A TALE OF TWO RECOMBINASES: CHARACTERIZATION OF PIV AND MOOV

by

CHANDRA D. CARPENTER

(Under the Direction of Anna C. Karls)

ABSTRACT

Specialized DNA recombination, DNA transposition and site-specific recombination, creates genetic diversity by altering the genetic information through DNA insertions, deletions, and inversions. The goal of this research was to determine the molecular mechanism by which recombinases of the novel Piv/MooV family mediate specialized recombination. While there are greater than 50 members in the Piv/MooV family, two recombinases were chosen for the focus of this dissertation, Piv (pilin invertase) and MooV (mover of IS*492* in oceanic variants). Piv catalyzes conservative site-specific inversion of a 2.1 kb chromosomal segment that encodes type 4 pili. MooV catalyzes movement of the insertion element, IS*492*, in *Pseudoalteromonas atlantica*. Site-specific insertion and precise excision of IS*492* within *epsG*, the gene for glucosyl transferase, controls phase variable expression of extracellular polysaccharide (EPS). To aid in the determination of the mechanism of Piv-mediated inversion, DNA sequence requirements at the recombination sites for Piv-mediated recombination were defined. In vitro DNA binding assays showed that Piv preferentially binds to single-stranded DNA instead of double-stranded *inv* sequence; in addition, this binding appears to be structure specific rather than sequence specific. To address the amino acid sequence requirements for MooV-mediated recombination, individual substitutions were generated within a DEDD motif in MooV which is

highly conserved in the Piv/MooV family of recombinases and is essential for Piv-mediated inversion. These MooV variants were assayed in vivo for IS*492* excision and insertion; each of the conserved acidic residues was shown to be required for both excision and insertion of IS*492*. A maltose binding protein MooV fusion protein was purified and used in DNA cleavage and strand transfer assays in order to isolate and characterize intermediates in MooV-mediated recombination; however, no in vitro activity of MooV was detected. Therefore, to improve the in vitro activity of MooV, an assay was designed to isolate hyperactive insertion variants of MooV. This approach is likely to allow the elucidation of the mechanism of MooV-mediated recombination of IS*492*.

INDEX WORDS: recombination, Piv, MooV, IS*492*, DEDD, *Moraxella*, *Pseudoalteromonas*

A TALE OF TWO RECOMBINASES: CHARACTERIZATION OF PIV AND MOOV

by

CHANDRA D. CARPENTER

B.S., University of Arkansas, Fayetteville, 2002

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2008

A TALE OF TWO RECOMBINASES: CHARACTERIZATION OF PIV AND MOOV

by

CHANDRA D. CARPENTER

Major Professor:	Anna C. Karls
Committee:	Timothy Hoover
	Sidney Kushner
	Robert Maier
	Ellen Neidle

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2008

DEDICATION

To my mother, Shelbia Willis, who has always believed in me when I needed it most and to my husband, James Carpenter for his unfailing support and encouragement.

ACKNOWLEDGEMENTS

Many people were involved in helping me complete my dissertation. Many thanks to my advisor, Anna Karls; I would not have made it without her wisdom and encouragement. I would like to thank my committee members: Timothy Hoover, Sidney Kushner, Robert Maier, and Ellen Neidle, for their helpful suggestions and support. I would also like to thank John Buchner; he was always willing to answer my questions and discuss recombination with me. I would like to acknowledge the "Thesis Reading Group" (Dawn Adin, John Buchner, Emily Henriksen, and James Henriksen) for their comments and suggestions.

TABLE OF CONTENTS

Page

Chapter 1

Introduction and Literature Review

From the sequencing of bacterial to human genomes, it is now obvious that mobile DNA is ubiquitous and abundant. In humans, mobile elements or remnants of elements constitute nearly half the genome. It is argued that mobile elements drive genome evolution, causing both detrimental and beneficial DNA rearrangements (22). Corresponding examples of DNA rearrangements in humans and bacteria are the deletions, insertions and inversions that generate the diversity of the immunoglobulins in B cells and produce the antigenic and phase variation of virulence factors in pathogenic bacteria. Thus, the progression of infection can be determined by these DNA rearrangements controlling phase or antigenic variation of structures such as capsular polysaccharides, pilin, and flagellar filaments, which enable a pathogen to evade the host immune defenses, as well as, change tissue tropism, or control cell invasion (13). For example, a representative DNA rearrangement with this key role in pathogenesis is the FimB/E-mediated site-specific inversion of a chromosomal segment encoding the promoter for the type I fimbriae gene in uropathogenic *E. coli* (UPEC) (39). The ability to switch on type I fimbriae expression allows UPEC to attach to urogenital epithelium after leaving the gastrointestinal tract and to invade bladder cells, which contributes to persistence in urinary tract infections (30). Rapidly switching off production of type 1 fimbriae facilitates survival of UPEC in the presence of the immune system, such as in infection of the peritoneum (28).

The two primary groups of specialized DNA recombinases that mediate DNA rearrangements in prokaryotic and eukaryotic cells are the Serine (S)- or Tyrosine(Y)- site-specific recombinases and the DDE-motif recombinases, which include retroviral integrases and

DNA transposases. The molecular mechanisms for recombination mediated by representative members of both of these groups have been characterized in exquisite detail and will be reviewed below. My research has focused on specialized recombinases of the DEDD-motif family, which do not fit in either the S/Y site-specific recombinase or the DDE-motif retroviral integrase/transposase groups. While the DEDD-motif recombinases have the structural component that comprises the catalytic domain of the DDE-motif recombinases, a ribonuclease H-like fold, they utilize a DEDD catalytic motif within this fold, suggesting a RuvC-like Holliday junction resolvase activity (37). In addition, the recombination systems of this family that have been defined in the most detail involve site-specific, conservative recombination, indicating that there is more to the reaction mechanism than transposase- or RuvC-like activities.

Specialized recombinases of the S-, Y-, and DDE-motif families

Specialized recombination mediated by site-specific recombinases or transposases alters genetic information through DNA insertions, deletions, and inversions. While site-specific recombinases and transposases both catalyze specialized recombination, the mechanisms for recombination can be vastly different.

Site-specific recombination occurs between two DNA sequences that share a short region of homology (2-12 bp). Site-specific recombination is conservative as no DNA sequence is lost or gained. In addition, host DNA replication/repair enzymes are not required. Site-specific recombination involves a two-step trans-esterification reaction in which a covalent recombinase-DNA intermediate is formed (Figure 1.1). The recombinases that mediate this reaction are divided into two families, Tyrosine (Y)-recombinases and Serine (S)-recombinases, based on the conserved amino acid involved in the covalent recombinase-DNA intermediate (3, 21). The covalent recombinase-DNA intermediate results from nucleophilic attack on the DNA

phosphodiester backbone by the hydroxyl group of the conserved tyrosine or serine of the recombinase. In the case of the Y-recombinase, there is a covalent DNA-3'phosphotyrosine intermediate (3). For the S-recombinase there is a covalent DNA-5'phosphoserine intermediate (21). The second step of the two-step trans-esterification reaction results in the restoration of the phosphodiester linkages of the exchanged DNA strands. This restoration occurs in the nucleophilic attack on the phospho-seryl or phospho-tyrosyl bonds by the 5'OH- or 3'OH-terminated DNA strand that was liberated in the first trans-esterification reaction (3, 21).

Transposition is the movement of a DNA segment to a non-homologous site, i.e. the ends of the transposable element that are recognized by the transposase show no homology with the target site for insertion. Transposition in prokaryotes and eukaryotes is primarily mediated by DDE-motif (classical) transposases. The conserved DDE (aspartic acid, aspartic acid, and glutamic acid) amino acid triad of the classical transposases comprises the catalytic pocket that binds metal cofactors, such as magnesium, to facilitate H_2O-mediated hydrolysis of donor DNA and coordinate strand transfer to target DNA (15). Transposition that is mediated by classical transposases is not conservative. Insertion of the transposable element results in duplication of target sequence and excision can result in DNA deletions in flanking sequences.

The process of DNA transposition can be divided into three steps: 1) cleavage of the donor site resulting in detachment of the transposon from the donor DNA, 2) transfer of the cleaved transposon ends to a target site, and 3) processing of the products by host-encoded enzymes (15). Transposition can result in excision of the transposable element from the donor site and insertion into a new site leaving a double-strand break in the DNA ("cut-and-paste" transposition). Alternatively, it can involve replication of the transposable element, thus leaving a copy at the donor site while inserting a copy at the target site (replicative transposition)

(Figures 1.2 and 1.3). Cut-and-paste versus replicative transposition is dependent on whether nicking at the 5' ends of the element occurs in either the first step or following the second step of transposition. 1) In the first step, hydrolysis of the DNA phosphate backbone at the 3' ends of the transposable element is mediated by the transposase (15). This cleavage leaves a 3'OH at each end of the element, which in cut-and-paste transposition can act as nucleophiles to attack the 5' end of the transposable element, thus creating hairpin ends on the element and releasing it from the donor DNA, as characterized in the transposition of Tn*10* and Tn*5* (5, 23). The hairpin is cleaved by transposase-mediated hydrolysis, leaving a 3'OH at each end of the element. The 5' ends of the element can be cleaved from the donor by a number of other mechanisms, but the key feature is that the whole element is released from the donor. 2) However, in replicative transposition, after cleavage at the 3' ends of the element at the donor site, the 3'OH ends act as the nucleophiles in the staggered attack at the target sequence (15). This strand transfer reaction creates new phosphodiester bonds between the opposing strands of the transposon and the target DNA and leaves 3'OH groups in the target DNA at the point of insertion.

An example of replicative transposition is the transposition of bacteriophage Mu. In Mu transposition, the transposase only nicks at the 3'ends of Mu, resulting in the element remaining connected to the donor backbone by its 5'end (7). A replication fork is created at each end of Mu and DNA polymerase extends from the 3'OH ends of the cleaved target site, thus the Mu DNA is replicated such that a cointegrate is formed containing the donor DNA and the target DNA linked by two copies of Mu (7). In the final step of both cut-and-paste and replicative transposition, host repair/replication machinery fills in the gaps created at the target site by the transposase-mediated staggered attack in the strand transfer step. The addition of DNA in this step shows the non-conservative nature of transposition (7).

DDE-motif transposases are the most common transposases involved in transposition; however, the mechanism of transposition mediated by DDE-motif transposases is not universal. Other transposases use alternative reaction mechanisms [for review see (9)]. The Y-, S-, and Y2-transposases do not contain a conserved DDE-motif, but have conserved serines or tyrosines that become covalently linked to DNA as an intermediate in the recombination reaction, similar to site-specific recombinases. The reaction mediated by these transposases is still considered transposition because they have multiple insertion sites that do not share sequence homology with the mobile element (9).

To give a short review of the differences between transposition and site-specific recombination, transposition involves the exchange of DNA between two non-homologous sites, while site-specific recombination involves the exchange of DNA between two sites with a limited region of homology. Transposition is often non-conservative and requires host DNA repair/replication machinery to repair regions in which the transposable element has created nicks or gaps, while site-specific recombination is conservative and does not require DNA repair/replication machinery to repair any gaps or nicks. DDE-motif transposases utilize H_2O, coordinated by divalent metal ions, as the attacking nucleophile in DNA cleavage, while site-specific recombinases utilize a conserved tyrosine or serine as the attacking nucleophile in DNA cleavage. DDE-motif transposases do not form covalent intermediates with donor DNA, but site-specific recombinases do form covalent DNA-recombinase intermediates in the recombination reaction.

Piv/MooV (DEDD-motif) family

Recombinases of the Piv/MooV family insert and excise IS elements of the IS*110* family or, as is the case with Piv, invert a segment of DNA. The recombination reactions mediated by these recombinases appear to be site-specific because the reaction is conservative (*i.e.* no DNA sequence is added or deleted). Although, these recombinases appear to be site-specific in their reaction, they are not related to the S- or Y-recombinases that catalyze site-specific recombination. Instead, based on primary amino acid sequence alignment, the majority of the recombinases of the Piv/MooV family have a conserved tetrad, DEDD, that is thought to constitute the catalytic site (6, 25). While every member of the family has 4 conserved acidic residues, a few members have a DDDD motif (25). The conserved tetrad is predicted to be arranged in the ribonuclease H fold motif that defines the catalytic domains of DDE-motif transposases and the DEDD-motif Holliday-junction resolvases (6, 40). This suggests that the recombinases mediate recombination through a hydrolysis and one-step transesterification reaction similar to the mechanism of transposition utilized by DDE-motif transposases.

Members of the DEDD-motif family are found in human pathogens (e.g. *Neisseria gonorrhoeae, Neisseria meningitis, Moraxella lacunata, Mycobacterium tuberculosis*, *Bordetella pertussis, Klebsiella pneumoniae, Shigella flexneri*), plant pathogens (*Agrobacterium tumefaciens, Erwinia chrysanthemi*), and bacteria that are important in soil and fresh water/marine environments (e.g. *Streptomyces coelicolor, Acinetobacter* species, *Pseudoalteromonas atlantica*). Although these recombinases direct DNA inversion and mobile element insertion and excision in at least 50 bacterial genera, the activity of only a few DEDD recombinases have been characterized to any extent. This introduction will focus on the two most extensively studied members of the family, MooV and Piv, while incorporating relevant

features of other members of the family. MooV mediates site-specific insertion and precise excision of the insertion sequence IS*492*, which directs phase variation of peripheral extracellular polysaccharide production in *P. atlantica*, and Piv is a DNA invertase controlling phase and antigenic variation of type 4 pili in *Moraxella lacunata* and *Moraxella bovis*, respectively (4, 19, 26, 27).

IS elements whose transposases are DEDD-motif recombinases

The insertion sequences that encode recombinases of the Piv/MooV family belong to the IS*110*/IS*492* family of insertion sequences. It has been suggested that the IS*110*/IS*492* family can be divided into two families, IS*110* and IS*1111*, based on sequences of the potential transposases of the members (24). However, because they are clearly linked by BLAST analysis, both groups belong to the IS*110*/IS*492* family. Members of this family have characteristics unusual to insertion sequences in that they lack terminal inverted repeats and generally do not duplicate target sequence upon insertion. The exception to these characteristics is found in the IS*1111* group; the elements closely related to IS*1111* have short, terminal or sub-terminal, and sometimes imperfect inverted repeats (24). The following section describes the precise excision of the elements, the site-preference for insertion of the elements, and possible reaction intermediates.

Excision of the IS elements

Two elements from the IS*110*/IS*492* family have been shown to excise precisely from donor DNA, IS*4321* and IS*492* (19, 32, 33). IS*4321* was first discovered in R751, a multidrug resistance plasmid, within one of the terminal inverted repeats (TIR) of Tn*4321* (32). IS*4321* belongs to the IS*1111* group within the family, and IS*4321* has sub-terminal inverted repeats that start 6-7 bp from the left end and 3-4 bp from the right end. Excision of IS*4321* from a donor

plasmid was detected in vivo, and DNA sequence data revealed that IS*4321* had precisely excised to recreate the 38 bp TIR of Tn*4321* (32).

IS*492*, another member of the family shown to excise precisely, controls the expression of peripheral extracelluar polysaccharide (^{p}EPS) in the marine bacterium *Pseudoalteromonas atlantica* by site-specific insertion and excision within a predicted glucosyl-transferase gene (*epsG*) (4, 19). Excision of IS*492* from *epsG* is precise, restoring the *epsG* sequence. Not only is the excision from *epsG* precise, it occurs at a very high frequency (2×10^{-2} per cell per generation on solid medium), which has been correlated to the level of transcription that initiates upstream of the element and passes through the *mooV* gene (19). Unlike IS*4321*, IS*492* does not contain terminal or sub-terminal inverted repeats. Even though IS*492* and IS*4321* belong to different sub groups within the IS*110*/IS*492* family, both have been shown to excise precisely from donor DNA. Precise excision from donor DNA is a hallmark of site-specific recombination, but as discussed earlier, the recombinases from this family of IS elements are not related to serine or tyrosine site-specific recombinases.

Insertion of the IS elements

Several members of IS*110*/IS*492* family show some site-preference for insertion. In fact, IS*117* has a preferred target site, but when this preferred target site is removed, IS*117* will insert at a lower frequency into secondary sites that have some sequence similarity with the preferred site (16, 38). Another element that shows a similar degree of site-preference is IS*492* which is found in at least 5 copies on the *P. atlantica* chromosome with the same 5 bp flanking sequence on the left end and 7 bp on the right end at all 5 locations (33). Interestingly, there is not much sequence similarity of the flanking sequence past the shared 5 bp and 7 bp between the 5 copies. Other elements of the family do not insert into just one sequence, but some elements have

insertion sites that are similar enough in sequence to be able to determine a consensus target site based on sequence alignment. (e.g. IS*Mpa1*, IS*900*, IS*902*, IS*Ec11*, and IS*1383*) (14, 24, 29, 31, 34).

Other members of the IS*110*/IS*492* family not only show site-preference, but also insert into repeated DNA sequences. For example, IS*621* and IS*Ppu10* are found in repetitive extragenic palindromes (REPs) (8, 35). IS*621* can insert into two different types of REPs, z1-type and z2 type, but the element is found at the same specific site in both REPs ($^{5'}$ CTTATCAGGC*CTAC, where * indicates insertion site) (8).

Some members of the IS*110*/IS*492* family do not target repeat sequences but do target other transposons. IS*Enfa110* was discovered in Tn*5382* and IS*5075* is found in 38 bp terminal inverted repeats (TIR) of transposons of the Tn*501*/Tn*21* family (10, 32). IS*5075* inserts at the same position in the same orientation in the 38 bp TIR, showing target site preference. IS*5075* is not the only element in this family to insert at a preferred site in the same orientation.

IS*492*, IS*110*, IS*900*, IS*Ec11*, and IS*1383* are just a few examples of IS elements that appear to insert within their target sites in the same orientation. When viewing the sequence flanking the inserted IS element, the asymmetrical nature of the target site can be seen. It has been suggested that the asymmetric structure of the target sites could be the reason that the elements insert in only one orientation (34).

IS*492*	CTTGT*TA
IS*110*	GGCACCCCCCC (insertion is somewhere in the string of C's)
IS*900*	$CATGN_{(4-6)}$*CNCCTT
IS*Ec11*	GTNAAAA*NANTG
IS*1383*	TTCAGATGGT*ATAAG

Thus far, I have discussed the degree of site-specificity of the elements, the targeting of repeated sequence and transposons, and the asymmetric nature of the target site. However, analysis of insertion sites of the elements by sequence or by Southern blot has shown that some of the elements can insert in tandem array.

IS*117* has been shown to insert next to itself forming a head-to-tail dimer; the sequence between each element is the same sequence as the junction of the circular form of the element (17). IS*1000* was also shown to be inserted as tandem repeats (2-3) into the chromosome of *Thermus thermophilus*, but, unlike IS*117*, there can be as much as 256 bp separating the IS elements (2). Preliminary data has also shown IS*492* inserted as a head-to-tail dimer with the same junction sequence between the elements as found on the circular form of the element (A. Cottrell, unpublished data).

Some elements have been proposed to move either replicatively or by cut-and-paste transposition with double-strand break repair from a sister chromosome such that the element would be found at both the donor site and the target site. For example, based on Southern blot analyses of EPS^- and EPS^+ phase variants of *P. atlantica*, the insertion of IS*492* into the *eps* site does not result in a loss of one of the other copies of IS*492* on the chromosome (4). Similar results have been seen with IS*1110* (18). In the case of IS*900*, when used to engineer a transposon with IS*900* flanking a kanamycin resistance gene, the engineered transposon has been shown to form cointegrates based on Southern blots which may have resulted from replicative integration (11).

How do the members of the IS*110*/IS*492* family target a specific DNA sequence? In the case of several members, such as IS*492*, their insertion results in what appears to be a target site duplication, but one of the duplicated sequences may actually be part of the element. The

element's recombinase may recognize and bind as a multimer to one of the duplicated sequences on the element and interact with matching sequence in the target site before insertion occurs. However, this similarity in sequences between the elements (if the duplicated sequence is carried with them) and their target sites does not occur in all elements in the family. Several members of the family are not flanked by direct repeats upon insertion. An important question is whether recombinases of this family that are not flanked by direct repeats, such as the IS*900* recombinase, have a binding site at the end of the element and a different binding site at the preferred target site.

One unique member of the Piv/MooV family of recombinases, Piv, does recognize and bind to two different sites. Piv is the pilin invertase found in *Moraxella bovis* and *M. lacunata* that catalyzes the inversion of a pilin segment. In *M. bovis* the inversion results in Q/I antigenic variation of the pilin, but in *M. lacunata* the inversion results in ON/OFF phase variation of the pilin (26, 27). Piv from *M. lacunata* and cloned as a fusion protein to MBP, binds two different DNA sequences, one strong upstream binding site (*sub*) is found at multiple sites within and outside of the invertible segment and the second binding site is within the invertible repeats (*inv*) at the ends of the pilin invertible segment, making it a bivalent protein (41). Perhaps the recombinases of the IS*110*/IS*492* family of insertion elements are also bivalent proteins which may explain their ability to target a specific DNA sequence that is different than the IS element being transposed.

Piv DNA invertase

The pilin invertase, Piv, catalyzes the inversion of a 2.1 kb chromosomal segment that alters the expression of type 4 pili in the human and bovine eye pathogens, *M. lacunata* and *M. bovis*, respectively (36). The invertible segment of *M. bovis* contains the pilin genes, *tfpQ* and

tfpI, and an additional open reading frame, *tfpB*, which is not required for inversion (25). The invertible segment found in *M. lacunata* is highly similar the pilin segment of *M. bovis*, but a notable difference is a 19 bp duplication in the *tfpI* gene, which results in expression of a truncated, non-functional pilin subunit (36). Therefore, *M. lacunata* exhibits an on/off pilin phase variation, while *M. bovis* alternately expresses Q and I type 4 pilin.

Unlike the IS elements whose recombinases belong to the Piv/MooV family, the pilin inversion segment is flanked by a 32 bp terminal inverted repeat (*invL* and *invR*). Piv, as a fusion to maltose binding protein (MBP), has been shown by electrophoretic mobility shift assays (EMSA) to bind to *invL*. The MBP-Piv protein also binds a strong upstream binding site (*sub1*) and additional *sub* sites within *piv* and the invertible element (41). MBP-Piv has a much higher affinity for *sub1* than it does for *invL*, and MBP-Piv interactions with *sub1* are quite different than with *invL*. The *sub1* site is bound cooperatively by two protomers of MBP-Piv; however, the *invL* site appears to only be bound by a multimeric form of MBP-Piv. This binding activity is not unusual for a bivalent protein that mediates DNA recombination. Other bivalent recombinases, such as bacteriophage lambda integrase, bind accessory sites at much higher affinities than the recombination site, or as in the case of the MuA transposase, bind one type of site as a monomer and then assemble as multimers before binding the ends of the mobile element [for review see (3, 7)]. It is thought that Piv binds the *sub* sites and then assembles into the multimeric form that binds to the recombination sites forming the synaptic complex. The formation of a synaptic complex assures that the recombinase is not active until a synaptic structure is assembled in order to prevent incorrect strand nicking and exchange reactions (41)

Amino acid substitutions within Piv indicate that the conserved tetrad, DEDD, is required for Piv-mediated inversion, but not for Piv binding, signifying the possibility that the DEDD

tetrad is a catalytic motif within Piv. Molecular modeling of the tertiary structure for the amino terminal region of Piv predicts an RNase H-like fold, found in DDE-motif transposases, with the completely conserved D9, E59, D101, and D104 positioned appropriately to coordinate two divalent metal ions (6, 40). The RuvC Holliday junction resolvases also utilize a DEDD catalytic tetrad within an RNase H-like fold for coordination of two divalent metal cations (1).

Interestingly, Piv contains a possible helix-hairpin-helix motif that is used by RuvC Holliday junction resolvases to bind the junction. Preliminary data suggest that Piv is capable of binding a Holliday junction in vitro (J. Buchner, personal communication). The current working model for Piv-catalyzed inversion involves a Holliday junction intermediate structure (6). The model in Figure 1.4 shows Piv bound to the synaptic complex as a dimer. Piv then coordinates a divalent cation to initiate hydrolysis of one strand. The released 3'OH then attacks the target DNA setting up a Holliday junction intermediate which is then cleaved by Piv in a second hydrolysis reaction using the catalytic DEDD-motif. Repair of the nicked DNA strands may occur by DNA ligase. This working model is currently being tested by the use of synthetic Holliday junctions.

While there are examples of DDE recombinases that contribute to site-specific recombination events, no characterized DDE recombinase catalyzes <u>conservative</u> site-specific recombination (12, 20). For example, RAG1/RAG2 is required for double strand cleavage at specific sites in immunoglobulin gene rearrangement; however, RAG1/RAG2 carries out only the cleavage, while the cellular non-homologous end joining functions repair the double-strand breaks to yield the rearranged gene. This repair process may delete or add base pairs at the recombined junction, thus recombination is not conservative (12). In contrast, the DEDD-motif transposase MooV of the Piv/MooV family mediates recombinase-dependent, <u>precise</u> excision of

IS*492*. As shown in Figure 1.5, precise excision of IS*110*-related elements, including IS*492*, cannot be accomplished by coordinated hydrolysis and one-step transesterification reactions because the donor DNA is left with either double nicks separated by a few base pairs (if there is a direct repeat flanking the element as with IS*492*) or a double strand break (if there is no direct repeat). It may be possible for the IS*110*-related elements to move by hydrolysis and one-step transesterification reactions if the recombinases hold the ends together and direct DNA ligase to seal the nicks, however, this would suggest that the transposases recognize different flanking host sequences. Similarly, the DEDD invertase Piv catalyzes intramolecular and intermolecular recombination between inversion sites (*invL/invR*) yielding products with no sequence changes in the exchanged DNA strands. Consequently, characterization of the Piv/MooV recombinases that are structurally similar to the DDE transposes, yet mediate apparently conservative, site-specific recombination, will explore a new area within the field of DNA recombination.

As described in this chapter, little is known about the mechanism of recombination mediated by the recombinases of the Piv/MooV family. The goal of my research has been to characterize defining members of the family, Piv, a site-specific invertase, and MooV, a recombinase that excises precisely and inserts its mobile element, IS*492*, in a site-specific manner. The research described in the following chapters of this dissertation focuses on the characterization of Piv-mediated inversion and MooV-mediated movement of IS*492* both in vivo, by mutational analysis of the recombinases and their DNA substrates, and in vitro, by isolating and characterizing recombination intermediates. In Chapter 2, the DNA sequence requirements for the Piv recombination site are further characterized. This chapter also contains studies examining the single-stranded DNA binding activity of Piv. The conserved DEDD-motif in MooV is shown to be required for the insertion, excision, and circularization of IS*492* in

Chapter 3. An assay for the isolation of hyperactive insertion variants of MooV to be used in in vitro binding, cleavage, and strand transfer assays is described in Chapter 3. Chapter 4 describes a variety of biochemical assays used to define the mechanism of MooV-mediated recombination of IS*492*. The results of these assays revealed that the system is intractable using wild-type MooV and it is proposed that isolation of hyperactive insertion variants of MooV will be necessary to detect in vitro activity or that there is a cofactor missing from the in vitro assays that is required for MooV-mediated recombination.

References

1. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, A. Fujishima, H. Shinagawa, and K. Morikawa.** 1994. Preliminary crystallographic study of *Escherichia coli* RuvC protein. An endonuclease specific for Holliday junctions. J Mol Biol **241:**281-2.
2. **Ashby, M. K., and P. L. Bergquist.** 1990. Cloning and sequence of IS*1000*, a putative insertion sequence from *Thermus thermophilus* HB8. Plasmid **24:**1-11.
3. **Azaro, M. A., and A. Landy.** 2002. Lambda Integrase and the Lambda Int family, p. 118-48. *In* N. Craig, Craigie R, Gellert M, Lambowitz, AM (ed.), Mobile DNA II. ASM Press, Washington D. C.
4. **Bartlett, D. H., M. E. Wright, and M. Silverman.** 1988. Variable expression of extracellular polysaccharide in the marine bacterium *Pseudomonas atlantica* is controlled by genome rearrangement. Proc Natl Acad Sci U S A **85:**3923-27.
5. **Bhasin, A., I. Y. Goryshin, and W. S. Reznikoff.** 1999. Hairpin formation in Tn*5* transposition. J Biol Chem **274:**37021-9.
6. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J Bacteriol **187:**3431-7.
7. **Chaconas, G., and R. M. Harshey.** 2002. Transposition of Phage Mu DNA. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II, vol. 384-402. ASM Press, Washington D.C.
8. **Choi, S., S. Ohta, and E. Ohtsubo.** 2003. A novel IS element, IS*621*, of the IS*110*/IS*492* family transposes to a specific site in repetitive extragenic palindromic sequences in *Escherichia coli*. J Bacteriol **185:**4891-900.

9. **Curcio, M. J., and K. M. Derbyshire.** 2003. The outs and ins of transposition: from mu to kangaroo. Nat Rev Mol Cell Biol **4:**865-77.

10. **Dahl, K. H., E. W. Lundblad, T. P. Rokenes, O. Olsvik, and A. Sundsfjord.** 2000. Genetic linkage of the *vanB2* gene cluster to Tn*5382* in vancomycin-resistant enterococci and characterization of two novel insertion sequences. Microbiology **146 (Pt 6):**1469-79.

11. **England, P. M., S. Wall, and J. McFadden.** 1991. IS*900*-promoted stable integration of a foreign gene into mycobacteria. Mol Microbiol **5:**2047-52.

12. **Gellert, M.** 2002. V(D)J Recombination, p. 705-29. *In* N. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.

13. **Glasgow-Karls, A.** 2001. Alternation of Gene Expression, The Encyclopedia of Genetics Academic Press.

14. **Green, E. P., M. L. Tizard, M. T. Moss, J. Thompson, D. J. Winterbourne, J. J. McFadden, and J. Hermon-Taylor.** 1989. Sequence and characteristics of IS*900*, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. Nucleic Acids Res **17:**9063-73.

15. **Haren, L., B. Ton-Hoang, and M. Chandler.** 1999. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol **53:**245-81.

16. **Henderson, D. J., D. F. Brolle, T. Kieser, R. E. Melton, and D. A. Hopwood.** 1990. Transposition of IS*117* (the *Streptomyces coelicolor* A 3 (2) mini-circle) to and from a cloned target site and into secondary chromosomal sites. Mol Gen Genet **224:**65-71.

17. **Henderson, D. J., D. J. Lydiate, and D. A. Hopwood.** 1989. Structural and functional analysis of the mini-circle, a transposable element of *Streptomyces coelicolor* A3(2). Mol Microbiol **3:**1307-18.

18. **Hernandez Perez, M., N. G. Fomukong, T. Hellyer, I. N. Brown, and J. W. Dale.** 1994. Characterization of IS*1110*, a highly mobile genetic element from *Mycobacterium avium*. Mol Microbiol **12:**717-24.

19. **Higgins, B. P., C. D. Carpenter, and A. C. Karls.** 2007. Chromosomal context directs high-frequency precise excision of IS*492* in *Pseudoalteromonas atlantica*. Proc Natl Acad Sci U S A **104:**1901-6.

20. **Izsvak, Z., E. E. Stuwe, D. Fiedler, A. Katzer, P. A. Jeggo, and Z. Ivics.** 2004. Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. Mol Cell **13:**279-90.

21. **Johnson, R. C.** 2002. Bacterial Site-Specific DNA Inversion, p. 230-71. *In* N. Craig, Craigie R, Gellert M, Lambowitz, AM (ed.), Mobile DNA II. ASM Press, Washington D. C.

22. **Kazazian, H. H., Jr.** 2004. Mobile elements: drivers of genome evolution. Science **303:**1626-32.

23. **Kennedy, A. K., A. Guhathakurta, N. Kleckner, and D. B. Haniford.** 1998. Tn10 transposition via a DNA hairpin intermediate. Cell **95:**125-34.

24. **Lauf, U., C. Muller, and H. Herrmann.** 1999. Identification and characterisation of IS*1383*, a new insertion sequence isolated from *Pseudomonas putida* strain H. FEMS Microbiol Lett **170:**407-12.

25. **Lenich, A. G., and A. C. Glasgow.** 1994. Amino acid sequence homology between Piv, an essential protein in site-specific DNA inversion in *Moraxella lacunata*, and transposases of an unusual family of insertion elements. J Bacteriol **176:**4160-4.

26. **Marrs, C. F., F. W. Rozsa, M. Hackel, S. P. Stevens, and A. C. Glasgow.** 1990. Identification, cloning, and sequencing of *piv*, a new gene involved in inverting the pilin genes of *Moraxella lacunata*. J Bacteriol **172:**4370-7.

27. **Marrs, C. F., W. W. Ruehl, G. K. Schoolnik, and S. Falkow.** 1988. Pilin-gene phase variation of *Moraxella bovis* is caused by an inversion of the pilin genes. J Bacteriol **170:**3032-9.

28. **May, A. K., C. A. Bloch, R. G. Sawyer, M. D. Spengler, and T. L. Pruett.** 1993. Enhanced virulence of *Escherichia coli* bearing a site-targeted mutation in the major structural subunit of type 1 fimbriae. Infect Immun **61:**1667-73.

29. **Moss, M. T., Z. P. Malik, M. L. Tizard, E. P. Green, J. D. Sanderson, and J. Hermon-Taylor.** 1992. IS*902*, an insertion element of the chronic-enteritis-causing *Mycobacterium avium* subsp. *silvaticum*. J Gen Microbiol **138:**139-45.

30. **Mulvey, M. A.** 2002. Adhesion and entry of uropathogenic *Escherichia coli*. Cell Microbiol **4:**257-71.

31. **Olsen, I., T. B. Johansen, H. Billman-Jacobe, S. F. Nilsen, and B. Djonne.** 2004. A novel IS element, IS*Mpa1*, in *Mycobacterium avium* subsp. *paratuberculosis*. Vet Microbiol **98:**297-306.

32. **Partridge, S. R., and R. M. Hall.** 2003. The IS*1111* family members IS*4321* and IS*5075* have subterminal inverted repeats and target the terminal inverted repeats of Tn*21* family transposons. J Bacteriol **185:**6371-84.

33. **Perkins-Balding, D., G. Duval-Valentin, and A. C. Glasgow.** 1999. Excision of IS*492* requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. J Bacteriol **181:**4937-48.

34. **Prosseda, G., M. C. Latella, M. Casalino, M. Nicoletti, S. Michienzi, and B. Colonna.** 2006. Plasticity of the P junc promoter of IS*Ec11*, a new insertion sequence of the IS*1111* family. J Bacteriol **188:**4681-9.

35. **Ramos-Gonzalez, M. I., M. J. Campos, J. L. Ramos, and M. Espinosa-Urgel.** 2006. Characterization of the *Pseudomonas putida* mobile genetic element IS*Ppu10*: an occupant of repetitive extragenic palindromic sequences. J Bacteriol **188:**37-44.

36. **Rozsa, F. W., and C. F. Marrs.** 1991. Interesting sequence differences between the pilin gene inversion regions of *Moraxella lacunata* ATCC 17956 and *Moraxella bovis* Epp63. J Bacteriol **173:**4000-6.

37. **Sharples, G. J.** 2001. The X philes: structure-specific endonucleases that resolve Holliday junctions. Mol Microbiol **39:**823-34.

38. **Smokvina, T., and D. A. Hopwood.** 1993. Analysis of secondary integration sites for IS*117* in *Streptomyces lividans* and their role in the generation of chromosomal deletions. Mol Gen Genet **239:**90-6.

39. **Sussman, M., and D. L. Gally.** 1999. The biology of cystitis: host and bacterial factors. Annu Rev Med **50:**149-58.

40. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. Mol Microbiol **39:**641-51.

41. **Tobiason, D. M., A. G. Lenich, and A. C. Glasgow.** 1999. Multiple DNA binding activities of the novel site-specific recombinase, Piv, from *Moraxella lacunata*. J Biol Chem **274:**9698-706.

Figure 1.1: Diagram of serine-recombinase pathway. The red and blue strands represent the sites of inversion with the core region represented by black dashes. The gray circles represent the serine-recombinase. Black arrows indicate nucleophilic attack. Site-specific recombination by a serine DNA recombinase, involves a two-step transesterification reaction in which a covalent recombinase-DNA intermediate is formed. The intermediate results from nucleophilic attack on the DNA phosphodiester backbone by a hydroxyl group of the conserved serine of the recombinase, forming a covalent DNA-5'phosphoserine intermediate. The restoration of the phosphodiester linkages of the exchanged DNA strands results from nucleophilic attack by the 3'OH-terminated DNA strand that has been liberated by the recombinase.

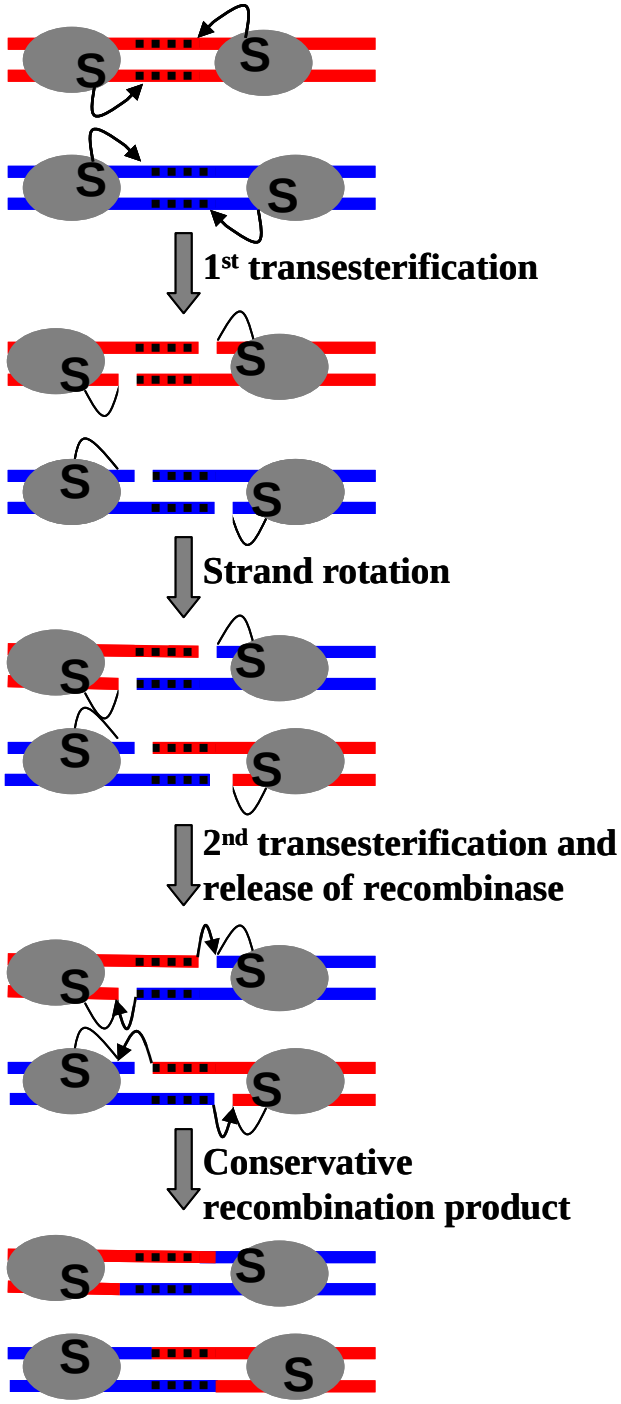

S
S
S
S
1st transesterification
S
S
S
S
Strand rotation
S
S
S
S
2nd transesterification and
release of recombinase
S
S
S
S
Conservative
recombination product
S
S
S
S

Figure 1.2: Diagram of cut-and-paste transposition by DDE-motif transposases. The black arrows indicate transposase-mediated nicking. The black lines indicate donor DNA, the red lines indicate the mobile element, and the blue lines indicate target DNA. The green lines indicate complementary sequence created by host DNA repair/replication machinery. In cut-and-paste transposition, the transposase mediates hydrolysis of the DNA backbone at the 3'end of the mobile element, liberating a 3'OH. The transposase also nicks the DNA backbone at the 5'end, thereby releasing the entire mobile element from the donor DNA backbone. Next, the transposase mediates strand transfer by using the liberated 3'OH as the nucleophile in attacking the target DNA. In the final step of transposition, host DNA repair/replication machinery repairs the regions in which the transposon has created gaps and/or nicks at the DNA insertion site.

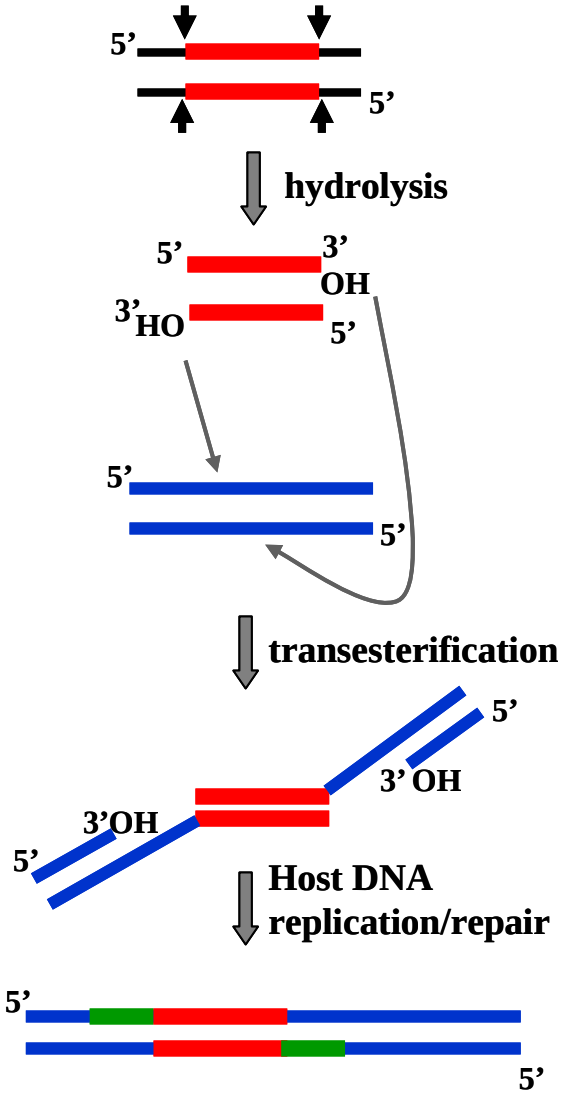

5’
5’
hydrolysis
5’
3’
OH
3’
HO
5’
5’
5’
transesterification
5’
3’ OH
3’OH
5’
Host DNA
replication/repair
5’
5’

Figure 1.3: Diagram of replicative transposition by DDE-motif transposases. The black arrows indicate transposase-mediated nicking. The black lines indicate donor DNA, the red lines indicate the mobile element, and the blue lines indicate target DNA. Green lines indicate complementary sequence generated by DNA replication. In replicative transposition, the transposase-mediated hydrolysis of the DNA backbone at the 3'end of the mobile element liberates a 3'OH, leaving the 5'end of the element attached to the donor backbone. The liberated 3'OH acts as the nucleophile in attacking the target in strand transfer. Because the 5'end of the mobile element is still attached to the donor backbone, strand transfer creates a replication fork in which the DNA polymerase extends from the exposed 3'OH ends on the target DNA.

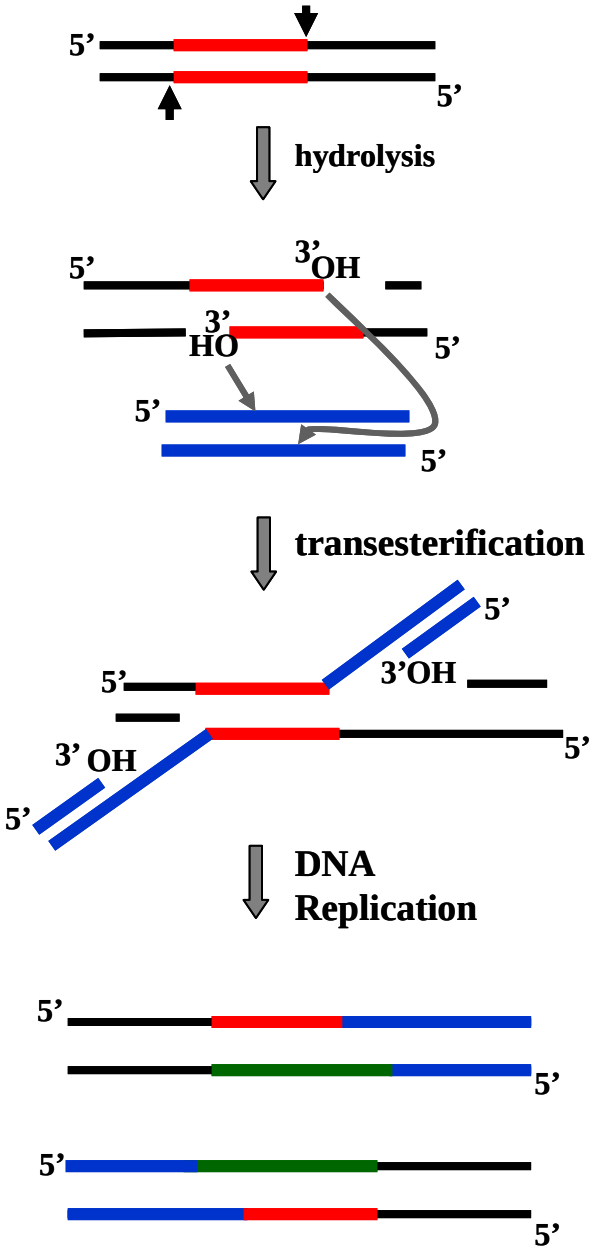
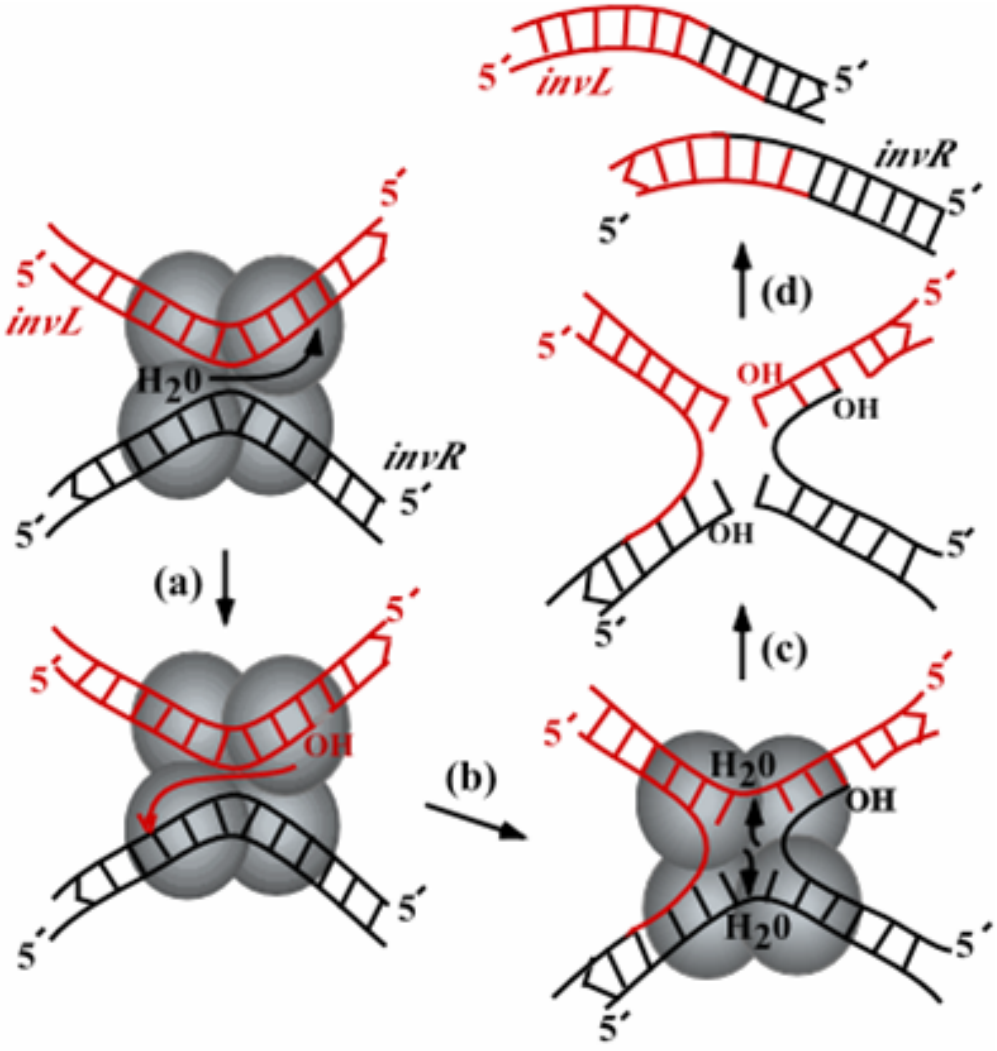
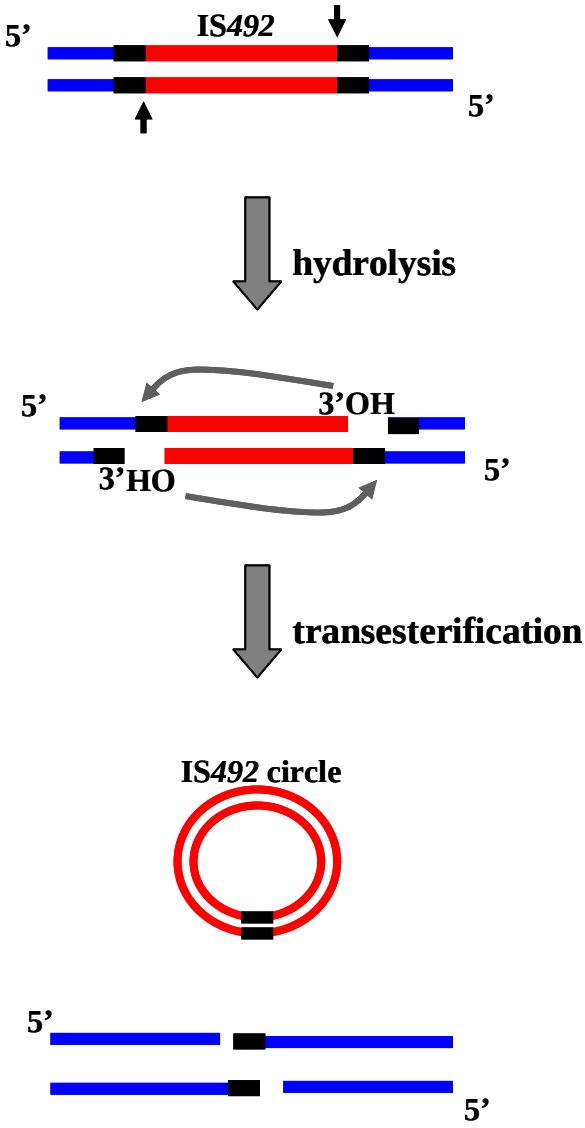

5’
5’
hydrolysis
3’OH
5’
3’
HO
5’
5’
5’
transesterification
5’
5’
3’OH
5’
3’ OH
5’
DNA
Replication
5’
5’
5’
5’

Figure 1.4: Working model for Piv-mediated inversion. The red lines indicate *invL* and the black lines indicate *invR*. Piv is represented as a gray sphere. In this model, Piv is shown bound as a dimer to both inverted repeats. After Piv mediates hydrolysis at one recombination site (depicted as *invL*) (a), Piv uses the released 3'OH as the attacking nucleophile in strand transfer (b) forming a Holliday junction structure. Repositioned Piv catalytic sites then cleave the outer strands of the junction (c) and the nicks are repaired by host DNA ligase (d).



Courtesy of A. Karls

Figure 1.5: Diagram of hydrolysis and one-step transesterification of IS*492*. Red lines indicate IS*492* sequence, black lines indicate 5 bp repeats, and blue lines indicate donor sequence. Black arrows indicate MooV-mediated hydrolysis of the DNA backbone releasing 3'OH. The 3'OH are used as nucleophiles in strand transfer, releasing an IS*492* circle, but leaving behind double-stranded breaks at the donor site.



Double-stranded breaks at donor site

Chapter 2

Definition of Sequence Requirements at the Recombination Site for Piv-Mediated Inversion and Characterization of Piv Single-Stranded DNA Binding Activity[1]

[1] Carpenter, C. D., Lenich, A. G., and Karls, A. C. 2008. To be submitted to the *Journal of Molecular Biology*.

Abstract

The site-specific recombinase, Piv, mediates the inversion of the type 4 pilin DNA segment in *Moraxella lacunata* and *Moraxella bovis*. Piv belongs to the Piv/MooV family of recombinases, which includes transposases of IS*110*/IS*492* family of insertion sequences, and does not show sequence similarity to site-specific recombinases of the tyrosine- or serine-recombinase families. Piv contains a DEDD catalytic motif that is predicted to be positioned within an RNase H-like fold similar to that found in DDE-motif transposases and retroviral integrases. Because Piv-mediated inversion has characteristics of recombination mediated by site-specific recombinases and DDE-motif transposases, Piv has been studied genetically and biochemically to characterize its mechanism of recombination. To aid in the elucidation of the mechanism of Piv-mediated recombination, we characterized the sequence requirement of the recombination sites for Piv-mediated inversion and the DNA binding activities of Piv. Fifteen base pairs of the perfect thirty-two base pair inverted repeats at the boundary of the pilin inversion segment are required for inversion. Electrophoretic mobility shift assays indicate that Piv binds single-stranded *inv* sequences with a higher affinity than double-stranded *inv* sequence. Surprisingly, Piv binding to single-stranded DNA is not sequence specific, suggesting that DNA structure determines the specificity of Piv binding.

Introduction

The site-specific invertase, Piv, catalyzes the inversion of a chromosomal segment encoding type 4 pilin genes, *tfpQ* and *tfpI*, of *Moraxella lacunata*, a human eye pathogen, and *Moraxella bovis*, a cow eye pathogen. Inversion switches the gene expressed from the pilin promoter, P_{tfp} (7, 9, 11). The inversion of this chromosomal segment results in *M. bovis* alternately expressing either TfpQ or TfpI pili. However, in *M. lacunata* the *tfpI* gene is non-functional due to a frame-shifting 19 bp duplication in *tfpI*. Thus, Piv mediated inversion of the pilin segment in *M. lacunata* controls phase variation of type 4 pili, where in one orientation *tfpQ* is expressed from P_{tfp}, and in the other orientation no pilin is expressed (Figure 2.1).

The pilin invertible segment is flanked by identical 32 bp inverted repeats, *inv,* where Piv mediates recombination. Piv, purified as a fusion to maltose binding protein (MBP) interacts with DNA fragments containing *invL* and *invR,* which include DNA sequences that flank both the left and right ends of the invertible segment. Further characterization of MBP-Piv interactions with double-stranded *invL* showed that in electrophoretic mobility shift assays (EMSA) less than 40% of the double-stranded labeled *invL* DNA substrate could be specifically bound, and MBP-Piv did not protect *invL* from DNase I or 1, 10-phenanthroline-copper cleavage (14). In addition, measurements of the dissociation rate for MBP-Piv bound to *invL* show that the half-life of the complex is less than 15 seconds indicating that the complex is unstable (14). Although, Piv interactions with double-stranded *invL* are quite weak, Piv interacts strongly with an accessory site, *sub1* (strong upstream binding site), which is located 5' to the pilin promoter (14). There are two other possible *sub* sites. One is located within a 535 bp DNA fragment (includes 3'end of *tfpQ* and sequence flanking 3'end), and the other is located within 250 bp DNA fragment (includes *piv* sequence). Both fragments were bound by MBP-Piv in vitro, and

sequence alignments with *sub* sequence using GCG Wisconsin Package Bestfit program indicate that these fragments contain *sub*-like sequences (14).

The *sub1* site is bound by at least two protomers of MBP-Piv. Dissociation rate analysis indicates that the first MBP-Piv protomer that binds to *sub1* forms a more stable complex (half-life of 12 minutes) than the multimeric complex (half-life of 15 seconds) (14). In contrast, Piv only binds *invL* as a multimer based on cooperativity assays and competition analyses (14). The nature of MBP-Piv binding to double-stranded *invL* and *sub1* suggest that it is possible that the *sub* site is an assembly site for the multimeric form of Piv that binds to the recombination sites (14).

Although Piv is capable of weakly binding the double-stranded *sub1* site, its poor interactions with its recombination site led us to determine if Piv binds other structures of DNA with better affinity than double-stranded DNA. Other recombinases have been shown to bind alternate structures of DNA better than double-stranded DNA, such as IntI1 integron integrase which binds single-stranded, but not double-stranded *attC* sequence (4). In fact, IntI1 integron integrase only binds one of the two strands of *attC*, the bottom strand. Another transposase that recognizes and binds only single-stranded DNA is the transposase encoded by ISHp*608*, transposase A (TnpA) (5, 15). Both IntI1 integron integrase and TnpA bind single-stranded DNA capable of forming 2° structures. We show here the MBP-Piv is capable of binding single-stranded *invR* sequence with higher affinity than double-stranded *invR*, but that this binding is not limited to only one strand of *invR*. MBP-Piv shows a preference for some single-stranded DNA sequences over others. We propose that the differences in possible 2° structures formed by these single-stranded DNA sequences affect MBP-Piv binding. We also further define the

essential DNA sequences within the inverted repeat, and use single-stranded binding assays to determine if the sequences essential for inversion are also essential for Piv binding.

Materials and Methods

Bacterial strains, plasmids, and media. *E. coli* DH5α [(Φ80*dlacΔ(lacZ)*M15*) sup*E44 *Δlac*U169 *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A] (obtained from C. Moran) and Top10 [F- *mcr*A Δ*(mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R) *end*A1 *nup*G] (Invitrogen, Carlsbad, CA) containing drug resistant plasmids were grown on Luria-Bertani (LB) agar plates or in LB media at 37°C with antibiotics at the following concentrations: ampicillin, 60 or 80 μg/ml; chloramphenicol, 34 μg/ml. Plasmids pMxL5 and pMxL1dL43 were described previously (9). Briefly, pMxL5 contains the *Moraxella lacunata* pilin inversion segment with *piv* inactivated by the insertion of a segment from the Ω interposon encoding streptomycin/spectinomycin resistance, and pMxL1dL43 is a derivative of pMxL1 (pMxL5 without the Ω interposon) that is phase-locked due to an Ω insertion that overlaps *invR*. pAG701 is pMxL5 with a deletion of *tfpB* as described previously (7). Plasmid pAG702 carries the *M. lacunata piv* under control of the *lac* promoter and LacI and its construction was described previously (13). All chemicals were purchased from Sigma Chemical Co (St. Louis, MO). T4 DNA ligase, polynucleotide kinase, and restriction enzymes were obtained from New England Biolabs (Beverly, MA).

Plasmid constructions. Sequences of all oligonucleotides used are listed in Table 2.1. Site-directed mutations from -7 position to position +21 (relative to the start of the inverted repeat) of *invR* (*tfpI* side) were derived from pAG701 using the QuikChange mutagenesis kit from Stratagene (La Jolla, CA). The site-directed mutation at position 14/17 for *invL* (*tfpQ)* was derived from pMxL1dL43 using the Quik Change mutagenesis kit. The oligonucleotide

Qkgchng was used for QuikChange mutagenesis with the substitutions made to the oligonucleotide to create the site-directed mutants indicated in the results.

Mutations at -7/-4, 22/25, 26/29, and 30/32 were created by overlap extension PCR. Mutant plasmid -7/-4 was derived from pAG701; however, mutant plasmids 22/25, 26/29, and 30/32 were derived from pCDC17. pCDC17 was created by replacing the AgeI/HindIII fragment of Litmus 29 with the AgeI/HindIII restriction fragment of pMxL5 (contains 3' end of *tfpQ*, *tfpB*, *tfpI* and part of *piv*). pCDC17 was created to decrease the chance of oligonucleotides annealing to both *invR* and *invL* because the *invL* sequence is not in this plasmid. Oligonucleotides -7/-4A and XL5PmlI or -7/-4B and XL5MfeI2 were used to produce the templates for the second round of PCR to create -7/-4. Oligonucleotides R2225A, R2629A, or R3032A and XL5AccI or R2225B, R2629B, or R3032B and Xl5SnaB1 were used to produce the templates for the second round of PCR to create 22/25, 26/29 and 30/32. Mutated PCR products at positions 22/25, 26/29, and 30/32 were ligated into pCR2.1 (Invitrogen), confirmed by sequence, and then subcloned into pMxL5.

All mutated *invR* substrates were subcloned into pMxL5 at the PmlI/MfeI sites on the *tfpI* sides. The double mutant, 14/17QI, was constructed by inserting the *tfpQ* recombination site mutation using the KpnI sites. Mutations were confirmed by sequencing using Sequeanse kit (Amersham) or University of Michigan Biomedical Research Core Facilities.

All PCR products created for cloning purposes were amplified using Pfu DNA polymerase (Stratagene). The thermal cycling conditions were as follows: 95°C for 3 min followed by 25 cycles of 95°C for 30 sec, the T_m of the oligonucleotide with the lesser T_m value minus 5°C for 30 sec and 72°C for 2.5 min. If PCR products were ligated into pCR2.1, they were incubated for an additional 10 min at 72°C with *Taq* DNA polymerase in the presence of

dATP. All ligations were performed at 16°C overnight with 400 U of T4 DNA ligase, or, for pCR2.1 cloning, with topoisomerase from the Invitrogen kit. Transformation of ligation mixtures into host strains was performed with chemically competent cells or by using the One Shot TOP10 competent cells from Invitrogen followed by plating on LB agar containing the appropriate antibiotics.

In vivo Piv inversion assays. For *invR* mutants at positions -7 to +21, pMxL5 plasmids containing the mutant DNA inversion substrates were transformed into *E. coli* DH5α cells containing the pAG702 Piv-expression plasmid. The transformants were resuspended in 2 ml each of LB containing the appropriate antibiotics and grown to an O.D. 600 of O.5. Cells were then induced with a final concentration of 0.1mM isopropylthio-β-D-galactoside (IPTG) and incubated with aeration at 37°C overnight. Alkaline lysis minipreps were performed, and inversion was assayed by restriction analysis with KpnI and gel electrophoresis (7, 12).

For *invR* mutants at positions 22 to 32, the inversion assays were performed as described previously (3). Briefly, chemically competent DH5α cells containing pMxL5 plasmids with the mutant *invR* sequences were transformed with pAG702 and plated on LB agar containing the appropriate antibiotics and 50 μM IPTG. After 24 to 36 hr incubation at 37°C, individual colonies were inoculated into LB broth containing the appropriate antibiotics and incubated at 37°C with aeration overnight. The inversion substrate and expression vector were isolated using GenElute Plasmid Miniprep Kit (Sigma-Aldrich), and inversion was assayed by restriction analysis with KpnI and agarose gel electrophoresis.

Purification of Piv. Piv was purified as a fusion protein to maltose-binding-protein (MBP) at its amino-terminus to produce MBP-Piv as described previously (14).

Labeling of DNA oligonucleotides used in DNA binding assays. 50 pmol invRB (bottom strand) or invRT (top strand) were labeled with 50 pmol of [γ-^{32}P]ATP using T4 polynucleotide kinase. To calculate the fraction of labeled oligonucleotide, after heat-inactivating T4 polynucleotide kinase, 0.5 µl of the labeling reaction was spotted onto a PEI strip and placed in 2N HCl. After the HCl had migrated up the PEI strip, the strip was cut in half and each half was placed into a scintillation vial containing scintillation fluid and counted on the Any Isotope program of Beckman LS6500 Multi-purpose Scintillation Counter. The bottom half of the PEI strip contained the labeled oligonucleotide, but the top half contained the unincorporated label; therefore, to calculate the fraction labeled, the CPM of the bottom half of the PEI strip was divided by the total number of CPM of the top and bottom halves. This number was multiplied by the specific activity of the [γ-^{32}P] ATP to calculate the number of molecules of labeled oligonucleotide. The number of molecules of unlabeled oligonucleotide was calculated by subtracting the number of molecules labeled from the total number of molecules in the reaction (50 pmol). Generally, the ratio of unlabeled to labeled oligonucleotide was between 2 and 3. The unincorporated label was removed from the sample using Centrisep Spin Columns (Princeton Separations, Adelphia, NJ).

Electrophoretic mobility shift assays. Electrophoretic mobility shift assays (EMSA) were performed by adding MBP-Piv to reaction mixture containing radiolabeled DNA oligonucleotide, 80 mM KCl, 20 mM Tris-HCl, pH 7.6, 5 mM $CaCl_2$, 250 ng/ml poly(C), 1 mM dthiothreitol, and 50 µg/ml BSA (14). The reactions were incubated for 20 min at room temperature, loaded onto a 5% non-denaturing polyacrylamide gel, and electrophoresed at 4°C in 0.5X TBE buffer. For the competition assays, 0.1 nM labeled oligonucleotide and 1140 nM MBP-Piv were incubated at room temperature in the above reaction buffer for 10 min before

being added to tubes containing the increasing amounts of competitor DNA ranging from 10 to 500-fold molar excess to labeled oligonucleotide followed by another 10 min incubation at room temperature before being loaded onto a 5% non-denaturing polyacrylamide gel.

The sequences of the top strand oligonucleotides used in the binding assays are listed in Table 2.1 with invRT being wild type and the positions of the mutations indicated by number. The oligonucleotides containing the bottom strand sequences (invRB) are complementary to the sequences of the top strand oligonucleotides listed in Table 2.1. The sequences of the non-specific oligonucleotides, LacZDra, U55, U42, U34, and U21 are also listed in Table 2.1. To create the double-stranded *invR* substrate, labeled invRB was annealed to unlabeled invRT.

Calculating relative percentage of bound oligonucleotide. The Image Quant TL software (GE Healthcare, Piscataway, NJ) was used to calculate the density of both the bound and unbound bands in each individual lane. The percentage of oligonucleotide bound was calculated by dividing the density of the bound band by the total density of both the bound and unbound bands in a single lane and multiplying by 100%. The percentages of bound oligonucleotides in the lanes containing unlabeled competitor were normalized to the percentage of bound oligonucleotide without competitor present. Each competition assay was repeated in triplicate.

Results

Sequence requirements of Piv-mediated site-specific inversion. There are perfect 32 bp inverted repeats, *invL* and *invR* (including flanking sequence to define the left versus right end), at both ends of the pilin invertible segment (11). Piv-mediated inversion of this segment is site-specific indicating that cleavage and strand exchange must occur within these long inverted repeats. It has also been shown that MBP-Piv binds to *invL* in vitro, but the interaction is too

weak to utilize DNase I or chemical protection assays. Therefore, scanning mutagenesis of *inv* was performed to define the sequences within the 32 bp inverted repeats that are essential for Piv-mediated inversion (Figure 2.2) (14). To determine if sequences flanking *inv* are essential for inversion, flanking sequences in *invR* (within the *piv* gene) were also mutated (Figure 2.2).

As seen in Figure 2.3, sequences beginning at position +2 and ending at position +17, with the exception of position +8, are required for Piv-mediated inversion. There are two possible explanations for the requirement of these sequences: (1) homology is required in the crossover region to facilitate strand exchange; (2) these sequences are required for Piv recognition and binding to the ends of the invertible segment.

Homology requirements within the crossover region. To investigate whether homology is required in the crossover region, the complementary mutations were created in the 14/17 mutant to restore homology to the recombination sites. In vivo inversion assays with this mutant (Figure 2.4) indicate that recreating the homology of the sites within mutants does not allow for inversion. This suggests that, at least for the 14/17 mutant, homology at the recombination sites is not essential.

Piv single-stranded DNA binding activity. To examine if the sequences corresponding to the inversion negative *invR* mutants are required for Piv binding, in vitro binding assays with MBP-Piv and *invR* mutants were performed. Surprisingly, MBP-Piv binding to a double-stranded (ds) oligonucleotide containing *invR* was not seen. However, MBP-Piv binding to single-stranded (ss) *invR* was detected. Figure 2.5 shows that MBP-Piv is unable to shift a ds *invR* substrate in the presence of the same amounts of protein that are able to shift a ss *invR* substrate. There is a shift in the assay employing ds substrate; however, some ss substrate is present in the reactions. The band corresponding to ds substrate was not shifted, instead, the band corresponding to ss

substrate was the one that was shifted by MBP-Piv. To determine if MBP-Piv binding is specific to the bottom and top strands of *invR*, a competition assay was performed using each of the oligonucleotides as both binding substrate and cold competitor. The oligonucleotide, LaczDra, was used as cold non-specific competitor for both bottom and top strands of *invR*. As seen in Figures 2.6 and 2.7, binding to the bottom and top strands of *invR* appears to be specific; since an increase in cold, specific ss competitor results in the decrease of binding to the same labeled oligonucleotide. The increase in cold, ss non-specific competitor results in a slight decrease of binding to the bottom and top strands of *invR*, but this change is dramatically less than that seen in the presence of specific competitor (Figure 2.10, LacZDra data points in comparison to WT data points). Based on these data, competition assays with ss oligonucleotides containing the *invR* mutations listed in Figure 2.2 were performed for both the bottom and top strands of *invR*.

Binding of Piv to single-stranded *invR* DNA with base substitutions. MBP-Piv binding to mutant *invR* substrates was investigated using competition assays in which each of the unlabeled substrates with specific base substitutions was added in increasing molar ratios to the labeled wild-type substrate in the presence of MBP-Piv. As expected, mutants that were capable of supporting inversion (-7/-4, -3/1, 8, 18/21, 22/25, 26/29, and 30/32) efficiently competed for MBP-Piv binding as well as the wild-type competitor (Figure 2.8). Unexpectedly, mutants that were unable to support inversion (2/5, (6, 7, 9), 10, 11, 12, 13, 14/17) also competed for MBP-Piv binding as well as the wild-type competitor (Figure 2.9). This result was puzzling as previous binding assays with specific competitor and non-specific competitor indicated that the binding of MBP-Piv to both bottom and top strands of *invR* is specific (Figures 2.6 and 2.7); therefore, we investigated the ability of MBP-Piv to bind other nonspecific competitors.

Piv has no apparent sequence preference for binding single-stranded DNA. One of the differences, besides sequence, between the nonspecific oligonucleotide that was used as a competitor, LacZDra, and the mutant oligonucleotides is that LacZDra is 34 nt long compared to the 42 nt length of the mutant oligonucleotides (the wild-type oligonucleotides are also 42 nt long). Therefore, we set up competition assays between the wild-type *invR* oligonucleotides and non-specific oligonucleotides ranging from 21 nt to 55 nt (U21 to U55). If MBP-Piv has a preference for longer ss DNA substrates, then the competition between non-specific competitor and the bottom and top strands of *invR* would increase as the length of the non-specific competitor increased. Figure 2.10 shows that, with the exception of LacZDra, all four non-specific competitors competed for MBP-Piv binding as well as wild-type competitor. Why is LacZDra unable to compete with bottom and top strands of invR for MBP-Piv binding? One explanation is the possibility of 2° structure in the middle of the LacZDra oligonucleotide (Figure 2.11). We used the DINAMelt Server (dinamelt.bioinfo.rpi.edu) to predict structures and their corresponding energies for the oligonucleotides used in the competition assays. As seen in Figure 2.11, invRB and invRT have predicted 2° structures containing a large hairpin at one end of the oligonucleotide with a smaller ring-structure located further to the interior of the oligonucleotide. The ΔG values for the proposed 2° structures for invRB and invRT are -2.0 and -1.8, respectively. Since these ΔG values are fairly high, it is likely that these structures are not formed at all or are formed intermittently. The ΔG value for the 55 nt non-specific oligonucleotide (U55) is more favorable at -5, but, like invRB and invRT, its hairpins are located at the ends of the oligonucleotide which may reduce the ability of U55 to stay folded in its predicted 2° structure. In contrast, LacZDra is predicted to fold into a 2°structure with a ΔG value of -4.3 and has a hairpin in the center of the DNA strand.

Discussion

Piv interactions with ds *invL* are weak resulting in an inability to get a "footprint" at the *invL* site in either nuclease or chemical protection assays; therefore, to define the sequence requirements within the inverted repeats we introduced specific base pair substitutions (transversions and transitions) in *invR* and tested the ability of the mutant substrates to invert (14). A 15 bp region of *invR* (positions 2 to 7 and 9 to 17) is required for Piv-mediated inversion of the invertible segment. We postulate that either homology is required at the crossover region or these sequences are required for Piv recognition prior to cleavage and strand transfer.

To create a mutant with homology at the crossover region, a substitution was introduced at *invL* that was complementary to the mutation at *invR* at positions 14 to 17. Because this mutant (14/17QI) was still inversion negative, homology at the crossover region was not enough to rescue an inversion minus mutant leading us to propose that the specific nucleotide sequence in this region is important for the inversion reaction.

To our surprise, when we set up in vitro binding assays to determine if MBP-Piv could bind the mutant *invR* as well as wild-type *invR*, we were unable to detect MBP-Piv binding to a ds oligonucleotide containing *invR* sequence, but we were able to detect MBP-Piv binding to single-stranded *invR* (both bottom and top strands). Previous data suggested that MBP-Piv could bind a ds oligonucleotide containing *invL* sequence, though this binding was very weak. Because both *invL* and *invR* contain the same 32 bp sequence, we expected MBP-Piv to bind ds *invR* (14). It is possible that what was reported as Piv bound to ds *invL* was actually Piv bound to ss *invL*. The polyacrylamide gels used to separate bound and unbound *invL* were 9% gels in contrast to the 5% gels used in this study. Figure 2.5 shows that even in a 5% polyacrylamide gel, the electrophoretic mobility of ss and ds oligonucleotides are close, and it is possible that

there was no separation of the ss and ds oligonucleotides in the 9% polyacrylamide gel. Because we were unable to reproduce MBP-Piv binding to ds DNA in vitro, we decided to pursue MBP-Piv binding to ss DNA.

MBP-Piv binding to both bottom and top strands of *invR* appeared to be site-specific as binding to the bottom and top strands decreased in the presence of increasing amounts of ss specific competitor. There was only a slight decrease in binding to bottom and top strands in the presence of increasing amounts of ss non-specific competitor (LacZDra). We decided to use ss mutant *invR* to determine if MBP-Piv was able to bind the *invR* in the inversion minus mutants as well as wild-type *invR*. The competition assays using cold bottom or top strand *invR* with labeled bottom or top strand of *invR* indicated that MBP-Piv binds the mutated *invR* sequences in the inversion minus mutants as well as the mutated *invR* sequences of inversion positive mutants and wild-type *invR*. This led us to investigate MBP-Piv binding to *invR* bottom and top strands in the presence of other non-specific competitors.

All of the non-specific competitors, except LacZDra, were able to compete with labeled bottom or top strands of *invR* for MBP-Piv binding as well as wild-type competitors. What is so different about the LacZDra sequence from the mutant *invR* sequences and the non-specific sequences? LacZDra is predicted to form a fairly stable hairpin in the inside of the oligonucleotide with short stretches of ss DNA on either side. In contrast the possible hairpins of invRB, invRT, and U55 are located at the ends of the oligonucleotide and may be less stable due to their location. These oligonucleotides also have longer stretches of ss DNA than LacZDra. Due to the differences between these oligonucleotides, we propose that MBP-Piv preferentially binds ss DNA that does not fold into 2° structures resulting in long spans of ds DNA and short spans of ss DNA. This is different than the ssDNA binding activity of IntI integron integrase and

TnpA of ISHp608 both of which require 2° structure within ssDNA in order to bind (4, 6, 10, 15). In fact, crystal structures of both IntI and TnpA bound to ssDNA show that the proteins interact with hairpins formed by ss DNA folding (2, 8, 10).

Another recombinase that has been shown to bind ssDNA with better affinity than dsDNA is Flp, a tyrosine site-specific recombinase (16). An in vitro binding selection assay recovered several ssDNA sequences that Flp binds in a sequence specific manner; however, the Flp recognition target (FRT) is only bound by Flp in vitro when it is ds (16). So far, there is no experimental evidence that Flp ssDNA binding activity has a role in recombination but it has been proposed that the binding of two Flp molecules to FRT induces a bend in the DNA that is severe enough to possibly separate the strands of the core which could then be bound by Flp to mediate strand exchange (16). It is possible that binding of Piv to the *sub* site could facilitate Piv-mediated opening of dsDNA at the recombination site allowing MBP-Piv to bind the released ssDNA and begin cleavage and strand exchange.

Another possibility is that Piv binding to ssDNA does not play a role in recombination. Instead, Piv binding to the *sub* site may be required to occur before Piv can bind the recombination sites in order for Piv to form a stable synaptic complex before DNA exchange occurs similar to λ integrase binding to arm and core sequences. λ integrase binding to arm sites is more stable than its binding to core sites and this binding to the arm sites allows the formation of a stable synaptic complex with the catalytic domain at the regions of DNA exchange [for review see (1)]. Although the *sub* site immediately upstream of the pilin promoter is not required for inversion, the two other possible *sub* sites may be required (9). In order to resolve how Piv interacts with *invL* and *invR*, we are currently utilizing in vitro sandwiching and looping assays with oligonucleotides containing both *sub* and *invR* sequences.

References

1. **Azaro, M. A., and A. Landy.** 2002. Lambda Integrase and the Lambda Int family, p. 118-48. *In* N. Craig, Craigie R, Gellert M, Lambowitz, AM (ed.), Mobile DNA II. ASM Press, Washington D. C.

2. **Barabas, O., D. R. Ronning, C. Guynet, A. B. Hickman, B. Ton-Hoang, M. Chandler, and F. Dyda.** 2008. Mechanism of IS*200*/IS*605* family DNA transposases: activation and transposon-directed target site selection. Cell **132:**208-20.

3. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J Bacteriol **187:**3431-7.

4. **Francia, M. V., J. C. Zabala, F. de la Cruz, and J. M. Garcia Lobo.** 1999. The IntI1 integron integrase preferentially binds single-stranded DNA of the *attC* site. J Bacteriol **181:**6844-9.

5. **Guynet, C., A. B. Hickman, O. Barabas, F. Dyda, M. Chandler, and B. Ton-Hoang.** 2008. In vitro reconstitution of a single-stranded transposition mechanism of IS*608*. Mol Cell **29:**302-12.

6. **Johansson, C., M. Kamali-Moghaddam, and L. Sundstrom.** 2004. Integron integrase binds to bulged hairpin DNA. Nucleic Acids Res **32:**4033-43.

7. **Lenich, A. G., and A. C. Glasgow.** 1994. Amino acid sequence homology between Piv, an essential protein in site-specific DNA inversion in *Moraxella lacunata*, and transposases of an unusual family of insertion elements. J Bacteriol **176:**4160-4.

8. **MacDonald, D., G. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul.** 2006. Structural basis for broad DNA-specificity in integron recombination. Nature **440:**1157-62.

9. **Marrs, C. F., F. W. Rozsa, M. Hackel, S. P. Stevens, and A. C. Glasgow.** 1990. Identification, cloning, and sequencing of *piv*, a new gene involved in inverting the pilin genes of *Moraxella lacunata*. J Bacteriol **172:**4370-7.

10. **Ronning, D. R., C. Guynet, B. Ton-Hoang, Z. N. Perez, R. Ghirlando, M. Chandler, and F. Dyda.** 2005. Active site sharing and subterminal hairpin recognition in a new class of DNA transposases. Mol Cell **20:**143-54.

11. **Rozsa, F. W., and C. F. Marrs.** 1991. Interesting sequence differences between the pilin gene inversion regions of *Moraxella lacunata* ATCC 17956 and *Moraxella bovis* Epp63. J Bacteriol **173:**4000-6.

12. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor.

13. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. Mol Microbiol **39:**641-51.

14. **Tobiason, D. M., A. G. Lenich, and A. C. Glasgow.** 1999. Multiple DNA binding activities of the novel site-specific recombinase, Piv, from *Moraxella lacunata*. J Biol Chem **274:**9698-706.

15. **Ton-Hoang, B., C. Guynet, D. R. Ronning, B. Cointin-Marty, F. Dyda, and M. Chandler.** 2005. Transposition of IS*Hp608*, member of an unusual family of bacterial insertion sequences. EMBO J **24:**3325-38.

16. **Zhu, X. D., and P. D. Sadowski.** 1998. Selection of novel, specific single-stranded DNA sequences by Flp, a duplex-specific DNA binding protein. Nucleic Acids Res **26:**1329-36.

Table 2.1

Oligonucleotide	Sequence
Qkchng	5'GAATAACATTTATTGGTATCCTAGCTGCAATCGCTCTA
XL5SnaB1	5'AAGAGATGTTAATAGTGGTTGC
XL5AccI	5'ACAACGTGAACACCACAAACGG
XL5PmlI	5'GTCTTTATTAGCACGTGTACC
XL5MfeI2	5'GGCAACTCATCAATTGCACG
R-7/-4A	5'GAATAATCGATATTGGTATCCTAGCTC
R-7/-4B	5'CCAATATCGATTATTCATGAATGCGTTTG
R2225A	5'AGGTAGATGACTTGCAGCTAGGATAC
R2225B	5'AGCTGCAAGTCATCTACCTAGCTGC
R2629A	5'AGCTAGGCTACGCGATTGCAGCTAG
R2629B	5'TGCAATCGCGTAGCCTAGCTGCAATCG
R3032A	5'TGCAGCTCAATAGAGCGATTGCAGC
R3032B	5'AATCGCTCTATTGAGCTGCAATCGCTC
invRT	5'TAGGTAGAGCGATTGCAGCTAGGATACCAATAAATGTTATTC
invRT-7-4	5'TAGGTAGAGCGATTGCAGCTAGGATACCAATAAATGCCTATC
invRT-3+1	5'TAGGTAGAGCGATTGCAGCTAGGATACCAATATCGATTATTC
invRT25	5'TAGGTAGAGCGATTGCAGCTAGGATACCCTACAATGTTATTC
invRT679	5'TAGGTAGAGCGATTGCAGCTAGGAAAGGAATAAATGTTATTC
invRT8	5'TAGGTAGAGCGATTGCAGCTAGGATTCCAATAAATGTTATTC
invRT10	5'TAGGTAGAGCGATTGCAGCTAGGTTACCAATAAATGTTATTC
invRT11	5'TAGGTAGAGCGATTGCAGCTAGAATACCAATAAATGTTATTC
invRT12	5'TAGGTAGAGCGATTGCAGCTAAGATACCAATAAATGTTATTC
invRT13	5'TAGGTAGAGCGATTGCAGCTTGGATACCAATAAATGTTATTC
invRT1417	5'TAGGTAGAGCGATTGCTAGAAGGATACCAATAAATGTTATTC
invRT1821	5'TAGGTAGAGCGAGCATAGCTAGGATACCAATAAATGTTATTC
invRT2225	5'TAGGTAGATGACTTGCAGCTAGGATACCAATAAATGTTATTC
invRT2629	5'TAGGCTACGCGATTGCAGCTAGGATACCAATAAATGTTATTC
invRT3032	5'TCCATAGAGCGATTGCAGCTAGGATACCAATAAATGTTATTC
LacZDra	5'TCACACACGTTGTGAGTTAGCTCACTCATTAGGC
U55	5'TACGCCACGGTATCGTCGGCTCTGTAACTATGTCTAAGTC ACTCAGGTGTGCGAC
U42	5'ACGGTATCGTCGGCTCTGTAACTATGTCTAAGTCACTCAGGT
U34	5'TATCGTCGGCTCTGTAACTATGTCTAAGTCACTC
U21	5'CGGCTCTGTAACTATGTCTAA

Figure 2.1: Piv controls phase variation of type 4 pili in *M. lacunata*. The 2.1 kb chromosomal invertible segment (indicated by the parentheses) encodes two pilin proteins encoded by the genes *tfpQ* and *tfpI*. The promoter and constant region of the pilin genes are located upstream of the invertible segment. If *tfpQ* is located immediately adjacent to the constant region (Q orientation), *M lacunata* produces TfpQ type 4 pili. When Piv, located immediately downstream of the invertible segment, site-specifically inverts the chromosomal segment placing *tfpI* immediately adjacent to the constant region (I orientation), *M. lacunata* does not produce type 4 pili. The 19 bp duplication in *tfpI* is indicated by the black box. The location of the *sub1* site is indicated by the pink box, and the locations of the other two possible sub sites are indicated by gold stars.

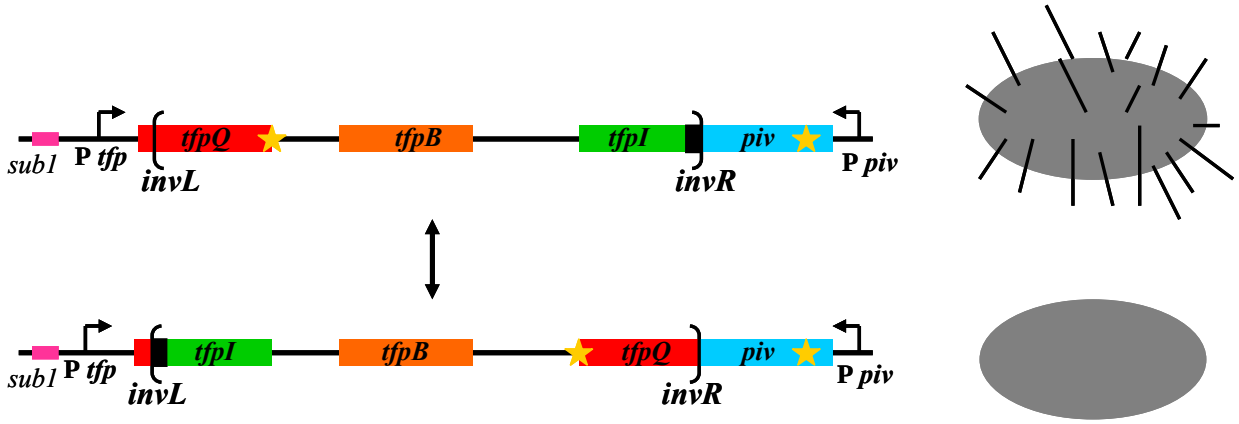

sub1
P tfp
tfpQ
tfpB
tfpI
piv
P piv
invL
invR
sub1
P tfp
tfpI
tfpB
tfpQ
piv
P piv
invL
invR

Figure 2.2: *invR* mutants. Top strand of *invR* is shown with 7 nucleotides of *piv* sequence at the 3' end. The inverted repeat sequence is between the +1 and +32 arrows. The sequence of each mutant is specified underneath with a dot indicating wild-type sequence. Names of the mutants are derived from the position of the substitution relative to the proposed +1 site of the inverted repeat. The bottom strand of *invR* has the complementary substitutions.

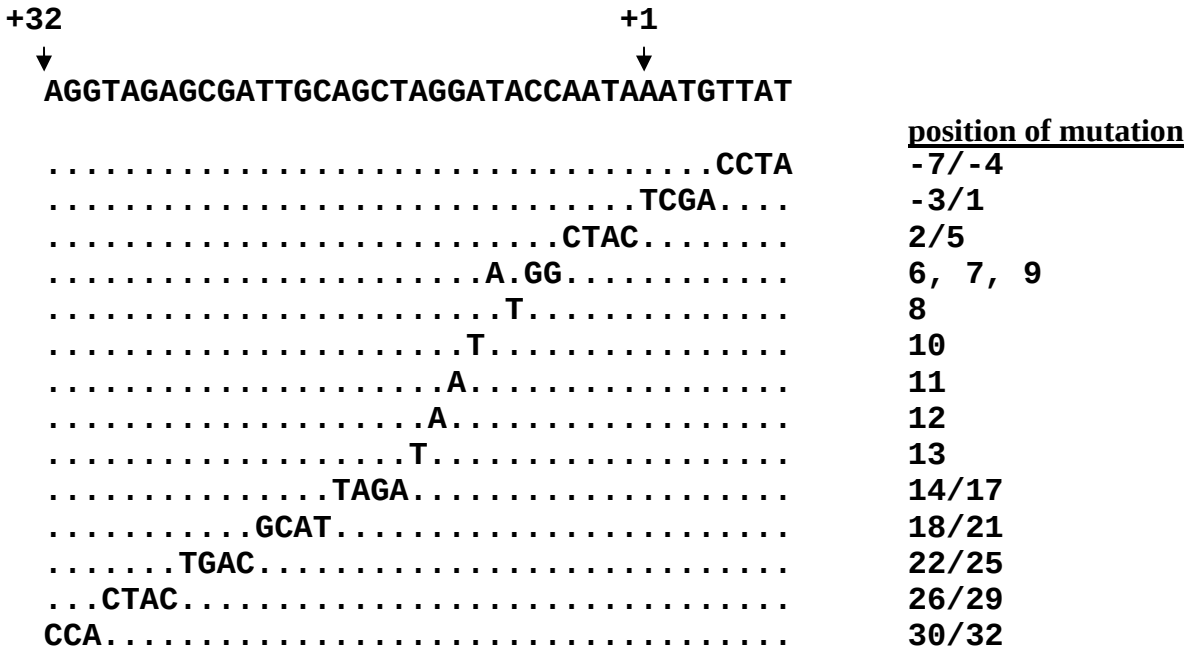
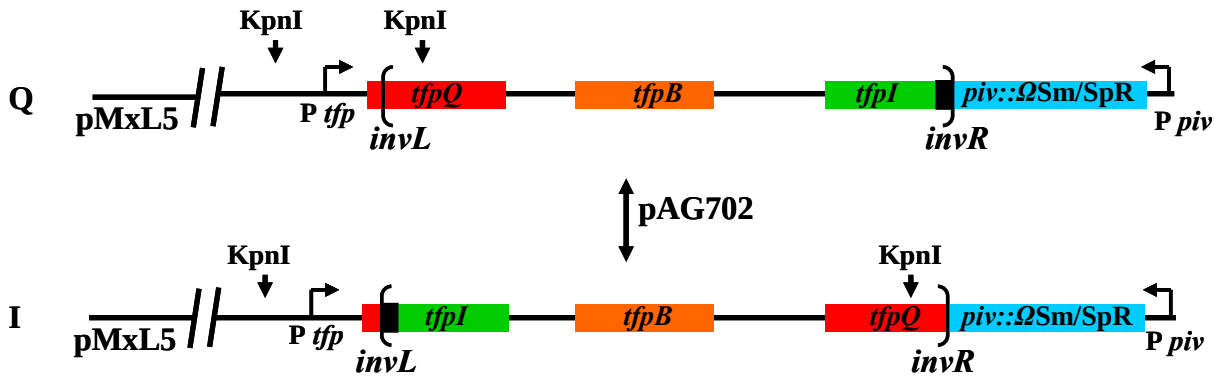
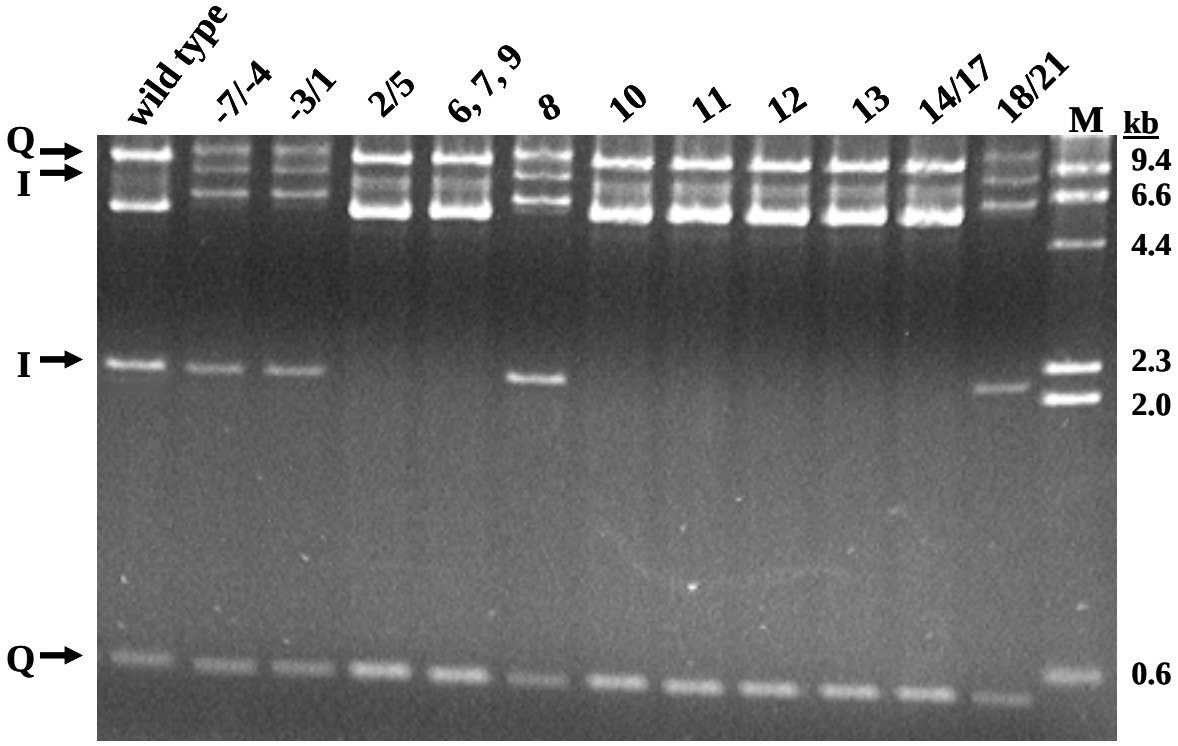

+32
+1
AGGTAGAGCGATTGCAGCTAGGATACCAATAAATGTTAT
position of mutation
.....CCTA -7/-4
.....TCGA.... -3/1
.....CTAC..... 2/5
.....A.GG..... 6, 7, 9
.....T..... 8
.....T..... 10
.....A..... 11
.....A..... 12
.....T..... 13
.....TAGA..... 14/17
.....GCAT..... 18/21
.....TGAC..... 22/25
...CTAC..... 26/29
CCA..... 30/32

Figure 2.3: In vivo inversion assay with substrates containing the mutant recombination sites. (A). Diagram of inversion substrate. pMxL5 starts in the Q orientation, and addition of the expression vector pAG702 results in the inversion of the pilin inversion segment resulting in the I orientation. The asymmetric placement of KpnI restriction digests allows for detection of inversion of the segment containing the pilin genes. The Q orientation yields 9.5 and 0.4 kb KpnI restriction fragments, and the I orientation yields 7.8 and 2.1 kb KpnI restriction fragments. (B). Ethidium bromide stained agarose gel of KpnI digested inversion substrates from position -7 to position +21. The assay was performed by growing transformants to an O.D. 600 of 0.5 before inducing with IPTG and incubating overnight (see Materials and Methods). The Q and I orientation bands are indicated by labeled arrows. (C). Ethidium bromide stained agarose gel of KpnI digested inversion substrates from position +22 to position +32. The assay was performed by plating transformants on LB agar plates containing IPTG, incubating overnight, and isolating plasmids from a tranformant grown in LB overnight (see Materials and Methods). The Q and I orientation bands are indicated by labeled arrows.

A.


KpnI
KpnI
Q
pMxL5
P tfp
tfpQ
tfpB
tfpI
piv::ΩSm/SpR
P piv
invL
invR
pAG702
KpnI
KpnI
I
pMxL5
P tfp
tfpI
tfpB
tfpQ
piv::ΩSm/SpR
P piv
invL
invR

B.


wild type
-7/-4
-3/1
2/5
6, 7, 9
8
10
11
12
13
14/17
18/21
M
kb
Q
I
I
Q
9.4
6.6
4.4
2.3
2.0
0.6

C.

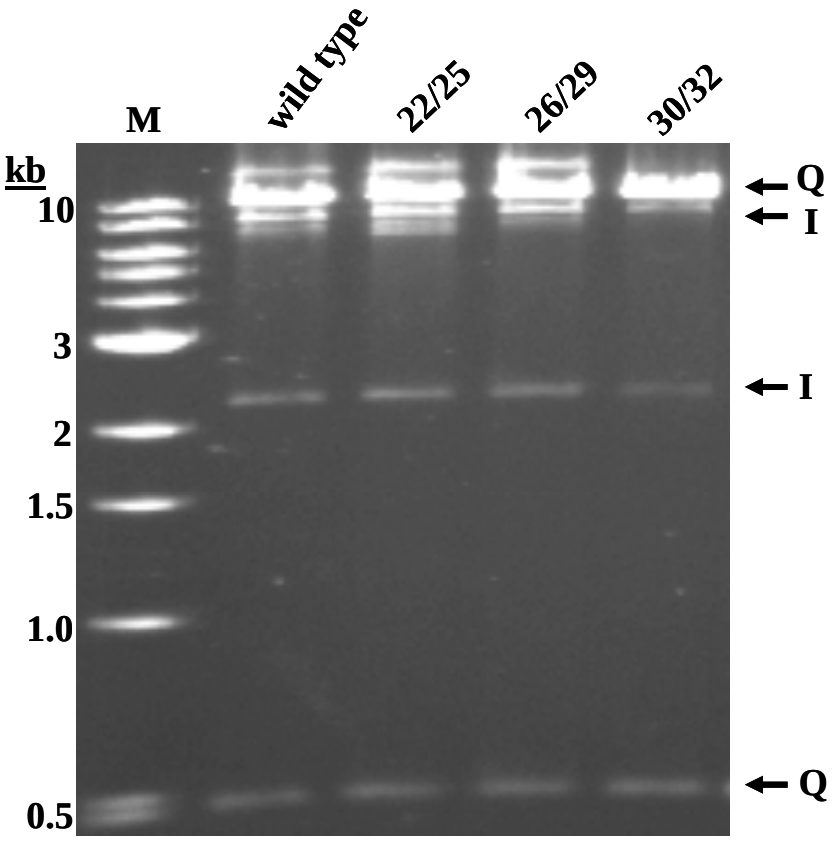


Figure 2.4: In vivo inversion assay with symmetric and asymmetric mutant inversion substrates. The asymmetric mutant, 14/17, has mutations located within *invR* (*tfpI*), while the symmetric mutant, 14/17QI, has mutations located in both *invL* (*tfpQ*) and *invR* (*tfpI*). The inversion substrate, pMxL5, starts in the Q orientation (9.5 and 0.4 kb KpnI fragments), and inversion (I orientation) results in 7.8 and 2.1 kb KpnI fragments. The expected inversion bands for 14/17 would be the same as pMxL5; however, 14/17QI, which has an additional 1 kb in the KpnI restriction fragment containing *tfpQ*. The expected inversion bands for 14/17QI are 7.8 and 3.1 kb (I orientation). The starting orientation for 14/17QI (Q) yields 9.5 and 1.4 kb KpnI fragments. The Q and I orientations bands are indicated by labeled arrows. The 1.4 kb KpnI fragment of 14/17QI is also indicated by a labeled arrow.

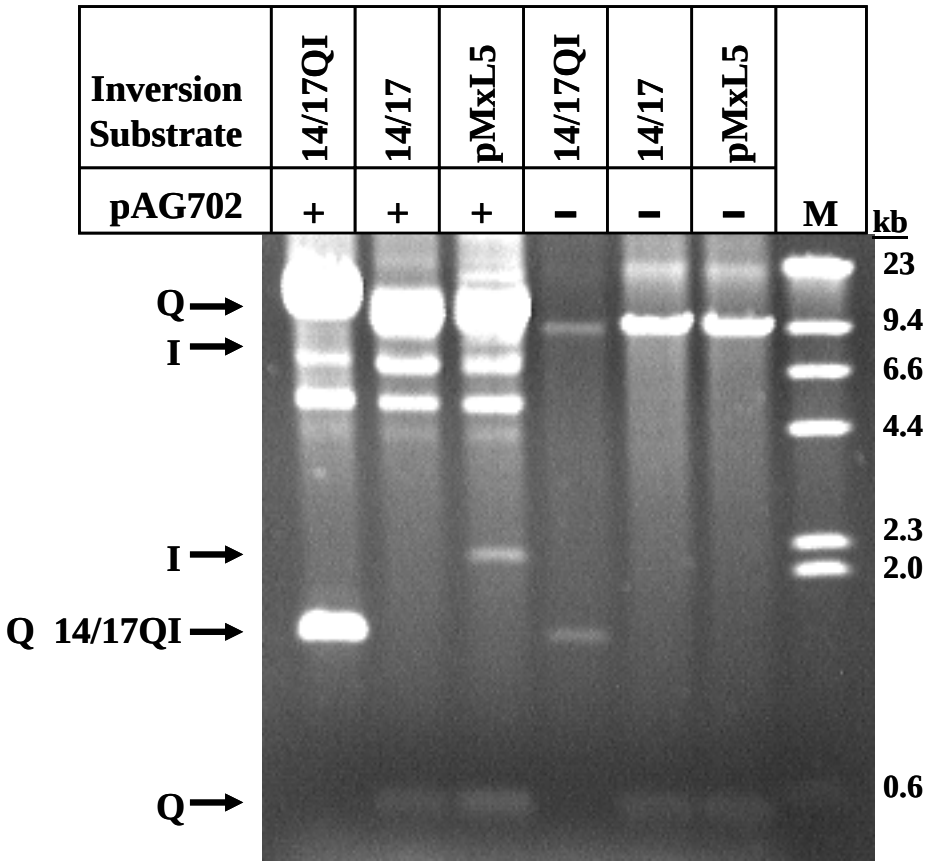

Inversion Substrate
14/17QI
14/17
pMxL5
14/17QI
14/17
pMxL5
pAG702
+
+
+
-
-
-
M
kb
23
9.4
6.6
4.4
2.3
2.0
0.6
Q
I
I
Q 14/17QI
Q

Figure 2.5: MBP-Piv binding to ds *invR* versus ss *invR*. The 5% non-denaturing gel on the left shows a binding assay with MBP-Piv and labeled ds invR (created by annealing oligonucleotides invRB and invRT). The molar excess of MBP-Piv to labeled ds oligonucleotide is indicated above the gel. The unbound and bound DNA is indicated by labeled arrows. As indicated by the arrow, there is some labeled invRB that is not annealed to invRT in the ds *invR* gel which results in ss DNA binding in the presence of ds and ss DNA. What appears to be a band between the unbound ds DNA and bound ssDNA is an artifact as it is seen in the absence of MBP-Piv. The 5% non-denaturing gel on the right shows a binding assay with MBP-Piv labeled ss bottom strand *invR* (invRB). The molar excess of MBP-Piv to labeled invRB is indicated above each lane. The unbound and bound DNA is indicated by labeled arrows. The bands above the bound ssDNA are MBP-Piv interacting with unincorporated [γ-^{32}P]ATP.

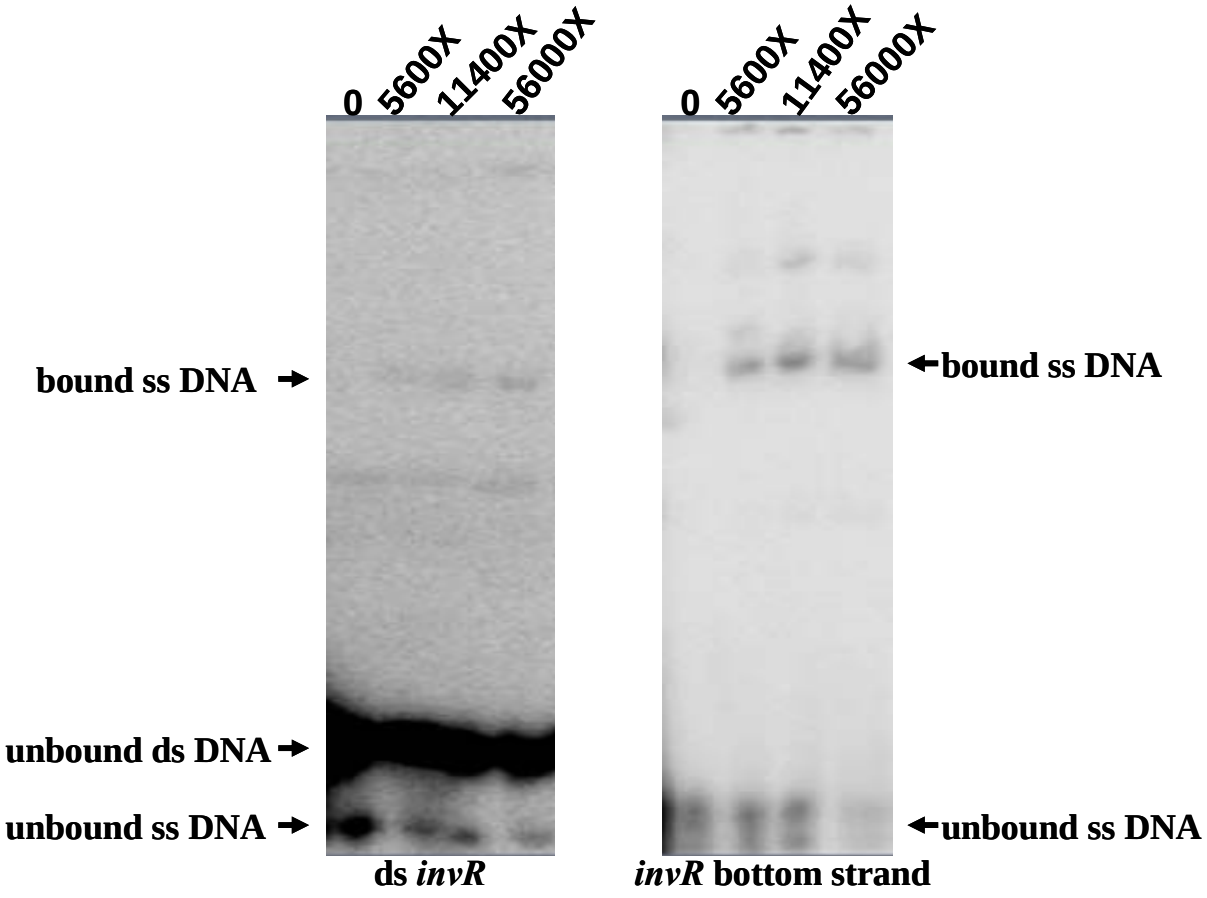

0
5600X
11400X
56000X
bound ss DNA
unbound ds DNA
unbound ss DNA
ds invR
0
5600X
11400X
56000X
bound ss DNA
unbound ss DNA
invR bottom strand

Figure 2.6: Specificity of MBP-Piv binding to *invR* bottom strand. Competition DNA binding assay was performed as described in Materials and Methods on invRB. A 5% nondenaturing gel is shown on which reactions containing ^{32}P-labeled invRB oligonucleotide, 1140 nM MBP-Piv and cold competitor DNA were electrophoresed. The –fold molar excess of unlabeled specific, invRB, or unlabeled non-specific, LacZDra, oligonucleotide to ^{32}P-labeled invRB is indicated above each lane. The addition of MBP-Piv is indicated by a + sign. Unbound and bound ^{32}P-labeled invRB is indicated by labeled arrows.

Non-specific competitor:Oligo	–	–	–	–	–	–	–	–	10	50	100	500
Specific competitor:Oligo	–	–	10	50	100	500	–	–	–	–	–	–
MBP-Piv	–	+	+	+	+	+	–	+	+	+	+	+

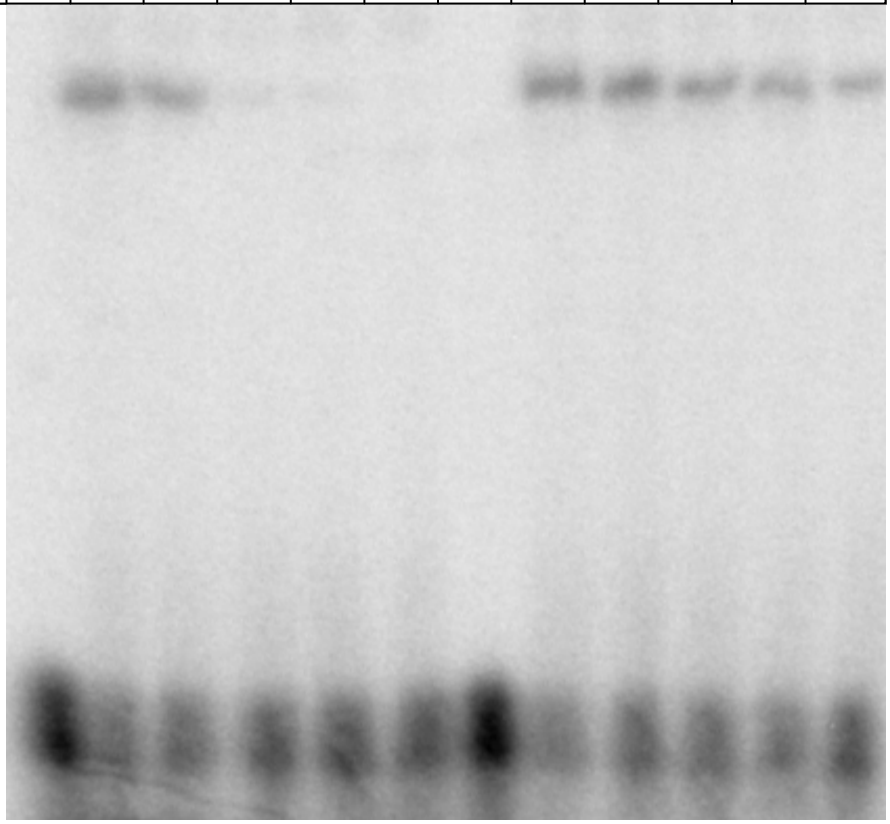


Figure 2.7: Specificity of MBP-Piv binding to *invR* top strand. Competition DNA binding assay was performed as described in Materials and Methods on invRT. A 5% nondenaturing gel is shown on which reactions containing ^{32}P-labeled invRT oligonucleotide, 1140 nM MBP-Piv and cold competitor DNA were electrophoresed. The –fold molar excess of unlabeled specific, invRT, or unlabeled non-specific, LacZDra, oligonucleotide to ^{32}P-labeled invRT is indicated above each lane. The addition of MBP-Piv is indicated by a + sign. Unbound and bound ^{32}P-labeled invRT is indicated by labeled arrows.

Non-specific competitor:Oligo	–	–	–	–	–	–	–	–	10	50	100	500
Specific competitor:Oligo	–	–	10	50	100	500	–	–	–	–	–	–
MBP-Piv	–	+	+	+	+	+	–	+	+	+	+	+

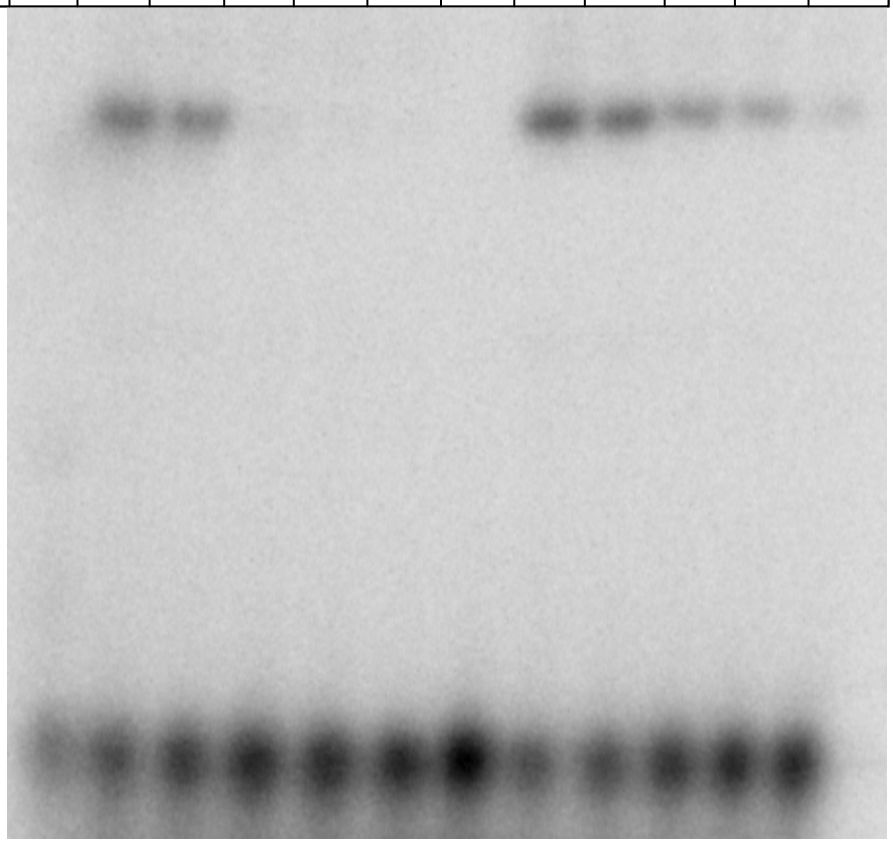
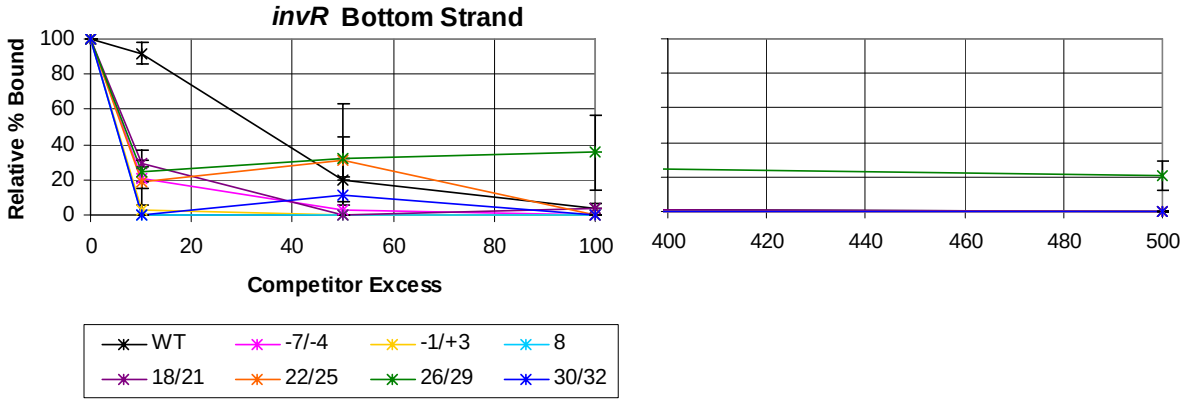


Figure 2.8: Graphs of competition assays using mutant *invR* substrates that are positive for inversion. (A). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* bottom strand, invRB, in the presence of mutant invRB oligonucleotides corresponding to the mutant *invR* substrates that are positive for inversion. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRB. The WT graph line indicates competition between unlabeled and ^{32}P-labeled invRB. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of each competitor is shown. (B). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* top strand, invRT, in the presence of mutant invRT oligonucleotides corresponding to the mutant *invR* substrates that are positive for inversion. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRT. The WT graph line indicates competition between unlabeled and ^{32}P-labeled invRT. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of each competitor is shown.

A.


invR Bottom Strand
Relative % Bound
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
-7/-4
-1/+3
8
18/21
22/25
26/29
30/32

B.

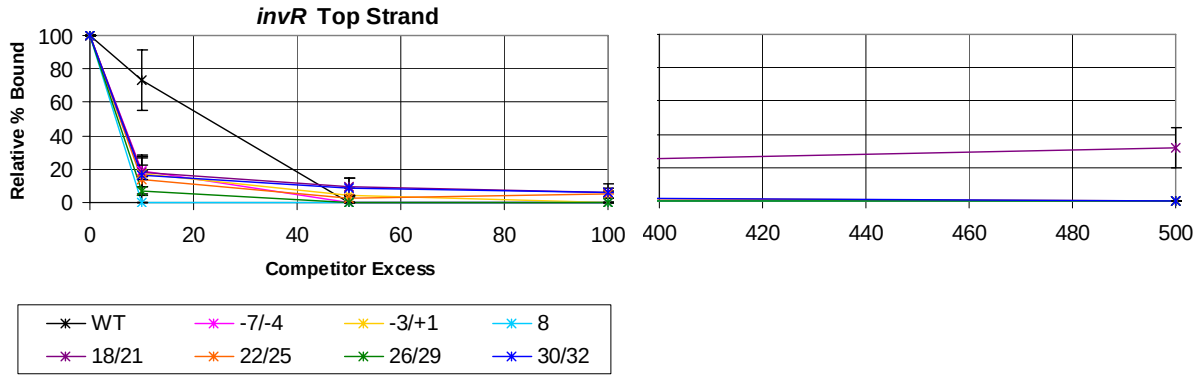
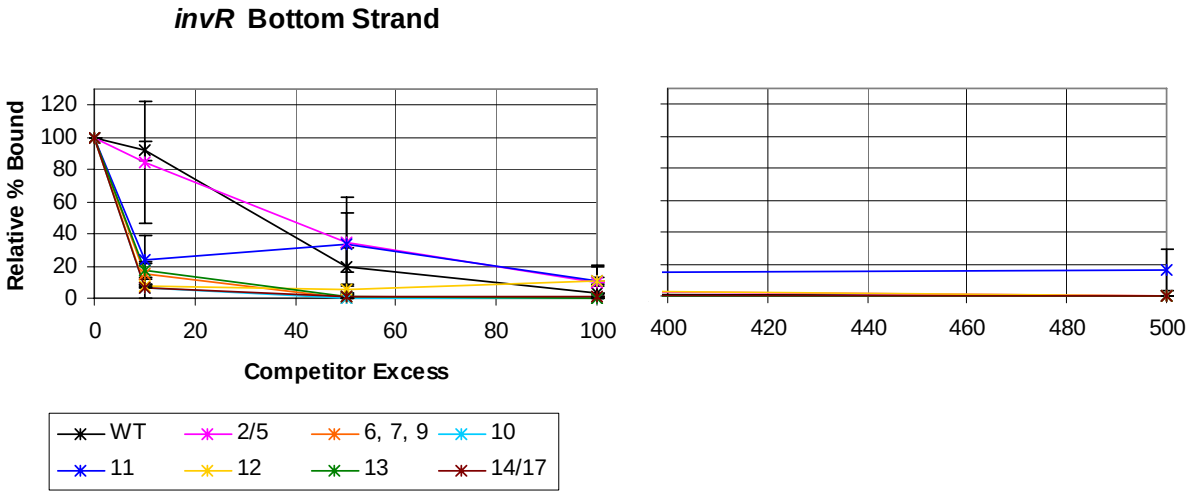

invR Top Strand
Relative % Bound
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
-7/-4
-3/+1
8
18/21
22/25
26/29
30/32

Figure 2.9: Graphs of competition assays using mutant *invR* substrates that are negative for inversion. (A). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* bottom strand, invRB, in the presence of mutant invRB oligonucleotides corresponding to the mutant *invR* substrates that are negative for inversion. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRB. The WT graph line indicates competition between unlabeled and ^{32}P-labeled invRB. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of each competitor is shown. (B). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* top strand, invRT, in the presence of mutant invRT oligonucleotides corresponding to the mutant *invR* substrates that are negative for inversion. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRT. The WT graph line indicates competition between unlabeled and ^{32}P-labeled invRT. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of each competitor is shown.

A.


invR Bottom Strand
Relative % Bound
120
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
2/5
6, 7, 9
10
11
12
13
14/17

B.

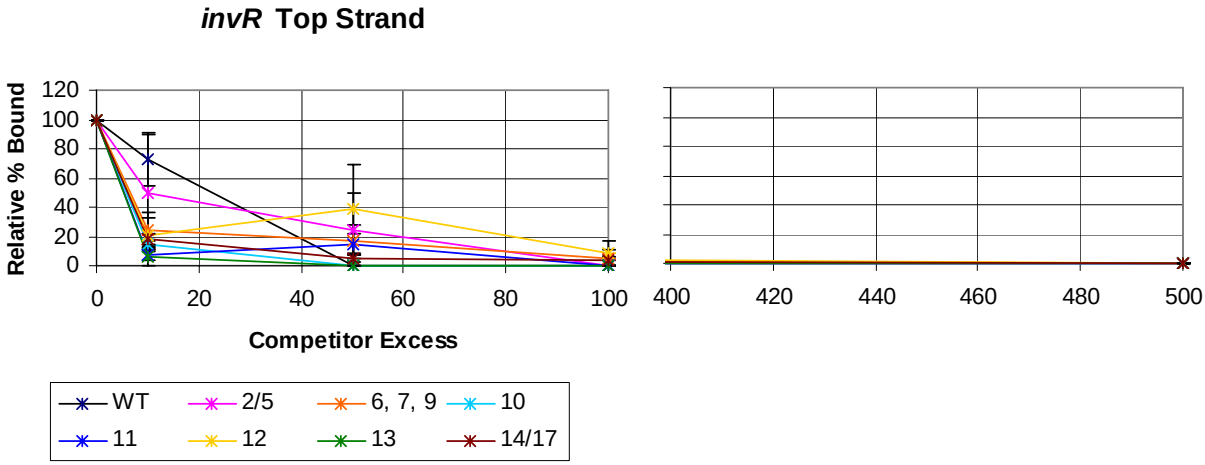
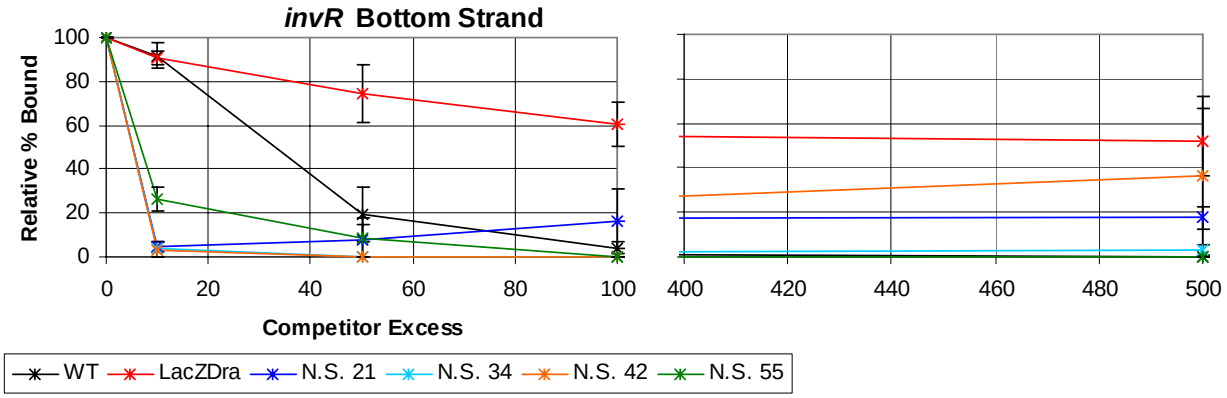

invR Top Strand
Relative % Bound
120
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
2/5
6, 7, 9
10
11
12
13
14/17

Figure 2.10: Graphs of competition assays using non-specific competitors of varying lengths. (A). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* bottom strand, invRB, in the presence of non-specific oligonucleotides. LacZDra is a non-specific competitor that is 34 nt in length. N.S. indicates non-specific oligonucleotide and the number indicates the length of the non-specific oligonucleotide (U21-U55). WT indicates competition between unlabeled and ^{32}P-labeled invRB. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRB. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of competitor is shown. (B). Graph of competition data derived by quantitation with Typhoon Imager and ImageQuant software of MBP-Piv binding to wild-type *invR* top strand, invRT, in the presence of non-specific oligonucleotides. LacZDra is a non-specific competitor that is 34nt in length. N.S. indicates non-specific oligonucleotide and the number indicates the length of the non-specific oligonucleotide. WT indicates competition between unlabeled and ^{32}P-labeled invRB. The relative % bound DNA (y-axis) was calculated as described in Materials and Methods. The x-axis indicates the –fold molar excess of unlabeled competitor to ^{32}P-labeled invRT. The competition assays were done in triplicate and the resulting standard error for the relative % bound in the presence of competitor is shown.

A.


invR Bottom Strand
Relative % Bound
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
LacZDra
N.S. 21
N.S. 34
N.S. 42
N.S. 55

B.

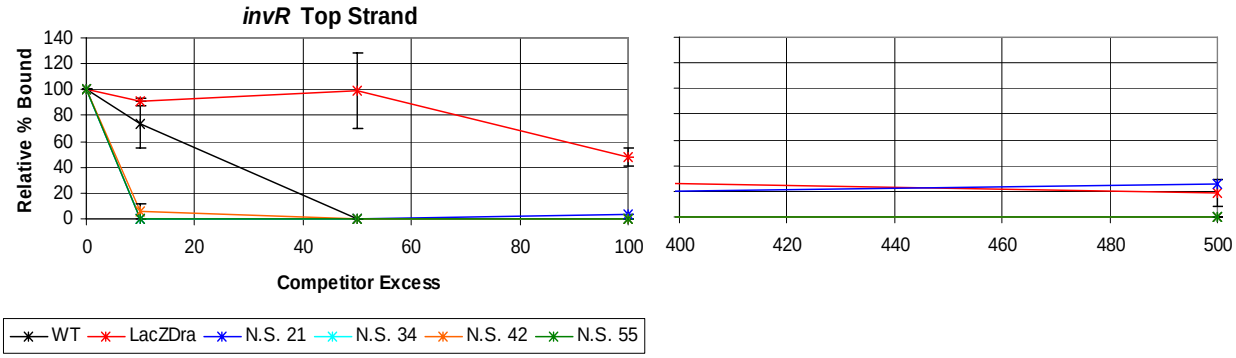
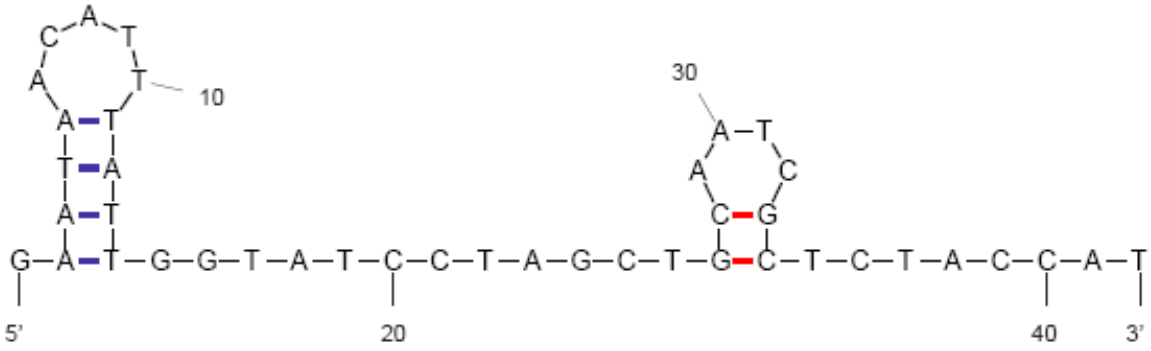
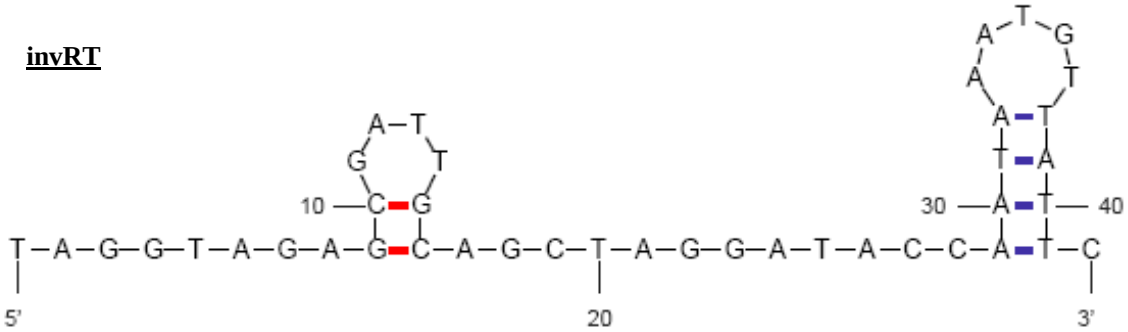

invR Top Strand
Relative % Bound
140
120
100
80
60
40
20
0
0
20
40
60
80
100
400
420
440
460
480
500
Competitor Excess
WT
LacZDra
N.S. 21
N.S. 34
N.S. 42
N.S. 55

Figure 2.11: Secondary structure predictions for specific and nonspecific oligonucleotides. The sequences of the oligonucleotides invRB, invRT, U55 (N.S. 55), and LacZDra were put into the Dinamelt Server at dinamelt.bioinfo.rpi.edu to predict their possible 2° structures. The most favorable structure for each oligonucleotide is shown below. The figures were isolated from the Dinamelt Server. Red lines indicate base pairing between guanine and cytosine. Blue lines indicate base pairing between adenine and thymine.

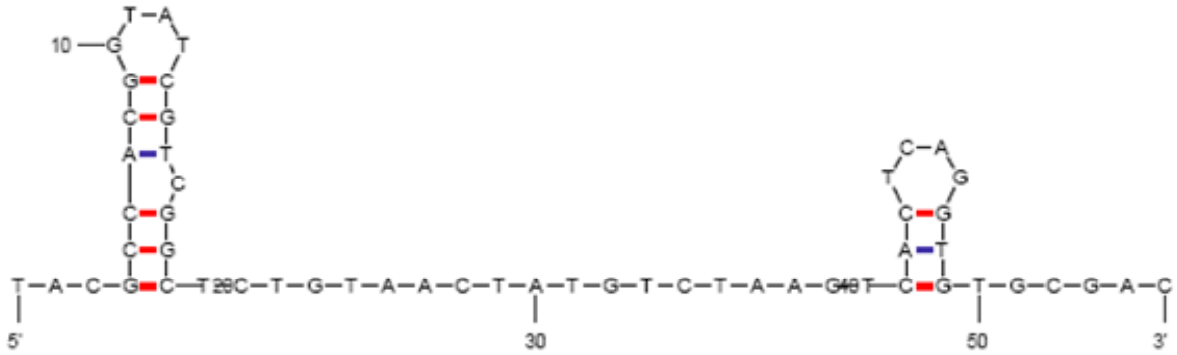
invRB



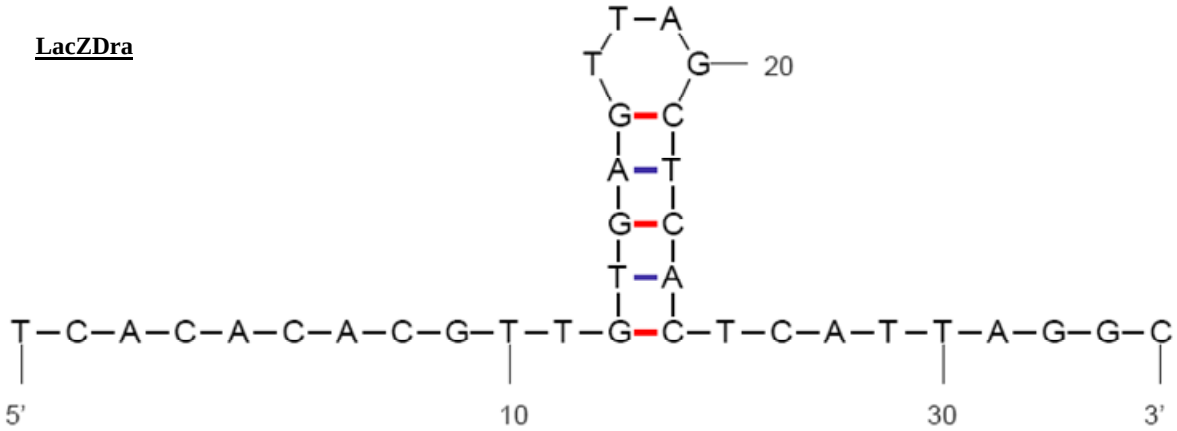
invRT



N.S. 55



LacZDra



Chapter 3

Conserved DEDD Motif is Required for MooV-Mediated Excision and Insertion of IS*492*[1]

[1] Carpenter, C. D. and Karls, A. C. 2008. To be submitted to the *Journal of Bacteriology*.

Abstract

MooV catalyzes the precise excision and site-specific insertion of the mobile element, IS*492*, at the *eps* locus in *Psuedoalteromonas atlantica* thereby controlling phase variation of extracellular polysaccharide. MooV is a member of the Piv/MooV family of DNA recombinases, which includes transposases of the IS*110*/IS*492* family of insertion sequences and the site-specific invertase Piv. The recombinases of the Piv/MooV family have a conserved tetrad of acidic amino acid residues, DEDD. The role of the conserved DEDD-motif in MooV was addressed by generating MooV variants that are substituted individually at these acidic residues with a non-polar amino acid. These variants have been shown to be incapable of excising and circularizing IS*492* and inserting IS*492* at its target site. Because we were unable to detect binding by wild-type MooV, we were unable to determine if the amino acid substitutions affected catalysis and/or DNA binding. We designed an assay to isolate hyperactive insertion variants of MooV that may have higher affinities for the ends of the element and/or target sites.

Introduction

The insertion and excision of IS*492*, catalyzed by the recombinase MooV, controls phase variation of extracellular polysaccharide (EPS) in *P. atlantica* (Figure 3.1) (1). Insertion of IS*492* into a glucosyl transferase gene, *epsG,* is site-specific, disrupting expression of EPS, thereby switching colony morphology from mucoid (EPS^+) to crenated (EPS^-) (1). IS*492* targets the 7 bp sequence $^{5'}$CTTGT/TA$^{3'}$ found within *epsG*, and, upon insertion, the 5 bp sequence $^{5'}$CTTGT$^{3'}$ is duplicated. Excision of IS*492* from *epsG* is precise, restoring expression of EPS and switching colony morphology from crenated (EPS^-) to mucoid (EPS^+). Excision results in the circularization of IS*492* with the right and left ends separated at the junction by one of the 5 bp direct repeats. Whereas IS*492* insertions and excisions are characteristic of site-specific recombination, MooV is a member of the Piv/MooV family of recombinases that mediate transposition of IS*110* family insertion sequences and site-specific DNA inversion.

Alignment of recombinases of the Piv/MooV family indicate a conserved DEDD tetrad (2, 15). Molecular modeling of the tertiary structure of Piv, a site-specific invertase with 25-35% amino acid identity and 45-55% similarity to the transposases of the Piv/MooV family, predicts a ribonuclease H (RNase H) fold like that found within DDE-motif transposases (15). The conserved DEDD tetrad in Piv is required for inversion, and the conserved residues are positioned suitably in the Rnase H-fold to coordinate divalent metal ions that are required for catalysis of DNA cleavage and strand transfer by DDE-motif transposases (2). Molecular modeling of the amino-terminus of MooV indicates a similar tertiary structure to that of Piv, suggesting that the recombinases may share a similar mechanism for recombination (15). Despite this similarity in conserved catalytic amino acid motif and structural domains with the DDE-motif transposases, MooV-mediated recombination has many features of site-specific

recombination. Therefore, elucidation of its mechanism of catalysis will contribute to the understanding of specialized DNA recombination.

This paper addresses whether the conserved DEDD residues are required in MooV-mediated excision and insertion of IS*492*. MooV variants with individual substitutions of alanine at each potential catalytic residue do not retain the ability to excise and circularize IS*492* nor do they retain the ability to insert IS*492*. These results indicate that the residues are required for MooV-mediated recombination. We cannot conclude that these residues are catalytic residues, however, because they may be essential for binding.

We also describe herein an assay to isolate MooV variants, generated by random mutagenesis, with an ability to insert IS*492* in *E. coli* at a much higher frequency than wild-type MooV. We expect that some variants may bind the target site with higher affinity thus allowing identification of the DNA recognition and binding determinants for MooV. The amino acid changes that result in higher frequency of insertion by MooV may also define the catalytic domain.

Materials and Methods

Bacterial strains. *E. coli* DH5α and TOP10 were cultured as described (12). *E. coli* DH5α-λ*pir*, which is DH5α lysogenized with λ*pir*, and *E. coli* BW23474 [Δ*lac-169*, *rpoS(am)*, *creC510*, *hsdR514*, *uidA* (Δ*Mlu*)::*pir116*, *endA*, *recA1*] were kindly provided by E. Stabb. *E. coli* CC118-λpir [Δ(*lac-pro*) *argE*(Am) *recA56*, *nalA* Rf^R (λpir)] was generously provided by L. Shimkets.

Media, enzymes, and reagents. All *E. coli* strains were grown on Luria-Bertani agar (LBA) or in LB broth (LB) at the following drug concentrations for plasmid selection and maintenance: ampicillin (Ap) at 60 to 100 μg/ml, chloramphenicol (Cm) at 50 μg/ml, kanamycin (Kan) at 50

μg/ml, nalidixic acid (Nal) at 30 μg/ml, spectinomycin (Spc) at 50 μg/ml, streptomycin (Sm) at 10 or 25 μg/ml, tetracycline (Tc) at 15 μg/ml, and trimethoprim (Tp) at 10 μg/ml. When indicated, 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) (40 μg/ml) was added to plates. When indicated, D-lactose monohydrate or IPTG were added to plates at 0.05% and 50 μg/ml, respectively. All chemicals and rabbit anti-chicken IgG alkaline phosphatase conjugate were purchased from Sigma Chemical Co (St. Louis, MO). T4 DNA ligase and restriction enzymes were purchased from New England Biolabs (Beverly, MA). Acrylamide, PVDF membrane, and protein molecular weight markers (Prestained SDS-PAGE Standards version) were purchased from BioRad (Hercules, CA). *Pfu* DNA polymerase used in PCR was obtained from Stratagene (La Jolla, CA).

Plasmids and plasmid construction. The sequences of all oligonucleotides used in this work are listed in Table 3.1. All PCR products created for cloning purposes were amplified by using Pfu DNA polymerase. The thermal cycling conditions were as follows: 95°C 3 min followed by 30 cycles of 95°C 30 sec, the T_m of the oligonucleotide with the lesser T_m value minus 5°C for 30 sec, and 72°C for 2.5 min. Products were digested with the appropriate restriction enzymes and ligated into the indicated plasmids using T4 DNA ligase.

Two plasmids containing an IS*492* derivative were used as donor plasmids: pDB22 and pAG957. IS*492*Δ*mooV::lacZα* was amplified from pAG994 (described previously, (12),) using oligos 5 and 6 and ligated into the *Bgl*II site of pDB20 to create pDB22. pDB20 was created by adding a unique *Bgl*II restriction site between the chloramphenicol acetlytransferase gene (*cat*) and its promoter in pACYC184 (New England Biolabs) as described previously (11). pAG957 containing IS*492*Δ*mooV::cat* was described previously (12). To create pAG951, 134 bp of *epsG*, including the 7 bp target site, was amplified from a mucoid colony of *P. atlantica* using

the primers, epsL 58 and epsR 76, and cloned into the TA-TOPO cloning vector pCR2.1 (Invitrogen, Carlsbad, CA).

Site-directed mutagenesis was used to change the targeted codon to encode an alanine within *mooV*. The creation of pAG990 [*mooV*(D14A)], pAG991 [*mooV*(D100A)], and pAG986 [*mooV*(E146A)] was described previously (11). *mooV* (E58A) and *mooV* (D103A) were created using overlap extension with the following primer sets: E58A2 (or D103A2) + MooVBamH1 and E58A1 (or D103A1) + MooVHindIII. The products of the previous PCR reactions were used as templates in the final extension reaction. The resulting PCR products were digested with HincII and BsmB1 and ligated into pAG991 digested with HincII and BsmB1 to create pCDC5 [*mooV*(E58A)] and pCDC6 [*mooV*(D103A)]. In order to provide the MooV mutants in trans, the mutated *mooV* genes were amplified using MooVBamH1 and MooVHindIII, digested with BamH1 and HindIII, and ligated into pAG900, the expression vector for wild-type MooV (12), replacing wild-type *mooV* with the mutant to create pCDC7 [*mooV*(D14A)], pCDC8 [*mooV*(E58A)], pCDC9 [*mooV*(D100A)], pCDC10 [*mooV*(D103A)], pCDC11 [*mooV*(E146A)]. The plasmids were sequenced to confirm inserted PCR products.

The conjugal helper plasmid, pEVS104, and the *oriT*-containing plasmid, pVSV107, were gifts from E. Stabb (13), (4). To create a mobilizable plasmid containing *epsG* sequence, pAG951 was digested with NotI and SpeI, and the resulting restriction fragment containing 134bp of *epsG* was ligated into pVSV107 replacing its NotI/SpeI restriction fragment. The resulting plasmid, pMMP1, was sequenced to confirm the inserted PCR product.

IS*492* circle junction PCR assay and western blot for wild-type and mutant MooV expression. Chemically competent DH5α containing pDB22 were transformed with pAG900, pCDC7, pCDC8, pCDC9, pCDC10, or pCDC11 and incubated at 25°C for 4 days on LBA plates

containing the appropriate antibiotics and 0.05% lactose to induce the expression of MooV. After 4 days of incubation, the transformants were placed at 4°C overnight. One transformant was picked from each plate, suspended in 30 µl of TE, boiled for 10 min, and the lysate was cleared by centrifugation. The cleared lysate was used in a PCR reaction with CJ50A and CJ250B primers to detect IS*492* circle junctions as previously described (7). Primers DB22EcorI and CJ250A were used in a PCR reaction to detect template. Aliquots (25 µl) of each reaction were electrophoresed on a 1.5 % agarose gel and visualized with ethidium bromide.

A western blot was performed as described previously to detect wild-type and mutant MooV expression (2, 7).

Random mutagenesis of *mooV* and isolation of possible hyperactive MooV variants. To create hyperactive variants of MooV, *mooV* was mutated by PCR mutagenesis using *Taq* DNA polymerase, $MnCl_2$, and an excess of one dNTP nucleotide over the others (5). The primers used in the PCR reaction were MooVBamH1 and MooVHindIII. After amplifying *mooV* from pAG900, the resulting PCR product and pAG900 were digested with BamHI and HindIII and ligated together so that the mutated *mooV* replaced wild-type *mooV*.

DH5α containing pDB22 were transformed with pAG900 containing mutated *mooV* (pAG900M), plated on LBA plates containing 0.05% lactose and 40 µg/ml X-gal, and incubated for 5 days at 25°C. The insertion of IS*492*Δ*mooV*::*lacZα* downstream of a promoter resulted in colonies with blue papillae or blue colonies. After isolating the plasmids from colonies containing insertions, pAG900M was isolated from pDB22 by restriction digest with BglII which digests pDB22 but not pAG900M, followed by plasmid dilution and transformation into DH5α.

Conjugation assays to determine insertion frequency of IS*492*Δ*mooV*::*lacZα* by wild-type and mutant MooV. To calculate the insertion frequency of wild-type MooV versus MooV

variants, CC118-λpir cells containing the donor plasmid, pAG957, and the mobilizable, target plasmid, pMMP1 were transformed with the expression vector, pAG900 or pAG900M, plated on LBA plates containing the appropriate antibiotics and 0.05% lactose, and incubated for 5 days at 25°C. Approximately 500 transformants were scraped from the plates, suspended in 500 µl of LB, pelleted, and resuspended in 40 µl of LB. The 40 µl of tranformants, 40 µl of 7.5 X 10^9 cells of BW23474 containing the helper plasmid, pEVS104, and 40 µl of 7.5 X 10^9 cells of the recipient, DH5α-λpir were spotted onto LBA plates and incubated at 25°C overnight.

The spots were suspended in 500 µl of LB, pelleted, and resuspended in 100 µl of LB. The suspension was plated on LBA plates containing trimethoprim and nalidixic acid to select for transcongugants and on LBA plates containing trimethoprim, nalidixic acid, and chloramphenicol to select for the insertion of IS*492*Δ*mooV*::*cat* into pMMP1.

Insertion frequency was calculated as the number of insertions per ml divided by the number of transcongugants per ml.

Results

DEDD motif residues are all required for MooV excision and circularization of IS*492*. It has been shown previously that the conserved tetrad, DEDD, is required for catalytic activity by Piv, a site-specific invertase belonging to the Piv/MooV family of recombinases (2). To determine if the conserved DEDD tetrad is required for catalytic activity of MooV, we substituted the conserved residues with alanine creating individual variants of MooV with D14A, E58A, D100A, or D103A. The non-conserved E146 residue was also changed to alanine because it is in the correct spacing with D14 and D100 to form a DDE-motif. The MooV variants, provided in trans, were assayed for their ability to excise and circularize

IS*492*Δ*mooV*::*lacZα* in vivo. PCR was used to detect the circular form of IS*492*Δ*mooV*::*lacZα* and as a control measure to determine the template concentration in the reaction mixture as described previously (7, 12). As shown in Figure 3.2, the MooV variants, D14A, E58A, D100A, D103A, and E146A were all negative for excision and circle formation of IS*492*. Western blot analysis indicates that D14A, E58A, D100A, and D103A are expressed at levels comparable to wild-type MooV (Figure 3.3), however, no MooV E146A was visible. This result suggests that substitution of E146 with alanine results in misfolding of MooV E146A such that it is targeted for degradation by host proteases (Figure 3.3).

DEDD motif residues are all required for MooV-mediated insertion of IS*492*. To determine if the DEDD motif residues are also required for insertion of IS*492*, an assay was designed to select for the insertion of IS*492* into its target. A mating-out assay was designed to select for the insertion of an IS*492* derivative with the majority of *mooV* replaced with the *cat* gene, including its promoter, into a mobilizable target plasmid (Figure 3.4). Transconjugants containing the mobilizable target plasmid with IS*492*Δ*mooV*::*cat* inserted were selected by plating LBA plates containing appropriate antibiotics. This assay is similar to that used with Tn*5* (16) and IS*903* (14) with the exception that the *tra* functions were provided by a helper plasmid that was transferred into the strain containing both the MooV expression plasmid and the target plasmid after growth at 25°C for 5 days to allow time for movement of IS*492*Δ*mooV*::*cat*. The insertion frequency of IS*492*Δ*mooV*::*cat* by wild-type MooV is 1.7×10^{-5}/cell. The insertion frequency in *E. coli* is higher than the insertion frequency of IS*492* into the *epsG* locus in *P. atlantica* (2.7×10^{-7}/cell/generation). The insertion frequency in *P. atlantica* takes into account the number of generations in a sample population, but the insertion frequency calculated in *E. coli* is calculated based on the number of transconjugants/ml (6). After determining the efficiency of the

conjugation assay to detect insertion in *E. coli*, the MooV variants were screened for their ability to catalyze insertion.

D14, E58, D100, and D103 are all required for the insertion of IS*492* (Table 3.2). As expected, the change of E146 to a non-acidic amino acid prevents the insertion of IS*492* most likely because of the effect the mutation has on the structure of the variant resulting in its degradation by host proteases (Figure 3.3).

Isolation of hyperactive MooV variants using a papillation assay. The conserved DEDD tetrad is required for both MooV-mediated excision and circularization of IS*492* and insertion of IS*492*; however, we have been unable to detect MooV binding to an IS*492* substrate in vitro. It is possible that the DEDD tetrad is essential for binding instead of catalysis of strand nicking and exchange. To facilitate MooV binding to its target site, we have designed an assay to isolate MooV variants that insert IS*492* at a higher frequency than wild type as it is likely that some of these hyperactive variants will bind the target site with higher affinity.

A papillation assay, similar to the one described for IS*903* (14), was used to screen for MooV derivatives that insert IS*492* at a higher frequency than wild type. The IS*492* derivative used in this assay has nearly all of *mooV* replaced by a promoter-less *lacZα* gene (insertion of IS*492*Δ*mooV*::*lacZα* downstream of a promoter in the chromosome or on a plasmid containing a target site is mediated by MooV variants that were generated by random mutagenesis and provided in trans from a compatible expression vector, pAG900M). Blue papillae or blue sectors on a colony grown on rich medium supplemented with lactose and Xgal indicated insertion of IS*492*Δ*mooV*::*lacZα* (Figure 3.5) (14). Cells containing IS*492*Δ*mooV*::*lacZα* in the absence of MooV were LacZ$^-$, and, in the presence of wild-type MooV only a small percentage (0.182% (+/-) 0.05%) of the colonies had papillae. Due to the small number of colonies

papillated in the presence of wild-type MooV, we isolated pAG900M from all colonies that were papillated to increase our chances of isolating a hyperactive insertion variant.

After isolating plasmids containing possible MooV variants, the variants were assayed for their frequency of insertion in the mating-out assay described above. To date, of the fifty MooV variants that have been assayed thus far, none exhibit hyperactive insertion of IS*492*.

Discussion

The conserved acidic tetrad DEDD is required for MooV-mediated recombination. It has been shown previously that the conserved acidic residue tetrad, DEDD, of the Piv/MooV family is required for the catalysis of Piv-mediated inversion (2). We investigated whether the conserved acidic residue tetrad is also required for the activity of the transposases in the Piv/MooV family by substituting the conserved acidic residues in the transposase, MooV, with the non-polar amino acid alanine. D14A, E58A, D100A, and D103A substitutions resulted in MooV variants that are unable to catalyze excision, circularization, or insertion of IS*492*. The MooV variant that substituted non-conserved amino acid E146 does not mediate excision and insertion of IS*492*, but the western blot of E146A indicates that the change of this acidic residue to a non-polar residue disrupts the structure of the variant such that it is rapidly degraded by host proteases. When the equivalent residue, E147, was mutated to alanine in Piv, it had very little effect on the inversion activity and was expressed as well as wild-type Piv in Western blot analysis indicating a difference in amino acid requirement for structure between the two recombinases.

Search for hyperactive MooV variants. Due to the inability to detect MooV binding to target sequences in vitro, we are unable to unequivocally state that the conserved DEDD motif in

MooV is a catalytic tetrad. To facilitate the binding of MooV to target sequences in vitro, we designed a papillation assay to detect hyperactive MooV variants that insert IS*492* at a higher frequency than wild type. It is likely that a hyperactive MooV variant will bind target sequences or the ends of the element with a higher affinity than wild type allowing the elucidation of the DNA recognition and binding sites for MooV. The amino acid changes resulting in a MooV variant with the ability to bind target sequences will be introduced into the DEDD tetrad MooV variants to determine the ability of those variants to bind target sequences. Because the DEDD residues are not essential for Piv binding to its target site, it is very likely that these residues will also not be essential for MooV binding based on the similarity between the two proteins. This approach is likely to yield a hyperactive variant that will improve our chances of detecting MooV activity in vitro and allow the isolation and characterization of reaction intermediates.

A working model for MooV-mediated recombination. If we assume that the four residues of the DEDD motif in MooV are required for catalysis of insertion and excision of IS*492* but not MooV binding, then, the MooV catalytic domain probably resembles that of RuvC. Like the catalytic domains of DDE-motif transposases, the catalytic domain of RuvC has its DEDD catalytic residues positioned within an RNase H-fold to coordinate divalent metal cations that facilitate the hydrolysis of the phosphodiester bond in substrate DNA, generating a free 3'OH (3). RuvC does not catalyze strand transfer following hydrolysis of substrate DNA; however, DDE-motif transposases use the 3'OH of cleaved donor DNA as the attacking nucleophile in strand transfer by retaining the 3' end of the cleaved donor DNA in the active site and binding the target DNA within the same active site as a substrate for the one-step transesterification reaction (8). It has been suggested that the third aspartic acid in the RuvC catalytic site may prevent the binding of both the cleaved 3'end and target DNA within the active site by

immobilizing the first nucleophile and/or the 3'OH end (2). An alternative reasoning for the lack of strand transfer by RuvC is the absence of a target strand for binding to the catalytic site (2).

In the working model for Piv-mediated inversion both hydrolysis-strand transfer and endonucleolytic activities for the DEDD active site are utilized in the reaction (2). Briefly, the model for Piv-mediated inversion begins with Piv mediated hydrolysis and transesterification leading to the formation of a Holliday junction structure which is then resolved by Piv endonucleolytic activity followed by ligation by host DNA ligase to repair the nicks (2). A similar model for MooV-mediated excision offers an explanation for how MooV precisely excises IS*492* without leaving behind the double-stranded breaks in the donor DNA that is a hallmark of classical transposition mediated by DDE-motif transposases. The model (Figure 3.6) shows MooV bound to one end of IS*492*, and, after mediating the hydrolysis of one strand of IS*492*, MooV uses the released 3'OH as the nucleophile to attack the other end of the element forming a single-stranded circle and a Holliday junction structure. Because the ends of IS*492* have no sequence homology with each other or the target site, we believe that MooV recognizes and binds the structure created by the hydrolysis and transesterification reactions. After binding to the DNA structure, MooV mediates the hydrolysis of the strands in the Holliday junction thereby resolving the Holliday junction structure. The other option for the resolution of the branched intermediate is for host functions to process the branch intermediates set up by transposase-mediated strand transfer as seen in the models for recombination mediates by IS*911* and IS*30* transposases (9, 10). DNA ligase provided by the host may repair the nicks created before and after the RuvC-like activity that leads to the resolution of the junction. We are currently testing the model utilizing both genetic and biochemical methods.

References

1. **Bartlett, D. H., M. E. Wright, and M. Silverman.** 1988. Variable expression of extracellular polysaccharide in the marine bacterium *Pseudomonas atlantica* is controlled by genome rearrangement. Proc Natl Acad Sci U S A 85:3923-27.
2. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J Bacteriol 187:3431-7.
3. **Craig, N., R. Craigie, M. Gellert, and A. M. Lambowitz.** 2002. Mobile DNA II, second ed. ASM Press, Washington D. C.
4. **Dunn, A. K., D. S. Millikan, D. M. Adin, J. L. Bose, and E. V. Stabb.** 2006. New rfp- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and *lux* expression in situ. Appl Environ Microbiol 72:802-10.
5. **Fromant, M., S. Blanquet, and P. Plateau.** 1995. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. Anal Biochem 224:347-53.
6. **Higgins, B. P.** 2006. Regulation of IS*492* Transposition in *Pseudoalteromonas atlantica*. University of Georgia, Athens.
7. **Higgins, B. P., C. D. Carpenter, and A. C. Karls.** 2007. Chromosomal context directs high-frequency precise excision of IS*492* in *Pseudoalteromonas atlantica*. Proc Natl Acad Sci U S A 104:1901-6.
8. **Kennedy, A. K., D. B. Haniford, and K. Mizuuchi.** 2000. Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity. Cell 101:295-305.

9. **Kiss, J., M. Szabo, and F. Olasz.** 2003. Site-specific recombination by the DDE family member mobile element IS*30* transposase. Proc Natl Acad Sci U S A 100:15000-5.

10. **Loot, C., C. Turlan, and M. Chandler.** 2004. Host processing of branched DNA intermediates is involved in targeted transposition of IS911. Mol Microbiol 51:385-93.

11. **Perkins-Balding, D.** 2000. Transposition of IS*492*: In vivo and in vitro characterization of a member of an atypical group of insertion sequences. Emory University, Atlanta.

12. **Perkins-Balding, D., G. Duval-Valentin, and A. C. Glasgow.** 1999. Excision of IS*492* requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. J Bacteriol 181:4937-48.

13. **Stabb, E. V., and E. G. Ruby.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae.* Methods Enzymol 358:413-26.

14. **Tavakoli, N. P., and K. M. Derbyshire.** 1999. IS*903* transposase mutants that suppress defective inverted repeats. Mol Microbiol 31:1183-95.

15. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. Mol Microbiol 39:641-51.

16. **Wiegand, T. W., and W. S. Reznikoff.** 1992. Characterization of two hypertransposing Tn*5* mutants. J Bacteriol 174:1229-39.

Table 3.1

Oligonucleotide	Sequence
moovBamH1	5’ GGCCGGATCCGCATTATGCAAGCAAGAGTT 3’
mooVHindIII	5’ CTCGAAGCTTAGACTATGGCGTCAATAGCTAA 3’
E58A1	5’ CGTATCGTCATTGCAGCAACTGG 3’
E58A2	5’ CCAGTTGCTGCAATGACGATACG 3’
D103A1	5’ GACAAACTAGCCGCTCAATTGATTG 3’
D103A2	5’ CAATCAATTGAGCGGCTAGTTTGTC 3’
CJ250A	5’ TTATCGGTCTCTCAACTGAGGAGC 3’
CJ250B	5’ TAAGAAAGTGGCTATTATTGCGTGC 3’
DB22EcoR1	5’ TTGCCATACGGAATTCCGG 3’
5	5’ GGCTCTCGAGCGGTACTGTCTTATCATCCTAATCG 3’
6	5’ CTCGCTCGAGCAGGAGGCTCTCTATTGTACAGC 3’

Table 3.2: Insertion frequencies of wild-type MooV and variants

MooV	**Insertion Frequency**
Wildtype	1.7×10^{-5}
D14A	Less than 1.3×10^{-7}
E58A	Less than 3.6×10^{-7}
D100A	Less than 7.1×10^{-8}
D103A	Less than 2.5×10^{-8}
E146A	Less than 5.9×10^{-8}

Figure 3.1: MooV-mediated insertion and excision of IS*492* controls phase variation of EPS in *P. atlantica*. MooV-mediated site-specific insertion of IS*492* into an essential *eps* gene, *epsG*, disrupts the expression of EPS resulting in a crenated (EPS-) colony morphology. A 5 bp direct repeat flanks IS*492* upon insertion (represented by black lines). Excision of IS*492* from *epsG* is precise, restoring *epsG* and switching colony morphology from crenated (EPS-) to mucoid (EPS+). When IS*492* excises, it forms a circle with the right and left ends brought together with one of the 5 bp direct repeats at the junction; although it is likely that the IS*492* circle is an intermediate in the reaction, it is unknown at this time whether or not this circular form of IS*492* can insert into *epsG*.

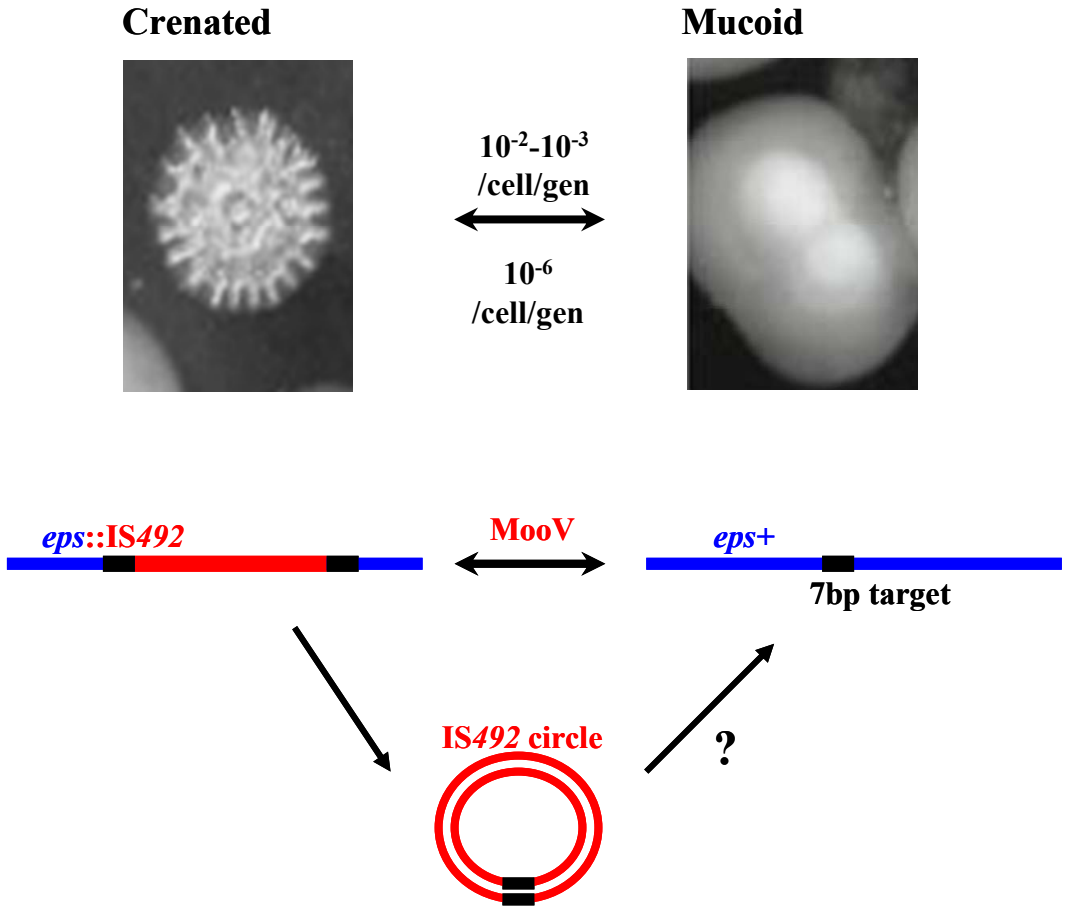
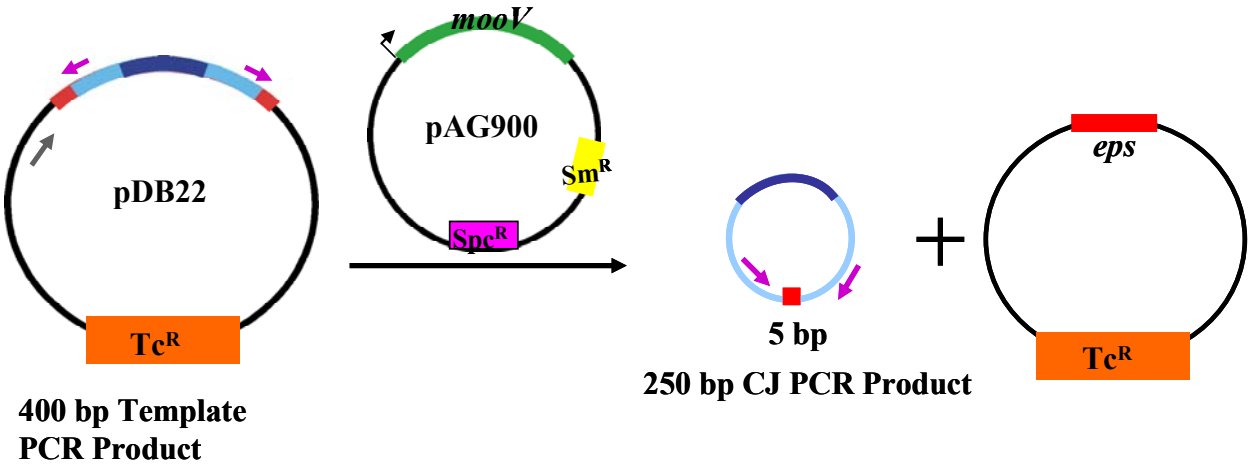

Crenated
Mucoid
10-2-10-3
/cell/gen
10-6
/cell/gen
eps::IS492
MooV
eps+
7bp target
IS492 circle
?

Figure 3.2: Excision and circle junction formation activity of MooV variants substituted within the DEDD motif and at a nonconserved glutamic acid residue. (A) Diagram of circle junction assay. MooV, provided in trans, catalyzes the excision and circularization of IS*492*Δ*mooV*::*lacZα* with the 5 bp flanking sequencing (red box) joining the ends. The pink arrows indicate the primers (CJ250A and CJ250B) used to detect the circle junction (CJ). The primers used in the template control reaction, DB22EcoR1 and CJ250A, are indicated by the gray and pink arrows, respectively. (B). 1.5% ethidium bromide stained agarose gel of circle junction PCR products. M = marker, no MooV = pDB22 not incubated with pAG900, WT = wild-type MooV, and D14A, E58A, D100A, D103A, and E146A = variant MooV. CJ = circle junction (C) 1.5% ethidium bromide stained agarose gel of template control PCR products. M = marker, no MooV = pDB22 not incubated with pAG900, WT = wild-type MooV, and D14A, E58A, D100A, D103A, and E146A = variant MooV. T = template.

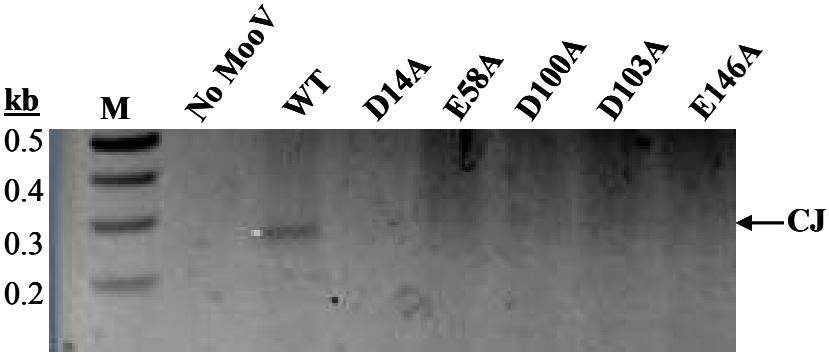
A.



400 bp Template PCR Product

250 bp CJ PCR Product

B.



C.

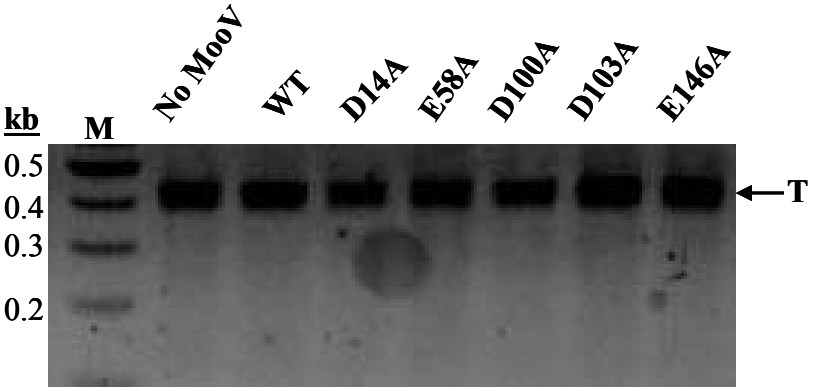


Figure 3.3: Expression of MooV variants in the strain used for circle junction assays.

Western blot utilizing anti-MooV antisera as primary antibody. pDB22 = negative control, no MooV expressed, WT = wild-type MooV, and D14A, E58A, D100A, D103A, and E146A = variant MooV. MooV is marked by an arrow.

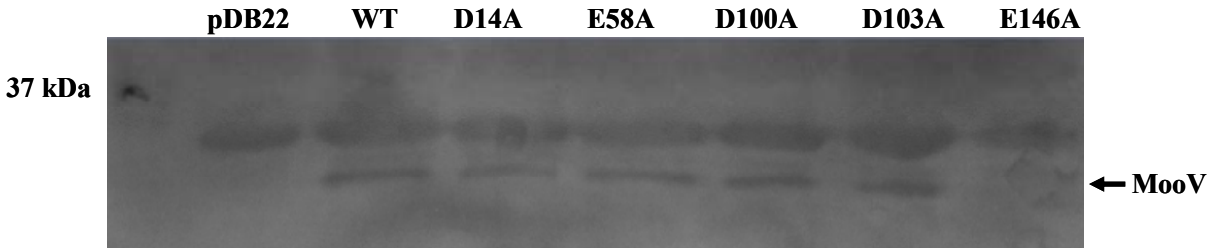

pDB22
WT
D14A
E58A
D100A
D103A
E146A
37 kDa
MooV

Figure 3.4: Diagram of mating out assay. The expression plasmid, pAG900 is transformed into an *E. coli* strain containing the target plasmid, pMMP1, and the donor plasmid, pAG957. MooV mediates insertion of IS*492*Δ*mooV*::*cat* into the *epsG* sequence on pMMP1. After incubation at 25°C, the transformants are spotted onto LB agar with the helper strain and the recipient strain. After incubation at 25°C overnight, the cells are suspended in LB and plated on LB agar plates containing trimethoprim (Tp) and nalidixic acid (Nal) to select for transconjugants and LB agar plates containing chloramphenicol (Cm), trimethoprim (Tp), and nalidixic acid (Nal) to select for insertion events. The insertion frequency is calculated by dividing the number of insertions per ml by the number of transconjugants per ml.

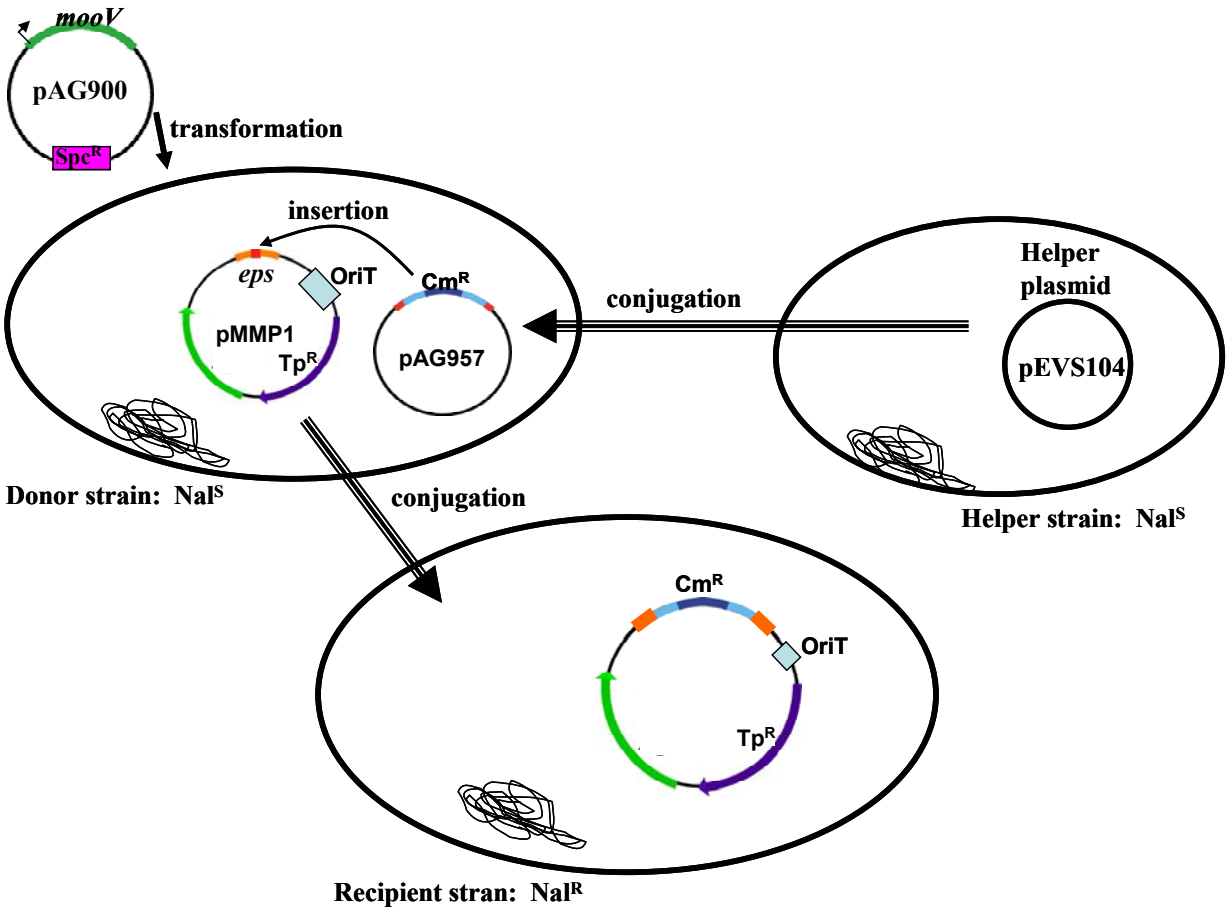
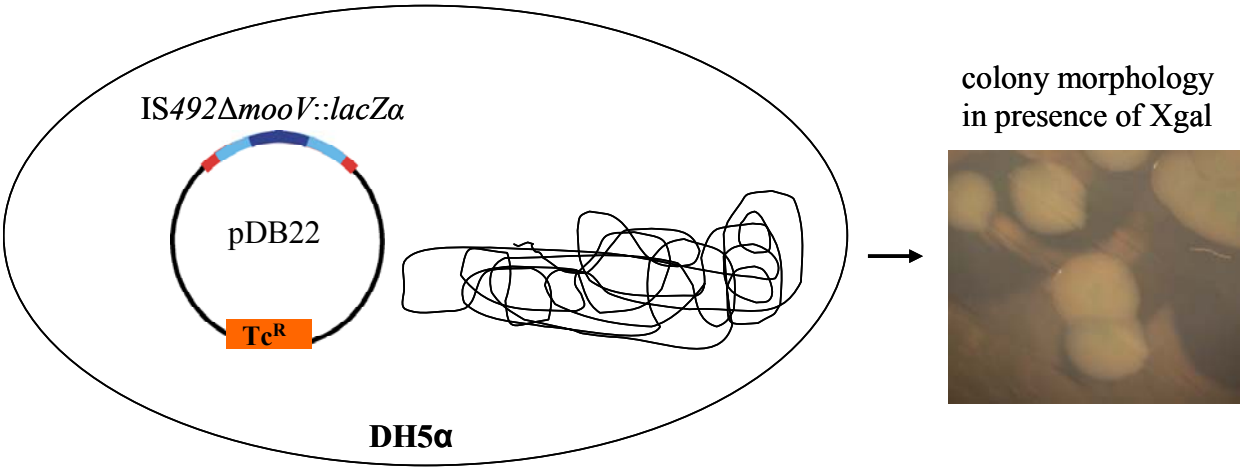

mooV
pAG900
SpcR
transformation
insertion
eps
OriT
CmR
pMMP1
TpR
pAG957
conjugation
Helper
plasmid
pEVS104
Donor strain: NalS
conjugation
Helper strain: NalS
CmR
OriT
TpR
Recipient stran: NalR

Figure 3.5: Diagram of papillation assay. (A). Diagram and picture of colony morphology of cells containing pDB22 (plasmids containing IS*492*Δ*mooV*::*lacZα*) without MooV provided. (B). Diagram and picture of colony morphology of cells with IS*492*Δ*mooV*::*lacZα* inserted in the chromosome downstream of a promoter. In some assays, *epsG* was provided on target plasmid, pAG951 downstream of a p*lac* promoter; therefore, insertion of IS*492*Δ*mooV*::*lacZα* into the target site would result in papillated colonies. It is also possible for IS*492*Δ*mooV*::*lacZα* to insert into pDB22 causing a head-to-tail dimer with a strong promoter formed at the junction of the dimer (A. Cottrell, unpublished data).

A.


IS492ΔmooV::lacZα
pDB22
TcR
DH5α
colony morphology
in presence of Xgal

B.

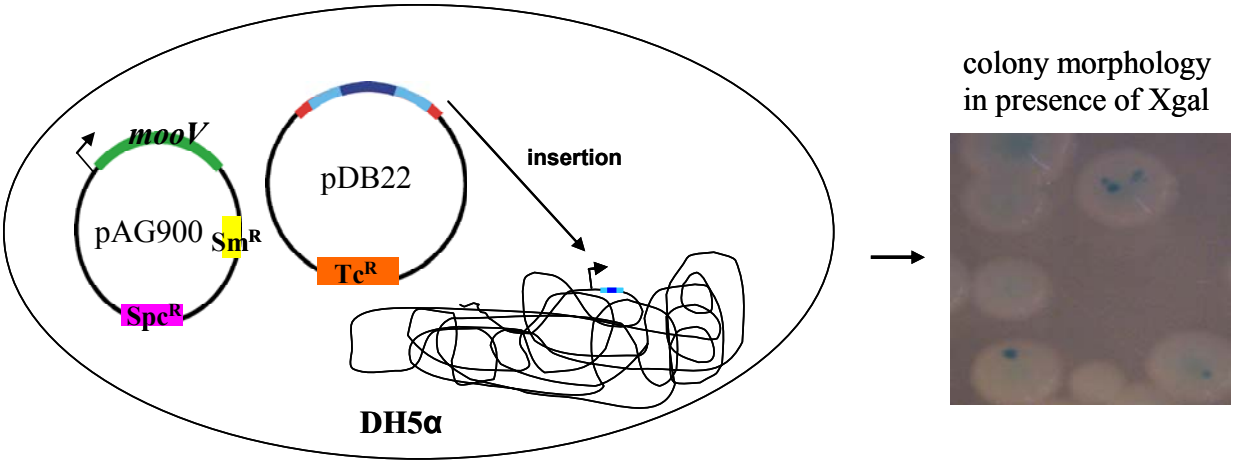
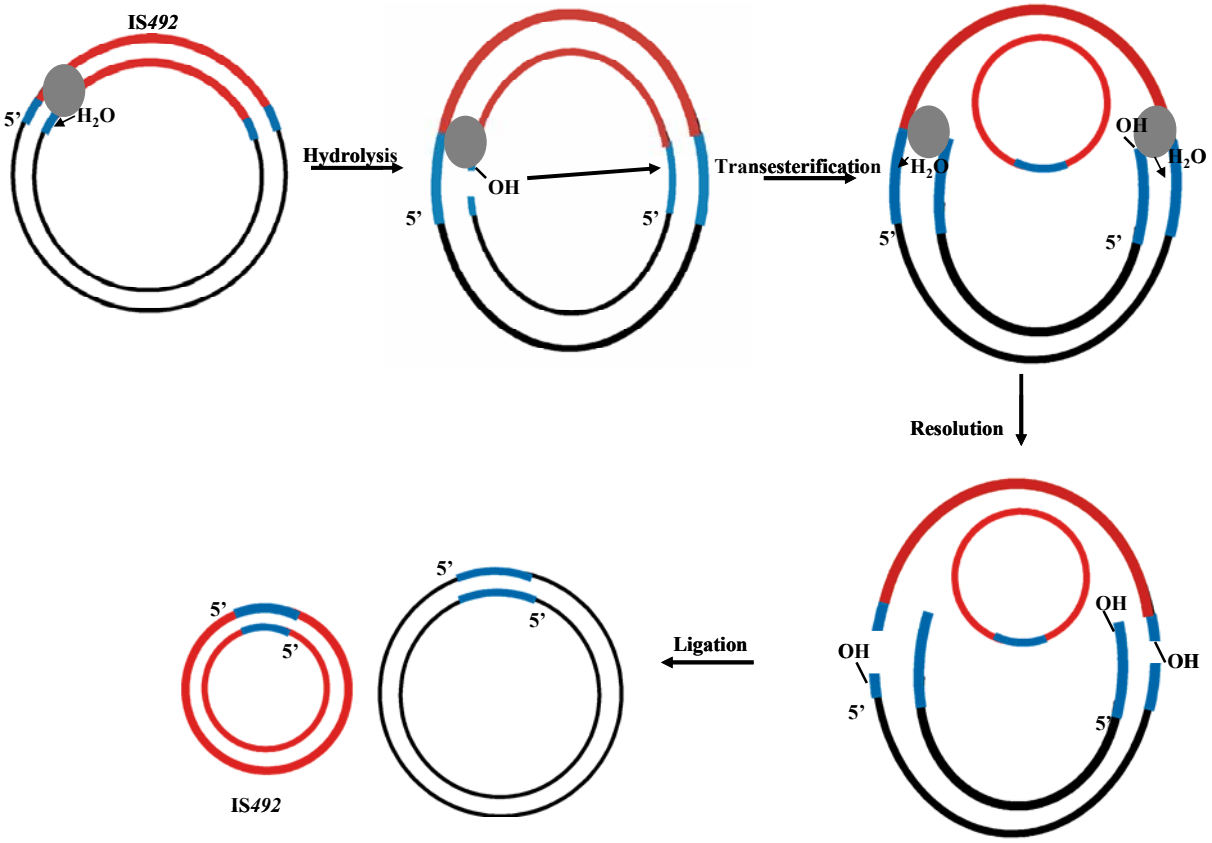

mooV
pAG900 SmR
SpcR
pDB22
TcR
insertion
DH5α
colony morphology
in presence of Xgal

Figure 3.6: Model of MooV-mediated excision. IS*492* is represented by red bars, the donor DNA is represented by black bars, and the 5 bp direct repeat is represented by blue bars. MooV is indicated by the gray circle. MooV-mediated hydrolysis is follow by one-step transesterification resulting in a Holliday junction intermediate. The Holliday junction intermediate is resolved by hydrolysis of the outer DNA strands of the junction followed by host DNA ligase activity to repair the nicks.


IS492
5'
H2O
Hydrolysis
OH
5'
5'
Transesterification
H2O
OH
H2O
5'
5'
Resolution
OH
OH
5'
5'
Ligation
5'
5'
5'
5'
IS492

Chapter 4

In Vitro Characterization of MooV-Mediated Recombination[1]

[1] Carpenter, C. D. and Karls, A. C. 2008. Unpublished.

Abstract

MooV catalyzes the site-specific insertion and precise excision of the insertion element, IS*492*, found in the marine bacterium *Pseudoalteromonas atlantica*. The conservative nature of MooV-mediated recombination suggests that MooV utilizes a recombination mechanism similar to the Serine- and Tyrosine- site-specific recombinases. However, MooV and the related DNA recombinases of the Piv/MooV family have a conserved DEDD-motif that is predicted to be positioned within a tertiary structure called an RNase H-like fold where these acidic residues can coordinate divalent metal ions like the DDE-motif of transposases in the retroviral integrase superfamily. This novel system of specialized recombination has elements of both site-specific recombination and DNA transposition and may reveal features relevant to other specialized recombination systems. We tried unsuccessfully to demonstrate MooV-mediated recombination of IS*492* in vitro using a variety of biochemical assays. We propose that either a required cofactor is missing from the in vitro assays or that wild-type MooV activity is inherently low in vitro.

Introduction

Specialized DNA recombination, which includes DNA transposition and site-specific recombination, creates genetic diversity through DNA insertions, deletions, and inversions. The study of model recombination systems is used in order to gain a better understanding of the molecular mechanisms for recombination in different systems and their regulation. Some examples of model systems that are currently being studied include: VDJ recombination that generates the diversity of immunoglobulin genes, transposition of numerous mobile elements in pathogenic bacteria that carry antibiotic resistance genes, and Fim-mediated site-specific DNA inversion that controls pilin phase variation and tissue tropism of uropathogenic *Escherichia coli* (reviewed in (6)). Just as specialized recombination is found in a diverse array of systems, the enzymes that mediate recombination are also diverse in their mechanism of catalysis. In order to gain an understanding of the reaction pathway in a particular specialized recombination system, the enzyme catalyzing the reaction (transposase or site-specific recombinase) and the DNA substrate can be studied at the biochemical level to define each step in the reaction.

Nearly all of the biochemically-characterized specialized recombinases utilize one of two distinct reaction mechanisms: two-step trans-esterification (site-specific recombinases) or DNA hydrolysis followed by one-step transesterification (transposases and retroviral integrases). DNA recombinases that mediate conservative site-specific recombination are divided into two families, tyrosine- (Y) and serine- (S) recombinases, based on the conserved amino acid that acts as the nucleophile in the first step of the two-step transesterification reaction (1, 12). The Y-recombinases form a covalent DNA-3'phosphotyrosine intermediate (1) and the S-recombinases form a covalent DNA-5'phosphoserine intermediate (12). The second step of the two-step trans-esterification reaction results in the restoration of the phosphodiester linkages of the exchanged

DNA strands via nucleophilic attack by the 5'OH- or 3'OH-terminated DNA strand that has been liberated by the recombinase.

In contrast to site-specific recombinases, most transposases and retroviral integrases catalyze recombination by the hydrolysis of the phosphodiester bonds at the ends of the mobile DNA element followed by one-step trans-esterification reactions in strand transfer (10). Many DNA transposases contain a conserved DDE (aspartate, aspartate, glutamate) triad of amino acids required for catalysis; the DDE triad comprises the catalytic pocket that binds metal cofactors such as magnesium, to facilitate the H_2O-mediated hydrolysis of donor DNA. During strand transfer, the 3'-OH groups from each end of the mobile element act as nucleophiles in the staggered attack of the phosphodiester backbone of the target DNA. This creates new phosphodiester bonds between 3' ends of the mobile element and the target DNA; in addition, gaps between the 3'-OH group in the target DNA and the 5' ends of the mobile element are generated. Host replication and/or repair machinery fill in the gaps, thus producing the target site duplication associated with the non-conservative recombination that is mediated by DNA transposases (10).

While both site-specific recombination and transposition are mechanisms for moving DNA, each contain intermediates in the reaction that are unique to that particular system; therefore, studying a specialized DNA recombination system at the biochemical level will define its mechanism as being similar to site-specific recombination or transposition. The transposases of the IS*110*-IS*492* family of insertion sequences, which include more that 50 recombinases, belong to the Piv/MooV family of recombinases whose primary amino acid sequences contain a conserved DEDD-motif instead of the DDE-motif shared by most DNA transposases and retroviral integrases (14). These DEDD-motif transposases share 25-35% amino acid identity

and 45-55% similarity with the site-specific DNA invertase Piv. Molecular modeling of the tertiary structure of Piv predicts a ribonuclease H (RNase H) fold similar to that found within DDE-motif transposases (22). The conserved DEDD tetrad in Piv is required for inversion, and the conserved residues are positioned suitably in the RNase H-fold to coordinate divalent metal ions like the acidic amino acid triad of DDE-motif transposases (4).

Alignment of predicted secondary structures for the MooV transposase with Piv indicates a similar tertiary structure, suggesting that the two recombinases may share a similar mechanism for recombination (22). MooV is the transposase for IS*492*, which is found in the marine bacterium, *Pseudoalteromonas atlantica* (17). IS*492* insertion and precise excision at a specific site within *epsG* controls phase variation of peripheral extracellular polysaccharide (EPS^p) in *P. atlantica* (3, 11). Although MooV is predicted to have an RNase H-fold similar to that found in DDE-motif transposases, MooV-mediated recombination has features of site-specific recombination as well as transposition reactions.

MooV-mediated excision of IS*492* from the *epsG* insertion site is precise, restoring the *epsG* gene and EPS^p expression (11, 17). Precise excision of a transposable element is highly unusual because the chemistry of DDE-motif transposase-mediated transposition dictates either release of the element from the donor DNA molecule, introducing double-stranded breaks at the donor site, or replication of the element, resulting in a copy at both the donor and target sites. The precise excision of IS*492* generates a circular form of the element that contains a 5 bp spacer sequence at the juncture of the element ends (17). The spacer corresponds to 5 bp of chromosomal target sequence (5'CTTGT3') that is duplicated at all IS*492* insertion sites in *P. atlantica* (11, 17). While the duplication of chromosomal target sequence is seen in transposition mediated by DDE-motif transposases, it also could result from site-specific insertion of the

circular form of IS*492* utilizing the 5-bp spacer (17). Unlike most transposable elements, IS*492* does not have terminal inverted repeat DNA sequences, therefore the 5 bp flanking sequence is likely to play an important role in MooV recognition and binding of the ends of IS*492* DNA.

Because the recombinase MooV appears to have the tertiary structure and catalytic residues of a transposase but mediates recombination events that resemble conservative site-specification recombination, the focus of this study is to define the molecular mechanism for MooV-mediated recombination. Experiments were performed to elucidate individual steps of the recombination reaction: DNA nicking or cleavage and strand transfer. The results of numerous attempts to detect MooV activity in vitro revealed that the system is currently intractable using the wild-type recombinase MooV and supercoiled or linear DNA substrates. It is possible that there is a cofactor missing from the in vitro assays that is required for MooV-mediated recombination, or that MooV variants with increased insertion activity are needed to detect in vitro activity.

Materials and Methods

Bacterial strains. *E. coli* strains, TB1 [F- *ara*Δ(*lac-proAB*) [Φ80*dlacΔ(lacZ)*M15] *rpsL*(Str^R) *thi hsdR* (New England Biolabs), Origami B (DE3) [F- *ompT hsdSB(rB- mB-) gal dcm lacY1 aphC (*DE3*) gor*522::Tn*10 trxB* (Kan^R,Tet^R)] (Novagen), and DH5α [(Φ80*dlacΔ(lacZ)*M15*) sup*E44 *Δlac*U169 *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A] (obtained from C. Moran) were used for the overexpression of MBP-MooV and MBP-MooV-His. *E. coli* strains DH5α and Top10 [F- *mcr*A Δ*(mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R) *end*A1 *nup*G] (Invitrogen) were used as hosts for IS*492*- and *eps*-containing plasmids.

Media, enzymes, and reagents. All *E. coli* strains were grown on Luria-Bertani (LB) agar or in LB broth at the following drug concentrations for plasmid selection and maintenance: ampicillin at 60 to 100 μg/ml, tetracycline at 15 μg/ml, and chloramphenicol at 50 μg/ml. Factor Xa protease, Klenow DNA polymerase, T4 DNA polynucleotide kinase, restriction enzymes and amylose resin were purchased from New England Biolabs. Acrylamide, PVDF membrane, disposable columns and protein molecular weight markers (Prestained SDS-PAGE Standards version) were purchased from BioRad. Metal chelation resin with immobilized nickel (Ni^{2+}) cations was purchased from Novagen. [α-^{32}P]dATP and [γ-^{32}P]ATP were purchased from Perkin Elmer. All chemicals and rabbit anti-chicken IgG alkaline phosphatase conjugate were purchased from Sigma Chemical Co. HALT protease inhibitor cocktail was purchased from Pierce.

Plasmids and plasmid constructions. All PCR products created for cloning purposes were amplified by using Pfu DNA polymerase (Stratagene). The thermal cycling conditions were as follows: 95°C for 5 min followed by 30 cycles of 95°C for 45 sec, T_m of the oligonucleotide minus 5°C for 45 sec, and 72°C for 1 min followed by one cycle at 72°C for 4 min. The products were digested with indicated restriction enzymes and purified away from reaction components with DNA Clean & Concentrator Kit (Zymo). All ligations were done at 16°C overnight with 400U of T4 DNA ligase. Transformation of ligation mixtures into host strains was performed with competent cells prepared by $CaCl_2$ treatment (19).

pAG949 is a pCR2.1(Invitrogen) derivative containing IS*492*; pAG957 is a derivative of pAG949 with 357-bp of IS*492* replaced with the 884-bp cloramphenicol acetlytransferase gene (cam) (17). pAG952 is a pCR2.1 derivative containing 400-bp of the circle junction (17).

pAG951 was generated by PCR amplification of 134bp of *epsG* target sequence from the pDB200 (3) using primers 1 and 2 and inserted into the TA cloning vector pCR2.1. pDB20 was created by adding a unique *Bgl*II restriction site between the chloramphenicol acetlytransferase gene (*cam*) and its promoter in pACYC184 (New England Biolabs); two bases (AT/TA) were introduced at position 234 using the QuickChange site-directed mutagenesis kit (Stratagene) and complimentary oligos 3 and 4. IS*492*Δ*mooV::lacZα* was amplified from pAG994 (17), using oligos 5 and 6 and ligated into the *Bgl*II site of pDB20 to create pDB22.

pMalE-MooV expresses a translational fusion of *malE* (maltose binding protein (MBP) gene) to the 5' end of *mooV* (17). pMBP-MooV-His was constructed by PCR amplification of *moov*-His_6 from pAG921 using oligonucleotides 5 and 6 followed by cloning into pMalE-MooV using BsrG1 and HindIII restriction sites. Sequencing data revealed that the fusion was in frame.

Western blot analysis of MBP-MooV degradation at increasing temperatures. DH5α cells containing pMalE-MooV were grown at 25°C, 32°C, or 37°C to an absorbance 600 nm of approximately 0.5 in LB broth containing ampicillin before the addition of 1 mM IPTG to induce the expression of MBP-MooV. After a two hour induction, approximately 1.25 x 10^9 cells were lysed by boiling and electrophoresed on a 10% SDS-polyacrylamide gel, transferred to PVDF membrane and subjected to Western blot analysis by anti-MooV anti-sera as described previously (4, 11).

Purification of MBP-MooV. The basic protocol used to purify MBP-MooV or MBP-MooV-His is described here. Modifications to this protocol are indicated in Table 2 and in the Results. DH5α cells containing pMalE-MooV were grown to an absorbance at 600 nm of approximately 0.5 in 1L of LB broth containing ampicillin (60 μg/ml) before the addition of 1 mM IPTG to induce the expression of MBP-MooV. The cells were grown for an additional 2 hours at 37°C.

The cells were harvested and resuspended in buffer containing 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1mM DTT, and 0.5 mM PMSF. Cells were lysed at 4°C by one to three passages through a French pressure cell at 1200 psi and centrifuged at 15,000 x *g* for 15 min at 4°C. The supernatant was diluted 1:5 in the resuspension buffer, loaded onto a 10 ml amylose resin column, and eluted as 2 ml fractions in resuspension buffer containing 10 mM maltose. The protein concentrations were determined by the Bradford assay (BioRad). Proteins were electrophoresed in a 10% SDS-polyacrylamide gel to determine which fractions contained MBP-MooV. Fractions were generally stored on ice at 4°C and used within 7 days of their elution from amylose resin.

Purification of MBP-MooV-His. After purification from an amylose column as described in previous section, MBP-MooV-His fractions were pooled and loaded onto a column containing metal chelation resin with immobilized nickel (Ni^{2+}) cations. Fractions were eluted with resuspension buffer plus 500 mM immidazole and electrophoresed on a 10% SDS-polyacrylamide gel (SDS-PAGE) to determine which fractions contained MBP-MooV-His.

Proteolytic cleavage of MBP-MooV and MooV isolation. MBP-MooV was cleaved into MBP (43kDa) and MooV (35kDa) by adding 1 μg of Factor Xa per 50 μg of MBP-MooV followed by incubation on ice at 4°C for approximately 14 hours. Separation of MBP and MooV was attempted using anion and cation exchange chromatography with the Q1 and S6 columns (BioRad), respectively. Factor Xa-cleaved MBP-MooV was loaded onto an ion exchange column either following elution from an amylose column or following elution from an amylose column and dialysis of sample from 200 mM NaCl to 20 mM NaCl. Stepwise elution was performed using resuspension buffer and salt concentrations ranging from 20 mM to 1 M. Fractions were collected in 2 ml aliquots and monitored by SDS-PAGE.

MBP-MooV nicking assays. To detect MBP-MooV nicking of either a donor plasmid (pAG949 or pAG952) or target plasmid (pAG951), 0.22 nmol of MBP-MooV were incubated with 1.5 μg of plasmids in a 200 μl reaction buffer containing 20 mM Tris-HCl, pH 7.5, 10 mM NaCl, 5 mM $MgCl_2$, 50 μg/ml tRNA, and 50 μg/ml BSA. Changes to this reaction include the addition of 5 mM $CaCl_2$, 10 mM $MgCl_2$, 0.15 mM $MnCl_2$, Factor Xa-cleaved MBP-MooV, and the mixing of donor and target plasmids in the reaction. Reactions were incubated for five hours at 37°C. Reactions were stopped by the addition of 0.5% SDS and 1 μg proteinase K, followed by incubation at 55 °C for 30 minutes. Alternatively, reactions were incubated at room temperature, for varying times, and stopped by the addition of only 0.5% SDS or only 1 μg of proteinase K. The DNA was concentrated by ethanol precipitation. Products were examined by electrophoresis on a 1% agarose gel in 1xTBE, followed by staining with a 0.5 μg/ml ethidium bromide solution. The relative amount of each DNA species was determined using Image Quant (BioRad) analysis of a digital gel image.

In a similar assay, 5×10^{-10} mol of MBP-MooV-His were incubated with 5×10^{-13} mol of pAG957 or pAG952 in a buffer containing 100 mM KCl, 20 mM Hepes, pH 7.5, 10 mM $MgCl_2$, and 0.1 mg/ml BSA. After incubation with MBP-MooV at room temperature for 30 minutes, the reaction was stopped by the addition of 0.5% SDS and 150 μg/ml of proteinase K, followed by phenol/chloroform extraction and ethanol precipitation. pAG957 or pAG952 were then digested with BsrG1 and HindIII or EcoR1, respectively, and end-labeled with Klenow and [α-^{32}P]dATP following protocols of Sambrook, *et al*. (19). After being labeled, the DNA fragments were examined by polyacrylamide gel electrophoresis (PAGE) on a 5% DNA-denaturing polyacrylamide gel, and exposed to a phosphoimaging screen or X-ray film (Kodak).

MBP-MooV cleavage and strand transfer. After digesting pAG952 with EcoR1, the resulting circle junction fragment was isolated and end-labeled with Klenow and [α-^{32}P]dATP. One nM of the labeled circle junction fragment was incubated with 1 nM or 2 nM target plasmid, pAG951, 94 nM MBP-MooV in a reaction buffer containing: 50 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 50 μg/ml BSA, and 50 μg/ml tRNA. After incubation at 37°C for 30 minutes, MBP-MooV was inactivated by the addition of 0.5% SDS and 1 μg proteinase K followed by incubation at 65°C for 20 minutes. Products were examined by electrophoresis on a 1% TBE agarose gel followed by exposure to X-ray film. Alternatively, 20% glycerol, 15% DMSO, 1 μg/ml sonicated salmon sperm DNA, or 0.5 mM EDTA was added to the reaction buffer. A second buffer containing 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 50 μg/ml BSA, 50 μg/ml tRNA, and 0.5 mM EDTA was also used in this assay.

MBP-MooV nicking and strand transfer assay. To detect MBP-MooV-mediated strand transfer, pre-processed oligonucleotides with 3'OH-ends were incubated with MBP-MooV and a target plasmid, pAG951. The oligonucleotides were of the left end, oligos 7, 8, 9, and 10, or of the right end, oligos 11, 12, 13, and 14 (see Table 4.1 for sequence). The oligonucleotides either contained 35 bp of end sequence including the 5 bp target duplication (8, 9, 11, and 14) or 30 bp of end sequence not including 5 bp target duplication (7, 10, 12, and 13). Oligonucleotides 7, 9, 11, and 13 were labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase following the protocols of Sambrook, *et al*. (19). The labeled oligonucleotides were annealed to their complimentary strand such that the 5 bp target duplication was not annealed to complimentary DNA. The resulting 5 bp overhang either contained the 3'OH, oligonucleotides 9/10 and 11/12 annealed, or the recessed end contained the 3' OH, oligonucleotides 7/8 and 13/14. Labeled ds oligonucleotides (0.12 pmol) were incubated with 0.03 pmol of target plasmid, pAG951, and 150

ng of MBP-MooV in a buffer containing 50 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 50 μg/ml BSA, and 50 μg/ml tRNA. After incubation at 37°C for 30 minutes, MBP-MooV was inactivated with 0.5% SDS and 1 μg proteinase K followed by incubation at 55°C for 20 minutes. Products were electrophoresed on a 1% TBE agarose gel and examined by exposure to X-ray film.

In a similar assay, the oligonucleotides were annealed such that no over-hangs were present, oligonucleotides 8 and 9 (left end) were annealed and oligonucleotides 11 and 14 (right end) were annealed. Twenty nM labeled, ds oligo was incubated with 1 nM target plasmid, pDB22, and 2 μM MBP-MooV-His in a reaction buffer containing 100 mM KCl, 20 mM Hepes pH 7.5, and 10 mM $MgCl_2$. Reactions were stopped by the addition of 1 μg of proteinase K followed by incubation at 55°C for 20 minutes. Half of the reaction was electrophoresed on a 1% TBE agarose gel to detect strand transfer, and half of the reaction was electrophoresed on a 20% DNA-denaturing polyacrylamide gel to detect nicking of DNA. Both gels were exposed to a phosphoimaging screen to examine products.

Primer extension. Primer extension analysis was performed on DNA products from the MBP-MooV nicking assays to determine where DNA cleavage was occurring. Two oligonucleotides (M13 forward and M13 reverse from Invitrogen) were labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase following the protocols of Sambrook, *et al*. (19). The Sanger sequencing reactions were performed according to manufacturer's instructions (Promega *fmol* DNA Cycle Sequencing system). Briefly, 40 fmol of template (pAG951) and 1.5 pmol primer was incubated in 1X fmol buffer and extension was performed with 5 U of Sequencing Grade *Taq* DNA polymerase. Cycling conditions: 1 cycle at 95°C for 2 minutes followed by 30 cycles of 95°C, 30 seconds, 42°C, 30 seconds, and 70 °C, 1 minute. The primer extension reaction was

performed similar to Sanger sequencing reactions with the exception of not including dideoxy nucleotide triphosphates in the reaction mixture. Products were electrophoresed on a 5 % DNA-denaturing polyacrylamide, and products were examined by exposing the gel to X-ray film.

Results

Purification of MBP-MooV

To facilitate purification of MooV for *in vitro* activity assays, a translational fusion of *malE* and *mooV* was created and the expressed maltose binding protein (MBP)-MooV was affinity purified on amylose resin. MBP-MooV was previously shown to be active in vivo (17). The fusion protein was engineered to have a short amino acid-linker between the proteins, which contains the recognition sequence for the protease Factor Xa. Figure 4.1 shows cleavage of MBP-MooV with Factor Xa. To separate MooV from free MBP, ion exchange chromatography was utilized based on the difference in the pIs of MBP and MooV, which are 5.28 and 9.94, respectively. In chromatography with the Q1 anion exchange column, MooV was expected to be in the flow-through eluate, while MBP should bind to the column, and with the S6 cation exchange column, MooV was predicted to bind the column while MBP eluted in the flow-through. Surprisingly, MBP and MooV co-eluted from the Q1 column and MooV did not elute from the S6 column in the salt buffer gradient. This inability to separate MooV from MBP may be due to denaturation and microprecipitation of MooV after being cleaved away from MBP.

To increase the solubility and activity of MBP-MooV, several purification protocols were investigated (Table 4.2). Western blot analysis of MBP-MooV expressed at 25°C, 32°C, and 37°C showed that decreasing the temperature during induction reduced degradation of the fusion protein (Figure 4.2). This may reflect that the optimum growth temperature for *P. atlantica,*

25°C, is also the optimum temperature for MBP-MooV. Therefore, most of the in vitro assays described here were performed with protein purified from cells incubated at 25°C.

As seen in Figure 4.1A (lane 4) and Figure 4.1B (lane 2), several proteins co-eluted with MBP-MooV from the amylose column. In order to add a second affinity chromatography step to the purification of MooV, a translational fusion of MooV with MBP at the amino-terminus and His(6) at the carboxyl-terminus was utilized. The MBP-MooV-His fusion protein is active in vivo as indicated by its ability to excise and circularize IS*492* (data not shown). As shown in Figure 4.1B, using both an amylose and a Ni^{2+} resin column to purify MBP-MooV-His does appear to decrease the amount of some co-eluting proteins in the sample, but does not remove the major contaminating polypeptide, which may be a degradation product.

MBP-MooV DNA cleavage assays

To assess MooV-mediated DNA nicking or double-strand cleavage, supercoiled plasmid DNA substrates containing the *eps* target DNA (pAG951), IS*492* inserted into the target DNA (pAG949), or the circle junction of the excised IS*492* element (pAG952) were incubated with MBP-MooV or Factor X-cleaved MBP-MooV. Donor-target plasmid combinations (pAG949-pAG951 or pAG952-pAG951), as well as each plasmid separately, were used in the assays in the event that DNA cleavage is linked to DNA strand transfer. Conversion of the supercoiled plasmid DNA to linear or open-circular DNA was determined by agarose gel electrophoresis and quantitation of each DNA species following incubation in the presence or absence of MBP-MooV (see Methods and Materials). In the assay shown in Figure 4.3B, comparison of treated and untreated plasmid DNA showed no change in the percent supercoiled versus open circular or linear substrate. To improve the sensitivity of the assay, >90% supercoiled substrate DNA was isolated by $CsCl_2$ gradients and used in the cleavage assays, but again no significant difference in

the percent open circular or nicked DNA was seen following incubation with or without MBP-MooV (data not shown).

Several changes to the reaction conditions were tried to optimize MooV-mediated cleavage of target DNA (see Materials and Methods). It was found that terminating the reaction with 0.5% SDS instead of proteinase K resulted in a measurable increase in cleaved DNA products in the presence of MBP-MooV (Figure 4.4). If MooV acts like a DDE-motif transposases, then the addition of 0.5% SDS should stop the reaction and still allow the MBP-MooV treated DNA to enter an agarose gel because DDE-motif transposases do not covalently link to their target DNA. The percentage of open-circular DNA compared to super-coiled DNA increased as the amount of MBP-MooV added to the reaction increased (Figure 4.5). Some enzymes with endonuclease activities, such as type I DNA topoisomerase and helicase I, nick DNA substrates when subjected to rigorous protein-denaturing treatment (9, 18). Helicase I nicks its supercoiled DNA substrate and becomes covalently attached to the 5' end of the nicked DNA substrate after being treated with SDS/EDTA and heated to 45°C (18).

A primer extension assay was performed on the DNA reaction products for pAG951 using primers that anneal outside the *epsG* insert and extend towards the insert. The primer extension products were compared with Sanger DNA sequencing of the plasmid using the same primers in order to determine the exact site of cleavage. The primers should stop once they reach a site of cleavage; however, no distinct products were visible from the primer extension reactions.

To determine if the increase of open-circular DNA was a result of an interaction between SDS and MBP-MooV, an assay was performed with pAG952 in which the reaction was stopped with 0.5% SDS, 1 μg proteinase K, or a combination of 0.5% SDS and 1μg proteinase K (Figure

4.6). As seen in Figure 4.6, the addition of 0.5% SDS to MBP-MooV treated DNA did not consistently result in an increase in the percentage of open-circular DNA compared to super-coiled DNA. Also, when proteinase K was added in the presence of SDS, no increase in open-circular DNA was detected. This variability in the plasmid cleavage assays led us to design a more sensitive assay to detect MooV-mediated nicking or cleavage.

To detect low levels of DNA cleavage, the plasmid DNA substrates were incubated with MBP-MooV-His followed by digest with restriction enzymes with recognition sites on either side of the target sequence and then labeled with Klenow and [α-^{32}P] dATP. Figure 4.7 represents one example of this assay using pAG952, containing the circle junction fragment, as the target DNA. If MBP-MooV-His nicked the DNA, then the nicked product would have a faster mobility on the DNA denaturing gel, and a similar result would be visualized if MBP-MooV-His cleaved both DNA strands. As seen in Figure 4.7, increasing amounts of MBP-MooV-His did not result in an increase of nicked or cleaved product. In addition, three reaction conditions were altered to determine if they would have an effect on cleavage activity. First, glycerol was added at 10% or 15% because glycerol stimulates in vitro Tn*10* transposition and Hin-mediated DNA inversion reactions (5, 13). Second, $CaCl_2$ was added in addition to $MgCl_2$ since $CaCl_2$ was shown by Tobiason et al. (23) to increase the binding affinity of the related recombinase Piv to its binding site. Third, integration host factor (IHF) was added to the reaction since it has been shown to be involved in both DNA transposition of Tn*10* (20) and site-specific integration of bacteriophage lambda (7). In addition, DN*ase* I protection assays indicate that IHF does interact with a predicted strong binding site on IS*492,* but it is not required for *in vivo* excision of IS*492* (2). No nicked or cleaved products were seen with the addition of glycerol, Ca_2Cl, or IHF (Figure 4.7). This assay was repeated using pAG957, containing

IS*492*Δ*mooV*::*cat* as the target DNA, and no nicked or cleaved products were detected with the different target DNA (data not shown).

MBP-MooV cleavage and strand transfer

To detect both MBP-MooV cleavage of a double-stranded DNA fragment and its subsequent transfer to a target plasmid, which may be linked reactions, an assay was designed in which a radiolabeled restriction fragment containing the circle junction was incubated in the presence of both MBP-MooV and a target plasmid, pAG951 (Figure 4.8A). If MBP-MooV cleaved the restriction fragment, the product would have a faster mobility on a 1% agarose gel than the original substrate, and, if MBP-MooV transferred the cleaved fragment to pAG951, the resulting strand-transfer, open-circular product, would have a slower mobility than the supercoiled target substrate. The reaction conditions used in this assay and the corresponding results are listed in Table 4.3. As in the nicking assays described above, the addition of various agents, including glycerol, DMSO, sonicated salmon sperm DNA (to stabilize the recombinase), and 0.5 mM EDTA (to chelate Mg_2Cl and reduce nuclease activity) did not result in any DNA cleavage or transfer activity for MBP-MooV.

Since the DNA cleavage and transfer assays with purified MBP-MooV did not reveal any catalytic activity, a cofactor from *P. atlantica* might be required for MooV activity. Therefore, DNA cleavage assays, like those shown in Figure 4.8A, were performed using reaction buffer A (Table 4.3) in the presence of added *P. atlantica* cell lysate. The results shown in Figure 4.8B suggest that the addition of cell lysate led to degradation of the DNA substrate, probably due to various nucleases in the crude protein preparation resulting in a lack of DNA substrate in lanes containing reactions with cell lysate added. Although the incubation time could be decreased to

reduce the loss of DNA substrate from nuclease activity, it was still likely that sensitivity was too compromised by loss of substrate, so the cell lysate was not further utilized.

Because strand transfer was not detected in the MBP-MooV-mediated cleavage and strand transfer assay described above, a strand transfer assay was initiated that provided double stranded oligonucleotides with the predicted ends of a DNA substrate that had been cleaved by MBP-MooV. This is a similar approach to that used in the study of strand transfer reactions by the recombinases RAG1 and RAG2, which mediate immunoglobulin gene rearrangements (16). This type of assay is useful as it allows the study one specific step in the reaction, strand transfer, and the isolation of reaction intermediates. As described in the introduction, molecular modeling of MooV suggests that it has a DEDD catalytic tetrad that functions similar to the DDE catalytic triad of classical transposases/retroviral integrases whose catalytic mechanism results in a free 3'OH upon cleavage of the donor DNA. Therefore, we provided MBP-MooV with double-stranded oligonucleotides encoding the ends of the element with a 3'OH. Because the site of MBP-MooV cleavage is unknown, 30 bp double-stranded oligonucleotides corresponding to the left and right end of IS*492* were designed to have a 3'OH at the last nucleotide of the element sequence and 35 bp double-stranded oligonucleotides were created with the 3'OH at the end of the 5 bp target sequence that flanks the ends of the element. Since, classical transposases/retroviral integrases usually cleave DNA in a staggered manner, the labeled double-stranded oligonucleotide contained a 3'OH on either the 5 bp target sequence overhang, or the recessed end of IS*492* (Figure 4.9A). A variation of this experiment involved mixing radiolabeled left end IS*492* sequence with unlabeled right end IS*492* sequence in the event that both the right and left end of IS*492* were required to activate the recombinase. After incubation with MBP-MooV, the products were subjected to electrophoresis on a 1% TBE agarose gel. If

MBP-MooV transferred the oligonucleotide to pAG951, a slowly migrating band would be visible on the agarose gel. As shown in Figure 4.9B, the presence of MBP-MooV resulted in a smear of the DNA into the agarose gel, but no distinct bands were present. This trend continued even when the samples were subjected to electrophoresis on a lower percentage, 0.8%, agarose gel. Due to the inability to visualize a single distinct band on the agarose gel, it was difficult to determine if MBP-MooV had indeed transferred the ds oligonucleotide to pAG951.

Another variation of this assay was to anneal the 35 nt oligonucleotides together to create a ds oligonucleotide donor containing 30 bp of the end of IS*492* and the 5 bp target duplication. Both strands of the 35 bp oligonucleotide were labeled with [γ-32P]dATP and T4 polynucleotide kinase (Figure 4.10A). The labeled, ds oligonucleotide was incubated in the presence of increasing amounts of MBP-MooV-His and a target plasmid, pDB22, which contains IS*492*Δ*mooV*::*lacZα* inserted in 60 bp of *epsG* sequence (Figure 4.10A). pDB22 was used as a target plasmid because head-to-tail dimers of IS*492* have been detected in vivo suggesting that MooV targets both *epsG* and IS*492* (A. Cottrell, personal communication). The products of the reaction were subjected to electrophoresis on an agarose gel to detect strand transfer and on a DNA denaturing gel to detect nicking of the ds oligonucleotide donor. No strand transfer products were visualized on the agarose gel (data not shown), and no nicked or cleaved oligonucleotide products were detected on the DNA denaturing gel (Figure 4.10B).

Discussion

MooV is an atypical transposase that has characteristics of both site-specific recombinases and transposases. Elucidation of its mechanism of catalysis will contribute to the understanding of specialized DNA recombination. Described in this chapter are in vitro assays

designed to isolate and characterize reaction intermediates in MooV-mediated DNA nicking or double-strand cleavage and strand transfer. Although we had some promising data in the MBP-MooV nicking assays treated with SDS and the strand transfer assay, we were unable to define the optimum reaction conditions to detect MooV activity in vitro or to isolate and characterize a reaction intermediate in MooV-mediated recombination.

It is not unusual to run into obstacles in the definition of conditions that support in vitro activity of a specialized DNA recombinase because these enzymes have adopted strategies to maintain a low level of activity so that DNA cleavage activity and movement of their elements is not deleterious to the host (15). In order to establish an in vitro assay for Tn*5* transposition the Reznikoff laboratory designed a screen to isolate hyperactive variants that were subsequently shown to be active in vitro (8, 24). We are taking a similar approach by screening for hyperactive MooV variants following PCR mutagenesis of *mooV*. In brief, the screen involves expressing mutagenized *mooV* from a plasmid in an *Escherichia coli* cell containing an IS*492* derivative with nearly all of *mooV* replaced by a promoter-less *lacZα* and screening for colonies with blue papillae to detect the insertion of IS*492*Δ*mooV*::*lacZα*. The *mooV* expression plasmid from colonies with higher numbers of papillae than seen with wild-type MooV are further characterized for IS*492* insertion frequency using a mating out assay, which is similar to the methods used by (21) and (24) to determine Tn*5* and IS*903* transposition frequencies, respectively. MooV variants that have a higher insertion frequency than wild-type MooV will be used in the in vitro assays described herein in order to further define the catalytic mechanism of this atypical recombinase.

References

1. **Azaro MA, L. A.** 2002. Lambda Integrase and the Lambda Int family, p. 230-271. *In* N. Craig, Craigie R, Gellert M, Lambowitz, AM (ed.), Mobile DNA II. ASM Press, Washington D. C.
2. **Balding, D. P.** 2000. Transposition of IS*492*: *In vivo* and *in vitro* characterization of a member of an atypical group of insertion sequences. Emory University, Atlanta.
3. **Bartlett, D. H., M. E. Wright, and M. Silverman.** 1988. Variable expression of extracellular polysaccharide in the marine bacterium *Pseudomonas atlantica* is controlled by genome rearrangement. Proc Natl Acad Sci U S A **85:**3923-3927.
4. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J Bacteriol **187:**3431-7.
5. **Chalmers, R. M., and N. Kleckner.** 1994. Tn*10*/IS*10* transposase purification, activation, and in vitro reaction. J Biol Chem **269:**8029-35.
6. **Craig, N., Craigie R, Gellert M, Lambowitz, AM.** 2002. Mobile DNA II, second ed. ASM Press, Washington D. C.
7. **Goodman, S. D., S. C. Nicholson, and H. A. Nash.** 1992. Deformation of DNA during site-specific recombination of bacteriophage lambda: replacement of IHF protein by HU protein or sequence-directed bends. Proc Natl Acad Sci U S A **89:**11910-4.
8. **Goryshin, I. Y., and W. S. Reznikoff.** 1998. Tn*5* in vitro transposition. J Biol Chem **273:**7367-74.
9. **Hamelin, C., and M. Yaniv.** 1979. Nicking-closing enzyme is associated with SV40 DNA in vivo as a sodium dodecyl sulfate-resistant complex. Nucleic Acids Res **7:**679-87.

10. **Haren, L., B. Ton-Hoang, and M. Chandler.** 1999. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol **53:**245-81.

11. **Higgins, B. P., C. D. Carpenter, and A. C. Karls.** 2007. Chromosomal context directs high-frequency precise excision of IS*492* in *Pseudoalteromonas atlantica*. Proc Natl Acad Sci U S A **104:**1901-6.

12. **Johnson, R. C.** 2002. Bacterial Site-Specific DNA Inversion, p. 230-271. *In* N. Craig, Craigie R, Gellert M, Lambowitz, AM (ed.), Mobile DNA II. ASM Press, Washington D. C.

13. **Johnson, R. C., and M. F. Bruist.** 1989. Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombinational enhancer in the strand exchange reaction. EMBO J **8:**1581-90.

14. **Lenich, A. G., and A. C. Glasgow.** 1994. Amino acid sequence homology between Piv, an essential protein in site-specific DNA inversion in *Moraxella lacunata*, and transposases of an unusual family of insertion elements. J Bacteriol **176:**4160-4.

15. **Nagy, Z., and M. Chandler.** 2004. Regulation of transposition in bacteria. Res Microbiol **155:**387-98.

16. **Neiditch, M. B., G. S. Lee, M. A. Landree, and D. B. Roth.** 2001. RAG transposase can capture and commit to target DNA before or after donor cleavage. Mol Cell Biol **21:**4302-10.

17. **Perkins-Balding, D., G. Duval-Valentin, and A. C. Glasgow.** 1999. Excision of IS*492* requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. J Bacteriol **181:**4937-48.

18. **Reygers, U., R. Wessel, H. Muller, and H. Hoffmann-Berling.** 1991. Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. EMBO J **10:**2689-94.

19. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor.

20. **Sewitz, S., P. Crellin, and R. Chalmers.** 2003. The positive and negative regulation of Tn*10* transposition by IHF is mediated by structurally asymmetric transposon arms. Nucleic Acids Res **31:**5868-76.

21. **Tavakoli, N. P., and K. M. Derbyshire.** 1999. IS*903* transposase mutants that suppress defective inverted repeats. Mol Microbiol **31:**1183-95.

22. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. Mol Microbiol **39:**641-51.

23. **Tobiason, D. M., A. G. Lenich, and A. C. Glasgow.** 1999. Multiple DNA binding activities of the novel site-specific recombinase, Piv, from *Moraxella lacunata*. J Biol Chem **274:**9698-706.

24. **Wiegand, T. W., and W. S. Reznikoff.** 1992. Characterization of two hypertransposing Tn*5* mutants. J Bacteriol **174:**1229-39.

Table 4.1: Oligonucleotide Sequences

Oligonucleotide	Sequence
1	5'CATTTTAGCTTCCTTAGATCTCTGAAAATC
2	5'GATTTTCAGGAGATCTAAGGAAGCTAAAATG
3	5'GACAGATCTACCTAACAGAGACTATGGC
4	5'GACAGATCTGAGAGCTTGATAGACCTAACAAG
5	5'GGCTCTCGAGCGGTACTGTCTTATCATCCTAATCG
6	5'CTCGCTCGAGCAGGAGGCTCTCTATTGTACAGC
7	5'TCAGGATGATTTACCTGAGTTATTATCCAT
8	5'CTTGTATGGATAATAACTCAGGTAAATCATCCTGA
9	5'TCAGGATGATTTACCTGAGTTATTATCCATACAGG
10	5'ATGGATAATAACTCAGGTAAATCATCCTGA
11	5'ACTACAAATTAGCTATTGACGCCATAGTCTCTTGT
12	5'AGACTATGGCGTCAATAGCTAATTTGTAGT
13	5'ACTACAAATTAGCTATTGACGCCATAGTCT
14	5'ACAAGAGACTATGGCGTCAATAGCTAATTTGTAGT

Table 4.2: Purification Protocols. Purification types A-I were performed with MBP-MooV. Purification types J-L were performed with MBP-MooV-His.

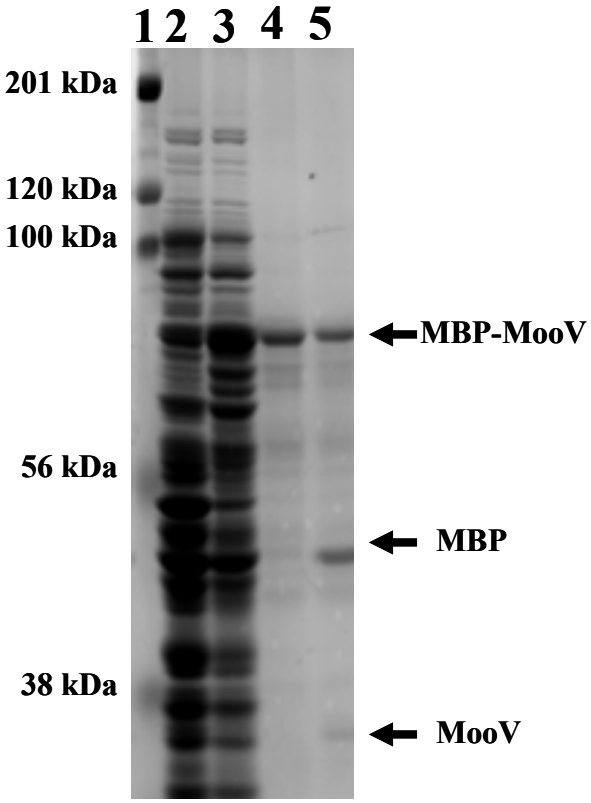
Purification/ Cell Strain	*Buffer*	*Salt*	*EDTA*	*DTT*	*Protease Inhibitor*	*Growth Temp.*	*IPTG*	*Induction Time*
A/ DH5α	20mM Tris-HCl, pH 7.4	200mM NaCl	1mM	1mM	0.5mM PMSF	37°C	1mM	2 hours
B/ OrigamiB	20mM Tris-HCl, pH 7.4	200mM NaCl	1mM	1mM	0.5mM PMSF	37°C	1mM	2 hours
C/ DH5α	20mM Tris-HCl, pH 7.4	200mM KSCN	1mM	1mM	1X HALT	37°C	1mM	2 hours
D/ DH5α	20mM Hepes, pH 7.5	200mM KSCN	1mM	1mM	1X HALT	37°C	1mM	2 hours
E/ DH5α	20mM Hepes, pH 7.5	200mM NaCl	1mM	1mM	1X HALT	37°C	1mM	2 hours
F/ DH5α	20mM Tris-HCl, pH 7.4	200mM NaCl	1mM	1mM	1X HALT	37°C	1mM	2 hours
G/ DH5α	20mM Tris-HCl, pH 7.4	200mM KSCN	1mM	1mM	0.5mM PMSF	37°C	1mM	2 hours
H/ DH5α	20mM Tris-HCl, pH 7.4	200mM KSCN	1mM	1mM	0.5mM PMSF	25°C	1mM	2 hours
I/ DH5α	20mM Tris-HCl, pH 7.4	200mM KSCN	1mM	1mM	0.5mM PMSF	32°C	1mM	2 hours
J/ TB1	20mM Hepes, pH 7.5	500mM KCl	n/a	n/a	0.5mM PMSF	25°C	1mM	2 hours
K/ TB1	20mM Hepes, pH 7.5	500mM KCl	n/a	n/a	0.5mM PMSF	25°C	0.3mM	16 hours
L/ TB1	20mM Hepes, pH 7.5	500mM KCl	n/a	n/a	0.5mM PMSF	18°C	0.3mM	16 hours

Table 4.3: Description of different reaction conditions used in MBP-MooV cleavage and strand transfer assay, and the results of the different reactions.

Reaction Conditions	Cleavage	Transfer
Buffer A (Tris-HCl, pH 7.9 100 mM NaCl, 10mM MgCl2, 1mM DTT, 50 µg/ml BSA, 50 µg/ml BSA, 50 µg/ml tRNA)	–	–
Buffer B (Buffer A + 20% glycerol)	–	–
Buffer C (Buffer A + 15% DMSO)	–	–
Buffer D (Buffer A + 1 µg sonicated salmon sperm DNA)	–	–
Buffer E (Buffer A excluding DTT + 0.5 mM BME)	–	–
Buffer F (Buffer A + 0.5 mM EDTA)	–	–

Figure 4.1: Factor Xa cleavage of MBP-MooV. (A) Whole-cell lysates of DH5α containing pMalE-MooV and affinity-purified MBP-MooV visualized on a 10% SDS-PAGE stained with Coomassie Brilliant Blue. Lane 1, marker; lane 2,whole-cell lysates of DH5α containing pMalE-MooV, uninduced; lane 3, whole-cell lysates of DH5α containing pMalE-MooV, induced; lane 4, full-length, partially purified MBP-MooV; lane 5, Factor Xa-cleaved MBP-MooV. (B) Purification of MBP-MooV-His visualized on a 10% PAGE stained with Coomassie Brilliant Blue. Lane 1, marker; lane 2, MBP-MooV-His purified from amylose column; lanes 3-6, MBP-MooV-His fractions purified from amylose column then from Ni^{2+} resin column.

A.


1 2 3 4 5
201 kDa
120 kDa
100 kDa
56 kDa
38 kDa
MBP-MooV
MBP
MooV

B.

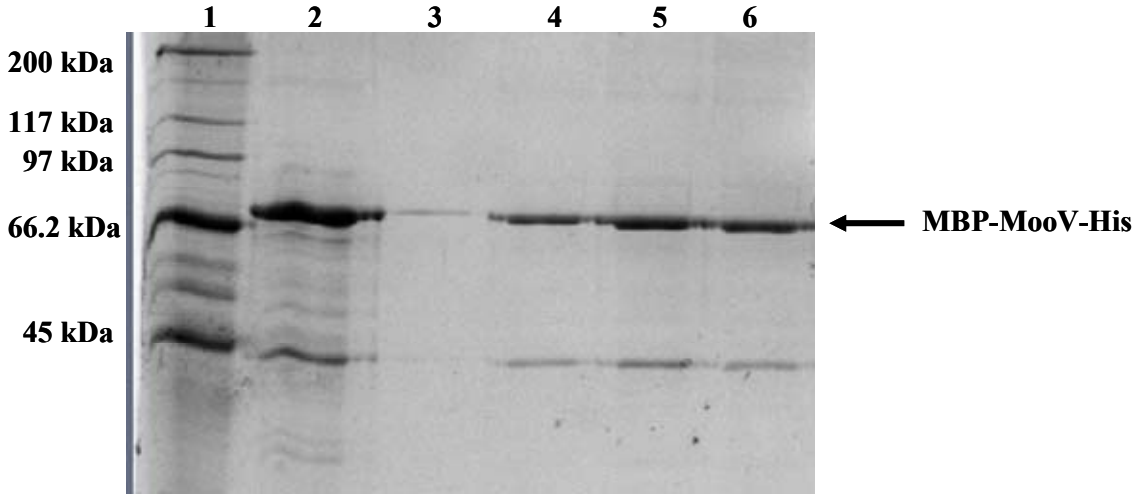

1 2 3 4 5 6
200 kDa
117 kDa
97 kDa
66.2 kDa
45 kDa
MBP-MooV-His

Figure 4.2: Affect of growth temperature on degradation of MBP-MooV. DH5α cells containing pMBP-MooV were grown to mid-log phase at 25°C, 32°C, or 37°C and induced with 1mM IPTG for 2 hours. Cell lysates from 1.25 x 10^9 cells were electrophoresed on a 10% SDS-PAGE followed by Western blot analysis using anti-MooV antisera. The growth temperature of the cells is indicated above the Western blot and MBP-MooV and degradation products are indicated by arrows.

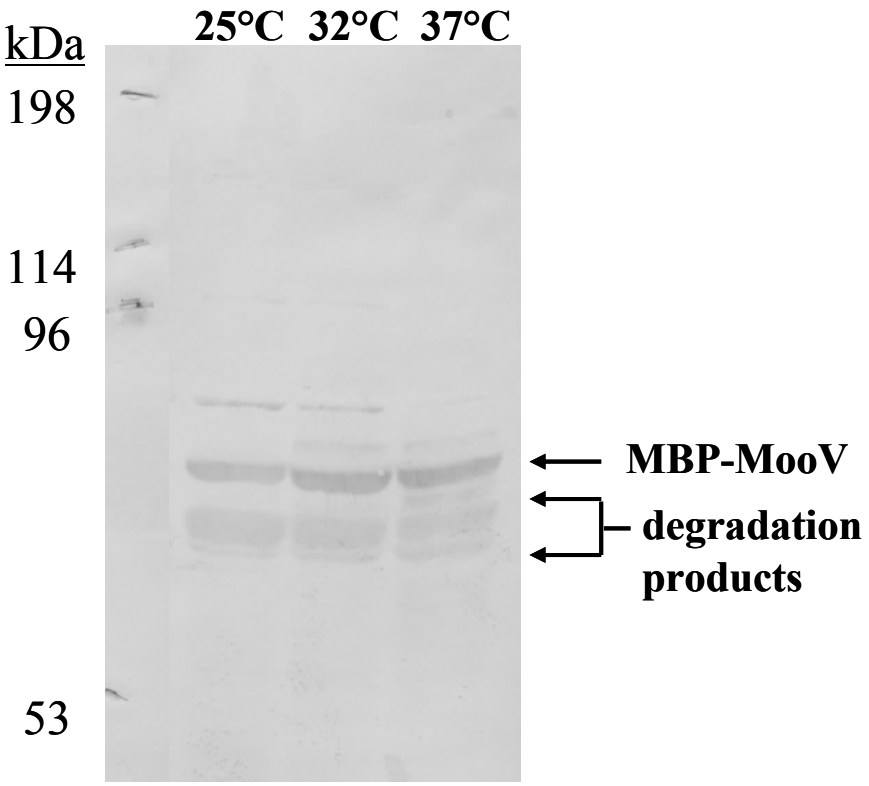
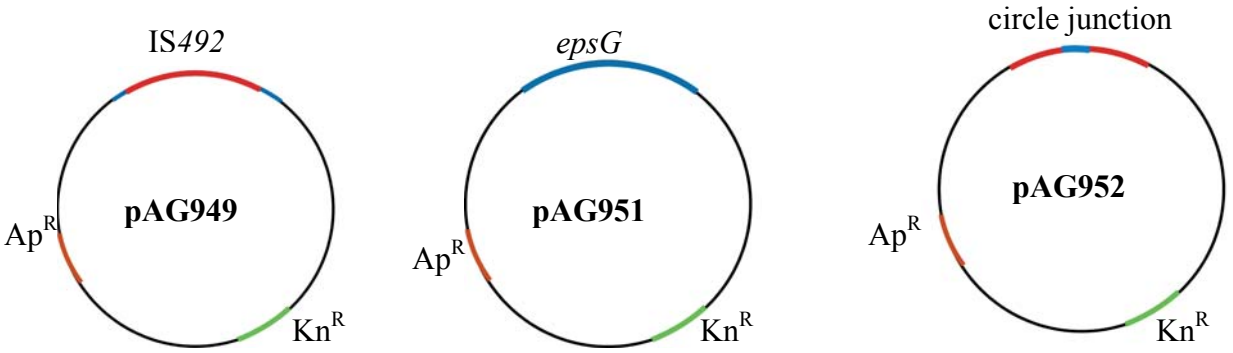

25°C 32°C 37°C
kDa
198
114
96
53
MBP-MooV
degradation products

Figure 4.3: *In vitro* DNA nicking activity of MBP-MooV. (A) The plasmids used in the DNA nicking assay are diagrammed here: pAG951 which has 124bp of *epsG* sequence including the IS*492* target site, pAG949 which has IS*492* inserted into the *epsG* sequence such that there is 58 bp on the left and 76 bp on the right, pAG952 which has sequence from the circle junction including 200bp left-end and 200bp right-end IS*492* and the 5bp junction sequence, and pCR2.1 which is the vector backbone of the three substrate plasmids. (B) Agarose gel electrophoresis of products of *in vitro* MBP-MooV nicking assays performed as described in Materials and Methods. The presence of MBP-MooV is indicated above each lane as + or -. SC, supercoiled plasmid; OC, open circular plasmid. Products were stained with ethidium bromide and the inverse image is shown.

A.


IS492
pAG949
Ap^R
Kn^R
epsG
pAG951
Ap^R
Kn^R
circle junction
pAG952
Ap^R
Kn^R

B.

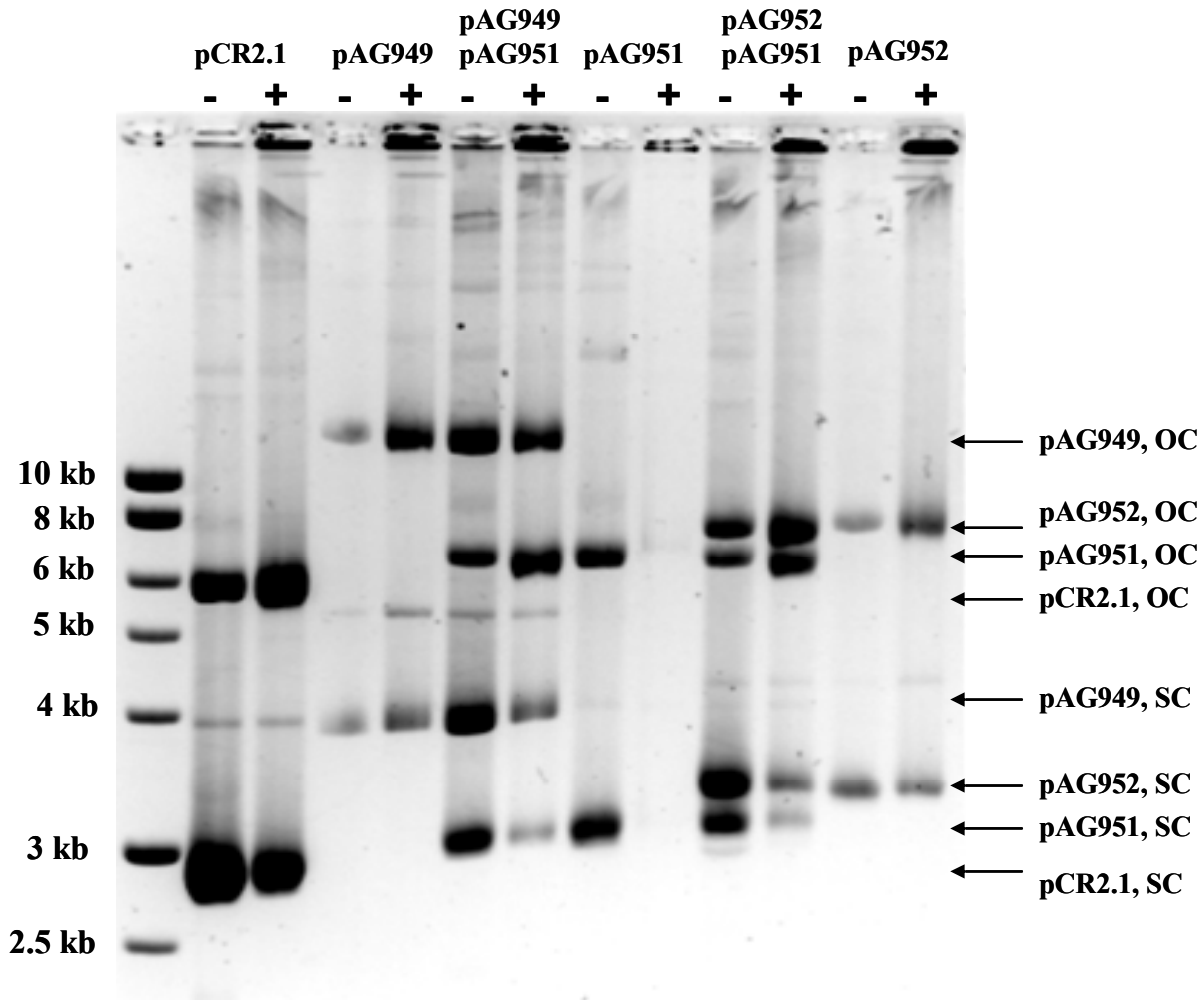

pCR2.1
pAG949
pAG949 pAG951
pAG951
pAG952 pAG951
pAG952
- + - + - + - + - + - +
10 kb
8 kb
6 kb
5 kb
4 kb
3 kb
2.5 kb
pAG949, OC
pAG952, OC
pAG951, OC
pCR2.1, OC
pAG949, SC
pAG952, SC
pAG951, SC
pCR2.1, SC

Figure 4.4: Effect of 0.5% SDS on *in vitro* MBP-MooV nicking assay. 1% TBE agarose gel of *in vitro* nicking assay using only 0.5% SDS to stop reaction. The presence or absence of 0.5% SDS and/or 22 nmol of MBP-MooV is indicated by + or -, U indicates untreated DNA. The percentage of open-circular DNA (OC) compared to super-coiled DNA (SC) within each lane is indicated at the bottom of the gel.

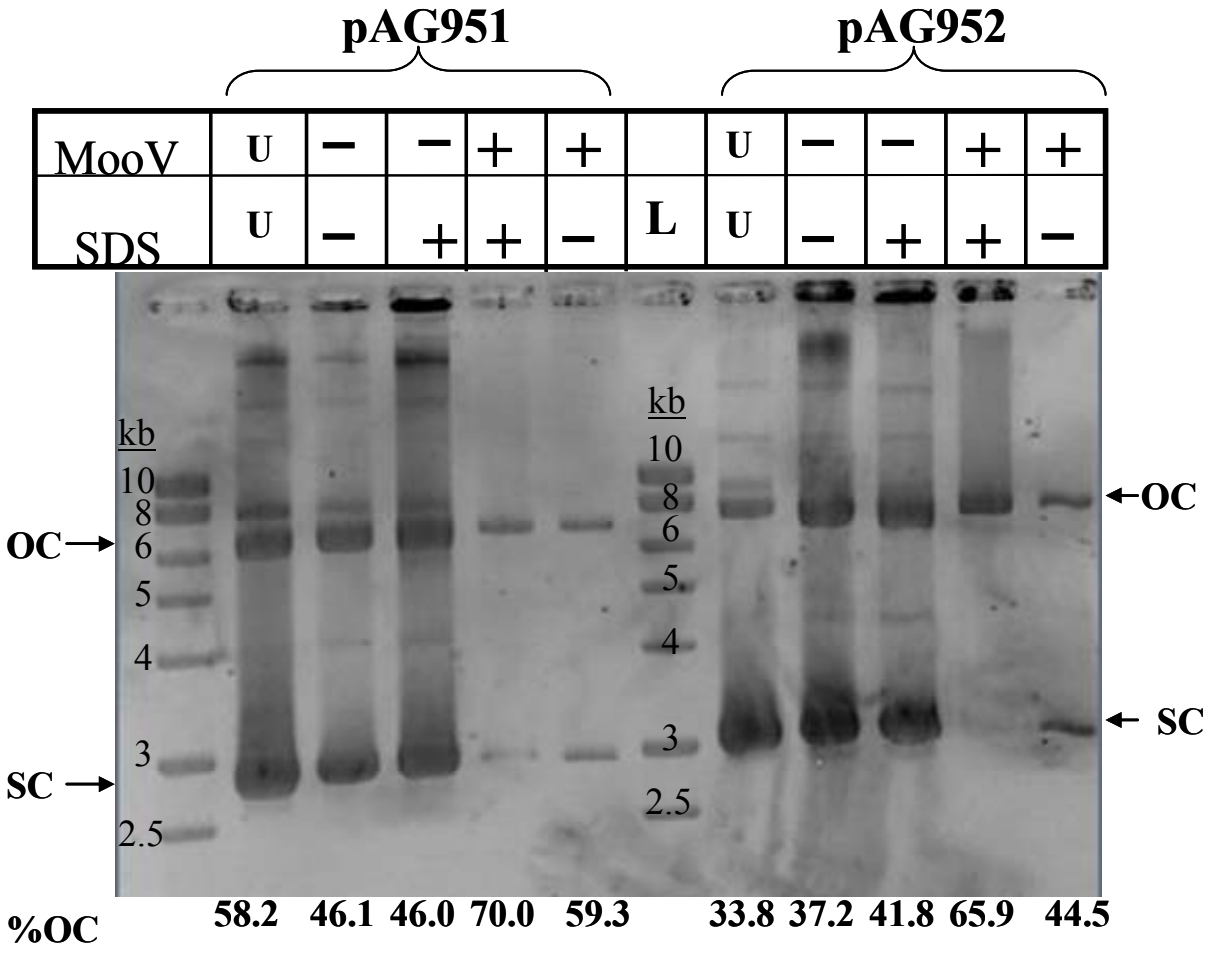

pAG951
pAG952
MooV
SDS
U
−
+
L
kb
10
8
6
5
4
3
2.5
OC
SC
%OC
58.2 46.1 46.0 70.0 59.3 33.8 37.2 41.8 65.9 44.5

Figure 4.5: Amount of OC DNA is dependent on amount of MBP-MooV when 0.5% SDS is used to stop the reaction. (A) 1% TBE agarose gel of in vitro nicking assay using decreasing amounts of MBP-MooV and 0.5% SDS to stop reaction. The absence of 0.5% SDS or MBP-MooV is indicated by -. U indicates untreated DNA. pAG951 was the plasmid used in this assay. The percentage of open-circular DNA (OC) compared to super-coiled DNA (SC) within each lane is indicated at the bottom of the gel.

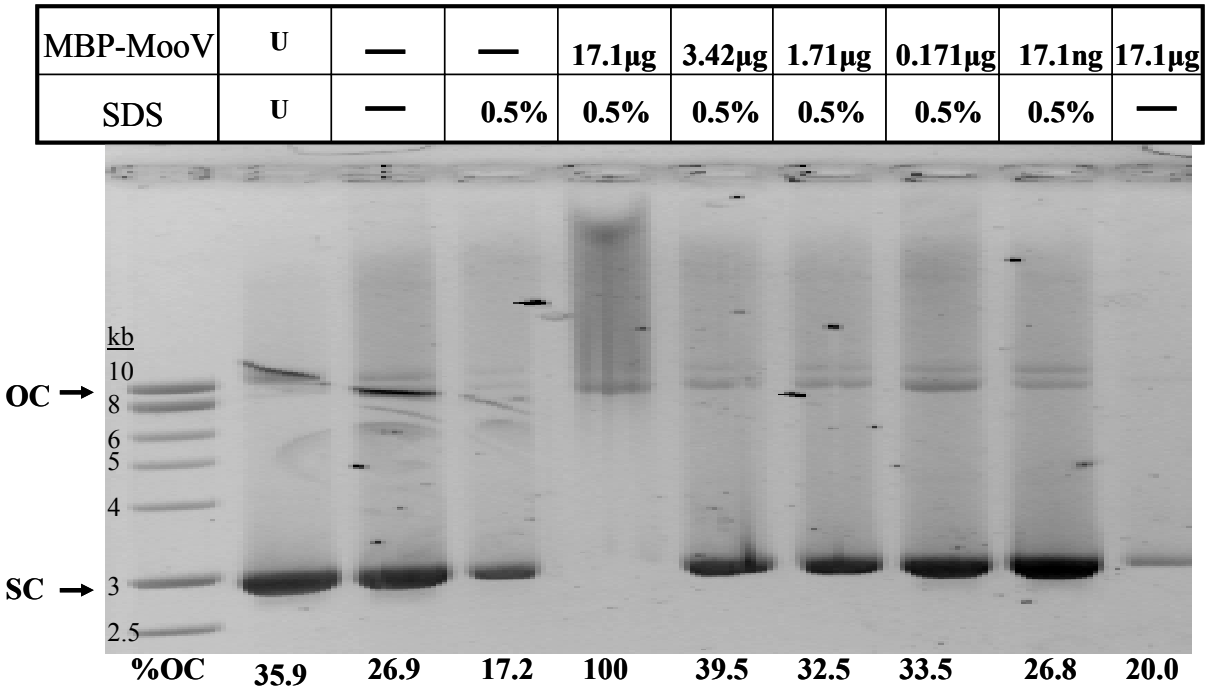
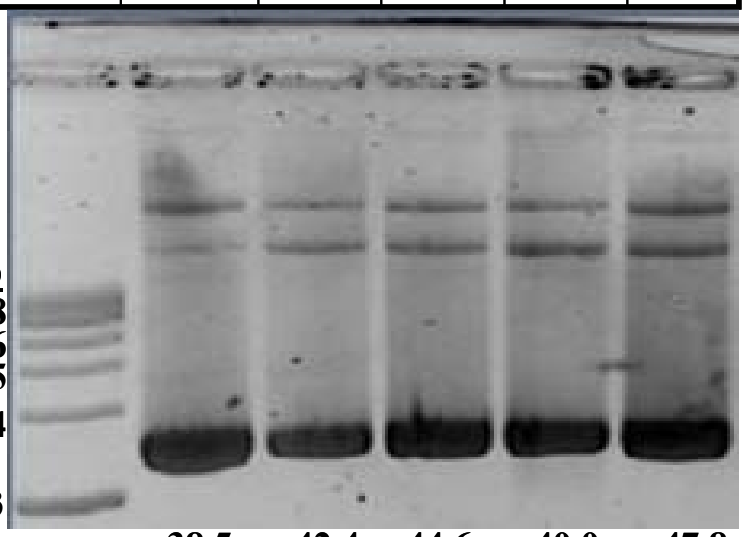

MBP-MooV
U
—
—
17.1µg
3.42µg
1.71µg
0.171µg
17.1ng
17.1µg
SDS
U
—
0.5%
0.5%
0.5%
0.5%
0.5%
0.5%
—
kb
10
8
6
5
4
3
2.5
OC
SC
%OC
35.9
26.9
17.2
100
39.5
32.5
33.5
26.8
20.0

Figure 4.6: *In vitro* DNA nicking assay using 0.5% SDS and proteinase K to stop reaction. 1% TBE agarose gel of *in vitro* nicking assay using either 0.5% SDS or 1 μg proteinase K or both to stop the reaction. Presence or absence of MBP-MooV, SDS, and proteinase K is indicated by + or -, and U indicates untreated DNA. pAG952 was the plasmid used in this assay. The percentage of open-circular DNA (OC) compared to super-coiled DNA (SC) within each lane is indicated at the bottom of the gel.

MooV	U	−	−	+	−
SDS	U	−	+	+	+
Proteinase K	U	−	+	+	−



%OC	38.5	42.4	44.6	40.0	47.8

Figure 4.7*: In vitro* MBP-MooV-His nicking assay followed by labeling with [α-^{32}P]dATP. 5% DNA denaturing gel of pAG952 incubated with MBP-MooV-His under various conditions followed by restriction digest by EcoR1 and labeling with [α-^{32}P]dATP and Klenow. The presence and amount of MBP-MooV-His is indicated in nmol. The presence of 11.2 U of IHF, or 5 mM $CaCl_2$ is indicated by +. The presence and amount of glycerol is indicated by %. The arrows point to starting substrate, which runs as two separate bands in this denaturing gel due to the different nucleotide content of the top strand versus bottom strand.

MBP-MooV-His (nmol)	–	0.025	0.05	0.25	0.5	1	–	1	–	1	–	1	–	1
IHF	–	–	–	–	–	–	+	+	–	–	–	–	–	–
Glycerol	–	–	–	–	–	–	–	–	10%	10%	15%	15%	–	–
$CaCl_2$ (5 mM)	–	–	–	–	–	–	–	–	–	–	–	–	+	+

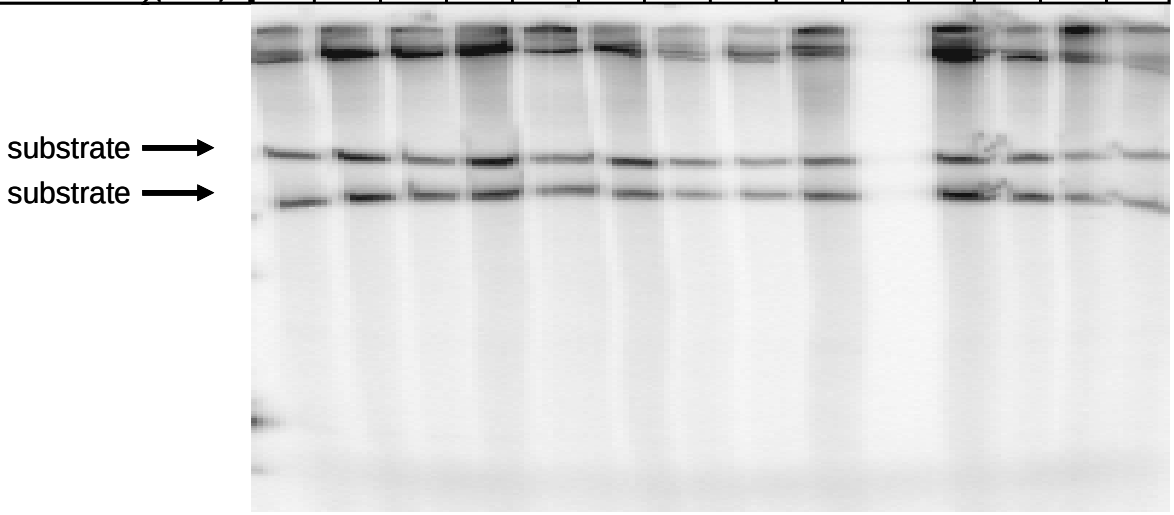
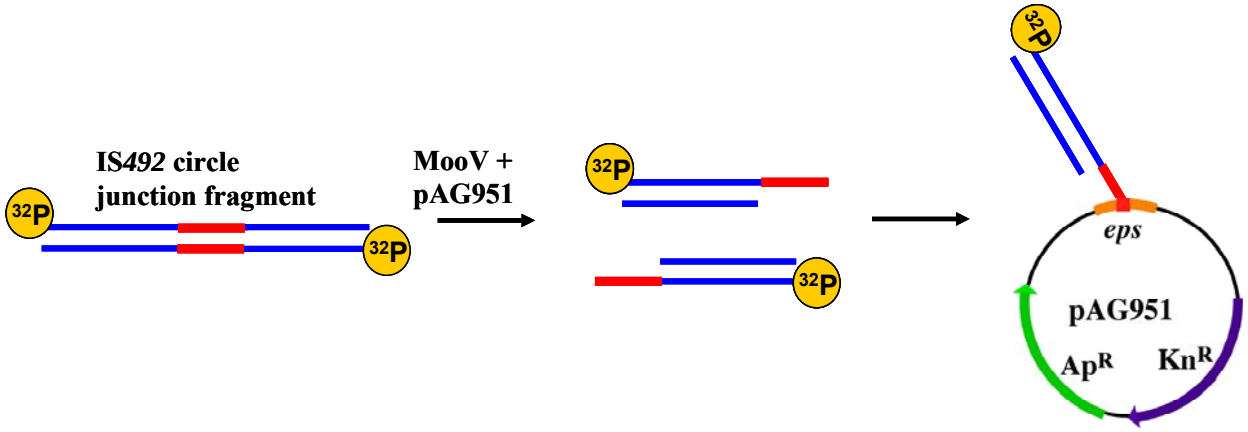


Figure 4.8: MBP-MooV cleavage and strand transfer using *P. atlantica* cell lysates. (A) Diagram of MBP-MooV cleavage and strand transfer. The circle junction fragment restriction fragment is isolated from pAG952 by EcoR1 restriction digest and labeled using [α-^{32}P]dATP and Klenow. (B) 1% TBE agarose gel of [α-^{32}P]-ATP labeled circle junction fragment after incubation with MBP-MooV and target plasmid, pAG951 in the presence or absence of *P atlantica* cell lysates. The presence or absence of MBP-MooV and pAG951 is indicated by a + or -, respectively. Un-reacted substrate is indicated by an arrow.

A.



B.

Target (pAG951)	-	-	-	-	+	+	+	+	+	+
MBP-MooV	-	-	+	+	-	+	-	+	-	+
Lysate (µg)	-	40	-	40	-	-	4	4	40	40

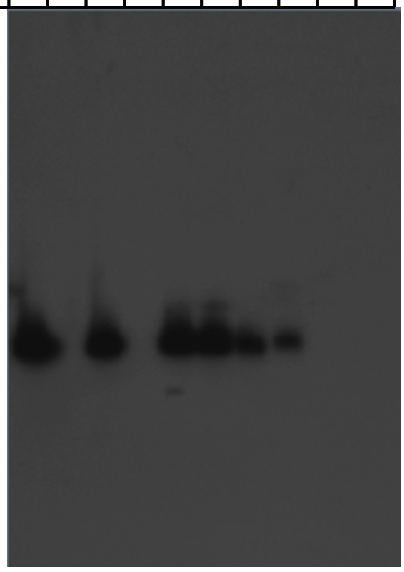
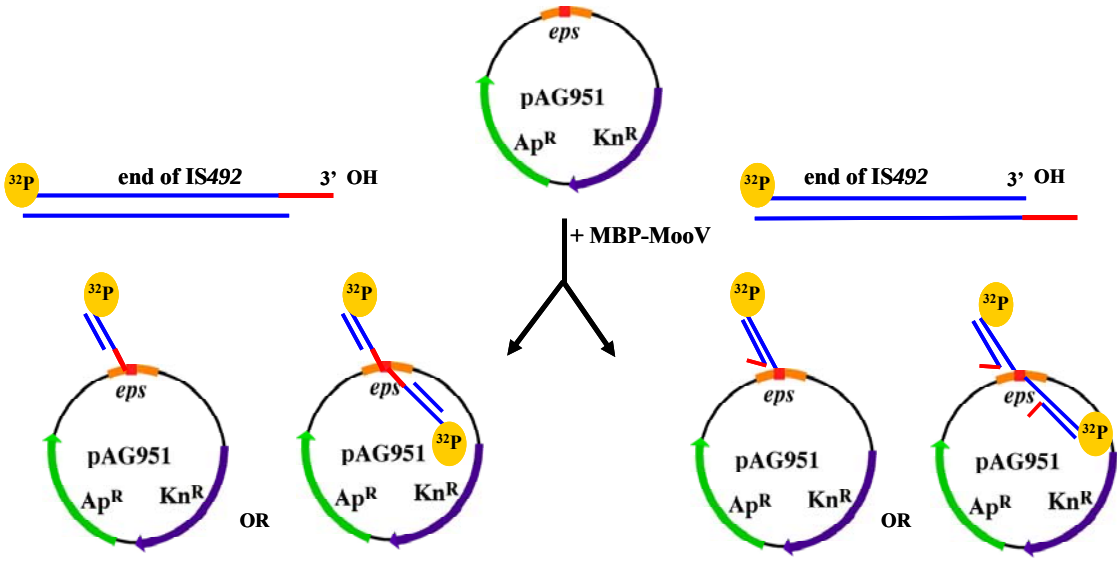


Figure 4.9: MBP-MooV strand transfer assay. (A) Diagram of strand transfer assay. Pre-processed oligonucleotides of the left and right end of IS492 were designed to have 3'-OH. Oligonucleotides were annealed such that the 3'OH was either on the recessed end or on the overhang end. The diagram shows just two possible reaction products with either one or two ds oligonucleotides being transferred. (B) 1% TBE agarose gel of strand transfer assay. ISL-R, left end of IS*492* with 3'OH on recessed end; ISL-O, left end of IS*492* with 3'OH on overhang end; ISR-O, right end of IS*492* with 3'OH on overhang end; ISR-R, right end of IS*492* with 3'OH on recessed end; ISL + R-R, labeled IS*492* left end with 3'OH on recessed end and unlabeled right end of IS*492* with 3'OH on the recessed end; ISL + R-O, labeled left end of IS*492* with 3'OH on overhang end and unlabeled right end of IS*492* with 3'OH on overhang end.

A.



B.

ds oligonucleotide	ISL-R		ISL-O		ISR-O		ISR-R		ISL+R-R		ISL+R-O	
MBP-MooV	-	+	-	+	-	+	-	+	-	+	-	+

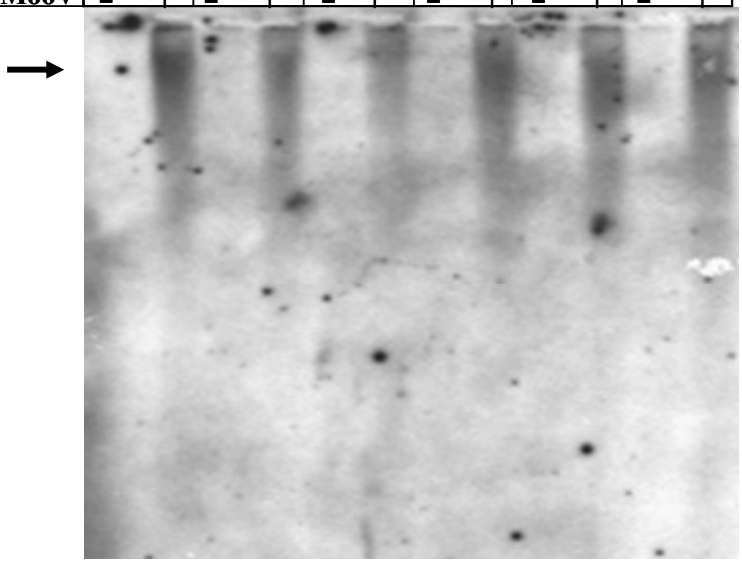
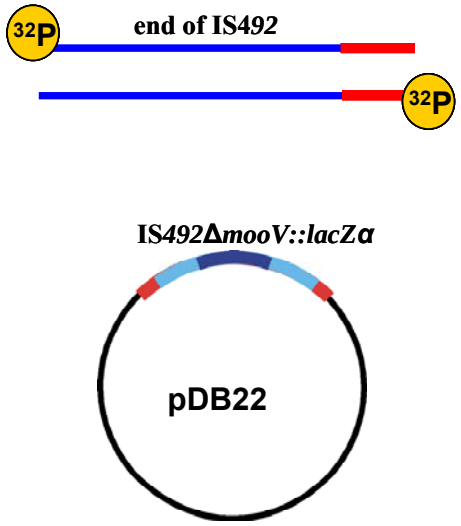
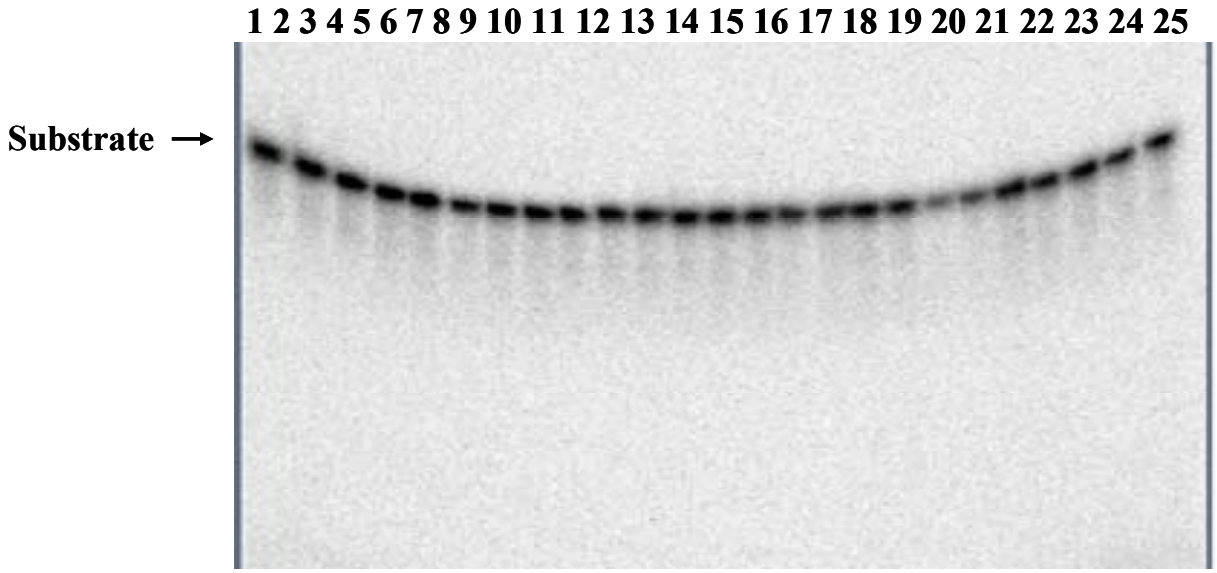


Figure 4.10: MBP-MooV-His nicking as a function of time and protein concentration. (A) Diagram of oligonucleotide and target plasmid used in nicking and strand transfer assay. Both strands of the left end of IS*492* are labeled before being annealed. The target plasmid is pDB22 which contains an IS*492*Δ*mooV*::*lacZα* derivative of IS*492* flanked by 23bp of *epsG* on either side. (B) 20% DNA-denaturing polyacrylamide gel of MBP-MooV-His nicking of labeled ds IS*492* left oligonucleotide incubated with Factor Xa-cleaved MBP-MooV-His and target, pDB22 to detect nicking of ds oligonucleotide. Un-reacted substrate is indicated by an arrow. Lanes 1-5 represent no protein added, lanes 6-10 represent 2 nM protein added, lanes 11-15 represent 20 nM protein added, lanes 16-20 represent 200 nM protein added, and lanes 21-25 represent 2000 nM protein added. For each protein amount used, time points were taken every 5 min starting with a time point at 0; therefore, lanes 1, 6, 11, 16, and 21 represent time 0, lanes 2, 7, 12, 17, and 22 represent 5 min incubation, lanes 3, 8, 13, 18, and 22 represent 10 min incubation, lanes 4, 9, 14, 19, and 24 represent 15 min incubation, and lanes 5, 10, 15, 20, and 25 represent 20 min incubation.

A.


32P
end of IS492
32P
IS492ΔmooV::lacZα
pDB22

B.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Substrate

Chapter 5

Summary and Discussion

Specialized DNA recombination creates genetic diversity by altering the genetic information through DNA insertions, deletions, and inversions. The two major groups of recombinases that mediate specialized DNA recombination are the serine- or tyrosine- site-specific recombinases, which mediate site-specific recombination, and the DDE-motif recombinases, which mediate transposition. As described in Chapter 1, the molecular mechanisms that these recombinases use are quite different. Briefly, site-specific recombination occurs between DNA strands with short regions of homology and proceeds through a two-step transesterification mechanism in which a covalent DNA-recombinase intermediate is formed. The site-specific recombination reaction is conservative, meaning no DNA is lost or gained in the reaction. Transposition occurs between DNA strands with no homology and proceeds through hydrolysis followed by one-step transesterification in which there is no covalent DNA-recombinase intermediate. The transposition reaction is non-conservative, requiring host DNA repair/replication machinery to repair regions containing transposase-mediated nicks or gaps.

Recombinases of the Piv/MooV family mediate recombination reactions that appear to be site-specific because their products reflect a conservative recombination mechanism. However, based on primary amino acid sequence alignment all of the recombinases in the Piv/MooV family have a conserved acidic residue tetrad that is within a ribonuclease H-fold similar to the catalytic domains of DDE-motif transposases and DEDD-motif Holliday-junction resolvases (3, 5, 8). Thus, these novel systems of specialized recombination have elements of both site-specific

recombination and DNA transposition and may reveal features relevant to other specialized recombination systems; therefore, my research has been to characterize the defining members of the family, Piv and MooV.

Scanning mutagenesis showed that a 15 bp region of *invR* (positions 2 to7 and 9 to 17) is required for Piv-mediated inversion of the pilin invertible segment, and, an inversion assay with a mutant that has complementary mutations within *invL* and *invR* (14/17QI) showed that homology at the crossover region does not rescue an inversion-minus mutant, suggesting that the specific nucleotide sequence is important for the inversion reaction. Although MBP-Piv has been shown to interact weakly with *invL* in vitro, we were never able to detect MBP-Piv binding to ds *invR* to determine if MBP-Piv could bind the mutant *invR* as well as wild-type *invR*, but we were able to detect MBP-Piv binding to ss *invR*.

Further investigation into MBP-Piv binding to ss DNA revealed that MBP-Piv ss binding activity is not sequence specific but could be structurally specific. MBP-Piv was unable to bind an oligonucleotide (LacZDra) that has a predicted 2° structure with a stable hairpin positioned in the middle of the oligonucleotide in competition assays with oligonucleotides without stable 2° structures. While there are recombinases that recognize and interact with hairpins formed by ssDNA folding (IntI integron integrase and TnpA of ISHp608), other recombinases, such as Flp, a tyrosine recombinase, can bind ssDNA in a sequence-specific manner (2, 6, 7, 10). There is no experimental evidence that Flp ssDNA binding activity has a role in recombination although it has been proposed that binding of two Flp molecules to its recognition target (FRT) induces a bend in the DNA that is severe enough to separate the strands at the cores releasing ssDNA that can be bound by Flp to mediate strand exchange (10). Perhaps Piv binding to the *sub* site facilitates Piv-mediated opening of dsDNA at its recombination site allowing Piv to bind ssDNA

before mediating cleavage and strand exchange. If Piv binding to ssDNA does not have a role in its recombination reaction, another possible reason we were unable to detect MBP-Piv binding to ds *invR* is that MBP-Piv must bind a *sub* site before binding *inv*. The ability of Piv to bind ssDNA brings an interesting view to how recombinases of the Piv/MooV family are able to recognize and bind their insertion sequences.

The insertion sequences that encode the recombinases of the Piv/MooV family belong to the IS*110*/IS*492* family of insertion sequences. One of the hallmarks of this family of insertion sequences is the lack of terminal inverted repeats at the ends of the elements. IS*492*, whose movement is catalyzed by MooV, lacks terminal inverted repeats, and there is no homology between the left and right ends of the element or between the element and its target sites. The DEDD-motif of MooV has been shown to be required for the excision, circularization and insertion of IS*492* (Chapter 3). If we assume that these residues are required for catalysis but not binding, then the catalytic domain of MooV resembles that of RuvC which uses its DEDD catalytic residues positioned within an RNase H-fold to coordinate divalent metal cations to facilitate the hydrolysis of the phosphodiester bond in substrate DNA in order to resolve Holliday junctions (1, 9).

The working model for Piv, whose DEDD motif is required for catalysis but not binding, involves the formation of a Holliday junction intermediate in the recombination reaction that is resolved by Piv followed by ligation by host DNA ligase to repair the nicks (3). A similar model has been proposed for MooV-mediated excision because it offers an explanation for how MooV precisely excises IS*492* without leaving double-stranded break in the donor DNA (see Chapter 3). A major issue with this model is that IS*492*, like other members of the IS*110*/IS*492* family of insertion sequences, does not have terminal inverted repeats, so how does MooV recognize both

ends of IS*492* and its target? Perhaps, MooV, like Piv, is capable of binding ssDNA. None of the in vitro cleavage or strand transfer assays described in Chapter 4 used ssDNA as substrate for MBP-MooV. The use of ssDNA and/or hyperactive insertion variants of MooV may elucidate the characteristics of its binding activity.

Another problem with the model described above is that unlike IS*492* which has direct repeats upon insertion that could base pair at the junction, some members of the IS*110*/IS*492* family do not have direct repeats flanking the inserted element. Perhaps some of the elements of the IS*110*/IS*492* recombine via a Holliday junction intermediate and others do not. Another possibility is that the other conserved amino acids within the Piv/MooV family are part of the catalytic motif. For example, one of the conserved amino acid motifs, PSG, has been shown to be required for Piv-mediated inversion but not Piv binding to *inv* (8). Site-directed mutagenesis of the conserved serine within PSG indicates that the serine is required for Piv-mediated inversion. It is not known whether the serine residue is required for binding (John Buchner, personal communication). This PSG motif is found within MooV and nearly all of the other members of the Piv/MooV family of recombinases, and it is possible that this serine residue plays a role in recombination. If the serine residue is required for catalysis then it could be possible that the recombinases have two separate catalytic domains where the conserved serine in one domain initiates strand cleavage and transfer of two strands to produce a Holliday junction and the DEDD motif of the other domain resolves the Holliday junction. Another possibility is that the serine residue is the actual residue involved in cleavage and strand transfer and the DEDD motif is involved in stabilizing the reaction, perhaps by donating protons to activate the serine for the first transesterification reaction. Although there are transposases that do have serine catalytic residues [for review see (4)], the recombinases of the Piv/MooV family are not

related to these transposases based on primary amino acid alignment and are site-specific unlike the S-transposases. If the serine residue is part of the catalytic domain in the recombinases belonging to the Piv/MooV family, the recombination reaction would proceed through different intermediates than that catalyzed by S-transposases.

There is still little known about the catalytic mechanism utilized by the recombinases of the Piv/MooV family. As described above, there are several possible reaction mechanisms that could be used. As a result, it is apparent that further development of biochemical assays is needed to elucidate the steps in both Piv- and MooV-mediated recombination. Some of these assays include: in vitro sandwiching and looping assays with oligonucleotides containing both *sub* and *inv* sequences to resolve how Piv interacts with *invL* and *invR*; isolation of hyperactive insertion variants of MooV for use in the in vitro assays described in Chapter 4; and, the use of ssDNA in electrophoretic mobility shift assays with MooV.

Piv recognition of and binding to its recombination site, *inv*, is an important first step in determining its catalytic mechanism. The looping and sandwiching assays will answer the question whether Piv requires binding to ds *sub* before binding to ds *inv* or if Piv binding to ds *sub* opens the DNA at *inv* to release ssDNA that is then bound and cleaved by Piv. It is important to determine if other recombinases in the family which do not have inverted repeats at their recombination sites are capable of binding ssDNA. To address this issue, ssDNA binding activity of MooV will be performed using both the left and right ends of IS*492* and the sequence at the junction of the circular form of IS*492*. If MooV binding to ssDNA is not site-specific, as is the case with Piv, then, like proposed for Piv, the ssDNA binding activity may not be important for MooV-mediated recombination, or MooV may bind to an accessory site causing the opening of the DNA at the recombination site. If ssDNA binding, that is not sequence

specific, is required for both MooV and Piv recombination, then binding to accessory sites must deliver the catalytic domains of the recombinases to the regions where strand exchange takes place.

When we isolate a recombinase with in vitro activity, the in vitro assays described in Chapter 4 will provide the information necessary to determine if the recombinases of the Piv/MooV family catalyze recombination through the conserved DEDD motif or through the conserved PSG motif. Isolation of a cleaved DNA intermediate will indicate whether MooV is covalently linked to the DNA (which would occur if the serine is the catalytic residue) or not covalently linked to the DNA (which would occur if the DEDD motif is involved in cleavage). If MooV uses both the PSG and the DEDD motif, then a cleaved DNA intermediate will be covalently linked to MooV, and strand transfer will result in the formation of a Holliday junction intermediate that will be resolved using the DEDD motif. The nicks at the site of resolution can then be ligated by adding ligase to the in vitro reaction.

In conclusion, the unusual characteristics of the recombinases of the Piv/MooV family raise several questions as to how these recombinases catalyze recombination. Further study of the defining members of the family, Piv and MooV, using the biochemical assays described in Chapter 4 will shed light on the reaction mechanism utilized by the recombinases of the Piv/MooV family.

References

1. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, A. Fujishima, H. Shinagawa, and K. Morikawa.** 1994. Preliminary crystallographic study of *Escherichia coli* RuvC protein. An endonuclease specific for Holliday junctions. J Mol Biol **241:**281-2.
2. **Barabas, O., D. R. Ronning, C. Guynet, A. B. Hickman, B. Ton-Hoang, M. Chandler, and F. Dyda.** 2008. Mechanism of IS*200*/IS*605* family DNA transposases: activation and transposon-directed target site selection. Cell **132:**208-20.
3. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J Bacteriol **187:**3431-7.
4. **Curcio, M. J., and K. M. Derbyshire.** 2003. The outs and ins of transposition: from mu to kangaroo. Nat Rev Mol Cell Biol **4:**865-77.
5. **Lenich, A. G., and A. C. Glasgow.** 1994. Amino acid sequence homology between Piv, an essential protein in site-specific DNA inversion in *Moraxella lacunata*, and transposases of an unusual family of insertion elements. J Bacteriol **176:**4160-4.
6. **MacDonald, D., G. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul.** 2006. Structural basis for broad DNA-specificity in integron recombination. Nature **440:**1157-62.
7. **Ronning, D. R., C. Guynet, B. Ton-Hoang, Z. N. Perez, R. Ghirlando, M. Chandler, and F. Dyda.** 2005. Active site sharing and subterminal hairpin recognition in a new class of DNA transposases. Mol Cell **20:**143-54.
8. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-

specific recombinases are required for catalysis of DNA inversion by Piv. Mol Microbiol **39:**641-51.

9. **Yang, W., and T. A. Steitz.** 1995. Recombining the structures of HIV integrase, RuvC and RNase H. Structure **3:**131-4.

10. **Zhu, X. D., and P. D. Sadowski.** 1998. Selection of novel, specific single-stranded DNA sequences by Flp, a duplex-specific DNA binding protein. Nucleic Acids Res **26:**1329-36.

Appendix A

Scanning Electron Microscopy of *Pseudoalteromonas atlantica*[1]

[1] Carpenter, C.D. and Karls, A.C. 2008. Unpublished

Introduction

Pseudoalteromonas atlantica is a Gram-negative marine bacterium capable of forming biofilms on solid surfaces in the marine environment. This organism was first isolated in association with marine algae and named *Pseudomonas*; however, after 16S rRNA sequencing, it was assigned to a new genus, *Pseudoatleromonas*, along with several *Alteromonas* species (4, 12). *P. atlantica* is associated with solid surfaces in the water column and eukaryotic hosts, such as algae and crab, where its production of extracellular compounds influences succession in marine communities (2, 4, 7, 12). For example, several *Pseudoalteromonas* species produce antibacterial, anti-fungal, and anti-viral compounds (7). *Pseudoaltermonas atlantica* produces several extracellular compounds including acidic extracellular polysaccharide (EPS), proteases, and enzymes capable of hydrolyzing agar and chitin (1, 3, 6).

The production of EPS is essential in the biofilm-forming abilities of *P. atlantica*. Microfouling studies performed in the oceans off the coasts of California, New Jersey, and Florida showed Pseudoalteromonas species were the first organisms to become attached and produced large amounts of EPS. After biofilm formation by *Pseudoalteromonas*, stalked and/or filamentous bacteria attached to the surface followed by the attachment of diatoms, microalgae, and sessile protozoa (3). The advantages of biofilm formation in oceans include the attraction of nutrients to the negatively charged surface of the biofilm and protection from bacteriophage (5).

We were awarded a grant from Department of Energy/Joint Genome Institute (DOE/JGI) to sequence the genome of *P. atlantica* and the grant required an electron micrograph of the bacterium. We report here scanning electron micrographs (SEM) of extracellular polysaccharide producing *P. atlantica* and describe the interesting observations from the SEMS.

Materials and Methods

Strains. *Pseudoalteromonas atlantica* T6c (2).

Scanning electron microscopy. *P. atlantica* cells from a mucoid (EPS-producing) colony were grown to mid-log phase in marine broth under aeration. A 100µl aliquot of cells was fixed in Parducz solution (3 parts 4% OsO_4, 3 parts distilled H_2O, and 1 part 2% saline:saturated $HgCl_2$) for 1 hour, washed twice with 2% saline and dehydrated with sequential washes in 25, 35, 50, 70, 85, 95, and 100% ethanol (8, 10). Dehydrated samples were applied to 0.2 µm Whatman filter and dried by a Samdri-780A critical point drying apparatus with liquid CO_2. Samples were coated with gold using SPI-Module Coater to approximately 76.4 Å in thickness examined with a LEO 982 FE-SEM at 5 kV.

Results

As seen in Figure 1, the extracellular matrix produced by *P. atlantica* is used to attach the cells to one another. Between the *P. atlantica* cells, there are small lengths of extracellular polysaccharide that stretch from cell to cell. This is similar to *V. cholerae* cells in biofilms that are surrounded by long finger-like projections of extracellular polymeric material (11), (9). Because the *P. atlantica* cells were aerated during growth, a biofilm did not form; however, it is possible that *P. atlantica* cells in a biofilm may form longer extracellular polysaccharide projections similar to that seen with *V. cholerae* cells in biofilms.

Discussion

We show here an SEM of EPS-producing *P. atlantica* that sheds more light on how the EPS is used in the attachment of *P. atlantica* to solid surfaces in the marine environment. *P.*

atlantica is the first organism present in biofilms in the marine environment, and, only after its establishment of a biofilm do other organisms add to the biofilm (3). The biofilm production on solid surfaces submerged in the sea causes large economic losses to the marine industry in which structures, such as ships, are exposed to sea water and become fouled (3). Further study of the extracellular matrix of *P. atlantica* and other members of marine biofilms will elucidate mechanisms to better prevent microfouling of structures in the marine environment.

References

1. **Akagawa-Matsushita, M., M. Matsuo, Y. Koga, and K. Yamasato.** 1992. *Alteromonas atlantica* sp. nov. and *Alteromonas carrageenovora* sp. nov., bacteria that decompose algal polysaccharides. Int J Syst Bacteriol **42:**621-27.
2. **Corpe, W. A.** 1980. Adsorption of Microorganisms to Surfaces, p. 105-44. Wiley, New York.
3. **Corpe, W. A.** 1973. Presented at the Third International Congress on Marine Corrosion and Fouling, Evanston, IL.
4. **Costa-Ramos, C., and A. F. Rowley.** 2004. Effect of extracellular products of *Pseudoalteromonas atlantica* on the edible crab *Cancer pagurus*. Appl Environ Microbiol **70:**729-35.
5. **Davey, M. E., and A. O'Toole G.** 2000. Microbial biofilms: from ecology to molecular genetics. Microbiol Mol Biol Rev **64:**847-67.
6. **Hoffman, M., and A. W. Decho.** 2000. Proteolytic enzymes in the marine bacterium *Pseudoalteromonas atlantica*: post-secretional activation and effects of environmental conditions. Aquat Microb Ecol **23:**29-39.
7. **Holmstrom, C., and S. Kjelleberg.** 1999. Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. FEMS Microbiol Ecol **30:**285-93.
8. **Kearns, D. B., P. J. Bonner, D. R. Smith, and L. J. Shimkets.** 2002. An extracellular matrix-associated zinc metalloprotease is required for dilauroyl phosphatidylethanolamine chemotactic excitation in *Myxococcus xanthus*. J Bacteriol **184:**1678-84.

9. **Mizunoe, Y., S. N. Wai, A. Takade, and S. I. Yoshida.** 1999. Isolation and characterization of rugose form of *Vibrio cholerae* O139 strain MO10. Infect Immun **67:**958-63.

10. **Parducz, B.** 1967. Ciliary movement and coordination in ciliates. Int Rev Cytol **21:**91-128.

11. **Wai, S. N., Y. Mizunoe, A. Takade, S. I. Kawabata, and S. I. Yoshida.** 1998. *Vibrio cholerae* O1 strain TSI-4 produces the exopolysaccharide materials that determine colony morphology, stress resistance, and biofilm formation. Appl Environ Microbiol **64:**3648-55.

12. **Yaphe, W.** 1957. The use of agarase from *Pseudomonas atlantica* in the identification of agar in marine algae (Rhodophyceae). Can J Microbiol **3:**987-93.

Figure 1: Scanning electron micrographs of *P. atlantica* T6c. P. atlantica cells producing EPS were grown to mid-log phase under aeration. Samples were prepared for scanning electron microscope as described in Materials and Methods. White Bar = 2 μm.