 7 in figure 3 – 25 (a)) and the other two were isolated from an intermediate glowing and glowing transformant (lane 3 and lane 11 in figure 3 – 25 (b)). However, multiple restriction patterns were observed for every *recG*<sup>-</sup> plasmid clone (figure 3 – 25 (a) and (b)), hinting at the possibility that perhaps the absence of a functional RecG somehow gave rise to multiple plasmid clones that were unstable.



(a)

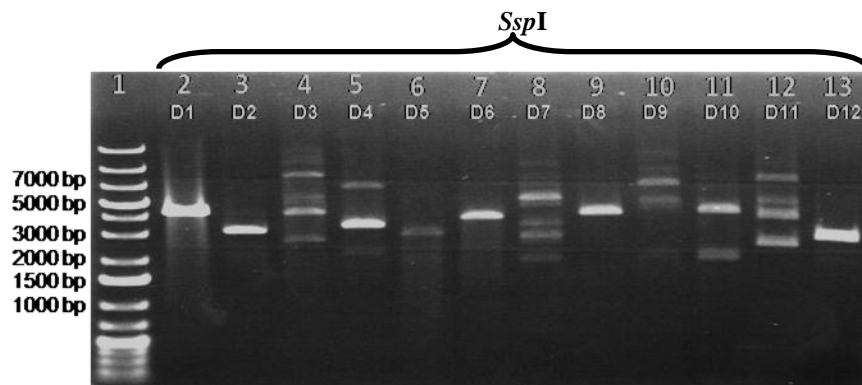


(b)

**Figure 3 – 25:** (a) *SspI* restriction digest of plasmid DNA isolated from *recG*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *recG*<sup>-</sup> transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct

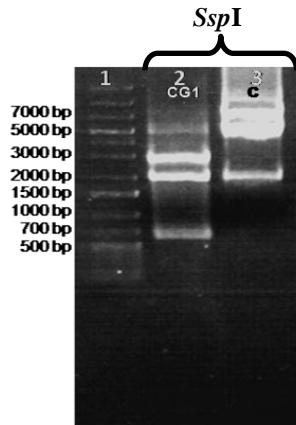
recombined clone – 2802 bp, 2008 bp and 598 bp). **(b)** *SspI* restriction digest of plasmid DNA isolated from *recG*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 11 – *recG*<sup>-</sup> transformant plasmid DNA samples digested with *SspI*; Lane 12 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).

For a *recD*<sup>-</sup> strain, the number of transformants was so high, that it was very hard to isolate single colonies to perform restriction analysis. One part of the problem was that somehow the number of satellite colonies was very high, probably brought about by plasmid instability (see discussion for further explanation). In spite of this, plasmid DNA was extracted from *recD*<sup>-</sup> transformants and analyzed by *SspI* restriction digestion (figure 3 – 26). The restriction profile of the *recD*<sup>-</sup> transformants suggest that none of the transformants analyzed had the expected band size of 2802 bp, 2008 bp and 598 bp, indicating possible gene rearrangements in the assay plasmid.



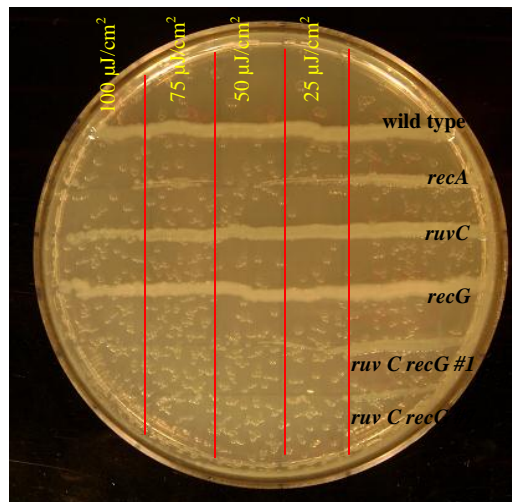
**Figure 3 – 26:** *SspI* restriction digest of plasmid DNA isolated from *recD*<sup>-</sup> transformants Lane 1 – 1 kb DNA ladder; Lanes 2 to 13 – *recD*<sup>-</sup> transformant plasmid DNA samples digested with *SspI*. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).

As a final screening test, plasmid DNA isolated from one of the *ruvC*<sup>-</sup> *recG*<sup>-</sup> double mutants was digested with *SspI* and the restriction pattern was compared to the restriction pattern of the starting plasmid (figure 3 – 27). As expected, they had different restriction patterns and although this clone was not sequenced, previous experiments had shown that the results of restriction analysis always concurred with DNA sequencing.



**Figure 3 – 27:** Confirmatory digest of plasmid DNA isolated from *ruvC<sup>-</sup> recG<sup>-</sup>* transformant using *SspI*. Lane 1 – 1 kb DNA ladder; Lane 2 - *ruvC<sup>-</sup> recG<sup>-</sup>* transformant plasmid DNA sample digested with *SspI*; Lane 3 – *SspI* digested *gfp* assay plasmid as control. (Expected band size for the correct recombined clone – 2802 bp, 2008 bp and 598 bp).

The image below (figure 3 – 28) shows the relative UV sensitivity for some of the different strains used throughout this study. The main purpose of testing UV sensitivity in these strains was to ensure that the *ruvC<sup>-</sup>*, *recG<sup>-</sup>* and *ruvC<sup>-</sup> recG<sup>-</sup>* strains have retained their characteristic UV sensitivity phenotype since these strains are much more likely to accumulate suppressors that can revert such strains to wild type phenotype.



**Figure 3 – 28:** Plate showing UV sensitivity of different assay strains tested. (UV exposure is in  $\mu\text{J}/\text{cm}^2$ .)

Although the *gfp* assay plasmid was intended to be used for a direct screening of transformants based on the assumption that the correctly recombined plasmid clone would allow expression of *gfp*, the presence of false positives arising during transformation rendered this screening process inaccurate. Moreover, screening and analysis of plasmid clones obtained from the pool of transformants indicated that the number of transformants which did not contain the correct recombined plasmid clone was very high. In fact, for every 20 transformants tested from different genetic backgrounds, only 1 to 3 transformants had the correct recombined plasmid clone. Due to the vast number of false positive transformants, there was no linear co-relation between the number of transformants arising from the linearized assay plasmid and the number of transformants containing the correct recombined plasmid clone. In spite of this, different isogenic strains of *E. coli* were transformed with the linear assay plasmid along with a non-homologous circular covalently closed plasmid as control. A ratio of the number of transformants (glowing, non-glowing and intermediate glowing) arising from the linear assay plasmid to the number of transformants arising from a covalently closed circular plasmid (control) has also been included. Results of these transformations have been summarized in table 3 – 2, table 3 – 3 and table 3 – 4 respectively. It was interesting to note that although the number of transformants obtained for each strain has been roughly in the same range, the ratio of linear to circular plasmid transformants changes drastically from each set of transformations, purely because of a large variation in circular plasmid transformants.

# CHAPTER FOUR

## DISCUSSION

### **4.1. Ends-in recombination – an overview**

An intensely studied area of research focuses on mechanisms of replication fork collapse and the subsequent recombinational repair of such collapsed replication forks in model organisms like *E. coli* and *S. cerevisiae* (Marians, 2004) (Lopes et al., 2001); (Cox, 2002). Two of the best characterized causes of replication fork collapse are the conversion of a single strand gap to a double strand break when a growing replication fork tries to traverse across the gap and when a double strand break in template DNA causes stalling and collapse of the replication fork (Liberi et al., 2006) (Cahill et al., 2006) (Kuzminov, 2001) (Nowosielska & Marinus, 2008). Such collapsed replication forks are fixed by the dsDNA repair machinery of the cell (Kuzminov, 1999). In *E. coli*, this is done through a subset of genes and proteins that include the RecBCD pathway of dsDNA break repair along with the primosomal pathways of replication restart. (Amundsen et al., 1990) (Brcic-Kostic et al., 1991) (Sandler, 2000) (Michel et al., 2004). RecBCD aids in processing blunt ended DNA to facilitate loading of the RecA recombinase which then leads to homologous strand pairing between one of the broken dsDNA strands and an intact duplex DNA (Amundsen et al., 1990). Branch migration and resolution of Holliday Junctions formed during homologous strand pairing is carried out by the RuvABC and/or RecG proteins. It has been hypothesized that in the presence of both RuvC and RecG, resolution of Holliday Junctions employs a “cut and paste” mechanism in which the recombining DNA molecule replaces the target duplex DNA. However, in the absence of RuvC and RecG, a “copy and paste” mechanism is employed in which DNA synthesis is primed from the ends of invading duplex DNA and proceeds across the entirety of the chromosome (Motamedi et al., 1999). Formation of successful recombinant products in the absence of Holliday Junction resolution proteins RuvC and RecG therefore has to rely on DNA synthesis across the entire chromosome (Motamedi et al., 1999). This could confound analysis of recombinant products. Elucidating the role of different replication restart proteins following Holliday Junction formation would be very difficult if there are different subsets of proteins required to a) resolve the Holliday Junction crossover and b) restart replication at sites blocked by the crossover event or otherwise blocked

during replication of the *E. coli* chromosome. One potential workaround to this problem could be the use of  $\lambda \times \lambda$  crosses. However, the use of  $\lambda \times \lambda$  crosses might confound analysis since the lambda genome might encode proteins that might possibly interfere with successful recombinant product formation. Moreover, using  $\lambda \times \lambda$  crosses might eliminate the requirements for RuvAB and RecG branch migration proteins possibly because the distance required for branch migration is shorter or because phage DNA packaging might influence branch migration. Hfr mating (crossovers) would be similarly inadequate since Hfr mating would also involve the need to replicate the entirety of the *E. coli* chromosome to obtain a scorable recombinant product. The ends-in recombination system was designed to address the shortcomings of conventional assay systems by providing the researcher with a substrate that would allow differentiation of Holliday Junction resolution between the previously demonstrated “cut and paste” RuvC and RecG dependent pathway and the “copy and paste” RuvC and RecG independent pathway (Motamedi et al., 1999). Moreover, the size of DNA needed to be synthesized for successful resolution of recombination intermediates via the “copy and paste” pathway could be manipulated (researcher defined).

In all the assay systems tested here for all strains of *E. coli*, there was evidence of DNA degradation at the ends of the regions of homology, consistent with nuclease degradation and recircularization. Analysis of the re-circularized *gfp* assay plasmid recovered following transformation showed that there was a consistent deletion/rearrangement of the assay plasmids, with all the plasmids analyzed showing a truncation flanking the dsDNA break site. It was also interesting to note that the number of transformants in a *recD*<sup>-</sup> strain was much higher than wild type or *recA*<sup>-</sup> strains. One possible reason for this could be that the loss of exonuclease activity in a *recD* mutant allows linear dsDNA to persist longer (Kuzminov & Stahl, 1997), thereby allowing sufficient time for recircularization. Interestingly, recircularized plasmid transformants were also observed in a RecA deficient background. These had lost some or whole of the region of homology flanking the dsDNA break site. Our working hypotheses about the outcome of the plasmid substrate following transformation did not provide an explanation as to the formation of truncated plasmids. Initially, two assumptions were made, both of which were mutually exclusive. The first assumption was that RecBCD would completely degrade the linear dsDNA and in cells where this occurred, no transformants would be seen. The second assumption was

that the dsDNA would recombine with the chromosomal target and produce recombinants as expected. Neither of these assumptions provided an explanation for the formation of the truncated plasmids. A search of previous literature revealed that certain exonucleases in *E. coli*, namely exonuclease III encoded by the gene *xthA*, exonuclease IX, exonuclease X and exonuclease XI were responsible for antagonizing the RecBCD pathway and preventing availability of substrates for RecBCD mediated recombination (Centore et al., 2008). This study also showed that the antagonistic effects of these exonucleases were dependent on exonuclease I encoded by *xonA* (also known as *sbcB*). More interesting was the result of two studies undertaken using transformation of linear plasmid pBR322 DNA molecules (Conley et al., 1986a); (Conley et al., 1986b). These studies showed that linear pBR322 plasmids transformed into wild type strains of *E. coli* were able to re-circularize in a RecA-independent manner. Moreover, a majority of these re-circularized plasmids had lost some part of their DNA near the site of linearization. It was suggested that even a region of homology as low as 4 to 10 bp could bring about plasmid re-circularization, albeit such events would be aberrant and result in plasmid modifications. However, transformation of the linear assay plasmid into strains deficient for various exonucleases did not eliminate the aberrant re-circularization of the assay plasmid that had been observed previously. It was possible that total elimination of the aberrant plasmid re-circularization would require inactivating two or more genes encoding the above exonucleases. DNA sequencing of plasmids isolated from wild type, *recA* and *recD* strains revealed no definite pattern of re-circularization and in each case the DNA sequence inserted between the arms of homology varied in size.

Following transformation in a *recD*<sup>-</sup> strain, a very large number of satellite colonies were observed, so much so that it was very hard to count the number of actual transformants in most cases. Previous studies by (Biek & Cohen, 1986; Cohen & Clark, 1986) had demonstrated that plasmid stability is affected by *recD* inactivation which allows the plasmid molecules to replicate through a rolling circle pathway, generating large linear multimers. If this was the case, it could explain how transformants arising from a *recD*<sup>-</sup> strain accumulate an overwhelming number of satellite colonies, most likely due to increasing production of beta-lactamase brought about by circumventing copy number controls on rolling circle plasmid replication. Biek & Cohen (1986) observed that this plasmid instability in a *recD* strain depends upon a functional RecA protein, likely because RecA-dependent recombination protects rolling circle replication intermediates

from degradation by nucleases like exonuclease III. In agreement with the above observation, multiple bands were not seen during restriction analysis of transformants arising on a *recA recD* strain. There was also no detectable evidence of satellite colonies and the numbers of transformants were comparatively much fewer.

#### **4.2. Location-specific recognition of *chi* sites**

The recognition of *chi* by RecBCD is of utmost importance during double strand break repair, due to the fact that *chi* sites modulate RecBCD activity and catalyze RecA loading onto the ssDNA filaments generated, thereby facilitating homologous recombination. It was therefore decided to test if the location of *chi* sites played a role in recognition by RecBCD. Only one study has focused on the importance of *chi* site location relative to the ends of dsDNA (Kulkarni & Julin, 2004), and that study was done using an *in vitro* system. Initially, the requirements of *chi* site location on dsDNA were assayed using a modified linear assay plasmid that had been PCR amplified with primers containing tandem *chi* sites at the ends of the primers, resulting in a linear plasmid molecule with correctly oriented tandem *chi* sites at the termini.

Using this substrate, all of the transformants obtained, irrespective of genetic background, were aberrantly re-circularized plasmids that had lost all or most of the region of homology flanking the dsDNA break site. This observation suggested that *chi* site(s) at the termini of dsDNA are not efficiently recognized by the RecBCD enzyme. The study by (Kulkarni & Julin, 2004) support this hypothesis. Their study shows that ssDNA oligonucleotides with the *chi* sequence at the ends of 14-mer or 20-mer DNA sequences failed to inhibit the exonuclease and helicase activities of RecBCD *in vitro* compared to 14 and 20-mer sequences with the *chi* site in the center. Thus far, this is the first *in vivo* study that demonstrates *chi* site recognition to be dependent upon location of the *chi* site with respect to a dsDNA break. To test the efficacy of *chi* sites internal to sites of dsDNA breaks, the (~2,100 bp) *sacB* gene was cloned between the two tandem *chi* sites present on the arms of homology. Using this assay plasmid, it was demonstrated that it was possible to obtain transformants that contain plasmids that have successfully recombined with the chromosomal target and re-circularized to generate the predicted recombinants.



### **4. 3. Role of RecG and RuvC during ends-in recombination**

One of the long term objectives for creating the ends in assay system was to determine if ends-in recombination could allow rescue of recombinants in the absence of Holliday Junction resolving pathways RuvABC and RecG. It has been previously shown (Lloyd, 1991) that *ruv* mutants of *E. coli* are recombination proficient and UV resistant to almost wild type levels, but when combined with a *recG* mutation, the recombinational proficiency and UV survival drops sharply. Thus, it was postulated that perhaps the RuvABC pathway has overlapping roles with RecG in resolving Holliday Junctions. Additional studies have also shown that once Holliday Junctions are formed and branch migration takes place by the RuvAB complex, the resolution of these Holliday Junctions formed depends on either RuvC or RecG (Muller & West, 1994) (Kuzminov, 1996). These resolution systems have different mechanisms of action. It has been demonstrated that the RuvC resolvase acts by endonucleolytic cleavage at Holliday Junctions. RecG is a branch translocase protein that is analogous to the RuvAB complex that could act to migrate DNA at Holliday Junctions. It was initially hypothesized that an as-yet unidentified endonuclease serves to resolve the Holliday Junctions in concert with RecG action. However, despite attempts to isolate such an enzyme, none have been identified (Zhang et al.). Current hypotheses suggest RecG does not participate in resolution, but rather its absence renders RuvABC activity essential for resolving recombination intermediates.

The first question that was addressed was – does the ends-in system require a functional RuvC protein to generate the correct recombined plasmid? Plasmid DNA extracted from the *ruvC*<sup>-</sup> transformants analyzed with restriction digestion showed that three out of twelve transformants had the correct restriction pattern. DNA sequencing confirmed that these three samples are indeed samples that have regained the missing *cfa* promoter region. This *ruvC*<sup>-</sup> mutation was also verified by transducing it into a *recG*<sup>-</sup> recipient strain and then measuring UV sensitivity of the resultant *ruvC*<sup>-</sup> *recG*<sup>-</sup> double mutant, which was as UV sensitive as a RecA-deficient strain. This shows that a functional RuvC protein was not required to generate the correct recombined plasmid.

Previous work by (Motamedi et al., 1999) has shown Holliday Junction processing in *E. coli* relies either on the break-join mechanism (mediated by RuvC and RecG) or the break-copy mechanism (which is independent of RuvC and RecG). According to their studies, it was possible to physically separate these two pathways by using cells deficient in the break-join

mechanism (i.e. *ruvC<sup>-</sup> recG<sup>-</sup>*). Hence, in order to ascertain if successful recombinant plasmid formation could occur in a strain lacking RuvC and RecG, a *ruvC<sup>-</sup> recG<sup>-</sup>* strain was transformed with the linear assay plasmid and the transformants obtained were analyzed using restriction digestion. It was found that even in a *ruvC<sup>-</sup> recG<sup>-</sup>* background, the linear plasmid was able to re-circularize after regaining the missing *cfa* promoter fragment. The transformants were also verified as being *ruvC<sup>-</sup> recG<sup>-</sup>* by UV sensitivity analysis (figure 3 – 28).

## **Future directions**

It is necessary to point out that while the assay system designed was a transcriptional fusion of *gfp* to the *cfa* region, evidently it did not work as expected. Following electroporation of plasmid molecules linearized in the *sacB* gene locus that contain tandem *chi* sites downstream of the regions of homology, transformants were observed that expressed *gfp* but were not the correct form of recombinants. Moreover, some recombinants that had gained the expected fragment of DNA did not express *gfp*. Following purification, some of these cells that previously expressed the green fluorescent protein seemed to lose the expression signal for *gfp* too. One reason for this could be that these plasmids could have re-circularized aberrantly and gained some fragment of DNA that has a functional promoter or gene coding sequence which would enable expression of *gfp*. Indeed, DNA sequencing analysis of plasmids recovered from transformants has revealed that these plasmids all possess random sequences of DNA between the arms of homology, which is very different from either the starting plasmid or the expected recombined plasmid. Some plasmids had lost both arms of homology, and the size of the plasmid was barely enough to contain an origin of replication and the *bla* gene. Other samples had lost one arm of homology, and the region of DNA upstream of the lost region of homology had somehow fused with part of the other region of homology. In short, there was rampant insertion, deletion and re-arrangement of almost all the plasmid samples recovered from the transformants, besides observing some transformant plasmids that had gained the correct fragment of DNA between the two arms of homology. Also, some plasmid samples recovered showed multiple restriction digestion patterns corresponding to either multiple plasmid molecules or high molecular weight linear concatemers. There has been documentation showing evidence that high molecular weight linear plasmid multimers are formed when plasmids containing *chi* sites are propagated inside *Escherichia coli* (Dabert et al., 1992), which would explain multiple bands seen during restriction digestion analysis. Unfortunately, it was not feasible to sequence the plasmid DNA isolated from all the transformants, but a few representative samples sequenced did show an abrupt termination of sequencing data and the DNA sequencing chromatogram sheet revealed the presence of multiple sequences of DNA (seen as multiple spikes of DNA sequence), implying that these transformants had more than one plasmid clone inside.

The ends-in system, which was initially designed as an *in vivo* linearization approach, was switched over to an *in vitro* based linearization. While this approach relieves the constraints

placed on the assay system by parameters like plasmid copy number and induction of I-SceI, it still does not accomplish the primary requirement of designing an assay system that can allow us to screen for the correctly recombined plasmid clone based on expression of *gfp*. Results show that certain transformants express *gfp* in spite of not having the correct fragment of *cfa* promoter DNA inserted between the two arms of homology, suggesting that although the assay plasmids are able to re-circularize following transformation, they generate a lot of background transformants that are somehow able to fluoresce. This renders screening of true (expected) recombinants via *gfp* expression totally unreliable. It is likely that the assay plasmid undergoes microhomology mediated recombination, allowing it to recombine with regions of DNA with non-specific homology to the arms of homology on the assay plasmid, thereby allowing the assay plasmid to re-circularize. If such a recombination event somehow leads to the insertion of a promoter between the two arms of homology on the assay plasmid, it might lead to promiscuous expression of *gfp*. One method to ensure that only true (expected) recombinants are able to express *gfp* would be to employ a translational fusion of *gfp* to the *cfa* promoter, instead of the transcriptional fusion that was used in this assay. The translational fusion would ensure a higher stringent control over promiscuous *gfp* expression from random re-circularization events. Another approach to correct the aberrant expression of *gfp* would be to clone the gene encoding *gfp* at a non-essential locus on the chromosome with a weak promoter upstream of the *gfp* gene. The two arms of homology could be designed such that they encode pieces of homology flanking the *gfp* gene on the chromosome. Following successful re-circularization of the (high copy number) assay plasmid, the assay plasmid would be expected to gain the missing *gfp* gene from the chromosome, thereby allowing a much more robust expression of *gfp* from the assay plasmid. Such recombination events could easily be visualized by screening for *gfp* expression of transformant colonies. This assay would ensure that non-specific recombination events arising due to microhomology mediated recombination would not promiscuously turn on *gfp* expression.

Once the assay system has been optimized, it can be used to test the genetic requirements for different recombination and replication restart proteins in *E. coli* and also other prokaryotes. For example, it is not yet clear how the three replication restart pathways PriA-PriB, PriA-PriC and Rep-PriC are co-ordinated by the cell to ensure faithful processing of all single stranded and double stranded DNA breaks. The nature of substrates that these replication pathways process is

also not clear, and using the ends-in assay system, the precise role of these replication pathways in processing different substrates might be elaborated.

## REFERENCES

- Amundsen, S. K., Neiman, A. M., Thibodeaux, S. M. & Smith, G. R. (1990) Genetic dissection of the biochemical activities of RecBCD enzyme. *Genetics*, **126**, 25-40.
- Amundsen, S. K., Taylor, A. F., Reddy, M. & Smith, G. R. (2007) Intersubunit signaling in RecBCD enzyme, a complex protein machine regulated by Chi hot spots. *Genes Dev*, **21**, 3296-3307.
- Bachmann, B. J. (1990) Linkage map of Escherichia coli K-12, edition 8. *Microbiol Rev*, **54**, 130-197.
- Biek, D. P. & Cohen, S. N. (1986) Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in Escherichia coli. *J Bacteriol*, **167**, 594-603.
- Birge, E. A. & Low, K. B. (1974) Detection of transcribable recombination products following conjugation in rec+, reCB- and recC-strains of Escherichia coli K12. *J Mol Biol*, **83**, 447-457.
- Bolt, E. L. & Lloyd, R. G. (2002) Substrate specificity of RusA resolvase reveals the DNA structures targeted by RuvAB and RecG in vivo. *Mol Cell*, **10**, 187-198.
- Bonner, C. A., Stukenberg, P. T., Rajagopalan, M., Eritja, R., O'Donnell, M., McEntee, K., Echols, H. & Goodman, M. F. (1992) Processive DNA synthesis by DNA polymerase II mediated by DNA polymerase III accessory proteins. *J Biol Chem*, **267**, 11431-11438.
- Bonura, T. & Smith, K. C. (1975) Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of Escherichia coli K-12. *J Bacteriol*, **121**, 511-517.
- Boonsombat, R., Yeh, S. P., Milne, A. & Sandler, S. J. (2006) A novel dnaC mutation that suppresses priB rep mutant phenotypes in Escherichia coli K-12. *Mol Microbiol*, **60**, 973-983.
- Braedt, G. & Smith, G. R. (1989) Strand specificity of DNA unwinding by RecBCD enzyme. *Proc Natl Acad Sci U S A*, **86**, 871-875.
- Brcic-Kostic, K., Salaj-Smic, E., Marsic, N., Kajic, S., Stojiljkovic, I. & Trgovcevic, Z. (1991) Interaction of RecBCD enzyme with DNA damaged by gamma radiation. *Mol Gen Genet*, **228**, 136-142.
- Bresler, S. E., Noskin, L. A. & Suslov, A. V. (1984) Induction by gamma irradiation of double-strand breaks of Escherichia coli chromosomes and their role in cell lethality. *Biophys J*, **45**, 749-754.
- Cahill, D., Connor, B. & Carney, J. P. (2006) Mechanisms of eukaryotic DNA double strand break repair. *Front Biosci*, **11**, 1958-1976.
- Centore, R. C., Lestini, R. & Sandler, S. J. (2008) XthA (Exonuclease III) regulates loading of RecA onto DNA substrates in log phase Escherichia coli cells. *Mol Microbiol*, **67**, 88-101.
- Cimino, G. D., Gamper, H. B., Isaacs, S. T. & Hearst, J. E. (1985) Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu Rev Biochem*, **54**, 1151-1193.
- Clark, A. J. & Margulies, A. D. (1965) Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12. *Proc Natl Acad Sci U S A*, **53**, 451-459.
- Clark, A. J. & Sandler, S. J. (1994) Homologous genetic recombination: the pieces begin to fall into place. *Crit Rev Microbiol*, **20**, 125-142.

- Cohen, A. & Clark, A. J. (1986) Synthesis of linear plasmid multimers in *Escherichia coli* K-12. *J Bacteriol*, **167**, 327-335.
- Conley, E. C. & Saunders, J. R. (1984) Recombination-dependent recircularization of linearized pBR322 plasmid DNA following transformation of *Escherichia coli*. *Mol Gen Genet*, **194**, 211-218.
- Conley, E. C., Saunders, V. A., Jackson, V. & Saunders, J. R. (1986a) Mechanism of intramolecular cyclization and deletion formation following transformation of *Escherichia coli* with linearized plasmid DNA. *Nucleic Acids Res*, **14**, 8919-8932.
- Conley, E. C., Saunders, V. A. & Saunders, J. R. (1986b) Deletion and rearrangement of plasmid DNA during transformation of *Escherichia coli* with linear plasmid molecules. *Nucleic Acids Res*, **14**, 8905-8917.
- Cox, M. M. (2002) The nonmutagenic repair of broken replication forks via recombination. *Mutat Res*, **510**, 107-120.
- Cromie, G. A., Connelly, J. C. & Leach, D. R. (2001) Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol Cell*, **8**, 1163-1174.
- Dabert, P., Ehrlich, S. D. & Gruss, A. (1992) Chi sequence protects against RecBCD degradation of DNA in vivo. *Proc Natl Acad Sci U S A*, **89**, 12073-12077.
- Dianov, G., Price, A. & Lindahl, T. (1992) Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Mol Cell Biol*, **12**, 1605-1612.
- Dickman, M. J., Ingleston, S. M., Sedelnikova, S. E., Rafferty, J. B., Lloyd, R. G., Grasby, J. A. & Hornby, D. P. (2002) The RuvABC resolvase. *Eur J Biochem*, **269**, 5492-5501.
- Dower, N. A. & Stahl, F. W. (1981) Chi activity during transduction-associated recombination. *Proc Natl Acad Sci U S A*, **78**, 7033-7037.
- Duncan, B. K. & Miller, J. H. (1980) Mutagenic deamination of cytosine residues in DNA. *Nature*, **287**, 560-561.
- Fiedler, S. & Wirth, R. (1988) Transformation of bacteria with plasmid DNA by electroporation. *Anal Biochem*, **170**, 38-44.
- Friedberg, E. C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
- Ganesan, A. K. (1974) Persistence of pyrimidine dimers during post-replication repair in ultraviolet light-irradiated *Escherichia coli* K12. *J Mol Biol*, **87**, 103-119.
- Ganesan, A. K. & Seawell, P. C. (1975) The effect of *lexA* and *recF* mutations on post-replication repair and DNA synthesis in *Escherichia coli* K-12. *Mol Gen Genet*, **141**, 189-205.
- George, J., Castellazzi, M. & Buttin, G. (1975) Prophage induction and cell division in *E. coli*. III. Mutations *sfiA* and *sfiB* restore division in *tif* and *lon* strains and permit the expression of mutator properties of *tif*. *Mol Gen Genet*, **140**, 309-332.
- Gumbiner-Russo, L. M., Lombardo, M. J., Ponder, R. G. & Rosenberg, S. M. (2001) The TGV transgenic vectors for single-copy gene expression from the *Escherichia coli* chromosome. *Gene*, **273**, 97-104.
- Hannan, M. A., Zimmer, S. G. & Hazle, J. (1984) Mechanisms of cisplatin (cis-diamminodichloroplatinum II)-induced cytotoxicity and genotoxicity in yeast. *Mutat Res*, **127**, 23-30.
- Harinarayanan, R. & Gowrishankar, J. (2004) A *dnaC* mutation in *Escherichia coli* that affects copy number of ColE1-like plasmids and the PriA-PriB (but not Rep-PriC) pathway of chromosomal replication restart. *Genetics*, **166**, 1165-1176.

- Heller, R. C. & Marians, K. J. (2005a) The disposition of nascent strands at stalled replication forks dictates the pathway of replisome loading during restart. *Mol Cell*, **17**, 733-743.
- Heller, R. C. & Marians, K. J. (2005b) Unwinding of the nascent lagging strand by Rep and PriA enables the direct restart of stalled replication forks. *J Biol Chem*, **280**, 34143-34151.
- Hevroni, D. & Livneh, Z. (1988) Bypass and termination at apurinic sites during replication of single-stranded DNA in vitro: a model for apurinic site mutagenesis. *Proc Natl Acad Sci U S A*, **85**, 5046-5050.
- Holton, T. A. & Graham, M. W. (1991) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res*, **19**, 1156.
- Horii, Z. & 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**, 327-344.
- Horsfall, M. J. & Lawrence, C. W. (1994) Accuracy of replication past the T-C (6-4) adduct. *J Mol Biol*, **235**, 465-471.
- Howard-Flanders, P., Theriot, L. & Stedeford, J. B. (1969) Some properties of excision-defective recombination-deficient mutants of Escherichia coli K-12. *J Bacteriol*, **97**, 1134-1141.
- Huisman, O., D'Ari, R. & Gottesman, S. (1984) Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc Natl Acad Sci U S A*, **81**, 4490-4494.
- Imlay, J. A. & Fridovich, I. (1991) Assay of metabolic superoxide production in Escherichia coli. *J Biol Chem*, **266**, 6957-6965.
- Inoue, J., Honda, M., Ikawa, S., Shibata, T. & Mikawa, T. (2008) The process of displacing the single-stranded DNA-binding protein from single-stranded DNA by RecO and RecR proteins. *Nucleic Acids Res*, **36**, 94-109.
- Iyer, V. N. & Rupp, W. D. (1971) Usefulness of benzoylated naphthoylated DEAE-cellulose to distinguish and fractionate double-stranded DNA bearing different extents of single-stranded regions. *Biochim Biophys Acta*, **228**, 117-126.
- Johnson, B. F. (1977) Fine structure mapping and properties of mutations suppressing the lon mutation in Escherichia coli K-12 and B strains. *Genet Res*, **30**, 273-286.
- Kastan, M. B. & Kuerbitz, S. J. (1993) Control of G1 arrest after DNA damage. *Environ Health Perspect*, **101 Suppl 5**, 55-58.
- Kellenberger, G., Zichichi, M. L. & Weigle, J. J. (1961) Exchange of DNA in the recombination of bacteriophage lambda. *Proc Natl Acad Sci U S A*, **47**, 869-878.
- Kleckner, N., Bender, J. & Gottesman, S. (1991) Uses of transposons with emphasis on Tn10. *Methods Enzymol*, **204**, 139-180.
- Kohn, K. W., Spears, C. L. & Doty, P. (1966) Inter-strand crosslinking of DNA by nitrogen mustard. *J Mol Biol*, **19**, 266-288.
- Kohn, K. W., Steigbigel, N. H. & Spears, C. L. (1965) Cross-linking and repair of DNA in sensitive and resistant strains of E. coli treated with nitrogen mustard. *Proc Natl Acad Sci U S A*, **53**, 1154-1161.
- Kowalczykowski, S. C. (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci*, **25**, 156-165.
- Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994) Biochemistry of homologous recombination in Escherichia coli. *Microbiol Rev*, **58**, 401-465.



- Kulkarni, A. & Julin, D. A. (2004) Specific inhibition of the E.coli RecBCD enzyme by Chi sequences in single-stranded oligodeoxyribonucleotides. *Nucleic Acids Res*, **32**, 3672-3682.
- Kuzminov, A. (1996) Unraveling the late stages of recombinational repair: metabolism of DNA junctions in Escherichia coli. *Bioessays*, **18**, 757-765.
- Kuzminov, A. (1999) Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol Mol Biol Rev*, **63**, 751-813, table of contents.
- Kuzminov, A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A*, **98**, 8241-8246.
- Kuzminov, A. & Stahl, F. W. (1997) Stability of linear DNA in recA mutant Escherichia coli cells reflects ongoing chromosomal DNA degradation. *J Bacteriol*, **179**, 880-888.
- Lawrence, C. W., Gibbs, P. E., Borden, A., Horsfall, M. J. & Kilbey, B. J. (1993) Mutagenesis induced by single UV photoproducts in E. coli and yeast. *Mutat Res*, **299**, 157-163.
- Lederberg, J. & Tatum, E. L. (1946) Gene recombination in Escherichia coli. *Nature*, **158**, 558.
- Lee, E. H., Masai, H., Allen, G. C., Jr. & Kornberg, A. (1990) The priA gene encoding the primosomal replicative n' protein of Escherichia coli. *Proc Natl Acad Sci U S A*, **87**, 4620-4624.
- Lennox, E. S. (1955) Transduction of linked genetic characters of the host by bacteriophage P1. *Virology*, **1**, 190-206.
- Liberi, G., Cotta-Ramusino, C., Lopes, M., Sogo, J., Conti, C., Bensimon, A. & Foiani, M. (2006) Methods to study replication fork collapse in budding yeast. *Methods Enzymol*, **409**, 442-462.
- Lin, P. F. & Howard-Flanders, P. (1976) Genetic exchanges caused by ultraviolet photoproducts in phage lambda DNA molecules: the role of DNA replication. *Mol Gen Genet*, **146**, 107-115.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
- Lindahl, T., Sedgwick, B., Sekiguchi, M. & Nakabeppu, Y. (1988) Regulation and expression of the adaptive response to alkylating agents. *Annu Rev Biochem*, **57**, 133-157.
- Lloyd, R. G. (1991) Conjugational recombination in resolvase-deficient ruvC mutants of Escherichia coli K-12 depends on recG. *J Bacteriol*, **173**, 5414-5418.
- Lloyd, R. G., Benson, F. E. & Shurvinton, C. E. (1984) Effect of ruv mutations on recombination and DNA repair in Escherichia coli K12. *Mol Gen Genet*, **194**, 303-309.
- Lloyd, R. G. & Sharples, G. J. (1993a) Dissociation of synthetic Holliday junctions by E. coli RecG protein. *Embo J*, **12**, 17-22.
- Lloyd, R. G. & Sharples, G. J. (1993b) Processing of recombination intermediates by the RecG and RuvAB proteins of Escherichia coli. *Nucleic Acids Res*, **21**, 1719-1725.
- Lopes, M., Cotta-Ramusino, C., Pellicoli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C. S. & Foiani, M. (2001) The DNA replication checkpoint response stabilizes stalled replication forks. *Nature*, **412**, 557-561.
- Macquet, J. P., Jankowski, K. & Butour, J. L. (1980) Mass spectrometry study of DNA-cisplatin complexes: perturbation of guanine-cytosine base-pairs. *Biochem Biophys Res Commun*, **92**, 68-74.
- Maity, A., McKenna, W. G. & Muschel, R. J. (1994) The molecular basis for cell cycle delays following ionizing radiation: a review. *Radiother Oncol*, **31**, 1-13.

- Mandal, T. N., Mahdi, A. A., Sharples, G. J. & Lloyd, R. G. (1993) Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of *ruvA*, *ruvB*, and *ruvC* mutations. *J Bacteriol*, **175**, 4325-4334.
- Marians, K. J. (1999) PriA: at the crossroads of DNA replication and recombination. *Prog Nucleic Acid Res Mol Biol*, **63**, 39-67.
- Marians, K. J. (2004) Mechanisms of replication fork restart in *Escherichia coli*. *Philos Trans R Soc Lond B Biol Sci*, **359**, 71-77.
- Masai, H., Asai, T., Kubota, Y., Arai, K. & Kogoma, T. (1994) *Escherichia coli* PriA protein is essential for inducible and constitutive stable DNA replication. *Embo J*, **13**, 5338-5345.
- Masters, M. (1977) The frequency of P1 transduction of the genes of *Escherichia coli* as a function of chromosomal position: preferential transduction of the origin of replication. *Mol Gen Genet*, **155**, 197-202.
- McCool, J. D. & Sandler, S. J. (2001) Effects of mutations involving cell division, recombination, and chromosome dimer resolution on a *priA2::kan* mutant. *Proc Natl Acad Sci U S A*, **98**, 8203-8210.
- McMilin, K. D. & Russo, V. E. (1972) Maturation and recombination of bacteriophage lambda DNA molecules in the absence of DNA duplication. *J Mol Biol*, **68**, 49-55.
- Meselson, M. (1964) On the Mechanism of Genetic Recombination between DNA Molecules. *J Mol Biol*, **9**, 734-745.
- Meselson, M. & Weigle, J. J. (1961) Chromosome breakage accompanying genetic recombination in bacteriophage. *Proc Natl Acad Sci U S A*, **47**, 857-868.
- Michel, B., Grompone, G., Flores, M. J. & Bidnenko, V. (2004) Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A*, **101**, 12783-12788.
- Modrich, P. (1994) Mismatch repair, genetic stability, and cancer. *Science*, **266**, 1959-1960.
- Mosig, G. (1987) The essential role of recombination in phage T4 growth. *Annu Rev Genet*, **21**, 347-371.
- Motamedi, M. R., Szigety, S. K. & Rosenberg, S. M. (1999) Double-strand-break repair recombination in *Escherichia coli*: physical evidence for a DNA replication mechanism in vivo. *Genes Dev*, **13**, 2889-2903.
- Muller, B. & West, S. C. (1994) Processing of Holliday junctions by the *Escherichia coli* RuvA, RuvB, RuvC and RecG proteins. *Experientia*, **50**, 216-222.
- Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y. & Sekiguchi, M. (1988) Expression of the *ada* gene of *Escherichia coli* in response to alkylating agents. Identification of transcriptional regulatory elements. *J Mol Biol*, **202**, 483-494.
- Newman, B. J. & Masters, M. (1980) The variation in frequency with which markers are transduced by phage P1 is primarily a result of discrimination during recombination. *Mol Gen Genet*, **180**, 585-589.
- Nowosielska, A. & Marinus, M. G. (2008) DNA mismatch repair-induced double-strand breaks. *DNA Repair (Amst)*, **7**, 48-56.
- Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z. & Woodgate, R. (2001) The Y-family of DNA polymerases. *Mol Cell*, **8**, 7-8.
- Preston, B. D., Singer, B. & Loeb, L. A. (1986) Mutagenic potential of O4-methylthymine in vivo determined by an enzymatic approach to site-specific mutagenesis. *Proc Natl Acad Sci U S A*, **83**, 8501-8505.

- Radman, M. (1975) SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci*, **5A**, 355-367.
- Radman, M., Cordone, L., Krsmanovic-Simic, D. & Errera, M. (1970) Complementary action of recombination and excision in the repair of ultraviolet irradiation damage to DNA. *J Mol Biol*, **49**, 203-212.
- Rothman, R. H., Kato, T. & Clark, A. J. (1975) The beginning of an investigation of the role of recF in the pathways of metabolism of ultraviolet-irradiated DNA in Escherichia coli. *Basic Life Sci*, **5A**, 283-291.
- Rupp, W. D. & Howard-Flanders, P. (1968) Discontinuities in the DNA synthesized in an excision-defective strain of Escherichia coli following ultraviolet irradiation. *J Mol Biol*, **31**, 291-304.
- Rupp, W. D., Wilde, C. E., 3rd, Reno, D. L. & Howard-Flanders, P. (1971) Exchanges between DNA strands in ultraviolet-irradiated Escherichia coli. *J Mol Biol*, **61**, 25-44.
- Sakai, A., Nakanishi, M., Yoshiyama, K. & Maki, H. (2006) Impact of reactive oxygen species on spontaneous mutagenesis in Escherichia coli. *Genes Cells*, **11**, 767-778.
- Sancar, A. (1994a) Mechanisms of DNA excision repair. *Science*, **266**, 1954-1956.
- Sancar, A. (1994b) Structure and function of DNA photolyase. *Biochemistry*, **33**, 2-9.
- Sancar, A. & Rupp, W. D. (1983) A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell*, **33**, 249-260.
- Sandler, S. J. (2000) Multiple genetic pathways for restarting DNA replication forks in Escherichia coli K-12. *Genetics*, **155**, 487-497.
- Sandler, S. J., Marians, K. J., Zavitz, K. H., Coutu, J., Parent, M. A. & Clark, A. J. (1999) dnaC mutations suppress defects in DNA replication- and recombination-associated functions in priB and priC double mutants in Escherichia coli K-12. *Mol Microbiol*, **34**, 91-101.
- Sandler, S. J., Samra, H. S. & 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.
- Selby, C. P. & Sancar, A. (1994) Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol Rev*, **58**, 317-329.
- Sharples, G. J., Chan, S. N., Mahdi, A. A., Whitby, M. C. & Lloyd, R. G. (1994) Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. *Embo J*, **13**, 6133-6142.
- Shevell, D. E., Friedman, B. M. & Walker, G. C. (1990) Resistance to alkylation damage in Escherichia coli: role of the Ada protein in induction of the adaptive response. *Mutat Res*, **233**, 53-72.
- Shiba, T., Iwasaki, H., Nakata, A. & 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**, 8445-8449.
- Shida, T., Iwasaki, H., Saito, A., Kyogoku, Y. & Shinagawa, H. (1996) Analysis of substrate specificity of the RuvC holliday junction resolvase with synthetic Holliday junctions. *J Biol Chem*, **271**, 26105-26109.
- Sinden, R. R. & Cole, R. S. (1978) Repair of cross-linked DNA and survival of Escherichia coli treated with psoralen and light: effects of mutations influencing genetic recombination and DNA metabolism. *J Bacteriol*, **136**, 538-547.

- Smith, G. R. (1989) Homologous recombination in *E. coli*: multiple pathways for multiple reasons. *Cell*, **58**, 807-809.
- Smith, K. C. & 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**, 459-472.
- Stahl, F. W. & Stahl, M. M. (1977) Recombination pathway specificity of Chi. *Genetics*, **86**, 715-725.
- Steenken, S. (1989) Structure, acid/base properties and transformation reactions of purine radicals. *Free Radic Res Commun*, **6**, 117-120.
- Takeshita, M. & Eisenberg, W. (1994) Mechanism of mutation on DNA templates containing synthetic abasic sites: study with a double strand vector. *Nucleic Acids Res*, **22**, 1897-1902.
- Taketo, A. (1988) DNA transfection of *Escherichia coli* by electroporation. *Biochim Biophys Acta*, **949**, 318-324.
- Taylor, A. F. & Smith, G. R. (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature*, **423**, 889-893.
- Teoule, R. (1987) Radiation-induced DNA damage and its repair. *Int J Radiat Biol Relat Stud Phys Chem Med*, **51**, 573-589.
- Tercero, J. A. & Diffley, J. F. (2001) Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature*, **412**, 553-557.
- Thaler, D. S. & Stahl, F. W. (1988) DNA double-chain breaks in recombination of phage lambda and of yeast. *Annu Rev Genet*, **22**, 169-197.
- Tomasz, M., Lipman, R., Chowdary, D., Pawlak, J., Verdine, G. L. & Nakanishi, K. (1987) Isolation and structure of a covalent cross-link adduct between mitomycin C and DNA. *Science*, **235**, 1204-1208.
- Tsaneva, I. R., Illing, G., Lloyd, R. G. & West, S. C. (1992) Purification and properties of the RuvA and RuvB proteins of *Escherichia coli*. *Mol Gen Genet*, **235**, 1-10.
- Wang, T. C. & Chen, S. H. (1992) Similar-sized daughter-strand gaps are produced in the leading and lagging strands of DNA in UV-irradiated *E. coli* uvrA cells. *Biochem Biophys Res Commun*, **184**, 1496-1503.
- Wang, T. C. & Smith, K. C. (1986) Postreplicational formation and repair of DNA double-strand breaks in UV-irradiated *Escherichia coli* uvrB cells. *Mutat Res*, **165**, 39-44.
- Ward, J. F. (1988) DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog Nucleic Acid Res Mol Biol*, **35**, 95-125.
- West, S. C. (1992) Enzymes and molecular mechanisms of genetic recombination. *Annu Rev Biochem*, **61**, 603-640.
- Whitby, M. C. & Lloyd, R. G. (1998) Targeting Holliday junctions by the RecG branch migration protein of *Escherichia coli*. *J Biol Chem*, **273**, 19729-19739.
- Whitby, M. C., Ryder, L. & 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**, 341-350.
- Wilkins, B. M. & Howard-Flanders, P. (1968) The genetic properties of DNA transferred from ultraviolet-irradiated Hfr cells of *Escherichia coli* K-12 during mating. *Genetics*, **60**, 243-255.
- Woodgate, R. (2001) Evolution of the two-step model for UV-mutagenesis. *Mutat Res*, **485**, 83-92.

- Yeung, A. T., Mattes, W. B., Oh, E. Y. & Grossman, L. (1983) Enzymatic properties of purified Escherichia coli uvrABC proteins. *Proc Natl Acad Sci U S A*, **80**, 6157-6161.
- Zavitz, K. H., DiGate, R. J. & Marians, K. J. (1991) The priB and priC replication proteins of Escherichia coli. Genes, DNA sequence, overexpression, and purification. *J Biol Chem*, **266**, 13988-13995.
- Zhang, J., Mahdi, A. A., Briggs, G. S. & Lloyd, R. G. Promoting and avoiding recombination: contrasting activities of the Escherichia coli RuvABC Holliday junction resolvase and RecG DNA translocase. *Genetics*, **185**, 23-37.