GENETIC CHARACTERIZATION OF IS*492* TRANSPOSITION IN

Escherichia coli

by

ADRIENNE MARIE COTTRELL

(Under the Direction of Anna C. Karls)

ABSTRACT

IS*492* transposition in the marine bacterium *Pseudoalteromonas atlantica* controls the phase variation of extracellular polysaccharide (EPS). IS*492* inserts into a single site in *epsG*, a predicted glycosyl transferase gene important for EPS production, and precisely excises to form a circular element. This reversible insertion of IS*492* into *epsG* is mediated by the DEDD recombinase MooV. Recombinases in the novel Piv/MooV family perform site-specific recombination and DNA transposition, an unprecedented characteristic for a single recombinase family, however the recombination mechanism used by these recombinases remains to be elucidated. The goal of this research was to characterize the recombination mechanism used in MooV-mediated IS*492* transposition. This study demonstrates that MooV-mediated IS*492*Δ*mooV*::*tet* insertion in *E. coli* is facilitated by using temperatures close to the growth optimum for *P. atlantica*. Single IS*492*Δ*mooV*::*tet* insertion into the *epsG* site duplicates the 5-bp 5'-CTTGT-3' sequence. Tandem insertion into the IS*492*Δ*mooV*::*tet* donor plasmid also targets this 5-bp sequence flanking the element, where the tandem element, (IS*492*Δ*mooV*::*tet*)2, is separated and flanked by 5'-CTTGT-3'. This suggests that the mechanism for insertion is either site-specific recombination or cut-and-paste transposition where the 5-bp sequence is the

common core of the recombination sites or the recognition sequence for the staggered cut by MooV, respectively. Chromosomal insertions were not observed, suggesting a target site preference. Experiments using the (IS*492*Δ*mooV*::*tet*)2 with the *epsG* target plasmid show that MooV-mediated transposition of (IS*492*Δ*mooV*::*tet*)2 is less dependent on lower temperatures and is more active for transposition compared to IS*492*Δ*mooV*::*tet*. Precise excision of one IS*492*Δ*mooV*::*tet* from the donor plasmid was observed, and the most frequent insertion products were cointegrates that had two directly-oriented IS*492*Δ*mooV*::*tet* elements flanking the target and donor plasmids with 5'-CTTGT-3' at the junctions. These results suggest that the inner ends of (IS*492*Δ*mooV*::*tet*)2 formed the active site for targeted insertion into the *epsG* site by either a site-specific or cut-and-paste mechanism, and that the rate-limiting step in IS*492* transposition is the formation of the circle-junction sequence.

INDEX WORDS: transposition, mobile genetic elements, specialized recombination, insertion sequence, IS*492*

GENETIC CHARACTERIZATION OF IS*492* TRANSPOSITION IN *Escherichia coli*

by

ADRIENNE MARIE COTTRELL

B.S., Towson University, 2001

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

GENETIC CHARACTERIZATION OF IS*492* TRANSPOSITION IN

Escherichia coli

by

ADRIENNE MARIE COTTRELL

Major Professor: Anna C. Karls

Committee: Timothy R. Hoover
Joy Peterson
Lawerence Shimkets
Eric V. Stabb

Electronic Version Approved:

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

DEDICATION

In loving memory of my grandparents Mr. Samuel Cottrell, Mr. Stewart McTeer, and Mrs. Fannie McTeer. I wish that you were here to see me now, but I know you are looking down from heaven with all smiles.

ACKNOWLEDGEMENTS

"Being confident of this, that he who began a good work in you will carry it on to completion until the day of Christ Jesus." Philippians 1:6.

To God be the glory for all of the great things He has done.

I would like thank my advisor Dr. Anna C. Karls for always believing in me, and encouraging me to finish my degree at times when I wanted to give up. I express gratitude to my committee members, Drs. Joy Peterson, Timothy Hoover, Larry Shimkets, and Eric Stabb for their wisdom in keeping my project moving forward. My appreciation to Dr. Anne Summers for recruiting me to The University of Georgia and to Dr. Ellen Neidle for her constant support. I have been blessed with additional mentors who took interest in my academic and professional development. Special thanks to Dr. Kojo Mensa-Wilmot for always providing good advice and always being available when I needed additional guidance; to Drs. Charlene Jackson (USDA) and Shawn Drew (NIH) for showing me that under-represented women can have successful careers in science. Members of the Karls group, past (especially Dr. Richard "Tad" Seyler, Dr. Brian Higgins, and Anne Robertson) and current, the Timothy Baptist Church, Church of the Nations, and my home church, Perkins Square Baptist Church.

To my husband Austin Berinyuy Yongye, Ph.D., thank you for all of your love, support, and words of encouragement to help me finish graduate school. Special thanks to my prayer partners Kahtonna Allen, Michelle Eason, Cynthia Grant, and Crystal Jackson, thank you for praying for

me and with me during all the good and difficult times. To Dr. Danielle Webster, thank you for your support all throughout graduate school. Special thanks to all the members of Scholars for Diversity in STEM Disciplines.

Words cannot express how thankful I am to my family: my mother Janice Cottrell and father Angelo Cottrell, thank for every prayer and always encouraging me to finish my degree. I know that all of my hard work "will pay off." To my sister Angela Cottrell, thank you for all of your support that only a sister can provide. To my grandmother Gladys Cottrell, thank you for your love, wisdom, and prayers.

TABLE OF CONTENTS

Page

CHAPTER 1

Introduction – Literature review

1.1. Introduction

DNA recombination plays a significant role in creating genetic diversity from bacteriophages to humans. General recombination of DNA in important cellular processes such as meiosis in sexually reproducing organisms[1] occurs via homologous recombination, which involves the highly conserved eukaryotic Rad51[2] and prokaryotic RecA enzymes.[3] In contrast to homologous recombination, which requires extensive sequence homology, many different enzymes, and a high energy cofactor (e.g. ATP),[3] specialized recombination via site-specific recombination and DNA transposition is based on limited or no sequence homology, and uses specific enzymes (recombinases) that cleave and exchange defined segments of DNA without the requirement of high energy cofactors.[4] In site-specific recombination predetermined loci are recognized by their cognate site-specific recombinases and DNA rearrangement involving insertion, deletion, or inversion typically produces a biological function, including integration of bacteriophage lambda into a bacterial chromosome[5] or alternate flagellin synthesis in *Salmonella* sp.[6] In DNA transposition the transposase mobilizes its associated transposable element from one genomic location to another. Transposable elements can interrupt or turn on gene expression and introduce gene products such as insertion sequence *256* (IS*256*) or transposon *916* (Tn*916*), which are involved in alternate gene expression of polysaccharide intercellular adhesin (PIA) in *Staphylococcus epidermidis*[7] and the spread of antibiotic resistant genes through conjugative transposition,[8] respectively. However, movement of mobile elements does not always alter

expression of a biological product or change cellular processes. Although site-specific recombination and transposition are similar in that these processes utilize special enzymes to mediate DNA cleavage and strand exchange in the absence of ATP, often require host encoded proteins as accessory factors, and are highly regulated,[4] they use distinct catalytic mechanisms that are reflected in their conserved catalytic motifs.

1.2. Catalytic mechanisms of specialized DNA recombination

DNA Transposition

In classical DNA transposition the transposase has a conserved acidic amino acid triad, aspartic acid, aspartic acid, and glutamic acid (DDE), in its catalytic core.[9] Integrases for retroviruses and retrotransposons also have this conserved DDE catalytic motif.[10] Transposition mediated by DDE transposases can be categorized into two pathways, (1) cut-and-paste or non-replicative (Figure 1.1) and (2) replicative (Figure 1.2), and for each pathway there are three main catalytic steps, (1) strand cleavage, (2) strand transfer, and (3) replication/gap repair. Typically the transposition reaction for both cut-and-paste and replicative pathways does not require homologous sequence between the transposable element and the target sequence and is non-reciprocal (DNA can be lost in the donor molecule during the excision reaction and gained in the target molecule during integration and repair/replication). In replicative transposition a copy of the element is left at the donor site while in cut and paste transposition movement of the transposable element from the donor to recipient molecule does not restore the donor molecule, leaving a double strand break in the donor DNA. Both transposition pathways follow a hydrolysis and one-step transesterification pathway, which does not involve a covalent DNA-protein intermediate (Table 1.1).

During the DNA cleavage step in transposition, divalent metal cations (e.g. Mg^{2+} and Mn^{2+}) bind in the catalytic pocket formed by the DDE residues and coordinate a water molecule as the attacking nucleophile.[11] The transposase-mediated hydrolysis at the 3'-ends of the transposable element releases free 3'-OH ends (Figure 1.1). In the cut-and-paste pathway the 5'-ends of the transposable element are also cleaved, freeing the element from the donor backbone. Various mechanisms are used to cut at the 5'-ends of the element, including a transposon-encoded endonuclease in the case of Tn*7*[12] and transposase-mediated hairpin formation for IS*50*.[13] Strand transfer via a transesterification reaction occurs when the free 3'-OH at each end of the transposable element attacks the target sequence in a staggered manner, resulting in insertion. This insertion event results in free 3'-OH ends in the target DNA and gaps between the target and transposable element. In the final step, host DNA replication machinery uses the free 3'-OH of the target DNA to initiate replication or repair of the gaps, creating the characteristic target site duplication seen in classical transposition; this replication or repair step does require ATP, as opposed to the DNA strand cleavage and exchange steps.

The first strand cleavage step in the replicative pathway is similar to the cut-and-paste pathway, except the 5'-ends of the transposable element are not cleaved (Figure 1.2). After transfer of the 3'-OH end of the transposable element to the 5'-end of the target sequence, the transposable element is linked to both the donor molecule and target DNA, forming a Shapiro intermediate.[9] At this point, the 5'-ends of the transposable element can be cleaved from the donor DNA, and the final outcome of the transposition reaction will resemble a cut-and-paste pathway (e.g. single transposable element insertion into a target molecule). If the 5'-ends of the transposable element are not cleaved from the donor backbone, then DNA replication initiating from the free 3'-OH group from the target DNA will not only repair the gaps between the target

and transposable element (creating target site duplications) but will also copy the entire transposable element. The end product of transposition would be replicon fusion or cointegrate formation between the donor and target molecules, with two copies of the transposable element in direct orientation flanking the two replicons.

Characterization of the *in vivo* end products of classical transposition cannot determine what pathway was used (Figure 1.3). The earliest example was the *in vivo* characterization of Tn*10* transposition, which showed movement of Tn*10* on the chromosome could result in a copy of Tn*10* at the donor and target site, suggesting replicative transposition.[14] It was then demonstrated that Tn*10* moves by a cut and paste mechanism following the movement of a replication fork through the element and repair of the double strand break at the donor site by gap repair with the sister chromosome replacing the Tn*10* element at the donor site. Homologous recombination between two identical transposable elements on different donor backbones can produce a replicon fusion, similar to what is seen in the replicative pathway.[9] Other examples are illustrated in Figure 1.3. If a transposable element is present in a tandem head-to-tail formation on a dimer donor plasmid, insertion of the dimer plasmid via a cut-and-paste pathway using the inner junction of the tandem elements as the cleavage site would result in cointegrate formation (Figure 1.3). This type of cointegrate formation via the cut-and-paste pathway has been observed for IS*21*[15] and IS*30*.[16] In addition, transposable elements in the Tn*3* family, which use the replicative pathway, encode their own site-specific resolvase to resolve the cointegrate insertion molecule into two products: one resembling the initial donor replicon; the other resembling a simple insertion into the target DNA (Figure 1.3).[17]

Conservative Site-Specific Recombination

Site-specific recombination is characterized by the precise breaking and rejoining of DNA in the region of limited sequence homology (core) between two defined sites.[18] In contrast to DNA transposition, which uses hydrolysis and a one-step transesterification reaction, does not require sequence homology, is non-reciprocal, and requires DNA replication or repair (Table 1.1), site-specific recombination uses a two-step transesterification reaction, is reciprocal (after one round of recombination, a second round of recombination would restore the exchanging partners back to their original pre-recombination state), DNA is not lost or gained, and replication is not involved (Table 1.1).[18] The recombinases that mediate site-specific recombination can be categorized into two families based on the amino acid that acts as the nucleophile in the first transesterification step of the recombination reaction: the tyrosine (Y-) or serine (S-) recombinases. In both families the recombination pathway follows a two-step transesterification reaction that involves a covalent DNA-protein intermediate, and can be summarized as follows: (1) synapsis; (2) strand breakage; (3) strand exchange/alignment; (4) ligation.

Members of the S-recombinase family, including the *Salmonella* invertase Hin and the resolvases of the Tn*3* family, execute a concerted breakage, followed by rotation, and ligation of four DNA strands in recombination. To initiate recombination the S-recombinase specifically recognizes and binds the recombination site, which exhibits dyad symmetry around a core sequence where strand exchange occurs, and the recombinase/DNA complexes are brought together during synapsis[18] (Figure 1.4). Within the synaptic complex the OH group of the conserved serine serves as the primary nucleophile and in the first transesterification reaction attacks the top and bottom strands of the core. The double strand breaks are staggered within the

core region, which is shared by the recombining sites, and yields 3'-OH overhangs and recessed 5'-ends (Figure 1.4). During this transesterification step the recombinase is covalently linked to the DNA via a 5' phospho-serine linkage, leaving free 3'-OH groups. In strand exchange a right-handed 180°C rotation occurs between one half of the paired cleaved recombination sites and the overlap regions between the new recombinant top and bottom strands are aligned based on sequence homology.[19] During the second transesterification reaction the free 3'-OH groups of the partner strands attack the 5' phospho-serine linkage, releasing the enzyme from the DNA and resealing the DNA backbone to complete the recombination reaction.

Bacteriophage λ Int, Cre of bacteriophage P1, XerC/XerD of *E. coli*, and other members of the Y-recombinase family mediate recombination through separate and sequential breakage and ligation of two pairs of DNA strands.[5] Recombinases in this family use the OH group of the conserved tyrosine residue as the primary nucleophile in the first transesterification reaction to attack the top strands within the paired recombination core complexes (Figure 1.5). This attack creates a covalent 3' phosphor-tyrosyl enzyme/DNA linkage and a free 5'-OH. The liberated 5'-OH of one strand attacks the 3' phosphor-tyrosyl linkage on the partner strand creating an intermediate, the Holliday junction, which contains two recombinant strands and two parental strands. The Holliday junction is resolved when the remaining two parental strands undergo the same strand cleavage, exchange, and resealing steps to complete the recombination reaction.

The chemical steps in the above described classical transposition and conservative site-specific recombination take place within higher-order nucleo-protein complexes. These complexes commonly known as transpososomes in DNA transposition, and invertasomes and intasomes for site-specific inversion and integration reactions, respectively, require the transposase or site-specific recombinase subunits to interact with each other and distant

recombination sites to mediate synaptic complex formation. Many specialized recombination systems employ host-encoded proteins as accessory factors to serve either an architectural role in facilitating the formation of active DNA/recombinase complexes or as regulatory proteins providing communication between the host cell and the recombination system to allow recombination to occur when cell physiological conditions are favorable. Some host-encoded proteins can play different types of roles in the specialized recombination reaction, and some proteins observed to influence the recombination reaction have undetermined roles (see below).

1.3. Host accessory factors involved in specialized DNA recombination

Host-encoded proteins associated with the bacterial nucleoid, including H-NS (histone-like nucleoid structuring protein), IHF (integration host factor), Fis (factor for inversion stimulation), and HU (heat-unstable nucleoid protein) are involved in condensing the bacterial chromosome and regulating a large variety of DNA transactions, including transcription and replication.[20] These proteins are able to bend or condense DNA *in vitro*,[20] and levels of these proteins vary under different physiological conditions.[21] The bending and/or regulatory functions of these proteins have been incorporated in a number of specialized DNA recombination reactions, as described below.

The global regulator H-NS binds DNA non-specifically, but prefers curved DNA or A/T rich sequences.[20] H-NS is required for efficient linear but not circular IS*1* transposition *in vivo*,[22,23] presumably by stabilizing the IS*1* transposase.[24] H-NS influences *in vivo* IS*1* and IS*903* target site selection,[23,25] and also affects *in vivo* Tn*10* transposition by affecting target capture.[25] Wardle *et al.*[26] demonstrated that H-NS directly interacts with the Tn*10* transpososome to promote intermolecular transposition (i.e. insertion into a target site other than itself) as opposed to affecting the transposase expression or the actual chemical steps in the strand transfer reaction.

In bacteriophage Mu transposition, H-NS affects the transposase (Mu A) expression by strengthening the repression mediated by the c protein. Thus in an *hns*$^-$ background, transposition of a mini-Mu construct is increased.[27]

IHF was first identified in being required for bacteriophage λ integration *in vitro*,[28] and binds a 30-35-bp sequence containing a consensus core element.[20] The λ *attP* site contains three IHF binding sites, and IHF serves an architectural role in the assembly of higher-order intasome complexes at the λ *att* site.[5] A number of transposable elements contain IHF binding sites at one or both termini, including γδ,[29] IS*1*,[30] Tn*10*,[31] and Tn*4652*.[32] Although the actual function of IHF binding to the IS*1* ends is not known,[30] IHF binding to γδ and Tn*4652* termini helps facilitates transposase binding[29,32] and modulates Tn*4652* transposase expression upon binding to a site upstream of the transposase transcription start site.[32,33] In Tn*10* transposition IHF, along with HU, stimulates *in vitro* circle formation.[34] IHF binding and bending of DNA at the Tn*10* outer end termini immediately adjacent to the primary transposase binding site plays a key role in the Tn*10* transpososome assembly and types of transposition products formed (intramolecular vs. intermolecular transposition).[35,36]

Fis binds a degenerate 15-bp consensus sequence,[20] and was identified as being required for efficient *in vitro* site-specific inversion mediated by the Gin (Mu) and Hin (Salmonella) recombinases.[37-39] Fis binds to two domains within the Hin enhancer sequence, a *cis*-acting sequence that strongly influences the inversion reaction[6] and the interaction between Fis and Hin subunits bound at the left and right recombination sites, *hixL* and *hixR* respectively, allows for a catalytically competent synaptic complex to form, which is stabilized by HU (see below).[40] Fis also stimulates the bacteriophage λ excision reaction *in vitro*[41] and *in vivo*,[42] and has been shown to stimulate *in vitro* recombination between *attB* and *attP* (the bacterial and phage attachment

sites, respectively) in reactions that resemble *in vivo* conditions (i.e. low integrase concentrations and low levels of DNA supercoiling).[43] Tn*5*/IS*50* contains a 15-bp Fis consensus binding sequence overlapping the 19-bp of the inner end sequence required for transposition, and Fis stimulates *in vivo* Tn*5*/IS*50* transposition over time.[44]

HU binds DNA non-specifically[20] and serves as an invertasome assembly factor in the Hin mediated inversion reaction.[45] HU is only required when the enhancer sequence is close to the left recombination site,[37,45] and participates in the Hin/Fis synaptic complex formation by stabilizing the DNA loop that occurs when the Hin/Fis subunits interact.[40] HU also plays an architectural role in the early steps of bacteriophage Mu transposition.[27-29] Binding of a single HU heterodimer to a 83-bp spacer region between the L1 and L2 Mu A binding sites is consistent with the DNA in this region looping out to allow the Mu A monomers to interact during the early steps in Mu transpososome formation.[46,47] HU is also implicated in playing a regulatory role in Mu transposition. HU can preferentially bind supercoiled DNA and is thought to serve as a "supercoiling sensor" to regulate Mu transposition as a function of level of DNA supercoiling within the cell.[48] Tn*916* excision is severely affected in strains lacking HU, and HU is thought to play in architectural role in Tn*916* transpososomes formation.[49]

The above nucleoid proteins are not the only host-encoded factors known to participate in specialized DNA recombination. DNA adenine methylase (Dam) regulates Tn*10*/IS*10* and Tn*5*/IS*50* transposition by methylating the adenines in the 5'-GATC-3' sequences within the transposase promoter region and element ends, decreasing transposase expression and availability of the element ends to transposase binding.[50,51] The methylation of Dam sites in Tn*10*/IS*10* and Tn*5*/IS*50* couples DNA transposition to DNA replication when the DNA is transiently present in a hemi-methylated form. DnaA, the initiator of chromosomal replication,[52]

binds to the outer end of Tn*5*[53] and stimulates Tn*5* transposition by an unknown mechanism.[54] The molecular chaperone ClpXP plays a dual role in facilitating Mu transposition. The ClpP protease promotes transition of Mu lysogenic growth to lytic growth by degrading the c repressor protein.[55] Once Mu is assembled into a transpososome/strand transfer complex, ClpX modifies the MuA transposase within the strand transfer complex to promote replication by host encoded proteins, resulting in cointegrate formation.[39,40] Interestingly, ClpP does not interfere with Mu replication, and even stimulates ClpX interaction with the Mu transposome.[56]

Target site selection and capture by Tn*7*, a bacterial transposon that displays insertion target site specificity[57] (see below), is stimulated by host ribosomal protein L29 and acyl carrier protein (ACP). These host proteins facilitate binding of the Tn*7*-encoded protein TnsD, a DNA-specific binding protein involved in Tn*7* target selection,[58] to the *att*Tn*7* site.[59]

1.4. Target selection and specificity

An important aspect of specialized recombination is the selection of an appropriate recombination site or insertion target mediated by the site-specific recombinase or transposase machinery, respectively. During site-specific recombination the recombinase recognizes a dyad symmetric site containing short regions of homology at the core between the recombining substrates. This specific sequence delineates the recombination substrates and facilitates the precise breaking and rejoining reactions that are hallmark to this type of specialized recombination.[4] Typically there is a lack of sequence homology between the ends of transposable elements and their target sites, and it was thought that transposable elements insert "randomly" into different chromosomal and episomal sites. Upon closer inspection of multiple target sites of different transposition systems (both *in vitro* and *in vivo*), insertion is not completely random and diverse factors can affect the use of a DNA site for insertion (see below).

Many of the characterized transposable elements display preferred target site selection. Consensus symmetrical target sequences have been described for *in vivo* and *in vitro* Tn*5*/IS*50* target sites (5'-A-GNTYWRANC-T-3' where N= any base, Y= T or C, W= A or T, R=A or G)[60] and known *in vivo* Tn*10* "hot-spot" insertion sites (5'-GCT-N-AGC-3').[61] There is some homology between the Tn*10* 6-bp consensus target sequence and Tn*10*/IS*10* ends (with more homology seen with the IS*10* ends), suggesting that one part of the transposase protein recognizes both the target sequence and element ends.[61] Analysis of independent IS*903* insertion sites into a conjugative plasmid did not reveal a common target sequence, but did reveal a symmetric consensus sequence (5'-T/A-T-T/A-Py-A-3', Py= pyramidine T or C) flanking many of the target sites.[62] These symmetrical sequences within or flanking the target sequence may provide symmetrical binding sites for transposase or form a unique structure making the insertion site more favorable for transposase binding, cleavage, and integration.[61,62] Some transposable elements, including IS*30*,[63] prefer to target their own terminal inverted repeats (IRs), creating an (IS*30*)2 junction that is very active in transposition[16] while other elements prefer to target the terminal IRs of another transposon including the *Bacillus thuringiensis* element IS*231*A.[64]

In addition to the actual target site used for insertion, sequences flanking the target can positively or negatively affect the preference of a transposase for an insertion site. Mutational analysis of the flanking sequence of the *Salmonella hisG*1 insertion hot-spot for Tn*10* demonstrated the flanking sequence along with the 6-bp symmetrical consensus target site play a major role in Tn*10* transposition.[65] Further analysis of flanking sequence in independent Tn*10* insertions did not reveal sequence-specific information, but may reflect the influence of the local DNA structure from base composition in target selection.[65] Ason and Reznikoff [66] showed that

the Tn*5* transposase favors A/T rich regions flanking the Tn*5* recognition/cleavage site, and in general A/T rich regions have a higher affinity for transposase binding than a G/C rich flanking region (with the implications that A/T double bonds are easier to separate compared to G/C triple bonds). While IS*30* and IS*231*A prefer to target terminal IRs of its own or another transposon, respectively, the sequence flanking the terminal IR can influence the use of that site for insertion.[63,64] Closer analysis of IS*231*A flanking sequence upstream of the proposed consensus sequence reveals alternate runs of A/T and G/C rich regions, which are likely to form an "S-shape" curve.[64] The authors suggest that the ability of a target region to have intrinsic bends or curves may serve a structural role in favoring transposase interactions with the flanking region and target site.[64] A role for DNA bending in target capture has also been described for Tn*10* transposition[67] and retroviral integration,[68] which is very similar to DNA transposition performed by DDE motif integrases.[69]

The ability of transposition proteins to access a potential insertion site also plays a role in target selection. For example, transcription and DNA topology can affect the target site's usage for insertion. Daniell *et al.*[70] first observed that high levels of transcription through the lactose operon resulted in fewer bacteriophage Mu insertions into *lacZ* as compared to when the operon was repressed. Subsequent experiments confirmed the inverse relationship between high transcription and bacteriophage Mu insertion, both at the operon and genome level.[71,72] Induction of the lactose operon was also shown to inhibit Tn*10* and Tn*5* insertion.[73] These observations suggest that the transposase can sense important coding regions of DNA via high transcription, and thus avoids insertion into essential genes.[73] Possible mechanisms for sensing the highly transcribed regions include the altered levels of supercoiling in transcribed regions of the chromosome[74] or more directly RNA polymerase could displace or occlude transposition

proteins from the target site.[72] DNA supercoiling has also been implicated in Tn*5* target selection. In addition to DNA gyrase being required for generating supercoiled targets for Tn*5* transposition,[75] Lodge and Berg[76] observed decreased Tn*5* insertions into its hot-spot I site on pBR322 when mutations reducing the high negative DNA supercoiling in pBR322 were introduced; leading to the proposal that regional melting of target DNA (i.e. hot-spot I) facilitated by high negative supercoiling may make it more accessible for recognition and/or cleavage during transposition.[76]

Site-specific target selection is usually not a common theme in DNA transposition; however some transposable elements display a marked preference for insertion target sites. One example is Tn*7* (Figure 1.6), an interesting transposable element that displays high frequency site and orientation specific insertion into a unique chromosomal sequence, *att*Tn*7*, found in *E. coli* and many different bacteria, and low frequency insertion into other non-related sites.[57] In addition to encoding transposase proteins (TnsA, TnsB), and the transposase activator protein (TnsC),[58] Tn*7* encodes two additional proteins (TnsD, TnsE) that influence target selection.[58,77-79] High frequency Tn*7* site-specific insertion into *att*Tn*7*[79] is mediated by the TnsABC+D pathway,[58] while the less-understood TnsABC+E pathway facilitates low frequency Tn*7* insertion into other non-related sites[79] and conjugative plasmids.[77]

*att*Tn*7* (Figure 1.6) in *E. coli* is located within the transcriptional terminator of *glmS*,[80,81] which is involved in cell-wall biosynthesis. *In vivo* and *in vitro* experiments have shown that the actual Tn*7* insertion site in *att*Tn*7* (i.e. the central position where the 5-bp duplication occurs as a result of Tn*7* insertion and designated position 0) is not required for *att*Tn*7* target activity, and target activity is actually dependent on specific binding of TnsD to a region upstream (+24 - +55) of the insertion site.[58,78,81,82] Bainton *et al.*[58] demonstrated *in vitro* that TnsC binding to

*att*Tn*7* is dependent on TnsD. Subsequent experiments indicate TnsC recognizes a DNA distortion in *att*Tn*7* caused by TnsD binding; this TnsC+D interaction with *att*Tn*7* covers a large region (+55 to -9), but does not include the actual insertion site.[82] Additional experiments performed by Kuduvalli *et al.*[82] suggest TnsD and TnsC interact with the major and minor grooves of *att*Tn*7*, respectively, and propose a model where TnsC binding to *att*Tn*7* in the minor groove would allow binding and activation of the transposase proteins TnsAB on the opposite side (major groove) of *att*Tn*7*.

Another transposable element, Tn*554*, also displays high-frequency site-specific insertion into a primary site within the *Staphylococcus aureus* chromosome.[83,84] Interesting features of Tn*554* are that it lacks terminal inverted repeats,[85] which typically delineate the ends of transposable elements,[9] it does not duplicate a target sequence upon insertion,[85] and its transposase is related to the lambda Int (Y-) family of site-specific recombinases.[86] In addition to Tn*554*, there are other transposable elements that have transposases more similar to either the Y- or S- site-specific recombinases, and have characteristics (e.g. target site specificity) comparable to site-specific recombination.[86] Insertion sequences in the IS*110*/IS*492* family are similar to elements like Tn*554* in that they do not cause target site duplication upon insertion, lack terminal inverted repeats, and show target site specificity.[87] However, this insertion sequence family is different from other characterized transposable elements since their transposases are members of a novel group of DNA recombinases that are not similar to the classical DDE motif transposases, or the Y- /S- site-specific recombinases. Interestingly, this novel recombinase family has members that perform DNA transposition and site-specific recombination (see below).

1.5. IS*110* family of insertion sequences

The IS*110* family of insertion sequences represents a unique group of elements that are grouped together based on sequence homology within their transposases and features that are not commonly found in other insertion sequence families. For example, the majority of members in the IS*110* family, with the exception of IS*492*[88] and IS*117*[89] do not create a target site duplication upon insertion (a hallmark of classical DNA transposition, see above) and do not have the terminal IRs that typically delineate the ends of the element and serve as binding sites for the transposase.[87] A subgroup of the IS*110* family, the IS*1111* subgroup, has members that possess subterminal 11-13-bp IRs.[90] Partridge and Hall[90] and Lauf *et al.*[91] have suggested that the IS*110* and IS*1111* subgroups be separated into two individual families based on the following features: (1) IS*1111* group members have 11-13-bp imperfect subterminal IRs; (2) target sites described for IS*1111* family members are different from target sites described for IS*110* group members; and (3) the IS*1111* subgroup transposases have more sequence homology to each other (>30% identity) compared to the IS*110* subgroup transposases (<20% identity). However, based on BLAST analysis the IS*110*/IS*1111* subgroups are not organized into two separate families.[87] The most interesting feature of the IS*110* family is that their transposases share high sequence homology with a novel site-specific recombinase, Piv, from *Moraxella bovis* and *Moraxella lacanata,*[92] and although little is known regarding the recombination mechanism of these IS elements, closer analysis of reports from IS*110*/IS*1111* group members suggests that these transposable elements may transpose via a site-specific recombination mechanism (see below).

Characterized members of the IS*110* subgroup include IS*900*, IS*901*, IS*902*, IS*1110* (*Mycobacteria* sp.), IS*110*, IS*116*, IS*117* (*Streptomyces* sp.), IS*1000* (*Thermus thermophilus*), and IS*492* (*Pseudoalteromonas atlantica*). Within this subgroup IS*900*, IS*901*, IS*902*, and IS*116*

are the most similar, with the sequences 5'-TCCT-3' (5'-TCCCT-3' for IS*116*) at the element left terminus, and 5'-CAT-3' 5-7 nucleotides (nt) from the right terminus of the element.[93-96] Comparison of the target sites for these elements reveals a consensus target sequence 5'-CAT$(N)_{6-9}$*CCT-3', (where * denotes the insertion site).[95] Insertion of the element into this insertion site positions the elements' left terminal 5'-CCT-3' sequence 5-9 nt downstream of the insertion target site sequence 5'-CAT-3' at the 5' element-insertion site junction (creating 5'-CAT$(N)_{5-9}$CCT-3') and positions the element right subterminal 5'-CAT-3' sequence 7-9 nt upstream of insertion target site sequence 5'-CCT-3' at the 3' element-insertion site junction (creating the sequence 5'- CAT$(N)_{7-9}$CCT-3'). These insertion events recreate the 5'-CAT$(N)_{6-9}$CCT-3' sequence found in the original insertion site, and may reflect a site-specific insertion mechanism if these elements excise to form a circular intermediate and aligns with the target site via short sequences of homology.[95,96] This type of insertion mechanism has also been proposed for IS*110*[97] and IS*117,*[98] in which a circular form of the element has been shown to be in intermediate in the transposition pathway.[98,99]

Some characteristics of the transposition pathway have been well-documented for the IS*110* subgroup member IS*117*. The 2.6 kilobase pair (kb) element (previously known as the 2.6-kb mini-circle) can be isolated from the *Streptomyces coelicolor* strain A3(2) as an extrachromosomal circular element and is integrated in the chromosome at two sites (copies A and B).[99] Although it is not found in the closely related *S. lividans* strain 66, the circular form of IS*117* was demonstrated to insert into the same chromosomal locus as the *S. coelicolor* strain A3(2) copy A locus,[99] providing the first evidence of a target site preference. Sequence analysis of the circular IS*117* element revealed the element's termini abutted with the sequence 5'-TAG-3' flanked by a 16-bp imperfect IR; this same sequence is flanked by an 11-bp imperfect IR in

the *S. lividans* strain 66 unoccupied preferred integration site.[98] IS*117* integration in *S. coelicolor* strain A3(2) and *S. lividans* strain 66 duplicates the target sequence 5'-TAG-3' at the chromosome/element insertion junctions.[98] The 5'-TAG-3' sequence present at the junction site in the circular IS*117* and at the *S. lividans* strain 66 target site and insertion junctions may indicate a type of site-specific insertion where the direct repeat is from integration and not DNA replication or repair as in classical transposition.[98] In addition, analysis of IS*117* chromosomal insertions in *Mycobacterium smegmatis* and into secondary sites (e.g. in the absence of the preferred target site) in *S. lividans* strain 66 revealed the sequence 5'-AG-3' conserved in all integration sites, suggesting that the A/G is where recombination takes place.[100,101]

Insertion sequences including IS*1111* (*Coxiella burnetti*), IS*1383* (*Pseudomonas putdia* strain H), IS*4321* (IncPβ plasmid R751) and IS*5075* (*Klebsiella pneumoniae*) are representatives of the IS*1111* subgroup of the IS*110* family. Characterized members of this subgroup, with the exception of IS*1111*, target terminal IRs of other transposable elements.[90,102,103] Analysis of IS*4321* and IS*5075* insertions within the terminal IRs of different Tn*501*/Tn*21* family members shows insertions of the elements in the same position and orientation within the 38-bp terminal IRs of different elements in this family, and IS*4321* excises precisely from its donor plasmid and inserts site-specifically into the terminal IRs of Tn*21*.[90] Orientation and position specific insertion of IS*1383* into the terminal IRs of IS*1384*, which is suggestive of site-specific insertion, has also been reported.[103]

Despite the above suggestions to separate the IS*1111* subgroup members into a different family, alignment of transposases from the IS*110*/IS*1111* subgroups shows that instead of having conserved DDE residues forming the catalytic motif found in classical transposases and retroviral integrases (see above), these transposases have conserved DEDD or DDDD

residues.[90,92] Interestingly, alignment of the site-specific recombinase Piv with these recombinases shows conservation of the same DEDD residues[92] as well as three amino acid motifs[104] found in the IS*110* family of recombinases. Piv and the transposases in the IS*110* family do not share any of the conserved amino acid motifs found in the Y- or S-recombinases.[105] Buchner *et al.*[106] demonstrated that the conserved DEDD residues in Piv are required for catalysis. Since site-specific recombination and DNA transposition involve two distinct catalytic mechanisms (see above) it is interesting that one family of DNA recombinases contain members that perform both types of reactions. For these reasons we hypothesize that Piv and transposases from the IS*110* family define a novel Piv/MooV family of DNA recombinases (DEDD motif recombinases). Therefore, it is important to further characterize recombinases in this family to determine if they share a common mechanism for DNA recombination, and to add new information to the field of DNA recombination. This dissertation research focuses on characterizing the MooV-mediated transposition of IS*492*, which controls extracellular polysaccharide (EPS) phase variation in *P. atlantica*, to ascertain if the transposases of the IS*110* family and Piv define a new family of DNA recombinases.

1.6. Previous characterization of IS*492* transposition

IS*492* was first discovered by Bartlett *et al.*[107] while investigating the genetic regulation of EPS synthesis in *P. atlantica*, a gram-negative marine bacterium,[108] which produces EPS to enable attachment and biofilm formation in the marine environment.[107] The *eps* phase variation observed in *P. atlantica* produces two distinct colony morphologies. EPS^+ colonies are large and mucoid while EPS^- colonies, indicated as crenated, have a smaller, wrinkled appearance (Figure 1.7).[107] Bartlett *et al.*[107] noted that EPS^- variants arose spontaneously at a frequency higher than expected for mutational events, and further analysis revealed that the addition and loss of 1.2-kb

of DNA from the *eps* locus was responsible for producing the two distinct EPS phenotypes. The 1.2-kb DNA sequence, later designated IS*492*, was isolated and further characterized. The 1,202-bp long mobile element lacks terminal IRs, and has a single open reading frame of 957-bp,[109] encoding the transposase MooV (Mover of IS*492* in oceanic Variants). Insertion of IS*492* results in a 5-bp target duplication, and excision of IS*492* and one of the target site duplications is precise.[109]

Genetic characterization of IS*492* within the *P. atlantica* genome indicated that there are five complete copies of IS*492* (including IS*492* in the *eps* locus) on the chromosome flanked by the same 5-bp left (5'-CTTGT-3') and 7-bp right (5'-CTTGTTA-3') sequences;[88] sequencing of the *P. atlantica* genome subsequently confirmed the five element copies and demonstrated an additional copy with the same flanking 5 and 7-bp sequences (ACK personal communication). The genome sequence also showed that the *eps* locus includes a predicted glucosyl transferase gene (*epsG*), which contains the target site for IS*492*. Further analysis of IS*492* excision in a naïve host, was accomplished by introducing IS*492* and ~130-bp of flanking sequence on a plasmid into *E. coli*; precise excision of IS*492* and circle formation were shown to be dependent on the transposase MooV, and require the 5-bp left and 7-bp right flanking *epsG* sequence (Perkins-Balding and Karls, unpublished results).[88] Perkins-Balding *et al.*[88] also demonstrated that the IS*492* circle junction sequence, which juxtaposes one of the 5-bp target duplication sequences with the ends of IS*492*, creates a strong promoter. This strong circle junction promoter may play a role in the up-regulation of MooV in a post-excision event increasing the likelihood of re-insertion, a mechanism proposed for IS*911*, which excises to form a circle intermediate.[110,111] Recently, Higgins *et al.*[112] have shown that in *P. atlantica* precise excision of IS*492* from *epsG* correlates with colony EPS phase variation, and host-initiated transcription

through IS*492* at *epsG* positively influences high-frequency IS*492* excision at this site. Further analysis suggests that the level of MooV production affected by outside transcription and not the actual process of transcription through the element affects high-frequency precise excision of IS*492*.[112]

IS*492* transposition may occur via an unusual pathway. As mentioned above, there are two characterized types of transposition reactions, non-replicative (or cut-and-paste), and replicative.[9] In the non-replicative pathway, the transposable element excises from the donor DNA and integrates into a target molecule. This is a simple insertion event, which leaves a double-strand break in the donor molecule. The replicative transposition pathway results in one copy of the transposable element in the donor and another copy in the target.[9] Southern analysis of IS*492* transposition in *P. atlantica* shows that precise excision of IS*492* from the *eps* locus results in loss of IS*492* from the *eps* site, but does not lead to insertion into a new chromosomal site.[107] However, insertion of IS*492* into the *eps* site does not correlate with a loss of IS*492* from another chromosomal site.[107] These results suggest transposition of IS*492* does not follow strictly a cut-and-paste or replicative pathway. Instead the transposition pathway may depend on the location of the IS element.

With the increased knowledge regarding IS*492* transposition there remain unknown determinants concerning the IS*492* transposition pathway. Although IS*492* excision has been characterized in *P. atlantica* and *E. coli*, insertion of IS*492* has not been previously observed in *E. coli*, suggesting requirements for *P. atlantica* regulatory or accessory recombination factors.[113] In addition, the IS*492* transposition pathway (e.g. cut-and-paste vs. replicative), preferred target sites for IS*492* insertion, and roles of the IS*492* excised circle/circle junction sequence in transposition also require further investigation. To address these questions genetic

assays were developed to study MooV-mediated IS*492* insertion in *E. coli*, and results in this dissertation research demonstrate that insertion of IS*492* in *E. coli* does not require *P. atlantica* specific protein-factors and displays target-site specificity. Analysis of MooV-mediated IS*492* insertion products in *E. coli* suggests that MooV utilizes a cut-and-paste transposition pathway, or a conservative site-specific mechanism. These and other observations described in this research suggests MooV-mediated IS*492* insertion in *E. coli* is more similar to site-specific recombination compared to classical DNA transposition, adding more evidence that recombinases in the proposed Piv/MooV family define a novel family.

Table 1.1 Summary of Specialized DNA Recombination

Characteristic	Classical DNA Transposition	Site-Specific Recombination
Catalytic Residues	DDE	S (resolvase/invertase family) Y (integrase family)
Initial Attacking Nucleophile	H_2O	OH (S/Y)
Chemical Mechanism	Hydrolysis and 1-step transesterification	2-step transesterification
Covalent DNA-Protein Linkages	No	Yes 5'Phospho-S (resolvases/invertases) 3'Phospho-Y (integrases)
Involvement of DNA Replication	Yes	No
Requirement of Sequence Homology	No	Yes

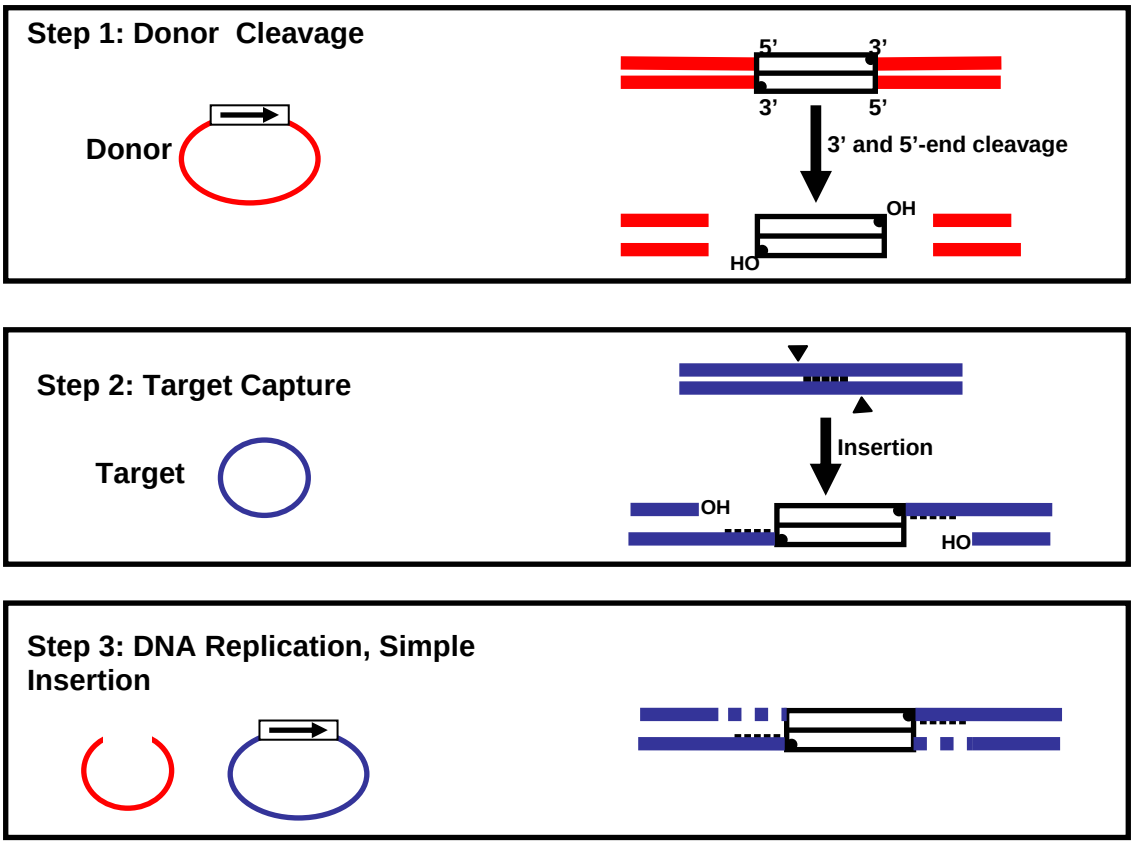


Figure 1.1 Cut-and-paste or non-replicative transposition. Step 1: The transposase coordinates a H_2O molecule (not shown) as the attacking nucleophile to cleave the 3'-ends of the transposable element (●). The 5'-ends of the element are also cleaved, releasing the element from the donor backbone. Step 2: The 3' –OH groups of the transposable element attack a 5-bp target sequence (represented by five black bars) in a staggered manner (▲▼), resulting in insertion and liberation of 3'-OH groups on the target DNA. Step 3: The free 3'-OH groups from the target DNA are used by host cell replication machinery to repair gaps created in Step 2. DNA replication causes duplication of the original 5-bp target sequence. The gapped donor molecule can be repaired from a sister chromosome via gap repair or is lost from the host cell.

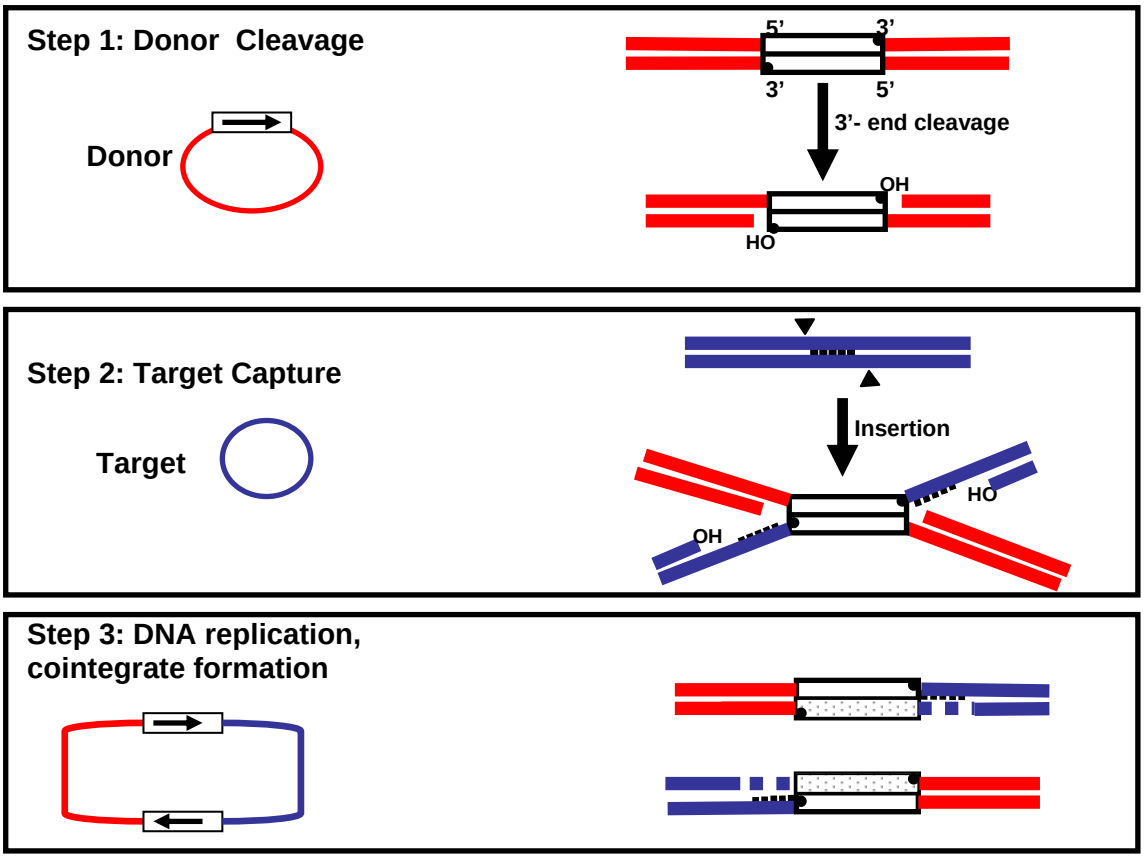


Figure 1.2 Replicative transposition. Symbols/representations are the same as in Figure 1. Step 1: The transposase coordinates a H_2O molecule (not shown) as the attacking nucleophile to cleave the 3'-ends of the transposable element. The 5'-ends of the element are not cleaved. Step 2: The 3' –OH groups of the element attack a 5-bp target sequence in a staggered manner, resulting in the element being linked to the donor and target backbone. Step 3: The free 3'-OH groups in the target DNA are used by the host cell replication machinery to repair gaps created in Step 2, and the 5-bp target sequence and transposable element are duplicated as a consequence of DNA replication. The end product creates a fusion, or cointegrate, between the donor and target molecules, with the duplicated transposable element, in direct repeats, at the donor/target molecule junctions.

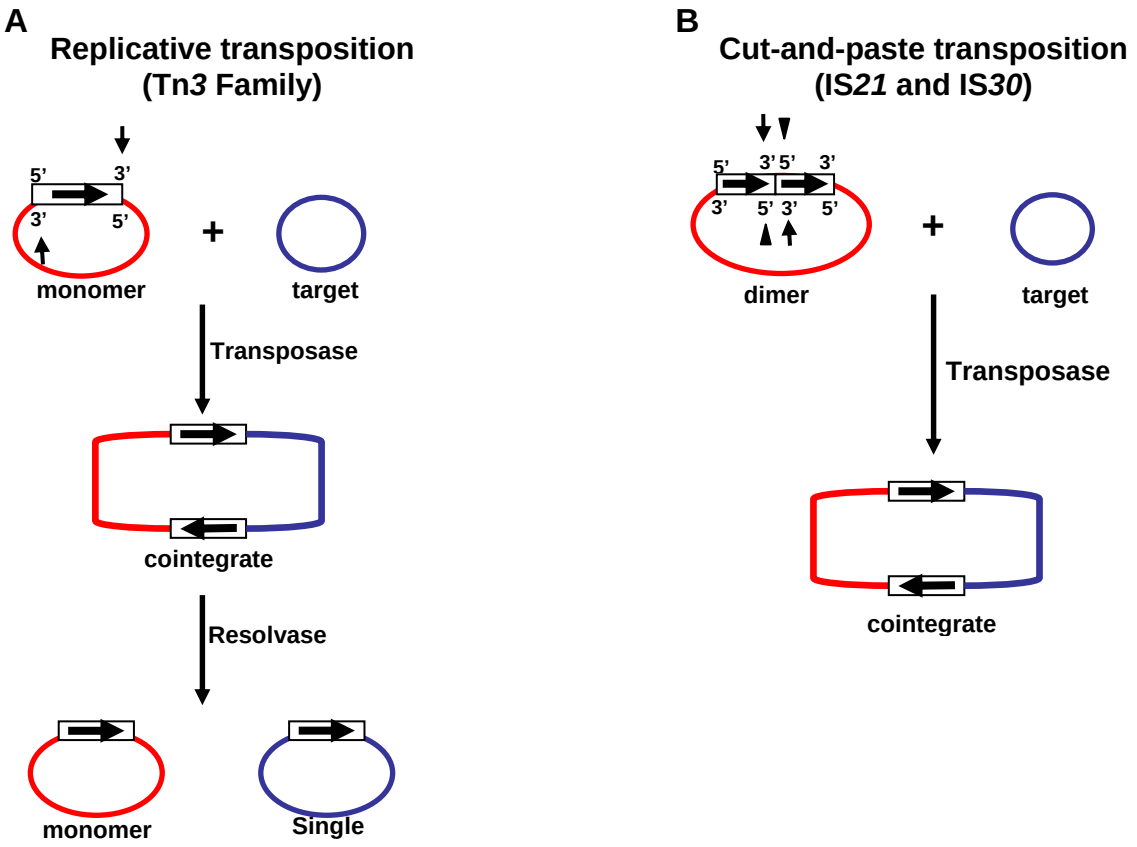


Figure 1.3 *In vivo* end products of DNA transposition. Panel A: Simple insertion from replicative transposition. Transposases from the Tn*3* family cleave the 3'-ends of the transposon (black arrows) and follow a replicative pathway of insertion (See Figure 2) resulting in a cointegrate. An element-encoded resolvase mediates site-specific recombination between the two transposons. This site-specific recombination resolves the cointegrate into a molecule resembling the starting transposon donor (blue) and simple insertion into the target DNA (red). Panel B: Cointegrate formation from cut-and-paste transposition. Insertion elements including IS*21* and IS*30* target their own ends via a cut-and-paste mechanism (Figure 1), forming a head-to-tail dimer insertion. The elements' transposases cleave the 3' and 5'-ends (black arrows and black triangles, respectively) in the inner junction of the head-to-tail dimer, and use a cut-and-paste

pathway to insert the dimer donor molecule into a target DNA, resulting in cointegrate formation.

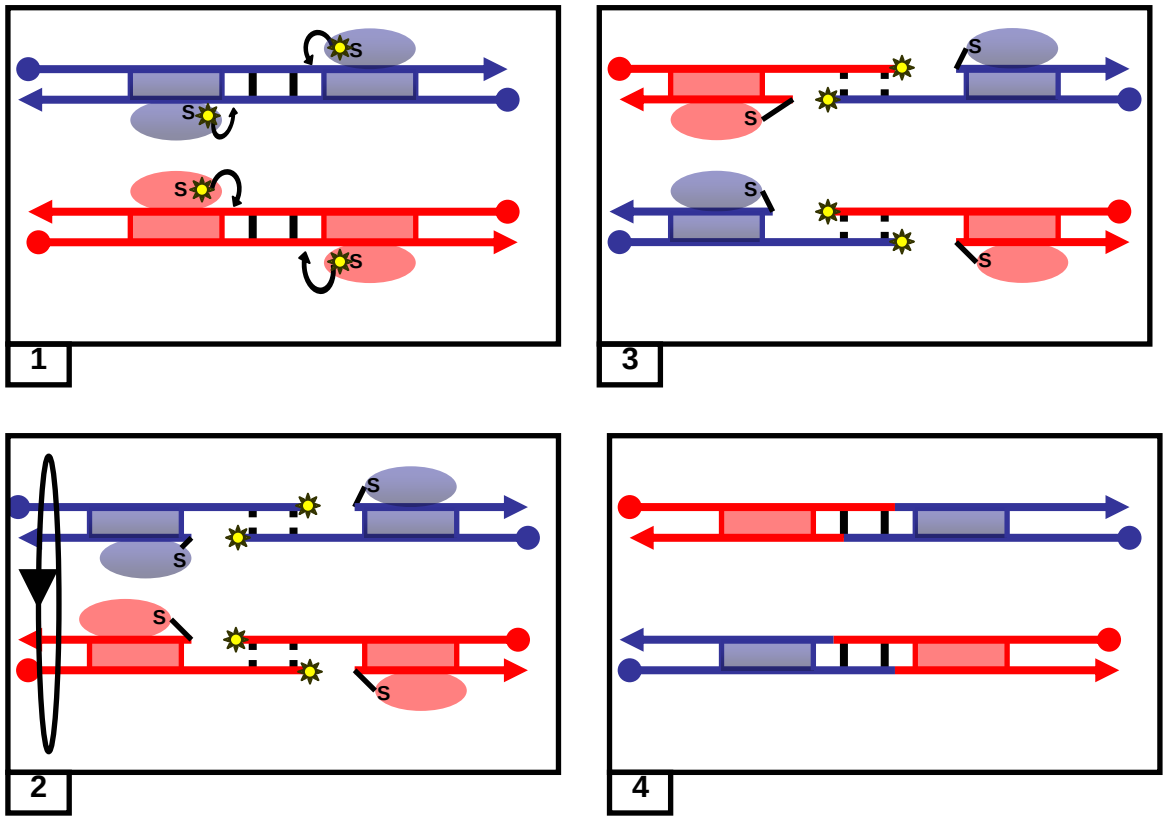


Figure 1.4. Site-specific recombination via Serine recombinases. Polarity of DNA is represented by colored ovals (5'-end) and arrowheads (3'-end). Black bars represent DNA base pairs. Box 1: Four recombinase subunits bind the inverted repeats (colored boxes) within the functional core of the recombining partners. During the 1st transesterification reaction the –OH groups (yellow stars) of the catalytic serine residues (S) attack the DNA backbone (black arrows) outside of the 2-bp overlap region in a concerted manner, resulting in 5'phospho-serine linkages and 3'overhangs (double strand breaks). Box 2: Rotation of one-half of the recombining substrates (circle with arrow) results in inversion. Box 3: In the 2nd transesterification reaction the free 3'-OH's (yellow stars) attack the neighboring 5'phospho-serine linkages, ligating the recombinant partners (Box 4).

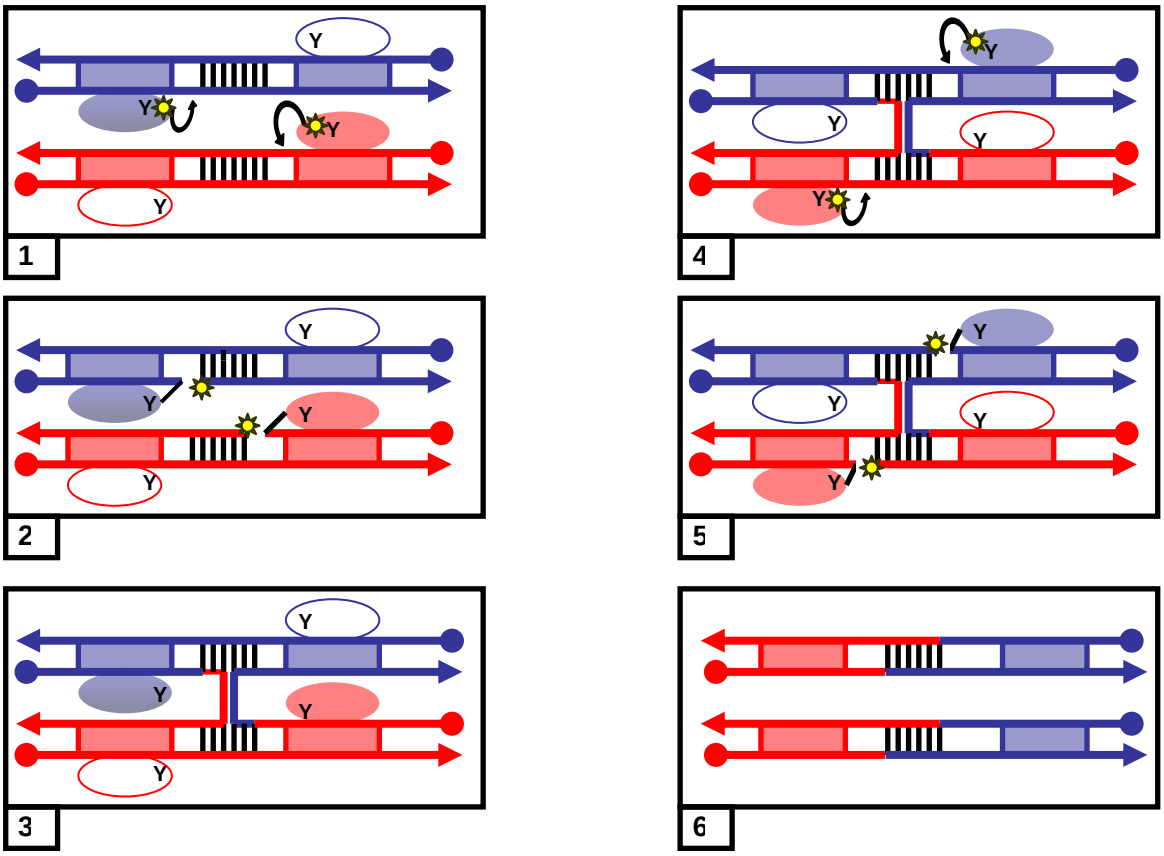


Figure 1.5 Site-specific recombination via Tyrosine recombinases. Symbols and representations are the same as in Figure 1.4. Box 1: Four recombinase subunits bind the inverted repeats (colored boxes) within the functional core of the recombining partners. The recombinase subunits on the bottom strands are activated (colored subunits) to perform the 1st transesterification reaction. The –OH groups (yellow stars) of the catalytic tyrosine residues (Y) attack the DNA backbone (black arrows) outside of the 6-bp overlap region, resulting in 3'phospho-tyrosoine linkages and 5'overhangs. Box 2: The free 5'-OH groups attack the 3'phospho-tyrosine linkages on the opposite strands, creating a Holliday junction intermediate (Box 3). Box 4: Isomerization of the Holliday junction intermediate (not shown) activates the recombinase subunits on the top strands to perform the 2nd transesterification reaction. Box 5:

The free 5’-OH groups attack the 3’phospho-tyrosine linkages on the opposite strands, resolving the Holliday junction intermediate. The end result is two recombined DNA strands.

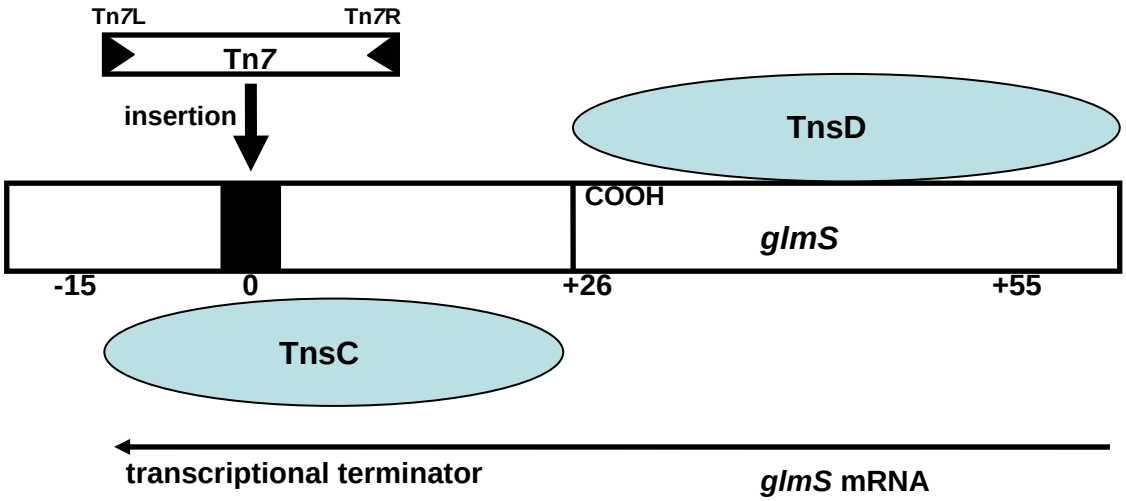


Figure 1.6 Tn*7* insertion into *att*Tn*7*. The Tn7 insertion site is designated 0, bases to the right are represented with a plus sign, and bases to the left are represented with a minus sign. The site and orientation specific insertion of Tn*7* into *att*Tn*7* (with the right end of Tn*7* inserting proximal to the C-terminal end of *glmS*) into position 0, occurs in the transcriptional terminator of the *glmS* gene. The TnsD protein interacts with sequences +25 - +55 in the C-terminal region of *glmS*, and causes a DNA distortion near position +27 (not shown) that recruits TnsC. TnsC+TnsD interaction with *att*Tn*7* covers sequences -15 - +55. Figure not drawn to scale.

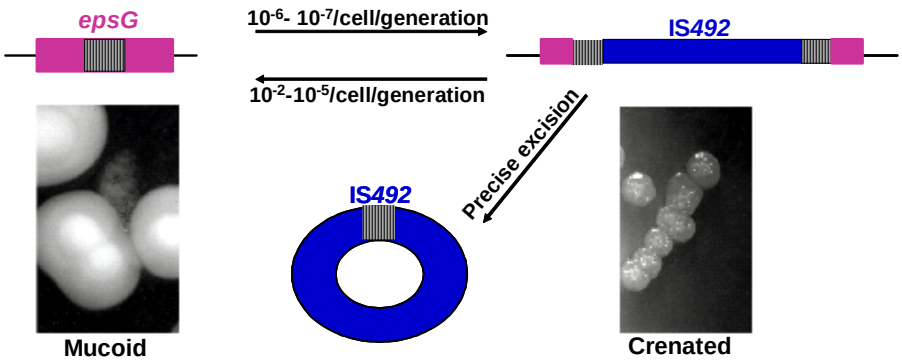


Figure 1.7 IS*492* transposition in *Pseudoalteromonas atlantica*. *P. atlantica* produces biofilms in the marine environment by expressing **e**xtracellular **p**oly**s**accharide (**EPS**). *P. atlantica* colonies expressing EPS have a mucoid phenotype on solid media. EPS production is inactivated via the site-specific insertion of IS*492* into a 5-bp sequence 5'-CTTGT-3', represented as a black and white box, in *epsG* (a predicted glycosyl transferase gene). IS*492* insertion duplicates this sequence, and *P. atlantica* colonies have a crenated phenotype on solid media. EPS expression is restored when IS*492* precisely excises to form a circular element, where the ends of IS*492* are juxtaposed by one of the 5-bp sequences.

CHAPTER 2

MooV-mediated IS*492* insertion in *Escherichia coli* is site-specific

Abstract

IS*492* controls the phase variation of extracellular polysaccharide (EPS) in the marine bacterium *Pseudoalteromonas atlantica*. The reversible insertion of IS*492* at a single site in *epsG* is mediated by the DEDD recombinase, MooV. The inserted element is flanked by direct repeats of the sequence 5'-CTTGT-3' and precise excision produces an IS*492* circle with the element ends joined by this 5-bp sequence. In experiments designed to identify factors required for IS*492* insertion in *E. coli*, transposition of IS*492*Δ*mooV*::*tet* (complemented by MooV *in trans*) did not require recombination host factors from *P. atlantica*, and was facilitated by using temperatures close to the growth optimum of *P. atlantica*. Although temperature affects insertion of IS*492*Δ*mooV*::*tet*, MooV is expressed at similar levels at 37°C and 25°C. Insertion of IS*492*Δ*mooV*::*tet* yielded single insertion into the *epsG* target plasmid, and tandem head-to-tail insertions into the donor plasmid; chromosomal insertions were not observed, suggesting a target site preference. Single insertion into the *epsG* site targeted the 5-bp 5'-CTTGT-3' sequence. The tandem insertion, (IS*492*Δ*mooV*::*tet*)2, is separated and flanked by this 5-bp sequence, suggesting that the mechanism for insertion is either site-specific or cut-and-paste where the 5-bp sequence is the common core of the recombination sites or the recognition sequence for the staggered cut by MooV, respectively. To characterize further the IS*492* transposition mechanism, the (IS*492*Δ*mooV*::*tet*)2 plasmid was used in insertion assays with a plasmid encoding the *epsG* target sequence. A cut-and-paste pathway would yield either a single IS*492*Δ*mooV*::*tet* insertion

into the target plasmid or, if the adjoining ends of (IS*492*Δ*mooV*::*tet*)2 are the active site for MooV, a cointegrate molecule with the IS elements joining the donor and target plasmids. A replicative pathway would yield a cointegrate plasmid with three IS*492*Δ*mooV*::*tet* elements. Although precise excision of one IS*492*Δ*mooV*::*tet* from the donor plasmid was observed, the most frequent insertion products detected in the assays were cointegrates that had two directly-oriented IS*492*Δ*mooV*::*tet* elements flanking the target and donor plasmids with 5'-CTTGT-3' at the junctions. MooV-mediated transposition of (IS*492*Δ*mooV*::*tet*)2 is less sensitive to temperature compared to IS*492*Δ*mooV*::*tet*. These results suggest that the inner ends of (IS*492*Δ*mooV*::*tet*)2 formed the active site for targeted insertion into the *eps* site by either a site-specific or cut-and-paste mechanism, and that the formation of the junction sequence is the rate-limiting step in MooV-mediated IS*492* transposition.

2.1. Introduction

The insertion sequence IS*492* controls the phase variation of extracellular polysaccharide (EPS) production in the biofilm-forming marine γ-proteobacterium *Pseudoalteromonas atlantica* via reversible excision from a predicted glucosyl transferase gene (*epsG*).[88,107,112] IS*492*, a member of the atypical IS*110* family of insertion sequences,[87] does not have terminal inverted repeats, but has a 5-base pair (bp) direct repeat flanking the inserted element in *P. atlantica*.[88,109] Previous characterization of IS*492* transposition suggests it may use an unusual pathway. High-frequency precise excision of IS*492* from a single site in *epsG* in *P. atlantica* results in an excised IS*492* circle, but does not correlate with re-insertion into another chromosomal site (Higgins *et al.* unpublished data).[88,107,112] However, IS*492* insertion into the *epsG* locus does not correlate with a loss of IS*492* from another chromosomal site[107] (Higgins *et al.* unpublished data), suggesting IS*492* transposition does not strictly follow a cut-and-paste or replicative

pathway[9] and the transposition pathway may depend on the location of the IS element. The characterized chromosomal copies of IS*492* in *P. atlantica* (EPS$^-$ phenotype) are flanked by the same left 5-bp sequence 5'-CTTGT-3' and right 7-bp sequence 5'-CTTGTTA-3'[88] suggesting a high-degree of target site-specificity.

Transposases of the IS*110* family, including the IS*492* transposase MooV, and the site-specific DNA invertase Piv from *Moraxella lacunata* and *Moraxella bovis* belong to the novel DEDD motif family of DNA recombinases.[92,106] These recombinases do not share the conserved amino acid motifs found in the Tyrosine (Y)- and Serine (S)- families of site-specific recombinases or the DDE motif family of transposases.[114] However the conserved acidic tetrad of the DEDD motif recombinases lies within a predicted RNaseH fold structural component like the catalytic triad (DDE) of classical transposases and retroviral integrases[114] and the catalytic tetrad (DEDD) of RuvC-like Holiday junction resolvases.[106] Since site-specific recombinases and classical DDE-motif transposases utilize two distinct catalytic mechanisms[115] it is atypical for one family of DNA recombinases to contain members that perform both types of reactions. For these reasons we propose that DEDD-motif recombinases, including the transposases from the IS*110* family, define a new family of DNA recombinases. The work reported herein focuses on the characterization of MooV-mediated transposition of IS*492* which provides insight into the catalytic mechanisms of this family.

IS*492* precise excision and circle formation in the naïve host *E. coli* is MooV-dependent, and requires the 5-bp left and 7-bp right flanking *eps* target sequence[88] (Perkins-Balding and Karls, unpublished data). In this work we have focused on characterizing the IS*492* insertion mechanism in *E. coli*. We demonstrate that an IS*492* construct, IS*492*Δ*mooV*::*tet*, exhibits MooV-mediated insertion in *E. coli* in the absence of *P. atlantica* specific recombination factors.

MooV-mediated insertion displays target site specificity, preferring a single site within the *epsG* sequence on a target plasmid or one end of the IS*492*Δ*mooV*::*tet* element on the donor plasmid to create a tandem head-to-tail insertion. Tandem IS*492*Δ*mooV*::*tet* insertions, [(IS*492*Δ*mooV*::*tet*)2], are very active in transposition in the presence of MooV, and analysis of (IS*492*Δ*mooV*::*tet*)2 transposition products, including single excision and cointegrate (fusion) products, has shown that MooV meditates precise excision and site-specific insertion of (IS*492*Δ*mooV*::*tet*)2. These observations are atypical for insertion elements from other families, and suggest that MooV uses a unique recombination pathway that combines aspects of both DNA transposition and site-specific recombination. The role of temperature in IS*492* insertion in *E. coli* is also examined.

2.2. Materials and methods

Bacterial strains. *E. coli* strain VCS257 [F$^-$ *tonA53 dapD8 lacY1 glnV44 supE44* Δ(*gal-uvrB*)*47 tyrT58 supF58 gyrA29* Δ(*thyA57*) *hsdS3*] (obtained from D. Bartlett) was used to maintain the *P. atlantica* cosmid library.[107] *E. coli* strains BW23473 [Δ*lac-169 robA1 creC510 hsdR514 uidA* (Δ*MluI*)::*pir endA* (*BT333*) *recA1*] (obtained from E. Stabb) and BW23474 [Δ*lac-169 robA1 creC510 hsdR514 uidA* (Δ*MluI*)::*pir-116 endA* (*BT333*) *recA1*][116] were used to maintain R6k *ori* plasmids in complementation experiments. *E. coli* strains DH5α [F- θ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deoR supE44 hsdR17 recA1 end*A*1*] (obtained from C. Moran) and TOP10 [F$^-$ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) θ80d*lac*ZΔM15 Δ*lacX74 deoR recA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL*(streptomycin resistance, Sm^r) *endA1 nupG*] (Invitrogen) were used as recipient strains in complementation experiments, as host strains in no-cosmid insertion assays (DH5α), and for plasmid maintenance.

Plasmid constructions. *Pfu* DNA polymerase (Stratagene) was used to amplify PCR products generated for cloning, and all *Pfu* amplified DNA was sequenced. All DNA enzymes, restriction enzymes, and ligases used in construction of plasmids were obtained from New England Biolabs (NEB). All oligonucleotides used in this study are listed in Table 2.1. The IS*492* derivative IS*492*Δ*mooV*::*tet* was constructed to study IS*492* transposition *in vivo*. The 1,255-bp fragment containing the promoter-less (with its ribosome binding site) tetracycline resistance gene (*tet*) was PCR amplified from pBR322 (NEB) using primers MfeTetI and NewBsmBI and ligated with T4 DNA Ligase into the MfeI-BsmBI region of IS*492* on pAG949,[88] creating pAMC1. IS*492*Δ*mooV*::*tet* was PCR amplified from pAMC1 using primers EPSL.DraIII and EPSR.PsiI and ligated into the DraIII-PsiI sites of pBADMooV.1 [contains *mooV* amplified with primers BADMooV.1 and BADMooV.2 and cloned into the XbaI and HindIII sites in pBAD33],[117] creating pAMC3. A 799-bp Acc65I-BsrGI fragment (containing 777-bp of *mooV* coding sequence) was removed from pAMC3 to create pAMC7 (see Figure 2.1). A tandem head-to-tail IS*492*Δ*mooV*::*tet* insertion into pAMC3 was isolated from *in vivo* insertion assays and designated pAMC10. The 799-bp Acc65I-BsrGI fragment (containing 777-bp of *mooV* coding sequence) was removed from pAMC10 to create pAMC13 (Figure 2.1).

The *eps* target plasmid, pAMC5, was created by PCR amplifying a 134-bp region of *epsG* from mucoid (EPS^+) *Pseudoalteromonas atlantica* DB27[107] colonies via colony PCR[88] using primers EPSL.58 and EPSR.76. The 134-bp PCR product was gel-extracted using the QIAquick Gel Extraction Kit (Qiagen), and 3' adenines were added to facilitate cloning into the TA cloning vector pCR2.1 (Invitrogen) following the manufacturer's instructions. Transformants were screened to identify pCR2.1 clones containing the *eps* target sequence in correct reading frame orientation relative to the *lac* promoter. The kanamycin resistance gene in pAMC5 was

knocked out via digestion with NcoI, followed by performing a fill-in reaction using the End-it DNA End Repair Kit (Epicentre) and ligation with T4 DNA Ligase.

PCR assay to detect IS*492*Δ*mooV*::*tet* circles. *E. coli* strain TOP10 was transformed with pAMC1, or co-transformed with pAMC1 and the MooV expression vector pAG900[88] via electroporation using the Bio-Rad micropulser electroporator with 0.2 cm cuvettes (Bio-Rad), outgrown 1 h at 37°C, and plated onto Luria-Bertani (LB) agar containing 100 μg/ml ampicillin (Ap_{100}), and LB agar containing Ap_{100}, 50 μg/ml spectinomycin (Sp_{50}), with 10 μM Isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce *mooV* expression from P_{tac}, respectively. Additionally, TOP10 was transformed with pAMC3, or co-transformed with pAMC3 and pAG900 as described above, and plated onto LB agar containing 34 μg/ml chloramphenicol (Cm_{34}) and LB/$Cm_{34}Sp_{50}$ + 10 μM IPTG, respectively. Following overnight incubation at 37°C and 4 h incubation at 4°C, colony PCR using primers CJ250A and ISR.51 was performed[88] using five pooled colonies resuspended in 30 μl per PCR template. Template control PCR was also performed using primers PCRAP.1 and PCRAP.2 (specific to the β-lactamase gene of pAMC1) or CamF and CamR (specific to the chloramphenicol acetyl-transferase gene of pAMC3). Reactions were done in 1X *Taq* polymerase Buffer B (Promega), 2.5 mM $MgCl_2$, 10 μM each dNTPs, and 1 unit *Taq* polymerase using the following thermal cycling conditions: 95°C for 2 min, followed by 30 cycles of 95°C for 45 seconds, 58°C for 45 seconds, and 72°C for 20 seconds. Products were detected on a 2% agarose gel (1% Seakem and 1%NuSieve), stained with ethidium bromide (EtBr), and visualized with UV.

Complementation assay to determine *P. atlantica* factors for IS*492*Δ*mooV*::*tet* insertion in *E. coli*. TOP10 or DH5α cells containing pAMC3 and pAMC5 (TOP10/pAMC3 pAMC5 or DH5α/pAMC3 pAMC5) was transformed with pAG900 via electroporation as described above,

outgrown 1 h at 37°C, and plated onto LB/$Cm_{34}Ap_{100}Sp_{50}$. Following overnight incubation at 37°C, transformant colonies were used to inoculate 2 ml LB/$Cm_{34}Ap_{100}Sp_{50}$ cultures, and 2 ml LB cultures supplemented with 50 μg/ml kanamycin (Kn_{50}) were started for the *P. atlantica* cosmid library donor strain VCS257/pMMB33::cosmid library (*mob*$^+$ *tra*$^-$)[107] and conjugation helper strain BW23473/pEVS104 (*mob*$^+$ *tra*$^+$).[116] Overnight cultures were harvested via centrifugation and cells were resuspended in 500 μl LB. 100 μl of each strain (donor, helper, and recipient) were mixed together for a total 300 μl conjugation mix, and spotted onto LB agar. LB mating plates were incubated overnight at 37°C or 48 h at 32°C. Cell growth was harvested using a sterile wooden stick and resuspended in 500 μl LB. Cells were diluted and plated on LB/$Cm_{34}Ap_{100}Sp_{50}$ (total number of recipients), LB/$Cm_{34}Ap_{100}Sp_{50}Kn_{50}$ and LB/$Ap_{100}Kn_{50}$ to select for total number transconjugant strains. To select for IS*492*Δ*mooV*::*tet* insertion events, cells were plated onto LB/$Ap_{100}Kn_{50}$ and 15 μg/ml tetracycline (Tet_{15}), with or without 10 μM IPTG or 0.2% glucose to induce *mooV* expression from P_{tac} on pAG900 and to decrease transcription from the P_{lac} across the *eps* target on pCR2.1, respectively. Plates were incubated overnight at 37°C or 48 h at 32°C, and the relative frequency of IS*492*Δ*mooV*::*tet* insertion was calculated by the total number of transconjugants resistant to ampicillin (Ap^r), kanamycin (Kn^r), and tetracycline (Tet^r) divided by the total number of Ap^r Kn^r transconjugants. To screen for IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid, plasmid DNA was isolated from Ap^r Kn^r Tet^r transconjugants using the Wizard® *Plus* Miniprep kit (Promega) and transformed into naïve TOP10 cells. Transformants were plated onto LB/$Ap_{100}Tet_{15}$ to select for *eps* target plasmids containing an IS*492*Δ*mooV*::*tet* insertion. Plasmids were isolated from Ap^r Tet^r transformants and analyzed via restriction and insertion PCR analysis (see below) to confirm IS*492*Δ*mooV*::*tet* insertion. Cosmids facilitating IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid were mated

out of Ap^r Kn^r Tet^r transconjugants as described above (all incubation temperatures performed at 37°C) into DH5α. DH5α recipients receiving the cosmid were selected with Kn_{50} and 30μg/ml nalidixic acid (Nal_{30}), and were used as the donor strain in subsequent complementation assays as described above.

IS*492*Δ*mooV*::*tet* insertion assays in *E. coli.* Protocol I: DH5α cells containing pAMC3 and pAMC5 were transformed with pAG900 via electroporation, outgrown for 1 h at 37°C, plated on LB/$Cm_{34}Ap_{80}Sp_{50}$ for plasmid maintenance, and incubated overnight at 37°C. Transformant colonies were inoculated into 2 ml LB/$Cm_{34}Ap_{80}Sp_{50}$ cultures and incubated with aeration overnight at 37°C. Cultures were diluted and plated onto LB/$Cm_{34}Ap_{80}Sp_{50}$, LB/$Ap_{80}Sp_{50}$, and LB/$Ap_{80}Sp_{50}Tet_{15}$ (with or without 10 μM IPTG or 0.2% glucose), and incubated overnight at 37°C or 48 h at 32°C. Relative transposition frequencies were calculated by dividing the total number of Ap^r Sp^r Tet^r colonies by the total number of Ap^r Sp^r colonies. Plasmid DNA was isolated from Ap^r Sp^r Tet^r colonies using the Wizard® *Plus* Miniprep kit (Promega), digested with HindIII and subjected to insertion PCR analysis to characterize IS*492*Δ*mooV*::*tet* insertion products. **Protocol II:** DH5α/pAMC3 pAMC5 was transformed with pAG900 as described for Protocol I, with a transformation outgrowth of 1 h at 37°C. 100 μl of transformed cells were inoculated into 2 ml LB/$Cm_{34}Ap_{80}Sp_{50}$ cultures, and incubated with aeration overnight at 37°C. Cultures were diluted and plated on selective media for insertion as described in Protocol I and incubated 72 h at 25°C. Transposition frequencies and analysis of Tet^r colonies were performed as described for Protocol I. **Protocol III:** DH5α/pAMC3 pAMC5 was transformed with pAG900 as described for Protocol I, except transformed cells were outgrown at either 32°C or 25°C for 6 h. 100 μl of transformed cells were inoculated into 2 ml LB/$Cm_{34}Ap_{80}Sp_{50}$ cultures, and incubated with aeration at 32°C (48 h) or 25°C (72 h). Cultures were diluted and plated on

selective media for insertion as described in Protocol I and incubated 48 h at 32°C or 72 h at 25°C. Transposition frequencies and analysis of Tet^r colonies were performed as described for Protocol I.

Insertion PCR assay to detect single IS*492*Δ*mooV*::*tet* into the *eps* target plasmid. Plasmids isolated from Ap^r Tet^r isolates were used as template in PCR assays using primers M13R + ISL.66 (214-bp product) or M13R + CJ250A (288-bp product) to detect the left IS*492*Δ*mooV*::*tet*/*eps* insertion junction or primers M13F + ISR.51 (243-bp product) to detect the right IS*492*Δ*mooV*::*tet*/*eps* insertion junction. Reactions were done in 1X *Taq* polymerase Buffer B (Promega), 2.5 mM $MgCl_2$, 10 μM each dNTPs, and 1 unit *Taq* polymerase using the following thermal cycling conditions: 95°C for 2 min, followed by 30 cycles of 95°C for 45 seconds, 58°C for 45 seconds, and 72°C for 20 seconds. Products were detected on a 2% agarose gel (1% GenePure LE and 1% GenePure 3:1), and cloned into pCR2.1 (Invitrogen) for subsequent sequence analysis at the University of Michigan Sequencing Core Facilities [(UM-BRCF)].

Inverse PCR to characterize IS*492*Δ*mooV*::*tet* chromosomal insertions. Chromosomal DNA was isolated from selected Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates using the Bio-Rad AquaPure Genome Kit, and 0.7 μg of DNA was digested with PstI, SacI, or AatII (contains no recognition sites in IS*492*Δ*mooV*::*tet*) and treated with T4 DNA Ligase to promote intramolecular ligation as previously described.[88] Inverse PCR[118] was performed using primers CJ250A and ISR.55, which anneal to IS*492*Δ*mooV*::*tet* with their 3'-ends directed outwardly to amplify DNA flanking IS*492*Δ*mooV*::*tet*. To increase product specificity, 1 μl of PCR DNA from the inverse PCR assays was used as template in nested PCR assays using primers EpsPrimer.1 and EpsPrimer.2. Reactions were done in 1X *Taq* polymerase Buffer B (Promega),

2.5 mM $MgCl_2$, 10 μM each dNTPs, and 1 unit *Taq* polymerase using the following thermal cycling conditions: 95°C for 2 min, followed by 30 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 5 min. Products were detected on a 0.8% agarose gel, and cloned into pCR2.1 (Invitrogen) for subsequent sequence analysis at UM-BRCF.

Temperature sensitive expression of Tet^r from IS*492*Δ*mooV*::*tet*. DH5α was transformed with pAMC3, pAMC7, or pBAD33 via electroporation as described above and outgrown 1 h at 37°C. Transformed cells were plated onto LB/Cm_{34} and LB/$Cm_{34}Tet_{15}$, and incubated overnight at 37°C or 48 h at 25°C. Frequency of Tet^r was calculated by dividing the total number of chloramphenicol resistant (Cm^r)Tet^r transformants by the total number of Cm^r transformants.

(IS*492*Δ*mooV*::*tet*)2 insertion assays in *E. coli*. Chemically competent DH5α/pAMC13 pAMC5 was transformed with either pAG900 or pAG900Δ*mooV* (obtained from C. Carpenter) and outgrown 1 h at 37°C or 6 h at 25°C. Transformants were plated on LB/$Cm_{34}Ap_{80}Sp_{50}$ with 50 μM IPTG. After overnight or 72 h incubation at 37°C or 25°C, respectively, individual transformant colonies were inoculated into LB broth/$Cm_{34}Ap_{80}Sp_{50}$ and incubated with aeration overnight at 37°C. Plasmid DNA was isolated from overnight cultures using the Fast Plasmid Mini Kit (Eppendorf), and isolated plasmid DNA was digested with DraIII or NcoI. Digested plasmid DNA was electrophoresed on 0.8% agarose gels, stained with EtBr, and visualized by UV.

Southern blotting. To screen for IS*492*Δ*mooV*::*tet* chromosomal and plasmid insertions from tandem (IS*492*Δ*mooV*::*tet*)2 experiments in DH5α, total DNA was isolated from selected DH5α/pAMC13 pAMC5 pAG900 isolates using a AquaPure Genomic Kit (BIO-RAD), and 4 μg of chromosomal DNA was digested with DraIII. Digested DNA was then electrophoresed on a (0.7%-0.8%) agarose gel and transferred onto a nylon membrane using the following method: the

gel was depurinated in 0.25 M HCl for 15 min, denatured in 0.5 M NaOH and 1.5 M NaCl for 30 min, and neutralized in 0.1 M Tris pH 7.5 and 1.5 M NaCl for 30 min. DNA was transferred onto a nylon membrane using a BIO-RAD model 785 vacuum blotter following the manufacturer's guidelines. The gel was probed using a hybridization temperature of 68°C with a digoxygenin (DIG)-labeled 1.3-kb tetracycline resistance gene fragment amplified from pAMC7 using primers MfeITet and NewBsmBITet and dNTPs labeled with DIG (Roche) according to the manufacture's guidelines. Hybridization was followed by two low stringency washes in 2X SSC (1X SSC is 0.15 M NaCl with 0.015 M sodium citrate)-0.1% (wt/vol) sodium dodecyl sulfate (SDS) at room temperature for 5 min and two high stringency washes in 0.1X SSC, 0.1% at 68°C for 15 min. Following the final washes, the blot was incubated with an anti-DIG antibody-alkaline phosphatase conjugate diluted 1:5000 (Roche) for 30 min at room temperature, and developed using a solution of Nitro blue tetrazolium chloride/ 5-Bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) (Roche).

Analysis of cointegrate, single insertion and precise excision plasmid junctions from *in vivo* (IS*492*Δ*mooV*::*tet*)2 insertion assays. Cointegrate plasmids between pAMC13 and pAMC5 (resulting in pAMC14) and single IS*492*Δ*mooV*::*tet* insertions into pAMC5 (pAMC1) isolated from *in vivo* (IS*492*Δ*mooV*::*tet*)2 transposition assays were gel extracted from 0.8% agarose gels using the QIAquick Gel Extraction Kit (Qiagen). Undigested isolated pAMC14 DNA was used in four separate junction PCR assays using the following primers: (i) Junction #1= M13R + ISL.66 (225-bp); (ii) Junction #2= M13F + ISR.55 (242-bp) or M13F + TetInv1 (301-bp); (iii) Junction #3= ExcR + TetInv1 (295-bp) and (iv) Junction #4= ExcF + ISL.66 (141-bp). Undigested pAMC1 DNA was subjected to Junction #1 and Junction #2 (M13F + TetInv1) PCR analysis. Reactions were done in 1X *Taq* polymerase Buffer B (Promega), 2.5 mM $MgCl_2$, 10

μM each dNTPs, and 1 unit *Taq* polymerase. Cycling conditions were 95°C for 2 minutes, followed by 40 cycles of 95°C for 45 seconds, 58°C (53.7°C for Junction #3) for 45 seconds, and 72°C for 20 seconds. PCR products were detected on a 2% agarose gel (1% GenePure LE and 1% GenePure 3:1), and purified PCR DNA [QIAquick PCR Purification Kit (Qiagen)] was sequenced at the UM-BRCF.

Single IS*492*Δ*mooV*::*tet* excision plasmids (resulting in pAMC7) from *in vivo* (IS*492*Δ*mooV*::*tet*)2 transposition assays were isolated via diluting plasmid DNA samples containing pAMC7 1:50 and transforming diluted plasmid DNA into $CaCl_2$ competent DH5α cells. Transformants were plated on LB/Cm_{34} to select for pAMC7, and after overnight incubation at 37°C were patched onto LB/Ap_{80} and LB/Sp_{50} to select for pAMC5 and pAG900, respectively. Plasmid DNA was isolated from colonies sensitive to Ap_{80} and Sp_{50}. Plasmids isolated from Cm^r colonies were digested with HindIII and analyzed on a 0.8% agarose gel (GenePure LE) to confirm presence of only pAMC7 DNA. Isolated pAMC7 DNA was sequenced using primers ExcF and ExcR at the UM-BRCF.

Western blot for MooV expression in *E. coli* at 37°C versus 25°C. $CaCl_2$ competent DH5α cells were transformed with pAG900 or pAG900ΔmooV. Transformations performed at 37°C were outgrown 1 h at 37°C, and transformations performed at 25°C were outgrown 6 h at 25°C. After transformation outgrowth 300 μl cells were inoculated into 2.7 ml LB/Sp_{50} and grown with aeration overnight at 37°C or for 72 h at 25°C. Turbid cultures were diluted to an optical density at 600 nm (OD_{600}) 0.1 in 6 ml LB/Sp_{50} or LB/Sp_{50} supplemented with 50 μM IPTG, grown to midlogarithmic phase, and 2 ml cells were harvested by centrifugation. Cells were resuspended in 54 μl 1X loading buffer,[119] incubated at 100°C for 5 min, and loaded onto a 12% sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel. Western blot analysis was performed as

described in Ausubel *et al.*[119] by using a primary anti-MooV chicken polyclonal α-MooV.His_6 antibody[88] at 1:2,000 dilution, and a secondary anti-chicken antibody alkaline phosphatase conjugate (Sigma-Aldrich) at 1:2,000 dilution. Membranes were developed with 5-bromo-4-chloro-3-indolyphosphate-nitorblue tetrazolium SigmaFast tablets (Sigma-Aldrich).

2.3. Results

Complementing IS*492*Δ*mooV*::*tet* for circle junction formation

Previous studies characterizing IS*492* transposition in *E. coli* demonstrated that precise excision and generation of a circular form of IS*492* is transposase dependent, but is not accompanied with insertion of the element into an *eps* target plasmid or *E. coli* chromosome.[88] It was hypothesized that either *P. atlantica* specific recombination factors are required for IS*492* insertion in *E. coli*, or other conditions non-specific to *P. atlantica*, including temperature, influence the IS*492* insertion activity in *E. coli*. An IS*492*-derivative, which contains a promoter-less tetracycline resistance gene with its own Shine-Dalgarno sequence (IS*492*Δ*mooV*::*tet*), was created to select for insertion events in *E. coli*. Strains containing pAMC1, with IS*492*Δ*mooV*::*tet* cloned downstream of the *lac* promoter, are resistant to 15 μg/ml tetracycline, whereas *E. coli* strains with IS*492*Δ*mooV*::*tet* cloned downstream of two transcriptional terminators on pAMC3 are sensitive to 15 μg/ml tetracycline at 37°C. IS*492*Δ*mooV*::*tet* circle junction formation was detected by colony PCR assays[88] when pAMC1 is provided with MooV *in trans* (Figure 2.2 A), indicating that the 595-bp MfeI-BsmBI fragment deleted within IS*492* did not remove any *cis*-acting sequence required for MooV-mediated circle junction formation. No circle junction formation by IS*492*Δ*mooV*::*tet* was detected in colony PCR assays with pAMC3 when 0.02% or 0.2% arabinose was included in the medium to induce *mooV* expression *in cis* from the *araBAD* promoter. Although low occurrence of IS*492*Δ*mooV*::*tet* circle junction

formation was detected from pAMC3 in the absence of arabinose (Figure 2.2 C) and from pAMC7 (a derivative of pAMC3 that has 777-bp of *mooV* coding sequence deleted) complemented with MooV *in trans*, reproducible circle junction formation was observed from pAMC3 when MooV is provided *in trans* (Figure 2.2 C). Subsequent insertion assays used pAMC3 (complemented with pAG900) as the IS*492*Δ*mooV*::*tet* donor.

Identifying factors required for IS*492*Δ*mooV*::*tet* insertion in *E. coli*

Experiments designed to identify factors required for IS*492* insertion in *E. coli* used the IS*492*Δ*mooV*::*tet* construct in pAMC3 to select for insertion of the element downstream of an actively transcribed promoter in the correct orientation to express the promoter-less tetracycline resistance gene. In addition, insertion assays used 134-bp of *epsG* target sequence from *P. atlantica* cloned downstream of P_{lac} on pAMC5, and MooV provided *in trans* from pAG900. Complementation experiments using a *P. atlantica* cosmid library to identify *P. atlantica* recombination factors required for IS*492* transposition did not yield insertion of IS*492*Δ*mooV*::*tet* in *E. coli* strain TOP10 at 37°C (Table 2.2). However when incubation temperatures were lowered to 32°C, Tet^r transconjugants were obtained at frequencies of 10^{-6}-10^{-4}. To confirm IS*492*Δ*mooV*::*tet* insertion into pAMC5, plasmid DNA was isolated from Tet^r transconjugants from Experiment III, introduced into naïve TOP10 cells, and transformants containing IS*492*Δ*mooV*::*tet* inserted into pAMC5 were selected with Ap and Tet. Characterization of plasmids isolated from Ap^r Tet^r transformants revealed IS*492*Δ*mooV*::*tet* insertion into pAMC5 via restriction (Figure 2.3 B lanes 2, 4-5) and PCR (Figure 2.3 C lanes 3, 5-6) analyses. Sequence analysis of insertion PCR products confirmed the site-specific insertion of IS*492*Δ*mooV*::*tet* into the 5-bp 5'-CTTGT-3' in the *eps* sequence. Cosmids that appeared to facilitate IS*492*Δ*mooV*::*tet* insertion into pAMC5 were mated out of the Tet^r TOP10

transconjugants (Figure 2.3 C Lanes 3, 5, and 6) into naïve DH5α, designated cosmids 3, 11, and 17, respectively, and used for confirmatory complementation experiments. Subsequent complementation experiments performed at 32°C yielded Tet^r frequencies from 10^{-7}-10^{-6}. Interestingly, the recipient strain containing the donor, target, and MooV-expression vectors gave Tet^r colonies at frequencies of 10^{-8}-10^{-7} in the absence of the cosmid library (Table 2.3). Although Tet^r frequencies were 10-fold higher in conjugation assays compared to the no-cosmid control experiments, restriction analysis of cosmids 3, 11, and 17 revealed the lack of *P. atlantica* chromosomal DNA cloned into the cosmid vector, and PCR analysis failed to detect IS*492*Δ*mooV*::*tet* insertion into these cosmids (data not shown). Subsequently, IS*492*Δ*mooV*::*tet* transposition in *E. coli* was further characterized without the use of the *P. atlantica* cosmid library.

Selected Tet^r colonies from the no-cosmid control experiments were screened to detect insertion into pAMC5. HindIII digests of plasmids isolated from selected Tet^r colonies revealed a 5.6-kb band corresponding to an IS*492*Δ*mooV*::*tet* insertion into pAMC5 (Figure 2.4 B lanes 5-6, 11). Insertion PCR confirmed the single insertion event for these three isolates (Figure 2.4 C, lanes 3, 4, 9), as well an additional isolate (Figure 2.4 C, lane 10) where insertion was not detected via restriction analysis (Figure 2.4 B lane 12). The additional 1.8-kb HindIII fragment in Figure 2.4 B lanes 11-12 was suggestive of a tandem insertion into the IS*492*Δ*mooV*::*tet* donor pAMC3. Additional products were not characterized further, but may be suggestive of insertion products mediated by the tandem IS*492*Δ*mooV*::*tet* (see Figure 2.8). The tandem insertion plasmid, designated pAMC10, was isolated and additional restriction analysis confirmed the presence and head-to-tail orientation of the tandem IS*492*Δ*mooV*::*tet* [(IS*492*Δ*mooV*::*tet*)2] insertion (Figure 2.5). To determine the spacer sequence between (IS*492*Δ*mooV*::*tet*)2 the 1.8-kb

HindIII fragment from pAMC10 was cloned into pUC18 and sequence analysis revealed that the elements are separated by the 5-bp sequence 5'-CTTGT-3', which flanks (IS*492*Δ*mooV*::*tet*)2 in pAMC10 and is the same spacer sequence found in IS*492* extrachromosomal circles.[88] Plasmids isolated from *E. coli* strains transformed with pAMC10 reproducibly gave two plasmid species: pAMC10 (tandem copy of IS*492*Δ*mooV*::*tet*) and pAMC3 (single copy of IS*492*Δ*mooV*::*tet*) (see Figure 2.5). It was hypothesized that the *mooV* copy on pAMC10 mediated the high-frequency excision of a single copy IS*492*Δ*mooV*::*tet* from pAMC10, and so 777-bp of internal *mooV* sequence was removed from pAMC10 (see Figure 2.1) to increase the plasmid's stability and to prevent homology-based recombination with the *mooV* copy on pAG900 in subsequent experiments. The resulting (IS*492*Δ*mooV*::*tet*)2 plasmid, pAMC13, was used in subsequent insertion assays in *E. coli* to characterize further the IS*492* transposition mechanism. These results indicate a reduction in incubation temperature permits detectable levels of insertion activity in *E. coli* in the absence of *P. atlantica* host factors, and that IS*492* insertion in *E. coli* displays target specificity.

Assay of temperature sensitivity to characterize monomer IS*492*Δ*mooV*::*tet* insertion in *E. coli*

To characterize further the effect of temperature on IS*492*Δ*mooV*::*tet* insertion in *E. coli* different temperatures for outgrowth (following transformation with the *mooV*-expression vector) and for selection of insertion were used. Insertion of IS*492*Δ*mooV*::*tet* was not observed using transformation outgrowth and selection temperatures of 37°C (data not shown), however combining 37°C transformation outgrowth with selection temperatures of 32°C or 25°C yielded Tet^r frequencies ranging from 10^{-8}-10^{-6} and 10^{-3}-10^{-2}, respectively (Table 2.4, lines 1 and 2). Outgrowth and selection of insertions at 32°C yielded Tet^r frequencies up to 10^{-4} (Table 2.4, line

3), while outgrowth and selection of insertions at 25°C resulted in the highest frequencies of Tet^r, at 10^{-1} (Table 2.4, line 4). We subsequently determined that background Tet^r in *E. coli* (DH5α) strains containing IS*492*Δ*mooV*::*tet* donor plasmids in the absence of MooV is observed at lower temperatures (Table 2.5), and thus prevents the accurate determination of IS*492*Δ*mooV*::*tet* insertion frequencies. The temperature sensitive expression of the *tet* gene in IS*492*Δ*mooV*::*tet* is linked to the putative *mooV* promoter; primer extension suggests it is located 220-bp upstream of the *tet* gene in IS*492*Δ*mooV*::*tet* (Heinrich and Karls, unpublished data). Despite the background resistance to Tet, single insertions into the *eps* target plasmid and tandem insertions into the IS*492*Δ*mooV*::*tet* donor plasmid were observed (Table 2.4). Out of a total 145 Tet^r colonies screened from all 32°C and 25°C experiments, 77 had single insertions into the *eps* target plasmid and 28 had tandem insertions into the IS*492*Δ*mooV*::*tet* donor plasmid. To determine if IS*492*Δ*mooV*::*tet* inserted into the chromosome, chromosomal DNA was isolated from selected Tet^r strains, including strains that did not contain IS*492*Δ*mooV*::*tet* plasmid insertion products, digested with PstI (which does not contain a recognition site in IS*492*Δ*mooV*::*tet*) and subjected to inverse PCR[118] and nested PCR analysis. Chromosomal insertions were not observed in selected Tet^r strains (Figure 2.6); the 1.7-kb and 198-bp products from inverse PCR assays were determined to be from the IS*492*Δ*mooV*::*tet* donor plasmid and an IS*492*Δ*mooV*::*tet* circle junction [from an IS*492*Δ*mooV*::*tet* circle or (IS*492*Δ*mooV*::*tet*)2], respectively. Similar products were observed from chromosomal DNA digested with SacI and AatII, which also do not contain recognition sites in IS*492*Δ*mooV*::*tet* (not shown). These results further suggest that IS*492* insertion in *E. coli* is sensitive to temperature, and displays target site-specificity.

Characterizing transposition activity of (IS*492*Δ*mooV*::*tet*)2 in *E. coli*

Tandem head-to-tail insertions have been observed for insertion elements IS*21*,[15,120] IS*30*,[16] IS*911*,[121] and the IS*110* family member IS*117*.[122] Tandem IS*21*, IS*30*, and IS*911* have been shown to be more active for transposition compared to the monomer element, and the most common transposition products include single excision of one element, and replicon fusions that resemble cointegrate structures from replicative transposition, with two copies the element in direct orientation at the donor and target plasmid junctions.[15,16,121] To determine if (IS*492*Δ*mooV*::*tet*)2 behaves in a similar manner as other characterized tandem insertion elements, and if it displays similar insertion characteristics as the monomer IS*492*Δ*mooV*::*tet* (i.e. temperature sensitive insertion and target site specificity), pAMC13 was used as the donor plasmid in insertion experiments. The MooV expression vector pAG900 was introduced into DH5α strains containing pAMC13, pAMC5, and DH5α strains containing pAMC7 (monomer IS*492*Δ*mooV*::*tet*), pAMC5. Plasmid DNA was isolated from cultures started from randomly picked transformant colonies and digested with DraIII, which does not have a restriction site in IS*492*Δ*mooV*::*tet*. At transformation outgrowth and final incubation temperatures of 37°C, transposition of (IS*492*Δ*mooV*::*tet*)2 was frequent enough to observe in randomly picked transformant colonies without the use of selection (Figure 2.7, Table 2.6), compared to transposition of monomer IS*492*Δ*mooV*::*tet* (Figure 2.7). The plasmid profiles observed from (IS*492*Δ*mooV*::*tet*)2 transposition are consistent with a single excision of one IS*492*Δ*mooV*::*tet* element, resulting in pAMC7, and additional bands at 2.3-kb and 10.7-kb are consistent with a fusion plasmid between pAMC13 and pAMC5 (Figure 2.7). The pAMC13/pAMC5 fusion plasmid, designated pAMC14, was isolated and additional restriction analysis confirmed the plasmid fusion (Figure 2.9). With the number of IS*492*Δ*mooV*::*tet* elements from pAMC13 being

conserved in pAMC14, the HindIII restriction profile of pAMC14 in Figure 2.9 also suggests MooV inserted (IS*492*Δ*mooV*::*tet*)2 into pAMC5 via a cut-and-paste mechanism rather than a replicative mode of transposition. Southern blot analysis performed on total DNA (chromosomal and plasmid) isolated from these strains and digested with DraIII did not reveal chromosomal insertions from the tandem or monomer IS*492*Δ*mooV*::*tet* donor, but did confirm the plasmid-based transposition products from (IS*492*Δ*mooV*::*tet*)2, which are dependent on MooV (Figure 2.7). The probe did not anneal to any chromosomal loci in a naïve DH5α strain (data not shown).

Different combinations of transformation outgrowth and final incubation temperatures were used to determine if temperature affects (IS*492*Δ*mooV*::*tet*)2 transposition. The results in Table 2.6 show that (IS*492*Δ*mooV*::*tet*)2 is active for both excision and insertion at all temperature combinations used, and that the primary transposition products observed with MooV provided *in trans* is single excision of one IS*492*Δ*mooV*::*tet* followed by cointegrates between the pAMC13 pAMC5. All pAMC14 cointegrate plasmids observed in Table 2.6 had the same DraIII restriction profile (data not shown), suggesting the presence of two IS*492*Δ*mooV*::*tet* elements in head-to-tail orientation at the plasmid junctions, and additional restriction analysis of independently isolated pAMC14 cointegrates confirmed the number and orientation of the IS*492*Δ*mooV*::*tet* elements (data not shown). Single IS*492*Δ*mooV*::*tet* insertion into pAMC5 was observed in the 37°C/37°C experiments and other types of plasmid insertion products were observed at all temperature combinations. Additional restriction and Southern blot analysis of other (IS*492*Δ*mooV*::*tet*)2 insertion products suggests pAMC13 replicon fusions with other plasmids, including pAG900, pAMC7, and pAMC1 (Figure 2.8). Other transposition products were not analyzed, but may reflect other types of intermolecular transposition. All

(IS*492*Δ*mooV*::*tet*)2 insertion products were dependent on MooV *in trans* (Table 2.6 lines 2, 4, 6), however MooV-independent deletion of an IS*492*Δ*mooV*::*tet* copy was occasionally observed (Table 2.6 line 6). (IS*492*Δ*mooV*::*tet*)2 is a long tandem repeat (3.6-kb in length), and MooV-independent deletion of an IS*492*Δ*mooV*::*tet* copy could possibly be due to pathways independent of RecA recombination, such as errors in DNA replication.[123] These data suggest that although (IS*492*Δ*mooV*::*tet*)2 is typically stable in the absence of MooV, when MooV is provided *in trans* it is very active in transposition in *E. coli* at 37°C and 25°C. These results also suggest that the junction formed by (IS*492*Δ*mooV*::*tet*)2, which is the same junction in the circular form of IS*492*, can be used by MooV to mediate insertion of (IS*492*Δ*mooV*::*tet*)2 in *E. coli* at 37°C and 25°C.

Bartlett *et al.*[107] and Perkins-Balding *et al.*[88] observed IS*492* precise excision from the *eps* site in *P. atlantica* and *E. coli*. To determine whether the single IS*492*Δ*mooV*::*tet* excision products from the (IS*492*Δ*mooV*::*tet*)2 experiments in Table 2.6 result from precise excision of one copy of the element, 16 single excision products from independent experiments were sequenced. This analysis confirmed precise excision of one IS*492*Δ*mooV*::*tet* element, leaving the remaining IS*492*Δ*mooV*::*tet* flanked by the 5-bp target sequence 5'-CTTGT-3' (data not shown). The low frequency MooV-independent excision products also showed "precise excision," which is mediated by host recombination and/or replication functions.

MooV mediates site-specific insertion of (IS*492*Δ*mooV*::*tet*)2

IS*492* insertion displays target-site specificity in *P. atlantica*,[109] and monomer IS*492*Δ*mooV*::*tet* targets the 5-bp 5'-CTTGT-3' at the ends of the element and within the *eps* target sequence in *E. coli* (this study). To determine if MooV mediates site-specific insertion of (IS*492*Δ*mooV*::*tet*)2 into the *eps* target sequence in pAMC5, 12 pAMC14 cointegrate plasmids

were isolated from independent (IS*492*Δ*mooV*::*tet*)2 experiments, and junction PCR analysis was performed (Figure 2.10). pAMC14 cointegrates analyzed had all four plasmid junctions, as seen with a representative junction PCR assay with one cointegrate plasmid (Figure 2.10). Sequencing analysis of the four junctions from all pAMC14 cointegrate plasmids analyzed via junction PCR confirmed the site-specific insertion of pAMC13 into pAMC5, with the 5-bp 5'-CTTGT-3' between all four plasmid and IS*492*Δ*mooV*::*tet* junctions (data not shown). The two single IS*492*Δ*mooV*::*tet* insertions into pAMC5, creating pAMC1, (Table 2.6), were also isolated and subjected to PCR analysis to confirm the insertion junctions (Data not shown). Sequence analysis of the pAMC1 junction products confirmed the site-specific insertion of IS*492*Δ*mooV*::*tet* into the *eps* target site, with duplication of the 5-bp target sequence (data not shown). These results indicate that MooV targets the 5-bp sequence in the *eps* site for insertion of IS*492*Δ*mooV*::*tet*. These results, along with the data from Figure 2.7, also suggest that MooV used the inner junction of the (IS*492*Δ*mooV*::*tet*)2 to insert pAMC13 into pAMC5 via either a cut-and-paste mechanism, where the 5-bp target in pAMC5 is the site of staggered cleavage, or via a conservative site-specific recombination, where the 5-bp sequence is the common core of the recombination sites.

MooV expression in *E. coli* at 37°C and 25°C

Previous experiments where monomer IS*492* was not observed inserting in *E. coli* at 37°C may indicate MooV has different levels of expression or stability at 37°C versus 25°C. Total cell proteins were isolated from cultures of DH5α strains containing pAG900 that were grown at 37°C or 25°C, and Western blot analysis (Figure 2.11) showed that when MooV expression is induced with the addition of IPTG, expression levels were comparable at 37°C and 25°C. This suggests that the insertion activity of monomer IS*492* could be linked to a

temperature sensitive insertion intermediate, or MooV having different enzymatic activities at the different temperatures.

2.4. Discussion

IS*492* is a member of the atypical IS*110* family of insertion elements, which lack terminal inverted repeats, and with the exception of IS*492* and IS*117*,[98] do not create target site duplications upon insertion.[87] An intriguing feature of MooV and the other transposases in this IS family is that they have high sequence homology to Piv, a site-specific invertase that uses a DEDD motif for catalysis[106] and is not related to the S- or Y- site-specific recombinases.[92] Initial observations of IS*492* transposition in *P. atlantica* revealed that excision of IS*492* from *epsG*, which results in circle formation,[88] does not correlate with insertion into another chromosomal site, and the other chromosomal copies of IS*492* do not excise when IS*492* re-inserts into *epsG*,[109] suggesting that the IS*492* transposition mechanism (cut-and-paste vs. replicative) may depend on the location of the element. Additional characteristics of IS*492* transposition in *P. atlantica* and *E. coli*, including high frequency precise excision,[107,112] target site specificity,[88,109] and the requirement of flanking DNA for precise excision[88] indicate this element displays unusual characteristics compared to other classical DDE transposable elements. The actual recombination mechanism used by IS*492* and other members of the IS*110* family is not well defined, and characterizing the IS*492* insertion mechanism will help elucidate the recombination mechanism used by this atypical family.

While many of the requirements for the IS*492* precise excision have been characterized in both *P. atlantica* and *E. coli*,[88,107,112] an IS*492* insertion assay in *E. coli* was necessary to analyze the insertion mechanism. Previous observations suggested the requirement for specific *P. atlantica* recombination factors for IS*492* insertion in *E. coli*,[88] however observations from

this study suggests lower incubation temperatures facilitates MooV-mediated insertion of monomer IS*492*Δ*mooV*::*tet* into an *eps* target and tandem head-to-tail insertion into the IS*492*Δ*mooV*::*tet* donor plasmid. Other transposable elements, including Tn*3*[124,125] and IS*911*[126] display temperature sensitive transposition, and it has been demonstrated that the production of truncated forms of IS*911* transposase that interfere with IS*911* transposition is sensitive to temperature.[127] Since IS*492* is isolated from a marine bacterium with an optimal growth temperature between 20-25°C,[128] it is possible that *mooV* transcripts or MooV protein are more stable at these temperatures; Western blot analysis of MooV expression from a heterogeneous promoter in *E. coli* at 37°C vs. 25°C did not reveal any significant differences in the levels of MooV (Figure 2.11). Since MooV is able to mediate precise excision of IS*492* in *E. coli* at 37°C, it is possible that IS*492* intermediate products may be sensitive to temperature. It is interesting to note that the IS*110* family member IS*1000*, isolated from *Thermus thermophilus* and whose transposase has the most homology with MooV,[92] was not observed to transpose in *E. coli.*[129] This observation may also be due to differences in optimal growth temperatures of the two organisms (37°C versus 70°C).

Tandem head-to-tail dimers are commonly found in insertion elements that excise to form circular products. Elements in this category include IS*21*,[130] IS*30*,[131] IS*911*,[121] and the IS*110* family member IS*117*.[122] (IS*492*Δ*mooV*::*tet*)2 displays similar transposition characteristics as the above elements: i.e. more frequent transposition activity compared to the monomer element, with the most common recombination products being single excision of an element and cointegrate insertion. Analysis of the most frequent tandem IS*492*Δ*mooV*::*tet* transposition products revealed precise excision of one IS*492*Δ*mooV*::*tet* element, and site-specific insertion of the (IS*492*Δ*mooV*::*tet*)2 donor plasmid into the *eps* target plasmid; precise excision and site-specific

insertion is not observed with the classical DDE motif tandem IS*21*,[15] IS*30*,[16] and IS*911* elements.[121] The observation of (IS*492*Δ*mooV*::*tet*)2 transposition in *E. coli*, in particular insertion into the *eps* target plasmid, being less sensitive to lower temperature may indicate the hyperactivity of the junction sequence that is formed in (IS*492*Δ*mooV*::*tet*)2 and the IS*492* excised circle.[88] Formation of this junction sequence may be the rate-limiting step in IS*492* transposition, as has been proposed for IS*21*[130] and IS*30.*[131] Indeed the circular forms of IS*117*[99] and IS*911*[111,132] serve as transposition intermediates. Transposition of (IS*492*Δ*mooV*::*tet*)2 may indicate the insertion role of the IS*492* circle junction. Additional experimentation will be done to identify the role of the IS*492* circle in transposition.

Site-specific insertion of either monomer or dimer IS*492*Δ*mooV*::*tet* in *E. coli* suggests that MooV is the main factor required for target site selection in *P. atlantica* and *E. coli*. Although IS*492* chromosomal insertions were not isolated in *E. coli*, there may be other factors influencing IS*492* target site specificity, as the 5-bp sequence is represented 2,204 times in the *E. coli* K-12 genome. The 5-bp sequence 5'-CTTGT-3' is common within in the *epsG* site and the ends of the IS*492*. The analysis of transposition products in *E. coli* suggest IS*492* uses either a cut-and-paste mechanism where the 5-bp sequence is the site of staggered cleavage, or via a conservative site-specific recombination, where the 5-bp sequence is the common core of the recombination sites. Other members of the IS*110* family also display target site specificity including IS*117*,[98,99] IS*900*,[94] IS*901*,[93] and IS*116*.[96] In particular it has been proposed that the 3-bp 5'-TAG-3' 'attachment sequence' found in the circular IS*117* element aligns with the same 3-bp sequence found in its preferred target site, and that the resulting duplication of this sequence upon IS*117* insertion is the result of a site-specific recombination mechanism.[98] These observations suggest that the recombinases in the Piv/MooV family use an unusual

recombination mechanism that combines features of both DNA transposition and conservative site-specific recombination. Additional IS*492* transposition experiments, including *in vitro* assays, will be developed to address further the IS*492* recombination mechanism.

Table 2.1. Oligonucleotides used in this study[a]

MfeTetI	ATTACAATTGATTGCTAACGCAGTCAGG
NewTetBsmBI	TAGAGCGGAGACGTGAATCCGTTAGCGAGG
EPSL.DraIII	TTCACGTAGTGTTTGGTGGCATTGAAAAC
EPSR.PsiI	TTTTATAAGAGAGCTTGATAGACCTAAC
EPSL.58	CGGTACTGTCTTATCATCCTAATCG
EPSR.76	CAGGAGGCTCTCTCTATTGTACAGC
ISL.66	AAGCTTGGGTCTCACTTTACC
ISR.51	GGTCATGTGGGAAGCGCCCG
ISR.55	ACGGGGTCATGTGGGAAGC
TetInv1	ACCTCGACCTGAATGGAAGC
ExcF	CCTTTAGGGTTCCGATTTAGTGC
PRACP.1	GTTACATCGAACTGGATCTCAACA
PCRAP.2	ATGATACCGCGAGACCCACGC
CamF	CATTCTGCCGACATGGAAGCC
CamR	GCAATGAAAGACGGTGAGCTGG
EPSPrimer.1	TTAGCTATTGACGCCATAGTA
EPSPrimer.2	CCTGAGTTATTATCCATCCTTG
M13R	GAGGAAACAGCTATGAC
M13F	GTAAAACGACGGCCAG

[a] Oligonucleotides listed in 5' to 3' orientation

Table 2.2. Complementation Assay with *P. atlantica* Library for IS*492*Δ*mooV*::*tet* Insertion in Recipient *E. coli* strain TOP10

		CFU/ml				
Experiment number[a]	**Temperature**[b]	**Total recipients ($Ap^r Cm^r Sp^r$)**	**Total transconjugants ($Ap^r Cm^r Sp^r Kn^r$)**	**Total transconjugants ($Ap^r Kn^r$)**	**Total transconjugants ($Ap^r Kn^r Tet^r$)**[c,d]	**Frequency of insertion ($Ap^r Kn^r Tet^r$/ $Ap^r Kn^r$)**
I	37°C	6.0×10^6	2.0×10^6	1.3×10^6	0 (IPTG)	N/A[e]
II	32°C	1.2×10^7	4.6×10^6	3.8×10^6	10 10 (IPTG)	2.6×10^{-6} 2.6×10^{-6}
III	32°C	6.7×10^5	2.2×10^5	1.5×10^5	20 (Glu) 60 (IPTG+Glu)	1.3×10^{-4} 4.0×10^{-4}

[a] Experiments I and II, cells from LB mating plates were resuspended in 500 µl LB before plating on appropriate selective media. Experiment III, cells from LB mating plates were resuspended in 1 ml LB.

[b] 37°C experiment, conjugation and final incubation/selection temperatures were performed overnight. 32°C experiment, conjugation plates were incubated overnight at 32°C, and final incubation/selection temperatures were 48 h at 32°C.

[c] Colonies counted at least 1 mm in size

[d] IPTG = 10 µM IPTG present in medium to induce transposase expression from P_{tac}; G = 0.2% glucose present in medium to decrease transcription from P_{lac} across *eps* target

[e] N/A = non-applicable; no Tet^r transconjugants isolated

Table 2.3. Complementation Assay with *P. atlantica* cosmids for IS*492*Δ*mooV*::*tet* insertion in *E.coli* strain DH5α at 32°C

	Insertion Frequencies[a]			
Cosmid donor[b]	**Ap Kn Tet (Ap Tet)**[d]	**Ap Kn Tet (ApTet) + 10 µM IPTG**	**Ap Kn Tet (Ap Tet) + 0.2% Glu**	**Ap Kn Tet (Ap Tet) + 10 µM IPTG, 0.2% Glu**
Cosmid Library	4.3×10^{-7}	2.6×10^{-6}	1.3×10^{-6}	2.1×10^{-6}
Cosmid 3	N/A[e]	6.2×10^{-6}	3.9×10^{-6}	7.5×10^{-6}
Cosmid 11	6.7×10^{-6}	7.5×10^{-6}	8.8×10^{-6}	9.6×10^{-6}
Cosmid 17	2.5×10^{-6}	2.4×10^{-6}	2.8×10^{-6}	6.6×10^{-6}
No Cosmid[c]	6.3×10^{-8}	6.3×10^{-8}	N/A[e]	2.5×10^{-7}

[a] Insertion frequencies are the relative number of Tet^r colonies among total transconjugants or transformants in no-cosmid control experiments (2-3 repetitions gave similar frequencies)

[b] VCS257 donor strain for cosmid library; DH5α donor strain for cosmids 3, 11, 17 (See Materials and Methods)

[c,d] No cosmid control: DH5α containing the IS*492*Δ*mooV*::*tet* donor and *eps* target plasmids following transformation with the *mooV* expression vector. Selective media for insertion did not contain Kn

[e] N/A = non-applicable; no Tet^r colonies isolated

Table 2.4. IS*492*Δ*mooV*::*tet* Insertion in DH5α at Decreased Temperatures

	Insertion Frequencies[a]							
Outgrowth/ final incubation temperature[b] (Protocol)[c]	**Ap Tet**	**Ap Tet + 10 µM IPTG**	**Ap Tet + 0.2% Glu**	**Ap Tet + 10 µM IPTG, 0.2% Glu**	**Total colonies screened**	**Single insertion into *eps* target plasmid**	**Tandem insertion into IS*492*Δ*mooV*::*tet* donor plasmid**	**Other products**
37°C/ 32°C Protocol I	5.4×10^{-8}	1.6×10^{-6}	7.1×10^{-7}	1.1×10^{-6}	41	14	6	4
37°C/ 25°C Protocol II	N/A[d]	1.4×10^{-2}	4.9×10^{-3}	1.2×10^{-2}	29	27	14	13
32°C/ 32°C Protocol III	6.5×10^{-6}	4.5×10^{-4}	3.6×10^{-6}	3.7×10^{-4}	44	19	0	0
25°C/ 25°C Protocol III	1.8×10^{-1}	1.6×10^{-1}	1.1×10^{-1}	1.8×10^{-1}	31	17	8	2

[a] Insertion frequencies are the relative number of Tet^r colonies among total number of transformants (see Material and Methods)
[b] DH5α/pAMC3pAMC5 cells were transformed with pAG900 to initiate IS*492*Δ*mooV*::*tet* insertion and outgrown at 37°C (1 h), 32°C (6 h) or 25°C (6 h). Following transformation outgrowth, cells were plated on selective media for insertion and incubated at 32°C (48 h) or 25°C (72 h).
[c] See Materials and Methods
[d] N/A = non-applicable; no Tet^r colonies isolated

Table 2.5. Expression of Tet^r from IS*492*Δ*mooV*::*tet* at decreased temperatures in *E.coli*

			CFU/ml		
Plasmid[a]	**Description**	**Selective temperature**[b]	**Number of Cm^r**	**Number of Cm^r Tet^r**	**Frequency of Tet^r (Cm^r Tet^r/Cm^r)**
pAMC3	pBAD33 with Wild type *mooV* and IS*492*Δ*mooV*::*tet*	37°C	1.7×10^4	0	N/A[d]
		25°C	1.2×10^4	1.3×10^3	1.1×10^{-1}
pAMC7	pAMC3 with internal *mooV* deletion	37°C	1.6×10^4	0[(c)]	N/A
		25°C	5.8×10^4	6.1×10^3	1.1×10^{-1}
pBAD33	Vector only	37°C	5.9×10^3	0	N/A
		25°C	5.7×10^3	0	N/A

[a] Plasmids were maintained in *E. coli* strain DH5α

[b] After transformation outgrowth at 37°C (See Materials and Methods), transformation plates were incubated overnight at 37°C or 48 h at 25°C

[c] 23 tiny colonies (< 1 mm) present in one quadrant of plate, not included in total colony count

[d] N/A = non-applicable; no Tet^r colonies isolated

Table 2.6. (IS*492*Δ*mooV*::*tet)2* transposition in *E. coli*

				Transposition Products				
Plasmids[a]	**Experimental temperatures**[b]	**Total experiments**	**Total transformants screened**	**Single excision**[c]	**Cointegrates**[d]	**Single insertion**[e]	**No changes**[f]	**Other products**[g]
pAMC13 pAMC5 pAG900	37°C/ 37°C	10	101	57	36	2	28	8
pAMC13 pAMC5 pAG900ΔmooV	37°C/ 37°C	5	48	0	0	0	48	0
pAMC13 pAMC5 pAG900	25°C/ 37°C	6	54	22	5	0	25	13
pAMC13 pAMC5 pAG900ΔmooV	25°C/37°C	3	38	0	0	0	38	0
pAMC13 pAMC5 pAG900	25°C/25°C	6	53	31	4	0	10	12
pAMC13 pAMC5 pAG900ΔmooV	25°C/ 25°C	3	30	2	0	0	30	0

[a] Plasmids were maintained in *E. coli* DH5α cells (see Materials and Methods). DH5α/pAMC13 pAMC5 cells were transformed with pAG900 or pAG900Δ*mooV* and incubated 1 h at 37°C or 6 h at 25°C. Transformants were plated on LB/$Cm_{34}Ap_{80}Sp_{50}$ media and incubated overnight at 37°C or 72 h at 25°C. Plasmid DNA was isolated from cultures started from randomly selected Cm^r Ap^r Sp^r transformant colonies and digested with DraIII.

[b] Transformation outgrowth temperature, followed by final incubation temperature

[c] Single IS*492*ΔmooV::*tet* excision from pAMC13 to create pAMC7

[d] Cointegrate plasmids between pAMC13 and pAMC5 to create pAMC14

[e] Single IS*492*ΔmooV::*tet* insertion into pAMC5 to create pAMC1

[f] No changes in pAMC13

[g] Other plasmid insertion products

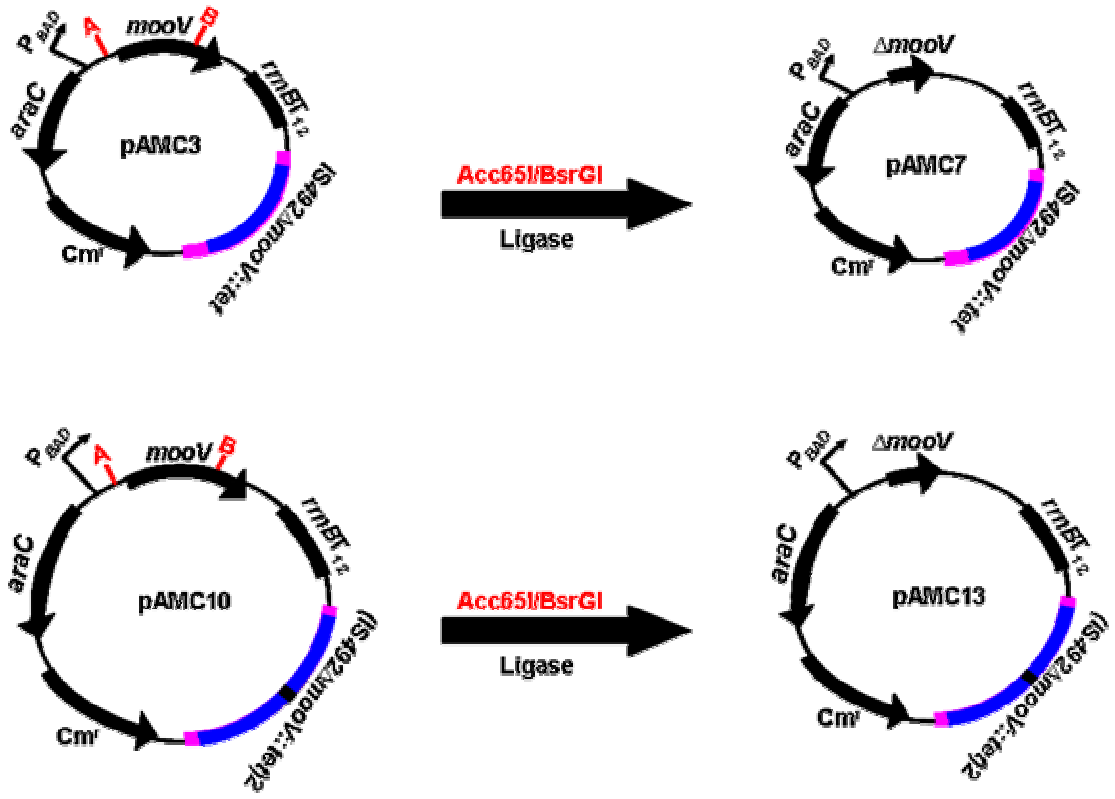


Figure 2.1. Creating pAMC7 and pAMC13. 777-bp of *mooV* coding sequence was deleted from pAMC3 and pAMC10 by digesting plasmids with Acc65I and BsrGI, which produce compatible cohesive ends. The plasmid backbones were self-ligated to produce pAMC7 and pAMC13, which have 182-bp of remaining *mooV* coding sequence.

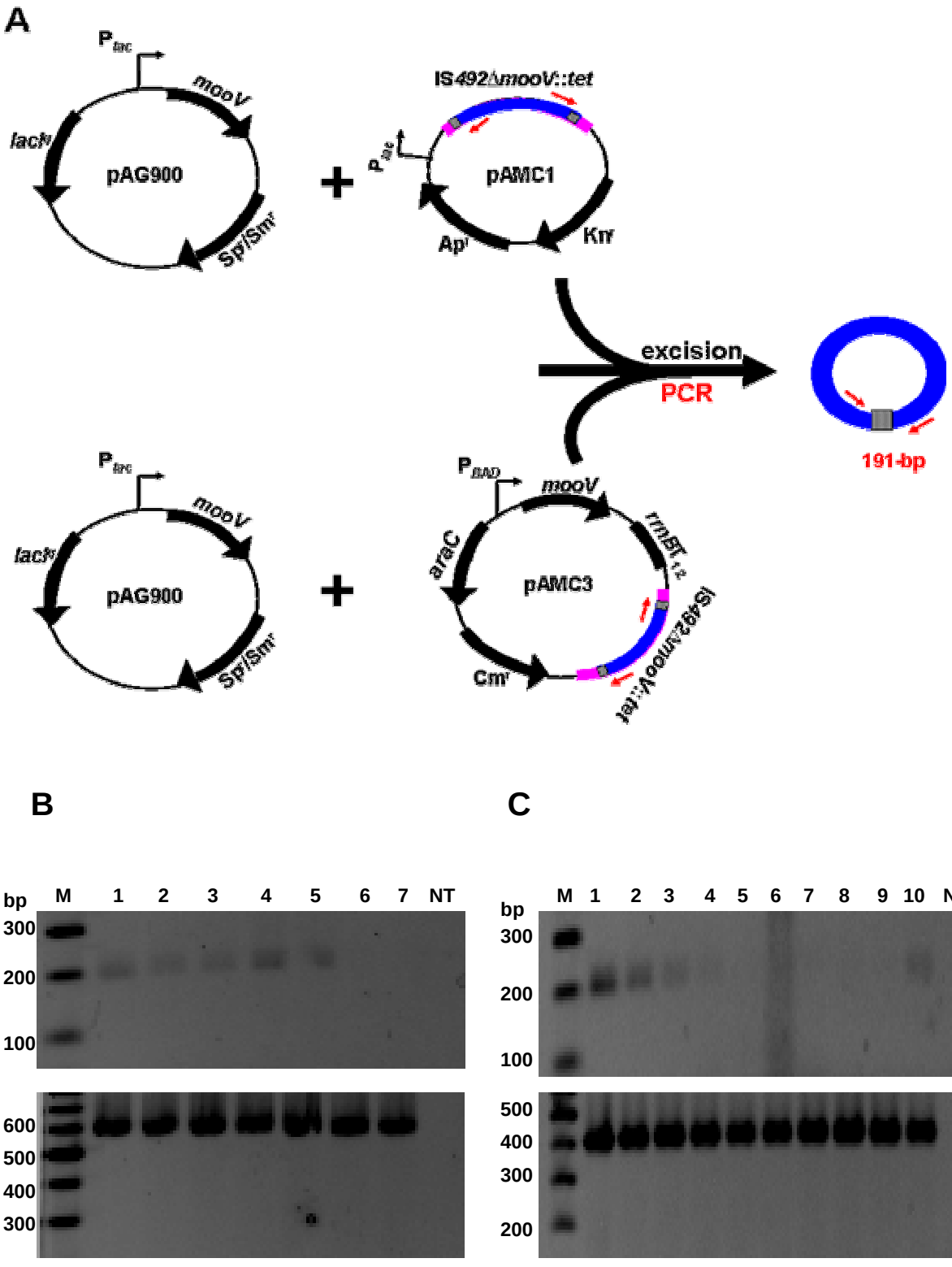
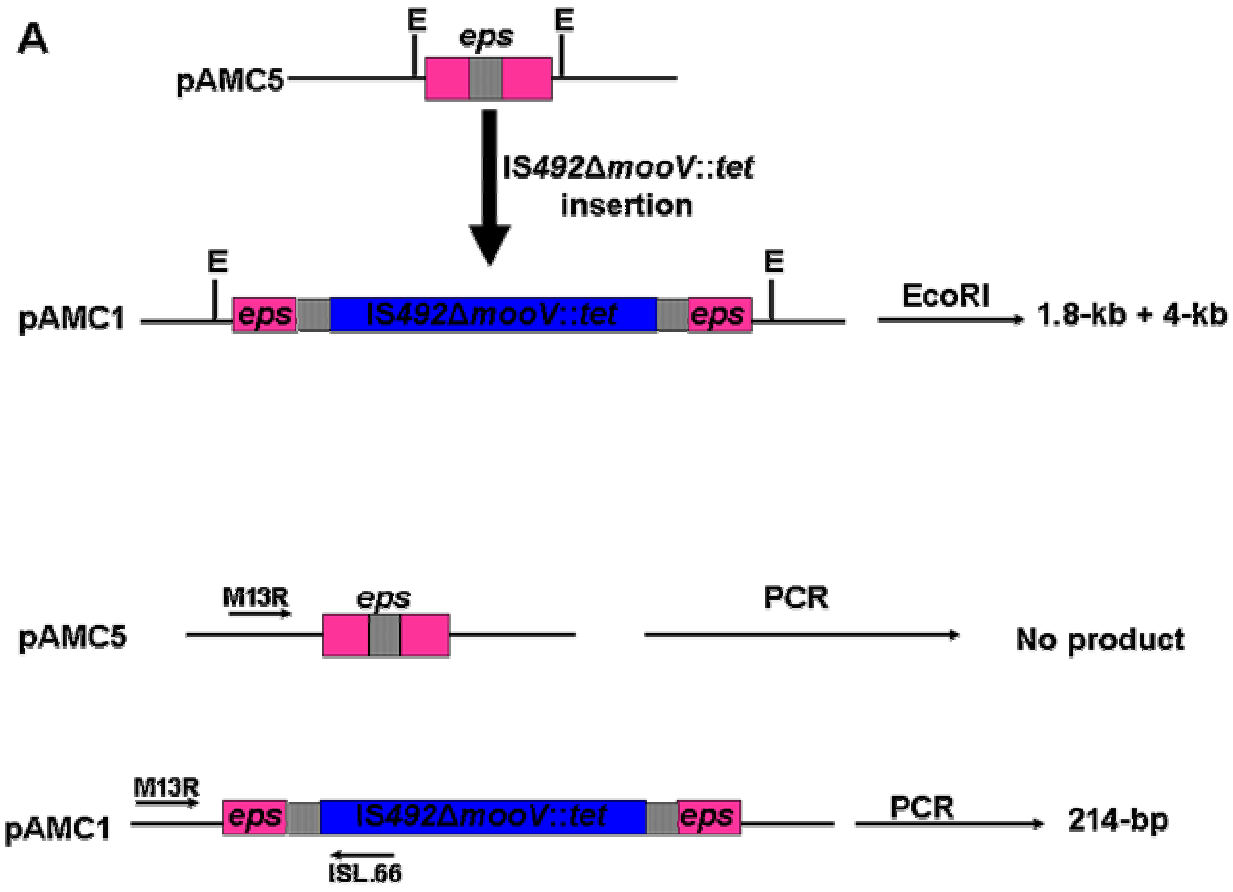


Figure 2.2. *In vivo* complementation assay for detecting IS*492*Δ*mooV*::*tet* circle junctions. (A) Cartoon illustrates IS*492*Δ*mooV*::*tet* complementation for excision and circle junction formation.

IS*492*Δ*mooV*::*tet* on pAMC1 is flanked by 58-bp left and 76-bp right *eps* sequence (pink boxes) and by 23-bp left and right *eps* sequence on pAMC3. The black and white striped box flanking IS*492*Δ*mooV*::*tet* represents the 5-bp duplicated sequence 5'-CTTGT-3'. (B) Complementation assay using pAMC1 as IS*492*Δ*mooV*::*tet* donor. 2% agarose gel electrophoresis of products generated from colony PCR with circle junction (CJ) primers (Top panel, expected CJ PCR product 191-bp), template control primers specific to the β-lactamase gene (Ap^r) on pAMC1 (Bottom panel, expected product 600-bp) and total DNA from five TOP10 colonies containing pAMC1 + pAG900 (lanes 1-5), and five TOP10 colonies containing pAMC1 (lanes 6-7). Lane M: Promega 100-bp marker; Lane NT: no template control. (C) Complementation assay using pAMC3 as IS*492*Δ*mooV*::*tet* donor. 2% agarose gel electrophoresis of products generated from colony PCR with circle junction (CJ) primers (Top panel, expected CJ PCR product 191-bp), template control primers specific to the chloramphenicol acetyl-transferase gene (Cm^r) on pAMC3 (Bottom panel, expected product 400-bp) and total DNA from five TOP10 colonies containing pAMC3 + pAG900 (lanes 1-5), and five TOP10 colonies containing pAMC3 (lanes 6-10). Lane M: Promega 100-bp marker; Lane NT: no template control.


A
E
eps
E
pAMC5
IS492ΔmooV::tet
insertion
E
E
pAMC1
eps
IS492ΔmooV::tet
eps
EcoRI
1.8-kb + 4-kb
M13R
eps
PCR
pAMC5
No product
M13R
pAMC1
eps
IS492ΔmooV::tet
eps
PCR
214-bp
ISL.66

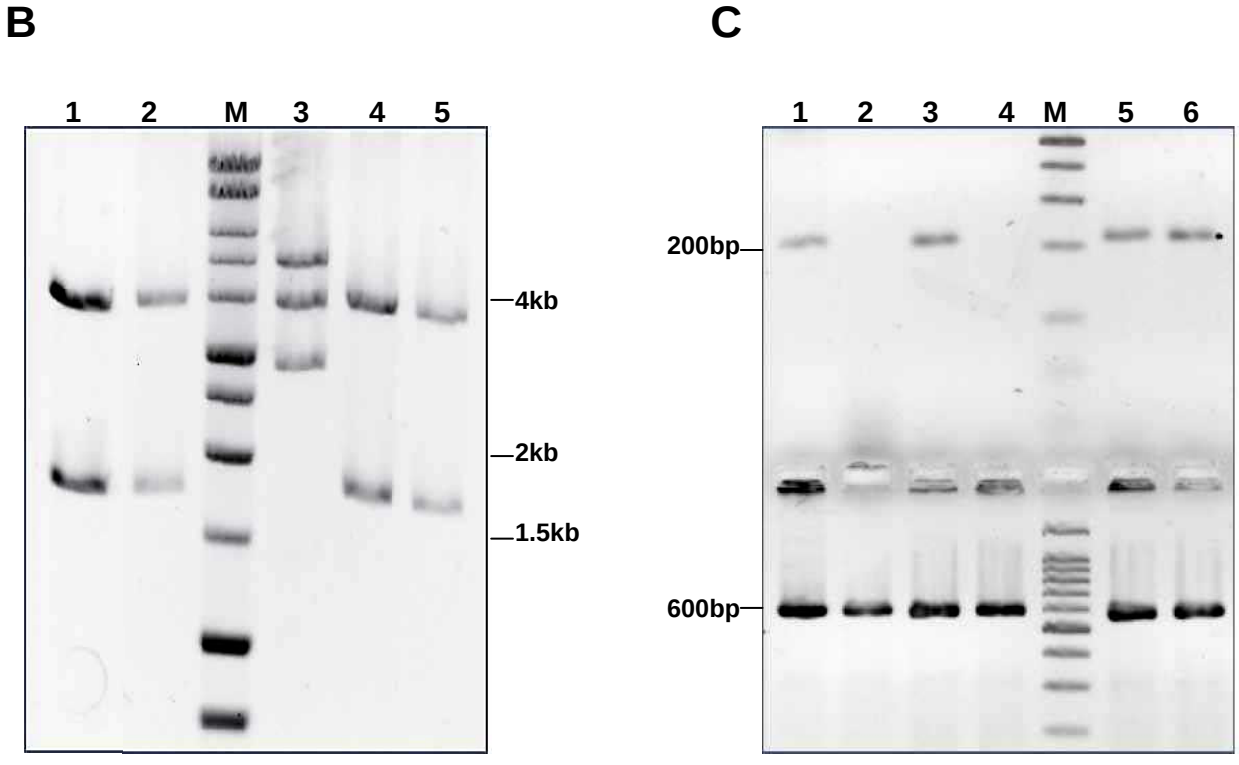
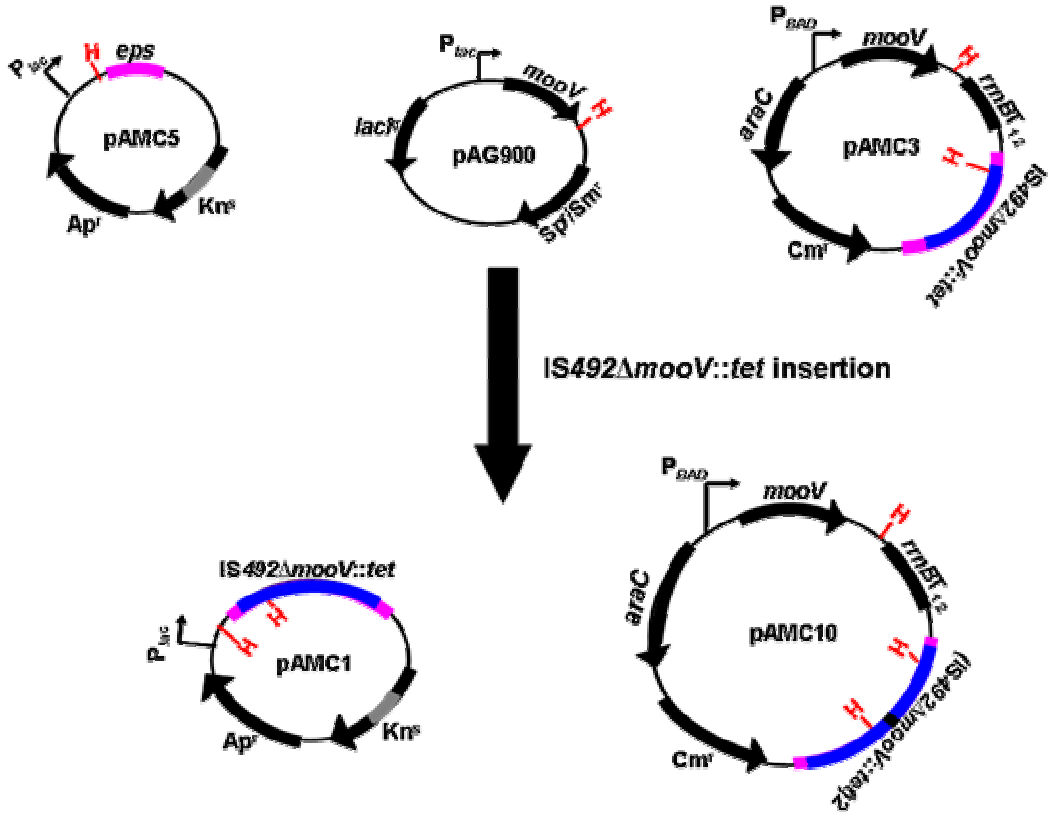
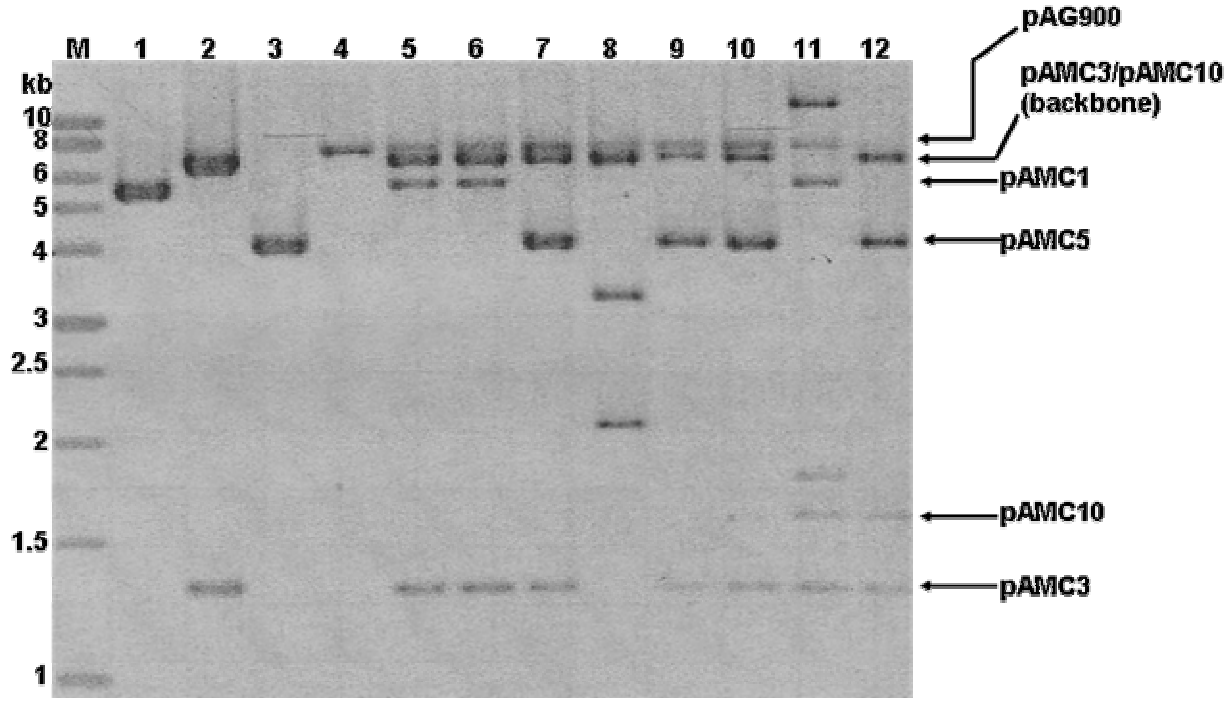

B
1 2 M 3 4 5
4kb
2kb
1.5kb
C
1 2 3 4 M 5 6
200bp
600bp

Figure 2.3. Confirmation of IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid in Tet^r transconjugants. (A) Illustration of EcoRI restriction and PCR analysis to detect IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid. Plasmid DNA was isolated from Tet^r transconjugants (Table 2.2, Experiment III) and used to transform naïve *E. coli* strain TOP10. (see Materials and Methods). (B) EcoRI digests of plasmids isolated from Ap^r Tet^r transformants. Lane 1: pAMC1 (IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid); Lanes 2-5: Plasmids isolated from Ap^r Tet^r transformants. Lane M: Promega 1-kb marker. 1.8-kb EcoRI fragment is IS*492*Δ*mooV*::*tet*; 4-kb EcoRI fragment is pCR2.1 backbone. (C) Plasmids isolated from Ap^r Tet^r transformants were used as template in insertion PCR analysis (Top panel) using primers ISL.66 and M13R to detect the left junction of IS*492*Δ*mooV*::*tet* insertion into the *eps* target plasmid, producing a 214-bp product. Lane 1: pAMC1 (positive control); Lane 2: pAMC5 (negative control); Lanes 3-6:Ap^r Tet^r transformants from Lanes 2-5 in (A). Lane M: Promega 100-bp marker. Bottom panel: Template control PCR. Primers PCRAP.1 and PCRAP.2 were used to amplify a 600-bp fragment from the β-lactamase gene on the pCR2.1 backbone (pAMC1 and pAMC5).


pAMC5
Ap^r
Kn^R
eps
pAG900
lacI^q
mooV
Sp/Sm^r
pAMC3
araC
Cm^r
rnfBT 1-2
IS492ΔmooV::tet
IS492ΔmooV::tet insertion
pAMC1
pAMC10

B


M
1
2
3
4
5
6
7
8
9
10
11
12
kb
10
8
6
5
4
3
2.5
2
1.5
1
pAG900
pAMC3/pAMC10
(backbone)
pAMC1
pAMC5
pAMC10
pAMC3

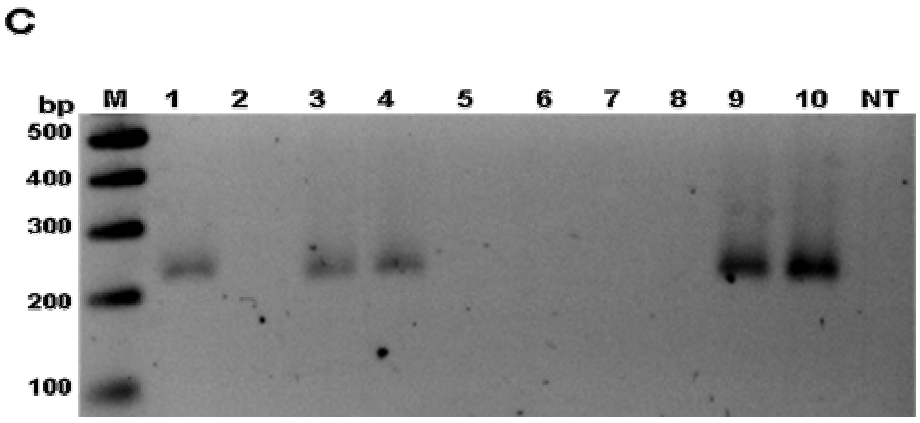


Figure 2.4. Characterization of IS*492*Δ*mooV*::*tet* insertions from Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates. (A) Schematic of plasmids used in insertion experiments. Cartoon illustrates IS*492*Δ*mooV*::*tet* single and tandem insertion into the *eps* target and IS*492*Δ*mooV*::*tet* donor plasmid, creating pAMC1 and pAMC10, respectively. H = HindIII restriction sites. Gray bar in Kn^r gene in pAMC5 represents inactivation of the gene via a NcoI digest and fill-in reaction (see Materials and Methods). Plasmids not drawn to scale. (B) Restriction analysis of plasmids isolated from Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates. Plasmid DNA was isolated from Ap^r Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates following 48 hr incubation at 32°C (Table 2.4 line 1), and digested with HindIII to analyze insertion products. Lane M: Promega 1-kb marker. HindIII digested control plasmids and their expected sizes are as follows: Lane 1: pAMC1 (IS*492*Δ*mooV*::*tet* inserted into the *eps* target plasmid) 5.6-kb + 0.134-kb (not shown); Lane 2: pAMC3 (IS*492*Δ*mooV*::*tet* donor) 6.7-kb + 1.3-kb; Lane 3: pAMC5 (*eps* target) 4-kb; Lane 4: pAG900 (*mooV* expression vector) 8-kb. Lanes 5-12: Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates. Lanes 7-8: 10 μM IPTG in selective media; Lanes 9-10: 0.2% glucose in selective media: Lanes 11-12: 10 μM IPTG and 0.2% glucose in selective media. The additional 1.8-kb fragment present in Lanes 11 and 12 was suggestive of a tandem IS*492*Δ*mooV*::*tet* insertion into the donor plasmid (creating pAMC10) and was further characterized. (C) PCR assay for insertion

of IS*492*Δ*mooV*::*tet* into the *eps* target sequence. Plasmid DNA in panel B was used as template in insertion PCR analysis using primers M13R and ISL.66 (see Materials and Methods, Figure 2.3 A). Lane M: Promega 100-bp marker; Lane 1: positive control pAMC1 (expected PCR product = 214-bp); Lane 2: negative control pAMC5; Lanes 3-10: Tet^r DH5α/pAMC3 pAMC5 pAG900 isolates in Panel A Lanes 5-12; NT = no template control.

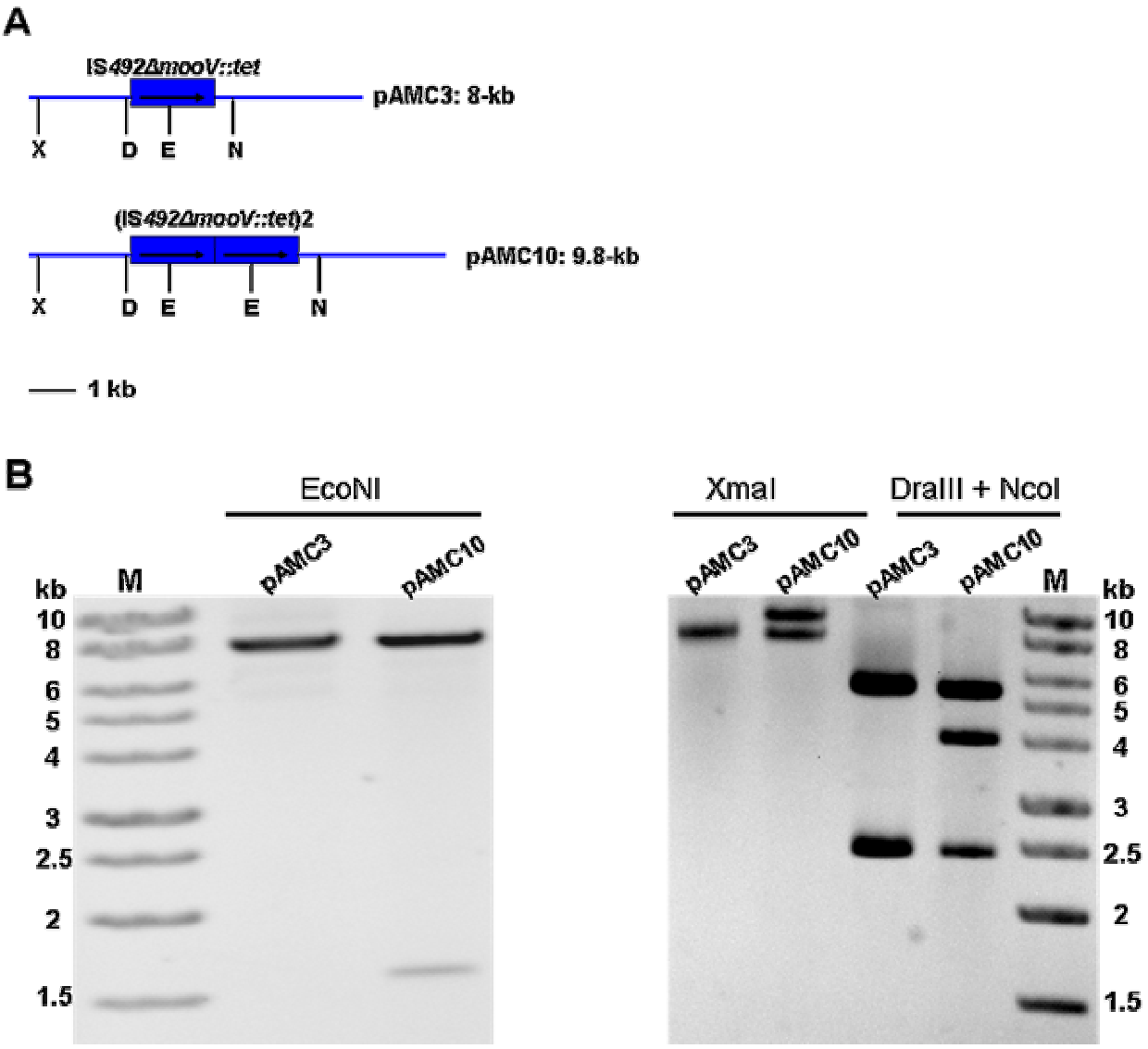
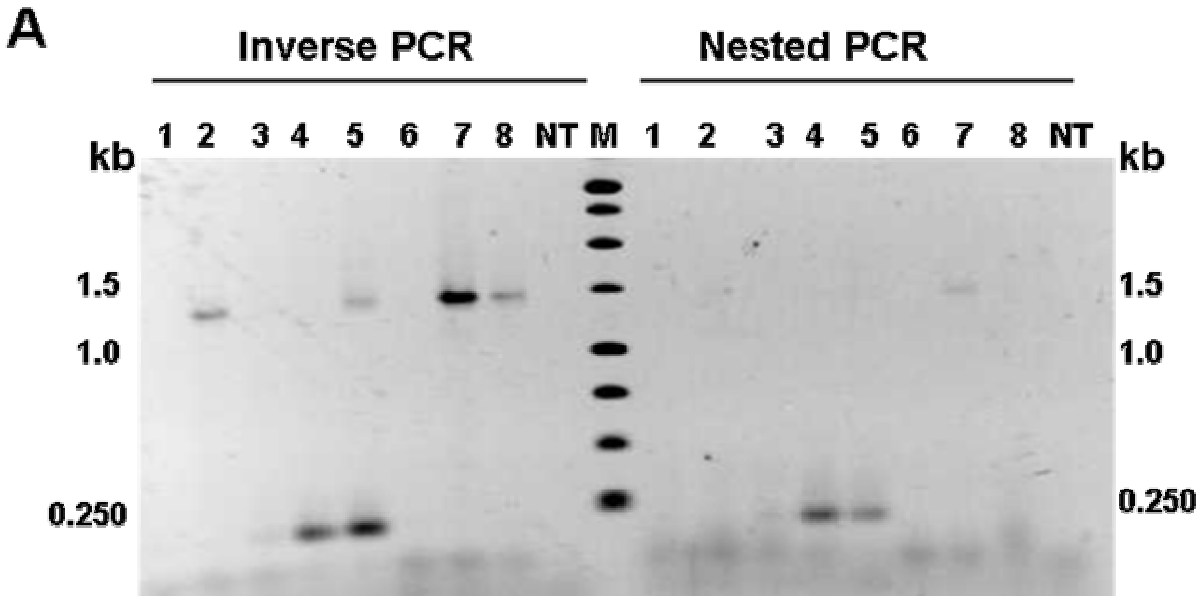


Figure 2.5. Characterization of tandem IS*492*Δ*mooV*::*tet* insertions. (A) Restriction profiles for pAMC3 (donor plasmid, single IS*492*Δ*mooV*::*tet* insertion) and pAMC10 (tandem IS*492*Δ*mooV*::*tet* insertion). Arrows represent orientation of IS*492*Δ*mooV*::*tet*. E = EcoNI; X = XmaI; D = DraIII; N = NcoI. Expected restriction profiles for pAMC3: EcoNI: 8-kb; XmaI: 8-kb; DraIII + NcoI: 2.6-kb + 5.4-kb. Expected restriction profiles for pAMC10: EcoNI: 8-kb + 1.8-kb; XmaI: 9.8-kb; DraIII + NcoI: 4.4-kb + 5.4-kb (B) Agarose gel electrophoresis of pAMC3 and pAMC10 digested with EcoNI, XmaI, and DraIII + NcoI. Lane M: Promega 1-kb marker.


A
Inverse PCR
Nested PCR
kb
1 2 3 4 5 6 7 8 NT M 1 2 3 4 5 6 7 8 NT
kb
1.5
1.0
0.250
1.5
1.0
0.250

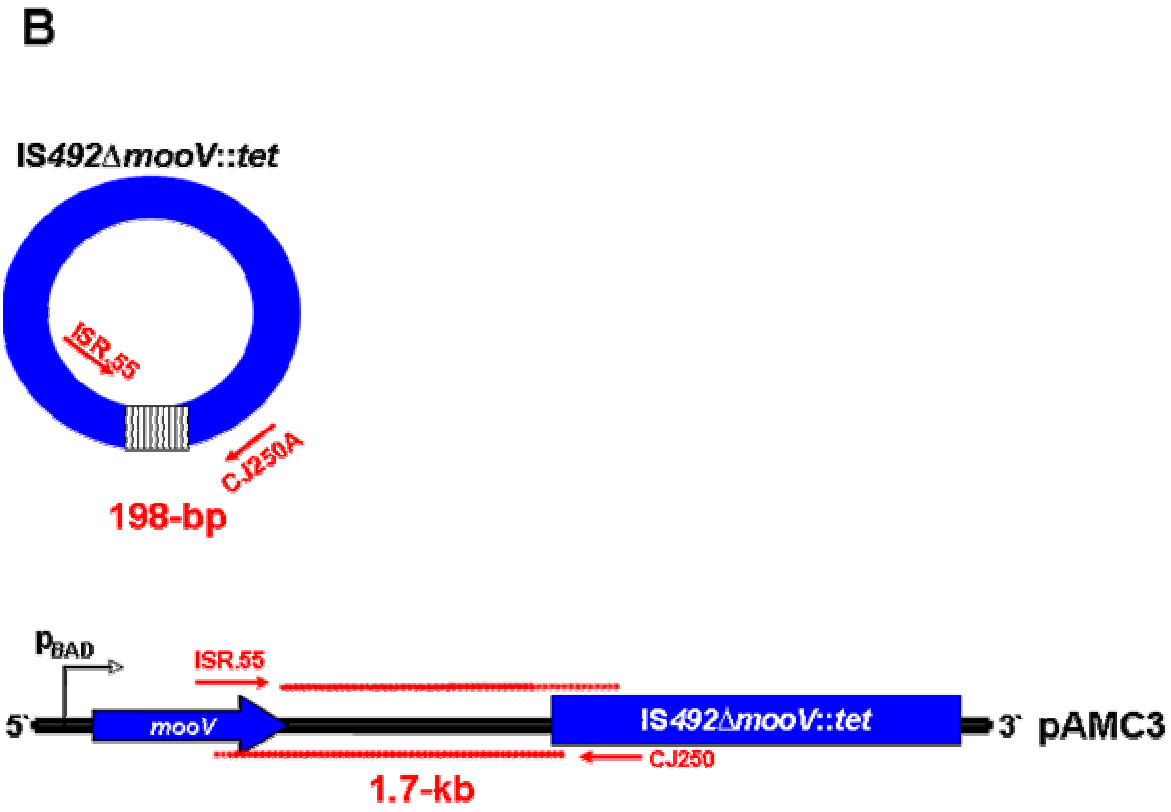

B
IS492ΔmooV::tet
ISR.55
CJ250A
198-bp
pBAD
ISR.55
5`
mooV
IS492ΔmooV::tet
3`
pAMC3
CJ250
1.7-kb

Figure 2.6. Screening of IS*492*Δ*mooV*::*tet* chromosomal insertions in Tet^r *E. coli* isolates. (A) Agarose gel electrophoresis of PCR products generated from inverse and nested PCR analysis (see Materials and Methods). Chromosomal DNA was isolated from Tet^r DH5α/pAMC3pAMC5pAG900 isolates (and one Tet^r TOP10/pAMC3pAMC5pAG900) selected for further analysis and naïve tetracycline sensitive (Tet^s) DH5α and digested with PstI, diluted, ligated and used as template in inverse PCR. PCR products from the inverse PCR assays were used as template for the nested PCR assay. Lane 1: naïve Tet^s DH5α negative control; Lane 2: Tet^r DH5α (see Figure 2.4 B lane 7); Lane 3:Tet^r DH5α (see Figure 2.4 B lane 8); Lane 4:Tet^r DH5α; Lane 5: Tet^r TOP10; Lanes 6-7: Tet^r DH5α; Lane NT: no template control; Lane M: Promega 1-kb marker. (B) Summary of sequence analysis of PCR products generated from inverse PCR. Lane 4: 198-bp IS*492*Δ*mooV*::*tet* circle junction product, resulting from an excised extrachromosomal IS*492*Δ*mooV*::*tet* circle or a tandem IS*492*Δ*mooV*::*tet* insertion (not shown). The black and white stripped box represents the 5-bp spacer sequence 5'-CTTGT-3'; Lane 6: 1.7-kb fragment amplified from the IS*492*Δ*mooV*::*tet* donor plasmid pAMC3. Figure not drawn to scale.

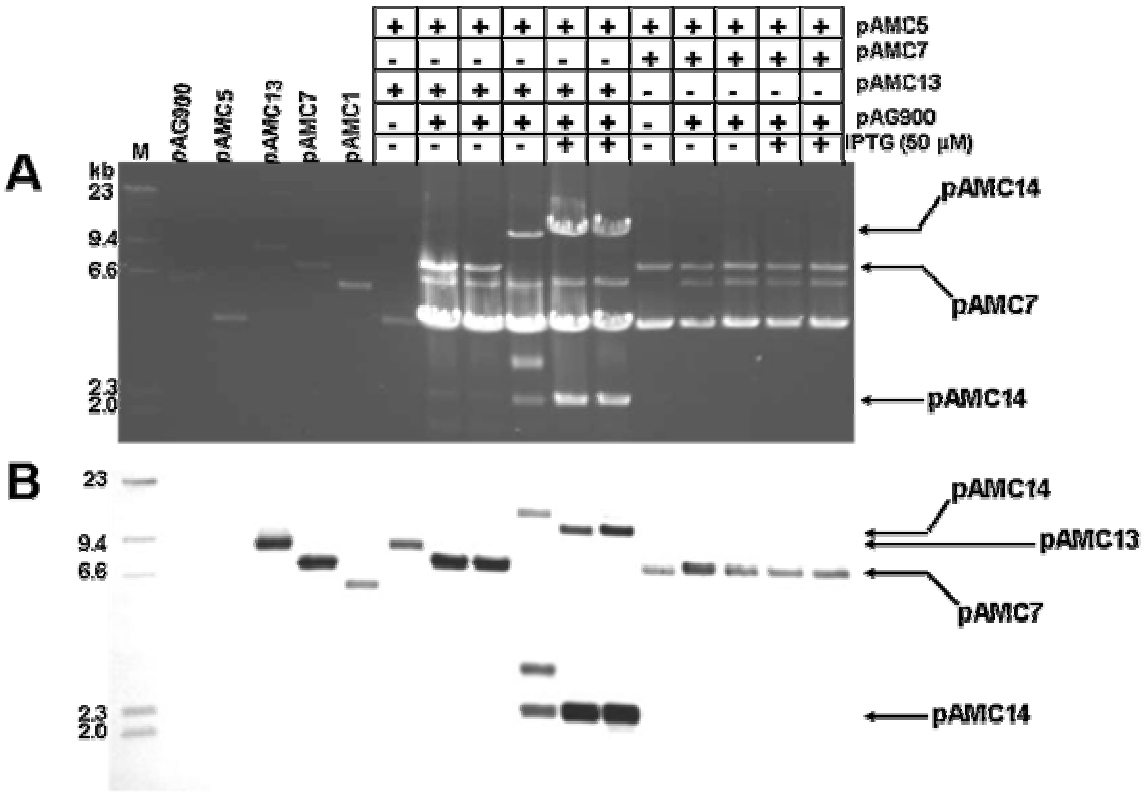


Figure 2.7. Characterizing *in vivo* (IS*492*Δ*mooV*::*tet*)2 transposition at 37°C. (A) Analysis of plasmid insertion products. DH5α containing pAMC13 [(IS*492*Δ*mooV*::*tet*)2] and pAMC5 and DH5α containing pAMC7 (IS*492*Δ*mooV*::*tet*) and pAMC5 were transformed with pAG900, and after outgrowing cells for 1 h at 37°C, cells were plated on LB/$Cm_{34}Ap_{80}Sp_{50}$ +/- 50 µM IPTG, and incubated overnight at 37°C. Plasmid DNA was isolated from overnight cultures of individual randomly selected Cm^r Ap^r Sp^r transformants and digested with DraIII to analyze insertion products. Digested products were electrophoresed on a 0.8% agarose gel and stained with EtBr. Control plasmids and their expected sizes are as follows: pAG900 [6.5-kb + 0.6-kb (not shown) + 0.4-kb (not shown)]; pAMC5 (4-kb); pAMC13 (9-kb); pAMC7 (7.2-kb); pAMC1 (5.8-kb) The 10.7-kb and 2.3-kb bands are from a pAMC13 and pAMC5 cointegrate, designated pAMC14. Panel B: Southern analysis of DH5α/ pAMC13 pAMC5 pAG900 and DH5α/ pAMC7 pAMC5 pAG900 transformants. Total DNA (chromosomal and plasmid) was isolated from the above Cm^r Ap^r Sp^r transformants, digested with DraIII, and after electrophoresis on a 0.8%

agarose gel was transferred to a nylon membrane. DNA was probed with a 1.3-kb DIG labeled tetracycline resistance gene (DIG High Prime DNA Labeling and Detection Kit, Roche) to probe for IS*492*Δ*mooV*::*tet* insertions. M: DIG-labeled Molecular Weight Marker II (Roche)

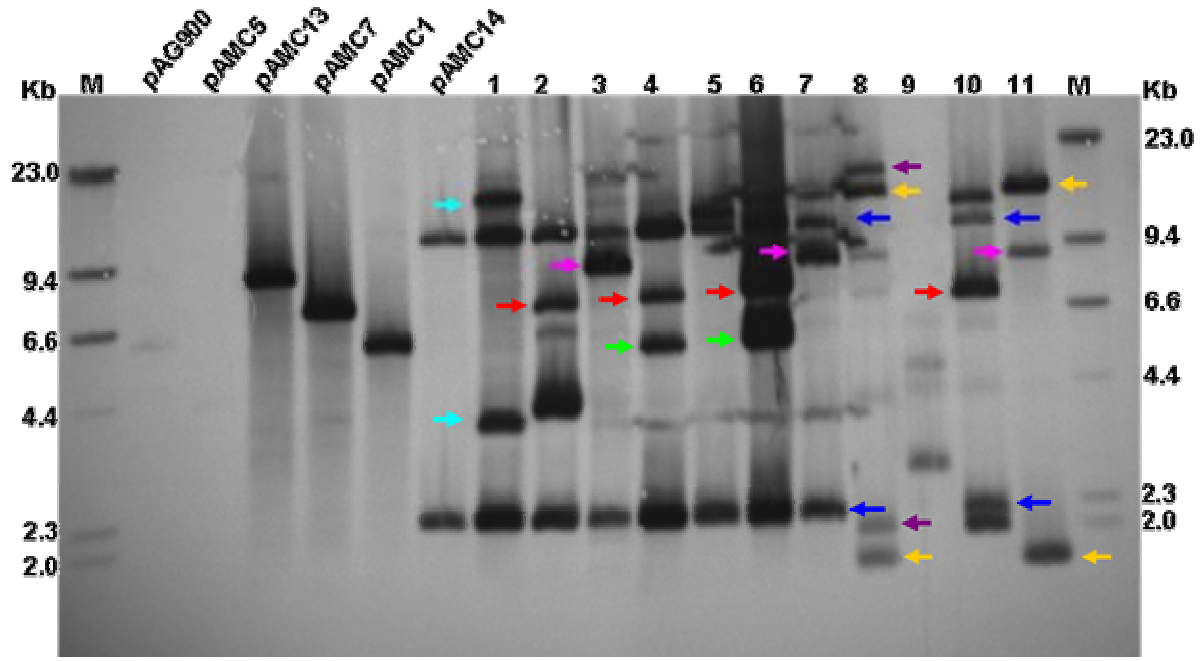


Figure 2.8. Characterizing other tandem IS*492*Δ*mooV*::*tet* mediated transposition products. Selected plasmid samples (#1-11) containing other types of tandem IS*492*Δ*mooV*::*tet* transposition products were digested with DraIII and subjected to Southern blot Analysis, using the 1.3-kb tetracycline resistance gene to probe for IS*492*Δ*mooV*::*tet* (DIG High Prime DNA Labeling and Detection Kit, Roche). Expected sizes for DraIII digested control plasmids are as follows: *mooV* expression vector pAG900 6.5-kb + 0.6-kb + 0.4-kb (negative control); *eps* target pAMC5 4-kb (negative control); tandem IS*492*Δ*mooV*::*tet* donor pAMC13 9-kb; single IS*492*Δ*mooV*::*tet* excision from pAMC13, pAMC7 7.2-kb; single IS*492*Δ*mooV*::*tet* insertion into *eps* target pAMC5, pAMC1 5.8-kb; pAMC13/pAMC5 cointegrate pAMC14 10.7-kb + 2.3-kb. MooV-mediated transposition products from tandem IS*492*Δ*mooV*::*tet* include single excision (**red arrow**), pAMC13/pAMC5 cointegrates (**blue arrows**), and single insertion into *eps* target (**lime green arrow**). No changes in pAMC13 (**pink arrow**) and other transposition products, including pAMC13/pAG900 cointegrates (**orange arrows**), pAMC13/pAMC7 (**purple arrows**)

and pAMC13/pAMC1 (**turquoise arrows**) were identified based on additional restriction analyses and antibiotic resistance markers on plasmid backbones. M, DIG-labeled Molecular Weight Marker II (Roche)

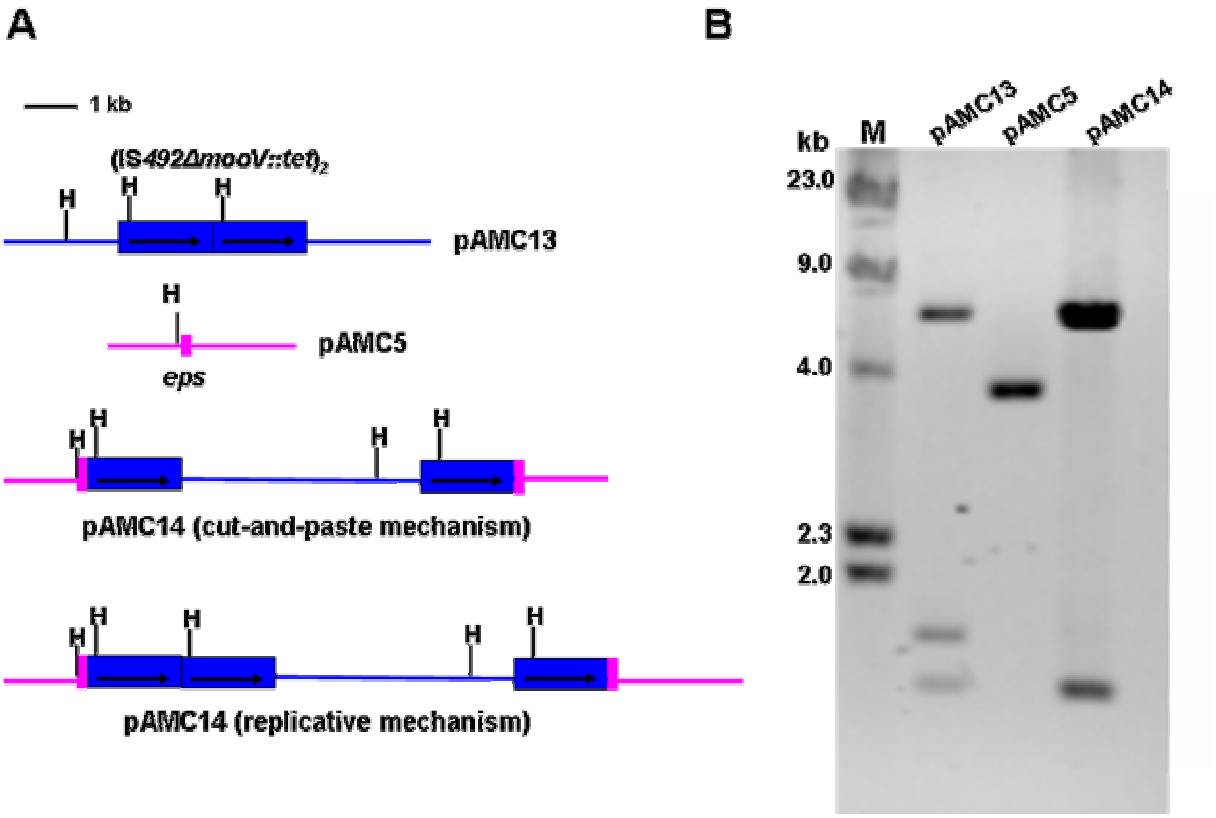
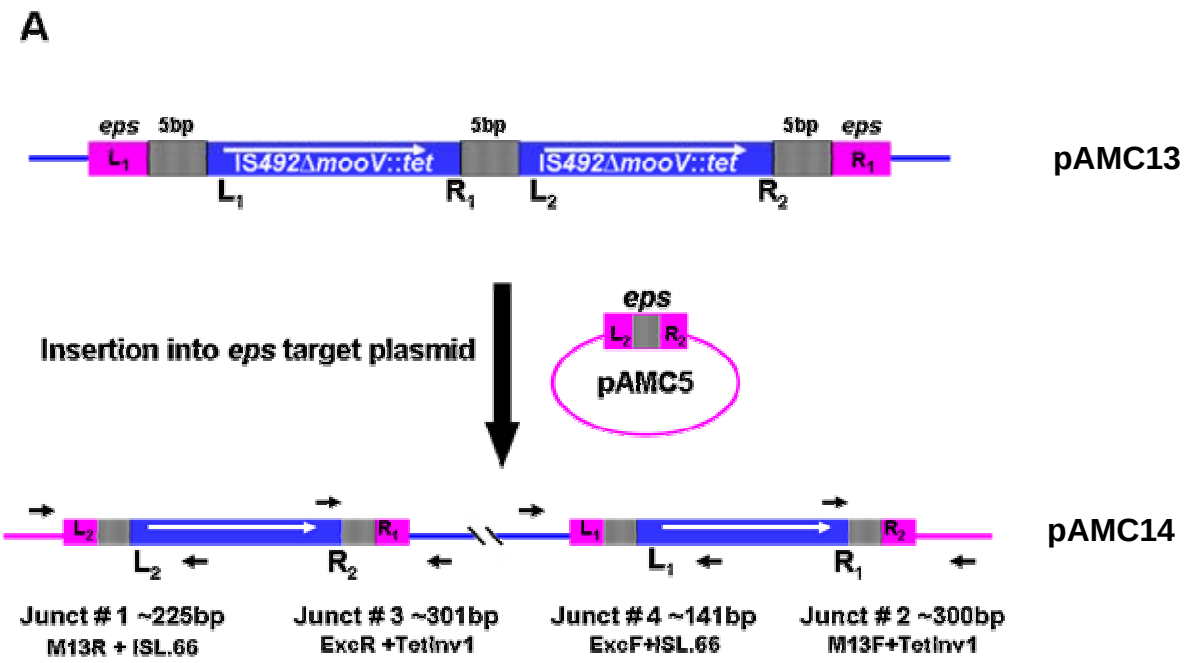


Figure 2.9. Characterizing pAMC5/pAMC13 fusion plasmids. (A) Restriction profiles for plasmids pAMC5, pAMC13, and pAMC14. Arrows represent orientation of IS*492*Δ*mooV*::*tet*. Expected HindIII profiles for control plasmids: pAMC13 = 6-kb + 1.8-kb+ 1.2-kb; pAMC5 = 4-kb. Predicted pAMC14 profile if pAMC13 inserted into pAMC5 via a cut-and-paste transposition mechanism (one IS*492*Δ*mooV*::*tet* element at each pAMC13/pAMC5 plasmid junction): 6.6-kb + 6.1-kb + 1.2-kb + 0.134-kp. Predicted pAMC14 profile if pAMC13 inserted into pAMC5 via a replicative transposition mechanism (two IS*492*Δ*mooV*::*tet* elements at one junction, one IS*492*Δ*mooV*::*tet* element at the other junction) : 6.6-kb + 6.1-kb + 1.8-kb + 1.2-kb + 0.134-kb. (B) HindIII digests of pAMC13, pAMC5 and a pAMC14 fusion plasmid isolated from a 37°C/37°C (IS*492*Δ*mooV*::*tet*)2 transposition experiment (Table 2.6, Figure 2.7). Lane M: λ-HindIII marker (New England Biolabs).



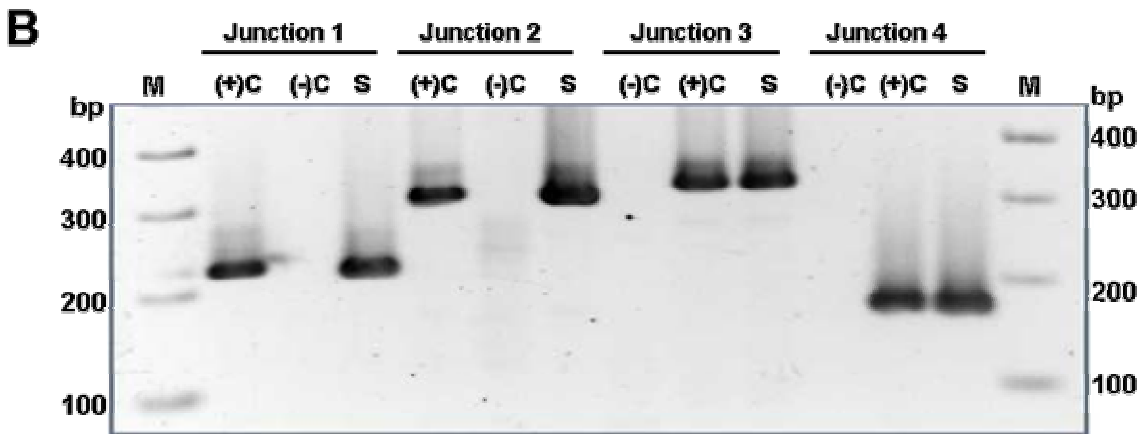


Figure 2.10. PCR analysis of pAMC14 cointegrate plasmids. (A) MooV mediated insertion of the tandem IS*492*Δ*mooV*::*tet* plasmid (pAMC13) into the *eps* target plasmid (pAMC5) creating the cointegrate plasmid, pAMC14. The 5-bp target sequence 5'-CTTGT-3' is represented as a striped box. The left (L) and right (R) sides of *eps* and ends of IS*492*Δ*mooV*::*tet* in each plasmid are indicated. Cointegrate junctions between pAMC13 and pAMC5 were confirmed in PCR analysis using 4 different primer combinations: PCR primer sets used for junctions #1 and #2

anneal to the IS*492*Δ*mooV*::*tet*/pAMC5 backbone insertion junction, and PCR primer sets used for junctions #3 and #4 anneal to the IS*492*Δ*mooV*::*tet*/pAMC13 backbone insertion junction. Only cointegrates between these two plasmids would have all four junctions. (B) Representative pAMC14 cointegrate junction PCR assay (for one sample, S). Positive/Negative controls for each junction PCR assay are as follows: pAMC1 (single IS*492*Δ*mooV*::*tet* insertion in pAMC5) positive control [(+)C] for Junction #1 (M13R + ISL.66) and Junction #2 (M13F +TetInv1), negative control [(-)C] for Junction #3 (ExcR + TetInv1) and Junction #4 (ExcF +ISL.66); pAMC7 (single IS*492*Δ*mooV*::*tet* insertion in pAMC13 backbone) negative control [(-)C] for Junctions #1 and Junction #2, positive control [(+)C] for Junction #3 and Junction #4.

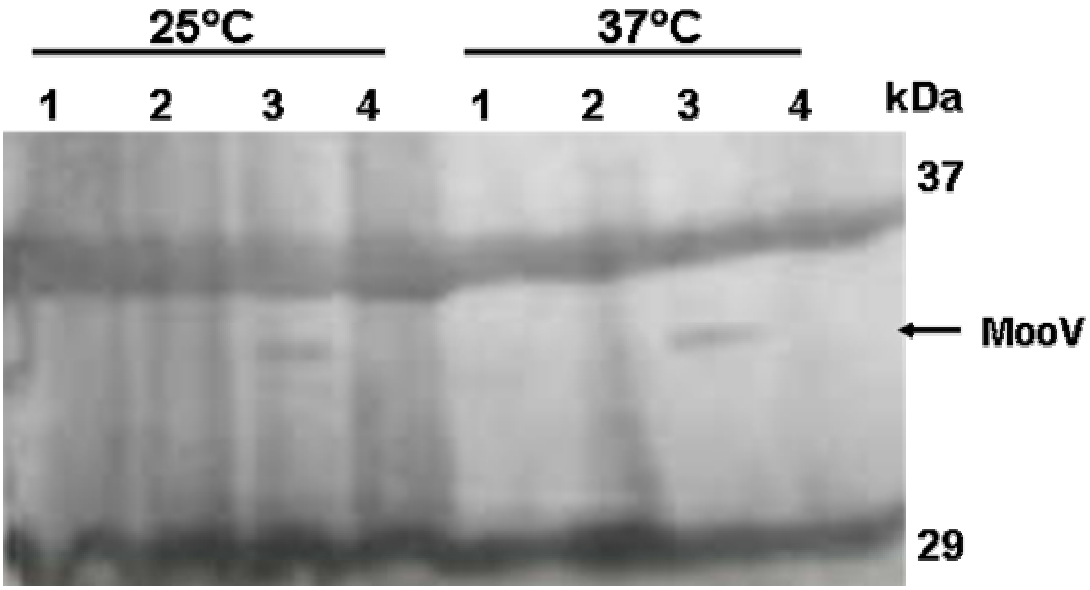


Figure 2.11. MooV expression in *E. coli* at 37°C versus 25°C. *E. coli* stain DH5α was transformed with pAG900 or pAG900Δ*mooV*, outgrown at 37°C or 25°C, and cells were inoculated into LB/Sp_{50} cultures and grown with aeration until turbid (37°C, overnight; 25°C, 72 h). Turbid cultures were diluted to a final O.D. $_{600}$ of 0.1 in 6 ml LB/Sp_{50} (+/- 50 μM IPTG), and grown with aeration at 37°C or 25°C to a final O.D. 600 0.5-0.7. Proteins were isolated from harvested cells, fractionated on a 12% SDS-PAGE gel, and subjected to Western blot analysis, using anti-MooV antisera as the primary antibody. MooV is indicated with an arrow. Lanes 1: DH5α/pAG900Δ*mooV* + IPTG; lanes 2: DH5α/pAG900Δ*mooV*; lanes 3 DH5α/pAG900 + IPTG; lane 4: DH5α/pAG900. Experiment performed in triplicate with comparable results.

CHAPTER 3

Development of a suicide vector-based assay to characterize IS*492* transposition in *Escherichia coli*

Abstract

A suicide vector based assay using a new IS*492* construct, IS*492*Δ*mooV*::*dfr*, was developed to investigate further the role of temperature and target specificity in IS*492* transposition in *E. coli*. The suicide vector based assay could potentially be used to study the role of known recombination host factors (IHF, Fis, H-NS, and HU) in IS*492* insertion. Insertion frequencies, as measured by frequency of trimethoprim resistant (Tp^r) isolates, were up to 37-fold higher at 25°C compared to 37°C, however analysis of Tp^r isolates from all 37°C and 25°C experiments revealed the occurrence of RecA- and MooV-independent plasmid fusion products with the suicide vector. Single IS*492*Δ*mooV*::*dfr* insertions into an *eps* target plasmid were only observed at 25°C, which correlate with previous observations of IS*492* transposition in *E. coli*. Different approaches to study IS*492* insertion in *E. coli* are discussed.

3.1. Introduction

Efficient assays are required to characterize transposition *in vivo*. In general, mobile elements transpose at a low frequency to prevent deleterious effects in the host cell, and characterizing *in vivo* transposition requires selecting or screening the phenotypes that are associated the movement of the element. The most common techniques to study transposition within a cell include monitoring antibiotic resistance associated with the element as it moves from one replicon to another, and screening for expression of a promoter-less *lacZ* gene within

an element that inserts downstream of an actively transcribed promoter.[9] To prevent false-positives in these assays, the reporter transposable elements are introduced into the host cell on replication-conditional (suicide) vectors so that the donor plasmid is not maintained; insertions into the chromosome or a stable plasmid can be selected or a "mating out" assay may be used to detect insertion of the element into a conjugative plasmid after transfer to a different recipient cell. [9]

Previous characterization of IS*492* insertion in *Escherichia coli* used the construct IS*492*Δ*mooV*::*tet* in a "promoter-trap" assay, where the expression of the promoter-less tetracycline resistance (Tet^r) gene was dependent on IS*492*Δ*mooV*::*tet* insertion downstream of an actively transcribed promoter in the correct orientation (Chapter 2). Using this approach, it was shown that MooV-mediated IS*492*Δ*mooV*::*tet* insertion displays target site specificity, and that insertion of monomer IS*492*Δ*mooV*::*tet* was sensitive to lower temperatures. Accurate insertion frequencies, however, could not be determined due to the temperature sensitive expression of Tet^r from the putative *mooV* promoter upstream of the Tet^r gene in IS*492*Δ*mooV*::*tet* (Chapter 2). This finding necessitates the development of a new approach to characterize IS*492* transposition in *E. coli* to elucidate further the recombination mechanism, and to identify DNA sequences and protein factors required for insertion.

The present study reports the use of a new IS*492* construct, IS*492*Δ*mooV*::*dfr*, and a suicide plasmid delivery system to examine IS*492* insertion in *E. coli*, to re-evaluate the role of temperature and target selection in IS*492* transposition. Single IS*492*Δ*mooV*::*dfr* insertion into the *eps* target plasmid was observed at 25°C, and was dependent on MooV. However, low frequency MooV-independent fusions of the IS*492*Δ*mooV*::*dfr* donor plasmid with the *eps* target plasmid and MooV expression vector were observed at 37°C and 25°C. These observations

support the previous findings that MooV-meditated IS*492* insertion into the *eps* site is sensitive to temperature. Modifications to this assay to prevent the detection of false-positive insertions are also discussed.

3.2. Materials and methods

Bacterial strains, plasmids, and oligonucleotides. All *E. coli* strains, plasmids, and oligonucleotides used in this study are listed in Table 3.1. *Pfu* DNA polymerase (Stratagene) was used to amplify PCR products generated for cloning, and all *Pfu* amplified DNA was sequenced. All DNA enzymes, restriction enzymes, and ligases used in construction of plasmids were obtained from New England Biolabs (NEB).

Characterizing IS*492*Δ*mooV*::*dfr* insertion in *E. coli*. $CaCl_2$ competent DH5α containing pAMC5 was transformed with pAG900 or pAG900ΔmooV and outgrown 1 h at 37°C. 200 µl of transformed cells were inoculated into 1.8 ml Luria-Bertani (LB) broth containing 80 µg/ml ampicillin (Ap_{80}) and 50 µg/ml spectinomycin (Sp_{50}), incubated with aeration overnight at 37°C, along with 2 ml LB broth supplemented with 50 µg/ml kanamycin (Kn_{50}) and 10 µg/ml trimethoprim (Tp_{10}) for helper strain BW23473/pEVS104 and donor strain BW23474/pAMC21, respectively. Cells were harvested via centrifugation, resuspended in 500 µl LB, and 100 µl of the donor, helper, and recipient strains were mixed together for total 300 µl conjugation mix, and spotted onto LB agar. Conjugations performed at 37°C were incubated overnight, and conjugations performed at 25°C were incubated 48 h. Cells were harvested from conjugation plates using a sterile wooden stick, and resuspended in 1 ml LB. Ten-fold serial dilutions were plated on LB/$Ap_{80}Sp_{50}$ to select for recipient strains and LB/$Ap_{80}Sp_{50}Tp_{10}$ to select for recipient strains with an IS*492*Δ*mooV*::*dfr* insertion. 50 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.2% glucose (glu) were included in LB/$Ap_{80}Sp_{50}Tp_{10}$ plates to induce transposase

expression from P_{tac} and decrease transcription across the *eps* target from P_{lac}, respectively. Plates were incubated overnight at 37°C, or 72 h at 25°C. Plasmid DNA was isolated from selected Ap^r Sp^r Tp^r colonies and subjected to restriction (BamHI) and insertion PCR analysis (previously described in Chapter 2) to screen for IS*492*Δ*mooV*::*dfr* insertion.

3.3. Results

Characterizing IS*492*Δ*mooV*::*dfr* insertion in *E. coli*. Previous studies characterizing IS*492* insertion in *E. coli* suggests MooV targets the 5-bp 5'-CTTGT-3' sequence located within the *P. atlantica epsG* sequence and at the ends of IS*492*, and that the insertion of the single/monomer element is sensitive to lower temperatures (Chapter 2). Insertion frequencies from the single IS*492* (IS*492*Δ*mooV*::*tet*) element could not be determined accurately due to high background expression of tetracycline resistance from IS*492*Δ*mooV*::*tet* at lower temperatures in the absence of MooV (Chapter 2). In order to decrease the probability of detecting false-positives and to more accurately determine the relative IS*492* insertion frequency in *E. coli*, an insertion assay using an IS*492* suicide donor vector was developed (Figure 3.1). The IS*492* element IS*492*Δ*mooV*::*dfr*, which contains the trimethoprim resistance cassette with its own promoter and Shine-Dalgarno sequence, was cloned into the R6k *ori* plasmid pEVS94. The resulting plasmid, pAMC21, requires the pi protein in order to replicate, and thus would be lost in a cell lacking the *pir* gene. Tp^r expressed from IS*492*Δ*mooV*::*dfr* would only be observed in DH5α (*pir*$^-$) if MooV mediates single insertion of IS*492*Δ*mooV*::*dfr* into a target plasmid or cointegrate formation between pAMC21 and a target plasmid.

In order to introduce pAMC21 into DH5α containing the *eps* target pAMC5 and MooV expression vector pAG900, a tri-parental mating assay was performed. Conjugation and selective insertion temperatures were performed at 37°C and 25°C to determine the effects of

temperature on MooV-mediated IS*492*Δ*mooV*::*dfr* insertion, and to examine the insertion products. Low frequency (10^{-8}) Tp^r was observed in 37°C experiments with and without MooV being present in the DH5α recipient strain (Table 3.2). Although low Tp^r frequencies (~10^{-8}) were observed at 25°C in the absence of MooV, in the presence of MooV Tp^r frequencies ranged from 2-fold (+ 0.2% glucose and 50 μM IPTG) to 37-fold (+ 50 μM IPTG) higher compared to the corresponding experiment performed at 37°C. Similar Tp^r frequencies were observed in additional experiments that included +/- 50 μM IPTG (not shown). BamHI, which has a single recognition site in all plasmids used and no recognition sites in IS*492*Δ*mooV*::*dfr*, was used to digest plasmids isolated from Tp^r colonies. Restriction analysis of plasmids isolated from 25°C experiments performed in the presence of MooV revealed three types of insertion products (Figure 3.2): a single 5.3-kb fragment indicative of single IS*492*Δ*mooV*::*dfr* insertion into pAMC5; 3-kb and 4.3-kb bands, indicative of a plasmid fusion between pAMC21 and pAMC5 (Figure 3.2, lane 5); 1.6-kb and 9.5-kb fragments indicative of a plasmid fusion between pAMC21 and pAG900 (Figure 3.2, lane 8). Insertion PCR analysis confirmed single insertion into the *eps* target plasmid in Tp^r isolates that contained the additional 5.3-kb BamHI fragment indicative of single insertion (not shown). The percentage of single insertions into the *eps* target plasmid in 25°C experiments with MooV *in trans* ranged from 14% - 80% with the addition of IPTG, and 0% - 30% without IPTG. The addition of glucose in the medium to decrease transcription across the *eps* target sequence from P_{lac} (Table 3.2) did not increase the occurrence of single IS*492*Δ*mooV*::*dfr* insertion into pAMC5, with only 13% single insertions with the addition of glucose, and no single insertions detected with the addition of glucose and IPTG.

The plasmid fusions between pAMC21 and pAMC5, and pAMC21 and pAG900 were also observed in the 37°C experiments with MooV (Figure 3.3) and without MooV (not shown)

and the 25°C experiments (without MooV) (not shown); all pAMC21/pAMC5 and pAMC21/pAG900 fusion plasmids had the same BamHI restriction profile, and no single insertion of IS*492*Δ*mooV*::*dfr* into the *eps* target plasmid was observed in 37°C or 25°C experiments without MooV provided *in trans*. These results indicate that the pAMC21 plasmid fusions were independent of MooV, and were the result of RecA independent recombination in the DH5α recipient strain. Control conjugation experiments mating-in pAMC21 into naïve DH5α did not result in Tp^r transconjugants, suggesting that the RecA independent recombination was plasmid based. Additionally introducing pAMC21 into DH5α/pAMC5 pAG900 or DH5α/pAMC5 pAG900ΔmooV via transformation did not yield Tp^r transformants, suggesting that the RecA independent pAMC21 plasmid fusions were dependant on how pAMC21 was introduced into the recipient strain (single stranded in conjugation vs. double stranded in $CaCl_2$ transformation). A pAMC21/pAMC5 fusion plasmid was further characterized via sequencing analysis: IS*492*Δ*mooV*::*dfr* is not inserted within the *eps* target sequence; recombination between the two plasmids occurred within a 39-bp region of sequence homology (IS*492*Δ*mooV*::*dfr* on pAMC21 is flanked by 39-bp left and 17-bp right pCR2.1 sequence, the plasmid backbone of pAMC5). Since this recombination occurred in a *recA*1 strain, this recombination event could be due to illegitimate recombination, which has been reported to be mediated by a number of enzymes in *E. coli*, including DNA gyrase (short homology independent recombination) and RecJ (short homology dependent recombination).[133]

3.4. Discussion

A suicide vector approach to study IS*492* insertion in *E. coli* was used to alleviate the detection a false-positives, to determine the relative insertion frequencies, and to validate the role of lower temperatures in single IS*492* insertion. IS*492*Δ*mooV*::*dfr* insertion frequencies, as

measured by the frequency of Tp^r in DH5α, were 2-fold to 37-fold higher at 25°C compared to 37°C when MooV was provided *in trans* (Table 3.2), however analysis of Tp^r isolates from all experiments revealed plasmid recombination products that are independent of RecA and MooV. Although the frequency of RecA and MooV independent recombination is low, the levels of Tp^r without MooV being present prevent the accurate determination of IS*492* insertion frequencies in *E. coli*. Despite the background Tp^r in these experiments, single IS*492*Δ*mooV*::*dfr* insertion into the *eps* target plasmid was observed only at 25°C, an observation that correlates with previous IS*492* studies in *E. coli* (Chapter 2).[88] Since MooV mediates precise excision and circle formation of IS*492* at 37°C[88] and the relative expression and stability of MooV is similar at 37°C versus 25°C (Chapter 2), the above observations along with the data from Chapter 2 suggest that other aspects of IS*492* transposition, including IS*492* transposition intermediates (i.e. the IS*492* circle) or enzymatic properties of MooV, are sensitive to temperature. For future IS*492* insertions assays in *E. coli*, we propose introducing the *epsG* target sequence into the *E. coli* chromosome or to use "mating-out" assays, where the *epsG* target sequence is cloned into conjugative plasmid and IS*492* insertions into the target plasmid is mated-out into a recipient strain.

Table 3.1. Bacterial strains, plasmids, and oligonucleotides

***E. coli* strains**	**Relevant characteristics[a]**	**Source or reference**
DH5α	F⁻ θ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deoR supE44 hsdR17 recA1 end*A*1 gyrA96 thi-1 relA1*	C. Moran
BW23474	Δ*lac-169 robA1 creC510 hsdR514 uidA* (Δ*MluI*)::*pir-116 endA* (*BT333*) *recA1*	116
BW23743	Δ*lac-169 robA1 creC510 hsdR514 uidA* (Δ*MluI*)::*pir endA* (*BT333*) *recA1*	E. Stabb
Plasmids		
pAMC5	134-bp *epsG* sequence from *P. atlantica* strain DB27 (EPS^+)[107] cloned downstream of P_{lac} on pCR2.1; ColE1 *ori*; Ap^r	Chapter 2
pAG900	MooV under control of P_{tac} promoter, pSC101 *ori*; Sp^r Sm^r	88
pAG900ΔmooV	MooV deleted from pAG900, pSC101 *ori*; Sp^r Sm^r	C. Carpenter
pEVS104	*oriT*; R6k *ori*; Kn^r	116
pAMC16	454-bp trimethoprim resistance cassette (*dfr*) with promoter and ribosome binding site, amplified from pJLB1[134] (with primers M13R and M13F primers) digested with EcoRI, and ligated into the MfeI sites of IS*492* on pAG989[88]	This work
pAMC21	IS*492*Δ*mooV*::*dfr* amplified from pAMC16 (with primers EPSL.SpeI and EPSR.PstI) and cloned into the SpeI and PstI sites in pEVS94;[116] *oriT*, R6k *ori*; Tp^r Em^r	This work
Oligonucleotides (5'-3')		
M13F	GTAAAACGACGGCCAG	
M13R	GAGGAAACAGCTATGAC	
EPSL.SpeI	ATTACTAGTCAGGTTTTGGTGGCATTG	
EPSR.PstI	TTACTGCAGCAATAGAGAGCTTGATAG	

[a] Ap^r, ampicillin resistance; Em^r, erythromycin resistance; Kn^r, kanamycin resistance; Sp^r, spectinomycin resistance; Sm^r, streptomycin resistance; Tp^r, trimethoprim resistance.

Table 3.2. Characterizing IS*492*Δ*mooV*::*dfr* transposition using the suicide vector pAMC21

		Insertion Frequencies[a]			
Selective temperature[b]	Recipient strain	$Ap^r Sp^r Tp^r$	$Ap^r Sp^r Tp^r$ + 50 μM IPTG[c]	$Ap^r Sp^r Tp^r$ + 0.2% Glu[d]	$Ap^r Sp^r Tp^r$ + 50 μM IPTG, 0.2% Glu
37°C	DH5α/pAMC5 pAG900	$3.5x10^{-8}$	$1.8x10^{-8}$	$9.2x10^{-8}$	$7.6x10^{-8}$
37°C	DH5α/pAMC5 pAG900ΔmooV	$4.5x10^{-8}$	N/A[e]	N/A[e]	N/A[e]
25°C	DH5α/pAMC5 pAG900	$4.6x10^{-7}$	$6.6x10^{-7}$	$1.2x10^{-6}$	$1.3x10^{-7}$
25°C	DH5α/pAMC5 pAG900ΔmooV	N/A[e]	$1.3x10^{-8}$	$1.7x10^{-7}$	$9.2x10^{-8}$

[a] Insertion frequencies are the relative number of $Ap^r Sp^r Tp^r$ colonies among the number of $Ap^r Sp^r$ colonies

[b] Conjugation and selective temperature for insertion (See Materials and Methods)

[c] 50 μM IPTG present in medium to induce transposase expression from P_{tac}

[d] 0.2% glucose present in medium to decrease transcription from P_{lac} across *eps* target

[e] N/A = non-applicable; no Tp^r colonies isolated

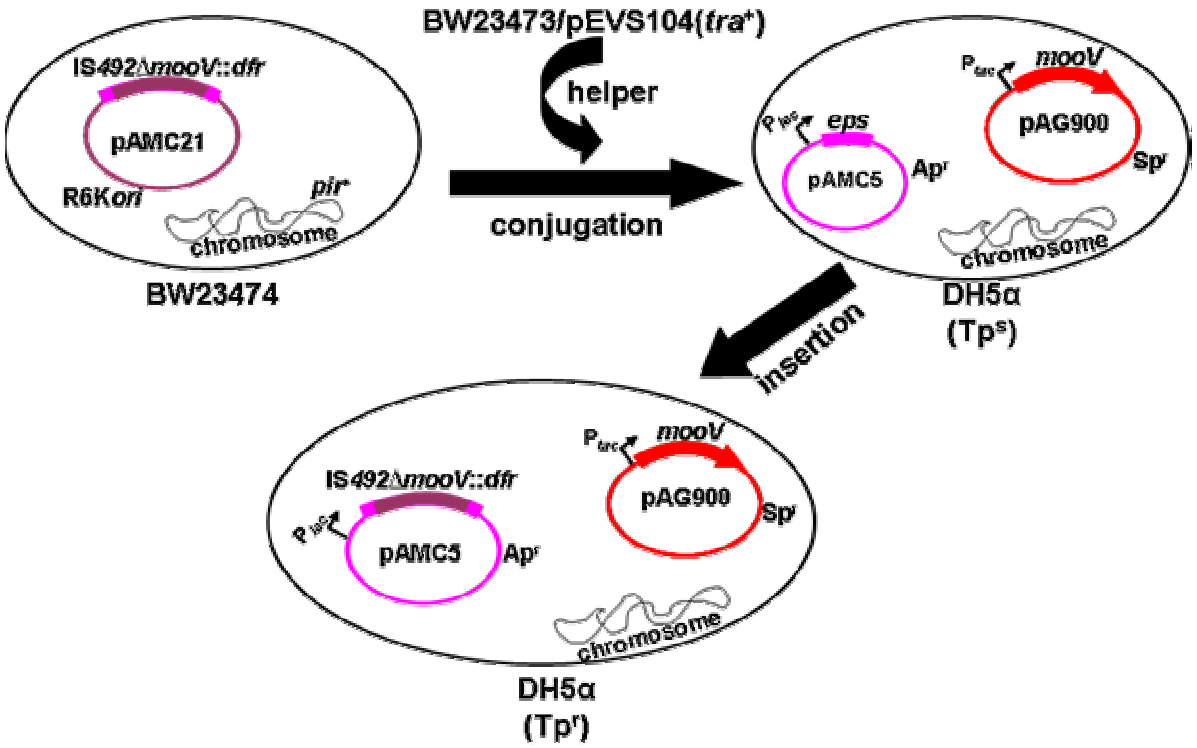


Figure 3.1. Suicide vector approach for studying IS*492* insertion in *E. coli*. IS*492*Δ*mooV*::*dfr* is flanked by 23-bp *eps* sequence on the suicide vector pAMC21, which is maintained in the *pir*+ strain BW23474. pAMC21 is mated into the recipient strain DH5α (*pir*-) containing the *eps* target plasmid pAMC5 and MooV expression vector pAG900 via a tri-parental mating with the helper strain BW23473/pEVS104. MooV-mediated insertion of IS*492*Δ*mooV*::*dfr* into pAMC5 would result in Tp^r DH5α cells.

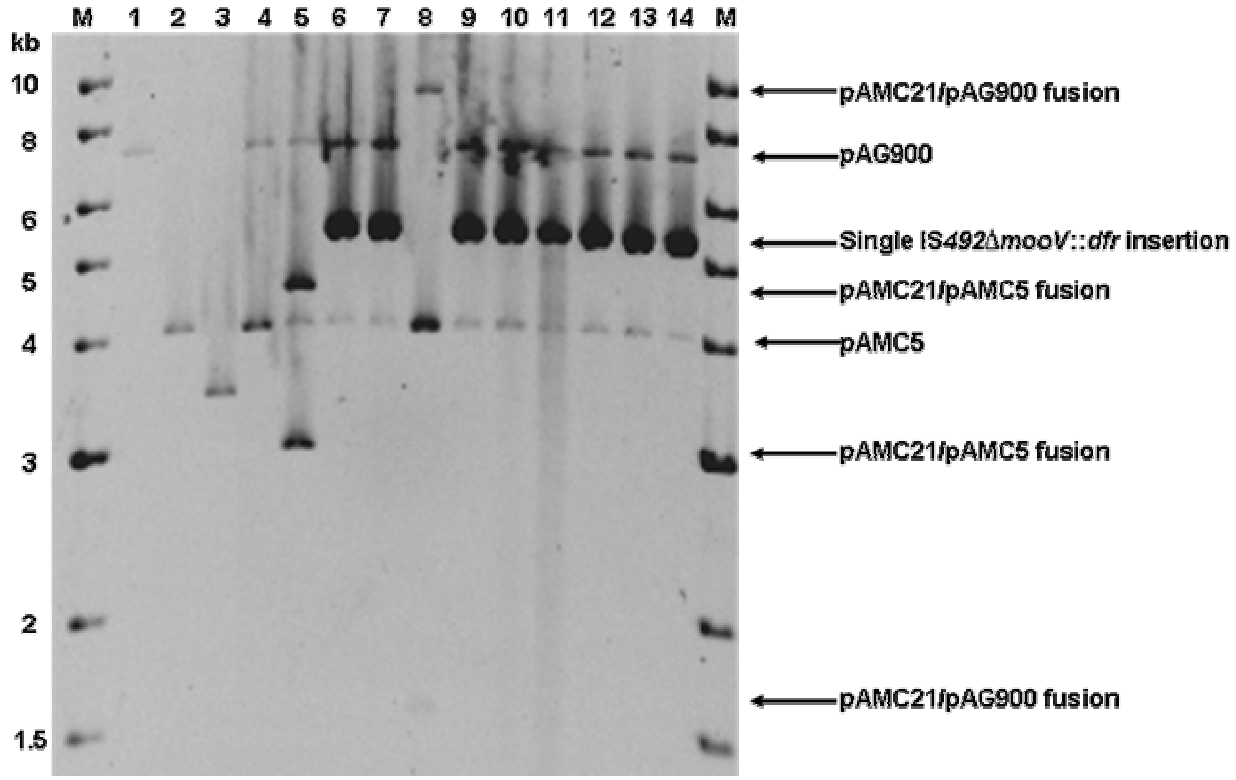


Figure 3.2. Characterizing Tp^r DH5α/pAMC5pAG900 isolates at 25°C. Plasmid DNA was isolated from Tp^r DH5α/pAMC5pAG900 isolates from a 25°C experiment (with 50 μM IPTG present in selective medium) digested with BamHI, and electrophoresed on a 0.8% gel (inverse image is shown). Controls and expected sizes: Lane 1 pAG900, 7.8-kb; Lane 2 pAMC5, 4-kb; Lane 3 pAMC21, 3.3-kb; Lane 4 Tp^s DH5α/pAMC5pAG900. Lanes 5-14: Tp^r DH5α/pAMC5pAG900 isolates. 5.3-kb bands in Tp^r isolates are indicative of single IS*492*Δ*mooV*::*dfr* (1.3-kb) insertion into pAMC5. pAMC21 fusion plasmids with pAMC5 (4.3-kb + 3-kb) and pAG900 (9.5-kb + 1.6-kb) are also indicted. Lane M: 1-kb marker (New England Biolabs).

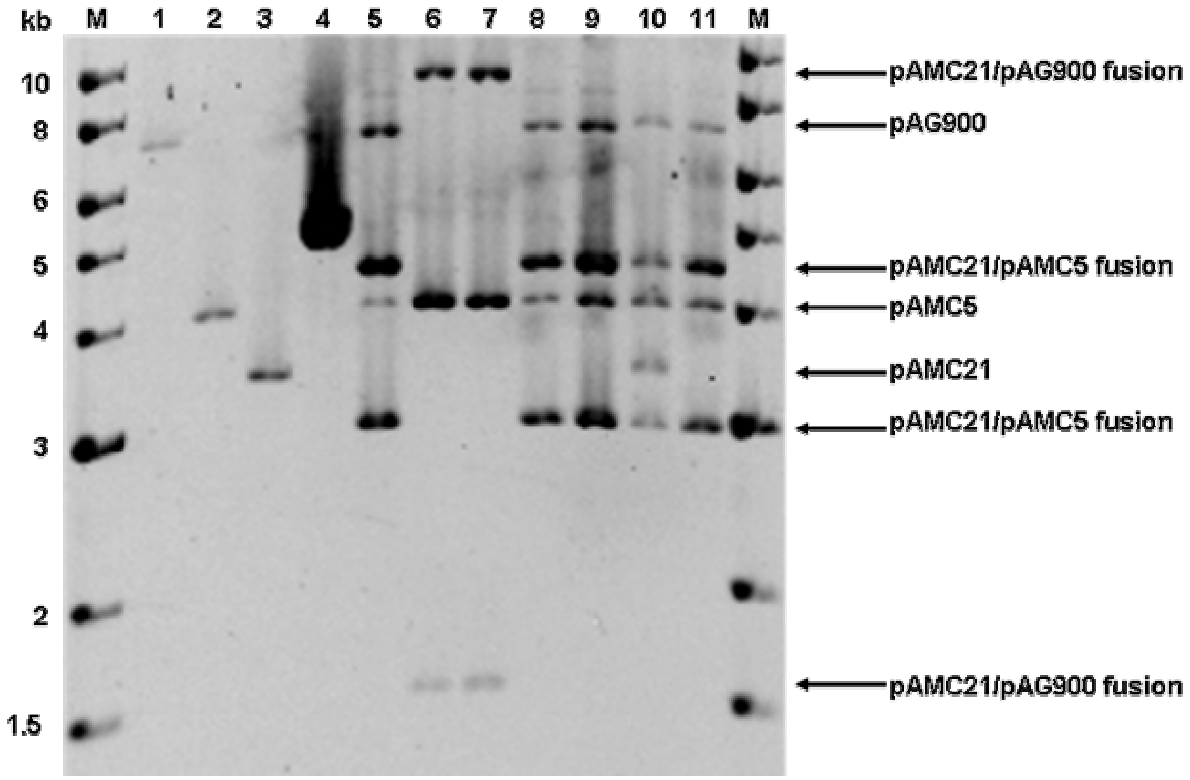


Figure 3.3. Characterizing Tp^r DH5α/pAMC5pAG900 isolates at 37°C. Plasmid DNA was isolated from Tp^r DH5α/pAMC5pAG900 isolates from a 37°C experiment (with 50 μM IPTG present in selective medium to induce MooV expression from P*tac*) digested with BamHI, and electrophoresed on a 0.8% gel (inverse image is shown). Controls and expected sizes: Lane 1 pAG900, 7.8-kb; Lane 2 pAMC5, 4-kb; Lane 3 pAMC21, 3.3-kb; Lane 4 pAG949 (positive control for single IS*492* insertion into pAMC5) 5.2-kb. Lanes 5-11: Tp^r DH5α/pAMC5pAG900 isolates. pAMC21 fusion plasmids with pAMC5 (4.3-kb + 3-kb) and pAG900 (9.5-kb + 1.6-kb) are indicted. Lane M: 1-kb marker (New England Biolabs).

CHAPTER 4

Conclusion

Great advancements have been made in understanding the molecular basis of specialized DNA recombination. From the ground-breaking work of Mizuuchi in developing the first *in vitro* transposition assay[135,136] to the examination of specialized DNA recombinases via X-ray crystallography studies,[137-139] much of what we currently know regarding specialized DNA recombination comes from the studies of tyrosine (Y-) integrases, serine (S-) invertases and resolvases, and DDE-motif transposases. With the increasing number of site-specific recombinases and transposases being characterized at the molecular level, it is now apparent that many of these enzymes do not completely "fit" into the above categories. For example, the IS*91* family of transposable elements is more closely related to the rolling circle replicating plasmids and single-stranded DNA phages than to the 'classical' DDE transposable elements, and transpose via a rolling-circle mechanism.[140] In addition, there are a number of transposable elements that have characteristics of both site-specific recombination and DNA transposition, including Tn*554* and Tn*916*, which have transposases that are closely related to the (Y-) integrases and have preferred target sites for insertion.[86] Identifying and characterizing new mobile elements and recombinases that do not completely fit into the three main recombinase categories described above would provide new insights into different and un-described mechanisms of DNA recombination.

There has been increasing evidence to suggest that the Piv/MooV recombinase family is a novel group of specialized recombinases that does not share significant sequence homology with

the Y- or S- site-specific recombinases or the classical DDE motif DNA transposases, and contains members that perform site-specific recombination and DNA transposition. Recombinases in this family share a conserved DEDD motif, which has been shown to be required for catalysis in Piv inversion.[106] Characterization of a number of transposable elements in this family has shown that they have preferred insertion target sites and may use a type of conservative site-specific recombination mechanism rather than a transposition mechanism for insertion.[98,103] Although these observations strengthen the idea that the Piv/MooV recombinase family defines a new category of recombinases, further characterization of mobile elements in this family is needed to determine how these enzymes mediate recombination. Early work characterizing IS*492* excision in *E. coli* suggested that MooV-mediated IS*492* transposition possesses features of both site-specific recombination and DNA transposition.[88,141] The goals of this research were to characterize IS*492* insertion in *E. coli* to elucidate further the mechanism used for recombination (i.e. cut-and-paste versus replicative transposition or conservative site-specific recombination). Observations from this study provide additional evidence that MooV-mediated IS*492* transposition has adopted a combination of site-specific recombination and DNA transposition for its mode of recombination.

The initial aim of this dissertation was to identify *P. atlantica* specific recombination factors that may be required for IS*492* insertion in *E. coli*. Using lower selective temperatures permitted the first observation of IS*492* insertion in *E. coli*, and suggests that MooV is the primary factor required for site-specific insertion in both *P. atlantica* and *E. coli*. As observed in *P. atlantica*, IS*492* insertion in *E. coli* displays site-specificity, targeting the 5-bp 5'-CTTGT-3' sequence found within the *epsG* target sequence and at the ends of IS*492*, creating single and tandem head-to-tail insertions, respectively. These types of insertions are commonly found with

transposable elements that use a cut-and-paste transposition mechanism. Head-to-tail tandem insertions are also common among transposable elements that excise to form circular products, where the junction sequence formed in the dimer is the same sequence found in the circular element. It has been demonstrated for IS*30*[131] and IS*911*[111,121,132] that the element junction sequence found in the excised circle and tandem dimer are competent for insertion in the presence of their transposase. IS*492* tandem dimers are also capable of insertion in the presence of MooV, and this observation suggests that the IS*492* "circle-junction" sequence not only plays a regulatory role by forming a strong promoter capable of driving *mooV* expression,[88] but that it may also serve as key intermediate in the IS*492* transposition pathway.

Although tandem IS*492* transposition is more frequent compared to monomer IS*492*, it mirrors that of the monomer element in having high frequency precise excision with detection of the repaired donor plasmid and site-specific insertion into the *epsG* target site. Analyses of MooV-mediated cointegrate formation between the tandem IS*492* donor and *eps* target plasmid suggests that MooV uses a cut-and-paste insertion mechanism; however the site-specific nature of the insertions could also arise from a conservative site-specific recombination, if the 5-bp "target" sequence acts as the core crossover site. The observed difference in sensitivity to temperature in monomer versus tandem IS*492* transposition in *E. coli* may further reflect the role of forming the circle-junction sequence to be the rate-limiting step in IS*492* transposition, and not necessarily differences in MooV expression or stability at different temperatures. Since MooV is capable of excising the monomer IS*492* to form a circular element at 37°C (with a required incubation step of 4°C between 2-24 hours to detect excised circles via PCR),[88] and insertion of the tandem IS*492* occurs at both 37°C and 25°C, I propose that the rate-limiting step

in IS*492* transposition is the formation of the circle junction. Once this junction is formed, MooV-mediated excision and insertion can be observed at 37°C.

The role of the IS*492* circle and the 5-bp 5'-CTTGT-3' sequence needs to be addressed further in IS*492* transposition. If the IS*492* circle is an intermediate in the transposition pathway, mutational analysis of the common 5-bp sequence within the IS*492* circle-junction sequence and the *epsG* target site can be performed to determine the requirement of sequence homology in IS*492* transposition, which would be a strong indicator that MooV uses a site-specific recombination versus a transposition mechanism. It has already been determined that 5-bp left and 7-bp right flanking *eps* sequences are required for IS*492* precise excision (Perkins-Balding and Karls, unpublished data), however the minimal *epsG* target sequence required for IS*492* insertion needs to be determined, and such studies would also help identify the necessary sequences required for MooV recognition and cleavage. Also, additional factors that influence site-specific insertion of IS*492* need to be analyzed; while all six chromosomal copies of IS*492* in *P. atlantica* are flanked by the same 5-bp left and 7-bp (5'-CTTGTTA-3') right sequence, these sequences are represented 5,645 and 396 times in the *P. atlantica* T6c chromosome, respectively. Examining the *in vivo* role of common host recombination factors in IS*492* insertion is also critical for developing an *in vitro* IS*492* transposition assay, where the recombination mechanism (cut-and-paste transposition versus site-specific recombination) can be accurately determined.

My research has laid the foundation for characterizing *in vivo* MooV-mediated IS*492* insertion in *E. coli*, and for determining essential proteins and DNA sequences required for IS*492* transposition. The findings in this study have moved characterization of the Piv/MooV

family of recombinases one step closer to revealing the mode of recombination used by this novel family.

REFERENCES

(1) Richardson, C.; Horikoshi, N.; Pandita, T. K. *DNA Repair* **2004**, *3*, 1149-1164.

(2) Baumann, P.; West, S. C. *Trends Biochem. Sci.* **1998**, *23*, 247-251.

(3) Kowalczykowski, S. C.; Dixon, D. A.; Eggleston, A. K.; Lauder, S. D.; Rehrauer, W. M. *Microbiol. Rev.* **1994**, *58*, 401-465.

(4) Hallet, B.; Sherratt, D. J. *FEMS Microbiol. Rev.* **1997**, *21*, 157-178.

(5) Azaro, M. A.; Landy, A. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 118-148.

(6) Johnson, R. C. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Soceity for Microbiology Press: Washington D.C., 2002, p 230-271.

(7) Ziebuhr, W.; Krimmer, V.; Rachid, S.; Lossner, I.; Gotz, F.; Hacker, J. *Mol. Microbiol.* **1999**, *32*, 345-356.

(8) Rice, L. B. *Antimicrob. Agents Chemother.* **1998**, *42*, 1871-1877.

(9) Craig, N. L. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*; 2 ed.; Neidhardt, F. C., Curtiss III, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E., Eds.; American Society for Microbiology Press: Washington D.C., 1996, p 2339-2362.

(10) Polard, P.; Chandler, M. *Mol. Microbiol.* **1995**, *15*, 13-23.

(11) Mizuuchi, K.; Baker, T. A. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 12-23.

(12) May, E. W.; Craig, N. L. *Science* **1996**, *272*, 401-404.

(13) Bhasin, A.; Goryshin, I. Y.; Reznikoff, W. S. *J. Biol. Chem.* **1999**, *274*, 37021-37029.

(14) Bender, J.; Kuo, J.; Kleckner, N. *Genetics* **1991**, *128*, 687-694.

(15) Reimmann, C.; Haas, D. *Genetics* **1987**, *115*, 619-625.

(16) Olasz, F.; Stalder, R.; Arber, W. *Mol. Gen. Genet.* **1993**, *239*, 177-187.

(17) Grindley, N. D. F. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 272-302.

(18) Nash, H. A. In *Escherichia coli and Salmonella:Cellular and Molecular Biology*; 2 ed.; Neidhardt, F. C., Curtiss III, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E., Eds.; American Society for Microbiology Press: Washington D.C., 1996, p 2362-2376.

(19) Dhar, G.; Sanders, E. R.; Johnson, R. C. *Cell* **2004**, *119*, 33-45.

(20) Johnson, R. C.; Johnson, L. M.; Schmidt, J. W.; Gardner, J. F.; Higgins, P. N., Ed.; American Soceity for Microbiology Press: Washington D.C., 2005, p 65-132.

(21) Azam, T. A.; Iwata, A.; Nishimura, A.; Ueda, S.; Ishihama, A. *J. Bacteriol.* **1999**, *181*, 6361-6370.

(22) Shiga, Y.; Sekine, Y.; Kano, Y.; Ohtsubo, E. *J. Bacteriol.* **2001**, *183*, 2476-2484.

(23) Shiga, Y.; Sekine, Y.; Ohtsubo, E. *Genes Cells* **1999**, *4*, 551-561.

(24) Rouquette, C.; Serre, M. C.; Lane, D. *J. Bacteriol.* **2004**, *186*, 2091-2098.

(25) Swingle, B.; O'Carroll, M.; Haniford, D.; Derbyshire, K. M. *Mol. Microbiol.* **2004**, *52*, 1055-1067.

(26) Wardle, S. J.; O'Carroll, M.; Derbyshire, K. M.; Haniford, D. B. *Genes Dev.* **2005**, *19*, 2224-2235.

(27) Falconi, M.; McGovern, V.; Gualerzi, C.; Hillyard, D.; Higgins, N. P. *New Biol.* **1991**, *3*, 615-625.

(28) Nash, H. A.; Robertson, C. A. *J. Biol. Chem.* **1981**, *256*, 9246-9253.

(29) Wiater, L. A.; Grindley, N. D. F. *EMBO J.* **1988**, *7*, 1907-1911.

(30) Gamas, P.; Chandler, M. G.; Prentki, P.; Galas, D. J. *J. Mol. Biol.* **1987**, *195*, 261-272.

(31) Huisman, O.; Errada, P. R.; Signon, L.; Kleckner, N. *EMBO J.* **1989**, *8*, 2101-2109.

(32) Ilves, H.; Horak, R.; Teras, R.; Kivisaar, M. *Mol. Microbiol.* **2004**, *51*, 1773-1785.

(33) Horak, R.; Kivisaar, M. *J. Bacteriol.* **1998**, *180*, 2822-2829.

(34) Morisato, D.; Kleckner, N. *Cell* **1987**, *51*, 101-111.

(35) Sakai, J. S.; Kleckner, N.; Yang, X.; Guhathakurta, A. *EMBO J.* **2000**, *19*, 776-785.

(36) Sewitz, S.; Crellin, P.; Chalmers, R. *Nucleic Acids Res.* **2003**, *31*, 5868-5876.

(37) Johnson, R. C.; Bruist, M. F.; Simon, M. I. *Cell* **1986**, *46*, 531-539.

(38) Koch, C.; Kahmann, R. *J. Biol. Chem.* **1986**, *261*, 5673-5678.

(39) Mertens, G.; Fuss, H.; Kahmann, R. *J. Biol. Chem.* **1986**, *261*, 5668-5672.

(40) Johnson, R. C.; Glasgow, A. C.; Simon, M. I. *Nature* **1987**, *329*, 462-465.

(41) Thompson, J. F.; Devargas, L. M.; Koch, C.; Kahmann, R.; Landy, A. *Cell* **1987**, *50*, 901-908.

(42) Ball, C. A.; Johnson, R. C. *J. Bacteriol.* **1991**, *173*, 4027-4031.

(43) Esposito, D.; Gerard, G. F. *J. Bacteriol.* **2003**, *185*, 3076-3080.

(44) Weinreich, M. D.; Reznikoff, W. S. *J. Bacteriol.* **1992**, *174*, 4530-4537.

(45) Haykinson, M. J.; Johnson, R. C. *EMBO J.* **1993**, *12*, 2503-2512.

(46) Lavoie, B. D.; Chaconas, G. *Genes Dev.* **1993**, *7*, 2510-2519.

(47) Lavoie, B. D.; Shaw, G. S.; Millner, A.; Chaconas, G. *Cell* **1996**, *85*, 761-771.

(48) Kobryn, K.; Lavoie, B. D.; Chaconas, G. *J. Mol. Biol.* **1999**, *289*, 777-784.

(49) Connolly, K. M.; Iwahara, M.; Clubb, R. T. *J. Bacteriol.* **2002**, *184*, 2088-2099.

(50) Roberts, D.; Hoopes, B. C.; McClure, W. R.; Kleckner, N. *Cell* **1985**, *43*, 117-130.

(51) Yin, J. C. P.; Krebs, M. P.; Reznikoff, W. S. *J. Mol. Biol.* **1988**, *199*, 35-45.

(52) Camara, J. E.; Crooke, E. In *The Bacterial Chromosome*; Higgins, N. P., Ed.; American Society for Microbiology Press: Washington D.C., 2005, p 177-191.

(53) Fuller, R. S.; Funnell, B. E.; Kornberg, A. *Cell* **1984**, *38*, 889-900.

(54) Yin, J. C. P.; Reznikoff, W. S. *J. Bacteriol.* **1987**, *169*, 4637-4645.

(55) Geuskens, V.; Mhammedialaoui, A.; Desmet, L.; Toussaint, A. *EMBO J.* **1992**, *11*, 5121-5127.

(56) Jones, J. M.; Welty, D. J.; Nakai, H. *J. Biol. Chem.* **1998**, *273*, 459-465.

(57) Craig, N. L. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 423-456.

(58) Bainton, R. J.; Kubo, K. M.; Feng, J. N.; Craig, N. L. *Cell* **1993**, *72*, 931-943.

(59) Sharp, P. L.; Craig, N. L. *EMBO J.* **1998**, *17*, 5822-5831.

(60) Goryshin, I. Y.; Miller, J. A.; Kil, Y. V.; Lanzov, V. A.; Reznikoff, W. S. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 10716-10721.

(61) Halling, S. M.; Kleckner, N. *Cell* **1982**, *28*, 155-163.

(62) Hu, W. Y.; Derbyshire, K. M. *J. Bacteriol.* **1998**, *180*, 3039-3048.

(63) Olasz, F.; Farkas, T.; Kiss, J.; Arini, A.; Arber, W. *J. Bacteriol.* **1997**, *179*, 7551-7558.

(64) Hallet, B.; Rezsohazy, R.; Mahillon, J.; Delcour, J. *Mol. Microbiol.* **1994**, *14*, 131-139.

(65) Bender, J.; Kleckner, N. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 7996-8000.

(66) Ason, B.; Reznikoff, W. S. *J. Mol. Biol.* **2004**, *335*, 1213-1225.

(67) Pribil, P. A.; Haniford, D. B. *J. Mol. Biol.* **2003**, *330*, 247-259.

(68) Muller, H. P.; Varmus, H. E. *EMBO J.* **1994**, *13*, 4704-4714.

(69) Craigie, R. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 613-630.

(70) Daniell, E.; Abelson, J.; Roberts, R. *J. Mol. Biol.* **1972**, *69*, 1-&.

(71) Wang, X. H.; Higgins, N. P. *Mol. Microbiol.* **1994**, *12*, 665-677.

(72) Manna, D.; Breier, A. M.; Higgins, N. P. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9780-9785.

(73) Casadesus, J.; Roth, J. R. *Mol. Gen. Genet.* **1989**, *216*, 204-209.

(74) Higgins, N. P.; Deng, S.; Pang, Z.; Stein, R. A.; Champion, K.; Manna, D. In *The Bacterial Chromosome*; American Soceity for Microbiology Press: Washington, D.C., 2005, p 133-153.

(75) Isberg, R. R.; Syvanen, M. *Cell* **1982**, *30*, 9-18.

(76) Lodge, J. K.; Berg, D. E. *J. Bacteriol.* **1990**, *172*, 5956-5960.

(77) Wolkow, C. A.; DeBoy, R. T.; Craig, N. L. *Genes Dev.* **1996**, *10*, 2145-2157.

(78) Waddell, C. S.; Craig, N. L. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 3958-3962.

(79) Kubo, K. M.; Craig, N. L. *J. Bacteriol.* **1990**, *172*, 2774-2778.

(80) Gay, N. J.; Tybulewicz, V. L. J.; Walker, J. E. *Biochem. J.* **1986**, *234*, 111-117.

(81) Gringauz, E.; Orle, K. A.; Waddell, C. S.; Craig, N. L. *J. Bacteriol.* **1988**, *170*, 2832-2840.

(82) Kuduvalli, P. N.; Rao, J. E.; Craig, N. L. *EMBO J.* **2001**, *20*, 924-932.

(83) Phillips, S.; Novick, R. P. *Nature* **1979**, *278*, 476-478.

(84) Krolewski, J. J.; Murphy, E.; Novick, R. P.; Rush, M. G. *J. Mol. Biol.* **1981**, *152*, 19-33.

(85) Murphy, E.; Lofdahl, S. *Nature* **1984**, *307*, 292-294.

(86) Derbyshire, K. M.; Grindley, N. D. F. In *The Bacterial Chromosome*; Higgins, P. N., Ed.; American Society for Microbiology Press: Washington D.C., 2005, p 467-497.

(87) Chandler, M.; Mahillon, J. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D.C., 2002, p 305-366.

(88) Perkins-Balding, D.; Duval-Valentin, G.; Glasgow, A. C. *J. Bacteriol.* **1999**, *181*, 4937-4948.

(89) Henderson, D. J.; Lydiate, D. J.; Hopwood, D. A. *Mol. Microbiol.* **1989**, *3*, 1307-1318.

(90) Partridge, S. R.; Hall, R. M. *J. Bacteriol.* **2003**, *185*, 6371-6384.

(91) Lauf, U.; Muller, C.; Herrmann, H. *FEMS Microbiol. Lett.* **1999**, *170*, 407-412.

(92) Lenich, A. G.; Glasgow, A. C. *J. Bacteriol.* **1994**, *176*, 4160-4164.

(93) Kunze, Z. M.; Wall, S.; Appelberg, R.; Silva, M. T.; Portaels, F.; McFadden, J. J. *Mol. Microbiol.* **1991**, *5*, 2265-2272.

(94) Green, E. P.; Tizard, M. L. V.; Moss, M. T.; Thompson, J.; Winterbourne, D. J.; McFadden, J. J.; Hermontaylor, J. *Nucleic Acids Res.* **1989**, *17*, 9063-9073.

(95) Moss, M. T.; Malik, Z. P.; Tizard, M. L. V.; Green, E. P.; Sanderson, J. D.; Hermon-Taylor, J. *J. Gen. Microbiol.* **1992**, *138*, 139-145.

(96) Leskiw, B. K.; Mevarech, M.; Barritt, L. S.; Jensen, S. E.; Henderson, D. J.; Hopwood, D. A.; Bruton, C. J.; Chater, K. F. *J. Gen. Microbiol.* **1990**, *136*, 1251-1258.

(97) Bruton, C. J.; Chater, K. F. *Nucleic Acids Res.* **1987**, *15*, 7053-7065.

(98) Henderson, D. J.; Lydiate, D. J.; Hopwood, D. A. *Mol. Microbiol.* **1989**, *3*, 1307-1318.

(99) Lydiate, D. J.; Ikeda, H.; Hopwood, D. A. *Mol. Gen. Genet.* **1986**, *203*, 79-88.

(100) Smokvina, T.; Hopwood, D. A. *Mol. Gen. Genet.* **1993**, *239*, 90-96.

(101) Bhatt, A.; Kieser, T. *Microbiology-Sgm* **1999**, *145*, 1201-1207.

(102) Thorsted, P. A.; Macartney, D. P.; Akhtar, P.; Haines, A. S.; Ali, N.; Davidson, P.; Stafford, T.; Pocklington, M. J.; Pansegrau, W.; Wilkins, B. M.; Lanka, E.; Thomas, C. N. *J. Mol. Biol.* **1998**, *282*, 969-990.

(103) Muller, C.; Lauf, U.; Herrmann, H. *Mol. Genet. Genomics* **2001**, *265*, 1004-1010.

(104) Mahillon, J. a. M. C. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 725-774.

(105) Tobiason, D. M., J.M. Buchner, W.H. Thiel, K.M. Gernert, and A.C. Glasgow-Karls. *Mol. Microbiol.* **2001**, *39*, 641-651.

(106) Buchner, J. M.; Robertson, A. E.; Poynter, D. J.; Denniston, S. S.; Karls, A. C. *J. Bacteriol.* **2005**, *187*, 3431-3437.

(107) Bartlett, D. H.; Wright, M. E.; Silverman, M. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 3923-3927.

(108) Gauthier, G.; Gauthier, M.; Christen, R. *Int. J. Syst. Bacteriol.* **1995**, *45*, 755-761.

(109) Bartlett, D. H.; Silverman, M. *J. Bacteriol.* **1989**, *171*, 1763-1766.

(110) Polard, P.; Prere, M. F.; Fayet, O.; Chandler, M. *EMBO J.* **1992**, *11*, 5079-5090.

(111) Bao, T. H.; Betermier, M.; Polard, P.; Chandler, M. *EMBO J.* **1997**, *16*, 3357-3371.

(112) Higgins, B. P.; Carpenter, C. D.; Karls, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 1901-1906.

(113) Perkins-Balding, D. a. A. C. G. *J. Bacteriol.* **1999**, *181*, 4937-4948.

(114) Tobiason, D. M.; Buchner, J. M.; Thiel, W. H.; Gernert, K. M.; Karls, A. C. G. *Mol. Microbiol.* **2001**, *39*, 641-651.

(115) Craig, N. L. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington,D.C., 1996, p 2339-2362.

(116) Stabb, E. V.; Ruby, E. G. In *Bacterial Pathogenesis, Pt C* 2002; Vol. 358, p 413-426.

(117) Guzman, L. M.; Belin, D.; Carson, M. J.; Beckwith, J. *J. Bacteriol.* **1995**, *177*, 4121-4130.

(118) Ochman, H.; Ajioka, J. W.; Garza, D.; Hartl, D. L. *Bio-Technology* **1990**, *8*, 759-760.

(119) Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl, K. *Current Protocols in Molecular Biology*; John Wiley & Sons, Inc.: New York, N.Y., 1991.

(120) Riess, G.; Masepohl, B.; Puehler, A. *Plasmid* **1983**, *10*, 111-118.

(121) Turlan, C.; Ton-Hoang, B.; Chandler, M. *Mol. Microbiol.* **2000**, *35*, 1312-1325.

(122) Smokvina, T.; Henderson, D. J.; Melton, R. E.; Brolle, D. F.; Kieser, T.; Hopwood, D. A. *Mol. Microbiol.* **1994**, *12*, 459-468.

(123) Bzymek, M.; Lovett, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8319-8325.

(124) Kretschmer, P. J.; Cohen, S. N. *J. Bacteriol.* **1977**, *130*, 888-899.

(125) Kretschmer, P. J.; Cohen, S. N. *J. Bacteriol.* **1979**, *139*, 515-519.

(126) Haren, L.; Betermier, M.; Polard, P.; Chandler, M. *Mol. Microbiol.* **1997**, *25*, 531-540.

(127) Gueguen, E.; Rousseau, P.; Duval-Valentin, G.; Chandler, M. *Mol. Microbiol.* **2006**, *62*, 1102-1116.

(128) Mikhailov, V. V.; Romanenko, L. A.; Ivanova, E. P. In *The Prokaryotes*; 3 ed.; Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K., Stackebrandt, E., Eds.; Springer: New York, 2006; Vol. 6, p 597–645.

(129) Ashby, M. K.; Bergquist, P. L. *Plasmid* **1990**, *24*, 1-11.

(130) Berger, B.; Haas, D. *Cell. Mol. Life Sci.* **2001**, *58*, 403-419.

(131) Kiss, J.; Olasz, F. *Mol. Microbiol.* **1999**, *34*, 37-52.

(132) Ton-Hoang, B.; Polard, P.; Chandler, M. *EMBO J.* **1998**, *17*, 1169-1181.

(133) Ikeda, H.; Shiraishi, K.; Ogata, Y. *Novel Developments on Genetic Recombination: DNA Double Strand Break and DNA End-Joining* **2004**, *38*, 3-20.

(134) Dunn, A. K.; Martin, M. O.; Stabb, E. V. *Plasmid* **2005**, *54*, 114-134.

(135) Mizuuchi, K. *Cell* **1983**, *35*, 785-794.

(136) Mizuuchi, K. *Cell* **1984**, *39*, 395-404.

(137) Davies, D. R.; Goryshin, I. Y.; Reznikoff, W. S.; Rayment, I. *Science* **2000**, *289*, 77-85.

(138) Yang, W.; Mizuuchi, K. *Structure* **1997**, *5*, 1401-1406.

(139) Rice, P. A.; Steitz, T. A. *Structure* **1994**, *2*, 371-384.

(140) Garcillan-Barcia, M. P.; Bernales, I.; Mendiola, M. V.; De La Cruz, F. In *Mobile DNA II*; Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., Eds.; American Society for Microbiology Press: Washington D. C., 2002, p 891-904.

(141) Perkins-Balding, D., Ph.D. thesis, Emory University, 2000.