The Development and Testing of a System

for Monitoring Site-Specific Lesions *In Vivo*

by

Melissa Pierce Asllani

Department of Biochemistry
Duke University

Date:_______________________
Approved:

___________________________
Kenneth N. Kreuzer, Supervisor

___________________________
Richard Brennan

___________________________
Arno L. Greenleaf

___________________________
Sue Jinks-Robertson

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Biochemistry in the Graduate School
of Duke University

2013
ABSTRACT

The Development and Testing of a System for Monitoring Site-Specific Lesions In Vivo

by

Melissa Pierce Asllani

Department of Biochemistry
Duke University

Date:_______________________
Approved:

___________________________
Kenneth N. Kreuzer, Supervisor

___________________________
Richard Brennan

___________________________
Arno L. Greenleaf

___________________________
Sue Jinks-Robertson

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2013
Abstract

Every day, cells face agents that generate lesions in genomic DNA, which can interfere with the processes of DNA replication and gene expression. These lesions can range from small abasic sites to alkylated bases to large proteins frozen on the DNA and can be caused by both endogenous and exogenous agents. These lesions must be repaired to maintain genomic stability, and multiple pathways exist to perform the necessary repairs or to bypass the damage. These pathways have been discovered and studied using a variety of experimental techniques, both in vitro and in vivo. While these studies have contributed valuable information about many of cellular processes, there are still gaps in the DNA repair field.

The goal of this study is to bridge some of those gaps by constructing a system to introduce DNA containing a site-specific lesion into *Escherichia coli* cells at high enough levels to monitor the lesion’s fate in vivo and in real-time. This system combines two separate DNA molecules to simplify the introduction of a site-specific lesion. The first molecule is the DNA from bacteriophage λ, a virus that is able to infect *E. coli* cells at a high level of efficiency. A typical commercial packaging reaction can yield titers of approximately 1.0 x 10⁹ plaque forming units (PFU)/mL. However, bacteriophage λ has a large genome of approximately 48.5 kb, which makes it a difficult substrate for extensive cloning and manipulation. In contrast, cloning and manipulation of a small
plasmid (~4 kb) is a much simpler endeavor, and small plasmids have been used previously to produce DNA containing a site-specific lesion. The problem with using a plasmid occurs when attempting to introduce it into cells, as the process of transformation is not very efficient and can cause unintended consequences in the cells. This new system allows for the incorporation of the lesion into the plasmid, which is then integrated into a bacteriophage λ vector, λ Kytos. The combination of these two molecules produces bacteriophage λ DNA containing a site-specific lesion, which can infect the cells at high efficiency, allowing the fate of the DNA to be monitored in real-time. Interestingly, repair of a single EthenoA lesion after infection appears to be a very inefficient process. Even if the repair system is activated by induction with methyl methanesulfonate (MMS) or if individual repair proteins are overexpressed, little to no repair occurs. As methylation of the DNA occurs upon injection, it does appear that the DNA is exposed to proteins in the cell, including any repair proteins present. These results indicate that other processes, perhaps replication or transcription, are required to repair a single EthenoA lesion in vivo.
## Contents

Abstract ......................................................................................................................................... iv

List of Tables ................................................................................................................................. ix

List of Figures ................................................................................................................................. x

List of Abbreviations ..................................................................................................................... xii

Acknowledgements ....................................................................................................................... xiv

1. Introduction ................................................................................................................................. 1

  1.1 DNA repair pathways ................................................................................................................ 1

    1.1.1. Base excision repair ........................................................................................................... 1

    1.1.2. Nucleotide excision repair ............................................................................................... 3

    1.1.3. Mismatch repair .................................................................................................................. 5

    1.1.4. Recombinational repair ..................................................................................................... 6

    1.1.5. Direct reversal pathways ................................................................................................. 8

  1.2 DNA lesions and repair ............................................................................................................ 11

    1.2.1. *In vitro* and *in vivo* studies with generalized damage ..................................................... 11

    1.2.2. *In vitro* studies with site-specific damage ..................................................................... 15

    1.2.3. *In vivo* studies with site-specific damage ..................................................................... 18

  1.3 DNA lesions and fork blockage .............................................................................................. 20

    1.3.1. *In vitro* studies ............................................................................................................. 20

    1.3.2. *In vivo* studies .............................................................................................................. 21

  1.4 Advantages of the new system ............................................................................................... 23
2. Cre-loc recombination ................................................................. 27
   2.1 Introduction ........................................................................... 27
   2.2 Materials and Methods ....................................................... 29
   2.3 Results .................................................................................. 34
   2.4 Discussion ............................................................................ 42
3. Addition of the abasic site lesion .................................................. 44
   3.1 Introduction ........................................................................... 44
   3.2 Materials and Methods ....................................................... 47
   3.3 Results .................................................................................. 50
   3.4 Discussion ............................................................................ 55
4. Construction of the working system .............................................. 57
   4.1 Introduction ........................................................................... 57
   4.2 Materials and Methods ....................................................... 60
   4.3 Results .................................................................................. 65
   4.4 Discussion ............................................................................ 79
5. In vivo repair of the EthenoA lesion .............................................. 84
   5.1 Introduction ........................................................................... 84
   5.2 Materials and Methods ....................................................... 89
   5.3 Results .................................................................................. 92
   5.4 Discussion ............................................................................ 105
6. Conclusions and future directions ............................................... 113
References ........................................................................................ 121
List of Tables

Table 1. DNA constructs. ......................................................................................... 30

Table 2. Probe constructs for Southern blotting. ...................................................... 92
List of Figures

Figure 1. Base excision repair .............................................................. 2
Figure 2. Nucleotide excision repair ...................................................... 4
Figure 3. Mismatch repair pathway ........................................................ 6
Figure 4. Recombinational repair of double strand breaks ...................... 7
Figure 5. Cre binding to the loxP site .................................................. 27
Figure 6. Irreversible recombination of mutant lox sites ......................... 29
Figure 7. Restriction site insert .............................................................. 34
Figure 8. Intramolecular Cre recombination ......................................... 37
Figure 9. Intermolecular Cre recombination .......................................... 38
Figure 10. Theoretical recombination of λgt11 vector and damaged plasmid 39
Figure 11. Maximum recombination efficiency with λgt11 vector and pMPA1 and pMPA2 41
Figure 12. New restriction site insert .................................................. 50
Figure 13. The dSpacer lesion and its incorporation .................................. 52
Figure 14. Repair of the dSpacer lesion in packaging extracts .................. 54
Figure 15. φC31 integration reaction ...................................................... 58
Figure 16. Creation of and infection with damaged DNA ......................... 67
Figure 17. Production of wild-type and damaged DNA .............................. 69
Figure 18. The EthenoA lesion and its incorporation ................................ 71
Figure 19. Stability of the EthenoA lesion in packaging extracts ............... 73
Figure 20. Schematic integration of λ Kytos and pMPA4 .......................... 74
Figure 21. *In vitro* integration reaction ................................................................. 76

Figure 22. Packaging of the integrated λ DNA ................................................................... 78

Figure 23. Test of *in vivo* repair of EthenoA-containing DNA ........................................... 94

Figure 24. Test of *in vivo* repair of EthenoA-containing DNA in the presence of AlkA or AlkB overexpression ........................................................................... 97

Figure 25. Methylation status of infecting DNA. ............................................................... 101

Figure 26. Observing DNA methylation and repair ............................................................ 104

Figure 27. System for monitoring DNA repair and replication ........................................... 117

Figure 28. Schematic gel of digests to study repair and replication .................................... 118
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ada</td>
<td>O(^6) alkyl guanine transferase I</td>
</tr>
<tr>
<td>BER</td>
<td>base excision repair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CAA</td>
<td>chloroacetaldehyde</td>
</tr>
<tr>
<td>CRAB</td>
<td>competitive replication of adduct bypass</td>
</tr>
<tr>
<td>dA</td>
<td>1,(N^6)-ethenoadenine</td>
</tr>
<tr>
<td>DMS</td>
<td>dimethylsulfate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EthenoA</td>
<td>1,(N^6)-ethenoadenine</td>
</tr>
<tr>
<td>Fpg</td>
<td>formamidopyrimidine DNA N-glycosylase</td>
</tr>
<tr>
<td>HPLC</td>
<td>high-performance liquid chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-beta-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Broth</td>
</tr>
<tr>
<td>MMR</td>
<td>mismatch repair</td>
</tr>
<tr>
<td>MMS</td>
<td>methyl methanesulfonate</td>
</tr>
<tr>
<td>MNNG</td>
<td>N-methyl-N'-nitro-N-nitrosoguanidine</td>
</tr>
<tr>
<td>MNU</td>
<td>N-methyl-N-nitrosourea</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>NER</td>
<td>nucleotide excision repair</td>
</tr>
<tr>
<td>Nfo</td>
<td>Endonuclease IV</td>
</tr>
<tr>
<td>Nth</td>
<td>Endonuclease III</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PFU</td>
<td>plaque forming units</td>
</tr>
<tr>
<td>RDF</td>
<td>recombination directionality factor</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>real-time polymerase chain reaction</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TCR</td>
<td>transcription-coupled repair</td>
</tr>
<tr>
<td><strong>Ung</strong></td>
<td>uracil DNA glycosylase</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
</tr>
<tr>
<td><strong>X-gal</strong></td>
<td>5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside</td>
</tr>
<tr>
<td><strong>Xth</strong></td>
<td>Exonuclease III</td>
</tr>
</tbody>
</table>
Acknowledgements

First and foremost, I would like to thank Ken for allowing me to join his lab after leaving the lab of my first advisor. I was extremely worried that I would not be able to find someone to work with, but Ken was quick to accept me into the lab. He then invested a large amount of time mentoring me and supervising my new project. He is always willing to listen to new ideas and look at new data but also encourages students to think for themselves, plan their own experiments, and draw their own conclusions. I think it is these critical thinking skills that will be most valuable in my future career. I am extremely grateful for his support and teaching during my time at Duke.

Next, I would like to thank my lab mates, both past and present, for their helpful discussions and ideas regarding my project. Additionally, I would like to thank them for providing a fun working environment.

I am extremely grateful to have been raised by two wonderful parents who have supported me in all the choices I have made. I have always been encouraged to learn and do well in school, but they never pressured me to do too much or to do things that did not truly interest me. They are also great listeners, willing to hear me talk about my day in lab or my results, even if they had no idea what I was talking about! My brother has also been very supportive and always urges me to “get done and get back to New England!” And my niece can always put a smile on my face if I am having a bad day. I
also have a great extended family, and I know they will be very excited for me as I complete my graduate career.

Finally, I am extremely thankful for my husband. I cannot imagine what graduate school would have been like without him. It is extremely reassuring to have someone at home who knows what you go through every day in lab and can provide support, both emotional and scientific. We have already had many great adventures together as a duo, and I look forward to continuing those adventures as a trio in February!
1. Introduction

1.1 DNA repair pathways

Cells face DNA damaging agents every day. These agents, whether endogenous or exogenous, can create DNA lesions that must be repaired to maintain genomic stability. If left unrepaired, these lesions can interfere with the processes of DNA replication and gene expression, which can be lethal to the cell. In *E. coli*, a variety of pathways exist to repair the many types of lesions that can arise.

The overall goal of this work was to develop a new system to study DNA repair and replication *in vivo*. This system was designed to incorporate larger amounts of DNA into cells, which will allow the DNA’s repair and replication to be followed in real-time. The capabilities of this system offer advantages over tools and techniques used in previous studies. These studies and the information that they provided will be described in the following sections.

1.1.1. Base excision repair

The base excision repair (BER) pathway is able to repair lesions in which the base of the nucleotide has been damaged or removed. These lesions are often small and non-helix distorting (*I*). This pathway begins with removal of the damaged or incorrect base by a DNA N-glycosylase. A variety of N-glycosylases are present in prokaryotes, and these enzymes can remove a specific type of damage or multiple types of damage.
Endonuclease III (Nth) is encoded by the nth gene and acts as an N-glycosylase to remove thymine glycol from DNA (2). Thymine glycol is the main product created by damage to thymine residues by oxidizing agents, oxygen-containing free radicals, or ionizing radiation (3-9). The formamidopyrimidine DNA N-glycosylase (Fpg) protein can remove imidazole ring-opened purines often created by alkylating agents (10). The AlkA N-glycosylase can also remove damage created by alkylating agents and has been shown to remove a variety of structurally diverse lesions (see (11) for review). One of those lesions is 1,N6-ethenoadenine (EthenoA or dA), a cyclized adenine formed upon exposure to alkylating agents. This lesion and its repair will be discussed in detail in Chapter 3.

![Figure 1. Base excision repair.](image)
The base excision repair pathway is responsible for the removal of small, non-helix distorting DNA lesions. The steps of the pathway are described in detail in the text. Adapted from (12).

Removal of the lesion by a N-glycosylase generates an abasic site. This site is the common intermediate acted upon by the downstream enzymes in the BER pathway (Figure 1). After base removal, endonucleases incise the DNA just 5’ to the lesion; the main endonucleases in E. coli are exonuclease III (XthA) and endonuclease IV (Nfo). After DNA incision, a short exonuclease reaction occurs, followed by DNA polymerization and ligation (13-15). The short patch subpathway of BER removes only one nucleotide, whereas long patch BER removes multiple nucleotides (see (16, 17) for review).

1.1.2. Nucleotide excision repair

Nucleotide excision repair (NER) is performed by the UvrABC complex and the pathway is illustrated in Figure 2. The UvrA:B complex forms to probe the DNA for distortions caused by damage. The UvrA dimer is thought to dissociate, which allows UvrC to bind UvrB and incise the DNA (18). All three subunits of the complex are required for repair. The NER system can remove a variety of lesions and is the main pathway involved in removing the bulky adducts that result from UV irradiation, such as pyrimidine dimers and 6-4 photoproducts. In contrast with BER, which removes only a few nucleotides, NER removes a string of nucleotides. The UvrBC complex cleaves the
eighth phosphodiester bond 5’ and the fifth phosphodiester bond 3’ to the lesion (19).

The combined action of UvrD, a DNA helicase, polymerase I, and ligase removes the
damage-containing string of nucleotides and fills the single strand gap (20).

Figure 2. Nucleotide excision repair.
The nucleotide excision repair pathway is responsible for the removal of large, bulky lesions. The steps of the pathway are described in the text. Adapted from (21).

Various studies have indicted that NER is linked to another repair pathway, transcription-coupled repair (TCR). TCR is necessary when unrepaired DNA lesions are located in actively transcribed regions; these lesions can lead to blockage of the RNA polymerase (22). This blockage is removed by Mfd (23). In addition to removing the stalled RNA polymerase, Mfd is able to recruit the UvrA protein to repair the lesion by TCR (22, 24, 25).

1.1.3. Mismatch repair

The mismatch repair (MMR) system (Figure 3) removes mispaired bases in DNA due to the incorporation of an incorrect nucleotide, the incorporation of a chemically damaged nucleotide, and the creation of loops resulting from strand slippage. The essential proteins involved in MMR in *E. coli* include MutS, MutL, MutH, and UvrD. MutS initiates repair by recognizing and binding to mismatched base pair(s). It then recruits MutL to the mismatch. This complex activates MutH, an endonuclease that nicks unmodified GATC sequences. This unmodified sequence is crucial in directing DNA repair. The GATC sequence is normally methylated in *E. coli* DNA. After replication, there is a brief delay before the daughter DNA strand is methylated. During this delay, the parental and daughter strands can be distinguished, and the daughter strand that contains the mismatch can be determined and nicked by MutH. UvrD then
unwinds the DNA helix in the direction of the mismatch and removes it. DNA polymerase III holoenzyme and ligase repair the gap (see (26, 27) for review).

Figure 3.Mismatch repair pathway.

The mismatch repair pathway is responsible for repairing mispaired DNA bases and DNA loops. The steps in the pathway are described in the text. Adapted from (26).

1.1.4. Recombinational repair

The recombinational repair pathway is the main pathway for the repair of double strand DNA breaks, and this repair is shown in Figure 4 (see (28, 29) for reviews). Upon the creation of a double strand break, the ends are degraded by RecBCD, a heterotrimeric helicase/nuclease, to produce 3’ single strand ends or tails. These ends are then coated
by RecA. The RecA/DNA complex then invades a homologous region to establish a D-loop structure with accessory proteins, including SSB and RecFOR, and form a double Holliday junction. The RuvAB complex and RecG protein then migrate the double Holliday junction, which is finally resolved by RuvC (30). Additionally, if the replication fork encounters a DNA lesion on the leading or lagging strand, it can cause the fork to stall. This stalling can result in the breakage of one of the DNA arms to create a double strand end, which can be used for fork restart via a modified homologous recombination pathway. Daughter strand gap repair is also considered a recombinational repair.

Figure 4. Recombinational repair of double strand breaks.
Recombinational repair is responsible for the repair of double strand breaks. The steps of the pathway are described in the text. Adapted from (28).

pathway. RecA and the RecFOR proteins pair and initiate strand exchange between two homologous strands, one of which contains a gap. This exchange results in the formation of a Holliday junction. The junction is then branch migrated by RecA, RuvAB, and RecG. The branch may be resolved by cleavage of the Holliday junction by RuvC or by reverse branch migration by RecG and RuvAB. This pathway acts to fill the gaps in a daughter strand of DNA generated by lesion bypass (31, 32).

1.1.5. Direct reversal pathways

Three different types of lesions can be chemically reversed. These lesions include pyrimidine dimers, methylated guanine bases, and some types of methylation of cytosine and adenine. These direct reversal pathways do not require a template and do not break the phosphodiester backbone.

Pyrimidine dimers are reversed in a process called photoreactivation, which, as the name implies, requires light of a specific wavelength. The enzymes that catalyze this process are known as photolyases. In E. coli, this enzyme is known as pyrimidine dimer-DNA (PD-DNA) photolyase. The reversal reaction consists of two distinct stages: a dark reaction and a light reaction. During the dark reaction, the enzyme binds to the dimer at least 100 times more effectively than it binds to unirradiated DNA (33). After binding to the lesion, the light reaction can occur if light at wavelengths between 300 and 500 nm is
present. This light is harnessed by the presence of chromophores in the PD-DNA photolyase. In *E. coli*, these chromophores are 1,5-dihydroflavin adenine dinucleotide (FADH$_2$ or FADH), a stable neutral radical (34), and 5,10-methenyltetrahydrofolyl polyglutamate (MTHF) (35). MTHF is believed to function as the primary light-harvesting cofactor; this harvest then triggers a series of photochemical reactions that result in the resolution of the dimer (36).

The methylation of guanine can also be repaired via a direct reversal pathway. O$^6$-methylguanine is one of many lesions that can be generated by alkylating agents, which include a wide variety of chemicals (37). This methyl group can be directly removed by two different alkyltransferases. O$^6$-alkylguanine DNA alkyltransferase I (*O*$^6$-AGT I), which is also known as Ada, is able to overcome an energy barrier to remove O$^6$-methylguanine because this lesion is relatively stable at neutral pH (38). The enzyme is able to achieve this function without a divalent cation or other high-energy cofactor. Additionally, there is no indication that peptid e bond cleavage is associated with removal of the methyl group (39). However, Ada is not a typical enzyme. Enzyme activity is consumed as the reaction occurs, and transfer of the methyl group will cease when a limiting amount of Ada is present (39, 40). These results indicate that Ada is a suicide enzyme. Other enzymes are inactivated by suicide inhibitors, but it is unusual that the suicide reaction occurs between the enzyme and its natural substrate as opposed
to a substrate analog (41). However, the alkylated version of the Ada enzyme does play a role as a transcription activator of a regulon of genes that is involved in the direct reversal of methylated cytosine and adenine. This regulon will be discussed in the next paragraph. $O^6$ AGT I can also remove larger alkyl groups (42, 43). The second alkyltransferase, $O^6$-alkylguanine DNA alkyltransferase II ($O^6$ AGT II), is encoded by the $ogt$ gene. Previous work indicates that $O^6$ AGT II is responsible for the constitutive alkyltransferase activity present in cells prior to optimal Ada induction (44). Like Ada, $O^6$ AGT II is a suicide enzyme (45), which makes the direct reversal of alkylation damage an expensive process. Ada has a distinct preference for $O^6$-methylguanine, while $O^6$ AGT II has a preference for $O^4$-methylguanine (46).

Alkylated Ada serves as a positive transcription factor of the ada regulon. This regulon consists of four genes, $ada$, $alkA$, $alkB$, and $aidB$, which will not be discussed (38). The $ada$ gene was described in the previous paragraph, and the $alkA$ gene was mentioned in the section on BER as a $N$-DNA glycosylase. The $alkB$ gene encodes the AlkB protein, which is able to catalyze the third direct reversal pathway, the repair of alkylated cytosine and adenine. AlkB was determined to repair the alkylation products $N^1$-methyladenine and $N^3$-methylcytosine (47). The biology of this enzyme and its additional substrates will be discussed in detail in Chapter 5.
1.2 DNA lesions and repair

Many different studies have been performed with many different types of DNA lesions, and these studies have provided extremely informative results that describe how the DNA lesions are repaired, the enzymes involved, and the consequences if repair cannot be completed. The following sections will detail some of the \textit{in vitro} and \textit{in vivo} studies performed to determine the fate of DNA lesions.

1.2.1. \textit{In vitro} and \textit{in vivo} studies with generalized damage

When DNA repair was first studied, early experiments involved fractionating \textit{E. coli} lysates and looking for specific enzymatic activities or isolating mutant \textit{E. coli} cells that were susceptible to agents that cause damage. Both approaches yielded important results and will be highlighted in the following examples. Additionally, these approaches often served to complement and inform one another.

One of these examples occurred during the isolation and identification of exonuclease III, which was originally named endonuclease II. In the original work performed by Friedberg and Goldthwait (48), a fraction from \textit{E. coli} was found to attack alkylated DNA and cause single-strand breaks. This activity was determined by incubating the \textit{E. coli} fractions with alkylated [\textsuperscript{3}H]T4 DNA and screening for the release of radioactivity. During early studies, the release of radioactive DNA was a common assay for endonuclease or exonuclease activity (49, 50). The fraction was further purified
to isolate the enzyme responsible for the activity, and some of the enzyme’s properties were determined. Once an enzyme of interest was found and an assay for its activity was established, genetic studies could be performed to determine the gene encoding the protein. By mutagenizing AB1157 cells with \( N \)-methyl-\( N' \)-nitro-\( N \)-nitroso-guanidine, screening for deficiencies in their ability to release radioactive DNA, and mapping the genetic locus for this deficiency, exonuclease III activity was found to be encoded by the \( \text{xth} \) gene (51), located between the \( \text{aroD} \) and \( \text{pncA} \) genes at approximately 33 min (14, 52). The use of these types of experiments to find and characterize DNA damage repair enzymes opened the door to many other studies, both \textit{in vivo} and \textit{in vitro}, with different DNA damaging agents.

A variety of \textit{in vivo} studies have been performed and yielded many different types of information. A study performed by Kuenmerle and Masker examined the effect of UV irradiation, which was known to form pyrimidine dimers, on wild-type \textit{E. coli} and \textit{E. coli} containing a mutation of the \textit{uvrD} gene (53). This UV hypersensitive mutant was used to study survival upon exposure to UV irradiation and the ability to catalyze host cell reactivation. Specific experiments were also performed to determine the fate of DNA from cells. Cells were grown in the presence of \textit{[^3H]thymine} and then irradiated. The DNA was harvested and analyzed for incision at UV damage sites, dimer excision, and DNA resynthesis. The results of these experiments demonstrated that the \textit{uvrD} mutant...
was able to incise the DNA but could not completely excise the dimer or resynthesize the DNA.

As more was learned about the enzymes involved in DNA repair, different types of questions could be asked and answered. By using a genetic reversion assay, one could examine the actual base changes produced by mutation of the ung gene, which encodes uracil DNA glycosylase (54). The effect of the mutant on cell survival was also determined. Studies that examined the effect of different DNA damaging agents on the survival of DNA repair mutants also provided important information about the type of DNA damage produced and the enzymes necessary to repair it (55, 56). Studies with various E. coli mutants also provided important information about which enzymes demonstrate overlapping functions in vivo. For example, it was shown that mutation of the xth or nfo genes resulted in wild-type survival levels after UV irradiation. However, the mutation of both enzymes resulted in reduced survival, which indicated that the enzymes function in the same pathway, as one enzyme can substitute for the other (57). In this same study, the authors could not construct the xthA nfo uvrA triple mutant, indicating that DNA repair in this strain was severely compromised, preventing cell survival. These many types of in vivo studies provided valuable information about the effect of DNA repair mutants on survival and the type of mutations that these mutants suffer after DNA damage.
A recent study by Jeiranian et al. (58) used a system that is most similar to the one described in this work. The authors transformed UV-irradiated plasmids into various *E. coli* knockout strains to determine which proteins were important in plasmid survival. Based on their previous results, they stated that the UV dosage would create one lesion per 10-kb strand of DNA, though the specific location of the lesion cannot be determined. Their results revealed that the RecA-mediated pathways (RecBC, RecF, and RuvABC) that contribute to cell survival after UV irradiation did not contribute to survival of the UV-irradiated plasmid. Only the lack of UvrA affected plasmid survival, and UvrA was found to be nearly essential, reducing plasmid survival by two orders of magnitude. Finally, the authors used two-dimensional (2D) electrophoresis to observe replication of the newly introduced plasmid and found that it did not immediately replicate upon transformation, which would explain the lack of involvement of RecA-mediated pathways.

*In vitro* studies with generalized damage also provided important information about DNA repair. An early study examined the incorporation and excision of 5-fluorouracil (FU) (59). First, cells were grown in the presence of 5-[6-14C]FU, and it was determined by paper chromatography that only the DNA of the dUTP pyrophosphatase (*dut*)/uracil DNA glycosylase (*ung*) double mutant contained a large amount of radioactive 5-FU. Additional experiments were performed with purified Ung protein to
determine its excision activity on DNA containing uracil or 5-FU, and the rate of release of uracil was found to be 20-fold faster than the release of 5-FU. The experiment provided a better idea of what was happening to the DNA lesion as excision of the lesion itself was measured, rather than the excision of DNA in general. Many other in vitro studies were performed with purified DNA repair proteins. Tag glycosylase, encoded by tag, was purified and determined to excise 3-methyladenine from DNA exposed to simple alkylating agents, such as methylmethane sulfonate (MMS) and N-methyl-N-nitrosourea (MNU). The release of 3-methyladenine was monitored by reverse phase high-performance liquid chromatography (HPLC). Additional studies have been performed, and there are too many to mention here. One feature is that they often involve purification of the repair enzyme and testing of its activity, both to determine characteristics of the enzyme such as its K_m or metal requirement (10) and to determine its possible substrates (60).

1.2.2. In vitro studies with site-specific damage

As more information was gathered about the DNA repair pathways, experiments were designed to examine a single, site-specific lesion rather than large amounts of generalized DNA damage. These types of experiments offer the advantage of studying a lesion in a defined sequence, which creates multiple strategies for monitoring the damage. Additionally, DNA damaging agents can cause multiple types of lesions, and
many of the previous studies looked at general measures of enzyme activity, such as the release of radioactivity, which is assumed to be caused by DNA excision. However, if there are multiple types of damage and a general monitoring technique is used, it is impossible to determine if one type of lesion is repaired or if all lesions are repaired. Using a site-specific lesion eliminates some of these caveats.

Several early studies used a site-specific lesion in an oligonucleotide. This type of experiment allowed the binding, activity, and other important properties of repair enzymes to be characterized. In one of these studies, a single psoralen adduct was used to examine the binding and activity of NER enzymes. Purified UvrA, UvrB, UvrC, and other proteins were used to determine the order of binding events, the requirement for ATP binding and hydrolysis, and postincision events (20). In a second study of this type, experiments were performed with a variety of DNA lesions, including an abasic site, a reduced abasic site, and various modified pyrimidines, to determine the activity of endonuclease VIII. The $K_d$, binding stoichiometry, and binding footprint were also determined (61). Another study used oligonucleotides depurinated by acid treatment to create an abasic site to directly measure DNA hydrolysis by endonuclease IV (62); only indirect evidence of the activity of endonuclease IV had been obtained before this study (63).
As more was learned about the DNA repair proteins, the questions asked became more complicated. *In vitro* experiments with site-specific lesions were able to solve some of this complexity. By testing a single type of lesion, the specificity of the enzyme can be probed, and inhibitors of the enzyme can be found. *E. coli* Fpg protein was known to excise a variety of DNA lesions. However, little was known about the complex between the enzyme and its DNA substrate. By determining that a reduced abasic site could be bound but not excised by Fpg protein, a substrate was developed that allowed the DNA-protein interactions to be probed without consumption of the DNA (64). Complex kinetic questions could also be answered. For example, the kinetic behavior of exonuclease III, endonuclease IV, and endonuclease III towards the normal abasic site and two oxidized abasic sites, the C4’-oxidized abasic site and 2-deoxyribo lactone, was determined. Researchers were able to determine the $K_m$, $k_{cat}$, and $k_{cat}/K_m$ of each enzyme for each lesion (65).

Finally, these types of experiments can be used to achieve one of the ultimate goals of understanding enzyme activity and mechanism – being able to “see” the enzyme as it binds its substrate (66). X-ray crystallography has been used to study the interaction of DNA repair enzymes with their substrates to provide a picture of how these enzymes “find” DNA damage (see (1) for a review of structures from the BER pathway).
1.2.3. *In vivo* studies with site-specific damage

It is commonly accepted that *in vivo* studies provide a more complete idea of the events that occur in cells as the DNA is exposed to all cellular components not just the proteins or factors included in an *in vitro* reaction. The use of a site-specific lesion provides the advantages of controlling the sequence surrounding the lesion and producing only one type of damage as opposed to a spectrum. Studying the fate of a site-specific lesion *in vivo* yields a more accurate picture of what is truly happening inside the cell as it encounters a single lesion.

Many studies have examined the effect of a single lesion on both survival and mutagenesis. In a study by Burnouf *et al.* (67), a single d(ApG)/cis-diamminedichloroplatinum(II) (cisDDP[d(ApG)]) adduct was added to the *lacZ* gene in a 3.9-kb plasmid, which was transformed into chemically competent *E. coli* cells. Repair of the adduct led to restoration of an in-phase stop codon that yielded a white phenotype on X-gal plates; the induction of mutagenesis led to loss of the stop codon and the production of a blue phenotype. The colonies were analyzed by oligonucleotide hybridization and sequencing to determine the exact base change caused by the lesion. The researchers determined that induction of the SOS pathway was required for mutagenesis and increased the mutation frequency by more than two orders of magnitude. Hybridization and sequencing revealed that > 80% of the mutants exhibited
a single A → T transversion of the 5’ A in the cisDDP[d(ApG)] adduct. A second set of mutants contained an A → G transition of the same 5’ A residue. Two similar studies were performed by the Essigmann lab to examine the effect of a lesion on both cell survival and mutagenesis. The first study examined the effect of a single 7-hydro-8-oxoguanine lesion (68), while the second examined the effects of the vinyl chloride-induced DNA lesions EthenoA, 3,N4-ethenocytosine, and 4-amino-5-(imidazole-2-yl)imidazole (69).

Another study examined cellular events such as excision repair, recombinational repair, and translesion DNA synthesis in response to a single EthenoA adduct (70). A plasmid containing this lesion was transformed into several different strains deficient in the activity of interest, and the progeny plasmid was analyzed by oligonucleotide hybridization. The key to this experiment was the inclusion of several strand-specific mismatches. These mismatches would result in the production of different restriction sites after a process like replication or NER. These different restriction sites allowed the researchers to monitor the fate of the two strands independently. The cell strains in the experiments were all deficient in mutS to assure that these mismatches would not be repaired by MMR. In a strain that was also excision repair deficient (alkA: tag uvrA), replication of the lesion-containing strand was greatly decreased and yielded repair in
only 2% of the total plasmids detected. A variety of strains were used to examine the cellular events and the DNA sequence that results.

1.3 DNA lesions and fork blockage

Over the past 40 years, it has been observed that DNA damage can inhibit progression of the replication fork (16, 71). Typically, this damage is massive and generalized. It is often created by exposure to UV irradiation. It is generated throughout the DNA and is not site-specific. This experimental approach has helped uncover many important pathways such as SOS induction, damage bypass, direct restart, and others (see reviews (72-74), respectively). More recent studies have been performed both in vitro and in vivo and have looked at the effects of site-specific and general lesions.

1.3.1. In vitro studies

One group of studies examined the effects of general damage on the activities of E. coli DNA polymerases I, II, and III. The template DNA contained random lesions generated by UV irradiation. These lesions were found to inhibit all three polymerases. They were also able to block replication of a single-stranded DNA phage (71).

Other studies examined the replication of DNA templates containing site-specific lesions. One set of experiments examined the effect of a single, site-specific cis,syn-cyclobutane dimer or a pyrimidine (6-4) pyrimidone dimer, which are commonly produced by exposure to UV irradiation. Eukaryotic replication forks from HeLa cell
extracts were blocked only when the lesion was in the leading strand of the DNA, with blockage observed by two-dimensional agarose electrophoresis (75). A second set of in vitro experiments studied replication of a gapped substrate containing a site-specific abasic site in the template strand. In these studies, the abasic site lesion inhibited replication by E. coli DNA polymerases I, II, and core polymerase III (76-78). In contrast, DNA polymerase V, an error prone (translesional) polymerase, was able to bypass the abasic site lesion at a rate of approximately 55%. This rate is based on quantitation of the bypass product in relation to uninitiated and blocked products (77). The DNA polymerase III holoenzyme was also capable of some bypass, with a rate of approximately 30% (78).

1.3.2. In vivo studies

Several studies have examined the in vivo effect of a site-specific lesion on replication, but these studies are often limited to a genetic analysis of the progeny. The common strategy for these experiments involved the transformation or transduction of plasmid or phage DNA into cells and the analysis of successfully transformed or transduced colonies or plaques the next day. Many of these types of studies have been performed (see reviews (79, 80)). One such study examined whether propagation of a single site-specific thymine dimer on a single-stranded phage molecule required repair or translesion synthesis and bypass. DNA was transduced into E. coli uvrA- mutant cells,
and plaque formation was measured. As previously mentioned, the UvrA protein is involved in the NER pathway. Successful propagation depended on whether the dimer was cis-syn (< 1%) or trans-syn (14%). However, when the same mutant cells were pre-induced for expression of the SOS system, the efficiency increased to 30%, regardless of the type of lesion (81, 82).

In a study by Pagès and Fuchs (83), plasmid DNA containing a single lesion was constructed by covalently linking N-2-acetylaminofluorene (AAF) to the C-8 position of a guanine residue. This plasmid was then introduced into cells by electroporation, and the effect of the lesion on replication, either of the leading or lagging strand, was examined by Southern blotting of the replication products. The two strands could be distinguished by the presence of a local sequence heterology, which is formed by a +3-nucleotide bulge that introduces a new restriction site in one strand after replication. The lesion was found to delay replication of the strand in which it was located by approximately 50 minutes but did not affect replication of the opposite strand. Additionally, bypass of the lesion required induction of the SOS system and translesion synthesis (TLS). The error-free pathway required DNA polymerase V (umuD′:C), whereas PolIII (polB) mediated a mutagenic, -2 frameshift pathway. The MMR and NER pathways were inactivated in the cell strains by introducing mutations in mutS and uvrA, respectively. These mutations prevent the plasmid from being repaired by a
pathway that does not involve replication. This study most closely resembles the new system described here and sets a precedent for studying the fate of DNA molecules immediately after introduction.

1.4 Advantages of the new system

The new system described in this work will combine several of the characteristics of the many studies described above to create a system that can be used to examine the fate of a single, site-specific lesion in vivo immediately after introduction. This system uses two separate DNA substrates to simplify introduction of the site-specific lesion. The genome of bacteriophage λ is encoded in a 48.5-kb DNA molecule, which makes it a difficult substrate for extensive cloning and manipulation. However, bacteriophage λ is able to efficiently infect E. coli cells; a typical in vitro packaging reaction can yield titers of approximately $10^8$ - $10^9$ plaque forming units (PFU)/mL. In contrast, the addition of a site-specific lesion is much simpler in a plasmid and has been described previously (84, 85). However, plasmid transformation is often not as efficient as bacteriophage λ infection, and the electroporation of cells, a very efficient transformation method, can cause cells to behave as RecBC phenocopies (86), which will affect repair behavior. By integrating damaged plasmid into bacteriophage λ vector, damaged bacteriophage λ DNA can be produced. Additionally, use of the large bacteriophage λ DNA molecule may result in an environment for the lesion that is more similar to its environment in the
chromosome. All these characteristics make integrated bacteriophage λ an excellent candidate for the introduction of a site-specific lesion.

The high titers of bacteriophage λ should make it possible to infect enough cells with the damaged DNA to then harvest the DNA at specific time points and observe DNA repair in real-time. The repair can be monitored with a variety of tools, including some that were described in the previous section. By infecting cells with relatively large amounts of DNA, restriction digests and Southern blotting can be used to monitor DNA repair. Southern blotting has been used in previous studies (83, 87), and the increased efficiency of bacteriophage λ will introduce enough DNA into the cells that it can be analyzed by physical methods. Other techniques, including real-time polymerase chain reaction (RT-PCR), could also be used to monitor DNA repair. Replication fork behavior as it encounters the DNA lesion could be observed using two-dimensional (2D) gel electrophoresis (88) and Southern blotting. By combining the ease of plasmid manipulation and the efficiency of bacteriophage λ, repair and replication of site-specific DNA lesions could be studied in vivo minutes after introduction.

The combination of a plasmid and bacteriophage λ also provides a tool to carefully control the replication and the location of the DNA damage relative to the origin. By introducing the damaged DNA into cells that have been lysogenized with wild-type bacteriophage λ, replication from the phage origin can be prevented. By using
a plasmid with a conditional origin such as R6Kγ, replication from the plasmid origin can be prevented. This combination will prevent the damaged DNA molecule from replicating when it enters the cells, which simplifies analysis of the resulting DNA. Additionally, if one wants to observe the effects of replication, either origin can be used to replicate the damaged DNA molecule. Use of the attλ site can be used to direct the damage into the E. coli chromosome by integrating the entire damaged DNA molecule into the genome. Directing the DNA would provide researchers with a simple tool to insert a site-specific lesion into the E. coli chromosome, which may provide a more accurate picture of how repair would occur in vivo.

This system will provide researchers with a new tool to answer many questions and to examine in vivo DNA repair and replication on a time scale that, in many cases, was not previously possible. By using techniques such as restriction digestion and Southern blotting or real-time PCR, the repair of a site-specific DNA lesion can be examined immediately after its introduction. The repair can be studied in wild-type cells and mutants of various DNA repair proteins. These types of experiments will allow researchers to determine in vivo which proteins are involved in DNA repair and which proteins/pathways are the preferred method of repair for a specific lesion. This system will also allow for the observation of DNA repair under different cellular conditions. For example, the SOS pathway could be induced before introduction of the DNA, which
would provide information about how this pathway affects DNA repair; other conditional pathways could also be tested. Replication past the DNA lesion can be examined using 2D gel electrophoresis and Southern blotting. This type of experiment will reveal whether the DNA lesion can block the replication fork and, if it does, how long the blockage persists. Sequencing of the DNA after replication will provide information about what DNA base is introduced opposite of the lesion and whether replication is error-free or not. Again, the effect of different *E. coli* mutants and conditional pathways on replication could also be examined.
2. Cre-lox recombination

2.1 Introduction

The Cre-lox recombination system has been used in different organisms as an effective system to introduce or remove DNA, both in vivo and in vitro. The greatest advantage of the Cre-lox system is the fact that the loxP recognition sites are small (34 bp in length), and Cre will recombine any DNA substrate that contains these sites, with no requirements for accessory proteins or substrate topology (89, 90). Additionally, no high energy cofactors are required, but the omission of simple salts such as MgCl$_2$ and spermidine decreases recombination by approximately 50% (91).

The lox sites are 34 bp inverted repeats that contain a central 8-bp crossover region. One Cre monomer binds to each inverted repeat to catalyze strand incision and form a Holliday junction (Figure 5) (92). The products produced after Holliday junction

![Figure 5. Cre binding to the loxP site.](image)

Cre specifically recognizes and binds the recombinase binding elements (RBEs), which consist of inverted repeats around a central 8-bp crossover region (shown in bold). The 8-bp region consists of an asymmetric sequence that defines the site’s directionality. Cleavage sites are indicated by the red arrows. Adapted from (93).
resolution depend on the DNA substrates involved. Intramolecular reactions can lead to inversion of the intervening DNA if the \textit{lox} sites are oriented in opposite directions or excision of the intervening DNA if the \textit{lox} sites are oriented in the same direction. Intermolecular reactions lead to production of a new DNA molecule produced by a combination of the two substrates. These many possible products highlight the usefulness of the Cre-\textit{lox} system for cloning and gene integration in a variety of organisms, even plants, which provides an important research and development tool (94).

The simplicity of the Cre-\textit{lox} system creates one of the possible problems encountered with its use. Because the recombination reaction requires only the Cre protein and the \textit{lox} sites and recombination does not destroy the \textit{lox} sites, the reaction is easily reversible. While the most efficient recombination is achieved with \textit{loxP} sites, other mutant \textit{lox} sites are able to recombine. The advantage of using mutant sites (Figure 6) is that recombination between the resulting \textit{lox} sites is extremely inefficient, which pushes the reaction toward the production of integration products (95). Both wild-type and mutant \textit{lox} sites were used to attempt to integrate plasmid DNA into a bacteriophage \textit{\lambda} vector.
Figure 6. Irreversible recombination of mutant \textit{lox} sites.

Recombination of the \textit{lox66} and \textit{lox71} sites produces a normal \textit{loxP} site and a mutant \textit{lox72} site. These sites recombine poorly and favor the formation of integration products. Adapted from (95).

2.2 Materials and Methods

\textit{Materials}

The pGPS2.1 plasmid and all the restrictions enzymes were purchased from New England Biolabs (Ipswitch, MA). The EcoRI-digested \textit{λgt}11 vector was obtained from Agilent Technologies (Santa Clara, CA). All oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). The \textit{in vitro} MaxPlax Lambda Packaging Extracts were purchased from Epicentre (Madison, WI). Both the Gel Extraction Kit and the Ni-NTA beads were from Qiagen (Valencia, CA). The protein assay reagent was purchased from Bio-Rad (Hercules, CA)
Table 1. DNA constructs.

These DNA constructs, their production, and their uses are described in the chapters that follow.

<table>
<thead>
<tr>
<th>Name</th>
<th>Starting vector</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMPA</td>
<td>pGPS2.1</td>
<td>Nt.BbvCI and Nb.BbvCI nicking sites</td>
</tr>
<tr>
<td>pMPA1</td>
<td>pGPS2.1</td>
<td>loxP site, Nt.BbvCI and Nb.BbvCI nicking sites</td>
</tr>
<tr>
<td>pMPA1.1</td>
<td>pBR322</td>
<td>loxP site</td>
</tr>
<tr>
<td>pMPA2</td>
<td>pGPS2.1</td>
<td>lox71 site, Nt.BbvCI and Nb.BbvCI nicking sites</td>
</tr>
<tr>
<td>pMPA2.1</td>
<td>pBR322</td>
<td>lox71 site</td>
</tr>
<tr>
<td>pMPA3</td>
<td>pGPS2.1</td>
<td>loxP site, Nt.BstNBI nicking sites</td>
</tr>
<tr>
<td>pMPA4</td>
<td>pGPS2.1</td>
<td>attB site, Nt.BstNBI nicking sites</td>
</tr>
<tr>
<td>pMPA4.1</td>
<td>pBR322</td>
<td>attB site</td>
</tr>
<tr>
<td>Bacteriophage λ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>λgt11</td>
<td>-----</td>
<td>lacZ, cI857, single EcoRI site</td>
</tr>
<tr>
<td>λ-loxP</td>
<td>λgt11</td>
<td>loxP site</td>
</tr>
<tr>
<td>λ-lox66</td>
<td>λgt11</td>
<td>lox66 site</td>
</tr>
<tr>
<td>λ Kytos</td>
<td>λgt11</td>
<td>attP site, kan^R gene</td>
</tr>
</tbody>
</table>

Plasmid production

The pGPS2.1 plasmid was digested with BglIII and NotI and ligated to duplex DNA produced by annealing two oligonucleotides (insert L – 5’

GATCTGCTGAGGTACACGTGATAAACTTAAGACATATGGAATCTCGAGCTGAGGG

30
GGCCGCCCTCAGCTCGAGATTCCATATGTCTTAAGTTATCAGTACCTCAGCA

3’) to produce the restriction site insert, which will be used to introduce and monitor the DNA lesion. The presence of the insert was verified by digestion and sequencing, and this plasmid was renamed pMPA. To insert the \textit{lox} site, the pBR322 plasmid was digested with NheI, gel-purified using a Qiagen Gel Extraction Kit, and ligated by T4 DNA ligase to duplex DNA produced by annealing two oligonucleotides to produce the wild-type or mutant \textit{lox} site (\textit{loxP}: L \textit{CTAGCATAACTTCGTATAGCATACATTATACGAAGTTA} R \textit{3'}, R \textit{CTAGCATAACTTCGTATAGCATACATTATACGAAGTTA} 5’

\textit{CTAGCATAACTTCGTATAGCATACATTATACGAAGTTA} 3’, R \textit{CTAGCATAACTTCGTATAGCATACATTATACGAAGTTA} 5’

\textit{AATTCTACCGTTCGTATAGCATACATTATACGAAGTTA} 3’, R \textit{AATTCTACCGTTCGTATAGCATACATTATACGAAGTTA} 5’

\textit{AATTCTACCGTTCGTATAGCATACATTATACGAAGTTA} 3’; \textit{lox71}: L \textit{3’}

\textit{AATTCTACCGTTCGTTCGTATAGCATACATTATACGAAGTTA} 3’, R \textit{AATTCTACCGTTCGTTCGTATAGCATACATTATACGAAGTTA} 5’

\textit{AATTCTACCGTTCGTTCGTATAGCATACATTATACGAAGTTA} 3’); these constructs were named pMPA1.1 and pMPA2.1, respectively. The resulting plasmids were transformed, and several chloramphenicol resistant colonies were mini-prepped and screened for the \textit{loxP} or \textit{lox71} site by sequencing. The pMPA1.1, pMPA2.1, and pMPA plasmids were then digested with AatIII and SphI and separated by agarose gel electrophoresis. The pMPA1.1 or pMPA2.1 fragment containing the \textit{lox} site and the pMPA fragment containing the restriction site insert were gel-purified, ligated together, and transformed. The new plasmids were verified by digestion and sequencing and
renamed pMPA1 and pMPA2, respectively. More information about the generation of damaged plasmid will be presented in the next chapter.

*Bacteriophage λ cloning*

The *lox* sites were constructed by annealing two oligonucleotides to produce duplex DNA *(loxP: L – 5’ AATTCATAACTTCGTATAGCATACATTATACGAAGTTATG 3’, R – 5’ (Phos)AATTCATAACTTCGTATAATGTATGCTATACGAAGTTATG 3’; lox66: L – 5’ AATTCATAACTTCGTATAGCATACATTATACGAACGGTAG 3’, R – 5’ (Phos)AATTCTACCGTTCGTATAGCATACATTATACGAAGTTATG 3’), which was then ligated to the EcoRI-digested λgt11 vector. The ligated DNA was packaged with *in vitro* MaxPlax Lambda Packaging Extracts according to the manufacturer’s instructions and used to infect Y1088 *E. coli* cells. Plaques were struck onto plates containing 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal) and isopropyl-beta-D-thiogalactopyranoside (IPTG) to screen for colonies containing vector with the *loxP* insert. These colonies appeared light blue. The *lox* sites were verified by sequencing, and the *lox* constructs were named λ- *loxP* and λ- *lox66* based on the *lox* site they contain.

*Cre protein purification*

One liter of *E. coli* BL21 cells containing the pShe6 plasmid (96), which expresses the Cre protein, was grown to an OD<sub>600</sub> of 1.0 and induced with 1 mM IPTG for four
hours at 37°C. The cells were harvested by centrifugation at 4,000 x g. The cell pellet was resuspended in 40 mL of lysis buffer (50 mM sodium phosphate pH 8.0 and 500 mM NaCl) and sonicated 6 times for 15 seconds on Power Level 6 using a Branson Digital Sonifier (Branson; Danbury, CT). Cell debris was removed by centrifugation at 20,000 x g and 4°C for 30 minutes. The supernatant was then incubated with rotation with 4 mL of equilibrated Qiagen Ni-NTA beads for 1 hour at 4°C. The beads were pelleted by centrifugation at 1,000 x g and 4°C for 15 minutes, and the supernatant was removed. The beads were resuspended in lysis buffer and poured to form a column, which was washed with 100 mL of wash buffer (50 mM sodium phosphate pH 8.0, 500 mM NaCl, 10% glycerol, and 50 mM imidazole). The column was eluted (elution buffer: 50 mM sodium phosphate pH 8.0, 500 mM NaCl, 10% glycerol, and 400 mM imidazole), and the fractions containing protein were pooled and dialyzed overnight (dialysis buffer: 50 mM sodium phosphate pH 5.7, 50 mM NaCl, and 20% glycerol). The final protein concentration was determined with Bio-Rad Protein Assay Reagent (Hercules, CA) and bovine serum albumin standards.

**Cre-lox recombination**

Equimolar amounts of the lox-containing λgt11 vector (loxP or lox71) and lox-containing plasmid (pMPA1 or pMPA2) were incubated with a titration of Cre concentrations to determine the extent of recombination. Cre protein was diluted in 1x
Cre buffer (50 mM Tris-HCl pH 8.0, 33 mM NaCl, 10 mM MgCl₂, 1 mM dithiothreitol (DTT), and 2% PEG), and 10 µL of protein was added to 20 µL of the DNA mixture in 1x Cre buffer. The reactions were incubated at 37°C for 30 minutes and deactivated at 70°C for 30 minutes. The reactions were digested with MluI at 37°C for 1 hour and analyzed by gel electrophoresis.

2.3 Results

Production of the various constructs

The restriction site insert was added easily to the pGPS2.1 plasmid to create pMPA. This insert introduces four unique restriction sites, PmlI, AflII, NdeI, and XhoI. The DNA lesion will be placed at key locations in these sites in order to monitor damage. Flanking the restriction sites, it also contains two Nt.BbvCI sites on the leading strand (relative to the R6Kγ origin) and two Nb.BbvCI sites on the lagging strand (relative to the R6Kγ origin) (Figure 7).

![Figure 7. Restriction site insert.](image)

The restriction enzyme sites are shown with their names and sequences in the same color. The nicking sites are indicated by red arrows. The BglIII and NotI overhangs generated upon annealing will be complementary to the overhangs generated by digestion of the pGPS2.1 plasmid by the BglIII and NotI enzymes.
The next step in the original strategy involved annealing two oligonucleotides to create a _lox_ site flanked by _NheI_ sites. This duplex could then be ligated into _NheI_-cut pMPA. For reasons that could not be determined, this small fragment could not be directly ligated into the pMPA plasmid. However, as the pGPS2.1 vector shares a large section of sequence with pBR322, the _lox_ site was first inserted into pBR322. A larger fragment of pBR322 could then be ligated with a fragment of pMPA containing the restriction site insert. This step produced the final pMPA1 and pMPA2 plasmids.

Addition of the _lox_ sites to the _λgt11_ vector was also challenging. The vector is provided as the two vector arms, which are produced by EcoRI digestion and phosphatase treatment. The fragments of interest that are to be inserted must be flanked by EcoRI sites. In this case, these sites were produced by the annealing of two oligonucleotides. The EcoRI site of _λgt11_ is located within the _lacZ_ gene, which can be used for blue-white screening when X-gal and IPTG are included in the medium. Addition of the _lox_ oligonucleotide duplex in the EcoRI site should interrupt the _lacZ_ gene and result in the production of white colonies. However, these white colonies were found not to contain the _lox_ site but a mutation of the _lacZ_ gene. The screen was then expanded to include light blue colonies, which were determined to contain the desired _lox_ sites. This result could be due to the fact that the _lox_ site is a small insertion that does not interrupt the reading frame of the _lacZ_ gene; the blue-white screening may be more
effective with larger insertions as the vector is able to accept insertions up to 7.2 kb in size. Once the necessary cloning was completed, the Cre protein was purified and recombination reactions were attempted.

**Cre-lox recombination**

In the next chapter, insertion of DNA damage into the plasmid will be discussed. First, test recombination reactions were performed with undamaged pMPA1 and pMPA2 plasmids to determine the extent of recombination between the λ and plasmid substrates.

Activity of the purified Cre protein was first tested with a plasmid containing two loxP sites, and the protein was found to be able to catalyze intramolecular recombination (Figure 8), though the reaction did not consume all the linearized plasmid. It should be noted that the addition of too much Cre protein can lead to the inhibition of recombination (Figure 8, last lane of gel). Additionally, experiments were
Figure 8. Intramolecular Cre recombination.

Linearized plasmid is labeled as linear unreacted. The addition of increasing amounts of Cre recombinase led to the conversion of linearized plasmid to linear Cre product and circular Cre product, though the reaction was never efficient enough to consume all the linearized plasmid. Cre recombinase concentrations of 0.05 to 3 µM were tested. L indicates the DNA ladder. The black triangle indicates increasing concentrations of Cre recombinase.

performed with the pMPA1 plasmid and a second plasmid containing the proper lox site to ensure that intermolecular recombination was possible. The purified Cre protein was also able to catalyze this reaction to produce intermolecular recombination products (Figure 9, green arrows). The two recombination products are the result of multimer formation by the plasmid. Again, the addition of too much Cre protein appears to inhibit
the reaction, as the recombination bands indicated by the green arrows become less intense (Figure 9, last two lanes).

![Figure 9. Intermolecular Cre recombination.](image)

The recombined plasmid is labeled in green. The addition of increasing amounts of Cre recombinase led to the conversion of two plasmids to the recombination product, though the reaction was never efficient enough to consume all the plasmids. L indicates the DNA ladder. DNA only includes only the two DNA substrates. Mock rxn includes the two DNA substrates and 1x Cre reaction buffer only. The black triangle indicates increasing concentrations of Cre recombinase. Cre recombinase concentrations of 25 to 200 nM were tested.
Recombination of the \textit{lox}-containing \textit{\lambda}gt11 vector and the \textit{lox}-containing damaged plasmid results in the production of a \textit{\lambda} vector containing a single, site-specific lesion that can be used to infect \textit{E. coli} cells with high enough efficiency that the fate of the lesion can be followed in a time dependent manner.

Once the Cre protein was found to be active and to be able to perform intermolecular recombination, recombination reactions were performed with the pMPA1 and pMPA2 plasmids and the \textit{\lambda-loxP} and \textit{\lambda-lox66} vectors (Figure 10), respectively. Many attempts were made to push the recombination reaction toward the production of recombined DNA (product); they included an increase in the concentrations of the DNA substrates, the inclusion of molecular crowding reagents, such as spermidine and polyethylene glycol (PEG), and the testing of many different
Cre:DNA ratios. Additionally, a time course of the Cre recombinase reaction was also performed. However, none of these factors produced a sufficient amount of recombined \( \lambda gt11 \) vector. Figure 11 shows two sets of recombination reactions that achieved the highest observed efficiencies. Unfortunately, neither of these reactions, with the wild-type or mutant \( \text{lox} \) sites, produced enough recombined \( \lambda gt11 \) vector for additional experiments.
Figure 11. Maximum recombination efficiency with λgt11 vector and pMPA1 and pMPA2.

Left - The DNA substrates used in this experiment both contain loxP sites. Unless otherwise indicated, the DNA was digested with MluI in order to observe the shift caused by recombination of the plasmid. Lane 1 – ladder, Lane 2 – undigested pMPA1 plasmid, Lane 3 – undigested pMPA1 plasmid and λ-loxP vector, Lane 4 – 1x Cre buffer only, Lane 5 – 100 nM Cre, Lane 6 – 150 nM Cre, and Lane 7 – 200 nM Cre.

Right - The DNA substrates used in this experiment both contain mutant lox sites. Unless otherwise indicated, the DNA was digested with MluI in order to observe the shift caused by recombination of the plasmid. Lane 1 – ladder, Lane 2 – undigested pMPA2 plasmid, Lane 3 – undigested pMPA2 plasmid and λ-lox66 vector, Lane 4 – 1x Cre buffer only, Lane 5 – 100 nM Cre, Lane 6 – 150 nM Cre, and Lane 7 – 200 nM Cre.

The red arrows indicate recombined DNA.
2.4 Discussion

Despite significant efforts to improve the recombination efficiency, we finally judged that Cre-\textit{lox} recombination could not be used to produce sufficient damage-containing bacteriophage \textit{\lambda} DNA. Even with improvements in the DNA concentrations, the reaction never produced enough recombined DNA for efficient packaging and infection (Figure 11, \textit{loxP + loxP} gel, red arrow). Based on the different results seen with the control plasmids and the desired substrates, the problem appears to be related to the simplicity of the Cre-\textit{lox} system. Recombination of \textit{loxP} sites does not destroy these sites; it simply creates a new DNA substrate that contains two \textit{loxP} sites. Simple logic implies that at certain DNA concentrations, those two \textit{loxP} sites in one molecule will be more likely to react with each other than two \textit{loxP} sites in two molecules. In a reversible reaction, the production of two products from one substrate will be favored over the production of one product from two substrates, which leads to the presence of more unrecombined DNA. While this characteristic is helpful in assays that use Cre-\textit{lox} recombination to eliminate DNA sequences (see review (94)), it becomes a problem when trying to recombine two DNA substrates.

This reversible recombination was observed previously, and a study was performed to isolate \textit{lox} sites that were able to resist excisional recombination after the initial reaction (95). Based on these previous observations, the problem with the wild-
type \( \text{loxP} \) sites was anticipated, and the DNA substrates were also produced with mutant \( \text{lox} \) sites. However, the mutant \( \text{lox} \) sites produced even less recombined DNA than the wild-type \( \text{loxP} \) sites (Figure 11, \( \text{lox66 + lox71} \) gel, red arrow). As attempts to improve the recombination were proving futile, the idea of using a new system was proposed. This new recombination system will be described in Chapter 4.
3. Addition of the abasic site lesion

An abasic site was initially chosen as the lesion with which to develop and test this system. This decision was influenced by several factors. The first factor was the ease with which the lesion could be obtained. A stable abasic site analog, 1',2'-dideoxyribose or dSpacer, has been used in previous studies and can be synthesized in an oligonucleotide. The second factor was the fact that the abasic site is a well-studied lesion. Much is known about the requirements for abasic site generation and repair. These previous studies can serve as a point of comparison for the results obtained with the new in vivo system. The final factor was the size of the abasic site lesion. The damaged DNA will still need to be packaged into bacteriophage λ heads. A large, bulky lesion could distort the DNA helix and affect packaging efficiency. A small lesion will serve as a proof of principle for the system, and once the system has been established, larger lesions can be tested.

3.1 Introduction

Abasic sites result from the depurination and depyrimidization of DNA bases and are repaired by the BER pathway (Figure 1). These depurination and depyrimidization reactions can remove incorrectly paired and damaged bases and can be spontaneous or catalyzed by enzymes known as DNA N-glycosylases. A variety of N-glycosylases exist in prokaryotes to remove different types of DNA damage, and several
examples were described in Chapter 1. The common thread between all these $N$-glycosylases is the production of an abasic site that serves as a common intermediate for downstream enzymes of the BER pathway.

After production of the abasic site, a DNA endonuclease cleaves the DNA just 5’ to the lesion; the main endonucleases in *E. coli* are Exonuclease III (XthA) and Endonuclease IV (Nfo). After DNA incision, a short exonuclease reaction occurs, followed by DNA polymerization and ligation (13-15). The short patch subpathway of BER removes only one nucleotide, whereas long patch BER removes multiple nucleotides (see (17) for review). Abasic sites must be removed from the DNA as they are noncoding and strongly block the replication fork (78). If the SOS system is upregulated in cells, abasic sites can be bypassed by DNA polymerase V (*umuD’sC*). However, this error prone polymerase often inserts an adenine nucleotide opposite the abasic site, which can lead to mutation (77).

As DNA damage can be problematic if left unrepaired, the BER pathway in *E. coli* cells contains redundancies. These redundancies allow BER to occur even if one protein in the pathway is disabled, as there is a second protein that can perform the same function as first disabled protein. For example, multiple DNA $N$-glycosylases can remove the same type of lesion. Endonuclease III is encoded by the *nth* gene in *E. coli* and serves to remove pyrimidine residues damaged by ring saturation, ring
fragmentation, or ring contraction (97-99). Though it can remove all these lesions, *E. coli* mutants lacking Nth are not hypersensitive to γ rays or hydrogen peroxide and demonstrate only a slight increase in their spontaneous mutation rate (61, 100). The moderate phenotype is due to the presence of Endonuclease VIII, which is encoded by the *nei* gene and can substitute for the activity of Nth. Only when *E. coli* are defective in both the *nth* and *nei* genes are the cell hypersensitive to ionizing radiation and hydrogen peroxide (100, 101). Redundancy also exists at the endonuclease step of the BER pathway. XthA is responsible for the majority of AP endonuclease activity in *E. coli*, as Nfo accounts for only approximately 10% of AP endonuclease activity in normal cells and was only identified by its activity in *xthA* mutant cells (102, 103). However, Nfo can substitute for XthA if XthA is not functioning (13, 14), and there is evidence that Nfo may recognize lesions that XthA does not (13).

The existence of these multiple enzymes make the abasic site an interesting candidate for the testing of this system. Different gene knockout strains can be constructed, and the fate of the abasic site lesion can be determined upon introduction into these knockout strains. These studies will yield actual kinetic information about the rate of repair by different proteins and pathways, which will help determine the most efficient pathway for the repair of abasic sites. The results can then be compared with
the data obtained from previous studies to determine if the different techniques yield the same results.

### 3.2 Materials and Methods

#### Materials

All the restrictions enzymes were purchased from New England Biolabs (Ipswich, MA). All oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). The in vitro MaxPlax Lambda Packaging Extracts were purchased from Epicentre (Madison, WI). Both the QIAquick PCR Purification Kit and Gel Extraction Kit were from Qiagen (Valencia, CA). The Site-Directed Mutagenesis Kit was purchased from Stratagene (La Jolla, CA). The Nytran transfer membrane was purchased from Whatman, Inc. (Piscataway, NJ). The Random-Primed DNA Labeling kit was obtained from Roche (Pleasanton, CA).

#### Plasmid production

A new restriction site insert was added to the pMPA1 plasmid. pMPA1 plasmid was digested with BglII and NotI and ligated to duplex DNA produced by annealing two oligonucleotides (insert L – 5’ GATCTGAGTCAATGACACGTGAGCCTTAAGATCTCGAGTCCTACGC 3’ insert R – 5’ GGCCGCGTAGGACTCGAGATCTTAAGGCTCACGAGTCATTGCAGACTCA 3’). An unwanted MlyI site in the vector portion of the plasmid was then destroyed by site-
directed mutagenesis using the primers 5’ GGGCTGCTTCTTAATGCAGTCCGCGCATAAGGGAGAGCGTCG 3’ and 5’ CGACGCTCTCCCTTATGCGCGGACTGCATTAGGAAGCAGCCC 3’. The final construct was verified by sequencing and named pMPA3.

Production of damaged DNA

The dSpacer lesion was used to provide a chemically stable abasic site lesion analog as a normal abasic site can undergo a spontaneous ring opening. pMPA3 was used for the preparation of dSpacer-containing and wild-type DNA by taking advantage of a previously described strategy (84, 85). In a typical preparation, 90 µg (34 pmol) of plasmid were incubated with 450 U of Nt.BstNBI at 55°C for two hours in 50 mM Tris-Cl (pH 7.9), 100 mM NaCl, 10 mM MgCl₂, and 1 mM dithiothreitol (DTT). The reaction was terminated by the addition of a 100-fold molar excess of the sequestering oligonucleotide 5’ GTAGGACTCGAGATCTTAAGGCTCACGTGTC 3’ and heating at 82°C for 20 minutes. This step inactivated the enzyme and melted out the 31-mer between the Nt.BstNBI nicking sites, which was unable to rehybridize to the gapped plasmid due to annealing to excess sequestering oligonucleotide. The gapped plasmid was purified using a QIAquick PCR Purification Kit according to manufacturer’s instructions. A 10-fold molar excess of synthetic oligonucleotide (5’ (Phos)GACACGTGAGCCCTTAAGATCTCGAGTCCTAC 3’, wild-type; 5’
(Phos)GACACGTGAGCCXTAAGATCTCGAGTCCTAC 3’, X represents the dSpacer lesion) was added to the gapped plasmid. The reaction was heated at 82°C for 10 minutes and slowly cooled to room temperature in a water bath. The reactions were then supplemented with 1x T4 DNA ligase buffer, 1 mM ATP, and T4 DNA ligase (40 U/µL) and incubated at room temperature for 1.5 hours (84). The reactions were separated on a 0.8% agarose gel containing ethidium bromide, and the closed circular DNA was excised and purified with the Gel Extraction Kit. The presence of the lesion was determined by restriction digests with SphI and AflIII.

Mock packaging reactions

To assure that exposure to the in vitro packaging extracts would not affect the abasic site lesion, mock packaging reactions were performed with unrecombined plasmid. A maximum volume of 10 µL of each plasmid was added to 25 µL of MaxPlax Lambda Packaging Extracts, and the reaction was incubated at 30°C for 1.5 hours. After packaging, the DNA was purified by successive extractions with phenol, phenol/chloroform, and chloroform and isopropanol precipitation.

DNA analysis by Southern blotting

The presence of the lesion was determined by restriction enzyme digestion and Southern blotting both before and after packaging. The extracted DNA was digested with SphI alone or SphI and AflIII. The digests were separated on a 0.8% agarose gel, and
the gel was transferred to a 0.45 µm Nytran transfer membrane by the downward sponge method (104). After transfer, the DNA was crosslinked to the membrane with a 120 mJ/cm² UV exposure. Radiolabeled probe was generated from singly-cut pMPA3 plasmid DNA using the Random-Primed DNA Labeling kit. All blots were visualized using a Storm 860 Phosphorimager (Molecular Dynamics; Sunnyvale, CA).

3.3 Results

Production and verification of damaged DNA

The method to produce the wild-type or damaged plasmid was based on a previous technique used to create gapped DNA substrates that could be further manipulated (84, 85). The nicking enzymes Nt.BbvCI and Nb.BbvCI were originally used to create the gapped DNA. However, it was determined that they could not efficiently produce gapped DNA (data not shown). A new enzyme, Nt.BstNBI was chosen and was able to create two nicks, separated by 31 bases, in the leading strand of pMPA3 (Figure 12).

**Figure 12. New restriction site insert.**
The restriction enzyme sites are shown with their names and sequences in the same color. The nicking sites are indicated by red arrows. Due to the nature of these nicking sites, only one strand will be nicked by Nt.BstNBI.

Gapped DNA was then generated by melting and sequestering the small, 31-bp oligonucleotide; the efficiency of the procedure was determined by ligation. When the resulting gapped DNA was incubated without additional oligonucleotide, only an extremely small amount of closed circular DNA was produced, indicating that the majority of the DNA was indeed gapped. Ligation of the wild-type or damage-containing oligonucleotide to the gapped DNA led to conversion to closed circular DNA.

This system was designed to study a variety of lesions. Ideally, the system could be used with any lesion that can be created in an oligonucleotide, either by synthesis or postproduction modifications. To develop and test the system, small lesions were first used to determine whether the idea was feasible. The first lesion selected was the dSpacer lesion (Figure 13A). This 1’2’-dideoxyribose base allows for the incorporation of a stable abasic site into the plasmid. The lesion was incorporated into a synthetic oligonucleotide at a location that would disrupt the AflII restriction site (Figure 13B). The isolated, closed circular undamaged and dSpacer-containing plasmids were digested with SphI alone or SphI and AflII. The wild-type plasmid was cut once by the SphI digest and cut twice with the SphI/AflII digest (Figure 13C, lanes 2+3), indicating
that the AflII site was intact. The dSpacer-containing plasmid was cut once by both the SphI and SphI/AflII digests (Figure 13C, lanes 5+6), indicating that AflII digestion was prevented by the dSpacer.

Figure 13. The dSpacer lesion and its incorporation.

A – Schematic of the dSpacer lesion. B – Sequence of the restriction site insert of pMPA3. The base replaced with the abasic site lesion is indicated by a red circle. C –
Digests of the wild-type and damaged plasmid to determine whether the dSpacer lesion interrupts the AflII site. Lanes 1-3 contain wild-type DNA and lanes 4-6 contain abasic site-containing DNA. Please see the text for a full description of the digests.

Determining lesion stability in packaging extracts

The stability of the dSpacer lesion was determined to assure that components of the packaging extracts would not repair the lesion. Repair of the lesion in the packaging extracts would compromise the system, preventing (or reducing) the subsequent introduction of damaged DNA into cells.

To determine lesion stability, wild-type or lesion-containing plasmid was incubated in a standard packaging reaction. Though the plasmid is too small to be packaged, this experiment will still determine whether the extracts are repair proficient or deficient. After incubation of the plasmid, the DNA was isolated by standard phenol/chloroform extraction and isopropanol precipitation and digested with either SphI alone (Figure 14; odd lanes) or SphI and AflII (Figure 14: even lanes). After incubation with the packaging extracts, the dSpacer-containing plasmid was no longer resistant to AflII digestion as seen with the original dSpacer-containing substrate (compare Figure 13C, lane 6 to Figure 14, lane 4). After packaging, the dSpacer-containing plasmid demonstrated the same digestion pattern as wild-type DNA (Figure 14, lanes 1+2 versus lanes 3+4), indicating that the dSpacer lesion was restored to a T residue. It can be concluded that the dSpacer lesion is efficiently repaired in the
packaging extracts. This result was the same whether homemade or commercial extracts were used.

<table>
<thead>
<tr>
<th>Lane</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>WT</td>
<td>abasic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SphI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>AflII</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 14. Repair of the dSpacer lesion in packaging extracts.

The wild-type and abasic site-containing plasmid DNA was incubated with *in vitro* packaging extracts. The DNA was purified and digested to determine whether the dSpacer lesion remained intact after incubation with packaging extracts.
3.4 Discussion

The dSpacer lesion’s lack of stability could be explained by the fact that it is repaired by both the BER and NER pathways, which are constitutively expressed (see review (105)). As packaging extracts are produced by lysing cells that overexpress the bacteriophage λ packaging proteins, they may also contain repair proteins. These proteins are then able to repair the lesion during the packaging reaction.

It is possible that this problem could be overcome by the use of homemade packaging extracts. These extracts can be produced by transforming cells with the pCM101 plasmid (terminase protein) or the pCM230 plasmid (heads and tail proteins) (106). To produce repair-deficient extracts, the plasmids would be transformed into cells lacking the protein(s) to repair the lesion of interest. In the case of the dSpacer lesion, the plasmids could be transformed into cells lacking the XthA and UvrA proteins, and the packaging extracts could then be produced. These extracts would lack the majority of the AP endonuclease activity required for BER and the damage sensing activity required for NER. While some mutants may face defects in growth, it is probable that the correct mutants could be constructed to produce packaging extracts that will not repair the lesion of choice. It is also possible that commercial extracts could be immunodepleted of the protein of interest. If a lesion of interest is found to be repaired in the packaging extracts, several strategies are available to eliminate this activity. In the interest of time,
these strategies were not attempted. Instead, a new lesion was selected for development of the system.
4. Construction of the working system

Based on the results achieved with the Cre-\textit{lox} recombination system and the abasic site lesion, the choice was made to change both these components. A new recombination system using the φC31 integrase and \textit{att} sites was selected for its unidirectional recombination. An EthenoA lesion was chosen because it had been shown to be repaired by a system that was induced rather than constitutively expressed, which was hypothesized to make the lesion more stable in \textit{in vitro} packaging extracts. This chapter will describe the φC31 integrase system and demonstrate its efficiency. It will also include a brief introduction about EthenoA and its repair and will illustrate the lesion’s stability.

4.1 Introduction

The genome of the Streptomyces phage φC31 is known to integrate into the host chromosome via the φC31 integrase, a serine recombinase. The mechanism of this family of recombinases appears to be similar to that used by resolvases/invertases (107-111). Cleavage of the DNA occurs at an extremely small 2-bp crossover sequence (TT; see Figure 15, orange boxes) and forms a covalent, transient phosphoserine bond (108, 111). Strand exchange then likely occurs by rotation of two recombinase subunits, which are bound to half sites of the DNA, relative to the other two subunits and reforming of the bond (see review (112)).
The φC31 integrase catalyzes the recombination of two att sites, attB and attP, to yield two new att sites, attL and attR (Figure 15). Unlike the Cre recombination reaction, which produces the same sites after recombination making the reaction reversible, the φC31 integrase will not catalyze the integration of the attL and attR sites, making the reaction irreversible (113). Though the integrase will not catalyze the reaction, it will still bind the attL and attR sites (114). Additional research found that the attL/attR reaction is catalyzed by the serine integrase Int and gp3, which serves as a recombination directionality factor (RDF) (115). The lack of reversibility, even with wild-type att sites, made this reaction an attractive one in the building of our system.

![Figure 15. φC31 integration reaction.](image)

The integration reaction of attB (blue) and attP (green) yields attL and attR (combination of blue and green), which are not acted upon by the φC31 integrase. This characteristic of the system makes it unidirectional. The 2-bp crossover site is indicated by an orange box. Adapted from (114, 116).
The EthenoA lesion is caused by the modification of an adenine base after exposure to vinyl chloride and other vinyl compounds (117) or lipid peroxidation in eukaryotes (118). DNA damaging agents can often cause a wide variety of lesions; therefore, multiple pathways and enzymes are necessary to repair the damage. The first pathway is BER, where multiple glycosylases can remove alkylated bases. Previous studies have indicated that EthenoA can be removed by AlkA glycosylase (119) or Mug glycosylase (120) to form the abasic site in the BER pathway. The second pathway involves the *E. coli* AlkB protein, an iron- and 2-oxoglutarate-dependent dioxygenase, which is able to directly repair the EthenoA lesion by opening the ring formed by the addition of an ethyl group. This opening results in the release of a glyoxal molecule and the restoration of the correct DNA base (121) without its removal. Both AlkA and AlkB are inducible enzymes controlled by the adaptive response in *E. coli*, which also regulates the *ada* and *aid* genes (see review (122)). Upon exposure to alkylating agents, the Ada protein acts as both a repair and a regulatory protein. The protein transfers a methyl group from certain methylated bases to itself. This act converts Ada to a transcription activator that induces expression of the other genes of the regulon; in some cases, removal of the methyl group also repairs the lesion. More information on the repair of this lesion will be presented in Chapter 5.
4.2 Materials and Methods

Materials

All the restrictions enzymes were purchased from New England Biolabs (Ipswitch, MA). Most oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). The oligonucleotide containing the EthenoA lesion was purchased from Sigma-Aldrich (Saint Louis, MO). The in vitro MaxPlax Lambda Packaging Extracts were purchased from Epicentre (Madison, WI). Both the QIAquick PCR Purification Kit and Gel Extraction Kit were from Qiagen (Valencia, CA). The Nytran transfer membrane was purchased from Whatman, Inc. (Piscataway, NJ). The Random-Primed DNA Labeling kit was obtained from Roche (Pleasanton, CA). The λ Kytos vector was a generous gift from Dr. Jody Plank (123).

Plasmid production

An attB site was first added to pBR322 plasmid by digesting the plasmid with NheI, gel-purifying the cut plasmid, and ligating it to a duplex oligonucleotide produced by annealing the oligonucleotides 5’ CTAGCGCGGTGCCAGGGCTGCCCTTTGGGCTCCCGGGCGCGTACTG 3’ and 5’ CTAGCAGTACGCCCGGGAGCCCAAGGCGACGCCTGCCCTGCGACCGACCACCACGCG 3’. This plasmid was named pMPA4.1. The pMPA4.1 and pMPA3 plasmids were then
digested with AatII and SphI and separated by agarose gel electrophoresis. The pMPA4.1 fragment containing the \textit{att} site and the pMPA3 fragment containing the restriction site insert were gel-purified, ligated together, and transformed. The presence and orientation of the \textit{attB} site was verified by sequencing. This construct was named pMPA4 and was used in integration reactions and in the production of damaged DNA.

\textit{Preparation of damaged DNA}

pMPA4 was used for the preparation of wild-type DNA and EthenoA-containing DNA by taking advantage of a previously described strategy (84, 85). In a typical preparation of 400 µL, 90 µg (34 pmol) of plasmid were incubated with 450 U of Nt.BstNBI at 55°C for two hours in 50 mM Tris-Cl (pH 7.9), 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT. The reaction was terminated by the addition of a 100-fold molar excess of the sequestering oligonucleotide 5' GTAGGACTCGAGATCTTAAGGCTCACGTGTC 3' and heating at 82°C for 20 minutes. This step inactivated the enzyme and melted out the 31-mer between the Nt.BstNBI nicking sites, which was unable to rehybridize to the gapped plasmid due to annealing to the excess sequestering oligonucleotide. The gapped plasmid was purified using a QIAquick PCR Purification Kit according to manufacturer’s instructions. A 10-fold molar excess of synthetic oligonucleotide (5’ GACACGTGAGCCCTAAGATCTCGAGTCCCTAC 3’, wild-type; 5’ ...
GACACGTGAGCCTTAAGATCTCGXGTCCTAC 3', EthenoA is represented as X) was added to the gapped plasmid. The reaction was heated at 82°C for 10 minutes and slowly cooled to room temperature in a water bath. The reactions were then supplemented with 1x T4 DNA ligase buffer, 1 mM ATP, and T4 DNA ligase (40 U/µL) and incubated at room temperature for 1.5 hours (84). The reactions were separated on 0.8% agarose gels containing ethidium bromide, and the closed circular DNA was excised and purified with a Gel Extraction Kit. The presence of the lesion was determined by restriction digests with EcoNI and XhoI.

Purification of φC31 integrase

The pHPhiC31Int plasmid, which expresses the φC31 integrase, was also created by Dr. Jody Plank by modifying the previously constructed pH62 plasmid (113). A 1-L culture of BL21(DE3) pLysS cells containing the pHPhiC31Int plasmid was grown at 37°C until an OD₆₀₀ of 0.6 was reached. The temperature was then shifted to 18°C, and the culture was grown to an OD₆₀₀ of 0.8. Protein production was induced with 0.1 mM IPTG and allowed to continue overnight at 18°C. The cells were then harvested by centrifugation at 3,500 x g for 20 minutes at 4°C and resuspended in 25 mL of lysis buffer (150 mM HEPES pH 7.0, 100 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol (BME), and 1 mM phenylmethylsulfonyl fluoride (PMSF; Fluka)). Lysozyme was added to a final concentration of 1 mg/mL. The cell suspension was frozen with liquid nitrogen,
incubated at -80°C for 15 minutes, and thawed on ice. The cells were sonicated 5 times on ice using a Branson Digital Sonifier (Branson; Danbury, CT) (5 seconds at 0.5-second intervals); the cells were allowed to cool on ice for 30 seconds between the 5 rounds of sonication. Sonication was monitored with Bio-Rad Protein Assay reagent (Bio-Rad; Hercules, CA). One volume of 2 M NaCl was added to the cells and mixed. A second volume of 1 M NaCl and 18% PEG-8000 was added and mixed; the mixture was incubated on ice for 30 minutes to allow the DNA to precipitate. The mixture was centrifuged at 50,000 x g for 30 minutes at 4°C. The supernatant was added to NiNTA beads equilibrated with lysis buffer and allowed to incubate with rotation at 4°C for 1 hour. The beads were spun at 1,000 x g for 15 minutes at 4°C, and the supernatant was decanted. The beads were resuspended in 10 mL of wash buffer I (50 mM HEPES pH 7.0, 1 M NaCl, 10% glycerol, 5 mM BME, and 1 mM PMSF) and poured into a column. The beads were allowed to settle, and the column was washed with an additional 90 mL of wash buffer I. The column was then washed with 100 mL of wash buffer II (50 mM HEPES pH 7.0, 100 mM NaCl, 5 mM imidazole, 10% glycerol, 5 mM BME, and 1 mM PMSF). The column was eluted in 1-mL fractions with 15 mL of elution buffer (50 mM HEPES pH 7.0, 100 mM NaCl, 400 mM imidazole, 10% glycerol, 5 mM BME, and 1 mM PMSF). The fractions were tested for the presence of protein using 1x Bradford reagent; those found to contain protein were pooled. The protein was mixed with an equal
volume of ice-cold glycerol, aliquoted, flash-frozen with liquid nitrogen, and stored at -80°C. The protein concentration was determined using Bio-Rad Protein Assay reagent and bovine serum albumin as a standard.

φC31 integrase reactions

Multiple 30-µL integrase reactions containing equimolar amounts of pMPA4 and λ Kytos vector in integrase buffer (10 mM Tris-Cl pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 100 mM NaCl, 4.5% glycerol, 5 mM DTT, 0.5 mg/mL bovine serum albumin (BSA), and 5 mM spermidine) were assembled. Ten microliters of 240 µM φC31 integrase were added to each reaction. The reactions were incubated at 30°C for 1 hour; the integrase was deactivated by incubating at 65°C for 20 minutes. The DNA was purified by successive extractions with phenol, phenol/chloroform, and chloroform and isopropanol precipitation. The recovered DNA was digested with EcoRI and run on a 0.8% agarose gel to determine the extent of recombination.

Packaging reactions

To assure that exposure to the in vitro packaging extracts would not affect the EthenoA, packaging reactions were performed with both unrecombined plasmid and recombined λ Kytos DNA. A maximum volume of 10 µL of each plasmid or recombined λ Kytos DNA was added to 25 µL of MaxPlax Lambda Packaging Extracts and incubated
at 30°C for 1.5 hours. After packaging, the DNA was purified by successive extractions with phenol, phenol/chloroform, and chloroform and isopropanol precipitation.

**DNA analysis by Southern blotting**

The presence of the lesion was determined by restriction enzyme digestion and Southern blotting both before and after packaging. The extracted plasmid DNA was digested with EcoNI alone or EcoNI and XhoI; the extracted λ Kytos DNA was digested with EcoRI alone or EcoRI and XhoI. The digests were separated on a 0.8% agarose gel, and the gel was transferred to a 0.45 µm Nytran transfer membrane by the downward sponge method \(^{(104)}\). After transfer, the DNA was crosslinked to the membrane with a 120 mJ/cm² UV exposure. Radiolabeled probe was generated from singly-cut pMPA4 plasmid DNA using the Random-Primed DNA Labeling kit. All blots were visualized using a Storm 860 Phosphorimager (Molecular Dynamics; Sunnyvale, CA).

**4.3 Results**

**Rationale and General Strategy**

Previous studies have transformed cells with plasmid DNA containing a single site-specific DNA lesion to determine the lesion’s fate *in vivo* (see reviews \(^{(79, 80)}\)). The underlying limitation of these studies is that the process of transformation is not very efficient \(^{(124)}\), and therefore relatively few cells are transformed. The results are generally limited to observation of the progeny or the end point of repair. Another
problem with the use of transformation is that it can alter cell behavior. Previous studies indicate that electroporated cells can behave as RecBC phenocopies (86), which can complicate the analysis. The process to produce competent cells could also be considered a mistreatment of the cells. These problems have led us to look for a new system to study DNA repair. The goal of this new system is to use bacteriophage λ Kytos as a tool to infect normal cells with much larger amounts of damaged DNA. By infecting more cells with the damaged DNA, cells can be harvested at different time points to observe DNA repair as it occurs.

This system uses two separate DNA substrates to simplify the introduction of a site-specific lesion. Bacteriophage λ is a 48.5-kb DNA molecule, which makes it a difficult substrate for extensive cloning and manipulation. However, it is able to efficiently infect E. coli cells; a typical in vitro packaging reaction can yield titers of approximately 1 x 10^9 PFU/mL. In contrast, the addition of a site-specific lesion is much simpler in a plasmid and has been done before (84, 85). The problem arises when attempting to introduce the plasmid into E. coli cells, as described above. By creating the DNA damage in a small plasmid, recombining the plasmid and bacteriophage λ vector, and packaging the DNA, we are able to exploit the advantages of each DNA molecule and avoid their pitfalls.
A brief outline of the process is outlined in Figure 16. Construction of the damaged DNA begins with digestion of pMPA4 (i) by the Nt.BstNBI nicking enzyme to generate gapped DNA (ii). The gapped DNA is then incubated with an oligonucleotide containing the site-specific damage and ligase to generate closed circular, damage-containing plasmid DNA (iii). The damaged plasmid is then recombined with \( \lambda \) Kytos vector (iv) to generate damaged \( \lambda \) Kytos DNA (v). That DNA is then packaged with \textit{in vitro} packaging extracts to generate phage particles (vi), which are used to infect \textit{E. coli} cells (vii). The fate of the DNA after introduction into cells can then be monitored.

![Diagram](image)

**Figure 16. Creation of and infection with damaged DNA.**

The steps to create each DNA species are described in the text above. Damage can be added to the smaller, easier to manipulate plasmid. The damaged plasmid is then recombined with the bacteriophage \( \lambda \) vector to produce the larger, damaged \( \lambda \) molecule for infection.
Production and verification of damaged DNA

The method to produce the wild-type or damaged plasmid was based on a previous technique used to study mismatch repair (84, 85). The enzyme Nt.BstNBI was used to create two nicks, separated by 31 bases, in the one strand of pMPA4. Based on the direction of the R6Kγ origin, this strand was designated as the leading strand. Gapped DNA (Figure 14, lane G) was then generated by melting and sequestering the small, 31-bp oligonucleotide. The efficiency of the procedure was determined by ligation: when the resulting gapped DNA was incubated without additional oligonucleotide, only an extremely small amount of closed circular DNA was produced, indicating that the majority of the DNA was indeed gapped (Figure 17, Lane L). Ligation of the wild-type or damaged oligonucleotide to the gapped DNA led to conversion to closed circular DNA (Figure 17, lanes WT and EthenoA, iii).
Figure 17. Production of wild-type and damaged DNA.

The starting substrate, various intermediates, and final products of the DNA production process are shown. The final closed circular product is indicated with an arrow and the numeral iii. Uncut plasmid, lane U; gapped DNA, lane G; gapped + T4 DNA ligase, lane L; wild-type ligation, lane WT; EthenoA ligation, lane EthenoA.

This system was designed to study a variety of lesions. Ideally, the system could be used with any lesion that can be created in an oligonucleotide, either by synthesis or postproduction modifications. To develop and test the system, small lesions were first used to determine whether the idea was feasible. As described in Chapter 3, the first lesion used was an abasic site mimic, the dSpacer lesion. This lesion was not stable in the \textit{in vitro} packaging extracts. The second lesion selected was EthenoA (Figure 18A). This
lesion was incorporated into a synthetic oligonucleotide in the place of an adenine residue to disrupt the XhoI restriction site (Figure 18B). The isolated, closed circular wild-type and EthenoA-containing plasmids were digested with EcoNI alone or EcoNI and XhoI. The wild-type plasmid was cut once by the EcoNI digest and cut twice with the EcoNI/XhoI digest, indicating that the XhoI site was intact (Figure 18C, compare lanes 2 and 3). The EthenoA plasmid was cut once by both the EcoNI and EcoNI/XhoI digests, indicating that the XhoI site was interrupted by the EthenoA lesion (Figure 18C, compare lanes 5 and 6).
Figure 18. The EthenoA lesion and its incorporation.

A – Structure of the EthenoA lesion. B – Sequence of the restriction site insert of pMPA4. The base replaced with the EthenoA lesion is indicated by a red circle. C – Digests of the wild-type and damaged plasmid to determine whether the EthenoA lesion interrupts the XhoI site. Lanes 1-3 contain wild-type DNA and lanes 4-6 contain EthenoA-containing DNA. Please see text for a full description of the digests.

**Determining lesion stability in packaging extracts**

The stability of the EthenoA lesion was determined to test whether components of the packaging extracts would repair the lesion. Repair of the lesion in the packaging
extracts would compromise the system, preventing (or reducing) the subsequent introduction of damaged DNA into cells.

To determine lesion stability, wild-type or lesion-containing plasmid was incubated in a standard packaging reaction. Even though the plasmid is too small to be packaged, this experiment will still determine whether the extracts are repair proficient or deficient. After incubation of the plasmid, the DNA was isolated using standard phenol/chloroform extraction and isopropanol precipitation and digested with EcoNI alone (odd lanes) or EcoNI and XhoI (even lanes). After incubation with the packaging extracts, the EthenoA-containing plasmid was still resistant to XhoI digestion as seen with the original EthenoA-containing substrate (compare Figure 18C, lane 6 to Figure 19, lane 4). It can be concluded that the EthenoA lesion is not efficiently repaired in the packaging extracts. This result was the same whether homemade or commercial extracts were used.
Figure 19. Stability of the EthenoA lesion in packaging extracts.

The wild-type and EthenoA-containing plasmid DNA was incubated with *in vitro* packaging extracts. The DNA was purified and digested to determine whether the EthenoA lesion remained intact after incubation with the packaging extracts.

The differences in stability from the abasic site could be explained by the fact that the repair system for the EthenoA lesion is an inducible one and would not be present in uninduced cells (125, 126). The dSpacer lesion is repaired by both the BER and NER pathways, which are constitutively expressed (see review (105)). From a practical
standpoint, we proceeded with the EthenoA-containing DNA for subsequent experiments

*Recombination of damaged plasmid with bacteriophage λ Kytos DNA*

The φC31 integrase catalyzes the nonreversible recombination of two DNA substrates, one containing the *att*<sub>B</sub> site and one containing the *att*<sub>P</sub> site. These molecules are depicted in Figure 20. The blue and purple circular plasmid contains the *att*B site, which is represented as a green arrow. The green linear bacteriophage DNA contains the *att*<sub>P</sub> site, which is represented as a yellow arrow.

![Figure 20. Schematic integration of λ Kytos and pMPA4.](image)

Integration occurs at the *attB* (plasmid DNA) and *attP* (λ DNA) sites to incorporate the plasmid DNA and the damage it contains into the λ Kytos construct. Upon digestion with EcoRI, it is possible to observe the disappearance of the bands from the plasmid (blue, 2431 bps) and λ Kytos (green, 1412 bps) and the appearance of the integration product (blue and green, 3603 bps).
When equimolar amounts of plasmid and λ Kytos DNA were mixed with increasing amounts of φC31 integrase, the plasmid was efficiently integrated into the λ Kytos. Integration was revealed upon digestion with EcoRI as the production of a 3603-bp fragment consisting of the 1412-bp fragment from the λ Kytos vector joined to the majority of the original 2431-bp fragment from the pMPA4 plasmid (Figure 20 and Figure 21, blue and green fragment). At the higher concentrations of φC31 integrase, the reactions were nearly complete, with a large majority of the λ Kytos DNA converted to recombinant form. The integration reaction is an extremely efficient one and can be used to produce large amounts of damaged λ Kytos DNA for packaging and infection.
Figure 21. *In vitro* integration reaction.

Integration of the λ Kytos construct and the pMPA4 plasmid can be completed in an *in vitro* reaction as indicated by the disappearance of the 1.4-kb green λ fragment. Large amounts of the integrated DNA can be produced (blue and green, lesion-containing fragment), which will yield enough substrate for packaging and infection. UC indicates uncut pMPA4 plasmid and λ Kytos DNA. The black triangle indicates increasing concentrations of φC31 integrase.

**Packaging of damaged λ Kytos DNA**

After the production of damaged λ Kytos DNA, the final hurdle was packaging the DNA *in vitro*. In the initial experiments, one packaging reaction was performed for
each construct (wild-type or EthenoA). These reactions yielded titers of approximately $10^8$ PFU/mL with the homemade packaging extracts and $10^9$ PFU/mL with the commercial extracts. The titers obtained using both packaging extracts are sufficient for performing infections to monitor DNA repair \textit{in vivo}. It may be possible to improve the phage titers by pooling packaging reactions and concentrating them by PEG-8000 precipitation.

Finally, the packaged DNA was examined to assure that the EthenoA lesion was still intact. First, the packaging reaction (~535 µL) was treated with 20 U of DNase for 20 minutes at room temperature. This step eliminated any unpackaged DNA. While the EthenoA-containing plasmid DNA was not affected by incubation with packaging extracts (Figure 19), it was necessary to verify that the plasmid behaved in the same manner once it was integrated into the $\lambda$ Kytos DNA. The wild-type DNA was digested by both EcoRI and XhoI. The EthenoA-containing DNA was digested by EcoRI but was still resistant to XhoI (Figure 22, compare lanes 2 and 4), indicating that the EthenoA lesion was intact after packaging.
Figure 22. Packaging of the integrated λ DNA.

The integrated wild-type and EthenoA-containing DNA was packaged with in vitro packaging extracts. The DNA was purified and digested to determine whether the EthenoA lesion remained intact after packaging.

Together, these various processes constitute a system that allows for the production and packaging of bacteriophage λ DNA with a single, site-specific lesion. A
total of $10^9$ cells could be infected with an average of 1 PFU/cell. The fate of this DNA can be monitored in vivo due to the high efficiency of infection, as shown by the in vivo data.

### 4.4 Discussion

Genomic DNA is often afflicted with lesions that can interfere with DNA replication, transcription, and gene expression. These lesions must be repaired to maintain genomic stability, and a variety of pathways exist to do so. These pathways were elucidated using a wide array of techniques that included both in vivo and in vitro studies of site-specific and general DNA damage (see Introduction). While all four types of studies have provided valuable information about DNA repair, the goal of this study was to produce a system to study site-specific lesions in vivo by direct physical analysis. By changing the integration system and the lesion, this system could be completed and used to examine in vivo DNA repair. The final system described above can be used to answer many of the same questions as in previous studies and can examine other factors that previous experiments could not.

A key advantage of this system is the amount of DNA introduced into the cells. This significant improvement in the amount of DNA increases the number of techniques that can be used to examine the DNA’s fate. By collecting larger amounts of DNA, it is possible to measure DNA repair by direct, physical methods. Isolation and digestion of
the DNA, one-dimensional (1D) gel electrophoresis, and Southern blotting allows direct measurement of how much DNA is repaired based on restoration of the restriction site originally interrupted by the lesion (XhoI site – EthenoA lesion). This technique and the results it yields will be the focus of the final chapter. Other techniques such as competitive PCR (127) or real-time PCR (128) could be used to obtain extremely quantitative answers about the amount of DNA repair, which will support the Southern blotting results. These techniques could also be performed with smaller amounts of DNA, which allow for repair measurements under difficult conditions where the amount of infecting DNA or the number of cells may be lower. By introducing larger amounts of DNA, the system can be used to monitor repair kinetics of the lesion. The DNA can be collected at various time points to examine the extent of repair. This feature is in contrast to many previous studies that examined only the endpoint of DNA repair (79, 80). A second key advantage of this system is that the cells do not have to survive in order to yield results. In assays that analyze the DNA from colonies that are allowed to grow overnight, treatments that kill the cells or a mutation that kills the cells would not yield usable results. By collecting the cells without a plating step, the DNA from “dead” cells can be harvested and analyzed to determine DNA repair after a harsh treatment or after the induction of a temperature sensitive mutant. This possibility greatly increases the conditions under which repair can be studied.
Two-dimensional gel electrophoresis and Southern blotting can be used to directly observe the behavior of the replication fork as it encounters the DNA lesion. While the prevention of replication simplifies the analysis of DNA repair, the study of DNA replication can be performed by either expressing the pir gene to allow replication from the R6Kγ origin in the plasmid or by infecting non-lysogen strains, which allows the bacteriophage λ construct to replicate using its own origin. Knockout mutants that prevent repair of the lesion would be constructed and infected with damaged DNA. After isolating the DNA and performing 2D electrophoresis and Southern blotting, DNA intermediates could be observed. The build-up of intermediates indicate blockage of the replication fork by the DNA lesion. If the lesion does not block the fork, the DNA can be sequenced to determine if the replication is error-free. If the lesion blocks the fork, but the blockage is resolved over time, additional mutants can be tested to determine the factors involved in resolving/restarting the stalled fork.

A variety of DNA lesions exist, and this system provides the flexibility to study them while using the main DNA components. The fate of different lesions can be studied simply by adding a new oligonucleotide to the starting pMPA4 plasmid. Any lesion that can be produced in an oligonucleotide, either by synthesis or chemical modification, should be incorporated into the plasmid. However, several factors may affect whether a lesion can be studied in this system. Lesions will have to first be
incubated with packaging extracts to ensure that the lesion is not repaired during packaging, as was observed with the abasic site lesion. For lesions that are repaired in the extracts, steps can be taken to avoid this repair. The extracts are produced by growth of cells containing the plasmids that express the packaging components. These plasmids could be transformed into repair knockout cell lines in an effort to produce extracts lacking the problematic repair activities. Repair proteins could also be immunodepleted from the extracts. Finally, the packaging reaction could be performed with purified proteins rather than extracts (129). Another factor to consider is whether the damaged bacteriophage λ DNA will be efficiently packaged. Large, bulky lesions may decrease packaging efficiency or prevent packaging completely.

With only a few changes, this system can be used to study several additional questions. By changing the sequence of the restriction site insert in the pMPA4 plasmid, the Nt.BstNBI nicking sites could be placed on the opposite strand of the plasmid. This change would allow the study of differences in both repair of and replication through the lesion when it was placed on the leading or lagging strand of the DNA when replication is allowed from the R6Kγ or λ origin. The lesion’s position in the leading or lagging strand may affect repair speed, replication blockage, mutagenesis, and other factors. Additionally, this system can be used to study transcription-coupled repair (TCR). By using a promoter already present in one of the DNA molecules or inserting a
new promoter into the pMPA4 plasmid, TCR of various lesions can be examined. Many questions still exist about TCR, including the diversity of lesions that can be repaired (130-132). By examining TCR in various knockout strains, the factors that may be recruited by the TCR machinery or may act downstream of TCR can be determined.

This system in its current state and with several modifications can be used to answer a variety of interesting questions about the repair and replication of DNA lesions. It can be used to verify studies that have already been performed in vitro, it can be used to find other unknown factors that may be involved in already studied processes, and it can be used to study new processes about which little is known, including TCR. It is particularly useful for comparing the efficiencies for competing repair pathways. It is a valuable system that will provide new insights into how E. coli deals with a variety of DNA lesions. The final data chapter will discuss the use of this system to study the repair of EthenoA.
5. *In vivo* repair of the EthenoA lesion

5.1 Introduction

The alkylation of DNA can lead to many different lesions that are repaired by different pathways. The difference in the lesions produced is dependent on the type of agent. For example, MMS and dimethylsulfate (DMS) are $S\equiv 2$ agents, which methylate DNA almost exclusively at the nitrogen atoms of the purine and pyrimidine rings. These agents can produce lesions such as 1-methyladenine (1-meA) (133). In contrast, $N$-methyl-$N'$-nitro-$N$-nitrosoguanidine (MNNG) and MNU are $S\equiv 1$ agents, which methylate nitrogen and oxygen atoms in the DNA bases and oxygen atoms in the sugar-phosphate backbone. These agents can produce lesions such as 7-methylguanine (7-meG), 3-methyladenine, and $O^6$-methylguanine ($O^6$-meG) (134). These lesions are simple additions of single methyl groups. More complex lesions can be generated upon exposure to additional compounds.

As stated in the previous chapter, EthenoA, a cyclized adenine lesion, is caused by the modification of an adenine base after exposure to vinyl chloride and other vinyl compounds (117) or lipid peroxidation (118). Early in its study, EthenoA was shown to lead to miscoding during *in vitro* DNA synthesis with *E. coli* DNA polymerase I (135). Additional studies have yielded some contrasting reports about which proteins are involved in EthenoA repair and the type of changes it causes in the DNA. In *E. coli*, this
lesion has both mutagenic and genotoxic effects in single-stranded DNA, though the
genotoxic effects are much stronger. An early study determined that the induced
mutations were mainly A to T and A to C transversions (136). This result differs from
that from a second group, which observed that the induced mutations were mainly A to
G transitions, and mutation occurred in only 0.1% of survivors. However, the EthenoA
lesion did reduce cell survival by 65% (69). It is important to note that the genotoxic and
mutagenic effect of the lesion were almost eliminated in double-stranded DNA; in fact,
the mutation rate was too low to be determined. It was speculated by the authors that
the DNA lesion was actually repaired in vivo and that the observed mutations were an
extremely small portion of the DNA. In contrast to this low mutagenesis rate in E. coli,
this lesion was highly mutagenic in mammalian cells and caused a mutation in 70% of
survivors, 63% of which were A to G transitions (137). Though it is only weakly
mutagenic in E. coli, this lesion must still be repaired as it has been shown to inhibit
DNA replication fork progression (138) and DNA synthesis by approximately 80 to 90%
(70). In addition to observing the inhibition of DNA synthesis, the study by Pandya et al.
used a plasmid containing a single EthenoA lesion, strand-specific markers, and
oligonucleotide hybridization to determine the fate of the DNA in various progeny after
replication. A full description of this study, including the strategies used can be found in
the last paragraph of Section 1.2.3. In this system, EthenoA was shown to be poorly
repaired by excision repair, to incorporate the correct nucleotide (dTTP) opposite the lesion, and to inhibit DNA synthesis by approximately 80 to 90% as previously stated. The DNA sequence results also indicate that the inhibition of DNA synthesis is rescued by daughter strand gap repair; the involvement of this pathway is supported by the observation that the recA, recF, recO, and recR genes are required for the production of these progeny (70).

Like many lesions, multiple pathways exist to repair EthenoA. The first pathway is BER, which begins with removal of the alkylated base by a DNA N-glycosylase. Again, different studies have led to different results. An early study indicated that the AlkA glycosylase can excise the EthenoA lesion to create an abasic site. Its activity was tested against a single EthenoA lesion in a 25-bp oligonucleotide duplex. The double-stranded DNA was treated with AlkA and other enzymes known to create abasic sites, including Tag, FGP, Nth, Nfo, and Xth. The DNA was then treated with sodium hydroxide to reveal any generated abasic site. AlkA and the human 3-methyladenine DNA glycosylase (ANPG) were the only proteins capable of acting on the EthenoA lesion (119). AlkA has been shown to act on a wide range of substrates (139). A second DNA glycosylase, Mug, has also been shown to remove the EthenoA lesion to generate an abasic site. However, this study showed that the effect of AlkA on EthenoA repair was negligible (120). The contrasting results are not surprising as the studies were
performed very differently. The first two studies were performed in vitro with purified AlkA protein, and the substrate was a duplex oligonucleotide containing a site-specific lesion. The last study was performed in vivo with plasmid DNA treated with chloroacetaldehyde (CAA).

The second pathway involves the E. coli AlkB protein, an iron- and 2-oxoglutarate-dependent dioxygenase, which is able to directly repair the EthenoA lesion by opening the ring formed by the addition of an ethyl group. This opening results in the release of a glyoxal molecule and the restoration of the correct DNA base (121) without its removal. The AlkB gene was first observed to control cell sensitivity to MMS (140). The researchers screened for cells sensitive to MMS but not UV irradiation and found seven mutants. Five of these mutants were mapped to the same region as the alkA gene; the other two mutants were mapped near the nalA gene and the responsible gene was named alkB. Almost twenty years later, two groups published studies on the function of purified AlkB. The study by Trewick et al. demonstrated that AlkB was able to repair 1-methyladenine and 3-methylcytosine in both single-stranded and double-stranded DNA in an oxygen-, α-ketoglutarate-, and Fe(II)-dependent manner (141). Repair was measured in vitro using several different techniques, including the release of radioactivity, HPLC, and GC-MS. A second study published in the same year also showed that AlkB released replication blocks and directly repairs 1-methyladenine in
single-stranded DNA \textit{in vitro} using the release of radioactivity and HPLC. The authors also demonstrated that the plasmid expression of AlkB in both wild-type cells and AlkB knockouts improved the transformation efficiency of MMS-treated single-stranded DNA \textit{in vivo} \cite{142}. In a study from the Essigmann lab \cite{121}, AlkB was also shown to repair EthenoA lesions in single-stranded viral DNA. The study demonstrated that while the rate of mutation caused by an EthenoA lesion in that single-stranded DNA was very low in wild-type cells (0.1%), the rate increased to 35% in AlkB knockout cells, with the majority being A to T mutations. Knocking out the \textit{alkB} gene also worsened the genotoxic effect. EthenoA-containing DNA was replicated 85% as efficiently as wild-type DNA in wild-type cells; however, replication efficiency decreased to only 5% in the AlkB knockout cells. Replication efficiency was determined using the Competitive Replication of Adduct Bypass (CRAB) assay. The authors were also able to demonstrate that repair of EthenoA by AlkB occurs by direct reversal as indicated by the appearance of glyoxal isomers in GC-MS experiments. The production of glyoxal comes from the loss of the etheno bridge in the EthenoA lesion. Finally, AlkA was shown to have a greater effect than AlkB on cell survival upon treatment with CAA. This treatment would presumably result in the production of etheno lesions in double-stranded DNA. Though the loss of AlkA resulted in a more severe phenotype, the loss of AlkB still caused a significant drop in cell survival.
Both AlkA and AlkB are inducible enzymes controlled by the adaptive response in *E. coli*, which also regulates the *ada* and *aid* genes (see review (122)). Upon exposure to alkylating agents, the Ada protein acts as both a repair and a regulatory protein. The protein transfers a methyl group from certain methylated bases to itself. This act converts Ada to a transcription activator that induces expression of the other genes of the regulon; in some cases, removal of the methyl group also repairs the lesion.

### 5.2 Materials and Methods

#### Materials

All the restrictions enzymes were purchased from New England Biolabs (Ipswitch, MA). Most oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). The oligonucleotide containing the EthenoA lesion was purchased from Sigma-Aldrich (Saint Louis, MO). The *in vitro* MaxPlax Lambda Packaging Extracts were purchased from Epicentre (Madison, WI). Both the QIAquick PCR Purification Kit and Gel Extraction Kit were from Qiagen (Valencia, CA). The Nytran transfer membrane was purchased from Whatman, Inc. (Piscataway, NJ). The Random-Primed DNA Labeling kit was obtained from Roche (Pleasanton, CA).

#### Strain construction

The *E. coli* knockout strains were constructed by P1 transduction. Lysogenized AB1157 cells served as the recipient strain; they were grown overnight, pelleted, and
resuspended in 0.5 volumes of 100 mM MgSO_4 and 5 mM CaCl_2. These cells were then mixed with various amounts of the desired P1 lysate and incubated without shaking at 30°C for 30 minutes. One hundred microliters of 1 M sodium citrate were added to stop the infection. One milliliter of Luria Broth (LB) was added, and the mixture was incubated at 37°C for one hour. The cells were then pelleted, resuspended in 100 µL of LB supplemented with 20 mM sodium citrate, plated on plates containing 50 µg/mL kanamycin and 4 mM sodium citrate, and grown at 37°C overnight. The presence of the knockout was confirmed by PCR.

*Preparation of damaged DNA*

The wild-type and EthenoA-containing DNA was produced as described in Chapter 4.

*φC31 integrase reactions*

The integrase reactions were performed as described in Chapter 4.

*Packaging reactions*

Recombined λ Kytos DNA was packaged with commercial MaxPlax Lambda Packaging Extracts. A maximum volume of 10 µL of DNA was added to 25 µL of extracts and incubated at 30°C for 1.5 hours. After packaging, each reaction was diluted with 500 µL of phage dilution buffer (10 mM Tris-Cl pH 8.3, 10 mM MgCl_2, and 100 mM NaCl) and pooled as appropriate. The packaged DNA was then treated with 20 U of
DNaseI for 20 minutes at room temperature to eliminate any unpackaged DNA. The packaging reactions were titered according to the manufacturer’s instructions to determine the overall titer of the reaction.

Infections with packaged DNA

The strains to be infected were grown in the presence of 20 mM MgSO₄ and 0.4% maltose to an OD₆₀₀ of 0.5. Additional treatments, including exposure to MMS and protein expression induction by arabinose, will be described. The culture was centrifuged at 6,000 x g to pellet the cells, which were then resuspended in 10 mM MgSO₄. For each time point desired, a total of 200 µL of phage was mixed with 150 µL of resuspended cells to give an MOI of approximately 0.5. This calculation is based on the observation that the in vitro MaxPlax Lambda Packaging Extracts normally yield titers of ~ 5 x 10⁸ PFU/mL.

DNA analysis by Southern blotting

The presence of the lesion was determined by restriction enzyme digestion and Southern blotting of DNA harvested from the infected cells. The extracted λ Kytos DNA was then singly or doubly digested. RNase was included in each of the digests. These digests were separated on a 0.8% agarose gel, and the gel was transferred to a 0.45 µm Nytran transfer membrane by the downward sponge method (104). After transfer, the DNA was crosslinked to the membrane with a 120 mJ/cm² UV exposure. The membrane
was blocked with hybridization buffer (0.5 M sodium phosphate pH 7.3, 4% sodium dodecyl sulfate (SDS), 1% BSA, 0.2 mg/mL salmon sperm DNA) while the radiolabeled probe was generated from various species of digested pMPA4 plasmid DNA or pMPA-derived PCR product using the Random-Primed DNA Labeling kit. The probes used are listed with the blots, and information about the probes can be found in Table 3. All blots were visualized using a Storm 860 Phosphorimager and quantitated with ImageQuant software (Molecular Dynamics; Sunnyvale, CA).

### Table 2. Probe constructs for Southern blotting.

<table>
<thead>
<tr>
<th>Probe name</th>
<th>Base position in pMPA4</th>
<th>Digest or PCR primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>homology-free</td>
<td>1162-2387</td>
<td>BsmI, AatII, and EcoRV</td>
</tr>
<tr>
<td>non-ASKA</td>
<td>1093-2798</td>
<td>5’ CGGCACCGTCACCCTGG 3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5’ CGCCCACATACGGATGGC 3’</td>
</tr>
<tr>
<td>methylation</td>
<td>3103-481</td>
<td>NcoI and AseI</td>
</tr>
</tbody>
</table>

### 5.3 Results

**Initial repair experiments**

A small-scale experiment was performed to gain an idea about the extent of DNA repair. AB1157 lysogens were used in the *in vivo* experiments because when they are infected with integrated λ Kytos, that DNA molecule is unable to replicate because the cells are lysogenic with wild-type bacteriophage λ, which is integrated into the chromosome and expresses the λ repressor. This repressor prevents the newly infecting
λ DNA from replicating, which could yield false positive results indicating repair. Wild-type AB1157 lysogens were left untreated or treated with 20 mM MMS for 20 minutes before infection. The untreated cells might not produce enough of the proteins required to repair the EthenoA lesion, leading to persistence of the lesion. Treatment with MMS should induce production of the AlkA, AlkB, and AidB proteins via Ada activation, leading to more efficient EthenoA repair. The cells were infected with either wild-type or EthenoA-containing DNA and digested with EcoRI alone (even lanes) or EcoRI and XhoI (odd lanes) to determine if the lesion was repaired. Digests of the wild-type DNA (Figure 23, lanes 1-7) yielded a single band upon digestion with EcoRI (lanes 2, 4, and 6), and that band was shifted upon digestion with EcoRI/XhoI, indicating its smaller size (lanes 1, 3, 5, and 7). The results were the same whether the cells were unexposed (lanes 1-3) or exposed to MMS (lanes 4-7). The 0 minute, EcoRI digested timepoint was excluded due to number of lanes available on the gel.

Not surprisingly, the EthenoA-containing DNA collected from cells that were not exposed to MMS was resistant to digestion by XhoI (Figure 23, lanes 8 vs. 9 and 10 vs. 11). Additionally, the EthenoA-containing DNA collected at the initial time of infection (0 min) from MMS-treated cells was not repaired (Figure 23, compare the digests of WT DNA in lanes 4+5 with the digests of EthenoA-containing DNA in lanes 12+13). Surprisingly, however, the EthenoA-containing DNA remained resistant to XhoI.
digestion 30 minutes after infection, indicating that the lesion was not repaired (Figure 23, compare the digests of WT DNA in lanes 6+7 with the digests of EthenoA-containing DNA in lanes 14+15). After seeing the results of this initial experiment, the decision was made to try to force repair by overexpressing repair proteins in the cells that were to be infected.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0</th>
<th>30</th>
<th>0</th>
<th>30</th>
<th>0</th>
<th>30</th>
<th>0</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Xhol</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>20 mM MMS</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lane</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 23. Test of in vivo repair of EthenoA-containing DNA.
The integrated wild-type (lanes 1-7) and EthenoA-containing DNA (lanes 8-15) was packaged with *in vitro* packaging extracts and used to infect AB1157 lysogens that were unexposed or exposed to 20 mM MMS. The DNA was purified and digested to determine whether the EthenoA lesion remained intact after infection. The digests are described in the text and listed in the table above the figure. The green arrow indicates the band that will be produced if the Xhol site is cut. The “homology-free” probe was used for this blot.

In 2005, Kitagawa *et al.* (143) published a paper describing the construction of a plasmid library from the *E. coli* open reading frame (ORF) archive. These plasmids allow for the overexpression of any *E. coli* protein using IPTG induction. It was hypothesized that induction of the *ada* regulon by MMS might be insufficient to repair the damage generated in both the genomic and infecting DNA, and that the overexpression of these proteins might be required to drive repair. Protein expression was induced by the addition of 1 mM IPTG, and the cells were infected as previously described. A new digest of SalI and XhoI was used to eliminate the star cutting that occurred upon digestion with EcoRI. Figure 24A shows the result of a mock infection, which yields no bands of the expected sizes upon SalI and SalI/XhoI digestion; there are some larger cross-reacting bands in the AB1157 and AlkA+ infections, which are due to interaction between the probe and the genomic DNA or AlkA overexpression plasmid. The wild-type DNA was cut by SalI to yield a single band (Figure 24B, lanes 1+3) and by SalI and XhoI to yield a second, smaller band (Figure 24B, lanes 2+4, green arrow). This result was expected, as there was no lesion in the XhoI site to prevent digestion. However, the
DNA containing the EthenoA lesion remained almost completely resistant to XhoI digestion, indicating that the lesion was not repaired (Figure 24B, lanes 6, 8, 10, 12, 14, and 16). This lack of repair was consistent when no DNA repair protein was overexpressed (Figure 24B, lanes 6+8), when AlkA was overexpressed (Figure 24B, lanes 10+12), and when AlkB was overexpressed (Figure 24B, lanes 14+16). There is a very faint band in lanes 6, 8, 10, 12, 14, and 16 that runs at the same size as repaired DNA and could be a small amount of repair. However, the band is difficult to see, and it could be argued that such a faint band is not due to repair but to aberrant XhoI digestion. There
Figure 24. Test of *in vivo* repair of EthenoA-containing DNA in the presence of AlkA or AlkB overexpression.

A. A mock infection of the three cell strains. B. The integrated wild-type (lanes 1-4) and EthenoA-containing DNA (lanes 5-16) was packaged with *in vitro* packaging extracts and used to infect AB1157 lysogens (lanes 5-8), AB1157 lysogens overexpressing AlkA (lanes 6-12), AB1157 lysogens overexpressing AlkB (lanes 13-16). The DNA was purified and digested to determine whether the EthenoA lesion remained intact after infection. The digests are described in the text and listed in the tables above the figure. The green arrow indicates the band that will be produced if the XhoI site is cut. The “non-ASKA” probe was used for this blot.

is also some smearing of the DNA collected from the AlkB overexpression strain (Figure 24B, lanes 13-16). This experiment will have to be repeated to determine if this result may be due to AlkB overexpression.
Testing DNA modification

The observation above of little to no repair caused concern, as there seemed to be no reason that AlkA and AlkB would not repair the DNA. A possible explanation that arose was that the DNA was never exposed to the repair proteins, because either the DNA was not injected into the cells and the phage was merely stuck to the outside of the cell or the DNA was not accessible even after injection by the phage. To test DNA accessibility, an experiment was designed to determine if the DNA was acted on by other DNA modifying enzymes, specifically EcoRII methylase. AB1157 lysogens were modified for this experiment. The dcm gene, which produces the dcm methylase, was knocked out by P1 transduction. The pBAD30-EcoRII plasmid, which overexpresses the EcoRII methylase was then transformed into the knockout strain. The dcm and EcoRII methylase proteins methylate the same recognition site in the DNA. These modifications resulted in the production of two strains with different methylation activities. In dcm AB1157 lysogens, the DNA will not be methylated, but in dcm AB1157 lysogens with pBAD30-EcoRII plasmid, the DNA should be methylated with induction of the expression plasmid. These cells were then infected with wild-type integrated λ Kytos containing the pMPA4 plasmid that was isolated from a methylase deficient strain and packaged either in vivo or in vitro. The in vivo packaged DNA serves as a control to
assure that the in vitro packaging extracts yield viable phage particles that can infect E. coli cells. This experiment was repeated to determine if the results were consistent.

These experiments yielded several interesting results. First, the in vivo and in vitro packaged DNA infected the lysogens with similar efficiencies, as the amount of harvested DNA was very similar from one packaging method to the other in both experiments (Figure 25A, compare lanes 1-8 and lanes 9-16; Figure 25B, compare lanes 1-8 and lanes 9-16). To determine the methylation status of the DNA, the samples were digested with AccI alone, which generates two bands of 2,190 bp and 1,638 bp. The 2,190-bp fragment contains a MscI site that will always be cut, leading to the production of a smaller 729-bp band and a 1,461-bp band. Neither of these bands could be visualized on the blot; the 729-bp band may run off the gel or be too diffuse and the 1,461-bp band only hybridizes with 38 bp of the probe, which will not yield enough signal to visualize. The disappearance of the 2,190-bp band does serve as an excellent control to ensure that the MscI enzyme is cutting properly. The 1,638-bp band is the informative one, as it contains a MscI site that is subject to blockage by methylation of the overlapping EcoRII methylation site. If the site is methylated, the 1,638-bp band will persist. If the site is not methylated, the 1,638-bp band will be cut into a 1,340-bp band that can be visualized and a 298-bp band that will run off the gel. The percentage of methylation is equivalent to the percentage of uncut DNA after digestion with
Accl/MscI. The intensity of the 1,638-bp band in each sample was determined using a median background correction method. The following equation was then used to determine the percentage of uncut DNA, which is equivalent to the percentage of methylated DNA.

\[
\text{% uncut} = \left( \frac{\text{intensity of 1,638-bp band}_{\text{Accl/MscI}}}{\text{intensity of 1,638-bp band}_{\text{Accl}}} \right) \times 100
\]

The methylase deficient strain, the \textit{dcm} AB1157 lysogen, serves as another good control. Upon digestion with Accl and MscI, both the 2,190-bp and 1,638-bp bands are cut into their smaller bands (Figure 25A and 25B, lanes 1+2, 3+4, 9+10, and 11+12). In the majority of the samples from the methylase deficient strains, the amount of methylated DNA is only a few percent (0.0 – 3.8%, Figure 25A, lanes 2, 4, 10, and 12; 0.0 – 4.9%, Figure 25B, lanes, 2, 4, and 10). The one exception is the \textit{in vitro} packaged 60 minute sample (Figure 25B, lane 12), in which 15.9% of the DNA appears to be methylated. However, closer examination of that lane and that region of the gel reveals that there is not a convincing 1,638-bp band present. The signal may come from the higher level of background in that area of the blot, and it does not appear that there is truly any signal in that area, indicating that the true amount of methylated DNA is much closer to the other controls (0.0 – 4.9%).
Two experiments (A+B) were performed to test the methylation status of the infecting DNA in AB11257 lysogens. Integrated, unmethylated wild-type DNA was packaged with \textit{in vitro} packaging extracts and used to infect \textit{dcm} AB1157 lysogens (lanes 1-4 and lanes 9-12) and \textit{dcm} AB1157 lysogens with pBAD30-EcoR1I methylase (lanes 5-8 and lanes 13-16). The DNA was purified and digested to determine the methylation status after infection. The digests are described in the text and listed in the tables above the figure. The “methylation” probe was used for this blot.

Infection of the \textit{dcm} AB1157 lysogens that contain pBAD30-EcoRII plasmid leads to very different results. In the infections with \textit{in vivo} packaged DNA, the percentage of methylated DNA ranges wildly, from 72.2 and 84.0% in experiment 1 (Figure 25A, lanes...
6+8) to 19.1 and 21.4% in experiment 2 (Figure 25B, lanes 6+8). However, as this system will not use in vivo packaged DNA, these results are not of great concern. More importantly, in the infections with in vitro packaged DNA, the percentage of methylated DNA ranges from 46% to 77%. While the methylation levels are less consistent than hoped for, they do not vary as much as with in vivo packaged DNA, and the results are still informative. They indicate that at least 46% of the DNA and often more is modified by EcoRII methylase, which demonstrates that the DNA has been successfully injected into the cell and is accessible to proteins. Because the DNA is accessible to proteins like EcoRII methylase, there is no reason to believe that it is not also accessible to repair proteins. Even if we postulate that only 50% of the DNA is accessible, in vivo experiments have never led to the observation of 50% repair. It would be generous to speculate that even 5% repair can be observed (Figure 24B, lanes 6, 8, 10, 12, 14, and 16, green arrow). There is a significant difference between the level of DNA methylation and the level of DNA repair. This difference suggests that additional processes, perhaps replication or transcription, might be required for DNA repair.

*Testing both DNA methylation and repair*

After observing methylation of some but not all of the infecting DNA, it was hypothesized that some of the DNA may be inaccessible to any in vivo process. To test this hypothesis, unmethylated plasmid, either wild-type or EthenoA-containing, was
integrated into \( \lambda \) Kytos, packaged, and used to infect methylase proficient lysogens with or without MMS treatment. Digests with SalI, XhoI, and PspGI, a methylation sensitive enzyme, were used to determine if the DNA is repaired and methylated. Digestion with SalI yields a 2,190-bp fragment. That fragment contains the XhoI site where the lesion is located and seven PspGI sites. If the PspGI sites are efficiently methylated, the 2,190-bp fragment will persist even if digested with both SalI and PspGI (Figure 26A, left schematic gel). If the PspGI sites are not efficiently methylated, the 2,190-bp fragment will be cut into many smaller fragments upon digestion with SalI and PspGI, and these fragments will not be visualized (Figure 26A, right schematic gel). Repair of the DNA lesion will be monitored by XhoI digestion as described previously.

Infection of \( dcm^{-} \) AB1157 lysogens that express EcoRII methylase yielded very interesting results. First, both the wild-type and EthenoA-containing DNA were efficiently methylated, as the 2,190-bp SalI band persisted upon digestion with PspGI (Figure 26B; WT gel, lanes 1 vs. 2 and 4 vs. 5; EthenoA gel, lanes 1 vs. 2 and 4 vs. 5). Surprisingly, PspGI digestion of the DNA from the MMS-treated cells resulted in the disappearance of the 2,190-bp SalI band (Figure 26B; WT gel, lanes 7 vs. 8 and 10 vs. 11; EthenoA gel, lanes 7 vs. 8 and 10 vs. 11), indicating that MMS treatment prevented methylation by EcoRII methylase, though the reason for prevention is unclear. Additionally, treatment of the cells with MMS reduces that amount of DNA recovered,
which reduces detection ability. The second surprising result, which was much more exciting, was that the EthenoA lesion appeared to be repaired in the absence of MMS treatment (Figure 26B; EthenoA gel, lanes 2 vs. 3 and 5 vs. 6). *In vivo* repair of the EthenoA lesion had not been observed in any other experiments with this system. The possible reasons for this repair will be described in the Discussion section. Control experiments confirmed that the starting EthenoA-containing DNA was resistant to XhoI digestion, as it should be (data not shown).

**Figure 26.** Observing DNA methylation and repair.

A. Schematic of the Sall, PspGI, and XhoI sites in the pMPA4 plasmid and the digestion patterns. Methylation of the PspGI sites results in persistence of the band after digestion with Sall alone or Sall/PspGI (left gel). No methylation of the PspGI sites results in disappearance of the band after digestion with Sall/PspGI (right gel).

B. Monitoring of methylation and repair of infecting DNA.
5.4 Discussion

The original lack of repair of the EthenoA lesion after infection was a very surprising result that was observed whether the *ada* regulon is induced by treating the cells with MMS or whether the repair proteins AlkA and AlkB from the *ada* regulon, which have been previously shown to repair EthenoA, are overexpressed from ASKA plasmids. It seemed possible that the treatment of cells with MMS might not lead to repair. Treatment with MMS will also induce lesions in the genomic DNA. It may be that the repair proteins Ada, AlkA, and AlkB preferentially repair the genomic DNA, a much more valuable substrate than the infecting bacteriophage λ DNA. The lack of repair when the AlkA or AlkB protein is overexpressed is more difficult to explain, especially as both proteins have been shown to repair the EthenoA lesion *in vivo* and *in vitro*.

Before drawing any conclusions about the lack of repair, a point of concern – whether the DNA was ever exposed to the cell cytoplasm – had to be addressed. There are two possibilities that could explain how bacteriophage λ DNA could be detected but not repaired. The first was that the phage particle could attach to the cell but never inject the DNA. Upon harvesting the cells, the phage particle would also be pelleted, still attached to the cell membrane. Lysing the cells with proteinase K treatment and phenol/chloroform extraction would also cause the phage particles to be destroyed and their DNA to be harvested. Therefore, the samples would appear to contain
“unrepaired” phage DNA when the DNA had actually never been exposed to the cytoplasm and given the chance for repair. The other possibility is that the phage DNA is injected into the cell but is inaccessible to proteins in the cytoplasm. Both possibilities were tested by determining the methylation status of the DNA. If the DNA can be methylated by the EcoRII methylase, it can be assumed that it is injected into the cell and that other proteins, including the repair proteins, can access the DNA. Unmethylated DNA was packaged and used to infect cells that were methylation deficient or expressed the EcoRII methylase. When the methylase was overexpressed, the nonreplicating infecting phage DNA was methylated to various degrees. This result demonstrates that the DNA can be acted upon by cellular proteins after infection, and there is no reason to believe that this does not include repair proteins.

Finally, DNA methylation and repair were tested in the same experiment, which yielded the most surprising result yet. The EthenoA lesion was repaired when the DNA was also methylated, even without exposing the cells to MMS (Figure 26B). There are several possible explanations for the results observed in this experiment. The first is that the methylation of DNA is an active process, comparable to replication or transcription, and that active process also allows the DNA to be repaired. The importance of active processes and their involvement in the previous studies will be discussed in the next paragraph. The second possibility is that overexpression of the EcoRII methylase leads
to aberrant methylation of pseudomethylation sites. This aberrant methylation leads to activation of the McrBC system, which leads to DNA restriction and activation of the SOS response (144). One of the proteins activated by the SOS response could then repair the EthenoA lesion. This possibility could be tested by looking for SOS induction. The other possibilities are less interesting but still important. The third possibility is that the P1 transduction used to create the dcm knockout could have also transduced a second linked mutation that is important for detecting repair. A simple experiment can be used to test this possibility. If the EthenoA lesion is not repaired in the dcm- AB1157 lysogens that do not overexpress the EcoRII methylase, then there is no second mutation responsible for repair. The final possibility is that there was inefficient digestion by XhoI in previous experiments and this inefficient digestion prevented the observation of repair. This possibility is unlikely as XhoI digestion was observed with wild-type DNA, though it was not complete in some cases.

It is important to consider the results obtained in the system described above and those observed in previous studies. In many of the previous studies, extremely efficient repair was observed under many different conditions. The system described here seems to yield repair under only very specific conditions. These differences could exist for a variety of reasons. The first explanation for some of the differences in the observed results could be the use of an in vivo versus an in vitro system. For example, several
studies have shown the repair of EthenoA by AlkA (119, 145). However, both of those studies were performed in vitro with a single EthenoA lesion in a small double-stranded oligonucleotide, a very different environment from the cell cytoplasm and a very different substrate from bacteriophage λ DNA. It is not unexpected to observe different results with these two systems.

Other studies have used multiple techniques to study repair. One of those studies examined the repair of DNA lesions by purified AlkB in vitro. The release of radioactivity by purified AlkB from single-stranded poly(dA) oligonucleotide treated with [3H]MNU was concluded to be repair of methylated bases by AlkB (142). However, this type of experiment does not examine EthenoA repair directly because treatment with MNU can lead to the production of different types of lesions. Another in vitro study looked at the repair activity of AlkB. However, in contrast to the previous study, the researchers examined the repair of a site-specific EthenoA lesion in a single-stranded oligonucleotide by following the release of glyoxal isomers (121). Again, the differences in environment and substrate could easily explain the differences in results. When experiments are performed in vitro, only a select number of components are included. Minimizing the complexity of the environment makes it easier for the enzyme and substrate to find each other. Additionally, the ability to add large amounts of both the
enzyme and substrate ensures that the products of the reaction can be detected. The use of a smaller substrate also makes it easier for the enzyme to find the single lesion.

*In vivo* experiments come closer to mimicking the conditions in the experiments described above, but there are still many differences between the previous studies and the work described here. For example, many *in vivo* experiments have looked at the repair of generalized DNA damage created by treatment with chemicals such as MMS or CAA. These compounds will not only generate multiple lesions in the DNA but will generate multiple types of lesion in addition to EthenoA. The study by Falnes et al. (142) looked at the repair of MMS-treated, single-stranded DNA by AlkB treatment. The level of repair was defined as the transformation efficiency of DNA in various cell strains. They found that treatment with AlkB led to the highest transformation efficiency in both wild-type and AlkB mutant cells. When the AlkB treatment was omitted, the transformation efficiency was decreased in wild-type cells and more severely decreased in AlkB mutant cells. However, MMS treatment is not an effective method for the generation of EthenoA lesions. The majority of lesions generated by MMS are simply methylated bases, \(N^7\)-methyguanine and \(N^3\)-methylguanine (146); therefore, this study may not even be observing EthenoA repair. In addition to the fact that this study did not use a site-specific lesion or even a single type of lesion, the substrate DNA was single-
stranded, which is a significant difference. Additionally, transformation efficiency is not the same as the direct observation of repair.

Another *in vivo* study looked at the mutation frequency of DNA treated with CAA. While the lesions generated by CAA are mainly EthenoA and 3,\(N^4\)-ethylcytosine (147, 148), this study also suffers from the caveat that different types of lesions at many locations could be created in the DNA, though this substrate is double-stranded (120). This study was interesting because it looked at CAA-induced mutations in a plasmid by testing reversion of the *lacZ* gene. Therefore, repair is being observed in an actively transcribed region of the plasmid; the damage in the plasmid used in this study is not located in a transcribed region of the plasmid. This mutation frequency was determined by transforming different cell strains and plating them on chloramphenicol plates to determine the total number of transformed cells and on lactose minimal media plates to determine the number of lac revertants. The plasmid was transformed into mutants of the *alkA*, *alkB*, and *mug* genes, both with and without induction of the Ada regulon with MMS. The researchers determined that AlkB and Mug glycosylase were involved in EthenoA repair, but the contribution from AlkA was negligible.

Several studies have examined the repair of a single, site-specific EthenoA lesion. Single-stranded viral DNA was used in a study by Delaney *et al.* (121) to examine lesion bypass frequency and lesion mutation frequency in a wild-type and AlkB mutant strain.
The EthenoA lesion was found to have a minimal effect on replication in wild-type cells (replication efficiency = 85%). However, when that DNA was introduced into AlkB cells, the replication efficiency decreased to only 5%. Similarly, the EthenoA lesion is negligibly mutagenic in wild-type cells but 35% mutagenic in AlkB cells. Again, the observation of characteristics like replication efficiency and mutagenesis is not the same as the direct observation of repair.

Another study observed no effect of the EthenoA lesion on plasmid survival in double-stranded DNA that was transformed and allowed to replicate, but plasmid survival decreased to approximately 35% in the single-stranded DNA (69). A second study observed the inefficient repair of EthenoA by NER, the inhibition of DNA synthesis by EthenoA, and the incorporation of the correct base (T) across from the lesion (70). However, this study never considered the involvement of the Ada regulon, focusing instead on NER, which is not thought to be involved in the repair of EthenoA.

There are many differences in these in vivo studies, including the number of strands in the DNA substrate, how the DNA damage is created, and how the “repair” of the lesion(s) is determined. The common characteristic that unites them is the ability of the DNA to replicate once it is introduced into the cells. In one case, the DNA is even transcribed. This characteristic may explain the different results obtained with the system constructed here and the previous studies. In all these previous studies, the DNA
is undergoing different processes in the cytoplasm. In many of the experiments performed with this new system, the molecule was infected into cells and simply sat in the cytoplasm. Additionally, the lesion was not located in a transcribed region. It is only when the DNA was methylated that it was also repaired. It is possible that the DNA molecule must be “active” in order for proteins to find and repair the EthenoA lesion, and these possible activities include DNA methylation, replication, and transcription. Additional experiments are necessary to investigate the effect of “active” processes.
6. Conclusions and future directions

The construction of this system has been more difficult than initially thought and has required considerable troubleshooting at virtually every step. However, at this point, the system can be used to incorporate a site-specific lesion into a plasmid with very high efficiency. Any lesion that can be commercially produced or achieved with postproduction modifications, either enzymatic or chemical, can be inserted into the plasmid. Whether that lesion can be studied in vivo is then determined by both its stability in in vitro packaging extracts and its ability to package. If the lesion is not repaired by the extracts, it can be studied. If it is repaired by the commercial in vitro packaging extracts, homemade extracts can be produced from repair protein knockouts in an attempt to stabilize the lesion. Additionally, repair proteins could be immunoprecipitated from the extracts if knockouts cannot be constructed. Finally, if the repair proteins for a specific lesion are unknown, the components of the packaging system can be purified and reconstituted in a purified protein system (149). Larger, bulkier lesions may not package or may package less efficiently. Careful analysis of packaging efficiencies will be required when testing the packaging of new lesions.

By using the φC31 integrase and the att sites, the damaged plasmid can be incorporated into the λ Kytos vector. Again, this process is a very efficient one and can be used to produce bacteriophage λ DNA containing a site-specific lesion of choice. The
use of this DNA in *in vivo* experiments was described in previous chapters, but this same DNA could also be used in *in vitro* experiments by other researchers who are using traditional or new *in vitro* techniques. This DNA molecule could be particularly useful in single molecule studies of DNA repair as the molecule provides a large, easily visualized substrate upon which the repair proteins can assemble. The inclusion of a site-specific lesion in a much larger molecule (~53 kb) may provide experimental results that are a more accurate reflection of the results that would actually be observed in genomic DNA. One would expect that this larger DNA molecule would provide more accurate results than those obtained with a traditionally sized plasmid (3-6 kb).

In terms of studying the *in vivo* repair of EthenoA, several stones remain unturned. Overexpression of the EcoRII methylase led to the first observation of EthenoA repair (Figure 26). However, there are several possible explanations for the observation of this repair, and they must be tested before any conclusions can be drawn. Repair must be tested in AB1157 lysogens, *dcm*-- AB1157 lysogens, and the *dcm*-- AB1157 lysogens that express EcoRII methylase from the pBAD30 plasmid to assure that the repair is truly methylase dependent. If it is dependent on methylation, other “active” processes, such as replication and transcription, can be tested. If it is not dependent on methylation, the explanation for the different results must be determined. As the
detected repair was less than 100% after 30 minutes, replication or transcription could greatly stimulate repair, even if the processes are not strictly required.

To test the requirement for replication, a new plasmid has been designed to monitor both repair and replication and to determine if both processes are occurring on the same DNA strand (leading or lagging) (Figure 27A). Two new restriction sites will be added to the plasmid. The first is an AatII site, which will allow the DNA to be digested in such a way that it produces a 3.6-kb fragment that is comprised of most of the damaged plasmid. Essentially, an AatII digest will allow the plasmid to be cut from the integrated λ Kytos DNA molecule. The second site is a new MscI site. This MscI site contains an overlapping EcoRII methylation site (GGACC) on the leading strand of the DNA, which is the same strand that contains the DNA damage. If this EcoRII methylase site is methylated, it will prevent the MscI site from being cut. There is an existing MscI site in the plasmid that also contains an overlapping EcoRII methylation site (GGACC), this time on the lagging strand of the DNA. These two sites will allow the methylation of the leading and the lagging strand to be followed separately. Another existing MscI site that is not affected by methylation will be mutated so it is no longer susceptible to digestion. This new plasmid can be fully methylated before integration into λ Kytos and then be used to infect methylation deficient strains. As the methylation pattern changes upon replication, different digests can be used to determine whether repair is occurring
only if the DNA is replicated or if it can occur without replication. These possible replication and repair DNA species are shown in Figure 27B; for simplicity, these DNA species are shown as the starting plasmids, not the integrated λ Kytos. The starting DNA will be methylated on both strands and contain the DNA damage (species A'B'). If there is no repair, only replication, then the leading parental strand will be replicated (i) to form species A'B and will still contain the damage. However, the lagging parental strand will be replicated (ii) to form species AB' and will not contain the damage; therefore, this species will be digested with XhoI. However, the difference in the methylation status will allow this species to be ignored as it is not truly repaired just replicated. The second round of replication (iii) will produce more of species A'B and AB', but will also produce a fourth species (AB) that is completely unmethylated and will not contain the damage. If the DNA is repaired before replication, species A'B'R will be produced. If the DNA is repaired during or after replication, species A'BR will be produced. Digests and gels of in vivo samples can distinguished between the various results.
Figure 27. System for monitoring DNA repair and replication.

A. The new pMPA5 plasmid contains a new AatII site (green), a new MscI site (blue) that is blocked by methylation of the leading strand, and an existing MscI site (orange) that is blocked by methylation of the lagging strand. B. The possible replication products and their methylation states are shown, in addition to the products produced if repair occurs.

A schematic of the banding pattern produced by the digestion of the various DNA species is shown in Figure 28. The design of a new probe will assure that only the desired bands can be visualized. It is important to note that the unrepaired species can be distinguished from the repaired species. For example, unrepaired species A’B’ (3.6 kb) shifts to form a new band (A’B’R; 1.2 kb) upon repair. The unrepaired species A’B
(2.8 kb) also shifts to form that same new band (A’BR; 1.2 kb) upon repair. By using the correct probe and digests, DNA that is digested by XhoI because it has been replicated (AB’ and AB) can be eliminated, and the processes of repair and replication can be studied.

Figure 28. Schematic gel of digests to study repair and replication.

The digests of each plasmid species are shown in pairs. The first band is generated by a SalI/MscI digest. The second band is generated by a SalI/MscI/XhoI digest. The probe used for the Southern blot will eliminate uninformative DNA bands.
By performing these additional *in vivo* experiments to examine repair with methylation and replication, it can be determined if the system can provide interesting *in vivo* results after *in vitro* packaging. It is possible that a single EthenoA lesion is not efficiently repaired *in vivo* if the DNA is not undergoing any active processes. In addition to the observation of repair during replication, this system could be used to study the behavior of the replication fork itself. 2D gel electrophoresis and Southern blotting of *in vivo* samples can be used to determine if the DNA lesion of interest will physically block progression of the replication fork. By examining replication fork behavior in various *E. coli* mutants, the proteins involved in replication fork resolution can also be studied.

The process of TCR can also be studied with this system. A previous graduate student in the laboratory recently constructed a plasmid similar to pMPA4 that can be used to construct a substrate with site-specific damage. However, unlike pMPA4, the damage can be placed in an actively transcribing gene and used to study TCR and transcription in general. The process of transcription could prove to be an important one if it is determined that the lesion is repaired if the DNA molecule is undergoing replication but is not repaired if the DNA molecule is simply sitting in the cytoplasm. There are times during the cell cycle when the DNA does not replicate and there are non-replicating cells; however, in both cases, the DNA in the cells can still be damaged.
If certain repair processes do not occur in an “inactive” DNA molecule, it would seem that another “active” process would be necessary for repair. Transcription could be that process.

The system could therefore be used to study additional lesions and prove a powerful tool. In previous studies, the processes of repair and replication were not separated. By using λ lysogens and a plasmid with a conditional origin, replication of the infecting DNA can be tightly controlled. This ensures that researchers can allow only certain processes to occur and can look at repair under very specific conditions. The ability to force the infecting DNA to be an “inactive” molecule may lead to results that are very different from those observed in previous studies and may lead to a more accurate picture of the requirements for DNA repair.
References


gives a different mutation spectrum and increased error rate compared with replication past this lesion in uninduced cells, *J Bacteriol* 172, 2105-2112.


reaction with some tRNA constituents. Stable intermediates, kinetics and selectivity of the reaction, *Nucleic acids research* 5, 789-804.


Biography

Melissa Pierce Asllani was born on November 24, 1981 in Augusta, Maine. After graduating from Erskine Academy in South China, Maine in 2000, she began attending Saint Anselm College in Manchester, New Hampshire. In 2004, she earned a Bachelor of Arts degree in Biochemistry and a certificate in Spanish. In the fall of 2004, she entered the Department of Biochemistry at Duke University where she was awarded a Kamin Fellowship. In the spring of 2006, she joined Dr. Kenneth Kreuzer’s lab where she began her doctoral research producing a system to study DNA repair \textit{in vivo} and in real-time.