MQP–BC–KNW–0035

THE ROLE OF RECN IN DNA RECOMBINATION AND REPAIR IN *ESCHERICHIA COLI*

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor in Science

in

Biochemistry

by

Christine Feulner

March 10, 2008

APPROVED:

Kenan Murphy, Ph.D. Molecular Genetics and Microbiology UMASS Medical School Major Advisor	Kristin Wobbe, Ph.D. Chemistry/Biochemistry WPI Project Advisor

ABSTRACT

The ability to repair damaged DNA is critical to cell survival. A double-stranded DNA (dsDNA) break is one of the key lesions that must be repaired in order for cells to survive. In *Escherichia coli*, dsDNA breaks are repaired by multiple pathways of DNA recombination. DNA breaks induce the SOS response, where a set of 40 to 50 genes are induced in response to DNA damage. Among these genes is the *recN* function, which encodes a protein belonging to a class of proteins known to be important for the structural maintenance of chromosomes (SMC proteins). The aim of this project was to determine the role of the RecN protein in *Escherichia coli* in repairing a dsDNA break. A plasmid-based gap repair assay was developed in order to test the recombinational repair function of the RecN protein in vivo. The results concluded that RecN was not needed for the repair of a dsDNA break when the break was repaired by the λ Red pathway of recombination. In addition, it was found that the RecA protein, a function important in virtually all known recombination repair pathways in *Escherichia coli*, was actually inhibitory to the Red-promoted repair of dsDNA breaks, suggesting a RecA-independent mode dsDNA break repair by the λ Red pathway.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my gratitude towards Dr. Kenan Murphy for the idea behind the project, the use his laboratory for performing the necessary experiments, and his encouragement throughout the project. Thanks are also in order to Dr. Anthony Poteete and Dr. Michael Volkert for their helpful ideas and assistance. Finally, I would like to thank Dr. Kristin Wobbe for her role as advisor on this project.

BACKGROUND INFORMATION

The integrity of genetic information is vital for cell survival. There are certain instances; however, when the genetic information of a cell is in danger. Exposure to DNA damaging agents such as ultraviolet light, gamma radiation, or chemical mutagens can alter the structure and function of DNA. One of the major threats to DNA from these exposures is the generation of double-stranded breaks (DSBs), which if not repaired, can lead to chromosomal abnormalities and/or cell death. Numerous cellular recombination and repair systems make it possible for all cells to repair the damage and accurately restore the DNA sequence to its original state.

An established repair and recombination system is necessary for survival for all organisms. The elaborate mechanisms of mending damaged DNA require a number of important recombination and repair proteins, each of which executes a distinctive function. The recombination and repair protein examined in this study is the *Escherichia coli* RecN protein, involved in the repair of DSBs. In order to understand the putative role of RecN in vivo, it is important to review the recombinational repair mechanisms known to be active in *Escherichia coli* that work on DSBs.

Repair of Double-Stranded DNA Breaks

It is vital that the cell contain repair mechanisms to restore the integrity of its replication forks. A stalled or collapsed replication fork can lead to the formation of DSB, which if not repaired can be lethal to the cell. Collapsed replication forks can occur spontaneously during DNA replication, which occurs at least once per generation (Voet et al., 2006). Alternatively, stalled replication forks can occur when the replisome encounters chemical or radiational-induced DNA damage. The blockage of the replisome at the site of damage can lead to the

generation of a DSB. These events are potentially lethal since a DSB can disrupt the coding sequence of a gene, disturb the linkage between coding and regulatory sequences, and alter chromosome organization. In addition, DSBs can disrupt DNA replication, chromosomal packaging, and chromosome segregation (Cromie et al., 2001). If not repaired, DSBs can lead to DNA rearrangements, such as translocations, inversions, and chromosomal deletions. Additionally, DSBs produce substrates for single- and double-strand exonucleases. These degradation enzymes can cause loss of significant genetic information if the lesion is not corrected.

Two Pathways for DSB Repair

There are two major pathways for repairing a DSB: non-homologous end-joining (NHEJ) and homologous recombinational repair (HR). NHEJ leads to the joining of two broken DNA ends, usually at sequences that contains little to no homology (Lieber, 1999 and Connelly et al., 2002). The mechanism of NHEJ is shown in Figure 1A and 1B.

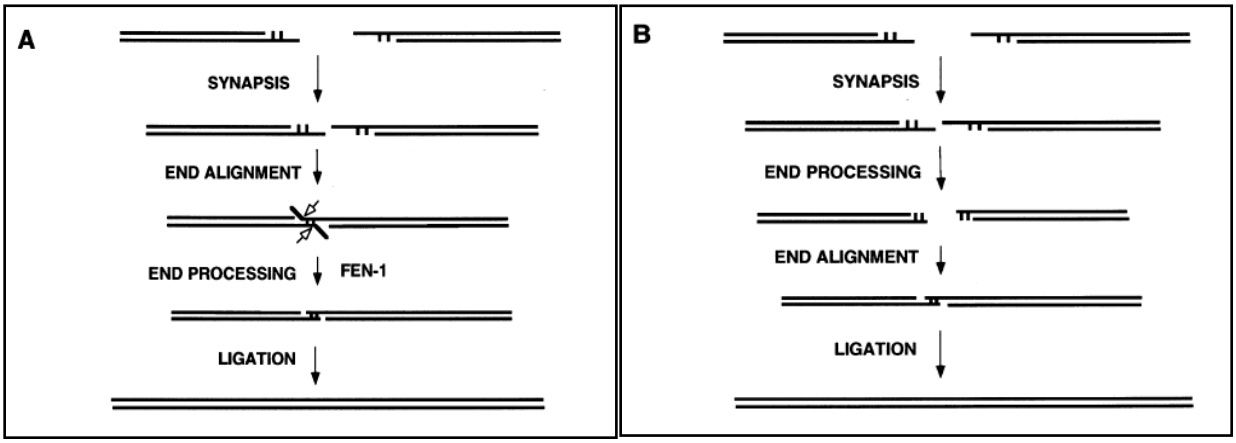


Figure 1A and 1B. Models for NHEJ. Figure 1A demonstrates the recombination mechanism when DSBs produce 5' or 3' overhangs. Figure 1B exhibits the same mechanism as Figure 1A where the difference is the removal of nucleotides at the overhangs by exonucleases (Lieber, 1999).

Genetic information can be lost if the alterations prior to end-joining eliminate a large part of DNA in order to rejoin the strands. NHEJ is thus error prone and can result in a lack of genomic integrity from generation to generation. While a common mechanism in eukaryotic cells, NHEJ does not occur in bacterial cells, except under special circumstances (Lieber, 1999). NHEJ has not been detected in *Escherichia coli*.

The second major pathway for DNA repair of DSBs occurs by the process of HR. During HR-promoted repair of DSBs, the ends of the break are aligned with an undamaged homolog (for example, a sister chromosome) and DNA strands are exchanged based on identical sequence homology. The undamaged template is used to restore DNA sequences that have been lost in or around the DNA break (Connelly et al., 2002). The recombinational mechanism can be viewed in Figure 2.

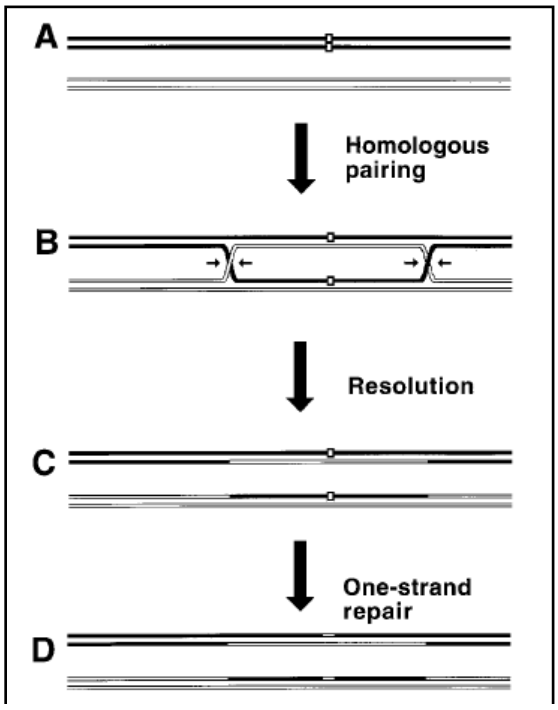


Figure 2. Model for HR. (A) DNA molecule with a two-strand lesion is shown (small rectangular boxes) aligned next to its homolog. (B) The two sequences exchanged strands at homologous regions of the DNA. (C) The two chromosomes are separated from each other by junction resolution. (D) The completed recombinational repair reaction with its final recombinant product (Kuzminov, 1999) is shown.

Integrity of the genetic content is maintained using HR because lost information is reinstated using the sequences within the complementary DNA strand. *Escherichia coli* cells perform HR to preserve its genetic integrity following damage to its chromosome.

Pathways of Homologous Recombination in *Escherichia coli*

The RecBCD pathway is the main HR pathway utilized in *E. coli*. The RecBCD enzyme is an ATP-dependent dsDNA exonuclease/helicase that acts on dsDNA to generate 3' single-stranded DNA (ssDNA) ends. It also plays a role in the loading of RecA onto these ssDNA ends. If the RecBCD pathway is unavailable (for example, if the *recBCD* genes are mutated), other pathways can be activated. The key alternate pathway of HR is *E. coli* is the RecF pathway, which can be defined as such since the *recF* gene was the first one to be shown important for this pathway. While considered a major pathway for the repair of ssDNA gaps in replicating DNA, the RecFOR pathway can perform dsDNA break repair as well. Both the RecBCD and RecF pathways are equally important for DNA damage repair. A third pathway for recombination in *E. coli* is known as the λ Red pathway. In this system, the bacteriophage λ Red recombination system can substitute for RecBCD (Murphy, 1998). The Red system is highly recombinogenic and has been used recently as an important tool for generating gene replacements in *E. coli* at high frequency (Murphy, 1998; Zhang et al., 1998; Yu et al., 2000; Datsenko and Wanner, 2000). Historically, the RecN recombination function (the subject of this study) is not required for HR via the RecBCD pathway, but is required for the RecFOR pathway. Its role in the Red recombination system has not been examined and is the subject of study in this project.

RecBCD Pathway

The principal damage recognized by the RecBCD pathway is DSBs. In *Escherichia coli*, RecBCD binds to the broken DNA ends and generates a 3' ssDNA overhangs by the action of its exonuclease/helicase activities (Kidane et al., 2004). How this end processing is accomplished is well defined with each of the three protein subunits in the RecBCD complex possessing a specific function. RecB is a 3'-5' helicase and a multifunctional nuclease, RecC recognizes the 8 base pair sequence known as Chi, and RecD is a 5'-3' helicase (Singleton et al., 2004). The exonuclease/helicase activities of RecBCD are modified by an encounter with a Chi site. The Chi site (5'-GCTGGTGG-3') is recognized as ssDNA within the double-stranded context as the DNA duplex is being unraveled (Singleton et al., 2004). After reaching the Chi site, the 3' exonuclease activity of RecBCD is suppressed, while the 5' exonuclease activity remains intact. Further unwinding at this point generates a ssDNA tail with a 3' end. Details of this mechanism can be observed in Figure 3.

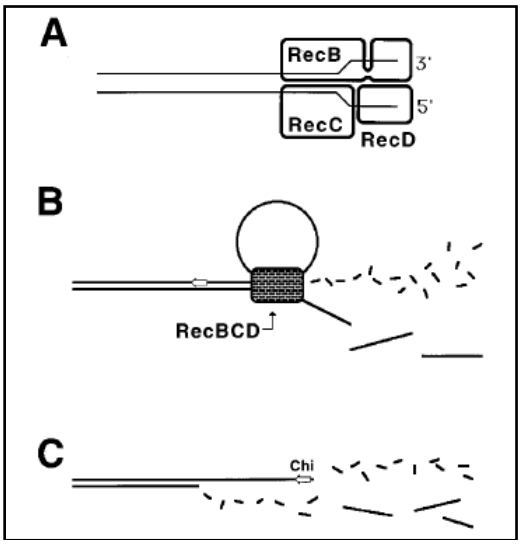


Figure 3. Damage recognition and end processing in the RecBCD Pathway. (A) RecBCD binds to a double-stranded DNA end and unwinds DNA using its helicase activities. (B) RecBCD promotes the degradation of DNA before reaching the Chi site, with higher rates of nicking activity on the 3' ended strand. After reaching the Chi site (C), degradation of the 3' strand is inhibited, while the 5'-3' end continues to be degraded by RecBCD, which is not shown in the figure (Kuzminov, 1999).

The key step of all HR systems is the invasion of a 3' ssDNA end into the homologous duplex (Kidane et al., 2004). In the RecBCD pathway, the RecBCD enzyme loads the RecA protein onto the 3' end. RecA is a single-stranded DNA binding protein that catalyzes ssDNA strand invasion and strand exchange, and is the major player in virtually all recombination pathways in *E. coli*. This terminus coated with RecA will invade the homologous duplex to initiate HR.

The final steps of HR are comprised of the extension of the DNA heteroduplex and its resolution. Following RecA-promoted strand invasion and exchange, a Holliday junction (four-stranded crossover structure of DNA) is generated by the action of RecG and/or RuvAB helicases. The RuvA protein forms tetramers that bind to the Holliday junction, and the RuvB protein binds duplex DNA as a hexamer (Kuzminov, 1999). RuvA and RuvB are known to form a complex that promotes branch migration. The Holliday junction is resolved by RuvC, and DNA ligase seals the necessary nicks (Kidane et al., 2004). Figure 4 demonstrates this last stage of HR.

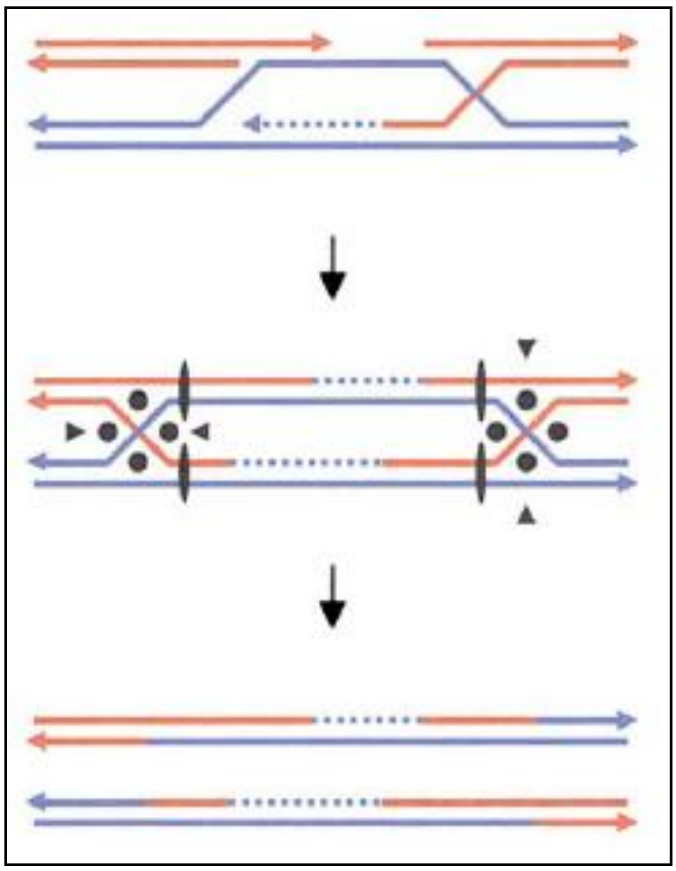


Figure 4. Formation and resolution of Holliday Junction. The first sketch demonstrates the generation of a double Holliday Junction. This Holliday Junction is resolved in the second sketch by the RuvABC resolvasome. RuvA is represented as the circles, RuvB as ovals, and RuvC as triangles. The resolved crossover products are shown in the third sketch (Cromie et al., 2001).

RecFOR Pathway

In the absence of RecBCD in *recBC sbcBC* mutant cells, the RecFOR pathway is activated (Sanchez et al., 2006). In this pathway, many genes not required in the RecBCD pathway now play a major role in recombinational repair. These genes include *recF*, *recO*, *recR*, *recQ*, *recJ*, and *recN*. To replace the function of the RecBCD protein, there are two different proteins available in the RecFOR pathway. These two proteins are important for creating the 3' ssDNA end that is bound by RecA: RecQ and RecJ. RecQ is a helicase that unwinds the duplex DNA end, where as RecJ acts on the unwound duplex and degrades the ssDNA with a 5'-to-3' polarity (Sanchez et al., 2006) generating a 3' ssDNA end. Unlike the RecBCD pathway, there

is not a specific DNA sequence that modulates RecJ exonuclease activity. It is thought that the affinity for DNA decreases as RecJ digests DNA, which causes the protein to dissociate from ssDNA.

Once the 3' ssDNA overhang is generated, the free 3' ssDNA end is coated with single-stranded DNA binding protein (SSB). The proteins, RecF, RecO, and RecR, promote displacement of SSB and load RecA onto the 3' ssDNA ends (Rocha et al., 2005). In an ATP-dependent manner, the RecA-ssDNA filament invades an homologous DNA duplex, where it anneals to the complementary strand. Following this step, the RecBCD and RecFOR pathways both follow identical paths, using RuvAB to generate a Holliday junction and RuvC to resolve the intermediate into recombinant products.

The contrasts and similarities between the RecBCD and the RecFOR recombination repair pathways can be observed in Figure 5. These two pathways can repair many lesions on one strand of DNA by making use of its corresponding homolog.

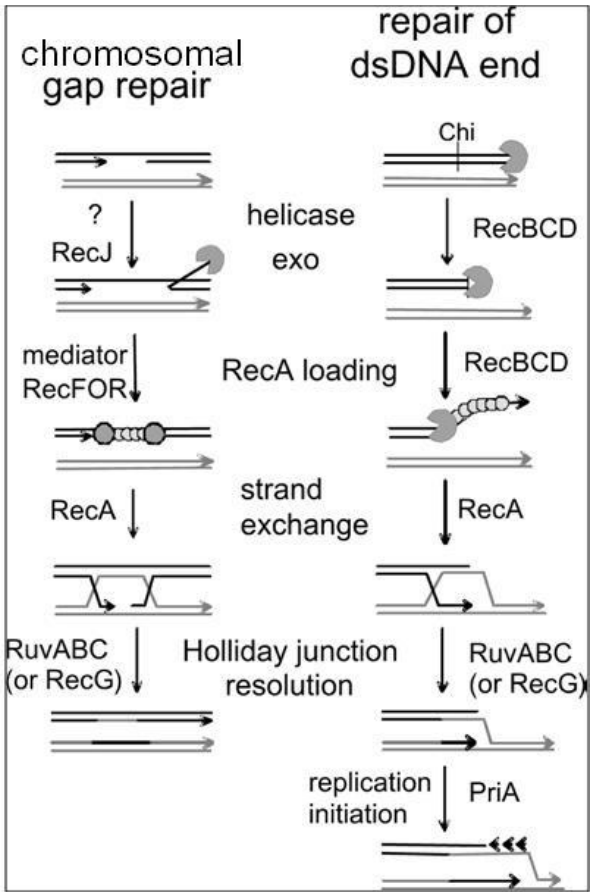


Figure 5. Comparison between HR using the RecFOR pathway (left) and the RecBCD pathway (right) (Rocha et al., 2005).

λ Red Pathway

The λ Red pathway takes precedent in *E. coli* cells during infection of *E. coli* with bacteriophage λ. During phage infection, the RecBCD pathway is inhibited by the λ Gam protein associated with the λ Red pathway, where the λ Gam protein inhibits the exonuclease function of the RecBCD protein (Murphy, 1990, 2007). The λ Red pathway consists of two major proteins: Exo and Bet. With the inhibition of the RecBCD enzyme, it is possible for these two phage proteins to access the dsDNA ends. Exo protein is a 5'-3' exonuclease that binds to dsDNA ends and digests the 5' strand, while the Bet protein is capable of annealing complementary ssDNA strands (Carter D. M. et al, 1971; Kmiec, E. et al., 1981; Little, J. W.,

1967; Muniyappa K. et al, 1986). Once the Bet protein has coated the ssDNA, there are both RecA-dependent and RecA-independent mechanisms by which this pathway can promote recombinational repair. It has been suggested that in the RecA-dependent pathway, RecFOR proteins are required to remove Bet and replace it with RecA (a role similar to the role of these proteins in replacing with RecA in the RecFOR pathway) (Poteete, A. R., 2004). The RecA-independent pathway has not been characterized.

SOS Response

While many recombination genes are expressed constitutively in *E. coli*, some repair proteins are turned on by the SOS response. The SOS regulatory system of *Escherichia coli* consists of a set of operons that are coordinately induced when DNA damage occurs (Finch et al., 1985). This chromosomal damage, which includes such modifications as methylated bases and thymidime dimers, can be seen following exposure of cells to alkylating agents, UV irradiation, or oxygen-derived free radicals. Other DNA damaging agents, such as gamma radiation and nalidixic acid, lead specifically to the generation of DSBs. The main goal of this emergency response is to repair a large amount of damage in the shortest time possible. If broken DNA is present in the cell for an extended period of time, potential degradation can occur through the action of exonucleases. Unrepaired DSBs are lethal in *E. coli*.

Two regulator genes of the SOS response are *recA* and *lexA*. In an uninduced system, the LexA protein represses the SOS response by binding to the operator sequence in front of each gene (Finch et al., 1985). However, in the presence of DNA damage, an intracellular inducing signal (most likely ssDNA) interacts with RecA, which is reversibly altered to an activated form (RecA*) that can mediate LexA repressor cleavage, thus derepressing the SOS regulon (Peterson

et al., 1988). Some of these LexA-controlled SOS genes include functions important for nucleotide excision repair (*uvrA*, *uvrB*), recombinational repair (*ruvAB*) and mutagenic replicative bypass (*umuC* and *umuD*). Also included in the SOS response is the *recN* gene, a function required for the normal repair of DSBs under the SOS response (Finch et al., 1985). What the actual function of RecN is in DSB repair is not known, and is the subject of this report.

Function and Structure of RecN

recN is one of the numerous SOS genes present on the *Escherichia coli* chromosome. The promoter for the *recN* gene contains recognizable -10 and -35 regions similar to the consensus promoter sequence for known *E. coli* promoters, which overlap two 16 bp repeat regions. These regions match the consensus for the SOS box and are known to bind the LexA protein. Thus a LexA protein binding site is found within the *recN* promoter region, which suggests that *recN* is regulated at the transcriptional level by the LexA protein (Rostas et al., 1987). The RecN protein has been shown to be highly induced in response to DNA damaging agents. RecN protein is expressed at low concentrations in the cell until the SOS response in induced; following induction, large amounts of RecN are generated (Meddows et al., 2005). In fact, *recN* is one of the most highly induced genes in the SOS regulon (McKenzie et al., 2000). Table 1 lists the characteristics of the *recN* gene and the RecN protein.

Characteristics of RecN Protein	
Rare codons	8.1% (*B. subtilis)*
Number of amino acid residues	576 (*B. subtilis)*
Molecular weight of the RecN protein	63.6 kDa (*B. subtilis)*
Isoelectric point	5.8 (*B. subtilis)*
Protein structure	Elongated head-coil-hinge structure (in *B. subtilis*; unknown for *E. coli*)
Subunit composition	Octamer in solution (in *B. subtilis*; unknown for *E. coli*)
Protein function	Unknown in both *B. subtilis* and *E. coli* – important in DNA recombination and repair

Table 1. Characteristics of the RecN protein present in *Escherichia coli* (Rostas et al., 1987; Kidane et al., 2004).

The RecN protein is member of the SMC (structural maintenance of chromosomes) class of proteins (Kidane et al., 2004). This protein family is found in all organisms as a DNA recombination and repair system. SMC proteins share a five-domain structure with globular N- and C- domains separated by a long coiled-coil segment in the center of a globular "hinge" domain (Haering et al., 2002). The coiled-coils are predicted to be antiparallel α–helices, bringing the N- and C- terminal globular domains together and generating an ATP binding site, which can be shown in Figure 6 (Connelly and Leach, 2002).

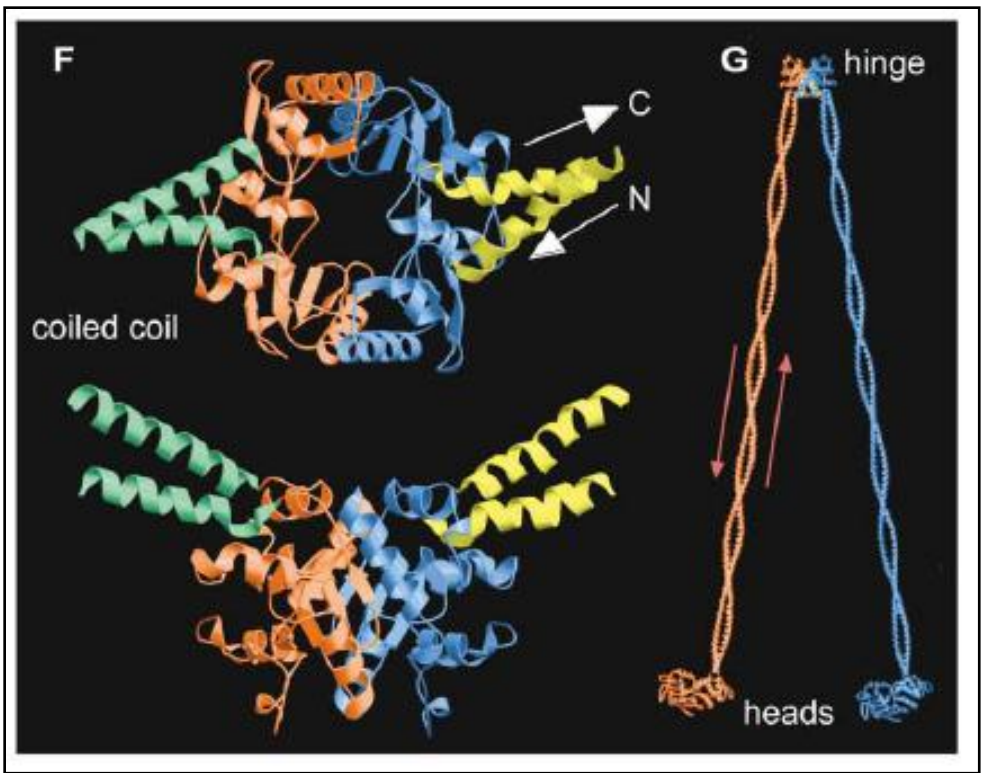


Figure 6. Sketch F demonstrates the crystal structure of the hinge domain in *Thermotoga maritime* SMC proteins. The yellow and green ribbon drawing shows the antiparallel coiled-coil regions of this protein. Sketch G displays the overall architecture of two eukaryotic SMC proteins interacting. In a prokaryotic system, there would be only one SMC protein (Haering et al., 2002).

These intramolecular coiled-coil regions are responsible for the flexibility of the globular protein. A conserved motif of the coiled-coil region is presumed to mediate protein dimerization and has been suggested to play a role in bringing broken ends of a DNA molecule together (Connelly and Leach, 2002).

In prokaryotic systems, such as *Escherichia coli*, the SMC proteins form heterodimers between similar SMC proteins. As a member of the SMC protein family, RecN was predicted to form an elongated head-coil-hinge structure. By means of electron microscopy, the *Bacillus subtilis* RecN protein has been shown to have a short rod-shaped or U-shaped extended flexible rod-like oligomer (Kidane et al., 2004). When present in solution, RecN forms homo-octamers, which are observed as four RecN dimers. RecN can also assemble into tetramers, but has not

been detected as monomers (Sanchez and Alonso, 2005). The coiled-coil arms of a RecN dimer form a ring with a diameter of 20 nanometers that can easily accommodate two DNA molecules (Kidane et al., 2004). RecN can also form connected structures, which have been proposed to enclose damaged DNA in an organized manner. This complex consists of multiple RecN oligomer and ssDNA molecules, where the RecN oligomers interact with the ssDNA backbone and the 3'-OH terminus, protecting the DNA end from exonucleases (Sanchez and Alonso, 2005).

Analogous to its SMC protein family, RecN has an ATP binding site. The active ATPase site of RecN is formed through the interaction of the globular N- and C- terminal regions, which associates with ssDNA regions (Sanchez et al., 2007). RecN can hydrolyze ATP to ADP and P_i in a ssDNA-dependent manner in the presence of Mg^{2+}. In the company of linear or supercoiled dsDNA, RecN ATPase activity was not stimulated (Sanchez and Alonso, 2005). The kinetics of RecN ATPase activity was measured to determine its affect on the protein structure. In the presence of ssDNA, RecN ATPase activity reached a plateau at 0.1mM ATP with a resulting K_m value of 0.05mM and an overall enzymatic catalytic rate of $18min^{-1}$ (Sanchez and Alonso, 2005). When ssDNA was not available, the overall enzymatic catalytic rate was $3min^{-1}$.

RecN is the first recombinational repair protein to localize to the site of a DSB. RecN arrives at the damaged DNA site in approximately 15 to 30 minutes after the inductions of DSBs, followed by recruitment of the RecO and RecA proteins roughly 15 minutes later (Sanchez et al., 2006). RecF is recruited by RecO shortly after it is relocated to the DSB. Figure 7 displays the first steps of recombination where RecN is localized to a DSB in DNA.

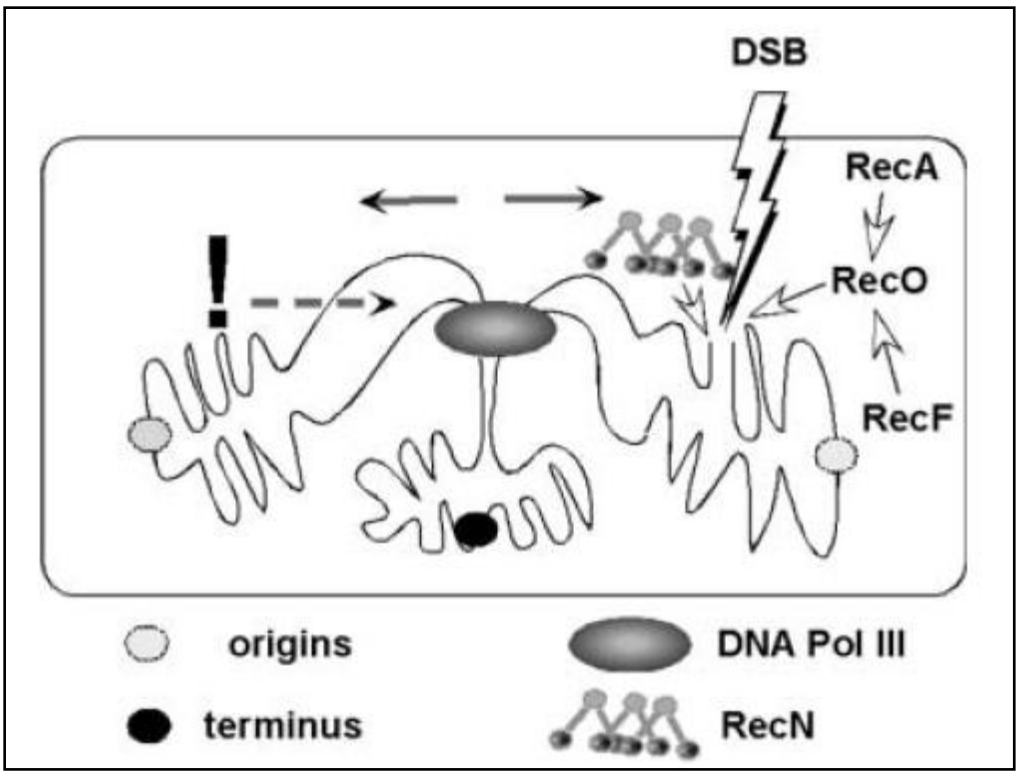


Figure 7. A model of DSB repairs in prokaryotic cells demonstrating the presence of RecN in the process as the first observed protein to localize at a discrete focus on DNA (Kidane et al., 2004).

The recruitment of these proteins by RecN forms active repair centers (RCs) that play an important role during DSB repair. Of the proteins utilized in active RCs, RecO is vital for the formation of the structures, where RecN is necessary for the organization of these dsDNA recombinational repair proteins (Kidane et al., 2004).

The number of RCs does not change with different amounts of DNA damage present in the cell. Typically, one RC is observed in the cell. Two or three RCs are seldom seen, regardless of the DNA damaging method used (DNA-modifying chemicals or irradiation) (Kidane et al., 2004). The concentration of RecN operates as a rate-limiting step in DSB repair. The number of RecN foci per cell does not increase with higher numbers of DSBs, signifying that several breaks are repaired within a single RecN focus (Sanchez et al., 2006).

Once completing its role in the cell, RecN begins to dissipate about an hour after first being induced (Kidane and Graumann, 2005). Since r*ecN* is regulated by the SOS response, the *recN* gene will attenuate once the signal for multiple DSBs diminishes. The RecN protein itself is degraded by a cytoplasmic protease, called ClpXP, present in the prokaryotic cell. ClpXP is a bacterial AAA+ protease that controls intracellular amounts of stress-response proteins, such as SOS proteins (Neher et al., 2006). ClpXP degrades the RecN protein by signal residues at the C-terminus (Nagashima et al., 2006) in order to remove the protein and obtain homeostasis in the prokaryotic cell following recovery from DNA damage.

The Putative Role of RecN in *Escherichia coli*

The RecN protein has been shown to be important for DNA recombination and DNA repair, but the exact function of RecN is uncertain (Kidane et al., 2004). The structure of RecN, its classification as a SMC-like protein, and its biochemical characteristics suggest a model of how RecN might participate in recombinational repair of DSBs. It has been hypothesized that the RecN protein congregates at the location of dsDNA breaks and facilitates recombination. Thus, the role of RecN might be to: 1) protect the DNA ends from degradation, 2) bring the two broke ends of DNA together, and/or 3) to juxtapose the broken DNA ends to the undamaged template to facilitate recombination repair. The main goal of this project was to develop a plasmid-based gap repair assay that could test these predictions in vivo. Both the RecFOR and λ Red pathways or recombination were tested for the ability to promote recombination of a linearized plasmid with the *E. coli* chromosome. The gap-repair assay substrate was designed to provide two DNA ends that must be brought together in close proximity to an undamaged template located on the *E. coli* chromosome. Once repaired, the plasmid would replicate and

confer tetracycline or chloramphenicol resistance to the host cell. The gap repair substrates were used to test predictions of this model regarding the role of RecN in dsDNA break repair in *E. coli*.

MATERIALS AND METHODS

Construction of Strains and Plasmids

Strains

The strains used and constructed in this study are listed in Table 2. Strain constructions were performed by P1 transductions or Red-mediated gene replacement (Maniatis et al, 1985; Murphy et al, 2000).

Strain	Relevant Genotype	Reference/Source
MG1655	wild type	K.C. Murphy
AB1157	wild type	K.C. Murphy
KM22	*ΔrecBCD::kan*	K.C. Murphy (1998)
TP404	*recB sbcB sbcC*	K.C. Murphy
KM126	*lacZ::Δtet254*	K.C. Murphy
KM129	*recA::kan lacZ::Δtet254*	K.C. Murphy
KM133	*recN::kan lacZ::Δtet254*	K.C. Murphy
KM139	*sulA::spec lexA::kan*	K.C. Murphy
KM149	*recB sbcB sbcC lacZ::cat*	K. C. Murphy
KM150	*recBC xonA::gen lacZ:cat*	K. C. Murphy
KM151	*lacZ::Δtet254 recN-LAA::kan*	K. C. Murphy
CF153	*lacZ::cat*	KM126 x pCF288
CF154	*recA::kan lacZ::cat*	KM129 x pCF288
CF155	*recN::kan lacZ::cat*	KM133 x pCF288
CF156	*sulA::spec lexA::kan lacZ::cat*	KM139 x pCF288
CF157	*ΔrecBCD::kan lacZ::cat*	KM22 x pCF288
CF158	*recBC sbcB sbcC lacZ::cat*	JC9387 x pCF288
CF159	*ΔrecBCD::Plac-red kan rpsL31 lacZ::cat288 ΔrecN::tet*	KM157 x P1 (TP664)
CF160	*ΔrecN::cat rpsL rpsL31 lacZ:: Δtet254*	KM126/pKM208 x PCR (recN::cat rpsL)
CF161	*ΔrecN::cat rpsL rpsL31 lacZ:: Δtet254*	KM126 x P1 (CF160)
CF162	*ΔrecBCD::Plac-red kan rpsL31 lacZ:: Δtet254*	KM126 x P1 (KM22)
CF163	*ΔrecBCD::Plac-red kan ΔrecN::cat rpsL rpsL31 lacZ:: Δtet254*	CF161 x P1 (KM22)

Table 2. Strains used and constructed in this study.

Plasmids

The plasmids used and constructed in this study are listed in Table 3.

Plasmid	Description
pKM208	λ Red + Gam-producing plasmid. Amp^R. Temperature sensitive replicon, must be kept at 30°C. (Murphy and Campellone, 2003).
pKM221	Used as a control to test electrocompetency and frequency of transformations following electroporation. Confers gentamicin resistance.
pKM254	Derivative of pBR322 (Boyer et al., 1973). Contains modified tetracycline resistance gene that creates two *XhoI* sites within *tet* sequence. Ampicillin resistance. (Murphy, unpublished).
pKM255	Derivative of pKM254 that contains *XhoI* deletion of the *tet* gene. Contains Amp^R. (Murphy, unpublished).
pKM258	Derivative of pKM132 (Murphy, 1998) containing a 2550 bp region of the *lacZ* gene inserted into the cloning vehicle pBR322 (Boyer, et al., 1973). A *NotI* site has been placed in the middle of the *lacZ* sequence. Used in this study as a gap repair substrate.
pCF288	The *NotI* fragment from pTP883 (consisting of the *cat* gene, Murphy et al., 2000) was cloned into the *NotI* site of pKM258. A 3 kb fragment is produced by *BglI* digestion that contains the *cat* gene flanked by the 1 kb of *lacZ* sequence. Used to measure recombination with linear DNA substrates and to construct strains containing the *cat* gene within the lacZ chromosomal loci.
pKM301	Derivative of pKM255 cut with *AflIII* then ligated with oligonucleotides to produce a TelN site. Used in short gap repair assay to measure recombination with a hairpin-containing DNA substrate.
pGB2	Used as a control plasmid, confers spectinomycin resistance; to test electrocompetency and frequency of transformation following electroporation.

Table 3. Plasmids used and constructed in this study.

Transformations

Cultures of *E. coli* $W3110^q$ cells were grown overnight in 10 ml of Luria-Bertani (LB) at 37°C. Fresh cultures (5 ml) were grown to 1 x 10^8 cells/ml, collected by centrifugation, and resuspended in 10 ml of CP-20 buffer. After 30 minutes on ice, the cells were collected by centrifugation and resuspended in 1-2 ml of CP-20 buffer. An aliquot of these competent cells (100 µl) was combined with 60 µl PCM buffer and 2 µl of plasmid or 15 µl of ligation mixture.

The transformation mixture was incubated in an ice bath for 30 minutes, followed by an incubation period of 5 minutes at 37°C (i.e. heat shock). The sample was incubated at room temperature for 5 minutes then diluted into 2 ml of LB. These cells were placed in a 37°C incubator and rolled for 1 hour. Following growth of the cells, 500 µl of the sample was plated (using sterile technique) on the LB plates containing 100 µg/ml ampicillin, 20 µg/ml kanamycin, 15 µg/ml tetracycline, or 10 µg/ml gentamicin. The plates were incubated overnight at 30°C or 37°C, depending on the plasmid.

Polymerase Chain Reaction (PCR)

The *recN* deletion allele used for the construction of CF160 was generated by PCR. The reaction mixture included: 24 µl H_2O, 4 µl 10x buffer, 2 µl dNTPs, 2 µl DMSO, 2 µl RecN1 primer, 2 µl RecN2 primer, 2 µl template (pTP1002b diluted 1:50), and 2 µl Taq polymerase at a 1:5 dilution (40 µl total). The primer sequence is listed in Table 4. The PCR tube was placed in the thermocycler and the following program was run:

Step 1 – Denaturing: 3 minutes at 95°C
Step 2 – Denaturing: 30 seconds at 95°C
Step 3 – Annealing: 30 seconds at 55°C
Step 4 – Extending: 30 seconds at 72°C
Step 5 – Repeat Steps 1 – 4 for 29 more cycles
Step 6 – Final Extension Step: 300 second at 72°C

Primer Name	Sequence (5'-3')
RecN1	ATGTTGGCACAACTGACCATCAGCAACTTTGCTATCGTTCAGAGTGTTGACTTGTGAGCG
RecN2	TTACGCTGCAAGCAGTTCTTTCGCATTCGCCAGTGTATTATAGAGGGCGGATTTGTCCTA

Table 4. The RecN1 and RecN2 primers used in the PCR to generate the RecN mutants.

The amplified DNA was purified using a Qiagen PCR purification kit. This DNA sample was used as a substrate for Red recombineering of KM126/pKM208 to construct CF160.

P1 Transductions

P1 lysates of P1 phage were prepared as followed: 5 ml cultures were grown overnight standing in a 37°C incubator in LB media. Plates were prepared by spreading 1 ml of 30 mM $MgSO_4$, 15 mM $CaCl_2$ onto LB plates, allowing the solution to soak into the plates for 20 minutes. A mixture of 0.2 ml cells, 10 µl P1 virus, and 3 ml of soft agar was combined and gently mixed by shaking. The mixture was quickly poured onto the plate and incubated overnight at 37°C. The next day, the soft agar was collected, washed in 3 ml of LB, and transferred to a 40ml Oakridge tub. Chloroform (0.2 ml) was added to the lysate and the mixture was vortexed thoroughly. The lysate was centrifuged for 5 minutes at 10,000 rpm and then the supernatant was collected. Prior to storage, 50 µl 30 mM $MgSO_4$, 15 mM $CaCl_2$ was added to the P1 lysate, which was mixed thoroughly, and placed into the refrigerator until further use.

P1 transductions were carried out as follows: cultures were grown standing at 37°C in 10 ml of LB with the appropriate antibiotics. The overnight culture was centrifuged at 6,000 rpm for 10 minutes and then resuspended in 2 ml of LB. The resuspended cells were placed on ice and 1 ml of 30 mM $MgSO_4$, 15 mM $CaCl_2$ was added; the mixture was then vortexed. Into four sterile microcentrifuge tubes, 0.6 ml of the cells was mixed with varying amounts of P1 lysate (0 µl, 1 µl, 5 µl, and 20 µl) and incubated for 45 minutes in a 37°C water bath. These cells were centrifuged at 6,000 rpm for 2 minutes at 4°C and resuspended in 50 µl 10 mM Tris-HCl, pH

7.5, 1 mM $MgSO_4$ (TM) buffer. This mixture was plated on the appropriate antibiotic plates and grown overnight at 37°C.

Electroporation Procedure

An overnight culture of *E. coli* (see Figure legends for strains used) was grown in 5 ml of LB with the appropriate antibiotics at either 30°C or 37°C (depending on the plasmid). Aliquots of the overnight culture was diluted into 20 ml LB containing the same antibiotic and shaken at 30°C in a 125 ml flask. Following 30 minutes, 200µl of 0.1 M IPTG was added to the culture. Once the culture reached a final density of 1.0×10^8 cells, the flask was placed in an ice bath and slowly stirred for 15 minutes. In some cases, the procedure included a heat shock step, where the cultures were placed in a 42°C water bath for 15 minutes prior to collection. The cells were centrifuged for 10 minutes at 6,000 rpm, resuspended in 1 ml of cold 1 mM MOPS, 20% glycerol, and transferred to a sterile 1.5 ml microcentrifuge tube. The cells were spun down for 1 minute at 10,000 rpm at 4°C and the supernatant was removed. The cell pellet was resuspended in 1 ml 1 mM MOPS, 20% glycerol. This step was repeated two more times. The cells were finally resuspended in 100 µl of cold 1 mM MOPS, 20% glycerol. This amount of electrocompetent cells was usually good for two electroporations (Murphy and Campellone, 2003).

Electroporation cuvettes were placed on ice for 10 minutes. DNA samples were used that contained 25 ng – 100 ng of intact plasmids or linear segments of DNA. A 50 µl sample of cells was combined with DNA substrate, and then transferred to a cooled electroporation cuvette. This cuvette was incubated on ice for approximately 1 minute, and then completely dried. The cells were shocked with a 2,000 V/cm discharge and quickly resuspended in 0.5 ml LB. The

mixture was placed into 3 ml LB and rolled at 37°C for 1.5 hours. The cultures were diluted and plated on the appropriate antibiotic plates, which were placed in a 37°C incubator overnight.

Gap Repair Assays

Two different strain/plasmid combinations were used to measure the ability of the λ Red and RecFOR recombination to promote repair of dsDNA breaks. In the first of these assays, plasmid pKM255 contains an internal deletion of the tetracycline gene (*tet*). Following linearization of pKM255 with *XhoI*, 33 ng of plasmid DNA was electroporated into various *E. coli* strains containing chromosomal sequences that, when recombined with the plasmid, would restore the *tet* gene and confer resistance to tetracycline (i.e., gap repair; see Figure 8). Gap repair frequency was found by determining the percentage of Tet^R transformants per survivor following electroporation of linear pKM255. Strains used in these experiments are listed in Table 1; strains used for each experiment are described in the Table and Figure legends. In some assays, 25 – 50 ng of intact control plasmids pGB2, pBR322 or pKM221 (conferring resistance to spectinomycin, ampicillin, and gentamicin, respectively) was used to determine the frequency of DNA uptake (i.e., competency). In these cases, gap repair frequency was determined by the number of Tet^R transformants per competent cell (Tet^R transformation frequency/control plasmid transformation frequency). Measuring gap repair using recombinants per competent cells was more reliable than measurements reported as recombinants per surviving cells. Measurements made using recombinants per surviving cells gave varying values for the linear and circular DNA substrates within a strain, which could not be used in the calculations. Recombinants per competent cell measurements gave more reliable values by basing the calculations on the competency of the cells instead of by the number of surviving cells.

UV-induced Induction of the SOS Response

A modified protocol for preparing electrocompetent cells was performed in order to include an exposure to UV during the growth of the cells, as a means to induce the SOS response in an effort to increase the intercellular concentrations of the RecN protein. Aliquots of overnight cultures (100 µl) were introduced into 4 different 125 ml flasks containing 20 ml of LB with the appropriate antibiotics. The flasks were shaken at 37°C. After 30 minutes, 200 µl of 0.1 M IPTG was added to the each culture. The samples were continuously shaken until the cultures reached a density of 5×10^7 cells. At this point, the samples were put on ice to hinder any further growth and centrifuge at 6,000 rpm for 10 minutes. The cells were resuspended in 5 ml cold phosphate buffer solution (PBS). The solutions were poured onto sterile plates and exposed to UV dose of 50 J/m^2. After UV radiation, the cells were collected by centrifugation (6,000 rpm for 10 minutes) and resuspended in 10 ml of warm LB broth. The samples were placed back into the flasks and grew for another 30 minutes or until they reached a final density of 1×10^8 cells/ml. The samples were once again placed on ice for 20 minutes to prevent further growth and then treated as described above for use in electroporations.

A previous UV exposure to cells containing a *sulA-β-gal* chromosomal function (*sulA* is an SOS-induced function) revealed that SOS induction is observed at this dose using a β-galactosidase assay. For the measurement of β-galactosidase activity, the procedure described was used (Miller, 1972). Briefly, the cell density was measured and recorded for each sample (1 ml) at an absorbance of 600 mµ. Following the absorbance reading, samples (0.1 ml and 0.5 ml) were aliquoted into Z buffer to a final volume of 1 ml. Under the hood, 2 drops of chloroform and 1 drop of 0.1% SDS solution were added to each tube, which were vortexed thoroughly for 10 seconds. The tubes were placed in a 28°C water bath for 5 minutes. To start the reaction, 0.2

ml of 4 mg/ml ONPG was added to each of the tubes (time zero). When a bright yellow color developed, the reaction was stopped by adding 0.5 ml of 1 M Na_2CO_3 solution and the time was recorded. The optical density was measured at two absorbances: 420 mμ and 550 mμ. To determine the concentration of β-galactosidase in units, the following equation was used:

$$Units = 1000\ x \frac{(OD_{420} - 1.75\ x\ OD_{550})}{(time\ x\ volume\ x\ \mathrm{OD}_{600})}$$

RESULTS

Gap Repair Assay

Plasmid-based Gap Repair Assay

The plasmid-based gap repair assay developed in this project was used to test the role of the RecN protein in the presence of broken DNA ends. This model consists of two major components: an undamaged DNA template and linear plasmid, the ends of which would act as a dsDNA break in vivo. With the method, it might be possible to mimic a dsDNA break that occurs in the *E. coli* chromosome in vivo. Two DNA substrates were used in this assay: linear pKM255 (testing for short homology, ~600 bp) and linear pKM258 (testing for long homology, ~1000 bp). In the short homology assay, pKM255 cut with *XhoI* was electroporated into an *E. coli* strain that contains an inactive fragment of the tet gene on its chromosome. If the two ends of the linear plasmid religate, the *E. coli* cells will not exhibit Tet^R (since an internal fragment of the gene is missing in the plasmid). However, if the ends undergo HR with the chromosome, the *tetracycline* resistance gene will be generated; making the host cells resistant to tetracycline. Figure 8 shows a picture of this plasmid-based gap repair assay using the short homology DNA substrate.

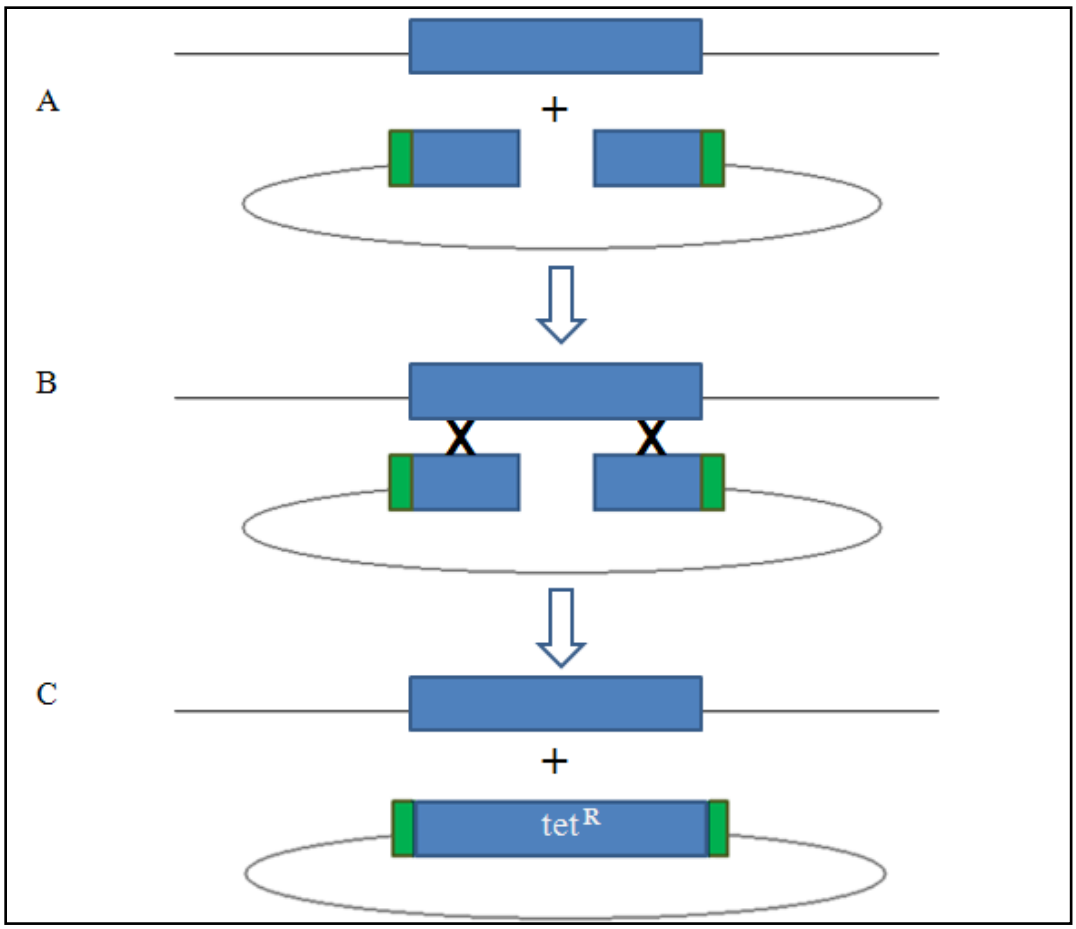


Figure 8. Short substrate gap repair assay using pKM255 cut with *XhoI*. (A) Chromosomal DNA is represented on top where a linearized plasmid (pKM255 cut with *XhoI*) is shown on the bottom. Both DNA fragments are missing essential portions of the *tetracycline* gene, thus neither is tetracycline resistance. The chromosomal DNA contains the middle portion of the gene and the linearized plasmid contains internal deletions of the gene. (B) HR occurs between similar DNA sequences on the chromosome and the linear pKM255. (C) After recombination, the *tetracycline* gene sequences from the chromosome will be integrated into the plasmid, resulting in a tetracycline resistance plasmid. Green boxed represent N- and C- terminal encoding regions of the *tet* gene (which are absent in the chromosome).

The recombinants per surviving cell, and in some experiments, the recombinants per competent cell were measured for each strain to determine the role of RecN.

The λ Red and RecFOR Pathways

First, it was necessary to determine the most effective pathway to use in the plasmid-based gap repair assay. The two pathways of interest included the λ Red and the RecFOR

pathways. One strain from each background was used in an electroporation with the same concentrations of identical plasmids to test and compare the ability of each pathway in gap-repair. The λ Red pathway was represented by CF157 and the RecFOR pathway was represented by CF158. The plasmids used in this experiment consisted of linear pKM258 (cut with *NotI*) and circular pKM221. The effectiveness of gap-repair in each pathway can be observed in Table 5.

Strain	**Relevant Pathway**	**Plasmid**	**Recombinants/ Competent Cell**
CF157	λ Red	pKM258 cut with *NotI* pKM221	8.89×10^{-3}
CF158	RecFOR	pKM258 cut with *NotI* pKM221	1.10×10^{-3}

Table 5. Comparison between the λ Red and RecFOR Pathways. A mixtures were used consisting of pKM258 cut with *NotI* (50 ng) and pKM221 (50 ng) in electroporations with both strains. The cells were plated on chloramphenicol (400 µg/ml) and gentamicin (10 µg/ml). The experiment was done once.

The data shows a 9-fold decrease in recombinants per competent cell in the RecFOR pathway when compared to the λ Red pathway. This drop in recombinant frequency demonstrates that the λ Red pathway was more efficient in performing the gap-repair assay. Further plasmid-based gap repair assays were conducted in the λ Red pathway.

Short Homology Plasmid-based Gap Repair Assay

The λ Red pathway was used in conjunction with the short homology plasmid (pKM255 cut by *XhoI)* in efforts to determine the function of RecN. Short homology refers to the homologous ~600 bp region in the *tetracycline* gene found between the chromosome and plasmid. Four strains (KM126, KM129, KM133, and KM151) were used in this series of experiments, which contained the plasmid encoding for the λ Red and Gam proteins (pKM208). Linear pKM255 was electroporated into the four strains where the *tetracycline* gene from the

chromosome and this plasmid would come together by HR, shown in Figure 8. Each strain was electroporated at least twice to confirm results, seen in Table 6.

Strain	Relevant Genotype	Recombinants/ Survivor (x 10^{-4})	Standard Error
KM126/pKM208	wild type	0.364	0.0398
KM129/pKM208	*recA*	8.37	2.35
KM133/pKM208	*recN*	0.995	0.236
KM151/pKM208	*recN-LAA*	0.285	0.005

Table 6. Short homology recombination used for linear recombination. Electroporated with pKM255 cut with *XhoI* (100 ng) and selected for tetracycline resistance (15 µg/ml tetracycline). KM126/pKM208 was electroporated 5 times, KM129/pKM208 and KM151/pKM208 were electroporated twice, and KM133/pKM208 was electroporated 4 times.

Unexpectantly, the number of recombinants per surviving cell shows a 3-fold increase in the *recN* mutant compared against the wild type strain. In addition to this, the *recA* mutant showed a 23-fold increase in the recombination rate compared to the wild type strain. These results demonstrate that the RecA or RecN proteins are not necessary for the λ Red-promoted gap repair. In fact, the data shows that the presence of each protein has an inhibitory effect on dsDNA break repair using the plasmid-based gap repair assay. The possible mechanisms for this observation are discussed below (see Discussion section).

Short Homology Plasmid-based Gap Repair Using Lower Amounts of λ Red Proteins

One possibility for the lack of an effect of RecN in the gap repair assay described above may be that over-expression of the Red recombination system promotes repair at such high levels that it can proceed without assistance from RecN. The KM22 strain background was used to determine if lowering the concentrations of the λ Red protein would affect RecN-mediated recombination. The KM22 background contains a single copy of the Red system driven by the

Plac promoter from the chromosome, whereas in the previous experiment, Red was expressed from the stronger Ptac promoter from a multi-copy plasmid.

The Red recombination system from KM22 was P1 transduced to strains containing the gap repair target (strains CF162 and CF163 in Table 2). The verification of the RecN phenotype in the KM22 background was tested by PCR analysis of the *recN* region of the chromosome, and a nalidixic acid assay sensitivity test. PCR products of the *recN* loci in strains (KM126, CF161, CF162, and CF163) were run on a 0.75% agarose gel. The *recN* mutants, containing the *cat rpsL* allele at the location of the *recN* gene, is predicted to give a slightly smaller PCR product relative to wild type in this test. This slight size difference between these PCR products can be seen in a 0.75% gel as shown in Figure 9.

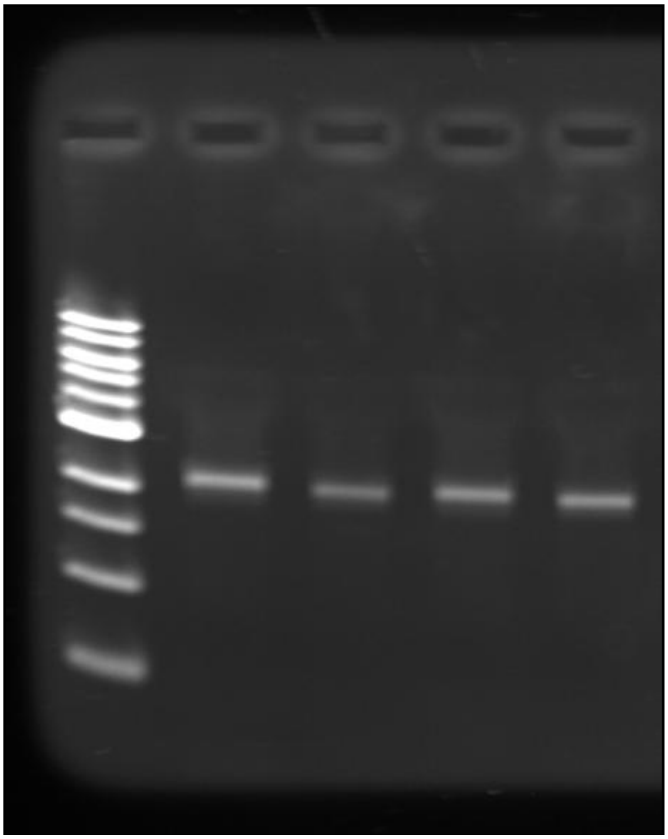
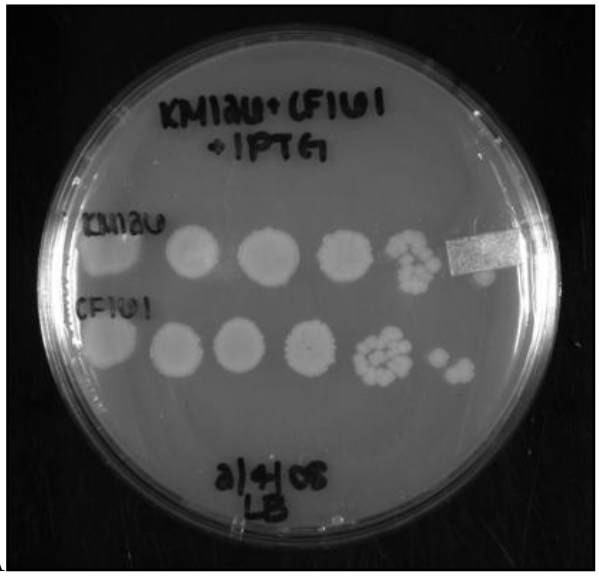


Figure 9. Comparison between wild type and *recN* mutants using gel electrophoresis. Lane 1 is 1kb DNA ladder, lane 2 is KM126 PCR product, lane 3 is CF161 PCR product, lane 4 is CF162 PCR product, and lane 5 is CF163 PCR product. A slight difference can be observed in lanes 2 and 4, showing the presence of *cat rpsL* replacing the *recN* gene.

The strains: KM126 (wild type) and CF161 (recN) were tested for sensitivity for nalidixic acid. Nalidixic acid is a chemical mutagen that induces dsDNA breaks. As such, a *recN* strain should be more sensitive to the effects of nalidixic acid relative to wild type. Both strains were spot tittered on LB and nalidixic acid plates of varying concentrations. Figure 10A demonstrates the spot titterings on LB plates and Figure 10B shows the spot titerings on increasing concentrations of nalidixic acid (1 µg/ml, 2 µg/ml, and 4 µg/ml). The *recN* phenotype was confirmed at 2 µg/ml nalidixic acid.

A 

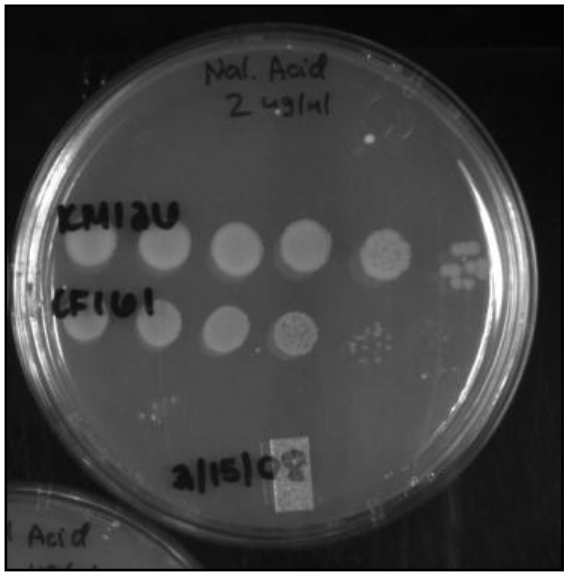
B 

Figure 10A and 10B. (A) KM126 and CF161 spot titered on LB plates. Each spot corresponds to a 10-fold dilution of the cells in PBS. The first spot on the left is the first 10-fold dilution. The number of colonies for each strain in each spot is approximately the same. (B) KM126 and CF161 spot tittered on varying concentrations of nalidixic acid (1 µg/ml, 2 µg/ml, and 4 µg/ml). In the presence of 2 µg/ml nalidixic acid, a decrease in colonies is observed in CF161, but not noticeable in KM126, demonstrating the presence of the *recN* mutant in CF161.

Once the *recN* mutant was verified in CF161, lower concentrations of the λ Red proteins were tested in the gap repair assay. The two strains CF162 and CF163 were used. These strains were exposed to 50 J/m^2 of UV light to turn on RecN and then electroporated with the short homology DNA substrate; results are shown in Table 7.

Strain	Relevant Genotype	Plasmid	UV Presence	Recombinants/ Competent Cell
CF162	wild type	pKM255 cut with *XhoI*	-	2.32×10^{-3}
		pKM221	-	
		pKM255 cut with *XhoI*	+	5.58×10^{-3}
		pKM221	+	
CF163	*recN*	pKM255 cut with *XhoI*	-	3.11×10^{-3}
		pKM221	-	
		pKM255 cut with *XhoI*	+	8.31×10^{-3}
		pKM221	+	

Table 7. Calculated recombinants per competent cell in CF162 and CF163 in the presence of UV light. Cells were introduced to UV light to turn on the SOS response and then electroporated with a mixture consisting of pKM255 cut with *XhoI* (100 ng) and pBR322 (50 ng). These cells were plated on tetracycline (15 µg/ml). This experience was done twice.

In these experiments, the recombination rate was calculated as the numbers of recombinants per competent cell by comparing transformation rates of the gap repair plasmid (pKM255) relative to a control plasmid (pKM221), which tested the electrocompetence of the cell preparation. Using this method, variabilites due to DNA uptake in each electroporation can be included in the recombination measurement. As seen in Table 7, recombinants per competent cell was slightly higher for the RecN mutants, showing that gap-repair was not dependent on the RecN protein.

Determination of Chloramphenicol Concentration

Two strains, a multicopy plasmid of the *cat* gene (AB1157/pKM208, pKM288) and a single copy of the same gene on the chromosome (AB1157/pKM208), were used to determine the concentration of chloramphenicol affecting *cat* gene on the chromosome or on a plasmid. Chromosomal sensitivity to chloramphenicol was established at 400 µg/ml. This concentration was used in the long homology plasmid-based gap repair to test on the plates. Figure 11 displays

the chloramphenicol concentrations used to establish the amount needed to show a difference between the *chloramphenicol* gene on the chromosome and on the plasmid.

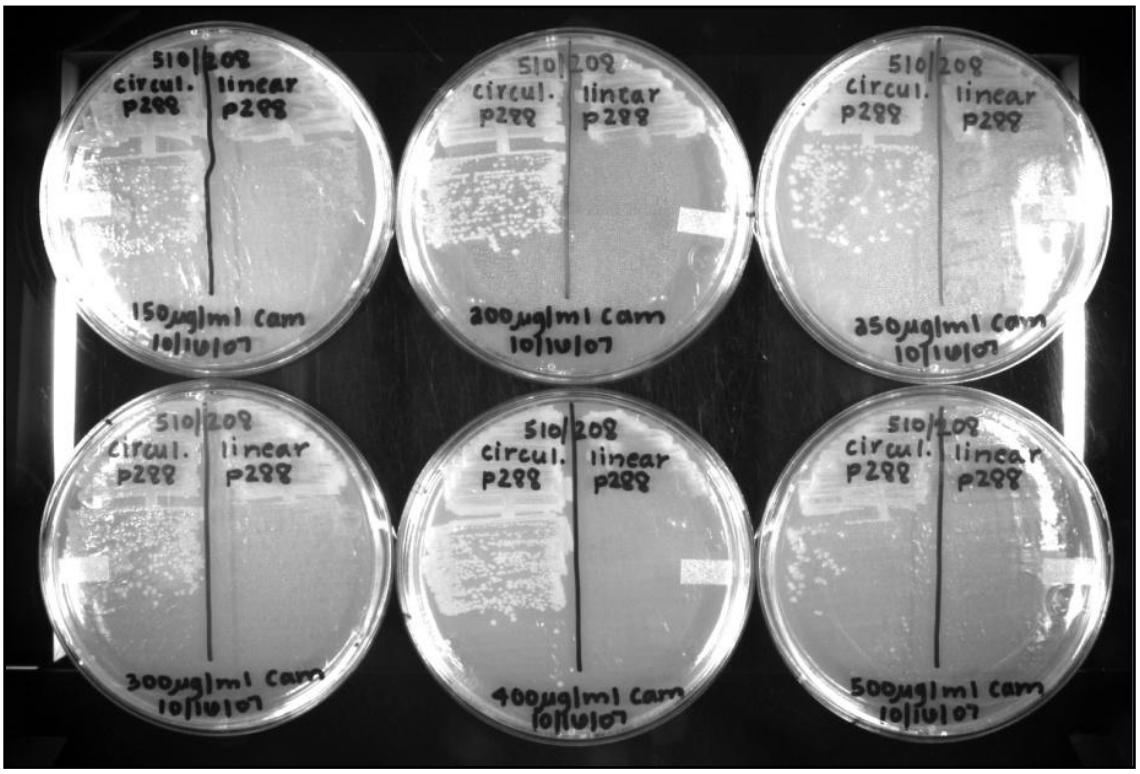


Figure 11. A range of chloramphenicol concentrations, starting at 150 µg/ml and ending at 500 µg/ml, were used to determine the concentration that could set apart the *cat* gene on the chromosome from the *cat* gene on a plasmid. On each plate, the plasmid *cat* gene was present on the left and the chromosomal *cat* gene was present on the right. The concentration of chloramphenicol that displayed the best difference was 400 µg/ml.

Determination of pCF288 Concentration

The circular plasmid, pCF288, was used as a control in electroporations involving the long homology gap repair system. Varying concentrations (in ng) of pCF288 were tested to determine the concentration needed to provide the maximum amount of DNA without saturating the system. Table 8 and Figure 12 show the results from this experiment. Transformant frequency saturated up to 50 ng was used for the electroporation. In order to use an amount of

substrate close to the linear range of transformation frequency, the resulting pCF288 concentration used in the long homology gap repair assay was 25 ng.

Strain	Concentration of pCF288 (ng)	Transformants/ µg pCF288
KM126/pKM208	50	9.16×10^6
	25	1.216×10^7
	12.5	3.47×10^6
	6.25	3.42×10^6

Table 8. Varying concentrations of pCF288 by 2-fold dilutions. These concentrations were electroporated into KM126/pKM208 and then selected for chloramphenicol resistance (400 µg/ml chloramphenicol).

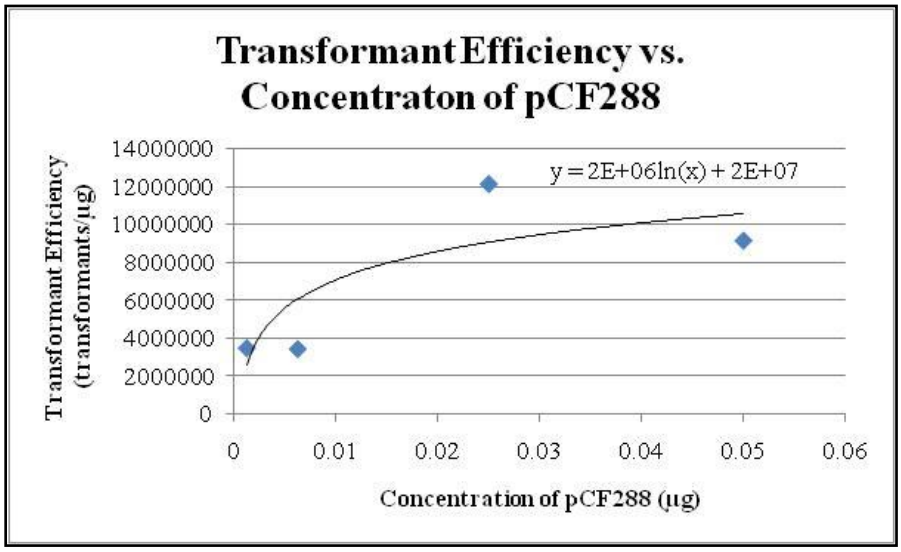


Figure 12. Transformant efficiency versus the concentration of linear pCF288 in micrograms.

Long Homology Plasmid-based Gap Repair Assay

The long homology plasmid (pKM258 cut with *NotI*) was also used in the λ Red background to test the function of RecN in *E. coli*. The linear pKM258 contains a homologous ~1000 bp regions homologous to the *lacZ* gene that flank the *chloramphenicol* (*cat*) gene. This DNA substrate is considered a long homology plasmid due to the flanking *lacZ* gene on each

side of the *cat* gene; thus increasing the rate at which HR could take place. Figure 13 shows the scheme behind this long homology plasmid-base gap repair assay.

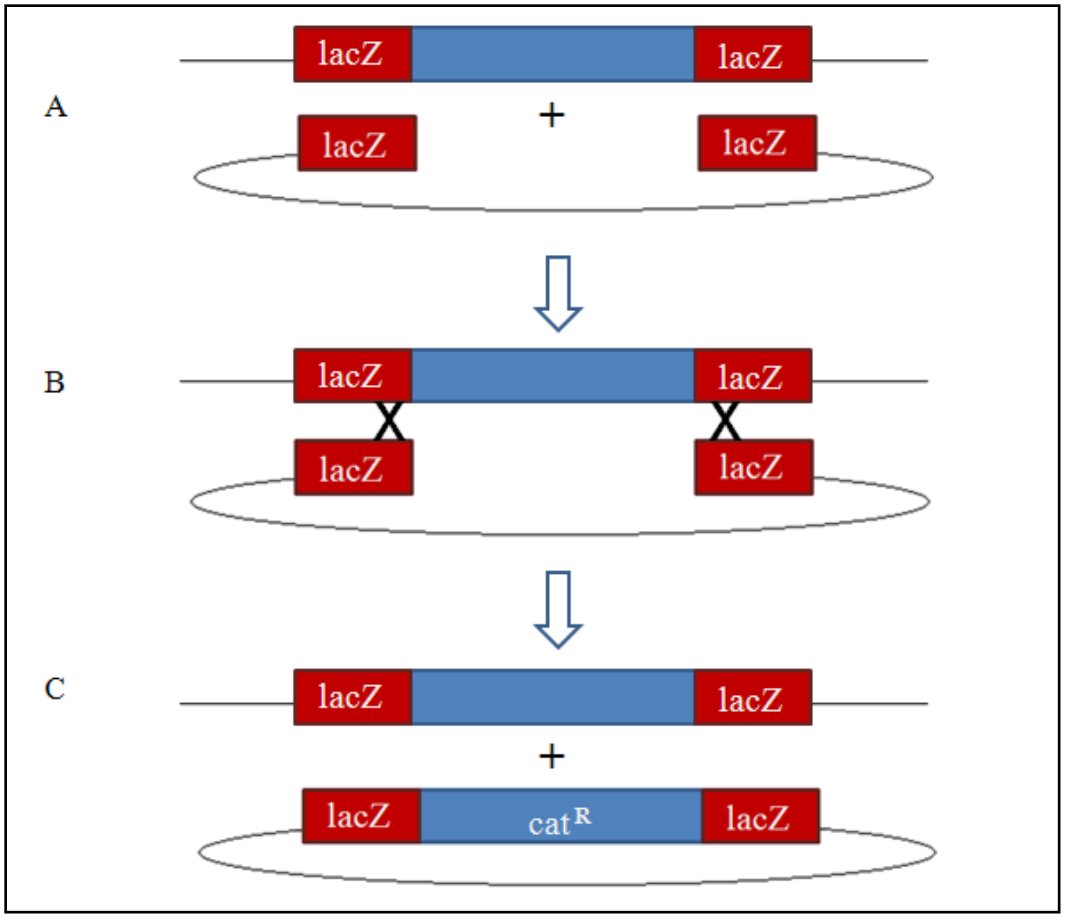


Figure 13. Long substrate gap repair assay using pKM258 cut with *NotI*. (A) Chromosomal DNA is represented on top where a linearized plasmid (pKM258 cut with *NotI*) is shown on the bottom. The chromosome *lacZ* regions contain an insertion of the *cat* gene. The plasmid contains *lacZ* sequences. (B) HR occurs between *lacZ* sequences on the chromosome and the linear pKM258. (C) After recombination, the *chloramphenicol* gene from the chromosome will be integrated into the plasmid, resulting in a chloramphenicol resistance plasmid. Plasmid-encoded *cat* function is distinguished from chromosomally-located *cat* by plating at high concentrations of chloramphenicol (400 µg/ml).

Two strains (CF153 and CF155) were used in this experiment. The control for this experiment consisted of pCF288 to measure recombination with linear pKM258 substrate. Tables 9 and 10 displays the results from the experiments.

Strain	Relevant Genotype	Plasmid	Recombinants/ Survivor (x 10^{-4})	Standard Error
CF153	wild type	pKM258 cut with *NotI*	0.40	0.28
CF155	*RecN*	pKM258 cut with *NotI*	1.08	0.17

Table 9. Long homology recombination for gap repair. Cells were electroporated with *NotI*-digested pKM258 (33 ng) and selected for chloramphenicol resistance (400 µg/ml chloramphenicol). CF153 was tested 3 times and CF155 was tested twice.

Strain	Relevant Genotype	Recombinants/ Competent Cell	Standard Error
CF153	wild type	3.13 x 10^{-3}	-
CF154	*recA*	19.5 x 10^{-3}	12.85
CF155	*recN*	3.69 x 10^{-3}	-

Table 10. Calculated recombinants per competent cell. Cell were electroporated with *NotI*-digested pKM258 (100 ng) and intact pCF288 (25 ng) and selected for chloramphenicol resistance (400 µg/ml chloramphenicol). The recombinants per competent cell were calculated by dividing the number of colonies by the amount of LB used (0.5 ml), then multiplying by the dilution factor. This was done for the both the plasmids used in each strain. CF153 and CF155 were electroporated once and CF154 was electroporated twice.

The number of recombinants per surviving cell shows a 3-fold increase in the *recN* mutant compared against wild type, similar to what was seen using the small homology substrate (see Table 6). These recombinant rates per surviving cell calculations further demonstrates that plasmid gap repair was not dependent on the RecN protein.

The recombinants per competent cell were also calculated with the long homology substrate. The calculated value was included the competency of the cells, rather than relying on the number of colonies that survived the electroporations. This method, as seen before, provides more reliable numbers for the gap repair assay. In this assay, there was no effect of RecN observed. The numbers obtained in Table 10 demonstrates that the RecN protein is not necessary in plasmid-based gap repair assays when Red is used for recombination. Interestingly,

the *recA* mutant results show a 6-fold increase in recombination rates compared to the wild type levels, which reveals the RecA protein may interfere with plasmid gap repair.

Tests of Substrates containing hair-pinned DNA Ends

A derivative of the short homology plasmid (pKM301) was used in the design of a substrate that would produce hair-pinned ends in *E. coli*. This main idea behind this design was to better duplicate the presence of dsDNA breaks in the cell by distancing both broken ends in solution. In the previous small homology gap repair technique, the broken dsDNA ends were attached to each other because they were the ends of a linear plasmid. This arrangement could possibly have an effect on dsDNA break repair. TelN sites (phage N15 telomerase site) were inserted into pKM301 by cloning, and then cut with the TelN telomerase to generate a linear DNA with hair-pinned ends. If this substrate were then cut by a restriction enzyme, two fragments would be generated that could serve as a pair of substrates for dsDNA break repair.

In test of such a substrate; however, the TelN-cut DNA substrate transformed with a 30-fold lower efficiency relative to intact pKM301. Table 11 displays the amount of transformants per 10^8 surviving cell for the TelN-cut substrate and a circular plasmid. It is likely that the hairpin interfered with replication of the linear plasmid; thus, TelN-treated DNA substrates could not be used in the analysis of RecN function in these assays, and were not further investigated.

Strain	**Transformation Titer (Transformants per 10^8 Survivor**
KM126/pKM301	2140
KM126/pKM301 *TelN* cut	73

Table 11. Titers following transformation of *E. coli* cells with pKM301 or pKM301 cut with TelN. DNA plasmid substrate (50 ng of pKM301 and pKM301 TelN-cut) were electroporated into KM126 and plated on LB plates containing tetracycline (15µg/ml). This experiment was done twice (representative experiment is shown).

DISCUSSION

The main objective of this project was to determine the role of RecN in vivo through the use of a plasmid-based gap repair assay. This type of assay was used to mimic the repair of a dsDNA break on the *E. coli* chromosome, where the linearized DNA substrate acted as the damaged dsDNA break and the chromosome was the undamaged template. The plasmid-based gap repair assay was easy to carry out by simple electroporation of purified DNA into *E. coli*, which added to its appeal.

The λ Red and RecFOR pathways were used to test the efficiency of the plasmid-based gap repair assay. The λ Red pathway was found to be more efficient in the plasmid-based gap repair assay than the RecFOR pathway. Wild type *E. coli* cells can be used for the expression of the λ Red pathway, whereas mutations (*recBC sbcBC* mutants) must be made in the wild type host to turn on RecFOR pathway. Mutations made in the RecFOR pathway may make the cells sick, which might be the reason this pathway is less efficient for the gap repair assay.

Two DNA substrates were used in the plasmid-based gap repair assay: a short homology plasmid (pKM255) and a long homology plasmid (pKM258). The short homology gap repair assay utilized a ~600 bp homologous region between the chromosomal and plasmid tetracycline-resistance gene (*tet*), which would result in tetracycline resistance when these two fragments joined by recombination. Three strains (wild type, the r*ecA* mutant, and the *recN* mutant) were used to test recombination proficiency, which was calculated based on the number of recombinants per surviving cell. The results from the short homology gap repair assay showed that there was a 3-fold increase in the recombination rate in the *recN* mutant compared to wild type cells, and a 23-fold increase in recombination rate in the *recA* mutant (these rates are measured by recombinants per surviving cell). These results clearly show that the RecN and

RecA proteins were not involved in λ Red-promoted gap repair. Since the absence of the RecN and RecA proteins resulted in higher rates of recombination, a conclusion can be made that these proteins are inhibitory to the λ Red-promoted gap repair in wild type cells.

The function of the RecN protein is not known; however this protein has been found to be expressed in high concentrations during the SOS response. It also has shown to bind to dsDNA ends (Sanchez and Alonso, 2005). As shown above, RecN has an inhibitory role in the λ Red pathway. The two proteins expressed in the λ Red pathway (Exo and Bet) also bind to dsDNA ends as a complex. Thus, inhibition due to the presence of RecN in the λ Red-promoted gap repair assay may arise because of competition between the λ Red proteins and RecN at broken dsDNA ends.

How could RecA, a known recombination pathway function, inhibit λ Red-promoted gap repair? The RecA protein is a single-stranded DNA binding protein that catalyzes ssDNA strand invasion during HR under the RecBCD pathway. The λ Red pathway is thought to promote ssDNA invasion at a replication fork (Poteete, manuscript submitted); on the contrary, RecA does not work by invading a replisome, but can promote strand invasion into non-replicating DNA. Thus, an explanation for the inhibition of λ Red-promoted gap repair may be that RecA competes with the λ Red proteins for dsDNA ends, which could potentially inhibit the replisome invasion pathway of λ Red.

A second substrate (the long homology substrate) used to measure gap repair contains a ~1000 bp homologous region within the chromosome in the lacZ region. The *lacZ* chromosomal target contains an insertion of the chloramphenicol-resistance gene (*cat*), which confers resistance to the antibiotic chloramphenicol (Cam). When recombination between the plasmid and the chromosome takes place, the *cat* gene is moved from the chromosome to the

plasmid. Due to its high copy number, the plasmid confers an increased resistance to Cam. To discriminate between the presence of the *cat* gene in the chromosome (in single copy) or on the plasmid (in multi-copy), different concentrations of chloramphenicol were tested with strains bearing the *cat* gene in both these locations. Chromosomal sensitivity to chloramphenicol was evident at 400 µg/ml, which was used in these experiments.

Three strains (wild type, *recA*, and *recN*) were again tested for both recombinants per surviving cell and recombinants per competent cell. In order to measure electrocompetence, the same *E. coli* preparation was used in the electroporations for the cut DNA substrate (pKM258 cut by *NotI*) and uncut (control) DNA substrate (pCF288). The results showed that there was a 3-fold increase in recombination rate in the *recN* mutant compared to wild type when measuring the recombinants per surviving cell. This was the same conclusion drawn from the short homology gap repair assay, as discussed previously. The results also show that the *recN* mutant and wild type had very similar recombinants per competent cell, demonstrating that RecN plays neither a stimulatory or inhibitory role in gap repair. However, there was also a 6-fold increase in the recombination rate for the *recA* mutant measured by recombinants per competent cell. This result supports the idea that the RecA protein may interfere in the plasmid-based gap repair assay. The measurements obtained by recombinants per competent cell are more accurate than the measurements from recombinants per surviving cell because the former values include a measurement of the electrocompetency of the cells.

Lower concentrations of the λ Red proteins were tested using the short homology plasmid. It was thought that lower concentrations of the λ Red proteins would bring about lower efficiencies of recombination activity. The pKM208 plasmid generates the λ Red proteins from multiple copy plasmid, while the KM22 strain produces the λ Red proteins from a single

chromosomal copy. The effect of the *recN* mutation in the KM22 genetic background were tested, where lower amounts of the λ Red proteins are expressed. The two strains (CF162 and CF163) used in this experiment were exposed to 50 J/m^2 of UV light (to induce RecN function) and then electroporated with a mixture of the short homology plasmid (pKM255 cut with *XhoI*) and a circular plasmid (pKM221). This internal control for competence (pKM221) was a better system for measuring the level of DNA substrate delivered to the desired strains since both test and control DNA substrates were in the same mixture. The strains isogenic to KM22 generated the same results as long homology gap repair; thus again, the RecN protein was not needed for λ Red- promoted gap repair.

In summary, the data presented in the study demonstrates that λ Red-promoted gap repair assay was not dependent on the RecN protein in *E. coli*. The data also showed that there was an inhibitory interaction between the λ Red and the RecA proteins in the λ Red-promoted gap repair assay. RecN may have a different function than previously believed. The real function of RecN could have to do with the alignment of chromosomes in the presence of multiple dsDNA breaks instead of aligning the breaks with the template chromosome. The gap repair assay would not work under this condition because representing one of these chromosomes with a cut plasmid is not likely to reveal the true role of RecN in vivo. To further investigate the function of RecN, the RecFOR pathway-promoted gap repair should be tested. RecN could have a role in the RecFOR pathway-promoted gap repair, where RecN is known to have a role in other types of RecFOR pathway-promoted recombination events, such as conjugational recombination.

BIBLIOGRAPHY

Carter, D. M. et al. (1971) Role of exonuclease and β protein of phage lambda in genetic recombination. II. Substrate specificity and the mode of action of lambda exonuclease. *J. Biol. Chem*., **246**: 2502 – 2510.

Connelly, John C. and Leach, David R. F. (2002) Tethering on the brink: The evolutionarily conserved Mre11-Rad50 complex. *TRENDS in Biochemical Sciences*., **27(8)**: 410 – 418.

Cromie, Gareth A., et al. (2001) Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Molecular Cell*., **8**: 1163 – 1174.

Datsenko, K.A. and B.L. Wanner. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA.*, **97(12)**: 6640 – 6645.

Finch, Paul Wet al. (1985) Identification of the *Escherichia coli recN* gene product as a major SOS protein. *Journal of Bacteriology*., **164 (2)**: 653 – 658.

Haering, Christian H., et al. (2002) Molecular architecture of SMC proteins and the yeast cohesin complex. *Molecular Cell*., **9**: 773 – 788.

Kidane, Dawit and Graumann, Peter L. (2005) Intracellular protein and DNA dynamics in competent *Bacillus subtilis* cells. *Cell*., **122**: 73 – 84.

Kidane, Dawit, et al. (2004) Visualization of DNA double-strand break repair in live bacteria reveals dynamic recruitment of *Bacillus subtilis* RecF, RecO, and RecN Proteins to distinct sites on the nucleoids. *Molecular Microbiology*., **52(6)**: 1627 – 1639.

Kmiec, E. et al. (1981) β protein of bacteriophage λ promotes renaturation of DNA. *J. Biol. Chem.* **250**: 7377 – 7387.

Kuzminov, Andrei. (1999) Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ. *Microbiology and Molecular Biology Reviews*., **63(4)**: 751 – 813.

Lieber, Michael R. (1999) The biochemistry and biological significance of non-homologous DNA end joining: An essential repair process in multicellular eukaryotes. *Genes to Cells*., **4**: 77 – 85.

Little, J. W. (1967) An exonuclease induced by bacteriophage λ. II. Nature of the enzymic reaction. *J. Biol. Chem*., **242**: 679.

McKenzie, Gregory J. et al. (2000) The SOS response regulates adaptive mutation. *PNAS*., **97(12)**: 6645 – 6651.

Meddows, Tom R., et al. (2005) RecN protein and transcription factor DksA combine to promote faithful recombinational repair of DNA double-strand breaks. *Molecular Microbiology*., **57(1)**: 97 – 110.

Miller, J.H. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory. Cold Springs Harbor, N.Y.

Muniyappa, K. et al. (1986) The homologous recombination system of phage λ: Pairing activities of β protein. *J. Biol. Chem*., **261**: 7472 – 7478.

Murphy, Kenan C. (1991) λ Gam protein inhibits the helicase and χ-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. *J. Bacteriol*. **173**: 5808 – 5821.

Murphy, Kenan C. (1998) Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *Journal of Bacteriology*., **180(8)**: 2063 – 2071.

Murphy, Kenan C. and Campellone, Kenneth G. (2003) Lambda red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. *BMC Molecular Biology*., **4**:1 – 12.

Murphy, Kenan C. (2007) The lambda Gam protein inhibits RecBCD binding to dsDNA ends. *J. Mol. Biol*., **371(1)**: 19 – 24.

Nagashima, Kohji, et al. (2006) Degradation of *Escherichia coli* RecN aggregates by ClpXP protease and its implications of DNA damage tolerance. *Journal of Biological Chemistry*., **281(41)**: 30941 – 30946.

Neher, Saskia B., et al. (2006) Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Molecular Cell*., **(22)**: 193 – 204.

Peterson, Kenneth R., et al. (1988) Derepression of specific genes promotes DNA repair and mutagenesis in *Escherichia coli*. *Journal of Bacteriology*., **170(1)**: 1 – 4.

Poteete, A. R. (2004) Modulation of DNA repair and recombination by the bacteriophage lambda Orf function in *Escherichia coli* K-12. *J. Bacteriol*., **186(9)**: 2699 – 2707.

Rocha, Eduardo P. C., et al. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genetics*., **1(2)**: 247 – 259.

Rostas, Katalin, et al. (1987) Nucleotide sequence and LexA regulation of the *Escherichia coli recN* gene. *Nucleic Acids Research*., **15(3)**: 5041 – 5049.

Sanchez, Humberto, and Alonso, Juan C. (2005) *Bacillus subtilis* RecN binds and protects 3'-single-stranded DNA extensions in the presence of ATP. *Nucleic Acid Research*., **33(7)**: 2343 – 2350.

Sanchez, Humberto, et al. (2006) Recruitment of *Bacillus subtilis* in RecN to DNA double-strand breaks in the absence of DNA end processing. *Journal of Bacteriology*., **188(2)**: 353 – 360.

Sanchez, Humberto, et al. (2007) Dynamic structures of *Bacillus subtilis* RecN-DNA complexes. *Nucleic Acids Research*., **36(1)**: 110 – 120.

Singleton, Martin R., et al. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*., **432**: 187 – 193.

Voet, Donald, Judith G. Voet, and Pratt, Charlotte W. (2006) "Recombination Repair." Fundamentals of Biochemistry: Life at the Molecular Level. Wiley. pg. 908.

Yu, D., et al. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*., **97(11)**: 5978 – 5983.

Zhang, Y., et al. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.*, **20(2)**: 123 – 128.

APPENDIX 1

Small Gap repair data

Date	Strain	Plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
9/11/07	126	p255	236	18		1000+	139	0.34 x 10^{-4}
9/11/07	126	p255	272	23		2000+	238	0.23 x 10^{-4}
9/11/07	133	p255	1000+	76		1000+	163	0.93 x 10^{-4}
9/11/07	133	p255	1000+	78		2000+	241	0.45 x 10^{-4}
9/14/07	126	p255	362	37		1000+	165	0.44 x 10^{-4}
9/14/07	126	p255	208	14		1000+	117	0.36 x 10^{-4}
9/14/07	129	p255	39	2		6	2	13.0 x 10^{-4}

9/14/07	129	p255	30	0		9	0	6.7×10^{-4}
(Replate)								
9/14/07	*129*	*p255*	*2000+*	*927*	*2000+*	*1000+*	*342*	*5.4×10^{-4}*
9/21/07	126	p255	242	27		1000+	107	0.45×10^{-4}
9/21/07	133	p255	112	11		137	8	1.6×10^{-4}
9/21/07	133	p255	159	58		318	56	1.0×10^{-4}
9/21/07	151	p255	82	7		471	55	0.29×10^{-4}
9/21/07	151	p255	107	10		598	76	0.28×10^{-4}

Recombinant / Competent cell data:

Plasmids electroporated are p255 (XhoI) 100 ng and pBR322 (50 ng)

Date	Strain	plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
2/6/08	162 (-)	p255	43	2	0			
		pBR322			1000+	224	24	0.19 x 10^{-3}
2/6/08	162 (+)	p255	1000+	314	37			
		pBR322			179	22	0	175 x 10^{-3}
2/6/08	163 (-)	p255	1000+	147	12			
		pBR322			190	26	4	77.4 x 10^{-3}
2/6/08	163 (+)	p255	528	57	2			
		pBR322			1000+	177	23	3.22 x 10^{-3}

Plasmids electroporated are p255 (XhoI) 100 ng and pKM221 (100 ng)

Date	Strain	plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
2/6/08	162 (-)	p255	28	2	0			
		p221			107	10	0	2.6×10^{-3}
2/6/08	162 (+)	p255	353	37	7			
		p221			447	50	3	7.89×10^{-3}
2/6/08	163 (-)	p255	303	36	1			
		p221			500+	114	15	2.7×10^{-3}
2/6/08	163 (+)	p255	365	43	4			
		p221			277	23	2	10.0×10^{-3}
2/12/08	162 (-)	p255	22	1	0			
		p221			108	14	1	2.04×10^{-3}
2/2/08	162 (+)	p255	122	8	0			
		p221			363	45	3	3.27×10^{-3}

2/12/08	163 (-)	p255	140	7	0			
		p221			399	48	4	3.55 x 10^{-3}
2/12/08	163 (+)	p255	47	6	0			
		p221			71	7	2	6.62 x 10^{-3}

Replate experiment of 2/10/08

Date	Strain	plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
2/6/08	162 (-)	p255	3	1		1000+	161	
2/6/08	162 (+)	p255	107	10		320	37	
2/6/08	163 (-)	p255	103	13		1000+	76	
2/6/08	163 (+)	p255	26	1		470	58	

Large Gap repair data

(33 ng)

Date	Strain	Plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
10/31/07	153	p258	125	2			443	0.056×10^{-4}
11/16/07	153	p258	300	17			62	0.96×10^{-4}
11/19/07	153	p258	82	5			90	0.18×10^{-4}
11/16/07	155	p258	130	10			27	0.96×10^{-4}
11/19/07	155	p258	47	8			8	1.2×10^{-4}

(100 ng)

Date	Strain	Plasmid	Undiluted	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Frequency
12/7/07	153	p258	1000+	98				
12/7/07	153	p288		2000+	1000+	313		3.13×10^{-3}

12/16/07	154	p258	1000+	73			
12/16/07	154	p288		1000+	225	48	32.4 x 10^{-3}
12/19/07	154	p258	1000+	67			
12/19/07	154	p288		2000+	1000+	97	7.0 x 10^{-3}
12/7/07	155	p258	1000+	95			
12/7/07	155	p288		2000+	1000+	257	3.69 x 10^{-3}