The Dissertation Committee for Zheng Lou certifies that this is the approved version of the following dissertation:

Transposable prophage Mu exists as an independent chromosomal domain in *E. coli*

Committee:

_____

Rasika M. Harshey, Supervisor

_____

Makkuni Jayaram

_____

Alan M. Lambowitz

_____

Tanya T. Paull

_____

Rick Russell

Transposable prophage Mu exists as an independent chromosomal domain in *E. coli*

by

Zheng Lou, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2012

Dedication

To my parents

Acknowledgements

First, I would like to express my deepest gratitude to my advisor, Dr. Rasika M. Harshey for her endless support and help throughout these years of my research. Her wisdom and enthusiasm, like a lighthouse in the dark, inspires and guides me. She sets up a perfect example of a great scientist that I believe will benefit me for the rest of my life.

I would like to extend my appreciation to my committee members: Dr. Makkuni Jayaram, Dr. Alan M. Lambowitz, Dr. Tanya T. Paull and Dr. Rick Russell for their suggestions and encouragement.

I also want to thank all the past and present lab members in Harshey and Jayaram lab for their help and friendship. My special thanks go to Dr. Jun Ge and Dr. Wonyoung Choi for their helpful advice and suggestion on my project.

I truly appreciated the friendship from my friends Dr. Xi Chen, Yang Wang, Song Han and their families.

I would like to thank my parents, Wenjun Lou and Xia Liu. Without their encouragement, I would not have a chance to be at the University of Texas at Austin. Finally, I thank my wife, Yu Tao, for her love and support.

Transposable prophage Mu exists as an independent chromosomal domain in *E. coli*

Zheng Lou, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Rasika Harshey

The 4.6 Mb circular *E. coli* chromosome is compacted by segregation into 400-500 supercoiled domains, created by both active and passive mechanisms like transcription and DNA-binding proteins. We find that transposable prophage Mu, transcriptionally silent by definition, is organized into an independent domain as determined by the close proximity of Mu termini L and R separated by a 37 kb Mu genome. Cre-*loxP* recombination is used in this study *in vivo* and *in vitro*. Critical to formation/maintenance of the Mu 'domain' configuration are a strong gyrase site SGS at the center of Mu, the Mu L end, the MuB protein, and the *E. coli* nucleoid-associated proteins IHF, Fis and HU. The Mu domain was observed at two structurally different chromosomal locations, and was specific to the Mu prophage, i.e. was not observed for the λ prophage. A model is proposed that by employing its *cis*-elements to create a domain barrier for segregation and compaction of its

genome, the large selfish DNA element Mu profits from the transposition-ready arrangement of its ends, while simultaneously providing a fitness advantage to the host.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

Transposable Elements

BACKGROUND

Transposable elements (TEs) or transposons are a class of DNA sequences that can change their relative positions within their host genomes. Barbara McClintock first discovered transposable elements in maize in the mid-1940s. Since then, TEs have been found to be widespread in nature. Examples range from the simple insertion sequences (IS) to the complex phage Mu in bacteria, P elements in fruit fly, and the retroviruses like HIV in humans (Craig, 2002). In eukaryotes, TEs occupy 50-80% of maize, 45% of human, 15% of fruit fly and 3% of yeast genomes, whereas in prokaryotes like *E. coli*, TEs only comprise 1-2% of the genome (Kidwell and Lisch, 2000; Mills et al., 2007).

Although in higher eukaryotes, most of the TEs are inactive copies of their ancestors and are labeled 'junk DNA', they have contributed to the structure, function, evolution and adaptation of genes and genomes (Kidwell and Lisch, 2000; Craig, 2002). The mobility of TEs can cause mutations in single genes or structural changes in the whole genome. For the single gene, the encoded protein would be changed if TEs insert into the DNA coding region, while gene expression patterns would be altered if TEs move into regions flanking the gene or into introns. Structural changes in the whole genome are caused by homologous recombination within the TE sequences, which result in genome rearrangements such as deletions, inversions and duplications. TEs can also be the force of molecular domestication, such as the mammalian V(D)J

recombination system which generates antigenic diversity (Hiom et al., 1998), and the non-LTR retrotransposons HetA and TART involved in the telomere maintenance system in *Drosophila melanogaster* (Levis et al., 1993). In bacteria, some TEs provide a selective advantage to their hosts by encoding antibiotic resistances (Kidwell and Lisch, 2000).

Years of research have developed transposable elements into powerful tools not only for basic biological research but also for biomedical use (Berg and Berg, 1996). Bacterial transposons, *Drosophila* P elements and animal *mariner* transposons have been utilized to study the function of genes through mutagenesis. Tn*3*, Tn*5*, and Mu have been exploited to sequence genomes (Strathmann et al., 1991). Recently, *Piggybac* and *Sleeping Beauty* transposons have been found to hold the promise of use in human gene therapy (Baudry et al., 2009).

CLASSIFICATION

According to their transposition intermediates, TEs can be divided into two major classes (Figure 1.1): RNA-mediated TEs or retrotransposons (Class I) and DNA transposons (Class II) (Finnegan, 1989; Curcio and Derbyshire, 2003). Retrotransposons replicate through intermediate RNA molecules, which are transcribed from their DNA genome and then reverse transcribed back into DNA. DNA transposons move directly without RNA intermediates.

RNA transposons are most common in eukaryotes, being especially abundant in plants and mammals. They consist of two classes: LTR (long terminal repeats) and non-LTR retrotransposons. LTR transposons share similar structural features with retroviruses such as HIV (human immunodeficiency virus) and MMTV (mouse mammary tumor virus) encoding reverse transcriptase (RT) and integrase (IN). The major difference between them is that, unlike

retroviruses, LTR transposons do not possess a functional *env* gene and thus cannot form viral particles to exit the cell; instead, they insert into the genome of the same cell. Non-LTR transposons lack LTRs but have a polyadenylate (polyA) sequence. They can be divided into two sub-types: long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). Both LINEs and SINEs are found in large numbers in eukaryotic genomes. In the human genome for example, SINEs occupy about 13% and LINEs make up 21% of the genome. LINEs encode their own endonuclease (EN) and RT required for retrotransposition, while SINEs take advantage of EN and RT from LINEs (Lander et al., 2001).

DNA transposons are most common in bacteria. Bacterial transposons can be divided into several classes: insertion sequences (IS), composite transposons, non-composite transposons and transposable phages (Kleckner, 1981; Curcio and Derbyshire, 2003). ISs, such as IS*1* and IS*2* etc., are the simplest transposons, consisting of only the elements necessary for transposition such as the transposase gene and a pair of 10-40 bp inverted repeats at both ends. Composite transposons, such as Tn*5* and Tn*10*, contain drug resistance or other beneficial genes between two IS elements. Non-composite transposons, like Tn*3* and Tn*7* have 38-40 bp inverted repeats at both ends like the IS elements, and carry in addition antibiotic resistance genes like the composite transposons. They also encode a transposase, and some carry a site-specific recombinase required for resolving the products of transposition (Craig, 2002). Bacteriophage transposons, such as phage Mu and its relative D108, are bacterial viruses that can replicate and insert into the bacterial genome through transposition. The ends of their genomes carry inverted repeats similar to those in other transposons, and they encode transposase and accessory functions that enhance transposition greatly. Their complex genomes also encode proteins required for phage morphogenesis.

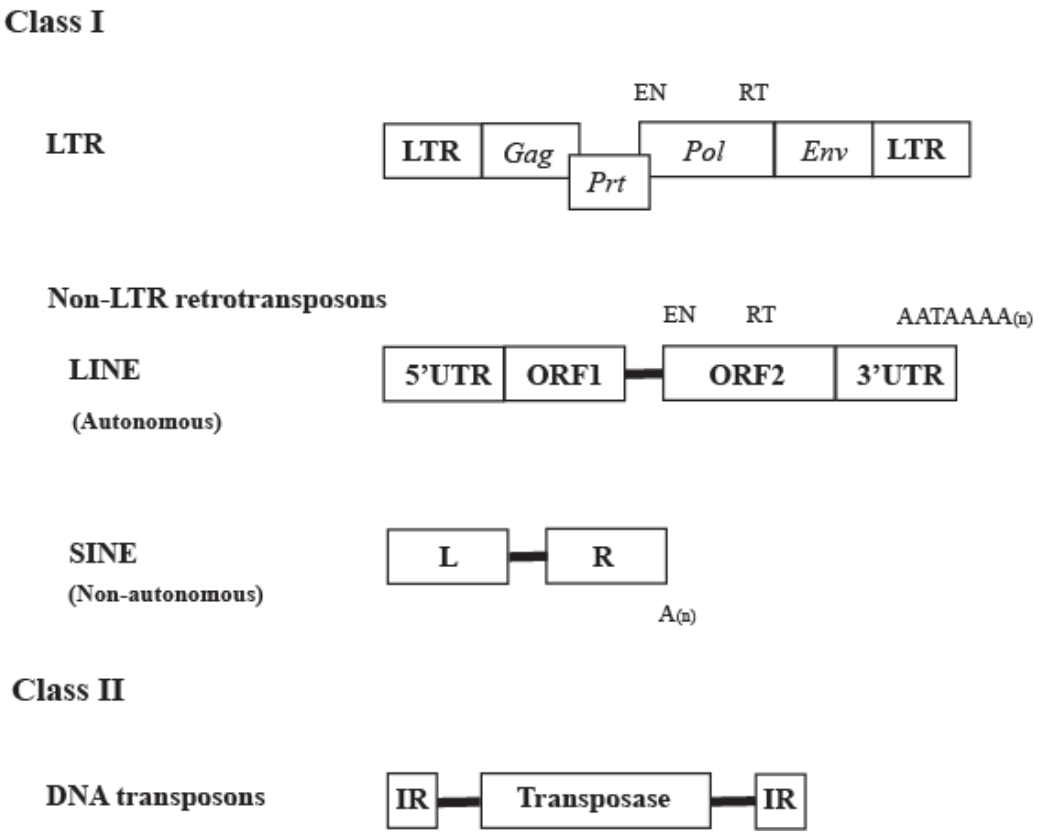


Figure 1.1 Classification of transposable elements.
Class I transposons are grouped into LTR retrotransposons and non-LTR retrotransposons. LTR retrotransposons have LTRs and slightly overlapping ORFs for group-specific antigen (*gag*), protease (*prt*), polymerase (*pol*) and envelope (*env*) genes. The reverse transcriptase (RT) and endonuclease (EN) domains are shown in the figure. Non-LTR retrotransposons are divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). LINEs are autonomous transposons and contain a 5'-untranslated region (5'UTR), an internal promoter, two ORFs, a 3'UTR and a poly(A) signal followed by a poly(A) tail (A_n). SINEs are non-autonomous elements containing two repeats, the left (L) and the right (R), and a polyA tail. Class II transposons are DNA transposons encoding a transposase, with inverted repeats at both ends.

TRANSPOSITION MECHANISMS

The consequence of a transposition event of a TE is either the duplication of the transposon or movement to a new site without duplication. During this process, the transposase is the soul of the transposition event, binding to the ends of the transposon and catalyzing the cleavage and joining reactions of transposition. Five major families of transposases have been reported based on their different catalytic mechanisms: DDE transposase, tyrosine (Y) transposase, serine (S) transposase, rolling circle (RC) or Y2 transposase and reverse transposase/endonuclease (RT/EN) (Curcio and Derbyshire, 2003). When they move out from their original site, some transposons copy, whereas others cut. Similarly when they insert into the target site, some transposons paste, whereas others copy.

Members of DDE transposase family include transposases of phage Mu, Tn*7*, and IS*630* family, retroviral integrase (IN), transposase of P elements, as well as RAG1 of the V(D)J recombination system. The signature of DDE transposases is a conserved DDE motif - aspartate, aspartate, 35 amino acid residues and glutamic acid – which is the active site that catalyzes transposition (Hiom et al., 1998; Craigie, 2001). Although there is little or even no similarity between the primary sequences of DDE transposases, the structural topology of their catalytic domains is similar to a conserved RNase H-like fold (Hickman et al., 2010). In spite of the utilization of the same cleavage and end-joining reaction, different DDE transposases exploit different ways for movement: ‘copy-in’, ‘cut-out, paste-in’ and ‘copy-out, paste-in’ (Figure 1.2 A, B).

For the ‘copy-in’ mechanism (Figure 1.2 A), a branched strand-transfer intermediate forms when the initially released 3’OH ends or transferred strands attack the target DNA. Host

replication resolves the intermediate generating a co-integrate molecule with two copies of the transposon. Mu and Tn*3* family are examples of this mechanism (Craig, 2002). For the 'cut-out, paste in' mechanism, both DNA strands are cleaved with formation of a terminal hairpin intermediate, which subsequently opens to release free 3'-OH ends which attack the target DNA. IS*10*/*50* and V(D)J recombination are examples of this mechanism (Jones and Gellert, 2004; Zhou et al., 2004). For the 'copy-out, paste-in', only one 3'-OH at one end of the transposon is generated by the initial asymmetric cleavage of the donor DNA. This 3'-OH attacks the same DNA strand at the opposite end forming a circular DNA intermediate. Host replication then resolves this intermediate into a new copy of the transposon. The newly formed circular transposon is nicked at repeated ends and the exposed 3'-OH at each terminus then joins to the target DNA. IS*2*, IS*3* and IS*911* families are examples of this mechanism (Loot et al., 2002).

Tyrosine or Serine transposases use a 'cut-out, paste-in' mechanism (Figure 1.2 F, G). These two families are related to either site-specific tyrosine or serine recombinases. They generate similar intermediates and products of transposition, in spite of different chemical mechanisms. The transposon is excised from its previous site, forming a circular DNA intermediate, with the original 'empty site' restored (Curcio and Derbyshire, 2003). After target capture, the transposon inserts into a new site by a reversal of the excision steps. These transposons are found almost exclusively in bacteria and are called conjugative transposons (Burrus et al., 2002).

The rolling circle or RC transposons utilize a 'copy-in' pathway (Figure 1.2 H). All the 'copy' pathways require a replicative step which is provided by the host DNA replication machinery during transposition. The RC transposons copy their DNA directly into the target by DNA replication such that each copy (new and old) of the transposon contains one newly

synthesized strand. A transposon-encoded protein nicks at the ends of the transposon and in the target DNA, resulting in free 3'OH ends to prime DNA synthesis by the host replication machinery using the transposon strand as a template (Curcio and Derbyshire, 2003).

Retrotransposons reverse transcribe their RNA genome into of a copy of cDNA. Afterwards, they use a different mechanism to transpose into the host genome. LTR or DDE-retrotransposons (such as Ty1) and retroviruses (such as HIV) utilize a 'copy-out, paste-in' pathway mechanistically similar to some of the DDE-transposons (Figure 1.2 C) (Haren et al., 1999). The DIRS1 family encoding Y-transposase, reverse transcribes into excised circular cDNA copies, followed by insertion of the circular intermediate into the host genome. This 'copy-out, paste-in' process is similar to the 'cut-out, paste-in' mechanism of the Y-transposons (Figure 1.2 E) (Duncan et al., 2002). Another family of retrotransposons is the TP (target-primed) retrotransposons encoding an endonuclease in addition to reverse transcriptase (RT/EN). The endonuclease nicks the target site which then serves as a primer for genome. The RNA intermediate of TP-retrotransposons is an integral component of the integration process (Figure 1.2 D) (Cost et al., 2002). Therefore, the TP-retrotransposons utilize a 'copy-out, copy-in' mechanism unlike the DDE-retrotransposons or the Y-retrotransposons.

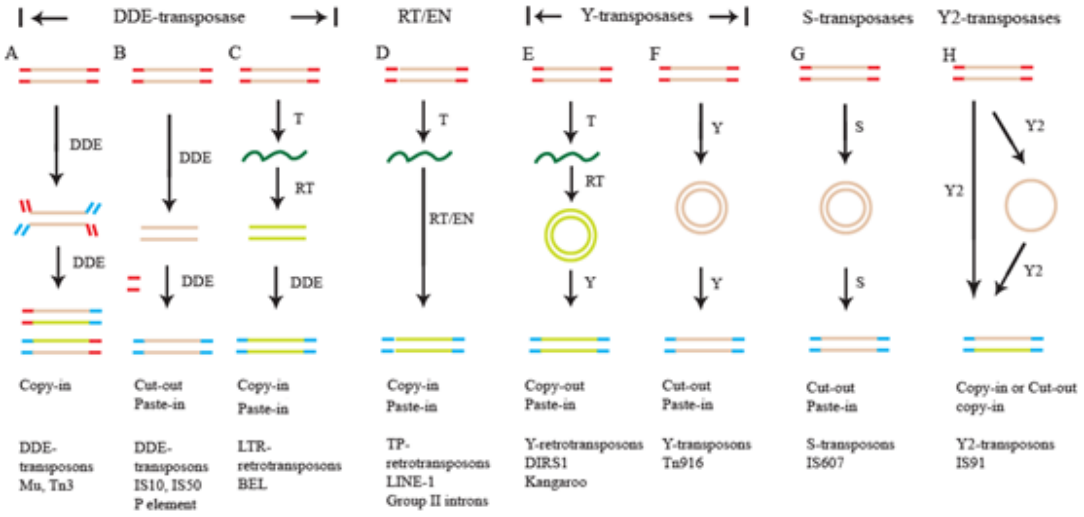


Figure 1.2 A variety of transposition mechanisms.

Modified from Curio and Derbyshire (2003). Transposition mechanisms mediated by five different transposase families are shown: DDE-transposase, reverse transcriptase/endonulcease (RT/EN), tyrosine (Y)-transposases, serine (S)-transposases and rolling-cycle (RC) or Y2-transposase. See text for details.

T: Transcription; RT: reverse transcription; EN: endonuclease

Bacteriophage Mu

Bacteriophage Mu is a transposable phage that infects enteric bacteria such as *E. coli*. Mu earns its name 'mutator' because of its ability to insert into essentially random sites in the chromosome and generate mutations at a high frequency (Taylor, 1963; Manna et al., 2004; Ge et al., 2011). Mu is a temperate phage. It can either reside on the *E. coli* genome as a prophage or enter into lytic cycle during which it transposes to multiple sites on the host chromosome by replicative transposition, amplifying its genome in the process. The Mu genome is 37 kb in length (Morgan et al., 2002). During packaging of the genome into phage particles, 50-150 bp of host DNA at the left end and 0.5-3 kb at the right end of the insertion sites of Mu are also packaged (Figure 1.3) (Bukhari et al., 1976). Because of its high efficiency of transposition and mutagenesis, as well as its similarity to other transposition systems, Mu has been an ideal model system to study TEs and DNA rearrangements. Indeed, the first *in vitro* transposition reaction was established for Mu (Mizuuchi, 1983), and subsequently replicated for other systems. The mechanism of integration of HIV and other retroviruses was found to closely resemble that of Mu (Engelman et al., 1991; Mizuuchi, 1992), including various DNA rearrangement mechanisms such as immunoglobulin V(D)J recombination (van Gent et al., 1996) (Figure 1.2).

LIFE CYCLE

As a temperate phage, Mu has two lifestyle choices to make: lysogenic or lytic growth (Figure 1.4). Upon phage infection, the linear Mu DNA is injected into the host cytoplasm together with phage protein MuN (Harshey and Bukhari, 1983; Puspurs et al., 1983). Unlike

other temperate phages such as λ or P1, which form covalently closed circular intermediates upon infection, the linear infecting Mu genome circularizes non-covalently through MuN binding to the DNA ends. This DNA then transposes or integrates into a random site on the host genome through a non-replicative or repair mechanism. In this pathway, the initial nick-join steps of transposition are the same as those established for replicative transposition, but the host DNA or flaps attached to either side of the infecting Mu genome are removed by an endonuclease activity in the C-terminal domain of the transposase MuA (MuA_{Nuc}) (Au et al., 2006; Choi and Harshey, 2010). The 5 bp gaps generated during integration are repaired by the host replication restart machinery and homologous recombination proteins (Figure 1.5) (Jang et al., 2012). Why the MuA_{Nuc} is activated only during the infection event is not yet known.

After the initial integration, only 1-10% of the infecting phage genomes become prophages, whereas the majority proceed to lytic growth (Howe and Bade, 1975). During the lytic cycle, over one hundred copies of Mu are made through replicative transposition (Chaconas et al., 1981). The strand transfer products (‘Shapiro intermediates’) of transposition are resolved by target-primed replication into a ‘mother’ copy in the original chromosomal location and a ‘daughter’ copy in a new location. In the late phase of the lytic cycle, phage Mu makes several morphogenesis proteins that package each Mu genome into viral particle. With the lysis of the cell, phage particles are released.

The lytic-lysogenic decision is controlled by the phage encoded repressor Rep, product of the *c* gene (Figure 1.3) (van de Putte et al., 1981). Binding of Rep to the operator region (which serves as a transposition enhancer in the absence of Rep) blocks the transcription of the early lytic promoter (Pe) and regulates the lysogenic promoter (P*c*), thus maintaining lysogeny. Induction of Mu prophage cannot be triggered by any known chemical or physical treatment, and

spontaneous induction is rare. Various Rep mutants have been developed to efficiently induce a Mu prophage. Mu*c*ts mutants have mutations in the N-terminal DNA-binding domain of Rep, which lower the binding affinity of Rep for the operator at higher temperatures and are therefore inducible at 42 ˚C (Lamrani et al., 1999).

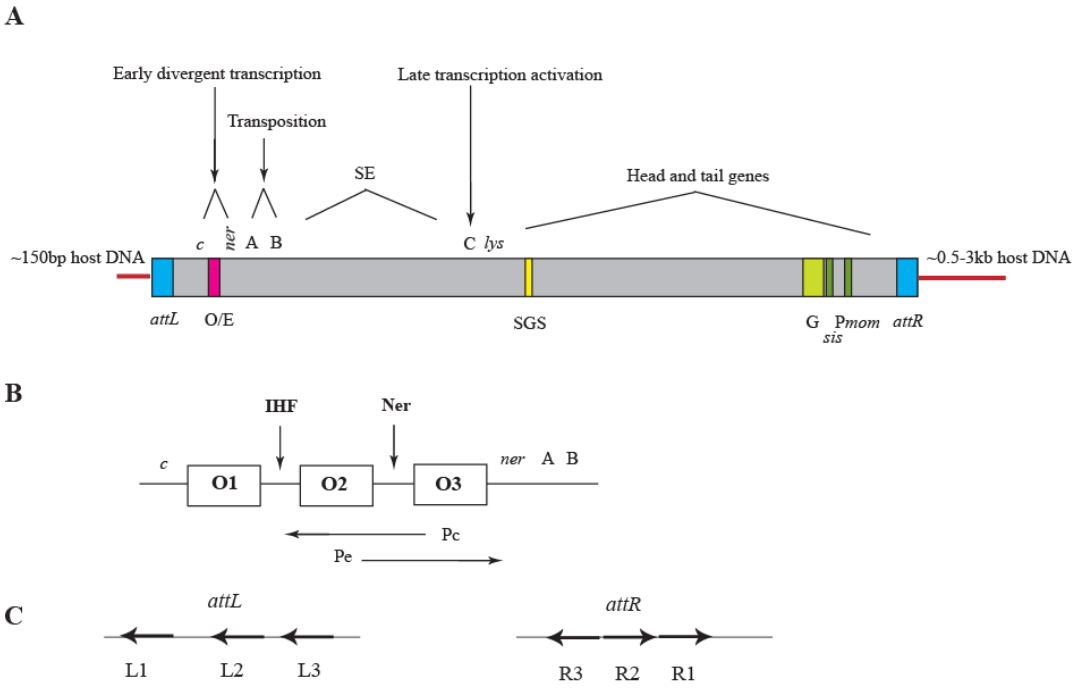


Figure 1.3 Mu genome organization.

(A) Schematic of a packaged Mu genome with co-packaged flanking host DNA. Major genes and DNA sites for phage development are shown. O/E, operator/enhancer; SE, semi-essential region; SGS, strong gyrase binding site; G, invertible G segment flanked by Fis-binding site *sis*. Pmom is a promoter for the *mom* gene, which has multiple Fis-binding sites.

(B) Enhancer region indicating IHF and Ner binding sites with two divergent promoters, Pe and Pc. The O sites bind Rep during lysogeny and MuA during lytic transposition.

(C) Mu ends *attL* and *attR* with arrows indicating the relative orientation of MuA transposase binding L and R subsites.

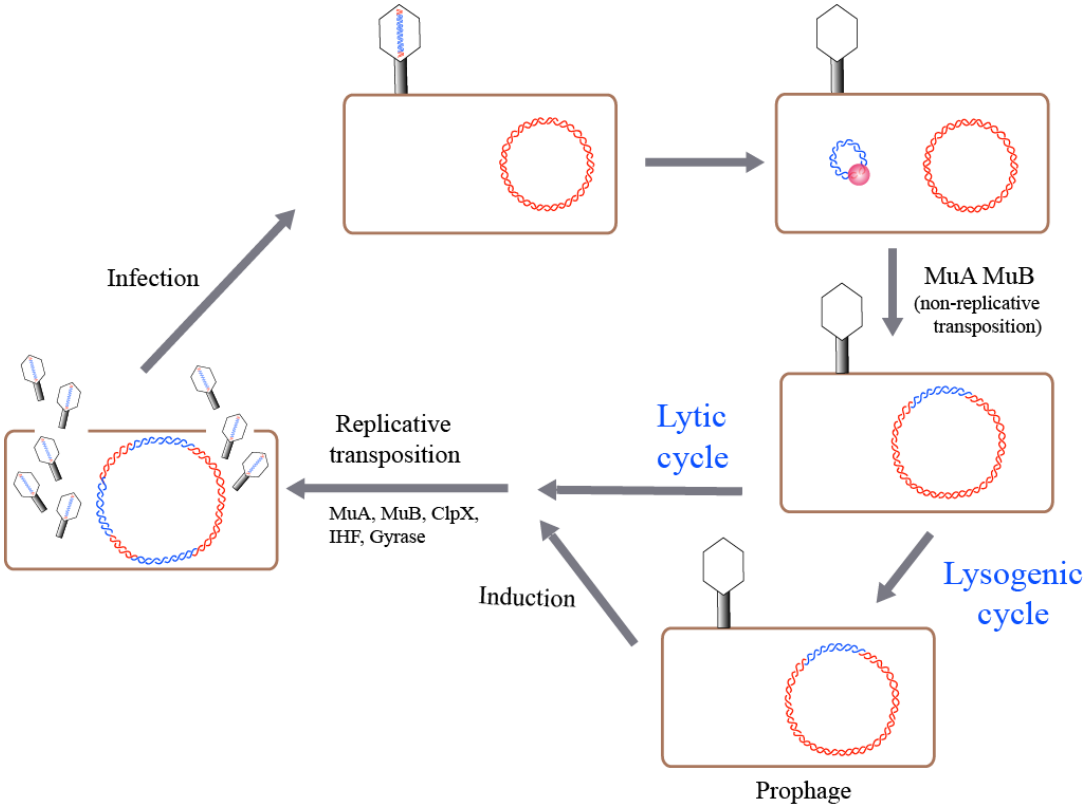


Figure 1.4 Mu life cycle
Lysogenic and lytic pathways as described in the text. Blue, Mu DNA; red, host DNA.

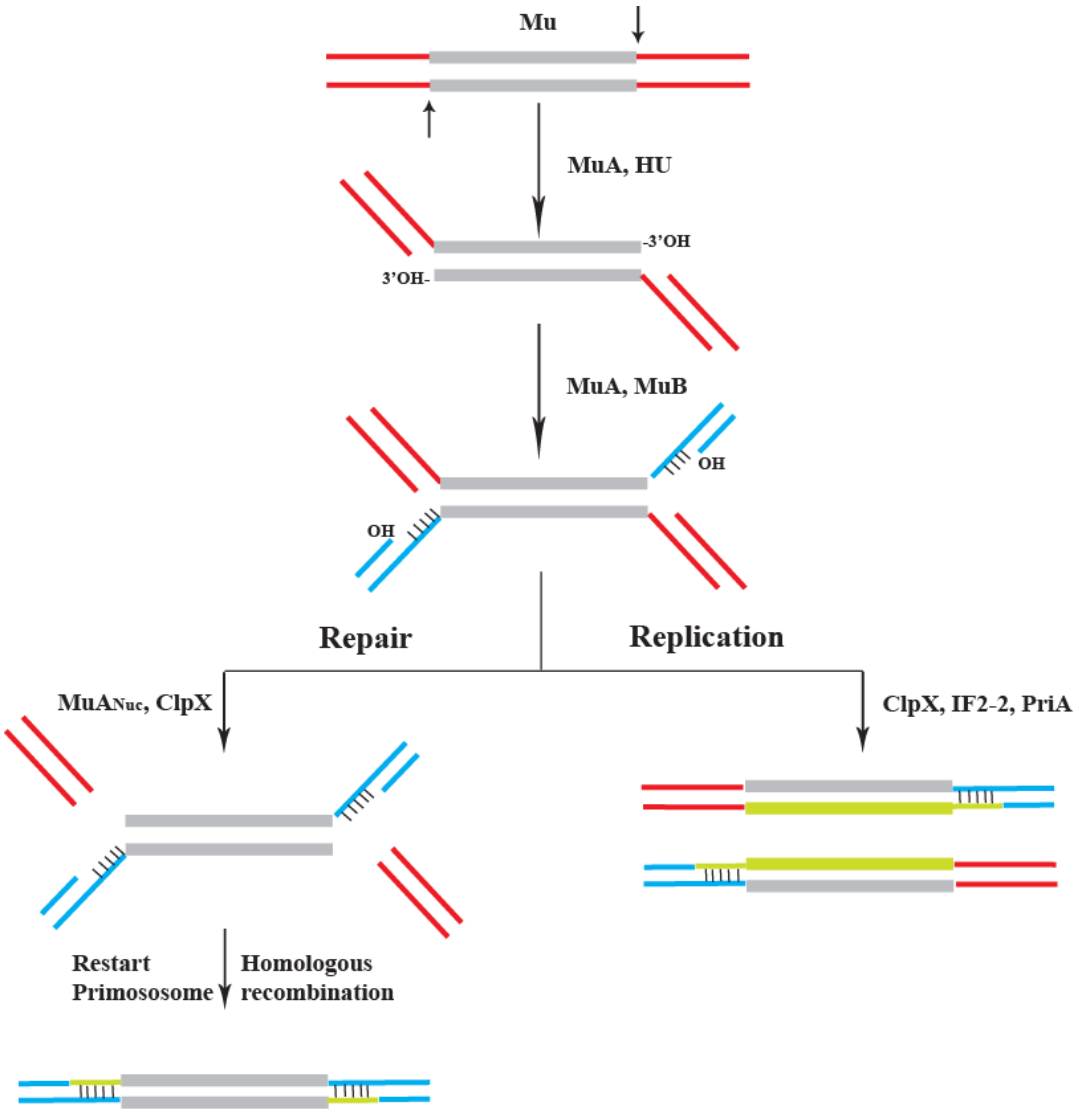


Figure 1.5 Steps in replicative and non-replicative (repair) pathways of Mu transposition. The transposase MuA, with the help of *E. coli* protein HU, makes single-stranded cleavages at the 3' ends of Mu. In the presence of MuB, the 3'OH groups at the cleaved ends are transfered to target DNA. The branched strand transfer intermediate is processed in two alternate pathways. During integration, the infecting Mu genome is separate from the chromosome, so the transposition is intermolecular. The host flap DNA (red line) of the branched strand transfer intermediate is removed by MuA$_{Nuc}$ in the presence of ClpX. The remaining gaps are repaired by host Restart Primososome (PriA) and Homologous Recombination proteins (RecABCD). During the lytic cycle, both Mu and the target DNA are in the chromosome, so transposition is intramolecular. The free 3'OHs in the strand transfer intermediate are used as primers to replicate Mu (pink line). ClpX and IF2-2 help to dissemble the transpososome and then the PriA Restart Primososome assembles on the Mu ends for replication.

MU TRANSPOSITION REQUIREMENTS

Cis-elements

Mu ends

Mu, like all other TEs, has a set of inverted repeats, *att*L and *att*R (also referred to as L and R ends), at the left and right ends of the Mu genome (Craigie et al., 1984; Groenen and van de Putte, 1986; Zou et al., 1991). These repeats are not as simple as those in other transposition systems. Each *att* end consists of three subsites, L1, L2, L3 and R1, R2, R3, respectively (Figure 1.3). The arrangement of the three sites is not the same at the two ends. L1, L2 and L3 are all in the same direction, with an 80bp spacer between L1 and L2 and a 25bp spacer between L2 and L3. The three R sites abut each other, with the R3 site in an opposite orientation with respect to R1 and R2 (Kuo et al., 1991; Lavoie et al., 1991).

Enhancer

Located about 1kb from the left end, the Mu operator serves as a unique transposition enhancer, found only in Mu and its relative phage D108. This site is capable of enhancing transposition efficiency by at least one hundred fold *in vivo* (Leung et al., 1989; Mizuuchi and Mizuuchi, 1989; Surette et al., 1989). Transposition does not require the enhancer for the chemical steps (Mizuuchi et al., 1992; Surette and Chaconas, 1992), but rather for assembly of the transpososome in a distance-independent but orientation-dependent manner. The enhancer consists of three subsites O1, O2 and O3. There is a binding site of *E. coli* integration host factor (IHF) between O1 and O2 (Higgins et al., 1989). Binary interactions between the various

enhancer and L and R subsites result in an ordered interaction first between E and R to form an ER complex, and then between ER and L to form the LER complex, which traps five DNA supercoils within it (Yin et al., 2007). This 3-site interaction is essential for maturation of the LER into a stable transpososome within which catalysis occurs (Mizuuchi et al., 1992; Namgoong et al., 1994). The enhancer remains associated within the transpososome from the establishment of ER complex until completion of the strand transfer reaction (Pathania et al., 2003).

DNA supercoiling

Under standard reaction conditions, a negatively supercoiled Mu donor substrate is strictly required *in vitro* (Mizuuchi, 1983). Although the target DNA can be linear or circular, supercoiled DNA is used with the highest efficiency. Supercoiling has both conformational and torsional effects on the transposition reaction (Chaconas and Harshey, 2002).

The conformational effects of DNA supercoiling can increase the local concentration of DNA sites dramatically, thus kinetically stimulating a reaction involving binding and interaction of multiple proteins and DNA sites. The binding affinity of MuA to the Mu ends is improved by negative supercoiling (Kuo et al., 1991). The torsional effects of DNA supercoiling are at both the enhancer and the Mu ends. At the enhancer, DNA bending is favored, because IHF binding and bending between O1 and O2 can reduce the requirement of the supercoiling level for optimal reaction *in vitro* (Surette and Chaconas, 1989). At the Mu ends, torsional effects are exhibited at two places, DNA bending between L1 and L2 that facilitates stable binding of HU (Kobryn et al., 1999) as well as DNA bending and helix destabilization at the Mu host junction (Wang et al., 1996).

The most critical role for supercoiling in the donor is for promoting transpososome assembly (Surette et al., 1987). Difference-topology studies of enhancer-independent Mu transposition indicate that not only the three-site interactions but also the two-site interactions are subjected to strict topological restrictions (Yin and Harshey, 2005). During synapsis, negative supercoiling in the donor DNA would favor the alignment of the two Mu ends in their correct orientation but not those in the wrong orientation (Gellert and Nash, 1987). Surprisingly, the supercoiling requirement in the donor can be by-passed under alternate solvent conditions (Me_2SO) (Savilahti et al., 1995).

Supercoiling in the target likely also favors a bent DNA conformation, as deduced by efficient targeting of Mu to linear substrates with DNA mismatches, which increase DNA flexibility (Yanagihara and Mizuuchi, 2002).

SGS (strong gyrase binding site)

The SGS is located right in the middle of the 37.2 kb Mu genome. It is required for Mu replication *in vivo* and its equidistant location from the Mu ends is essential for this requirement (Pato et al., 1990; Pato, 1994; Pato et al., 1995; Pato and Banerjee, 1996). Thus, it is believed that SGS promotes efficient synapsis of the Mu ends during transposition. DNA gyrase binds strongly to SGS and nicks the site *in vitro* (Pato et al., 1990). SGS is the strongest gyrase binding site known. The replacement of Mu SGS with SGS of phage D108, a close relative of Mu, can still promote efficient Mu replication, while replacements with gyrase binding sites on plasmids pSC101 or pBR322 cannot (Pato and Banerjee, 1999). The gyrase-binding function of SGS has been proposed to localize Mu DNA into a single loop of plectonemically supercoiled DNA. The SGS is not required for *in vitro* transposition, probably because the mini-Mu plasmid substrates

used in *in vitro* studies have a relatively small distance between the two Mu ends.

Trans-elements

MuA

Mu encodes its own transposase MuA, which is essential for transposition. The active form of MuA is a tetramer, which catalyzes the cleavage and joining steps of transposition (Lavoie et al., 1991). MuA is 663 amino acids long, and can be divided into three major globular domains to which a variety of different functions have been mapped (Figure 1.6A) (Harshey et al., 1985; Naigamwalla and Chaconas, 1997). A crystal structure for the central catalytic domain is solved (Rice and Mizuuchi, 1995), while NMR structures for two N-terminal DNA-binding domains are available (Clubb et al., 1994; Clubb et al., 1997; Schumacher et al., 1997). The dual DNA-binding capacity of MuA (to *att* and enhancer sites) is critical for LER formation and assembly of the active transpososome. The endonuclease (Nuc) activity in domain III is essential for the repair pathway of Mu transposition. The C-terminal domain of MuA interacts with the molecular chaperone ClpX, an interaction necessary for disassembling the transpososome and transitioning to replication (Levchenko et al., 1997).

MuB

The MuB protein has multiple roles in transposition, the most prominent being its ability to stimulate intermolecular transposition over one hundred fold (Symonds et al., 1987). It does so by capturing target DNA and delivering it to the transpososome. MuB is a protein with 312 amino acids and can be divided into two domains, both of which have non-specific DNA-binding

activity (Figure 1.6A). MuB binds ATP within the N-terminal domain and (Teplow et al., 1988) has an ATPase activity which is stimulated by MuA and DNA (Maxwell et al., 1987; Adzuma and Mizuuchi, 1991). An NMR structure of the C-terminal domain is available (Hung et al., 2000). Through binding to the C-terminal domain of MuA, MuB also regulates the disassembly of transpososomes after strand transfer by competition with ClpX for MuA interaction (Levchenko et al., 1997).

MuB is also critical to the phenomenon of transposition immunity. *cis*-immunity is the name given to a phenomenon where targets containing Mu-end sequences are immune to new insertions of Mu. *In vitro* data suggests that this is likely due to end-bound MuA interacting with MuB bound near its vicinity and promoting dissociation of MuB via MuA-MuB interaction, which stimulates the ATPase activity of MuB (Mizuuchi et al., 1995). The MuB-free DNA is a poor target for new insertions (Adzuma and Mizuuchi, 1988; Adzuma and Mizuuchi, 1989; Greene and Mizuuchi, 2002). Mu genome-immunity is the name given to the phenomenon that prevents Mu from inserting into itself during replicative transposition. This is different from *cis*-immunity because MuB binds strongly throughout the Mu genome during Mu replication (Ge et al., 2010). Mu genome-immunity is likely due to a property of MuB which blocks Mu integration within MuB-bound regions, but promotes integration in its vicinity (Ge and Harshey, 2008). The strong binding of MuB within the Mu genome is puzzling because MuB prefers AT-rich DNA (Greene and Mizuuchi, 2004), but the Mu genome is AT-poor. It is possible that SGS-promoted supercoiling of Mu creates a unique Mu domain structure that promotes MuB binding (Ge et al., 2010).

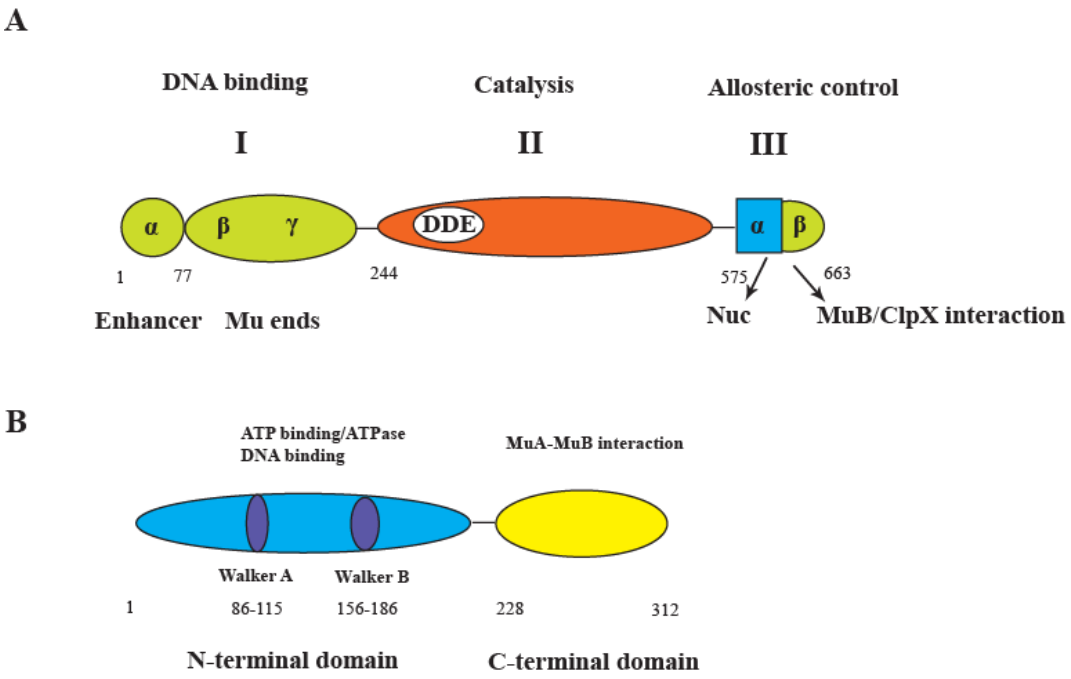


Figure 1.6 Domain structures of MuA and MuB .
(A) The N-terminal domain I of MuA has separate subdomains that bind to the enhancer element and Mu ends. (Leung et al., 1989; Mizuuchi and Mizuuchi, 1989). The central domain II has the catalytic DDE motif. The C-terminal domain III has nuclease activity (Nuc domain) in the α domain, which is masked in the full-length protein (Wu and Chaconas, 1995). Domain IIIβ has an overlapping binding site to which MuB and host ClpX protein compete to bind (Chaconas and Harshey, 2002). The numbers refer to amino acid residues.
(B) The N-terminal domain of MuB has both non-specific DNA binding and ATP binding (Miller et al., 1984). The C-terminal domain also has non-specific DNA binding activity and is thought to interact with MuA (Hung et al., 2000).

Host proteins

HU

E. coli HU (Heat Unstable) protein, a member of histone-like protein family, is a small heterodimer consisting of the 9.5kDa products of genes *hupA* and *hupB*. It binds DNA non-specifically and bends it sharply, with a higher affinity for supercoiled DNA and distorted regions (Swinger and Rice, 2007). HU-induced flexibility contributes to the formation of DNA loops, a property important in compacting the chromosome (Becker et al., 2007). On the Mu genome HU binds site-specifically within the 83 bp spacer region between L1 and L2 at the L end, introducing a 155˚ DNA bend (Lavoie and Chaconas, 1990). This binding is an essential step to the formation of LER complex, a precursor of the transpososome. HU is not required for the catalytic steps of Mu transposition (Craigie and Mizuuchi, 1987).

IHF

E. coli IHF (Integration Host Factor) protein is a member of the histone-like protein family as well. IHF is composed of two subunits, the products of genes *himA* (11kDa α-subunit) and *himD* (9.5kDa β-subunit), with about 25% homology between the subunits (Miller et al., 1984; Flamm and Weisberg, 1985; Rice et al., 1996). IHF has sequence similarity to HU but has a different interaction with DNA. Unlike HU, IHF has a consensus binding sequence. The binding of IHF bends DNA by 140˚ or into a U-turn within 30bp of DNA (Thompson and Landy, 1988; Swinger and Rice, 2004). It's ability to contort and alter DNA shape is important for a variety of DNA processes including site-specific recombination, transcription, replication, and compacting the *E. coli* nucleoid (Dillon and Dorman, 2010). On the Mu genome IHF binds between O1 and

O2 of the enhancer region (Figure 1.3). Through altering the architecture of the DNA, the binding of IHF not only reduces the supercoiling requirement of donor DNA during transposition but also influences early Mu transcription. IHF enhances transcription from Pe about three to five fold and decreases transcription from P*c* five to ten fold (Krause and Higgins, 1986). Mu cannot replicate in host strains deficient in either subunit of IHF protein; this effect can be partially alleviated by supplying MuA and MuB in *trans* (Berg and Howe, 1989).

IN VITRO MU TRANSPOSITION

An *in vitro* Mu transposition system was established in 1983 (Mizuuchi, 1983), which dissected this complicated process thoroughly. Not only were the chemical steps of transposition deduced using this system, but a series of high-order protein-DNA complexes called transpososomes were isolated, within which these reactions take place (Surette et al., 1987; Baker and Mizuuchi, 1992; Chaconas and Harshey, 2002) (Fig. 1.7). The recruitment of transposase MuA to multiples DNA binding sites at the Mu ends, and their ordered interaction, was seen to result in topologically unique transposition-competent complexes (Pathania et al., 2002; Pathania et al., 2003; Yin and Harshey, 2005; Yin et al., 2005; Yin et al., 2007).

The *in vitro* addition of Mu proteins A and B, host protein HU and divalent metal ions (usually Mg++) to a mini-Mu plasmid donor containing the L and R ends and the enhancer (E), recapitulates all the catalytic steps of transposition (see Figure 1.2). Using a difference-topology methodology that employed the Cre recombinase, pairs of *loxP* sites were positioned at strategic locations flanking L, E and R. Analysis of the recombination products established that a 2-site ER complex forms first, followed by joining of the L end to form a 3-site LER complex with

assistance from HU (Watson and Chaconas, 1996; Pathania et al., 2002; Pathania et al., 2003). This complex traps five DNA supercoils. Addition of supercoiled target DNA, MuB and ATP results in a strand transfer complex within which the enhancer remains associated (Pathania et al., 2003).

The LER complex is unstable, and is rapidly converted to the more stable Type 0 transpososome, in which MuA forms a catalytically active tetramer (Lavoie et al., 1991; Watson and Chaconas, 1996). L1 and R1-bound MuA monomers within the Type 0 complex catalyze the cleavage of 3' ends *in trans* (Namgoong and Harshey, 1998), giving rise the stable cleaved Type I complex. The cleavage relaxes the vector domain of the plasmid but the Mu DNA domain remains supercoiled (Surette et al., 1987). In the presence of MuB protein, ATP and the target DNA, the cleaved Mu ends are joined to the target DNA forming the most stable strand-transfer complex Type II. This complex is also referred to as the θ or Shapiro intermediate, which is resolved by the host replication machinery (Figure 1.7) (Nakai et al., 2001).

While DNA supercoiling is essential for generating the 5-noded LER complex where the L and R ends of Mu are synapsed, MuA can recreate a similar topology on nicked substrates in the presence of Me_2SO, conditions that obviate the need for supercoiling, suggesting that formation of the interwrapped architecture is an active process rather than a fortuitous trapping of the DNA crossings already present in supercoiled DNA (Yin et al., 2005). Plectonemic interaction of sites requires that they find each other by slithering rather than by random collision. Topological analysis of the Mu synapse under different enhancer-independent condition have shown that all three sites – L, E and R – impose a topological selectivity in establishing the MuA-mediated specific interwrapped synapse (Yin and Harshey, 2005).

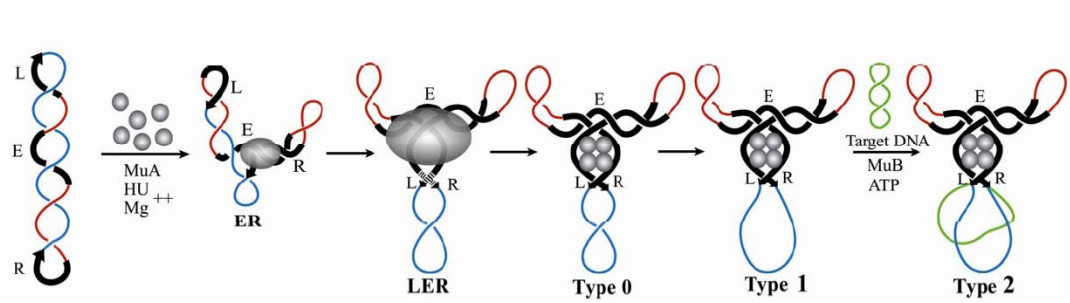


Figure 1.7 Mu transposition *in vitro*.
A series of nucleoprotein complexes and their DNA topology have been identified and characterized as described in the text. This figure is modified fromYin et al., 2005.

Phage λ

BACKGROUND

Phage Mu and phage λ are both temperate phages. Unlike Mu, however, phage λ can only integrate into a specific attachment site located between the *gal* and *bio* genes on the *E. coli* chromosome in the bacterial genome (Gottesman and Weisberg, 2004). The mechanisms of integration and replication of phage λ are very different from those of phage Mu (Azaro and Landy, 2002).

λ DNA in the phage head is linear, 48 kb in length, with 12 bp single-stranded cohesive ends. Upon infection, the DNA is converted into a circular molecule by base-pairing between cohesive ends (*cos* sites), which are then ligated by the host ligase (Friedman and Court, 2001). Following circularization, λ can either enter the lytic cycle or integrate into the host genome. Unlike Mu, however, integration is not an obligatory step for replication. During the lytic cycle, λ DNA first utilizes a θ replication mechanism resembling that of the bidirectional chromosome replication mechanism, which originates at a specific origin. After a few rounds of θ replication, λ switches to rolling-circle replication. Here, after one DNA strand is nicked, the free 3' end primes the synthesis of a new DNA strand that displaces the old strand. When replication around the circle is completed, the circle keeps rolling, generating long tandem repeats of individual λ DNA molecules called concatemers, which are then cut at the *cos* sites and packaged (Gottesman and Weisberg, 2004).

In the lysogenic pathway, integration of λ into the bacterial *attB* site requires the phage-encoded integrase (Int) and the attachment sites *attP* on the circular phage DNA, generating

hybrid sites *attL* and *attR* flanking the integrated λ genome after the site-specific recombination event (Azaro and Landy, 2002). Prophage λ is induced by conditions that damage DNA such as radiation or mitogens. Induction triggers the expression of several genes including the integrase (Int) and excisase (Xis), which promote excisive recombination between *attL* and *attR* sites. The excised the circular form of phage λ then starts replication and enters the lytic cycle (Friedman and Court, 2001).

DNA SUPERCOILING AND λ RECOMBINATION

The λ Int-dependent recombination between *attP* and *attB* sites requires the IHF protein, and a supercoiled *attP* DNA substrate (Azaro and Landy, 2002). The *attP* site is wrapped around the recombination proteins in a specific geometry as judged by the unique topological products of recombination (Nash and Pollock, 1983). This nucleosome-like complex is required to capture the *attB* site. This same study showed that in contrast to integrative recombination between *attP* and *attB*, supercoiling is not required for the excision recombination between prophage *attL* and *attR* sites. Synapsis of sites is simply dependent on random collision between the two recombining loci. Indeed, this observation has been used to probe chromosome structure in *Salmonella* using long range *attL-attR* promoted recombination (Garcia-Russell et al., 2004).

Bacterial Chromosomal Structure

Bacterial chromosomal DNA does not have an envelope-enclosed nucleus as the eukaryotic cell does. The chromosome of an *E. coli* cell is packed into a structure called a nucleoid, which occupies only a small fraction (~0.5μm^3) of the 2μm x 1μm cell. If completely unconstrained, the *E. coli* DNA would form a random coil with a length of 1.6mm and a volume approximately 1000-times the volume of the nucleoid (Dame, 2005). The circular bacterial chromosome can be marked by *oriC* where replication originates and Ter where replication terminates. Several levels of condensation mechanisms combine to compact the circular DNA, including basic forces such as macromolecular crowding and DNA supercoiling (de Vries, 2010). Nucleoid-associated proteins (NAPs) such as HU and IHF distribute uniformly in the entire nucleoid and condense the chromosome by DNA bending over short distances (Azam et al., 2000). SMC-like proteins can condense DNA with a span of 170 bp (Case et al., 2004). These NAPs likely interact to form higher-order ~10 kbp DNA loops that are dynamic (i.e. not fixed) as well as topologically independent (i.e. nicking one domain maintains integrity of neighboring domains (Sinden and Pettijohn, 1981)). Another level of organization has been called 'macro-domain' as defined by limited the mobility of marked DNA sites within these domains (Valens et al., 2004). Four such macro-domains have been deduced as described below, with Ori and Ter regions having a non-random location through the cell cycle (Gordon et al., 1997).

CHROMOSOMAL LOOPS/DOMAINS

In 1976, dramatic electron microscopy (EM) images of isolated *E. coli* nucleoids showed a

large number of DNA loops attached to a central structure (Kavenoff and Bowen, 1976; Kavenoff and Ryder, 1976). The average size of the loops was estimated to be ~100 kbp. As technology improved, a re-examination of the loops by both EM and other methods in *E. coli* and *Salmonella* reached similar conclusions regarding existence of the loops, but estimated their average size to be about 10 kb. The new methodology included a γδ resolvase-based recombination assay in *Salmonella*, which depends on the recombining sites being able to find each other by slithering, unhindered by topological barriers (Higgins et al., 1996; Stein et al., 2005). This study concluded that topological domains (loops) are variable in both size and the placement of barriers. Gyrase mutants increased the number of domains, indicating an important role of negative supercoiling in maintaining the looped domain structure. In *E. coli*, in addition to high-resolution EM, the methodology employed investigation of the activity of >300 supercoiling-sensitive reporter genes after nicking the chromosome at specific locations around the chromosome (Postow et al., 2004). This study also concluded that domains have a range of sizes with an average size of 10 kbp, that domain barriers are not located sequence-specifically, and that chromosomes are structurally fluid.

MACRO-DOMAINS

Rather than being folded randomly, compacted bacterial chromosome regions appear to be located at specific spatial positions. Fluorescence *in situ* hybridization (FISH) and fluorescence labeled DNA binding proteins that bind specifically to chromosomal regions revealed different localization of Ori and Ter domains during the cell cycle (Gordon et al., 1997). In the nucleoid of new-born cells, the localization of the Ori region was at the old-pole-proximal nucleoid border

whereas the Ter region was at the new-pole-proximal border. During replication and cell division, one copy of the duplicated *oriC* remained at the old-pole-proximal border, whereas the other one migrated to the opposite border in an early stage of the cell division cycle. Duplicated DNA Ter regions were separated at the last stage of chromosome separation and migrated from the nucleoid border to mid-cell (Niki and Hiraga, 1998).

Based on both cytological and genetic studies, a higher-order organization of the bacterial chromosome has been suggested, in which the whole genome is divided into 4 macro-domains (Valens et al., 2004). λ Int-mediated recombination between *attL* and *attR* sites placed at a variety of locations on the chromosome has shown a strong bias in recombination frequency between DNA sites in the different locations, indicative of a higher-order compaction/organization. FISH microscopy indicates that the local DNA dynamics differs in each of the macro-domains, which might be a consequence of distinct physical structures of these domains (Espeli et al., 2008; Valens et al., 2004). Regions in which fluorescent markers do not collide with each other have been defined as macro-domains. Four such macro-domains have been designated (Ori, Ter, Right and Left) together with two less structured regions flanking the Ori domain (Figure 1.8).

NAPs

A group of small histone-like proteins that can bind and bend DNA and help to compact and stabilize it are known as nucleoid-associated proteins (NAPs). Some NAPs may serve as boundary elements or domain barriers that mark the boundary of chromosomal domain loops (Hardy and Cozzarelli, 2005). These proteins also contribute to the control of global transcription

of the host genome (Travers and Muskhelishvili, 2005).

Fis

E. coli Fis (factor for inversion stimulation) protein is a 22kD homo-dimeric protein. It was first found to enhance the site-specific inversion recombination reaction (Hin system in *Salmonella* and Gin system in phage Mu) and thus got its name (Johnson et al., 1986). Fis binding results in DNA bending between 50˚ and 90˚. However when several Fis molecules bind to DNA in the same plane, the bending of DNA can be up to 160˚ (Travers and Muskhelishvili, 2005). Although Fis binds specifically to a 15bp degenerate consensus sequence it also binds non-specifically to DNA with relatively high affinity (Pan et al., 1996). Fis, like H-NS, also binds throughout the genome and is especially abundant at intergenic sites (Grainger et al., 2006). It has also been reported to form DNA-protein-DNA bridges (Schneider et al., 2001). The expression pattern of Fis is in keeping with the dynamic nature of looped domain boundaries, in that Fis is most abundant in the early exponential phase of growth, when the looped domains are most plentiful (Browning et al., 2010).

Fis enhances the Mu Gin-mediated inversion reaction by binding to a *sis* site outside the invertible G segment on the Mu genome (Figure 1.3). A recent study also showed that Fis has several binding sites in the promoter region of the *mom* gene (Figure 1.3), the product of which modifies Mu DNA to avoid digestion by the host restriction enzyme system (Karambelkar et al., 2012).

H-NS

The histone-like nucleoid-structuring protein (H-NS) is a 15.4kDa protein. Two DNA

binding domains in dimeric H-NS can interact with two DNA duplexes simultaneously (Luijsterburg et al., 2006). Chromatin immunoprecipitation-on chip (ChIP-on-chip) studies have shown that H-NS binds AT-rich DNA throughout the genomes of *E. coli* (Grainger et al., 2006). The distribution of these binding sites is consistent with the likely locations of domain loop boundaries. However, bacteria depleted of NAPs, including H-NS and Fis, can maintain their nucleoid structure (Zimmerman, 2006), perhaps indicating a redundant role of NAPs.

HU and IHF

These proteins play important roles in chromosome compaction, and have been described under 'Mu transposition requirements' in a previous section.

Knowledge of the NAPs and their contribution to the higher-order structure of the bacterial chromosome is growing rapidly and new models of the bacterial nucleoid are likely to soon emerge.

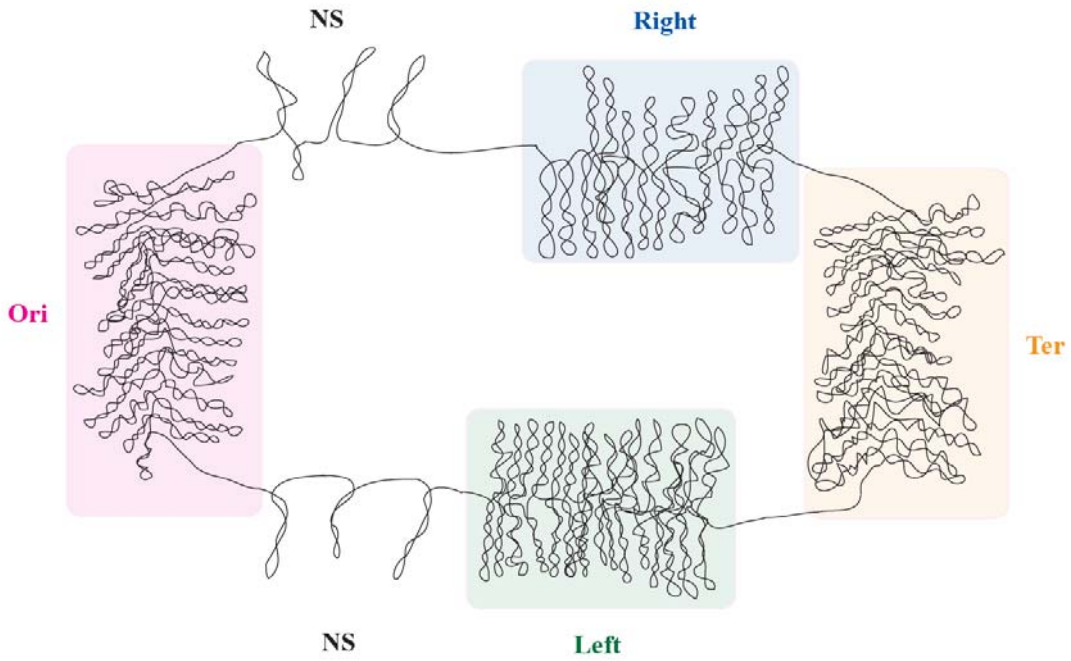


Figure 1.8 *E. coli* macro-domain structure.
Four macro-domains (Ori, Ter, Left, Right) together with two less structured region (NS) are shown on the figure.

CHAPTER 2: MATERIALS AND METHODS

STRAINS

All strains used in this work were derivatives of *E. coli* K-12 and are listed in Table 2.1. Gene disruptions and point mutations on the chromosome were made using the phage λ red-mediated homologous recombination methodology (Sawitzke et al., 2007). In strains with the temperature-inducible Muc*t*s prophage, all incubation steps were at 30 °C. All gene disruptions with the kanamycin gene *kan* replaced the start codon to the stop codon of the gene to be disrupted in the same direction by amplifying the flanking regions of the gene using primers with 50 nt homology extensions using pKD4 as the *kan* template, and selecting for kanamycin resistance (50μg/ml) (Datsenko and Wanner, 2000). The 1.6 kb *kan* cassette was left in the Mu*A* and Mu*B* deletion strains in order to maintain the symmetry of the SGS site. Other chromosomal mutations were introduced with a two-step procedure. First, the chromosomal DNA was replaced by a dual selection cassette *cat-sacB* amplified from strain SIMD30 (Datta et al., 2008), selecting for chloramphenicol (Cam) resistance (100μg/ml). Next, appropriate PCR products carrying the desired mutations were introduced, selecting for sucrose resistance on Luria broth (LB) plates supplemented with 6% sucrose, which eliminates the *cat-sacB* cassette (See Table 2.2 for position of each mutation). The gyrase mutant was made by moving the *gyrB402* ts allele flanked by a *cat-sacB* cassette into the experimental strain by P1 transduction, selecting for Cam^R. To make the strain with *loxP* sites flanking λ prophage, *loxP* sites were first inserted flanking the *attB* site in MG1655 and this strain was infected by cell lysate of phage λcI857 (Cam) and selected for chloramphenicol resistance. All constructs were confirmed by DNA sequencing.

Table 2.1 Strains

E. coli

Strains	Genotype	Source
AB1157	*thr-1, araC14, leuB6, lacY1, tsx-33, qsr'-0, glnV44, galK2, LAM-, Rac-0, hisG4, rfbC1, mgl-51, rpoS396, rpsL31(strR), kdgK51, xylA5, mtl-1, argE3, thi-1*	Pato (1996)
MP1999	AB1157, *recB, recC, sbcB*, *malF*::Mucts62	Pato (1996)
SIMD30	W3110 *lacDU169 galKTYR145UAG* (λc1857Δ*cro-bioA*) *[l int'*::*cat-sacB]*	Datta (2008)
ZL524	MP1999, *loxP* sites flanking Mu ends (61 bp and 70 bp outside *attL* and *attR,* respectively)	This study
ZL530	ZL524, Mucts *B*::*kan*	This study
ZL536	ZL524, Mucts *A*::*kan*	This study
ZL552	ZL524, Mucts Δ*attL*	This study
ZL556	ZL524, Mucts Δ*attR*	This study
ZL562	ZL524, Mucts ΔSGS	This study
ZL573	ZL562, Mucts *gp23* <SGS> *gp24*	This study
ZL 578	ZL562, Mucts *gp41*::SGS	This study
ZL580	AB1157, *malF*::*loxP*1	This study
ZL 582	ZL580, *malF*::*loxP*2 (2nd site next to *loxP*1)	This study
ZL592	ZL580, *jycF*::*loxP*	This study
ZL594	ZL580, *purH*::*loxP*	This study
ZL598	ZL582, *pspG*::SGS	This study
ZL604	ZL524, *himA*::*kan*	This study
ZL614	ZL524, *fis*::*kan*	This study
ZL624	ZL524, *hns*::*kan*	This study
ZL634	ZL524, *hupA*::*kan*	This study
ZL652	ZL524, 74 bp deletion of the spacer between *L1* and *L2* sites in Mucts *attL*	This study

ZL656	ZL524, Mucts Δ IHF site	This study
ZL660	ZL524, Mucts Δ*sis*	This study
ZL662	ZL660, Mucts Δ*attR*	This study
ZL670	ZL524, Mucts ΔP*mom*	This study
ZL672	ZL662, Mucts ΔP*mom*	This study
ZL704	ZL580, *lamB*::*loxP* (5kb downstream of *malF*)	This study
ZL706	ZL704, *malE*::SGS	This study
ZL708	ZL580, *ubiC*::*loxP* (9kb downstream of *malF*)	This study
ZL710	ZL708, *lamB*:: SGS	This study
ZL810	ZL580, *dinF*<*loxP*>*yjbJ* (15kb downstream of *malF*)	This study
ZL712	ZL580, *yjbS* < *loxP* > *aphA* (25 kb downstream of *malF*)	This study
ZL714	ZL712, *plsB*::SGS	This study
ZL732	MP1999, Mucts *E7*:: *loxP*, *gp49*::*loxP*	This study
MG1655	F^- λ^- *ilvG*$^-$ *rfb*-50 *rph*-1	Lab stock
ZL808	MG1655, λcI857Cm, *loxP* sites flanking λ (80 bp and 121 bp outside *attL* and *attR* respectively)	This study
ZL901	MG1655, *lacZ*::Muc+	This study
ZL911	ZL901, *loxP* sites flanking Muc+ (50 bp and 60 bp outside *attL* and *attR* respectively)	This study
ZL921	ZL911, Muc+ ΔSGS	This study
ZL931	ZL911, Δ*hns*	This study
EC1512	*argE thi ilv gyrB402*	Filutowicz (1983)
ZL940	EC1512, *yidX*::*cat-sacB*	This study
ZL941	ZL901, *yidX*::*cat-sacB gyrB402*	This study

< > intergenic region between two genes

:: DNA deletion/ addition

Table 2.2 Exact position of insertions and deletions in Mu and *E. coli*

Mu mutations*	**Position**
SGS site	17774-18000 (for addition of deletion)
Δ*sis*	35155-35215
Δ*attL*	1-180
Δ*attR*	36538-36717
Δ IHF site	947-958
Δ HU site	31-104
ΔP*mom*	35564-35779
gp23 <SGS> *gp24*	11414-11474 replaced by SGS
gp41::SGS	23987-24043 replaced by SGS
E7:: *loxP*	5111-5171 replaced by *loxP*
gp49::*loxP*	31507-31571 replaced by *loxP*

*The numbers indicate base-pairs starting at position 1 at the L end of the Mu genome

E. coli* mutations	**Position**
malF::*loxP* 1	4242073-4242233 replaced by *loxP*
malF::*loxP* 2	4241803-4241883 replaced by *loxP*
lamB::*loxP*	4246725-4246783 replaced by *loxP*
ubiC::*loxP*	4250927-4250976 replaced by *loxP*
dinF<*loxP*>*yjbJ*	4257145-4257194 replaced by *loxP*
yjbS < *loxP* > *aphA*	4267036-4267089 replaced by *loxP*
jycF::*loxP*	4279831-4280154 replaced by *loxP*
malE::SGS	SGS inserted into 4244289
lamB::SGS	4246725-4246783 replaced by SGS
plsB::SGS	4253635-4253674 replaced by SGS
pspG::SGS	4260863-4261094 replaced by SGS
purH:loxP	4204046-4204952 replaced by *loxP*

*The numbers indicate position of *E.coli* genome

PLASMIDS

All plasmids used in this study are listed in Table 2.3.

Table 2.3 Plasmids

Plasmid	Expressed protein	Resistance	Replication Origin	Induction	Source (ref.)
pKD4	Source for Kan cassette	Kanamycin	oriR6K gamma		Datsenko and Wanner (2000)
pKD46	Lamda Red recombinase	Ampicillin	repA101ts & oriR101	Arabinose	Datsenko and Wanner (2000)
pBAD24-his-Cre	Cre recombinase	Ampicillin	pACYC	Arabinose	Guzman et al. (1995)
pJG8	MuB	Kanamycin	p15A		Ge et al. (2011)
pUC19		Ampicillin			NEB

OLIGONUCLEOTIDES

PCR primers used in this dissertation are listed in Table 2.4.

Table 2.4 Oligonucleotides

Name	Sequence
Primers for making mutations	
attL loxP 1t	cgcggtgccatcgttcatttcaatcttgattttcaaagggttgtttaacctgtgacggaagatcacttcg
attL loxP 1b	tttaatcagcagcggcagcgtaatcttaaagaagttctggaacggacctgatcaaagggaaaactgtccatat
attL loxP 2t	cgcggtgccatcgttcatttcaatcttgattttcaaagggttgtttaaccataacttcgtataatgtatgctatacgaagttat
attL loxP 2b	tttaatcagcagcggcagcgtaatcttaaagaagttctggaacggacctgataacttcgtatagcatacattatacgaagttat
attR loxP 1t	tttaactttaacaacttcgtgctgattcaactgttaaccaacggcggccctgtgacggaagatcacttcg
attR loxP 1b	cgccttcaaaagcgatgcggtaggtgtagttaacaagcaggtcggtataaatcaaagggaaaactgtccatat
attR loxP 2t	tttaactttaacaacttcgtgctgattcaactgttaaccaacggcggcccataacttcgtataatgtatgctatacgaagttat
attR loxP 2b	cgccttcaaaagcgatgcggtaggtgtagttaacaagcaggtcggtataaataacttcgtatagcatacattatacgaagttat
malF loxP 1t	ttttaccagtccattaccgccgacggcaactggggtgatgaaaagctaagtgtgacggaagatcacttcg
malF loxP 1b	gttaaaaagacagtgatcagcgagaacaccacggtccagacgaaaatggcatcaaagggaaaactgtccatat
malF loxP 2t	ttttaccagtccattaccgccgacggcaactggggtgatgaaaagctaagataacttcgtataatgtatgctatacgaagttat
malF loxP 2b	gttaaaaagacagtgatcagcgagaacaccacggtccagacgaaaatggcataacttcgtatagcatacattatacgaagttat
5kR loxP 1t	atactcagagtgtcctgaagggctttaacaagtttgttgttcagtacgcttgtgacggaagatcacttcg
5kR loxP 1b	aaatgcaacgccagaaccctgcgacagccctttaccctgcgaggtcatcgatcaaagggaaaactgtccatat
5kR loxP 2t	atactcagagtgtcctgaagggctttaacaagtttgttgttcagtacgctataacttcgtataatgtatgctatacgaagttat
5kR loxP 2b	aaatgcaacgccagaaccctgcgacagccctttaccctgcgaggtcatcgataacttcgtatagcatacattatacgaagttat
9kR loxP 1t	tcacatcatcgacattaacccgggactttattgagataggccgtgatgcc tgtgacggaagatcacttcg
9kR loxP 1b	ttagtacaacggtgacgccggtaaaaacagttctgttagcaacagcggtt atcaaagggaaaactgtccatat
9kR loxP 2t	tcacatcatcgacattaacccgggactttattgagataggccgtgatgcc ataacttcgtataatgtatgctatacgaagttat
9kR loxP 2b	ttagtacaacggtgacgccggtaaaaacagttctgttagcaacagcggtt ataacttcgtatagcatacattatacgaagttat
15kR loxP 1t	ctatctggcggcgtcactggcgcaatggtacctggtttgccgcaacgtga tgtgacggaagatcacttcg

15kR loxP 1b	gttaggttgtgaatgagagtacgttcacttttcttctgaacgtgagatta atcaaagggaaaactgtccatat
15kR loxP 2t	ctatctggcggcgtcactggcgcaatggtacctggtttgccgcaacgtga ataacttcgtataatgtatgctatacgaagttat
15kR loxP 2b	gttaggttgtgaatgagagtacgttcacttttcttctgaacgtgagatta ataacttcgtatagcatacattatacgaagttat
25kR loxP 1t	tcaacaggtaaagatgtcgttttgttgctattcacatatgatatattcat tgtgacggaagatcacttcg
25kR loxP 1b	acagcattacaaaaacaatcgaagtttataaagatgatttctgattgacc atcaaagggaaaactgtccatat
25kR loxP 2t	tcaacaggtaaagatgtcgttttgttgctattcacatatgatatattcat ataacttcgtataatgtatgctatacgaagttat
25kR loxP 2b	acagcattacaaaaacaatcgaagtttataaagatgatttctgattgacc ataacttcgtatagcatacattatacgaagttat
37kR loxP 1t	cgcggttaacaaatacgttcctcgaaatggttatatgtaccgactctaattgtgacggaagatcacttcg
37kR loxP 1b	cagtaatcttcatagcgactgcatgaagattatcctcagtttgtgacatatcaaagggaaaactgtccatat
37kR loxP 2t	cgcggttaacaaatacgttcctcgaaatggttatatgtaccgactctaatataacttcgtataatgtatgctatacgaagttat
37kR loxP 2b	cagtaatcttcatagcgactgcatgaagattatcctcagtttgtgacatataacttcgtatagcatacattatacgaagttat
2.5kR SGS 1t	ccagtttatccggatgctcaacggtgactttaattccggtatctttctcgtgtgacggaagatcacttcg
2.5kR SGS 1b	attaacggcgataaaggctataacggtctcgctgaagtcggtaagaaattatcaaagggaaaactgtccatat
2.5kR SGS 2t	ccagtttatccggatgctcaacggtgactttaattccggtatctttctcggcccggcaggcgttcctgcg
2.5kR SGS 2b	attaacggcgataaaggctataacggtctcgctgaagtcggtaagaaattttgcgggaggtgcgggaaaa
5kR SGS 2t	atactcagagtgtcctgaagggctttaacaagtttgttgttcagtacgctgcccggcaggcgttcctgcg
5kR SGS 2b	cgcagcatgtgaccgttgttgttgatattgtaggcaaatttttcgttatcttgcgggaggtgcgggaaaa
12kR SGS 1t	gcgagcagcttattaaacagatcctgacgagcaggcaaacgtgggcctactgtgacggaagatcacttcg
12kR SGS 1b	atggcggatgaacacggcacggataaaactatcgctcagaaactggcgcgatcaaagggaaaactgtccatat
12kR SGS 2t	gcgagcagcttattaaacagatcctgacgagcaggcaaacgtgggcctacgcccggcaggcgttcctgcg
12kR SGS 2b	atggcggatgaacacggcacggataaaactatcgctcagaaactggcgcgttgcgggaggtgcgggaaaa
18kR SGS 1t	ggcatgattcttgtaatgccagcaagagatttcatatttgggagagcatctgtgacggaagatcacttcg
18kR SGS 1b	ccctttctttttatagttcggctgtatgtagggtacagcacgatgaatctatcaaagggaaaactgtccatat
18kR SGS 2t	ggcatgattcttgtaatgccagcaagagatttcatatttgggagagcatcgcccggcaggcgttcctgcg
18kR SGS 2b	ccctttctttttatagttcggctgtatgtagggtacagcacgatgaatctttgcgggaggtgcgggaaaa
37kL loxP 1t	ttagcagaaaaaggtctgccggtaaccgaagtttccgattacaccggttttgtgacggaagatcacttcg
37kL loxP 1b	cgcaataatcacttcgacaaactgacgagaaatgatggcctgtgcggtttatcaaagggaaaactgtccatat
37kL loxP 2t	ttagcagaaaaaggtctgccggtaaccgaagtttccgattacaccggtttataacttcgtataatgtatgctatacgaagttat
37kL loxP 2b	cgcaataatcacttcgacaaactgacgagaaatgatggcctgtgcggtttataacttcgtatagcatacattatacgaagttat

del SGS 1t	gtggggcggtacacccgatatggcaccccgcccggcaggcgttcctgcgc tgtgacggaagatcacttcg
del SGS 1b	gctgctcatcatggtttgtgagcagtcagtttgcgggaggtgcgggaaaa atcaaagggaaaactgtccatat
del SGS 2t	acacccgatatggcaccccgcccggcaggcgttcctgcgcttttcccgca
del SGS 2b	tggtttgtgagcagtcagtttgcgggaggtgcgggaaaagcgcaggaacg
del sis 1t	atgtaagggtatcaacaaatgaccagaatacagacctgcaacgaaacgcttgtgacggaagatcacttcg
del sis 1b	cagtgtgtcacctttttgaaggcgctttaaagcgcgttttaatcccggtcatcaaagggaaaactgtccatat
del sis 2t	atcaacaaatgaccagaatacagacctgcaacgaaacgctgaccgggatt
del sis 2b	ccttttgaaggcgctttaaagcgcgttttaatcccggtcagcgtttcgtt
del pmom 1t	ggaatcccccgcaagcaggttgcattgatctacgatgtggccctgtcaac tgtgacggaagatcacttcg
del pmom 1b	cgcggacacctgatttcaatgtgatcaaaggaatccgccttaaataacag atcaaagggaaaactgtccatat
del pmom 2t	gcaagcaggttgcattgatctacgatgtggccctgtcaacctgttattta
del pmom 2b	caaaggaatccgccttaaataacaggttgacagggccacatcgtagatca
del IHF site 1t	agtattattcttttctataaagttacttttcaaaatttaaactccttatttgtgacggaagatcacttcg
del IHF site 1b	aaaagctttttgtaaagctgctttttggtaaattcctttgattactgattatcaaagggaaaactgtccatat
del IHF site 2t	agtattattcttttctataaagttacttttcaaaatttaaactccttattaatcagtaatcaaaggaatttaccaaaaag
del IHF site 2b	aaaagctttttgtaaagctgctttttggtaaattcctttgattactgattaataaggagtttaaattttgaaaagtaact
del attR 1t	caaggttcagccataccctaagtgatccccatgtaatgaataaaaagcag tgtgacggaagatcacttcg
del attR 1b	gggccgccgttggttaacagttgaatcagcacgaagttgttaaagttaaa atcaaagggaaaactgtccatat
del attR 2t	caaggttcagccataccctaagtgatccccatgtaatgaataaaaagcagtttaactttaacaacttcgt
del attR 2b	gggccgccgttggttaacagttgaatcagcacgaagttgttaaagttaaactgctttttattcattacat
del attL 1t	caggtccgttccagaacttctttaagattacgctgccgctgctgattaaa tgtgacggaagatcacttcg
del attL 1b	gttatcggttttgaacgttttttgaagctgttattgaaatgatttgcagtatcaaagggaaaactgtccatat
del attL 2t	caggtccgttccagaacttctttaagattacgctgccgctgctgattaaaactgcaaatcatttcaataa
del attL 2b	gttatcggttttgaacgttttttgaagctgttattgaaatgatttgcagttttaatcagcagcggcagcg
del HU site 1t	tgattaaaccgctgacgccg tgtattgattcacttgaagtacgaaaaaaatgtgacggaagatcacttcg
del HU site 1b	aacaaaagcaatttttttactatctttcgcgtttcattgattaacgactaaatcaaagggaaaactgtccatat
del HU site 2t	tgtattgattcacttgaagtacgaaaaaaattagtcgttaatcaatgaaa
del HU site 2b	aacaaaagcaatttttttactatctttcgcgtttcattgattaacgactaa
del MuA t	aaactaaaccggaagtaatctggcctagccgatatcaagcaggtgaataa gtgtaggctggagctgcttc

del MuA b	ggaaatattcatctgaataacctcgtaattaaattttgtttaaacgttaa catatgaatatcctcctta
del himA t	atggcgcttacaaaagctgaaatgtcagaatatctgtttgataagcttgtgtaggctggagctgcttc
del himA b	ttactcgtctttgggcgaagcgttttcgacccggctttttaacttctgcatatgaatatcctccta
del MuB t	aaagccatttaattaacgtttaaacaaaatttaattacgaggttattcag gtgtaggctggagctgcttc
del MuB b	gcatccgttgccatttttatatttcgggccatcataatttcttctccttc catatgaatatcctcctta
del fis t	ggcatacttcgaaaattttgcgtaaacagaaataaagagctgacagaact gtgtaggctggagctgcttc
del fis b	ttccccatgccgagtagcgccttttttaatcaagcatttagctaacctgaacatatgaatatcctcctta
del hupA t	ccgtcgcactcgatgcttagcaagcgataaacacattgtaaggataactt gtgtaggctggagctgcttc
del hupA b	aaaaggggtgaaaccacccccttcgttaaaactgttcactgccacgcaatc catatgaatatcctcctta
del hns t	tctattattacctcaacaaaccaccccaatataagtttgagattactaca gtgtaggctggagctgcttc
del hns b	aaaaaatcccgccgctggcgggattttaagcaagtgcaatctacaaaaga catatgaatatcctcctta
SGS L 1t	tacagaactgacggaagcaacgcccgttccgcccatgccacagcccctgatgtgacggaagatcacttcg
SGS L 1b	aaatttctttccttcggatctcgtgttcccggactgatgccttgtcgcggatcaaagggaaaactgtccatat
SGS L 2t	tacagaactgacggaagcaacgcccgttccgcccatgccacagcccctgaaggcataaaatcagccgcacagatttttttaaaacgcg ccacggga
SGS L 2b	aaatttctttccttcggatctcgtgttcccggactgatgccttgtcgcggacaccgttaaataccggtttaaaaatcccgtggcgcgttttaa aa
SGS R 1t	cggcatcgctgacacaagaacgggggcgagtggctgcggtgccggaatga tgtgacggaagatcacttcg
SGS R 1b	ttttagcctgctggcagtgccggaaaaagcgggacagcgggagagacaag atcaaagggaaaactgtccatat
SGS R 2t	cggcatcgctgacacaagaacgggggcgagtggctgcggtgccggaatga aggcataaaatcagccgcacagatttttttaaaacg cgccacggga
SGS R 2b	ttttagcctgctggcagtgccggaaaaagcgggacagcgggagagacaag acaccgttaaataccggtttaaaaatcccgtggcgcgttttaaaa
Mu L loxP 1t	aatgtcaaaatccgaaatgaaattcaggcattaatccgaattcaggatgtgacggaagatcacttcg
Mu L loxP 1b	cgcacagatgctgtaatggtcgaaagtaatcaggtaagttttttcaccatatcaaagggaaaactgtccatat
Mu L loxP 2t	aatgtcaaaatccgaaatgaaattcaggcattaatccgaattcaggaataacttcgtataatgtatgctatacgaagttat
Mu L loxP 2b	cgcacagatgctgtaatggtcgaaagtaatcaggtaagttttttcaccatataacttcgtatagcatacattatacgaagttat
Mu R loxP 1t	atgttttatattgataacgacagcggcgtaaccgtcatgccgcccgtatctgtgacggaagatcacttcg
Mu R loxP 1b	gcgtgtttaataactccgcctgcacaatattaaaccagtccatgccgggcatcaaagggaaaactgtccatat
Mu R loxP 2t	atgttttatattgataacgacagcggcgtaaccgtcatgccgcccgtatcataacttcgtataatgtatgctatacgaagttat
Mu R loxP 2b	gcgtg tttaa taact ccgcc tgcacaatattaaaccagtccatgccgggc ataacttcgtatagcatacattatacgaagttat
λ L loxP 1t	gggatgcaaaatagtgttgagcatcgaaattctgcgcttcttttgccgactgtgacggaagatcacttcg
λ L loxP 1b	tagcccggagcaacctgtgaacacattttcagtttcccgtctggcgctggatcaaagggaaaactgtccatat

λ L loxP 2t	gggatgcaaaatagtgttgagcatcgaaattctgcgcttcttttgccgacataacttcgtataatgtatgctatacgaagttat
λ L loxP 2b	tagcccggagcaacctgtgaacacattttcagtttcccgtctggcgctggataacttcgtatagcatacattatacgaagttat
λ R loxP 1t	catctggcagagtgattaactaaacatcgcagtaatcgaggcgcttgccatgtgacggaagatcacttcg
λ R loxP 1t	tgatacccgcgtattaccgcaagggtttggctctggtctggtagcaatgcatcaaagggaaaactgtccatat
λ R loxP 1t	catctggcagagtgattaactaaacatcgcagtaatcgaggcgcttgccaataacttcgtataatgtatgctatacgaagttat
λ R loxP 1t	tgatacccgcgtattaccgcaagggtttggctctggtctggtagcaatgcataacttcgtatagcatacattatacgaagttat
lacZ L loxP 1t	cagcgcagcaccatcaccgcgaggcggttttctccggcgcgtaaaaatgctgtgacggaagatcacttcg
lacZ L loxP 1b	atccgacgggttgttactcgctcacatttaatgttgatgaaagctggctaatcaaagggaaaactgtccatat
lacZ L loxP 2t	cagcgcagcaccatcaccgcgaggcggttttctccggcgcgtaaaaatgcataacttcgtataatgtatgctatacgaagttat
lacZ L loxP 2b	atccgacgggttgttactcgctcacatttaatgttgatgaaagctggctaataacttcgtatagcatacattatacgaagttat
lacZ R loxP 1t	gcgcatcgtaaccgtgcatctgccagtttgaggggacgacgacagtatcg tgtgacggaagatcacttcg
lacZ R loxP 1b	gactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatcc atcaaagggaaaactgtccatat
lacZ R loxP 2t	gcgcatcgtaaccgtgcatctgccagtttgaggggacgacgacagtatcgataacttcgtataatgtatgctatacgaagttat
lacZ R loxP 2b	gactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccataacttcgtatagcatacattatacgaagttat
yidx catsac t	gtagtaccagcgtgatgacgttcgcgtttgccgtgcgtgtaatgtagtactgtgacggaagatcacttcg
yidx catsac b	cagatttctgcccgacgcatgtcatttttttatgcattgttctttttgtgatcaaagggaaaactgtccatat

Primers for real-time PCR

RT attL loxP t	ctacgcggtgccatcgtt
RT attL loxP b	acggcgtcagcggtttaa
RT LR loxP t	tcgcgtcctgctgattctg
RT LR loxP b	aaagcgatgcggtaggtgtag
RT malF b	gttaaaaagacagtgatcag
RT 5kR b	aaatgcaacgccagaaccct
RT 9KR b	ttagtacaacggtgacgccg
RT 15KR b	gttaggttgtgaatgagagt
RT 25KR b	tacagcagataaatgttcta
RT 37KR b	agcgactgcatgaagattat
RT 37kL t	aaggtctgccggtaaccgaa

RT lacZ L t	cagcgcagcaccatcaccgc
RT lacZ L b	atccgacgggttgttactcg
RT lacZ R b	gactgggaaaaccctggcg
RT λ L t	gggatgcaaaatagtgttga
RT λ L b	tagcccggagcaacctgtga
RT λ R b	tgatacccgcgtattaccgc

Primers for primer efficiency test

pUC19 t	tgcgcaacgttgttgccatt
pUC19 b	aacactgcggccaacttact
PE attL loxP t	ctacgcggtgccatcgtt tgcgcaacgttgttgccatt
PE attL loxP b	acggcgtcagcggtttaa aacactgcggccaacttact
PE LR loxP t	tcgcgtcctgctgattctg tgcgcaacgttgttgccatt
PE LR loxP b	aaagcgatgcggtaggtgtag aacactgcggccaacttact
PE malF b	gttaaaaagacagtgatcag aacactgcggccaacttact
PE 5kR b	aaatgcaacgccagaaccct aacactgcggccaacttact
PE 9KR b	ttagtacaacggtgacgccg aacactgcggccaacttact
PE 15KR b	gttaggttgtgaatgagagt aacactgcggccaacttact
PE 25KR b	tacagcagataaatgttcta aacactgcggccaacttact
PE 37KR b	agcgactgcatgaagattat aacactgcggccaacttact
PE 37kL t	aaggtctgccggtaaccgaa tgcgcaacgttgttgccatt
PE λ L t	gggatgcaaaatagtgttga tgcgcaacgttgttgccatt
PE λ L b	tagcccggagcaacctgtga aacactgcggccaacttact
PE λ R b	tgatacccgcgtattaccgc aacactgcggccaacttact

IN VIVO CRE RECOMBINATION

Overnight cell culture in M9 minimal media was diluted 1:100 into 20 ml of minimal media. The cells were then incubated in a 30°C shaker for 4-5 hrs until OD_{600} reached 0.5-0.6. They were washed three times with ice cold 10% glycerol and brought to the final volume of 200ul. 40µl of the cells were electroporated (Biorad Gene pulser™, 1.8kV, 1mm cuvette) with 90ng of the Cre plasmid. After recovery for an hour, cells were transferred to minimal media with ampicillin (100µg/ml) and grown for 40 hrs to allow maximal Cre-*loxP* recombination. Chromosomal DNA was purified using the Wizard Genomic DNA purification kit from Promega. Cre recombination products were assayed by real-time PCR as described below.

Cre recombination assays in the gyrase mutant were carried out differently. Wild type and gyrase mutant strains were grown at 37°C in LB. The transformation steps were the same as described above. After transformation, cells were only grown for 12 hrs in 37°C to avoid *loxP* sites being completely recombined, because Cre recombination was much more effective in LB than in minimal media, before genomic DNA is purified. Cre recombination products were assayed by real-time PCR as described below.

IN VITRO CRE RECOMBINATION

In vitro Cre recombination was performed on cellular DNA that was crosslinked with formaldehyde as follows. Overnight cell culture in LB was diluted 1:100 into 50 ml LB. The cells were then incubated in a 30°C shaker for 3 hrs until OD_{600} reached 0.5-0.6. After that, 1.35 ml of formaldehyde crosslinker (Fisher-Scientific, 37%) was added to 50 ml of cell culture and

incubated for 20 min at room temperature with slow shaking. Then 12 ml of 2.5 M glycine was added to the culture and incubated for another 5 min at room temperature to stop the crosslinking reaction. Cells were then harvested and washed twice in 10 ml ice-cold PBS pH 7.5. Each 100 µl sample was spun down at 13,000 rpm, 4°C for 5 min and washed in 200 µl 1x reaction buffer (50mM Tris-Cl pH7.5, 50mM NaCl, 10nM MgCl2, 1mM DTT). Cells were centrifuged and resuspended in 192 µl digestion buffer (1x reaction buffer, 1x complete protease inhibitor from Roche). Next, 2 µl of 35 KU/µl Ready-Lyse lysozyme (Epicenter Biotechnologies) was added to the resuspension and incubated at room temperature for 20 min. After that, 6 µl of 10% SDS solution was added for overnight incubation at 37°C to denature and dissociate the loosely bound host proteins.

To do the *in vitro* Cre recombination, the cell lysate (100 µl) obtained from above was added to 200 µl reaction buffer and 30 µl of 20% Triton X-100 was added to the mixture to denature various host proteins. After incubation at 37°C for 1.5 hrs, 10 µl of Cre (1µg/µl) was added for 3 hrs at 37°C. Cre protein was a gift from Dr. Makkuni Jayaram (Ma et al., 2009). To remove RNA from the samples, 30 µl of 10 mg/ml RNase A (Promega) was added for 45 min incubation at 37°C for each sample. DNA was extracted by an equal volume of phenol/ chloroform/ isoamyl alcohol (25:24:1, pH 6.7, Fisher-Scientific) twice and then pure chloroform once. To precipitate DNA, glycogen (final concentration 50 µg/ml, Affymetrix), 10 µl 3M sodium acetate, and 200 µl isopropyl alcohol were added to 100µl DNA solution. Finally, the air-dried DNA pellet was hydrated with 100 µl of 10 mM Tris-HCl pH 7.5 for real-time PCR quantification.

REAL-TIME PCR

The real-time PCR conditions were: 50 ng of DNA, 10 µl SYBR master mix (ABI; includes dNTPs, enzyme and buffer), 0.4 µl of each primer (10 µM) and 8.2 µl of double distilled H2O. The reactions were held for 10 minutes at 95°C followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C (7900HT; Applied Biosystems). Three independent biological replicates were tested, and for each biological replicate three independent technical replicates were performed. The product integrity was checked using the dissociation curve. Cycle Threshold (C_t) was read out, and the starting template amount was quantified based on the value of C_t assuming exponential growth at early stages of amplification.

Recombination efficiencies were calculated based on the threshold cycle (Ct). The relative threshold cycles of each sample were calculated as ΔCt= Ct(P)-Ct(S), where (P) means recombination product and (S) means substrate before recombination. The recombination efficiencies (RE) of different samples are normalized to set the recombination efficiency of *loxP* sites flanking wild type Mu as one. The relative recombination efficiency of each sample is calculated as $RE= 2^{\Delta Ct\,(WT)-\Delta Ct(Mutant)}$.

Primer efficiencies were determined as follows: primer pairs were linked to pUC18f and pUC19r; the pUC primers anneal to the pUC19 plasmid and amplify a common 180 bp fragment. The PCR products were purified (Qiaquick PCR purification Kit®; Qiagen) and used as templates for real-time PCR in the following reaction: 12.5 µl SYBR mix (Qiagen), 0.75 µl of MuR (10 µM) paired with various primers (10 µM), 1 µl of template (10-5 ng/µl) and 10 µl of ddH2O. The reactions were held for 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C (7900HT; Applied Biosystems). The template input amount was

controlled by real-time PCR using the primer pair pUC19f and pUC19r. The efficiencies of the other primer pairs were normalized to Mu left end primer pairs.

CHAPTER 3: TRANSPOSABLE PROPHAGE MU EXISTS AS AN INDEPENDENT CHROMOSOMAL DOMAIN IN *E. COLI*

INTRODUCTION

The bacterial chromosome, is spatially organized and condensed ~1000 fold to fit inside a bacterial cell (Thanbichler et al., 2005; Travers and Muskhelishvili, 2005). As introduced in Chapter 1, the *E. coli* chromosome, referred to as a nucleoid, is organized in a series of negatively supercoiled loops or domains (Sinden and Pettijohn, 1981; Postow et al., 2004), segregated by dynamic domain barriers (defined as entities that prevent the free diffusion of supercoils) and compacted by several nucleoid-associated proteins (NAPs) including HU, IHF, Fis and H-NS (Dillon and Dorman, 2010; Macvanin and Adhya, 2012). The chromosome is not randomly condensed, but rather has a ring organization with four structured macro-domains and two less-structured regions (Figure 1.8); interactions between these regions are highly restricted as determined by cytological and genetic analyses (Valens et al., 2004). Macro-domains are thought to organize and orchestrate chromosome movements during the cell cycle (Willenbrock and Ussery, 2004; Boccard et al., 2005).

DNA supercoiling not only plays a vital role in compacting the chromosome, but a proper degree of supercoiling is crucial for all DNA-related processes (Travers and Muskhelishvili, 2005). Segregation of supercoils into topological domains protects these processes by preventing DNA breaks from relaxing the entire chromosome (Worcel and Burgi, 1972; Sinden and Pettijohn, 1981; Stein et al., 2005). The level of DNA superhelicity is tightly controlled by the

combined activities of topoisomerases (Witz and Stasiak, 2010) and nucleoid-associated proteins (Dillon and Dorman, 2010); the latter not only constrain negative supercoils (Luijsterburg et al., 2006) and generate diffusion barriers for the formation of topological domains (Hardy and Cozzarelli, 2005), but are also global regulators of gene transcription (See Chapter 1) (Dillon and Dorman, 2010). That DNA supercoiling is the master function that interconnects chromosome structure and global gene transcription was also seen in evolution experiments where supercoiling was observed to be under strong selection in *E. coli* populations propagated in parallel for several thousand generations (Crozat et al., 2010).

As described in the introductory chapter, transposable phage Mu is a temperate phage that integrates into essentially random locations on the *E. coli* chromosome (Figure 1.4) (Chaconas and Harshey, 2002; Manna et al., 2004; Ge et al., 2011). Transposition from mini-Mu plasmids *in vitro* requires DNA supercoiling for formation of a high-order transpososome within which the two Mu ends are interwound and synapsed (Figure 1.7) (Wang and Harshey, 1994; Au et al., 2006). Supercoiling is inferred to be similarly important *in vivo* as well (Sokolsky and Baker, 2003), where synapsis of prophage Mu ends additionally requires a centrally located gyrase binding site SGS (Figure 1.3) (Pato, 2004). This site is the strongest among such sites studied, and is found only in Mu-like prophages (Oram and Pato, 2004). It has been proposed that highly processive supercoiling by gyrase bound at the SGS promotes the formation of a supercoiled loop, with SGS at the apex, allowing reactive Mu ends to be synapsed at the base of the Mu genome, thereby facilitating the chemical steps of transposition (Pato and Banerjee, 1996).

The present study was first undertaken to determine if the transposase-mediated ordered assembly of three *cis*-acting sites - Mu L and R ends and an enhancer element E - observed *in vitro* (Kobryn et al., 2002; Pathania et al., 2002; Pathania et al., 2003), would also be seen *in vivo*.

Previous *in vitro* studies in Harshey lab employed Cre recombinase-mediated exchange at two strategically placed *loxP* sites to determine the topology of interactions between a given Mu site and the other two sites (Figure 1.7) (Pathania et al., 2002; Au et al., 2006). Using this same strategy *in vivo*, the ends were observed to be already paired in a prophage, and that the transposase was not involved in their pairing. The Cre-*loxP* system has been used to investigate the basis of this pairing (Dekker et al., 2002; Wang et al., 2011). It is shown that Mu SGS, Mu L end and host NAPs play an important role in the process, and that the MuB protein likely provides a NAP-like function. The results show that Mu exists as an independent chromosomal domain, and both phage and host factors contribute to its formation/maintenance. The implications of this work for the maintenance of Mu on the bacterial genome are discussed.

RESULTS

Mu ends at the termini of a 37 kb prophage genome behave as if they are paired.

Recombination catalyzed by Cre does not require any cofactors or special target topology; only two *loxP* sites are required (Hoess et al., 1984). The distance-dependence of *loxP* recombination efficiency on the *E. coli* chromosome was assessed first by placing a pair of directly oriented *loxP* sites from 189 bp to 37 kb apart (Figure 3.1; Tables 2.1 and 2.2). Cre recombinase was provided on a plasmid, and both the starting substrate and deletion product of recombination were amplified by appropriate primers and quantified by qPCR (Figure 3.2A). Recombination efficiency or RE was calculated as the ratio of the recombination product to the starting substrate (see Chapter 2). The RE of *loxP* pairs was seen to decrease exponentially when the distance between them increased from 189 bp to 10 kb, plateauing after ~20 kb (Figure 3.2B). In this experiment, *loxP* sites were placed in and around *malF* (Figure 3.1) because the Mu prophage monitored was located here.

loxP sites were next engineered within *E. coli* DNA flanking the L and R ends of a *malF*::Mu lysogen, ~70 bp from each end. The RE of these sites was closest to the control *loxP* sites placed 189 bp apart in the same region of *malF* (Figure 3.2C); the 2-fold RE difference between the two pairs is not statistically significant. To control for recombination at similar distances around this region of the chromosome, a *loxP* site was placed 37 kb upstream (*yjcF*) and one downstream (*purH*) of a *loxP* site in *malF* (Figure 3.1). The RE of both these pairs of *loxP* sites was similar (Figure 3.2C), and reflected their linear distance as determined from the graph shown in Figure 3.2B. The conclusion is that reduction of the linear 37 kb distance

between the L and R Mu ends to a distance equivalent to 189 bp as measured by Cre recombination, is indicative of some form of ‘synapsis’ of the Mu ends.

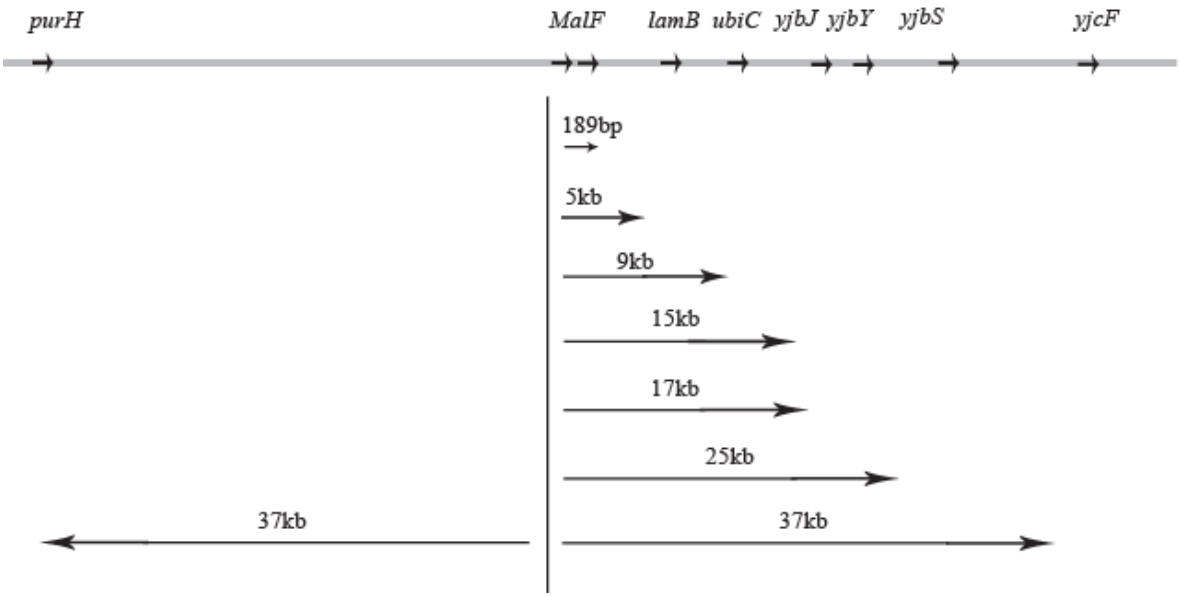


Figure 3.1. Position of pairs of *loxP* sites at different distances on the *E. coli* chromosome. The strains are listed in Table 2.1 and details of strain construction can be found in Chapter 2.

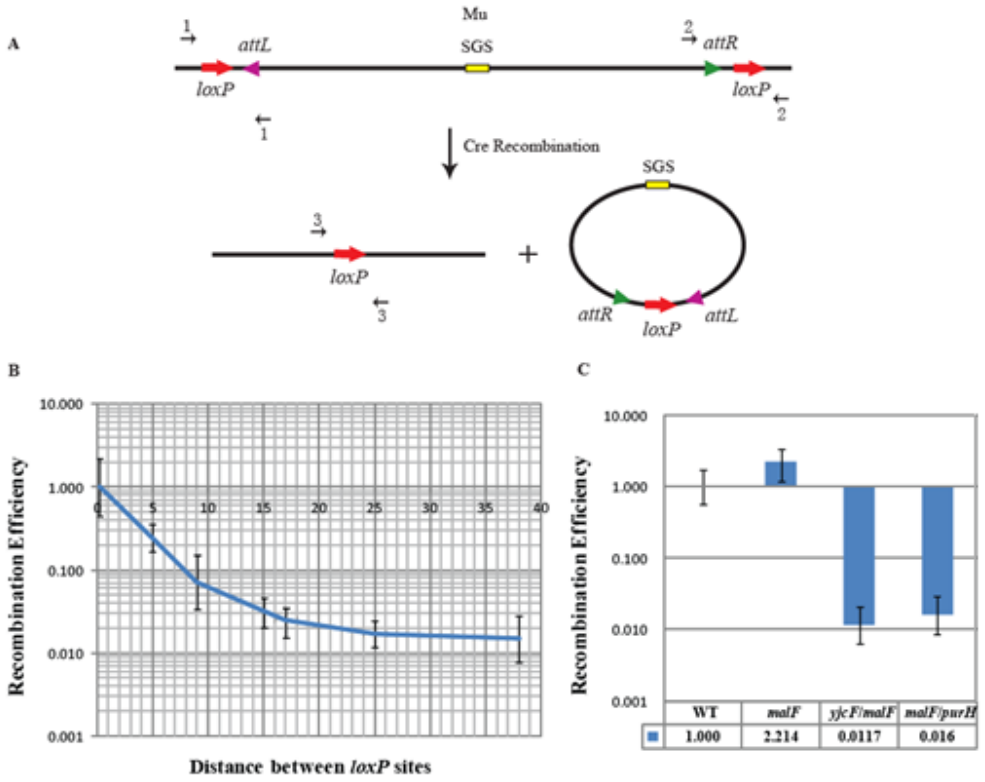


Figure 3.2 Cre-*loxP* recombination efficiency as a function of distance.

(A) Schematic showing *loxP* sites directly oriented on the chromosome, and the deletion products of Cre recombination. The diagram shows these sites flanking the *attL* and *attR* ends of Mu (refered to simply L and R ends in the text) with the centrally located SGS. The amount of substrate and product was assessed by PCR amplication across the *loxP* sites in the substrate (using either primer pairs 1 or 2) and in the product (primer pair 3). The efficiency of recombination was calculated as described in Chapter 2.
(B) Recombination efficiency (RE) of *loxP* sites with distance (in kb). The RE value obtained for the *loxP* pair with the smallest distance (189 bp) was arbitrarily set at 1. Error bars are S.D. of mean. Chromosomal location of the *loxP* sites is shown in Figure 3.1.
(C) RE of *loxP* sites flanking the *malF*::Mu prophage in MP1999 (WT), compared to that of *loxP* sites placed 189 bp-apart at the same locus (*malF*; the RE of this pair is set at 1 in Figure 3.2B), and those placed 37 kb downstream (*yjcF/malF*) and upstream (*malF/purH*) of the *loxP* site in *malF*. The bottom panel shows the precise fold differences in RE between these pairs of sites.

The centrally located strong-gyrase-site SGS within prophage Mu is important for end-synapsis.

A central location of SGS is obligatory for optimal replication of Mu monitored after prophage induction (Pato et al., 1990; Pato, 1994). Deletion of this site results in inefficient pairing of Mu ends *in vivo*, as judged by a low efficiency of transposase-mediated 3' nicking at the ends (Pato et al., 1995; Pato, 2004). Pato and colleagues have proposed that the requirement for SGS *in vivo* but not *in vitro* (where the distance between the Mu ends is typically ~2 kb on mini-Mu plasmid substrates), is an adaptation for aiding synapsis of transposition sites located at larger distances (Pato and Banerjee, 1996). Therefore the importance of both the presence and position of SGS on RE of *loxP*s flanking Mu was tested (Figure 3.3A top; O refers to *loxP* sites outside Mu). Deletion of the central SGS decreased *loxP* recombination by 30-fold, while asymmetric location to the left or right of center had a 10- and 3-fold reduction, respectively (Figure 3.3B); the latter two values are not significantly different from each other.

To test whether SGS would reduce the effective distance between sites present on non-Mu DNA, SGS was placed at the center of a 37 kb segment of *E. coli* DNA upstream of *malF* (*yjcF-malF*). SGS did not significantly increase the RE of *loxP* sites bordering this segment (Figure 3.3C). An SGS effect was also not seen over smaller regions (5-25 kb) in non-Mu DNA (Figure 3.4).

Because the effect of SGS is proposed to be mediated via gyrase-promoted supercoiling, the importance of gyrase was tested by introducing the temperature sensitive allele *gyrB402ts* into the Mu lysogen; at the non-permissive temperature, the *loxP* sites behaved as if they were unpaired (Figure 3.3D).

The conclusion is that DNA supercoiling is important for reducing the distance between the Mu ends, that SGS plays a critical role in this process when located centrally on a 37 kb Mu genome, but that SGS does not similarly contribute when located within non-Mu DNA. These results support the end-pairing function of SGS as deduced by the transposition/replication results of Pato and colleagues. However, the data were derived in the absence of prophage induction i.e. presumably in the absence of the transpososome proposed to stabilize the synapsed ends.

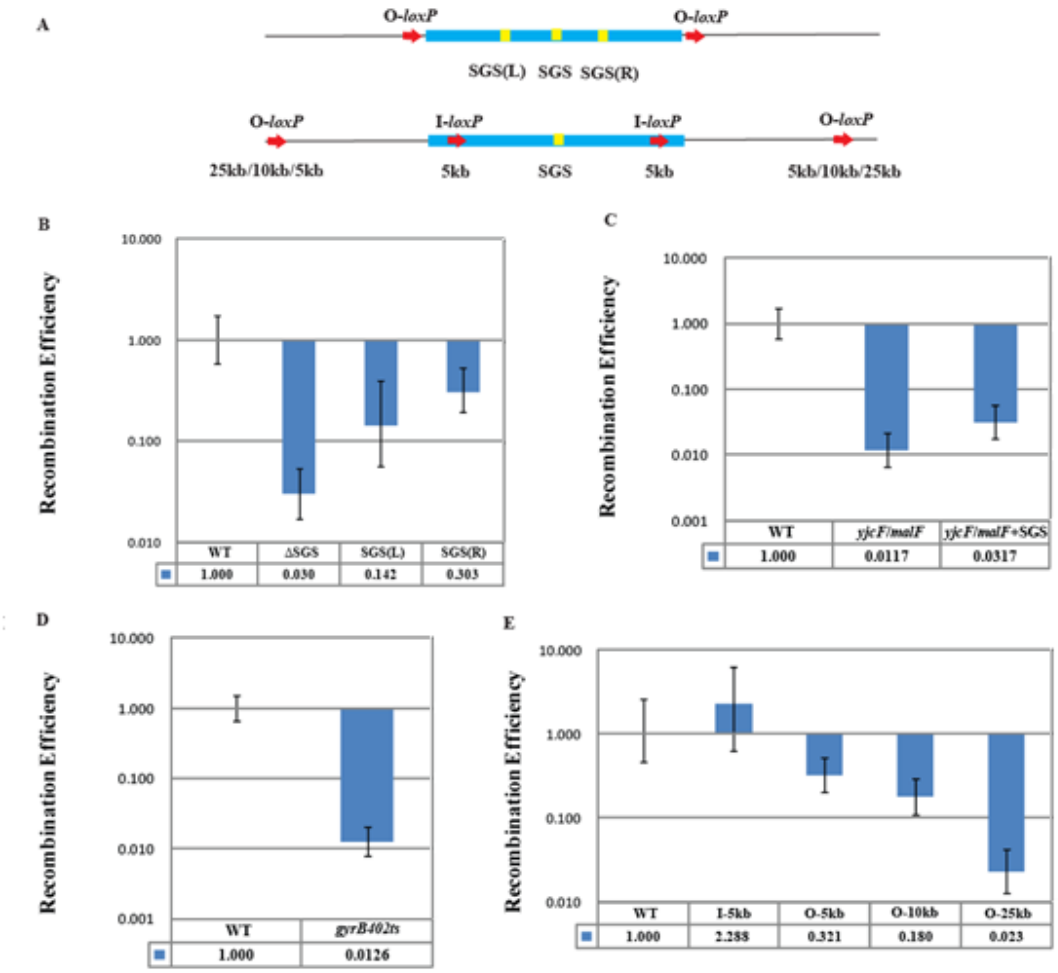


Figure 3.3. Importance of presence and position of SGS and DNA supercoiling to prophage Mu-end synapsis assayed by Cre-*loxP* recombination.
(A) Top: SGS was either deleted or displaced to the left and right of center, as described under Chapter 2. Bottom: *loxP* site pairs were placed either inside (I) or outside (O) Mu at indicated distances in different strains. The exact location of the displaced SGS and *loxP* sites is given in Table 3.3.
(B) RE of *loxP* sites flanking Mu in SGS-deleted, SGS-displaced strains was measured.
(C) The SGS site was engineered at the center of a 37 kb *E. coli yjcF-malF* segment and RE of flanking *loxP* sites measured.
(D) The *gyrB402ts* allele was introduced into the *malF*::Mu lysogen and RE measured at the non-permissive temperature as described in Methods.
(E) RE of *loxP* site pairs placed 5 kb inside Mu (I-*loxP*) or at varying distances outside Mu (O-*loxP*) as diagrammed in Fig. 1B. I-5kb, O-5kb etc. refers to RE of pairs of symmetrically placed sites within and outside the L and R ends of Mu in different strains.

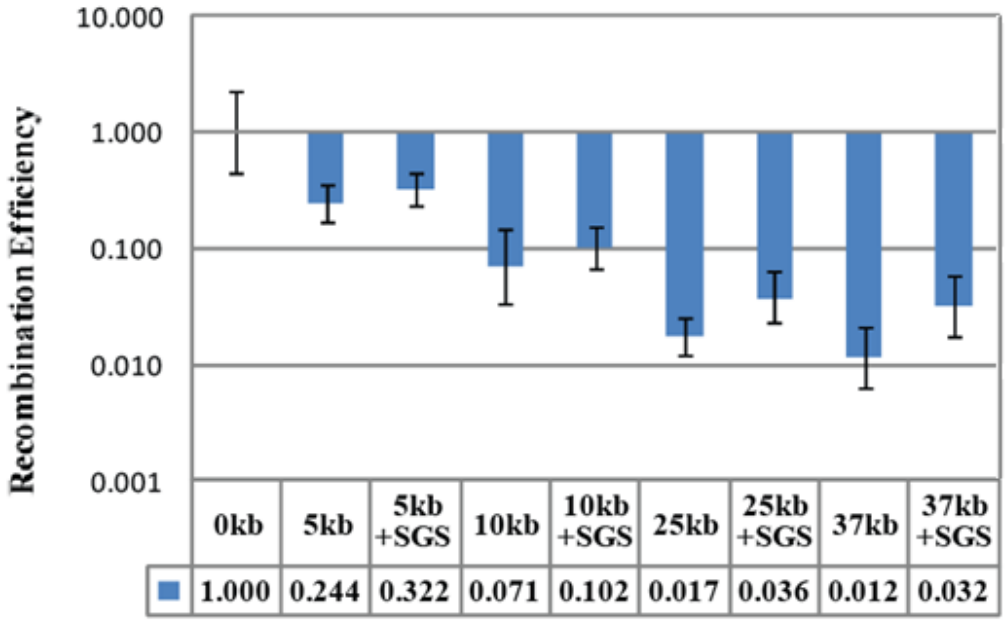


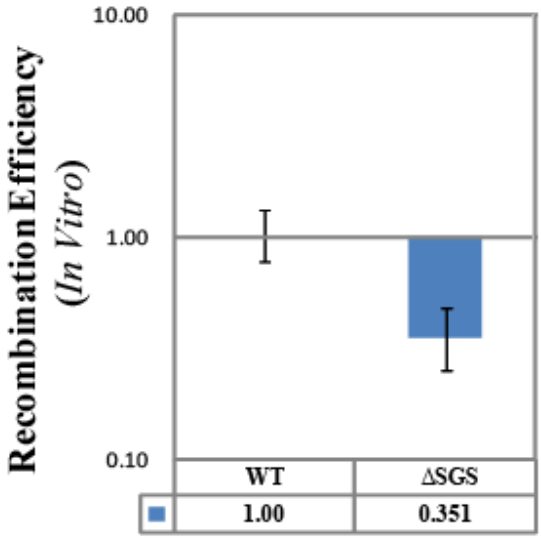
Figure 3.4. Effect of SGS on the RE of *loxP* site pairs at varying distances in non-Mu DNA. The SGS site was introduced in the center of the *loxP* pairs separated by 5 – 37 kb shown in Figure 3.1, and the RE with and without SGS was measured.

SGS-mediated Mu DNA synapsis does not extend far outside the Mu ends.

To test whether Mu ends define the base of the SGS-mediated Mu DNA synapsis loop, the *loxP* site pairs are moved symmetrically from 5 kb inside Mu (I-*loxP*; Figure 3.3A, bottom) to 5 - 25 kb outside Mu (O-*loxP*) in separate strains. RE of the internal *loxP*s (I-5 kb) was similar to WT, while that of external *loxP*s (O-5 kb and O-10 kb) decreased 3-5 fold over 5 and 10 kb distances, which is a total linear distance of 47 and 57 kb between the *loxP* pairs, respectively (Figure 3.3E). Synapsis was no longer evident between the O-25 kb pair (87 kb linear distance), as judged by O-*loxP* RE values which fell in the plateau region of the standard graph shown in Figure 3.2B. The conclusion is that the SGS effect extends 5 -10 kb outside Mu ends into the flanking *E. coli* DNA, but not beyond.

***In vitro* Cre recombination assay also reveals an interaction between Mu prophage ends.**

In this method, formaldehyde is used to cross-link proteins to other proteins and to DNA in intact cells (See Chapter 2 for details). Cre-*loxP* recombination is assayed *in vitro* in the cross-linked product. While the *in vitro* efficiency cannot be directly compared to the *in vivo* efficiency, the RE of *loxP* sites without SGS was lower than of the wild type substrate (Figure 3.5). These results are an independent confirmation of the close spatial proximity of Mu ends. This apparent Mu-loop is henceforth referred to as a 'Mu domain'.



Fugure 3.5 *In vitro* Cre-*loxP* recombination of the cross-linked DNA.

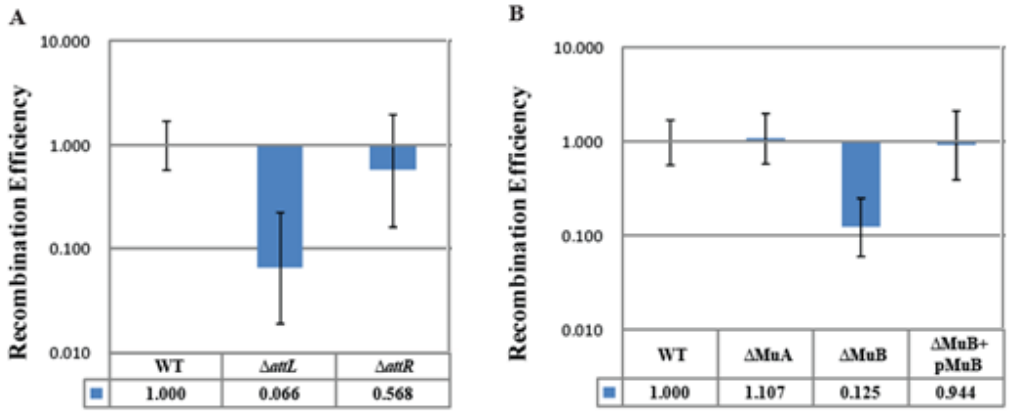


Figure 3.6. *cis*- and *trans*-acting Mu transposition factors required for formation of the Mu domain.
(A) RE of *loxP*s in Δ*attL* and Δ*attR* prohage strains compared to WT.
(B) RE of *loxP*s in strains deleted for Mu *A* and *B* genes, and in Δ*B* strain complementated with MuB

Importance of *cis*- and *trans*-acting Mu transposition factors.

Given that the SGS effect is Mu-specific, it is of interest to test if the L and R ends of Mu are important for closing the Mu loop at its base, and if so, whether the Mu transposase (A protein), known to bind to the Mu ends is expressed in the prophage to carry out this function. Therefore individually deletions of the L and R ends as well as the MuA and MuB proteins from the Mu prophage (MuB regulates MuA function allosterically; (Chaconas and Harshey, 2002)) were made and assayed RE of *loxP*s flanking Mu ends in these constructs. Deletion of the L end showed a 15-fold reduction of RE, but deletion of the R end showed no significant effect (Figure 3.6A). Deletion of the Mu *A* gene had no effect, but deletion of Mu *B* had an 8-fold effect (Figure 3.6B). To confirm the effect of the *B* gene deletion, MuB was supplied to this strain from a plasmid; the RE levels were restored to wild-type in this strain. The conclusion is that the L end is important but that the R ends is dispensable to the Mu domain. MuB but not MuA protein is also important to the Mu domain.

Cellular NAPs are critical for maintenance of the Mu domain.

E. coli NAPs such as H-NS, IHF, FIS, and HU are implicated in maintaining chromosomal supercoiled domains via their DNA bending and bridging properties. Recent study using 3C(chromosome conformation capture) in combination with super-resolution fluorescence microscopy found that in contrast to the major *E. coli* NAPs which were found largely scattered throughout the nucleoid, H-NS formed two compact clusters per chromosome, sequestering and juxtaposing numerous H-NS regulated DNA segments broadly distributed throughout the chromosome into these clusters; deleting H-NS led to substantial chromosome reorganization (Wang et al., 2011). IHF, FIS and HU have in addition, specific binding sites on the Mu genome (Chaconas and Harshey, 2002) (Figure 3.7A).

In order to determine the importance of these NAPs to the Mu domain, changes in Cre RE of *loxP* sites flanking Mu in strains individually deleted for genes expressing these proteins were assayed (Figure 3.7B). Absence of H-NS (Δ*hns*) had no effect on the Mu domain. Absence of HU (*hupA* encodes one of the two HU subunits; (Laine et al., 1980)) and of Fis (Δ*fis*) had 10- and 25-fold effects on RE, respectively. The strongest effect was observed in the absence of IHF, which essentially abrogated the Mu domain (*himA* encodes the A subunit of the heterodimeric IHF protein; (Friedman, 1988)).

To test if the effects of IHF, Fis and HU were exerted at the specific binding sites for these proteins on the Mu genome, these sites were deleted individually within the prophage. The dramatic reduction in RE seen in the IHF mutant was not observed with deletion of the IHF binding site (Figure 3.7B). Deletion of the HU-binding site at the Mu L end had a small effect, while deletion of the Fis-binding site (Δ*sis*) had no effect. Recent experiments have identified a set of three Fis-binding sites within the promoter region of the *mom* gene near the R end (Karambelkar et al., 2012). A deletion spanning all three sites (ΔP*mom*) also had no effect. However, a combination of *sis*-*attR* or *sis*-P*mom*-*attR* deletions reduced RE 5-10 fold. The conclusion is that IHF, HU and Fis affect the Mu domain configuration either directly or via their global effects on chromosome structure.

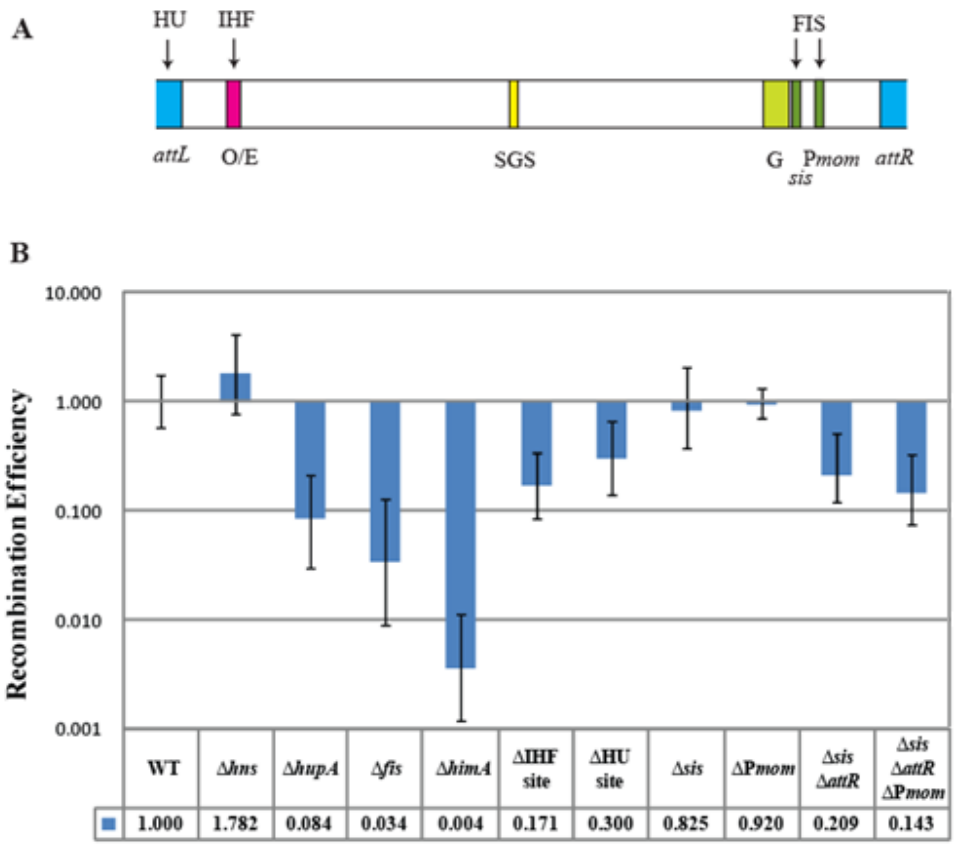


Figure 3.7. Role of *E. coli* NAPs and their binding sites on the Mu genome in formation of the Mu domain.
(**A**) Schematic showing binding sites for IHF, Fis and HU on the Mu genome.
(**B**) RE of *loxP* sites flanking Mu in strains either deleted for the indicated NAPs or for their binding sites on the Mu genome.

A domain organization is unique to Mu, and is not observed for prophage λ.

The four structured macro-domain regions of the *E. coli* chromosome are called Ori (3.76-0.04 Mb), Right (0.59-1.2 Mb), Ter (1.2-2.18 Mb), and Left (2.18-2.87 Mb); the two less structured regions are located on either side of Ori (Boccard et al., 2005) (Figure 3.8A). The *malF*::Mu prophage used in all the experiments thus far is located in the Ori macro-domain. In order to test if the Mu domain organization is specific to its location in a macro-domain, its presence was also tested in a *lacZ*::Mu prophage located in the less-structured NS region between Ori and Right. The results were similar to those seen with the *malF*::Mu prophage, with similar negative effects of deletion of SGS, and IHF on the domain structure (Figure 3.8B, compare to similar data in Figures. 3.3B and 3.7B). Deleting H-NS had no effect at this location as well.

If the domain organization of Mu were designed to pre-engage the Mu ends in a transposition-ready mode for lytic growth, might a similar arrangement be expected for other prophages which depend on pairing of their ends at the start of lytic growth? Prophage λ is an example of an insertion element which must pair its *attL* and *attR* ends for excisive recombination from the *E. coli* chromosome (Campbell, 1962; Azaro and Landy, 2002). The location of the λ prophage is between the *gal* and *bio* genes and within the *attB* site, which resides in the Right macro-domain (Figure 3.8A). To determine if λ ends were paired, *loxP* sites were engineered flanking the ends of the 48.5 kb λ genome, as what had done for Mu. In contrast to Mu, however, the Cre-mediated RE of these sites reflected the linear distance between the λ ends (Figure 3.8B). Thus, domain organization is specific to Mu.

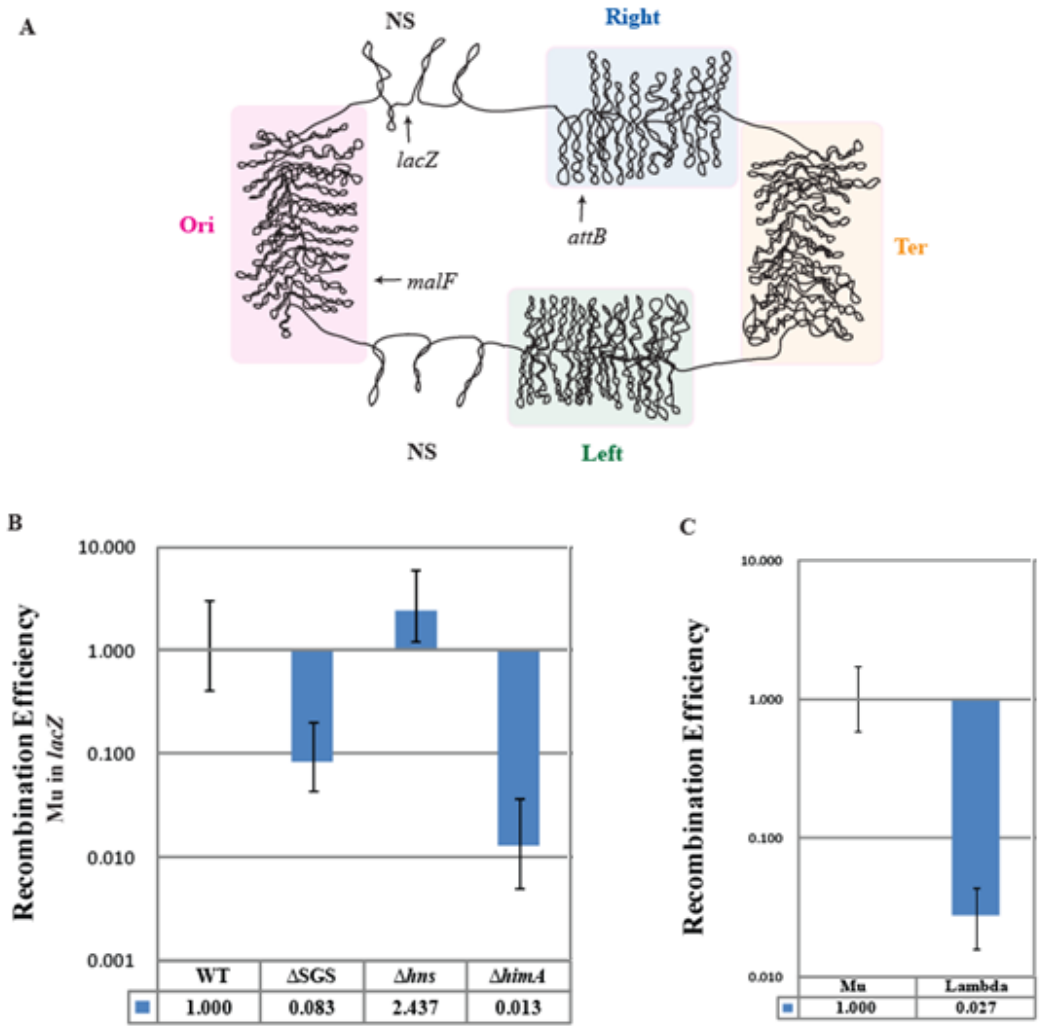


Figure 3.8. Determination of domain organization for Mu at the NS chromosomal location, and testing for a similar domain organization for prophage λ.
(A) Ring organization of the macrodomains on the *E. coli* chromosome and position of prophages at the loci examined in this study.
(B) RE of *loxP* sites flanking Mu located in *lacZ* and its isogenic Mu and host mutants.
(C) RE of *loxP* sites flanking prophage λ compared to Mu in *lacZ*.

DISCUSSION

In this study, the Cre-*loxP* system is employed to probe the domain organization of prophage Mu in the *E. coli* chromosome. Previous studies of chromosome supercoiling domains used methods such as trimethylpsoralen binding, direct observation via electron microscopy, transcription of supercoiling sensitive genes, or efficiency of site-specific recombinases Res and Int (Sinden and Pettijohn, 1981; Higgins et al., 1996; Garcia-Russell et al., 2004; Postow et al., 2004). Cre recombinase was chosen because of the simplicity of its substrate requirements (Hoess et al., 1984; Ghosh and Van Duyne, 2002), and the ease of targeting *loxP* sites to specific locations on the chromosome. The distance-dependence of Cre recombination reported has not been examined *in vivo* before. Using this method, it was determined that the Mu prophage occupies a separate supercoiled domain, and that both Mu and host factors contribute to its formation and maintenance. There is no example to date of such a large stable chromosomal domain in *E. coli*.

Cre-*loxP* recombination: a reporter for chromosomal domains?

Two site-specific recombination systems have been used before to study bacterial chromosome organization. The γδ resolvase system was used to examine recombination in a non-essential region spanning ~100 kb of the *S. typhimurium* (Higgins et al., 1996). Recombination in this system depends on pairing of *res* sites within the same plectonemically supercoiled domain (Watson et al., 1996). The efficiency of recombination was found to decrease linearly with distance between *res* sites over the entire region studied. That study concluded that topological domains are variable in both size and placement of barriers, and that

domains average 25 kb in length. A later study in *E. coli* using an entirely different set of experiments agreed with the γδ study with regard to the variability of size and placement of barriers, but arrived at a smaller average domain size; electron microscopy measured the sizes of the supercoiled loops to vary exponentially between 2 – 66 kb, with an average of 10 kb (Postow et al., 2004). In other studies, recombination of *attL* and *attR* sites of prophage λ, which is independent of supercoiling and location (Nash and Pollock, 1983; Azaro and Landy, 2002), was used to probe interaction of these sites placed over different regions of the *E. coli* chromosome (Garcia-Russell et al., 2004; Valens et al., 2004). These studies found that chromosomal locations are not equally accessible to each other, and concluded that the chromosome is not completely fluid but rather organized in a way that makes some pairs of sites less accessible than others. The macro-domain organization of the *E. coli* chromosome deduced by these studies is shown in Figure 3.8A (Valens et al., 2004).

Like the λ Int system, the P1 Cre system also does not require that *loxP* sites be present on supercoiled DNA or on the same DNA molecule. Unlike the γδ *res* sites, which find each other by slithering on a plectonemic supercoil, *loxP* sites find each other by random collision. It was expected therefore that *loxP* sites would show a distance-dependence for recombination at larger distances, although it was not known *a priori* what these distances would be. The data show that, in the *malF* region of the *E. coli* chromosome, Cre RE decreases exponentially when *loxP* sites are separated by distances ranging from 189 bp to 10 kb, plateauing after ~20 kb (Figure 3.2B). If short distances restrict the sites to the same supercoil branch, these sites would incur higher recombination. At longer distances, sites are likely to be in different supercoil branches; the higher the number of branches, the greater might be the steric interference of inter-branch collisions. If the biphasic behavior of *loxP* sites, demarcated around 10 kb, reflects intra- versus

inter-domain recombination, then the domain size is consistent with the average domain size of ~10 kb derived by Postow et al. (Postow et al., 2004).

Sequestration of Mu into an independent supercoiled domain: Mu and host factors.

The efficiency of recombination of *loxP* sites flanking a *malF*::Mu prophage mimics that of sites placed 189 bp apart within *malF*, indicative of a close proximity of Mu ends (Figure 3.2C). A second biochemical method, where interacting sites are cross-linked *in vivo* via their bound proteins, and monitored *in vitro* after Cre recombination, supported this close arrangement of Mu ends (Figure 3.5). The centrally-located gyrase binding site SGS, as well as gyrase, were critical to the synaptic arrangement of Mu ends (Figure 3.4B,D), leading to the conclusion that prophage Mu is segregated into a separate supercoiled domain on the *E. coli* chromosome. Such an SGS-dependent supercoiled Mu domain has been hypothesized to assist Mu end synapsis via the known end-bridging properties of the Mu transposase (Pato, 1994; Chaconas and Harshey, 2002), and indeed been deduced to exist throughout Mu replicative growth (Ge et al., 2010). The interesting aspect of supercoiled Mu-loop/domain formation inferred in the present work is that it exists in a prophage i.e. prior to induction of Mu transposition functions, and is independent of Mu location at the two positions tested (Figures. 3.3, 3.7 and 3.8), positions reported to have different levels of compaction (Ori macro-domain and a less-structured region) (Boccard et al., 2005; Dame et al., 2011). Pairing of the two Mu DNA arms was seen to extend 5-10 kb outside Mu into the *E. coli* DNA (Figure 3.3E).

***Mu factors*.**

The observation was that the Mu L end but not the R end to be important for domain formation, and the MuB protein but not the MuA protein to make a substantial contribution (Figure 3.6A,B). Thus, the Mu domain is not stabilized by the end-bridging interactions of the transposase expected in a Mu loop formed during replication-transposition. It is possible that in the prophage, the ends are bridged not by the transposase but by Mu repressor Rep, which could not be tested because it is essential for prophage maintenance. Cellular NAPs expected to bind the AT-rich regions at the Mu ends appear to play an important role (Figure 3.7B). Given that deletion of the R end, in combination with deletion of the Fis-binding sites at the R end, reduces efficiency of the domain to some degree, the R end can perhaps be substituted by A-tract elements to which Fis is known to bind (Cho et al., 2008; Ge et al., 2011). A-tracts are over-represented in the *E. coli* genome, and distributed 'quasi-regularly' with a 10–12 bp periodicity throughout the genome; they are proposed to constitute the 'structural code' for DNA compaction (Tolstorukov et al., 2005).

The contribution of MuB to domain maintenance suggests that the protein is expressed in the prophage state (Symonds et al., 1987). Unpublished studies in Harshey lab do indeed find low levels of MuB as monitored by fluorescence of an EGFP-MuB fusion engineered into the prophage, as well as transcripts originating from both the lytic promoter Pe and a new promoter within MuA named Pcon. Deletion of both promoters was equivalent to deletion of Mu*B* for domain maintenance (Rudra Saha, unpublished data). MuB has several different functions in transposition (Chaconas and Harshey, 2002). It is essential for target selection, for modulating the activity of the transposase, and for a phenomenon known as *cis*-immunity, where Mu insertions are discouraged for some distance outside Mu (Adzuma and Mizuuchi, 1988; Manna

et al., 2004). MuB is also implicated in 'Mu genome immunity', where its strong binding within the Mu genome is hypothesized to prevent use of Mu as target, thus avoiding self-integration during replication (Ge et al., 2010). One might ask why MuB is being made under conditions that repress transposition. Because MuB has a preference for AT-rich sequences (Adzuma and Mizuuchi, 1991; Greene and Mizuuchi, 2004; Ge and Harshey, 2008), it is possible that MuB may serve as an additional NAP for maintenance of the prophage domain. Both the N- and C-terminal regions of MuB have non-specific DNA-binding activity (Chaconas and Harshey, 2002), and could be involved additionally in tethering the two Mu DNA arms of loop.

Host factors.

Of the dozen or so NAPs recognized in *E. coli* thus far, the most abundant are Fis, HU, H-NS and IHF (Ali Azam et al., 1999). These proteins bind A-T rich DNA segments, and tend to have overlapping binding sites (Grainger et al., 2006; Cho et al., 2008). Interestingly, MuB competes with NAPs for binding AT-rich sequences (Ge et al., 2011), which are also found at the ends and center of the Mu genome (Ge et al., 2010). NAPs compact DNA by their bending, wrapping and bridging activities (Dillon and Dorman, 2010). Bacterial cells deficient in one of the NAPs usually have subtle phenotypes, which indicate that (some of) the roles of one NAP can be fulfilled by another. Double mutations of HU/Fis, HU/IHF often have more severe effects, while H-NS/Fis are not viable (Ge et al., 2010). Single deletions of the cellular NAPs showed that absence of IHF had the most severe consequence, effectively abrogating the Mu domain (Figure 3.7B). Deletion of Fis and HU had substantial effects, but H-NS had none. Thus, the chromosome compaction attributed to H-NS (Wang et al., 2011) is not involved in formation of the Mu domain.

IHF, HU and Fis all have specific binding sites on the Mu genome (Chaconas and Harshey, 2002). IHF binds in the operator region and regulates transcription of the early promoter Pe (Higgins et al., 1989), HU binds within the Mu L end subsites L1 and L2 and is important for assembly of the transpososome (Lavoie and Chaconas, 1993; Chaconas and Harshey, 2002), and Fis is important for site-specific recombination of the invertible G segment (Koch et al., 1988). Although deletion of these sites affected formation of the Mu domain, none of them had the severe impact seen with deletion of IHF. This suggests that the effect of IHF may be manifested through its general NAP-related property, rather than through binding to its specific site. A curious aspect of NAP-binding on the Mu genomeis that except for the Mu ends and center, the rest of the Mu genome is not AT-rich (Ge et al., 2010). Thus, either these NAPs all bind to the ends and center, or that some other property of the supercoiled loop propagated from SGS (higher superhelical density?) increases the affinity of these proteins throughout Mu DNA.

***Model*.**

Based on the results presented in Figures 3.2-3.8, a model is proposed for the organization of the prophage Mu domain (Figure 3.9). The model incorporates the initial suggestion of Pato and colleagues that gyrase-mediated processive supercoiling creates a Mu loop with SGS at the apex and Mu ends at the base. In the prophage state, it is AT-rich character of the Mu ends that defines the base. Both site-specific and non-specific binding of Mu Rep, and host IHF, HU and Fis, stabilize this structure and possibly seal it at the base. Other NAPs not tested in this study, such as SMC-like proteins, may also be involved. It is suggestted that MuB, which binds strongly throughout the Mu genome during replication (Ge et al., 2010), contributes a NAP-like function as well (Ge et al., 2011). A second potential Mu function that might be involved in Mu-

end binding is the Mu lysogenic repressor Rep, which shares homology with the Mu transposase and binds with a lower affinity to the same end-sites as the transposase (Chaconas and Harshey, 2002);

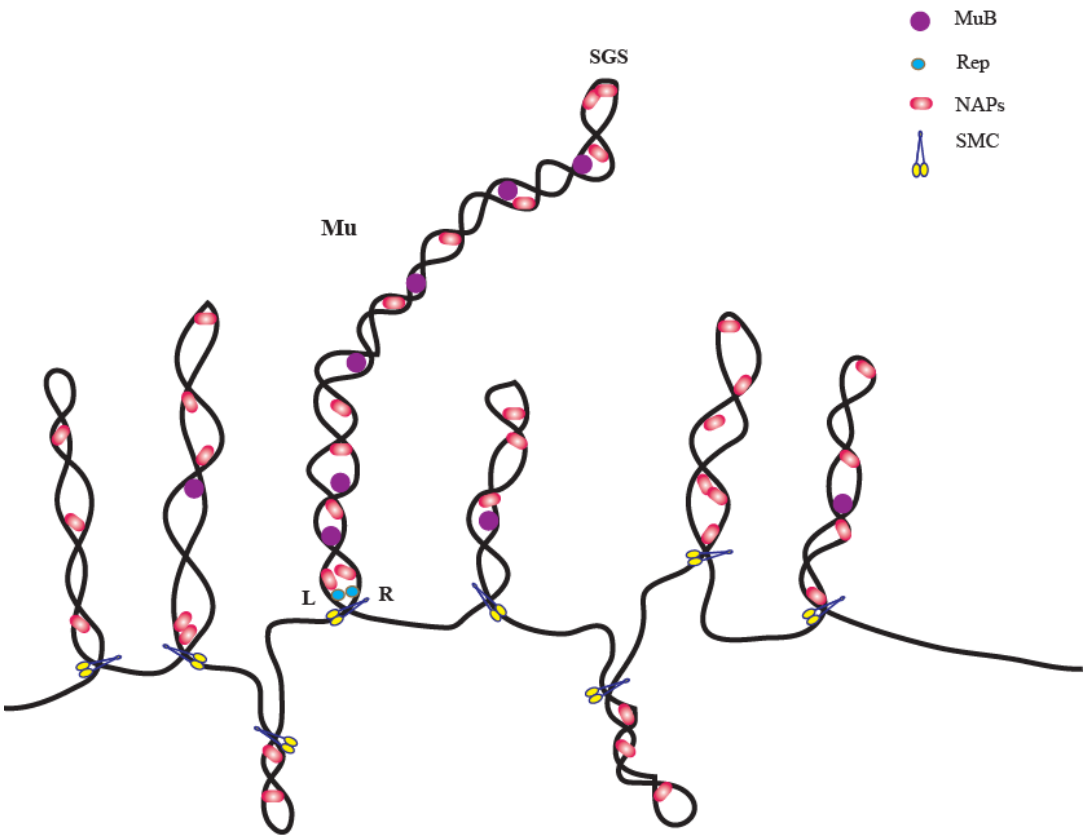


Figure 3.9. Model of the Mu prophage domain.
Segregation of Mu into a separate domain on the *E. coli* chromosome is aided by the centrally located SGS, which initiates loop formation, and is sealed by NAP-binding to AT-rich elements at the Mu ends (see text for details). The NAPs could also regulate the Mu domain indirectly. The single supercoiled Mu loop shown is not intended to imply absence of branching. A variation of this model was proposed earlier for replicating Mu, to account for strong MuB binding only within Mu (Ge et al., 2010).

Why is the domain organization seen for Mu but not λ prophage?

If the domain organization of Mu aids synapsis of the Mu ends and promotes transposition, one might have expected a similar organization of prophage λ ends, which pair during excisive recombination. However, λ ends can interact by random collision from any DNA substrate (Azaro and Landy, 2002), while Mu requires that its DNA be supercoiled and the ends be present on the same DNA molecule (Craigie and Mizuuchi, 1986; Au et al., 2006). Thus, SGS is a unique adaptation for end-synapsis of Mu-like phages. But why are the ends synapsed in the prophage state where transposition is repressed?

Prophages are the major contributors to genome diversification in some species (Canchaya et al., 2004). They often ensure against self-elimination by conferring varied fitness traits to their hosts including virulence factors for infectious diseases and antibiotic resistance (Frost et al., 2005). While a majority of phage λ genes are not transcribed in the prophage state, some genes such a *rex*, *lom* and *bor* are transcribed at a low level and confer resistance to lytic phages or to serum (Lin et al., 1977; Barondess and Beckwith, 1990). Prophage genes without selective value to the host are likely to be deleted, and Mu and other prophages have been demonstrated to increase fitness of *E. coli* grown in chemostats under glucose-limiting conditions (Edlin et al., 1977). The data showing an effect of Mu *B* gene deletion on the Mu domain (Figure 3.6B), as well as unpublished data of Harshey lab showing that the early promoter Pe is expressed at a low level in a domain-dependent manner in the prophage, suggests that transcription of these genes must provide some benefit. Pe controls transcription of a large early transcript, which includes not only the *A* and *B* genes, but also a large number of genes referred to as 'semi-essential' because they reduce the Mu plaque size but are otherwise not essential, whose function is largely unknown (Symonds et al., 1987). However, one of the known genes *lig* has a ligase function, and

the gene *gam* is an orthologue of the eukaryotic protein Ku, which participates in double strand break repair (Symonds et al., 1987; d'Adda di Fagagna et al., 2003). The Mu domain is suggested to be beneficial to both Mu and its host – to Mu in providing a transposition-ready end-configuration, and to the host in providing an extra NAP in the form of MuB as well as beneficial functions such as *lig, gam* and others that are regulated from the early promoter Pe and the newly mapped MuB promoter P_{con} (lab unpublished data). Thus, the SGS site, and its role in structuring the domain configuration of Mu, is an evolutionary acquisition that benefits both the phage and its host.

REFERENCES

Adzuma, K. and Mizuuchi, K. (1988) Target immunity of Mu transposition reflects a differential distribution of Mu B protein. *Cell*, **53**, 257-266.

Adzuma, K. and Mizuuchi, K. (1989) Interaction of proteins located at a distance along DNA: mechanism of target immunity in the Mu DNA strand-transfer reaction. *Cell*, **57**, 41-47.

Adzuma, K. and Mizuuchi, K. (1991) Steady-state kinetic analysis of ATP hydrolysis by the B protein of bacteriophage mu. Involvement of protein oligomerization in the ATPase cycle. *J Biol Chem*, **266**, 6159-6167.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *Journal of Bacteriology*, **181**, 6361-6370.

Au, T.K., Agrawal, P. and Harshey, R.M. (2006) Chromosomal integration mechanism of infecting mu virion DNA. *J Bacteriol*, **188**, 1829-1834.

Azam, T.A., Hiraga, S. and Ishihama, A. (2000) Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid. *Genes Cells*, **5**, 613-626.

Azaro, M.A. and Landy, A. (2002) *Lambda Integrase and the Lambda Int family*. American Society of Microbiology, Washington, DC.

Baker, T.A. and Mizuuchi, K. (1992) DNA-promoted assembly of the active tetramer of the Mu transposase. *Genes Dev*, **6**, 2221-2232.

Barondess, J.J. and Beckwith, J. (1990) A bacterial virulence determinant encoded by lysogenic coliphage lambda. *Nature*, **346**, 871-874.

Baudry, C., Malinsky, S., Restituito, M., Kapusta, A., Rosa, S., Meyer, E. and Betermier, M. (2009) PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate Paramecium tetraurelia. *Genes Dev*, **23**, 2478-2483.

Becker, N.A., Kahn, J.D. and Maher, L.J., 3rd. (2007) Effects of nucleoid proteins on DNA repression loop formation in Escherichia coli. *Nucleic Acids Res*, **35**, 3988-4000.

Berg, C.M. and Berg, D.E. (1996) *Transposable element tools for microbial genetics.* American Society for Microbiology, Washington, D.C.

Berg, D.E. and Howe, M.M. (1989) *Mobile DNA*. American Society of Microbiology, Washington, D.C.

Boccard, F., Esnault, E. and Valens, M. (2005) Spatial arrangement and macrodomain organization of bacterial chromosomes. *Molecular Microbiology*, **57**, 9-16.

Browning, D.F., Grainger, D.C. and Busby, S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol*, **13**, 773-780.

Bukhari, A.I., Froshauer, S. and Botchan, M. (1976) Ends of bacteriophage mu DNA. *Nature*, **264**, 580-583.

Burrus, V., Pavlovic, G., Decaris, B. and Guedon, G. (2002) Conjugative transposons: the tip of the iceberg. *Mol Microbiol*, **46**, 601-610.

Campbell, A. (ed.). (1962) *Episomes*. Academic, New York.

Canchaya, C., Fournous, G. and Brussow, H. (2004) The impact of prophages on bacterial chromosomes. *Molecular Microbiology*, **53**, 9-18.

Case, R.B., Chang, Y.P., Smith, S.B., Gore, J., Cozzarelli, N.R. and Bustamante, C. (2004) The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure. *Science*, **305**, 222-227.

Chaconas, G. and Harshey, R.M. (2002) *Transoposition of phage Mu DNA*. American Society of Microbiology, Washington D.C.

Chaconas, G., Harshey, R.M., Sarvetnick, N. and Bukhari, A.I. (1981) Predominant end-products of prophage Mu DNA transposition during the lytic cycle are replicon fusions. *J Mol Biol*, **150**, 341-359.

Cho, B.K., Knight, E.M., Barrett, C.L. and Palsson, B.O. (2008) Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. *Genome Res*, **18**, 900-910.

Choi, W. and Harshey, R.M. (2010) DNA repair by the cryptic endonuclease activity of Mu transposase. *Proc Natl Acad Sci U S A*, **107**, 10014-10019.

Clubb, R.T., Omichinski, J.G., Savilahti, H., Mizuuchi, K., Gronenborn, A.M. and Clore, G.M. (1994) A novel class of winged helix-turn-helix protein: the DNA-binding domain of Mu transposase. *Structure*, **2**, 1041-1048.

Clubb, R.T., Schumacher, S., Mizuuchi, K., Gronenborn, A.M. and Clore, G.M. (1997) Solution structure of the I gamma subdomain of the Mu end DNA-binding domain of phage Mu transposase. *J Mol Biol*, **273**, 19-25.

Cost, G.J., Feng, Q., Jacquier, A. and Boeke, J.D. (2002) Human L1 element target-primed reverse transcription in vitro. *Embo J*, **21**, 5899-5910.

Craig, N.L. (2002) *Mobile DNA II*. American Society for Microbiology, Washington D.C.

Craigie, R. (2001) HIV integrase, a brief overview from chemistry to therapeutics. *J Biol Chem*, **276**, 23213-23216.

Craigie, R. and Mizuuchi, K. (1986) Role of DNA topology in Mu transposition: mechanism of sensing the relative orientation of two DNA segments. *Cell*, **45**, 793-800.

Craigie, R. and Mizuuchi, K. (1987) Transposition of Mu DNA: joining of Mu to target DNA can be uncoupled from cleavage at the ends of Mu. *Cell*, **51**, 493-501.

Craigie, R., Mizuuchi, M. and Mizuuchi, K. (1984) Site-specific recognition of the bacteriophage Mu ends by the Mu A protein. *Cell*, **39**, 387-394.

Crozat, E., Winkworth, C., Gaffe, J., Hallin, P.F., Riley, M.A., Lenski, R.E. and Schneider, D. (2010) Parallel genetic and phenotypic evolution of DNA superhelicity in experimental populations of Escherichia coli. *Molecular biology and evolution*, **27**, 2113-2128.

Curcio, M.J. and Derbyshire, K.M. (2003) The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol*, **4**, 865-877.

d'Adda di Fagagna, F., Weller, G.R., Doherty, A.J. and Jackson, S.P. (2003) The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep*, **4**, 47-52.

de Vries, R. (2010) DNA condensation in bacteria: Interplay between macromolecular crowding and nucleoid proteins. *Biochimie*, **92**, 1715-1721.

Dame, R.T., Kalmykowa, O.J. and Grainger, D.C. (2011) Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. *PLoS genetics*, **7**, e1002123.

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

Datta, S., Costantino, N., Zhou, X. and Court, D.L. (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc Natl Acad Sci U S A*, **105**, 1626-1631.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306-1311.

Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol*, **8**, 185-195.

Duncan, L., Bouckaert, K., Yeh, F. and Kirk, D.L. (2002) kangaroo, a mobile element from Volvox carteri, is a member of a newly recognized third class of retrotransposons.

Genetics, **162**, 1617-1630.

Edlin, G., Lin, L. and Bitner, R. (1977) Reproductive fitness of P1, P2, and Mu lysogens of Escherichia coli. *Journal of virology*, **21**, 560-564.

Engelman, A., Mizuuchi, K. and Craigie, R. (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, **67**, 1211-1221.

Finnegan, D.J. (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet*, **5**, 103-107.

Flamm, E.L. and Weisberg, R.A. (1985) Primary structure of the hip gene of Escherichia coli and of its product, the beta subunit of integration host factor. *J Mol Biol*, **183**, 117-128.

Friedman, D.I. (1988) Integration host factor: a protein for all reasons. *Cell*, **55**, 545-554.

Friedman, D.I. and Court, D.L. (2001) Bacteriophage lambda: alive and well and still doing its thing. *Curr Opin Microbiol*, **4**, 201-207.

Frost, L.S., Leplae, R., Summers, A.O. and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nature reviews. Microbiology*, **3**, 722-732.

Garcia-Russell, N., Harmon, T.G., Le, T.Q., Amaladas, N.H., Mathewson, R.D. and Segall, A.M. (2004) Unequal access of chromosomal regions to each other in Salmonella: probing chromosome structure with phage lambda integrase-mediated long-range rearrangements. *Mol Microbiol*, **52**, 329-344.

Ge, J. and Harshey, R.M. (2008) Congruence of in vivo and in vitro insertion patterns in hot E. coli gene targets of transposable element Mu: opposing roles of MuB in target capture and integration. *J Mol Biol*, **380**, 598-607.

Ge, J., Lou, Z., Cui, H., Shang, L. and Harshey, R.M. (2011) Analysis of phage Mu DNA transposition by whole-genome Escherichia coli tiling arrays reveals a complex relationship to distribution of target selection protein B, transcription and chromosome architectural elements. *J Biosci*, **36**, 587-601.

Ge, J., Lou, Z. and Harshey, R.M. (2010) Immunity of replicating Mu to self-integration: a novel mechanism employing MuB protein. *Mob DNA*, **1**, 8.

Gellert, M. and Nash, H. (1987) Communication between segments of DNA during site-specific recombination. *Nature*, **325**, 401-404.

Ghosh, K. and Van Duyne, G.D. (2002) Cre-loxP biochemistry. *Methods*, **28**, 374-383.

Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W. and Wright, A. (1997) Chromosome and low copy plasmid segregation in E. coli: visual

evidence for distinct mechanisms. *Cell*, **90**, 1113-1121.

Gottesman, M.E. and Weisberg, R.A. (2004) Little lambda, who made thee? *Microbiol Mol Biol Rev*, **68**, 796-813.

Grainger, D.C., Hurd, D., Goldberg, M.D. and Busby, S.J. (2006) Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome. *Nucleic Acids Res*, **34**, 4642-4652.

Greene, E.C. and Mizuuchi, K. (2002) Target immunity during Mu DNA transposition. Transpososome assembly and DNA looping enhance MuA-mediated disassembly of the MuB target complex. *Mol Cell*, **10**, 1367-1378.

Greene, E.C. and Mizuuchi, K. (2004) Visualizing the assembly and disassembly mechanisms of the MuB transposition targeting complex. *J Biol Chem*, **279**, 16736-16743.

Groenen, M.A. and van de Putte, P. (1986) Analysis of the ends of bacteriophage Mu using site-directed mutagenesis. *J Mol Biol*, **189**, 597-602.

Hardy, C.D. and Cozzarelli, N.R. (2005) A genetic selection for supercoiling mutants of Escherichia coli reveals proteins implicated in chromosome structure. *Mol Microbiol*, **57**, 1636-1652.

Haren, L., Ton-Hoang, B. and Chandler, M. (1999) Integrating DNA: transposases and retroviral integrases. *Annu Rev Microbiol*, **53**, 245-281.

Harshey, R.M. and Bukhari, A.I. (1983) Infecting bacteriophage mu DNA forms a circular DNA-protein complex. *J Mol Biol*, **167**, 427-441.

Harshey, R.M., Getzoff, E.D., Baldwin, D.L., Miller, J.L. and Chaconas, G. (1985) Primary structure of phage mu transposase: homology to mu repressor. *Proc Natl Acad Sci U S A*, **82**, 7676-7680.

Hickman, A.B., Chandler, M. and Dyda, F. (2010) Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Crit Rev Biochem Mol Biol*, **45**, 50-69.

Higgins, N.P., Collier, D.A., Kilpatrick, M.W. and Krause, H.M. (1989) Supercoiling and integration host factor change the DNA conformation and alter the flow of convergent transcription in phage Mu. *J Biol Chem*, **264**, 3035-3042.

Higgins, N.P., Yang, X., Fu, Q. and Roth, J.R. (1996) Surveying a supercoil domain by using the gamma delta resolution system in Salmonella typhimurium. *J Bacteriol*, **178**, 2825-2835.

Hiom, K., Melek, M. and Gellert, M. (1998) DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell*, **94**, 463-470.

Hoess, R., Abremski, K. and Sternberg, N. (1984) The nature of the interaction of the P1 recombinase Cre with the recombining site loxP. *Cold Spring Harbor symposia on quantitative biology*, **49**, 761-768.

Howe, M.M. and Bade, E.G. (1975) Molecular biology of bacteriophage mu. *Science*, **190**, 624-632.

Hung, L.H., Chaconas, G. and Shaw, G.S. (2000) The solution structure of the C-terminal domain of the Mu B transposition protein. *Embo J*, **19**, 5625-5634.

Jang, S., Sandler, S.J. and Harshey, R.M. (2012) Mu insertions are repaired by the double-strand break repair pathway of Escherichia coli. *PLoS Genet*, **8**, e1002642.

Johnson, R.C., Bruist, M.F. and Simon, M.I. (1986) Host protein requirements for in vitro site-specific DNA inversion. *Cell*, **46**, 531-539.

Jones, J.M. and Gellert, M. (2004) The taming of a transposon: V(D)J recombination and the immune system. *Immunol Rev*, **200**, 233-248.

Karambelkar, S., Swapna, G. and Nagaraja, V. (2012) Silencing of toxic gene expression by Fis. *Nucleic Acids Res*.

Kavenoff, R. and Bowen, B.C. (1976) Electron microscopy of membrane-free folded chromosomes from Escherichia coli. *Chromosoma*, **59**, 89-101.

Kavenoff, R. and Ryder, O.A. (1976) Electron microscopy of membrane-associated folded chromosomes of Escherichia coli. *Chromosoma*, **55**, 13-25.

Kidwell, M.G. and Lisch, D.R. (2000) Transposable elements and host genome evolution. *Trends Ecol Evol*, **15**, 95-99.

Kleckner, N. (1981) Transposable elements in prokaryotes. *Annu Rev Genet*, **15**, 341-404.

Koch, C., Vandekerckhove, J. and Kahmann, R. (1988) Escherichia coli host factor for site-specific DNA inversion: cloning and characterization of the fis gene. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 4237-4241.

Kobryn, K., Lavoie, B.D. and Chaconas, G. (1999) Supercoiling-dependent site-specific binding of HU to naked Mu DNA. *J Mol Biol*, **289**, 777-784.

Kobryn, K., Watson, M.A., Allison, R.G. and Chaconas, G. (2002) The Mu three-site synapse: a strained assembly platform in which delivery of the L1 transposase binding site triggers catalytic commitment. *Mol Cell*, **10**, 659-669.

Krause, H.M. and Higgins, N.P. (1986) Positive and negative regulation of the Mu operator by Mu repressor and Escherichia coli integration host factor. *J Biol Chem*, **261**, 3744-3752.

Kuo, C.F., Zou, A.H., Jayaram, M., Getzoff, E. and Harshey, R. (1991) DNA-protein complexes during attachment-site synapsis in Mu DNA transposition. *Embo J*, **10**, 1585-1591.

Laine, B., Kmiecik, D., Sautiere, P., Biserte, G. and Cohen-Solal, M. (1980) Complete amino-acid sequences of DNA-binding proteins HU-1 and HU-2 from Escherichia coli. *European journal of biochemistry / FEBS*, **103**, 447-461.

Lamrani, S., Ranquet, C., Gama, M.J., Nakai, H., Shapiro, J.A., Toussaint, A. and Maenhaut-Michel, G. (1999) Starvation-induced Mucts62-mediated coding sequence fusion: a role for ClpXP, Lon, RpoS and Crp. *Mol Microbiol*, **32**, 327-343.

Lander, E.S., Linton, L.M. and et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Lavoie, B.D. and Chaconas, G. (1990) Immunoelectron microscopic analysis of the A, B, and HU protein content of bacteriophage Mu transpososomes. *J Biol Chem*, **265**, 1623-1627.

Lavoie, B.D. and Chaconas, G. (1993) Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. *Genes Dev*, **7**, 2510-2519.

Lavoie, B.D., Chan, B.S., Allison, R.G. and Chaconas, G. (1991) Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu-host junction of the Mu type 1 transpososome. *Embo J*, **10**, 3051-3059.

Leung, P.C., Teplow, D.B. and Harshey, R.M. (1989) Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. *Nature*, **338**, 656-658.

Levchenko, I., Yamauchi, M. and Baker, T.A. (1997) ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. *Genes Dev*, **11**, 1561-1572.

Levis, R.W., Ganesan, R., Houtchens, K., Tolar, L.A. and Sheen, F.M. (1993) Transposons in place of telomeric repeats at a Drosophila telomere. *Cell*, **75**, 1083-1093.

Lin, L., Bitner, R. and Edlin, G. (1977) Increased reproductive fitness of Escherichia coli lambda lysogens. *Journal of virology*, **21**, 554-559.

Loot, C., Turlan, C., Rousseau, P., Ton-Hoang, B. and Chandler, M. (2002) A target specificity switch in IS911 transposition: the role of the OrfA protein. *Embo J*, **21**, 4172-4182.

Luijsterburg, M.S., Noom, M.C., Wuite, G.J. and Dame, R.T. (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular

perspective. *J Struct Biol*, **156**, 262-272.

Ma, C.H., Kachroo, A.H., Macieszak, A., Chen, T.Y., Guga, P. and Jayaram, M. (2009) Reactions of Cre with methylphosphonate DNA: similarities and contrasts with Flp and vaccinia topoisomerase. *PLoS One*, **4**, e7248.

Macvanin, M. and Adhya, S. (2012) Architectural organization in E. coli nucleoid. *Biochimica et biophysica acta*.

Manna, D., Breier, A.M. and Higgins, N.P. (2004) Microarray analysis of transposition targets in Escherichia coli: the impact of transcription. *Proc Natl Acad Sci U S A*, **101**, 9780-9785.

Maxwell, A., Craigie, R. and Mizuuchi, K. (1987) B protein of bacteriophage mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. *Proc Natl Acad Sci U S A*, **84**, 699-703.

Miller, J.L., Anderson, S.K., Fujita, D.J., Chaconas, G., Baldwin, D.L. and Harshey, R.M. (1984) The nucleotide sequence of the B gene of bacteriophage Mu. *Nucleic Acids Res*, **12**, 8627-8638.

Mills, R.E., Bennett, E.A., Iskow, R.C. and Devine, S.E. (2007) Which transposable elements are active in the human genome? *Trends Genet*, **23**, 183-191.

Mizuuchi, K. (1983) In vitro transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction. *Cell*, **35**, 785-794.

Mizuuchi, K. (1992) Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu Rev Biochem*, **61**, 1011-1051.

Mizuuchi, M., Baker, T.A. and Mizuuchi, K. (1992) Assembly of the active form of the transposase-Mu DNA complex: a critical control point in Mu transposition. *Cell*, **70**, 303-311.

Mizuuchi, M., Baker, T.A. and Mizuuchi, K. (1995) Assembly of phage Mu transpososomes: cooperative transitions assisted by protein and DNA scaffolds. *Cell*, **83**, 375-385.

Mizuuchi, M. and Mizuuchi, K. (1989) Efficient Mu transposition requires interaction of transposase with a DNA sequence at the Mu operator: implications for regulation. *Cell*, **58**, 399-408.

Morgan, G.J., Hatfull, G.F., Casjens, S. and Hendrix, R.W. (2002) Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in Haemophilus, Neisseria and Deinococcus. *J Mol Biol*, **317**, 337-359.

Naigamwalla, D.Z. and Chaconas, G. (1997) A new set of Mu DNA transposition intermediates: alternate pathways of target capture preceding strand transfer. *Embo J*, **16**, 5227-5234.

Nakai, H., Doseeva, V. and Jones, J.M. (2001) Handoff from recombinase to replisome: insights from transposition. *Proc Natl Acad Sci U S A*, **98**, 8247-8254.

Namgoong, S.Y. and Harshey, R.M. (1998) The same two monomers within a MuA tetramer provide the DDE domains for the strand cleavage and strand transfer steps of transposition. *Embo J*, **17**, 3775-3785.

Namgoong, S.Y., Jayaram, M., Kim, K. and Harshey, R.M. (1994) DNA-protein cooperativity in the assembly and stabilization of mu strand transfer complex. Relevance of DNA phasing and att site cleavage. *J Mol Biol*, **238**, 514-527.

Nash, H.A. and Pollock, T.J. (1983) Site-specific recombination of bacteriophage lambda. The change in topological linking number associated with exchange of DNA strands. *J Mol Biol*, **170**, 19-38.

Niki, H. and Hiraga, S. (1998) Polar localization of the replication origin and terminus in Escherichia coli nucleoids during chromosome partitioning. *Genes Dev*, **12**, 1036-1045.

Oram, M. and Pato, M.L. (2004) Mu-like prophage strong gyrase site sequences: analysis of properties required for promoting efficient mu DNA replication. *Journal of Bacteriology*, **186**, 4575-4584.

Pan, C.Q., Finkel, S.E., Cramton, S.E., Feng, J.A., Sigman, D.S. and Johnson, R.C. (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J Mol Biol*, **264**, 675-695.

Pathania, S., Jayaram, M. and Harshey, R.M. (2002) Path of DNA within the Mu transpososome. Transposase interactions bridging two Mu ends and the enhancer trap five DNA supercoils. *Cell*, **109**, 425-436.

Pathania, S., Jayaram, M. and Harshey, R.M. (2003) A unique right end-enhancer complex precedes synapsis of Mu ends: the enhancer is sequestered within the transpososome throughout transposition. *Embo J*, **22**, 3725-3736.

Pato, M.L. (1994) Central location of the Mu strong gyrase binding site is obligatory for optimal rates of replicative transposition. *Proc Natl Acad Sci U S A*, **91**, 7056-7060.

Pato, M.L. (2004) Replication of Mu prophages lacking the central strong gyrase site. *Research in microbiology*, **155**, 553-558.

Pato, M.L. and Banerjee, M. (1996) The Mu strong gyrase-binding site promotes efficient synapsis of the prophage termini. *Mol Microbiol*, **22**, 283-292.

Pato, M.L. and Banerjee, M. (1999) Replacement of the bacteriophage Mu strong gyrase site and

effect on Mu DNA replication. *J Bacteriol*, **181**, 5783-5789.

Pato, M.L., Howe, M.M. and Higgins, N.P. (1990) A DNA gyrase-binding site at the center of the bacteriophage Mu genome is required for efficient replicative transposition. *Proc Natl Acad Sci U S A*, **87**, 8716-8720.

Pato, M.L., Karlok, M., Wall, C. and Higgins, N.P. (1995) Characterization of Mu prophage lacking the central strong gyrase binding site: localization of the block in replication. *J Bacteriol*, **177**, 5937-5942.

Postow, L., Hardy, C.D., Arsuaga, J. and Cozzarelli, N.R. (2004) Topological domain structure of the Escherichia coli chromosome. *Genes Dev*, **18**, 1766-1779.

Puspurs, A.H., Trun, N.J. and Reeve, J.N. (1983) Bacteriophage Mu DNA circularizes following infection of Escherichia coli. *Embo J*, **2**, 345-352.

Rice, P. and Mizuuchi, K. (1995) Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. *Cell*, **82**, 209-220.

Rice, P.A., Yang, S., Mizuuchi, K. and Nash, H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295-1306.

Savilahti, H., Rice, P.A. and Mizuuchi, K. (1995) The phage Mu transpososome core: DNA requirements for assembly and function. *Embo J*, **14**, 4893-4903.

Sawitzke, J.A., Thomason, L.C., Costantino, N., Bubunenko, M., Datta, S. and Court, D.L. (2007) Recombineering: in vivo genetic engineering in E. coli, S. enterica, and beyond. *Methods Enzymol*, **421**, 171-199.

Schneider, R., Lurz, R., Luder, G., Tolksdorf, C., Travers, A. and Muskhelishvili, G. (2001) An architectural role of the Escherichia coli chromatin protein FIS in organising DNA. *Nucleic Acids Res*, **29**, 5107-5114.

Schumacher, S., Clubb, R.T., Cai, M., Mizuuchi, K., Clore, G.M. and Gronenborn, A.M. (1997) Solution structure of the Mu end DNA-binding ibeta subdomain of phage Mu transposase: modular DNA recognition by two tethered domains. *Embo J*, **16**, 7532-7541.

Sinden, R.R. and Pettijohn, D.E. (1981) Chromosomes in living Escherichia coli cells are segregated into domains of supercoiling. *Proc Natl Acad Sci U S A*, **78**, 224-228.

Sokolsky, T.D. and Baker, T.A. (2003) DNA gyrase requirements distinguish the alternate pathways of Mu transposition. *Mol Microbiol*, **47**, 397-409.

Stein, R.A., Deng, S. and Higgins, N.P. (2005) Measuring chromosome dynamics on different time scales using resolvases with varying half-lives. *Mol Microbiol*, **56**, 1049-1061.

Strathmann, M., Hamilton, B.A., Mayeda, C.A., Simon, M.I., Meyerowitz, E.M. and Palazzolo, M.J. (1991) Transposon-facilitated DNA sequencing. *Proc Natl Acad Sci U S A*, **88**, 1247-1250.

Surette, M.G., Buch, S.J. and Chaconas, G. (1987) Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. *Cell*, **49**, 253-262.

Surette, M.G. and Chaconas, G. (1989) A protein factor which reduces the negative supercoiling requirement in the Mu DNA strand transfer reaction is Escherichia coli integration host factor. *J Biol Chem*, **264**, 3028-3034.

Surette, M.G. and Chaconas, G. (1992) The Mu transpositional enhancer can function in trans: requirement of the enhancer for synapsis but not strand cleavage. *Cell*, **68**, 1101-1108.

Surette, M.G., Lavoie, B.D. and Chaconas, G. (1989) Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). *Embo J*, **8**, 3483-3489.

Swinger, K.K. and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol*, **14**, 28-35.

Swinger, K.K. and Rice, P.A. (2007) Structure-based analysis of HU-DNA binding. *J Mol Biol*, **365**, 1005-1016.

Symonds, N., Toussaint, A., Van de Putte, P. and Howe, M. (1987) *Phage Mu*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Taylor, A.L. (1963) Bacteriophage-Induced Mutation in Escherichia Coli. *Proc Natl Acad Sci U S A*, **50**, 1043-1051.

Teplow, D.B., Nakayama, C., Leung, P.C. and Harshey, R.M. (1988) Structure-function relationships in the transposition protein B of bacteriophage Mu. *J Biol Chem*, **263**, 10851-10857.

Thanbichler, M., Wang, S.C. and Shapiro, L. (2005) The bacterial nucleoid: a highly organized and dynamic structure. *Journal of cellular biochemistry*, **96**, 506-521.

Thompson, J.F. and Landy, A. (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic Acids Res*, **16**, 9687-9705.

Tolstorukov, M.Y., Virnik, K.M., Adhya, S. and Zhurkin, V.B. (2005) A-tract clusters may facilitate DNA packaging in bacterial nucleoid. *Nucleic Acids Res*, **33**, 3907-3918.

Travers, A. and Muskhelishvili, G. (2005) Bacterial chromatin. *Curr Opin Genet Dev*, **15**, 507-

514.

Valens, M., Penaud, S., Rossignol, M., Cornet, F. and Boccard, F. (2004) Macrodomain organization of the Escherichia coli chromosome. *Embo J*, **23**, 4330-4341.

van de Putte, P., Giphart-Gassler, M., Goosen, N., Goosen, T. and van Leerdam, E. (1981) Regulation of integration and replication functions of bacteriophage Mu. *Cold Spring Harb Symp Quant Biol*, **45 Pt 1**, 347-353.

van Gent, D.C., Mizuuchi, K. and Gellert, M. (1996) Similarities between initiation of V(D)J recombination and retroviral integration. *Science*, **271**, 1592-1594.

Wang, Z. and Harshey, R.M. (1994) Crucial role for DNA supercoiling in Mu transposition: a kinetic study. *Proc Natl Acad Sci U S A*, **91**, 699-703.

Wang, W., Li, G.W., Chen, C., Xie, X.S. and Zhuang, X. (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*, **333**, 1445-1449.

Wang, Z., Namgoong, S.Y., Zhang, X. and Harshey, R.M. (1996) Kinetic and structural probing of the precleavage synaptic complex (type 0) formed during phage Mu transposition. Action of metal ions and reagents specific to single-stranded DNA. *J Biol Chem*, **271**, 9619-9626.

Watson, M.A., Boocock, M.R. and Stark, W.M. (1996) Rate and selectively of synapsis of res recombination sites by Tn3 resolvase. *J Mol Biol*, **257**, 317-329.

Watson, M.A. and Chaconas, G. (1996) Three-site synapsis during Mu DNA transposition: a critical intermediate preceding engagement of the active site. *Cell*, **85**, 435-445.

Willenbrock, H. and Ussery, D.W. (2004) Chromatin architecture and gene expression in Escherichia coli. *Genome biology*, **5**, 252.

Witz, G. and Stasiak, A. (2010) DNA supercoiling and its role in DNA decatenation and unknotting. *Nucleic Acids Research*, **38**, 2119-2133.

Worcel, A. and Burgi, E. (1972) On the structure of the folded chromosome of Escherichia coli. *Journal of Molecular Biology*, **71**, 127-147.

Wu, Z. and Chaconas, G. (1995) A novel DNA binding and nuclease activity in domain III of Mu transposase: evidence for a catalytic region involved in donor cleavage. *Embo J*, **14**, 3835-3843.

Yanagihara, K. and Mizuuchi, K. (2002) Mismatch-targeted transposition of Mu: a new strategy to map genetic polymorphism. *Proc Natl Acad Sci U S A*, **99**, 11317-11321.

Yin, Z. and Harshey, R.M. (2005) Enhancer-independent Mu transposition from two

topologically distinct synapses. *Proc Natl Acad Sci U S A*, **102**, 18884-18889.

Yin, Z., Jayaram, M., Pathania, S. and Harshey, R.M. (2005) The Mu transposase interwraps distant DNA sites within a functional transpososome in the absence of DNA supercoiling. *J Biol Chem*, **280**, 6149-6156.

Yin, Z., Suzuki, A., Lou, Z., Jayaram, M. and Harshey, R.M. (2007) Interactions of phage Mu enhancer and termini that specify the assembly of a topologically unique interwrapped transpososome. *J Mol Biol*, **372**, 382-396.

Zhou, L., Mitra, R., Atkinson, P.W., Hickman, A.B., Dyda, F. and Craig, N.L. (2004) Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature*, **432**, 995-1001.

Zimmerman, S.B. (2006) Cooperative transitions of isolated Escherichia coli nucleoids: implications for the nucleoid as a cellular phase. *J Struct Biol*, **153**, 160-175.

Zou, A.H., Leung, P.C. and Harshey, R.M. (1991) Transposase contacts with Mu DNA ends. *J Biol Chem*, **266**, 20476-20482.

VITA

Zheng Lou was born in Nanjing, China. After graduated from Nanjing Foreign Language School in Nanjing, he entered Fudan University in Shanghai, China, in 2001. He received the degree of Bachelor of Science in biological sciences in June, 2005. In August, 2005, he entered the Graduate School at the University of Texas at Austin.

Email address: zhenglou@gmail.com

This dissertation was typed by the author.