STRUCTURAL AND FUNCTIONAL ANALYSES OF A NOVEL RECOMBINASE: PIV

by

JOHN M. BUCHNER

(Under the Direction of Anna C. Karls)

ABSTRACT

The DNA site-specific invertase Piv catalyzes the inversion of a 2.1 kb segment in *Moraxella lacunata* and *Moraxella bovis* altering the expression of type 4 pilin genes in two phases. The switch is a biphasic system, alternating between expression the *tfpI* and *tfpQ* pilin genes. In *M. lacunata* this switch is an on/off phase variation. Outside but adjacent to the invertible region, *piv* encodes the DNA invertase (Piv) that interacts with *invL* and *invR* as the loci of recombination. Other sites of Piv DNA binding have been observed, and one, *sub1*, has been characterized. Piv does not have homology with the traditional site-specific tyrosine or serine recombinases. Instead, Piv has homology to the transposases of the IS*110*/IS*492* family of insertion elements. Modeling of Piv revealed three structures of interest: 1) the ribonuclease H-like fold, a catalytic domain associated with DDE-motif transposases, retroviral integrases, and RuvC Holliday junction resolvases, 2) a potential leucine zipper that could be the site of protein dimerization, 3) a predicted helix-hairpin-helix (HhH) DNA binding motif, best studied in RuvA, a protein that binds Holliday junctions in a non sequence-specific manner. We determined that four acidic residues that are conserved among Piv and the recombinases of the IS*110*/IS*492* family of recombinases are required for catalysis of inversion and comprise a DEDD motif in Piv like that of RuvC. There was no requirement for the predicted leucine zipper in our *in vivo*

inversion system. We also mutated residues that are conserved in a consensus HhH sequence and observed that they are required for DNA inversion and may play a role in binding to Holliday junctions *in vitro*. Through electrophoretic mobility shift assays, we observed an affinity of Piv for Holliday junctions and branched DNA substrates containing mismatched DNA. This leads us to our hypothesis: Piv is unique among enzymes in that it is a site-specific invertase that catalyzes recombination through hydrolysis and transesterification, like the DDE-motif transposases, and generates a Holliday junction intermediate that is resolved by Piv with a pair of hydrolysis reactions, like RuvC.

INDEX WORDS: Predicted Structure, Protein Structure, Tertiary Structure

STRUCTURE AND FUNCTIONAL ANALYSIS OF A NOVEL RECOMBINASE: PIV

by

JOHN M. BUCHNER

B.S., University of Wisconsin - Madison, 1996

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

STRUCTURE AND FUNCTIONAL ANALYSIS OF A NOVEL RECOMBINASE: PIV

by

JOHN M. BUCHNER

Major Professor: Anna Karls
Timothy Hoover

Committee: Ellen Neidle
Claiborne Glover

Electronic Version Approved:

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

DEDICATION

I dedicate this to my wife, Dawn, who without her, I would've never arrived here. From giving me room to work, being a sounding board for all my ideas an inspiration, with her support, and understanding, I could not ask for a better partner to go through all of this with.

ACKNOWLEDGEMENTS

First, I owe my gratitude to my graduate Mmntor Anna Karls for her support throughout the years, her tenacity in asking me twice to work for her, and bringing me out to Athens. I'd like to acknowledge my committee members Claiborne Glover, Ellen Neidle, Tim Hoover, and Larry Shimkets for their suggestions and advice and words of encouragement and motivation when needed. I've appreciated the help of co-workers I've had throughout the years, including Donna-Perkins Balding, Wei-Ping Fu, Richard "Tad" Seyler, Anne Robertson, Brian Higgins, Chandra Carpenter, Adrienne Cotrell, How-Yi Chang, William Thiel, David Poynter, Shelby Denniston, Rhaisa Castrodad, and every member of the Karls laboratory, past and present who've supported my work in any way. I thank all my classmates for helping my class work look good, particularly a core group of people I started graduate school with, namely, Dawn, James, Patrick, Praveen, Lara, and Malani. Any time a class required a group effort, I always felt that everyone around, even when being critical, was doing it in a supportive way. Thank you to the members of the Stabb and Hoover laboratories for having joint group meetings with us. I've appreciated the help of many of the office staff, particularly Pat, Nancy, and Rebecca, who are always patient no matter how many times I've tried to flaunt the rules, and Dale for helping out with any technological requests, no matter how absurd I sound to him. In the last stretch of this phase of my life, I'd like to thank our Dissertation Reading Group, Dawn, Emily, James, and Chandra, in particular for straightening out my presentation that's taking me to the next step.

On a personal note, I acknowledge the love and support of my family, Ruth, Leo, Peg, Dave, Ed, Jenna, Sara, Bill, Nancy, Claude, Rosie, Tanya, Edward, and “Bud”, and all the relations too numerous to name here of both my birth and my extended families. I thank all of them for taking an interest in my work, giving support and encouragement, and trying to understand this process. I also want to thank all my friends; the ones I’ve made since coming to Athens, those from Atlanta, Madison, and growing up in New Berlin. They’ve seen me change, and see me stay the same. Thanks to the ones from Athens who’ve shared recent highs and lows, and Eric and Janice for hosting many seasons of Packer games. From Atlanta I’d like to thank Jim, Bill, Jon (Tex), Steve, Buddha, Major Dan, Duggan, Gruppo, Pete, and Kevin, and the rest of the Renegades R.F.C. and the High Point R.F.C., for giving me a chance to play a team sport at a level higher then I’ve ever played, and showing me more adventure than I’ve ever had. I still don’t know how to play Rugby, but I do know how to get up after being knocked down. Friends from Mike & Angelo’s: Tanner, Emily, Dreb, Alex, and the rest. From Madison, Neha, in particular, amongst all the people I met there and know now, you understand me the most. Finally from New Berlin, thanks Mark, Jason B, and everyone else I’ve stayed in touch with throughout the years.

Last, but not least, I thank once again, Dawn, for her love and patience through some of the most difficult times of my life, and some of the greatest moments of our lives together. I look forward for what the future holds for us together, and towards the next steps of our dual adventure. It is a wonderful thing to find someone who knows who you are, how you work, and understands what you do.

TABLE OF CONTENTS

Page

CHAPTER 1: Introduction and Literature Review

DNA recombination is a ubiquitous phenomenon in nature and is seen in many forms.

Chromosomes can be altered by horizontal transfer of genes through conjugation or the integration of bacteriophage or viral chromosomes into host DNA. Transposition can result in the silencing of genes through the integration of elements into coding regions, or it can stimulate the expression of silenced genes by putting transposons adjacent to promoters. Transposition, and the accompanying repair processes requiring host-encoded recombinases that take place after transposition, can result in the rearrangement (inversion, deletion, duplication) of DNA. Even in a cell not otherwise under attack from the rearrangements of transposable elements requires recombination for the restart of replication or the segregation of chromosomes. General and specialized recombination and DNA rearrangements play roles in the life cycles of all organisms in processes such as chromosome replication and segregation, spread of drug resistances, carcinogenesis, creation of antibodies, and variation of surface antigens in pathogenic bacteria (10), to name just a few processes.

Enzymes catalyzing what on the surface appear to be similar reactions, integrations/excisions, inversions, transpositions, or Holliday junction resolution, may have divergent tertiary structures and have different mechanisms for the breaking and/or joining of the phosphate backbones in recombination. Here, I present structure and function relationships for four groups of enzymes, the tyrosine recombinases, the serine recombinases, the DDE-motif transposases and retroviral integrases, and the Ribonuclease H-like Holliday junction resolvases.

A brief review of tyrosine recombinases, serine recombinases, DDE-motif transposases, and Ribonuclease H-like Holliday junction resolvases

Nearly all site-specific recombinases have been grouped into either the tyrosine (λ integrase) or serine (Hin/Resolvase) recombinases based on catalytic amino acid motifs and mechanistic similarities (4, 27, 59). The reactions mediated by site-specific recombinases include the rearrangement of promoters (Hin), chromosomal segregation (Xer), the introduction of new genes (IntI1), and the integration or excision of extra-chromosomal elements (λInt) (4). Enzymes of both families utilize a two-step *trans*-esterification recombination reaction (59). A hydroxyl group on tyrosine (λ-integrase) or serine (Hin/Resolvase) residues of the recombinase acts as a nucleophile to the attack on the phosphate backbone creating a covalently-linked DNA-protein intermediate, leaving either a 5'-OH or 3'-OH DNA end (Figure 1.1) (4, 59). Strand exchange occurs either through isomerization of a Holliday junction (HJ) intermediate (λ integrase) (4) or concerted strand cleavage and rotation (Hin/Resolvase) (27). The free hydroxyl attacks the phosphotyrosyl or phosphoseryl bond for the second step in the *trans*-esterification that restores the phosphate backbone of the DNA strands. The recombinase can cleave and rejoin DNA strands without the need of nucleoside triphosphates (59). Members of the λ integrase and Hin/Resolvase families utilize accessory proteins, such as Fis, IHF, or HU, that function to bind and bend DNA in assembly of the nucleoprotein complexes required for recombination (4, 27). The protein and DNA topological requirements of the site-specific recombination reaction allow for cellular control of recombination to prevent errant reactions such as a lethal chromosomal break (27).

Previously, the classification of an enzyme as a site-specific recombinase or a transposase has been based on the type of mechanism used: catalytic serine or tyrosine, for site-specific

recombinases, *versus* DDE transposases (5, 6, 51). It has been suggested that prokaryotic transposons be defined as "…genetic entities which are capable of inserting as discrete, non-permuted DNA segments at many different sites in prokaryotic genomes (13)". This definition states that transposons integrate into numerous locations in the chromosome regardless of what enzyme or mechanism is used. The definition is designed specifically to exclude phage λ, which integrates into a specific location, include the classic DDE-motif transposases, and accommodate new groups such as the tyrosine (Y) transposases (e.g.; Tn*916* Int), serine (S) transposases (e.g.; Tn*5397* TndX), and the Y2 transposases (e.g.; IS*91* TnpA) (11). Molecular mechanisms for the three new groups involve a covalently bound DNA-protein intermediate, and the Y or S transposases have homology to either λ integrase (Y transposase) or Hin resolvase (S transposase) respectively (13).

In both eukaryotes and prokaryotes, transposases mediate the relocation of mobile DNA elements such as transposons, insertion sequences; and retroviral elements, within or between chromosomes and extrachromosomal DNA. The majority of DNA transposases and retroviral integrases (11, 13) are members of the DDE-motif family. These recombinases also share tertiary structures (ribonuclease H-like motif), metal ion requirements, and biochemical mechanisms. The metal ions that serve as the reaction center of the transposase are often divalent cations coordinated by the DDE motif. The mechanism of transposition mediated by DDE-motif transposases and retroviral integrases is hydrolysis followed by a one step *trans*-esterification (Figure 1.2). A water molecule, coordinated in the reaction center, is the nucleophile used to attack the phosphate backbone leaving a free 3'-OH on either end of the element. The integration step in transposition uses this 3'-OH as the nucleophile to attack the phosphate backbone at the point of insertion in a one-step *trans*-esterification reaction (21). With

DDE-motif transposases, the transfer of each transposon DNA strand is staggered by a few to several bases so when the host DNA replication or repair machinery is used to close the gaps created by transposition, a direct duplication of the target sequence is created flanking the insertion element. The DNA cleavage and strand transfer steps of transposition are carried out without the consumption of nucleoside triphosphates (21). However, the repair of nick and gaps created by the transposition process do require host enzymes (ligase, polymerase) and the consumption of ATP or other nucleoside triphosphates.

The reactions mediated by DDE-motif transposases and site-specific recombinases differ in four aspects. First, the hydrolysis of the DNA phosphate backbone is catalyzed by a water molecule (DDE-motif transposase) or by a serine or tyrosine (S or Y recombinases). Second, the S and Y recombinase reaction mechanism utilize a covalently bound recombinase-DNA intermediate. Third, strand transfer mediated by DDE-motif transposases results in only one DNA strand at each end of the transposon linked to the target site. Finally, there is a DNA homology requirement (core sites) for site-specific DNA recombinases whereas even if transposases exhibit a site preference there is no homology requirement (4, 21, 27, 59).

The Holliday junction, a four arm branched DNA structure (24), is an intermediate of homologous DNA recombination (60) and plays a role in the restart of stalled DNA replication forks (34). Though Holliday junction resolvases are united in the resolution of Holliday junctions by cleavage of opposing phosphate backbones of the DNA strands at the junction (53), Holliday junction resolvases are a diverse group of enzymes found in eukaryotes, viruses, bacteria, and bacteriophages (1, 32), with differences in primary sequence, tertiary structure (1), cleavage site specificity, DNA bending effects, and DNA structure specificity (53). Here the focus will be on the bacterial chromosomally-encoded Holliday junction resolvase RuvABC.

Historically the first observations of the recombination of DNA introduced into *Escherichia coli* by conjugation (37), transformation (8), or bacteriophage infection indicated covalent linkages between the introduced and resident DNA, and host encoded factors were required for these linkages to occur (9). Another research group suggested that the persistence of UV-induced pyrimidine dimers in excision-deficient strains of *E. coli*, a potentially lethal phenotype because of the arrest of replication due to pyrimidine dimers, did not result in lethality because of a systematic recombination system involving Holliday junctions between chromosomes (48). The proteins identified that catalyzed the translocation and resolution of those Holliday junctions were RuvABC (14). As described later in this introduction, the function of RuvAB is to bind the Holliday junction and catalyze its translocation (38), and RuvC is the Holliday junction resolvase (5, 15, 51). Using synthetic Holliday junctions, *in vitro* resolution of Holliday junctions only requires RuvC and manganese or magnesium ions (5, 51). In the absence of RuvAB, RuvC still shows specificity in binding to Holliday junctions over duplex DNA. This DNA binding occurs in the absence of divalent cations, but RuvC requires divalent Mg^{2+} or Mn^{2+} to catalyze resolution (5, 7, 51). The crystal structure of RuvC revealed a RNaseH-like motif (2), similar to that found in the DDE-motif transposases. The catalytic domain of RuvC was not a triad but a quartet of acidic residues, D7, E66, D138, and D141. Variants of RuvC with replacements of these acidic residues, D7, E66, D138 and D141, results in loss of Holliday junction resolution, as measured by UV-light sensitivity when the RuvC variants are expressed *in trans* in RuvC deficient *E. coli* (49). *In vitro* synthetic Holliday junction cleavage assays revealed a requirement for homology (15) and a site preference, 5'WTTS3' where W is A or T and S is G or C. Cleavage of opposite strands in the HJ occurs through hydrolysis (Figure 1.3) between the third and fourth bases of the preferred sites,

resulting in two nicked strands of DNA, which are repaired by ligase (5, 51). DNase I and chemical protection assays reveal that RuvC protects the crossover region of the Holliday junctions and twelve bases more on each arm (5, 6, 51).

The site–specific DNA invertase Piv of *Moraxella lacunata*

Piv mediates the DNA inversion of type 4 pilin genes, controlling phase variation in *Moraxella lacunata*. *M. lacunata,* a member of the delta Proteobacteria, attaches to human eye epithelial cells and conjunctiva through type-4 pili causing conjunctivitis and keratitis (44, 47). *M. lacunata* phase-varies expression of type-4 pili by reversible inversion of a 2,197 bp chromosomal segment that encodes the type four pilin genes *tfpQ* and *tfpI* (33) (Figure 1.4). Immediately adjacent to the invertible region on one side is the 5'-end of the pilin gene sequence and the pilin promoter (P_{PILIN}). In the Q (on) orientation P_{PILIN} drives the expression of *tfpQ*. The *tfpQ* gene encodes functional type-4 pilin, allowing the establishment of new eye infections (23). Co-transcribed with *tfpQ*, *tfpB* has no known function and deletion of *tfpB* does not affect inversion. In the I (off) orientation, *tfpI* is expressed from P_{PILIN} (23). A 19 bp duplication in the *tfpI* sequence results in expression of a truncated protein that is not functional for pili assembly (33). On the other side of the invertible region is *piv* (pilin inversion) encoding a 322 amino acid site-specific DNA invertase (33). In *E. coli* DH5α, Piv is the only *M. lacunata* gene product required for inversion of this pilin segment (54). Three sites of Piv binding have been characterized: two recombination sites (*invL* and *invR*) and the strong upstream binding site (*sub1*) (55). The *sub1* site is not required for inversion, (54) but in other inversion systems alternate recombinase binding sites are involved in regulation of recombinase expression, or for the assembly of an active recombinatorial complex (59). Two additional accessory Piv binding

sites (one at the 5'-end of *piv* the other at the 3'-end of *tfpQ*) of unknown function are indicated by gel shift and sequence homology to *sub1* (55) (Figure 1.4).

Piv catalyzes site-specific recombination but exhibits 25%-35% amino acid sequence identity and 45%-55% similarity to the transposases of the IS*110*/IS*492* family of insertion sequences, including MooV, the IS*492* transposase (31), which is also being characterized in our laboratory and BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgisearch) results of Piv protein reveal a relationship to two PFAM families, Transposase 9 and Transposase 20. The reaction mechanisms utilized by the traditional categories of site-specific recombinases and transposases are dramatically different, so it is interesting that one family of enzymes, the Piv/MooV family, contain enzymes catalyzing both transposition and site-specific recombination. Piv and transposases of the IS*110*/IS*492* family share three completely conserved amino acid motifs: the glycine, aspartate, lysine (GDK) motif; the lysine, threonine, aspartate, aspartate, alanine (KTD--DA) motif; and the proline, serine, glycine (PSG) motif. The GDK, KTD--DA, and PSG motifs are required for Piv-mediated site-specific inversion *in vivo* but not for binding the *inv* or *sub1* sites (Figure 1.5) (54).

Although Piv appears to function as a site-specific recombinase and exhibits no significant sequence identity or homology to the DDE-motif transposases, predictions of Piv secondary structure and molecular modeling of Piv's tertiary structure revealed similarity with the ribonuclease H-like structural motif of the DDE-motif transposases. To predict regions of Piv secondary structure (α-helices or β-strands), five secondary structure prediction programs were used: GOR IV, NNPREDICT, PREDICTPROTIEN, SOPMA, and PSI DIRECT (12, 16-18, 28, 29, 46). The results of these prediction programs were assembled to give the total number of secondary structure α-helix or β-strand hits per residue. Compiling all the predictions created

"windows" of secondary structures. Using GENDOC (35) the predicted Piv structure was aligned with each of the then known secondary structures of DDE-motif transposases, ASV integrase, HIV integrase, and Tn*5* transposase inhibitor (Brookhaven Protein Data Bank accession numbers: 1VSD, 1ITG, 1B7E) and with a multiple alignment of the known structures. The multiple alignments of ASV integrase, HIV integrase, and Tn*5* transposase inhibitor accomplished using the SYBYL FIT MONOMER (57) program using the catalytic DDE residues as anchor points. The alignments were scored using GENEDOC scoring function under the BLOSUM 62 matrix score columns. The best scores for the key secondary structures of the DDE-motif transposases were the basis for adjustments of alignments. Using the co-ordinates of 1VSD, 1ITG, 1B7E, and composite structures from the multiple alignments, the three-dimensional modeling of each alignment was performed with MODELLER and analyzed using SYBYL PROTABLE (57). Each bond was inspected and adjustments made to avoid unacceptable molecular interactions. After each round of adjustment, the model was analyzed again. The result is a predicted Piv structure that is energetically possible and shares structural homology to the catalytic domain of the DDE-motif transposases. Therefore, the site-specific invertase Piv may have the structure and mechanism of a transposase (54). The two aspartic acids and the glutamic acid residues of Piv proposed to be the reaction center are modeled to be D9 of the GDK motif, E59, and D101 or D104 of the KTD--DA motif, thus constituting a DED catalytic motif instead of a DDE motif of transposases (54, 56). The E59 residue, required for Piv inversion (54), is conserved as either a glutamate or aspartate residue in all recombinases of the Piv/MooV family (31) possibly forming either a DED or DDD motif. A more recent modeling of Piv secondary structure is shown in Figure 1.6.

A helix-hairpin-helix (HhH) is a motif that allows proteins to bind DNA in a non-sequence specific manner and is found in many proteins that interact with DNA. Enzymes observed to have this motif include endonucleases, glycosylases, helicases, exonucleases, DNA and RNA polymerases, and DNA repair proteins such as Rad51 and RuvA (52). Perhaps the best-studied protein with this motif is RuvA which contains two HhH motifs (22, 36, 50) and associates with RuvBC to generate a Holliday junction "Resolveasome" to resolve Holliday junctions introduced into the cell through recombination or repair activities (61). The function of RuvA is to localize and bind to Holliday junctions (40), distorting them slightly in the process by catalyzing the denaturation of the DNA bases at the junction (43). RuvA interacts with RuvB, creating a machine that consumes ATP to rotate the DNA strands, translocating the Holliday junction though regions of homology (38-41) or through short regions of non-homology (58). RuvC joins the complex as the resolvase and symmetrically cleaves the junction through hydrolysis of the phosphate backbone (5, 15, 51). RuvA has been observed to form a structure with four "channels" where one face of a square planar Holliday junction could nest (36). In this model RuvC, known to form dimers in the presence (63) or absence (2) of RuvAB, could then bind to the other face of the Holliday junction to facilitate cleavage possibly through interaction with RuvAB (63). However, there are steric arguments against this model and evidence from *Mycobacterium leprae* that RuvA subunits form an octomeric structure, two tetramers stacked upon each other, creating "pipes" instead of "channels" for the Holliday junction strands to travel through, and a more secure attachment point for RuvB, such that the RuvA structure acts as a stator for the RuvB rotor. This model eliminates access for RuvC, as both faces and the junction are covered. Without dissociation of this structure, RuvC cannot catalyze resolution (45). Furthermore, the crystal structure of an *E. coli* RuvAB complex revealed a similar RuvA

octomeric structure associated with four RuvB units also arranged covering both faces of the HJ, and preventing obvious RuvC access to the junction (62).

With such diversity in the enzymes that have a HhH, there are few amino acids completely conserved in the primary sequence of these proteins. Many of these structures are identified only in the crystal structures of the proteins or through the use of bioinformatics tools (52). As a result, the most conserved feature among HhH motifs is Glycine-Hydrophobic Residue-Glycine (52). One study defined the primary sequence motif as: bxuxxupGuGpxxAxxuuxx, where b is D, E, R, K, N, Q, S, or T at least 65% of the time, u is M, L, I, V, F, T, A or Y at least 65% of the time, and p is P at least 35% of the time (42). The relevant residues in Piv that match this pattern are S195, L197, I200, P201, G202, I203, G204, L211, and L212 (Figures 1.5, 1.6). Two residues do not match this pattern. Piv has a lysine at position 206 instead of a proline, and Piv has a leucine at position 208 instead of alanine. Therefore the potential HhH motif in Piv is SxLxxIPGIG**K**xx**L**xxLLxx where the non-consensus residues of this sequence in bold. The Piv primary sequence, including resides L197, I200 through L208, and L211 through V213 (L--IPGIGKKTL--LLV) closely matches the RuvA HhH sequence L110, L116 though A121, and L124 through V126 (L--LPGIGK<u>K</u>TA--LIV). K119 (underlined) was found to be required for DNA binding (42) and in a position to be able to directly interact with the phosphate backbone in a crystal structure of RuvA-DNA complex (22). The equivalent lysine from Piv, K206, is studied to determine if it is required for catalysis of inversion or DNA binding. The known structure of RuvA has a helix from V107 to K112, a loop from L113 to K118, and a helix from K119 to F132. The most recent Piv secondary structure predictions using Predict Protein (http://www.predictprotein.org/, Figure 1.6) of the equivalent Piv residues predict a long leading to residue A198, a loop from T199 to G204, and a helix from

K205 to V214. In the third chapter of this dissertation, I examine the possibility of an HhH motif being conserved in the IS*110*/IS*492* family.

Another feature of Piv and related transposases indicated by sequence alignments was a potential leucine zipper (Figures 1.5, 1.6). Leucine zippers are often the location of protein-protein interactions and have been found as a common motif in transcription factors and recombinases (25, 30). A leucine zipper is the site of transposase subunit interactions in IS*911* transposition (19, 20). *In vitro* assays of Maltose Binding Protein-Piv (MBP-Piv) and Piv-His_6 peptide/nickel-binding fusion proteins indicated that Piv binds as a multimer to the recombination site but binds independently as monomers to the *sub1* site (55). The source of Piv monomer interactions may be centered on this leucine zipper. Protein-protein interactions are important because many site-specific recombinases and transposases are active only as dimers or tetramers, and interactions with other proteins frequently regulate the recombination process (3, 19, 26, 56, 59).

The focus of my dissertation research is the mechanism of inversion of Piv. First, residues predicted to be part of the catalytic motif were targeted for *in vivo* studies of Piv binding and inversion activities. Those results were used to develop the hypothesis expanded upon in the second part of my thesis, Piv is a site specific DNA invertase that utilizes hydrolysis and one-step transesterifaction with a Holliday junction intermediate. To this end, I built synthetic Holliday junctions and tested the binding properties of MBP-Piv fusions, also testing a K206E variant, to determine the requirement for that residue. I also further examine the possibilities of a HhH motif in Piv, and explore the possibility of a leucine zipper in Piv.

References:

1. **Aravind, L., K. S. Makarova, and E. V. Koonin.** 2000. SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. Nucleic Acids Res **28:**3417-32.

2. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, H. Nakamura, H. Shinagawa, and K. Morikawa.** 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. Cell **78:**1063-72.

3. **Azaro, M. A., and A. Landy.** 2002. λ Integrase and the λ Int Family, p. 118-148. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.

4. **Azaro, M. A., and A. Landy.** 2002. λ Integrase and the λ Int Family, p. 118-148. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.

5. **Bennett, R. J., H. J. Dunderdale, and S. C. West.** 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell **74:**1021-31.

6. **Bennett, R. J., and S. C. West.** 1995. Structural analysis of the RuvC-Holliday junction complex reveals an unfolded junction. J Mol Biol **252:**213-26.

7. **Benson, F. E., and S. C. West.** 1994. Substrate specificity of the *Escherichia coli* RuvC protein. Resolution of three- and four-stranded recombination intermediates. J Biol Chem **269:**5195-201.

8. **Bodmer, W. F., and A. T. Ganesan.** 1964. Biochemical and Genetic Studies of Integration and Recombination in *Bacillus Subtilis* Transformation. Genetics **50:**717-38.

9. **Clark, A. J.** 1971. Toward a metabolic interpretation of genetic recombination of *E. coli* and its phages. Annu Rev Microbiol **25:**437-64.

10. **Craig, N. L., R. Craigie, M. Gellert, and A. M. Lambowitz (ed.).** 2002. Mobile DNA II. ASM Press, Washington, D.C.

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

12. **Deleage, G., and B. Roux.** 1987. An algorithm for protein secondary structure prediction based on class predicition. Protein Eng **1:**28-294.

13. **Derbyshire, K. M., and N. D. F. Grindley.** 2005. DNA Transposons: DIfferent Proteins and Mechanisms but Similar Rearrangements, p. 467-498. *In* N. P. Higgins (ed.), The Bacterial Chromosome. ASM Press, Washington, D.C.

14. **Donaldson, J. R., C. T. Courcelle, and J. Courcelle.** 2006. RuvABC is required to resolve holliday junctions that accumulate following replication on damaged templates in *Escherichia coli*. J Biol Chem **281:**28811-21.

15. **Dunderdale, H. J., F. E. Benson, C. A. Parsons, G. J. Sharples, R. G. Lloyd, and S. C. West.** 1991. Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. Nature **354:**506-10.

16. **Garnier, J., J.-F. Gibrat, and B. Robson.** 1996. GOR secondary structure prediction method, Version IV. Methods Enzymol **266:**540-553.

17. **Geourjon, C., and G. Deleage.** 1995. SOPMA: significant improvements in protein seondary structure prediction by consensus prediction from multiple alignmetns. Comput Appl Biosci **11:**681-684.

18. **Guermeur, Y.** 1997. Combinaison de classifieurs statistiques. Application a la prediction de structure secondaire des proteines. Normale Superieure, Lyon, France.

19. **Haren, L., C. Normand, P. Polard, R. Alazard, and M. Chandler.** 2000. IS*911* transposition is regulated by protein-protein interactions via a leucine zipper motif. J Mol Biol **296:**757-68.

20. **Haren, L., P. Polard, B. Ton-Hoang, and M. Chandler.** 1998. Multiple oligomerisation domains in the IS*911* transposase: a leucine zipper motif is essential for activity. J Mol Biol **283:**29-41.

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

22. **Hargreaves, D., D. W. Rice, S. E. Sedelnikova, P. J. Artymiuk, R. G. Lloyd, and J. B. Rafferty.** 1998. Crystal structure of *E.coli* RuvA with bound DNA Holliday junction at 6 Å resolution. Nat Struct Biol **5:**441-6.

23. **Heinrich, D. W., and A. C. Glasgow.** 1997. Transcriptional regulation of type 4 pilin genes and the site-specific recombinase gene, piv, in Moraxella lacunata and Moraxella bovis. J Bacteriol **179:**7298-305.

24. **Holliday, R.** 1964. A Mechanism for gene conversion in Fungi. Genet. Res. Camb. **5:**282-304.

25. **Hu, J. C., E. K. O'Shea, P. S. Kim, and R. T. Sauer.** 1990. Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. Science **250:**1400-3.

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

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

28. **Jones, D. T.** 1999. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol **292:**195-202.

29. **Kneller, D.** 1991. NNPREDICT.

30. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759-64.

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

32. **Lilley, D. M., and M. F. White.** 2000. Resolving the relationships of resolving enzymes. Proc Natl Acad Sci U S A **97:**9351-3.

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

34. **Michel, B., H. Boubakri, Z. Baharoglu, M. LeMasson, and R. Lestini.** 2007. Recombination proteins and rescue of arrested replication forks. DNA Repair (Amst) **6:**967-80.

35. **Nicholas, K. B., H. B. Nicholas, Jr, and D. W. I. Deerfield.** 1997. GeneDoc: analysis and visualization of genetic variation. EMBNEW News **4:**14.

36. **Nishino, T., M. Ariyoshi, H. Iwasaki, H. Shinagawa, and K. Morikawa.** 1998. Functional analyses of the domain structure in the Holliday junction binding protein RuvA. Structure **6:**11-21.

37. **Oppenheim, A. B., and M. Riley.** 1966. Molecular recombination following conjugation in *Escherichia coli*. J Mol Biol **20:**331-57.

38. **Parsons, C. A., A. Stasiak, R. J. Bennett, and S. C. West.** 1995. Structure of a multisubunit complex that promotes DNA branch migration. Nature **374:**375-8.

39. **Parsons, C. A., A. Stasiak, and S. C. West.** 1995. The *E.coli* RuvAB proteins branch migrate Holliday junctions through heterologous DNA sequences in a reaction facilitated by SSB. Embo J **14:**5736-44.

40. **Parsons, C. A., I. Tsaneva, R. G. Lloyd, and S. C. West.** 1992. Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. Proc Natl Acad Sci U S A **89:**5452-6.

41. **Parsons, C. A., and S. C. West.** 1993. Formation of a RuvAB-Holliday junction complex in vitro. J Mol Biol **232:**397-405.

42. **Rafferty, J. B., S. M. Ingleston, D. Hargreaves, P. J. Artymiuk, G. J. Sharples, R. G. Lloyd, and D. W. Rice.** 1998. Structural similarities between *Escherichia coli* RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the holliday junction. J Mol Biol **278:**105-16.

43. **Rice, D. W., J. B. Rafferty, P. J. Artymiuk, and R. G. Lloyd.** 1997. Insights into the mechanisms of homologous recombination from the structure of RuvA. Curr Opin Struct Biol **7:**798-803.

44. **Ringvold, A., E. Vik, and L. S. Bevanger.** 1985. *Moraxella lacunata* isolated from epidemic conjunctivitis among teen-aged females. Acta Ophthalmol (Copenh) **63:**427-31.

45. **Roe, S. M., T. Barlow, T. Brown, M. Oram, A. Keeley, I. R. Tsaneva, and L. H. Pearl.** 1998. Crystal structure of an octameric RuvA-Holliday junction complex. Mol Cell **2:**361-72.

46. **Rost, B.** 1996. PHD: predicting one-dimentional protein structure by profile based on neural networks. Methods Enzymol **266:**525-539.

47. **Ruehl, W. W., C. F. Marrs, R. Fernandez, S. Falkow, and G. K. Schoolnik.** 1988. Purification, characterization, and pathogenicity of *Moraxella bovis* pili. J Exp Med **168:**983-1002.

48. **Rupp, W. D., C. E. Wilde, 3rd, D. L. Reno, and P. Howard-Flanders.** 1971. Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli*. J Mol Biol **61:**25-44.

49. **Saito, A., H. Iwasaki, M. Ariyoshi, K. Morikawa, and H. Shinagawa.** 1995. Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. Proc Natl Acad Sci U S A **92:**7470-4.

50. **Sedelnikova, S. E., J. B. Rafferty, D. Hargreaves, A. A. Mahdi, R. G. Lloyd, and D. W. Rice.** 1997. Crystallization of *E. coli* RuvA gives insights into the symmetry of a Holliday junction/protein complex. Acta Crystallogr D Biol Crystallogr **53:**122-4.

51. **Shah, R., R. J. Bennett, and S. C. West.** 1994. Activation of RuvC Holliday junction resolvase *in vitro*. Nucleic Acids Res **22:**2490-7.

52. **Shao, X., and N. V. Grishin.** 2000. Common fold in helix-hairpin-helix proteins. Nucleic Acids Res **28:**2643-50.

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

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

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

56. **Ton-Hoang, B., P. Polard, L. Haren, C. Turlan, and M. Chandler.** 1999. IS*911* transposon circles give rise to linear forms that can undergo integration *in vitro*. Mol Microbiol **32:**617-27.

57. **Tripos.** 2000. SYBYL 6.6. Tripos Inc., St. Louis USA.

58. **Tsaneva, I. R., B. Muller, and S. C. West.** 1993. RuvA and RuvB proteins of *Escherichia coli* exhibit DNA helicase activity *in vitro*. Proc Natl Acad Sci U S A **90:**1315-9.

59. **Van Duyne, G. D.** 2002. A Structural View of Tyrosine Recombinase Site-Specific Recombination, p. 93-117. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.

60. **West, S. C.** 2003. Molecular views of recombination proteins and their control. Nat Rev Mol Cell Biol **4:**435-45.

61. **West, S. C.** 1996. The RuvABC proteins and Holliday junction processing in *Escherichia coli*. J Bacteriol **178:**1237-41.

62. **Yamada, K., T. Miyata, D. Tsuchiya, T. Oyama, Y. Fujiwara, T. Ohnishi, H. Iwasaki, H. Shinagawa, M. Ariyoshi, K. Mayanagi, and K. Morikawa.** 2002. Crystal structure of the RuvA-RuvB complex: a structural basis for the Holliday junction migrating motor machinery. Mol Cell **10:**671-81.

63. **Zerbib, D., C. Mezard, H. George, and S. C. West.** 1998. Coordinated actions of RuvABC in Holliday junction processing. J Mol Biol **281:**621-30.

Figure 1.1 Mechanisms of tyrosine and serine site-specific recombinases. Though united by a requirement for target homology and a covalently bound protein-DNA intermediate, there are differences between the two groups. Integration catalyzed by the tyrosine recombinases (panel A and B) use the hydroxyl of a tyrosine residue to attack the phosphate backbone (panel A). This phospho-tyrosyl intermediate is subjected to the target DNA 5' hydroxyl attack. Because the activation of the tyrosine recombinase subunits is sequential, a Holliday junction intermediate is formed before resolution by the second pair of tyrosine recombinases. Inversion catalyzed by the serine recombinases (panel C and D) use the hydroxyl of a serine residue to attack the DNA phosphate backbone, leaving a 3' hydroxyl (panel C). Strand exchange is accomplished by a rotation of the DNA (panel D). Serine recombination is a concerted cleavage and rejoining with all four recombinases activated at once.

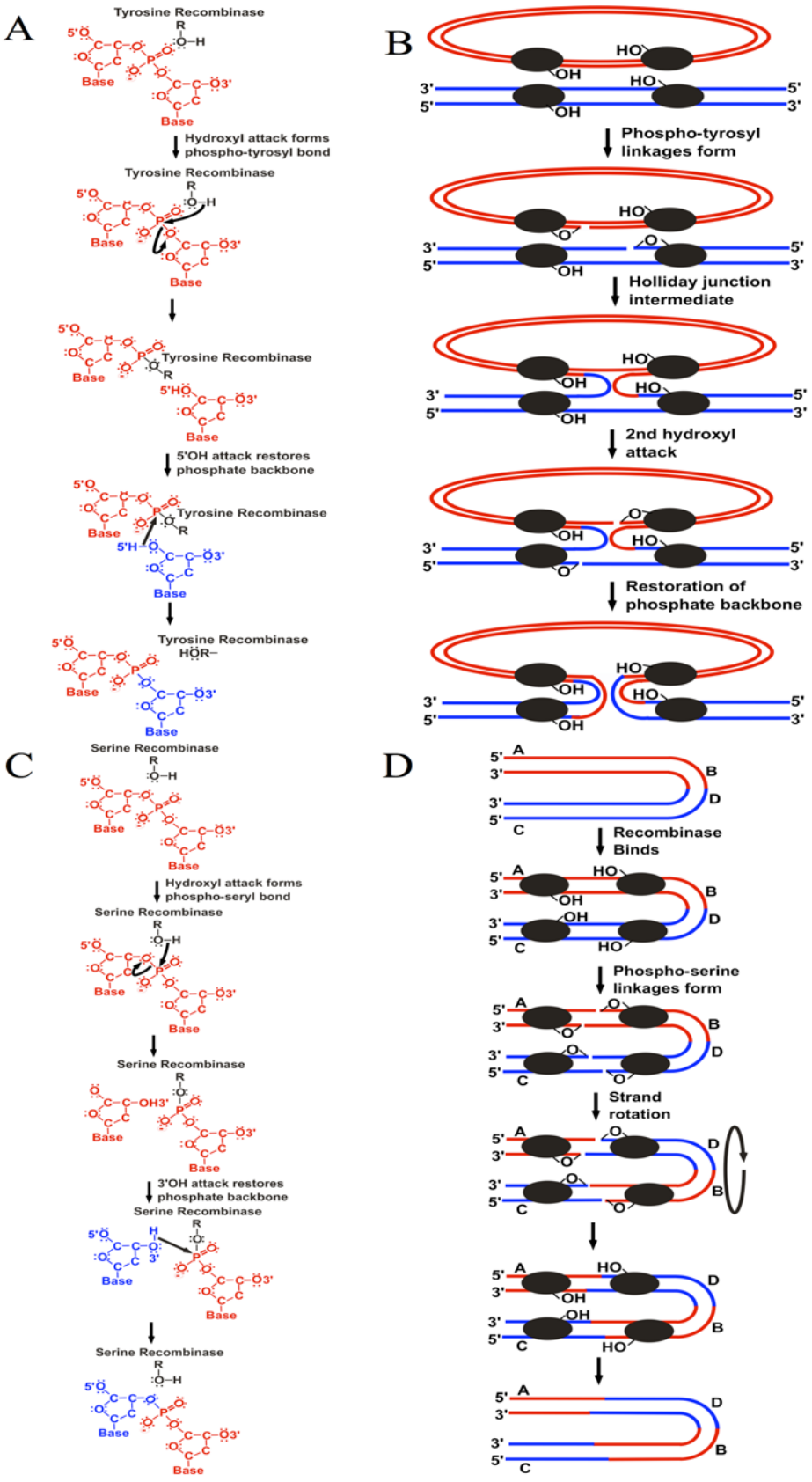

A
Tyrosine Recombinase
Hydroxyl attack forms phospho-tyrosyl bond
5'OH attack restores phosphate backbone
B
Phospho-tyrosyl linkages form
Holliday junction intermediate
2nd hydroxyl attack
Restoration of phosphate backbone
C
Serine Recombinase
Hydroxyl attack forms phospho-seryl bond
3'OH attack restores phosphate backbone
D
Recombinase Binds
Phospho-serine linkages form
Strand rotation

Figure 1.2. Transposition mediated by the DDE-motif recombinases. Hydrolysis is used to break the DNA-phosphate backbone (panel A) and release the element in cut and paste transposition (panel B). In the one-step transesterification, the 3'hydroxyls created in the hydrolysis reaction are used to attack the target DNA phosphate backbone (panel A). These attacks are staggered on the target backbone, therefore, after the element is covalently linked to the target, gaps are left to be filled in by host machinery, creating target site duplications.

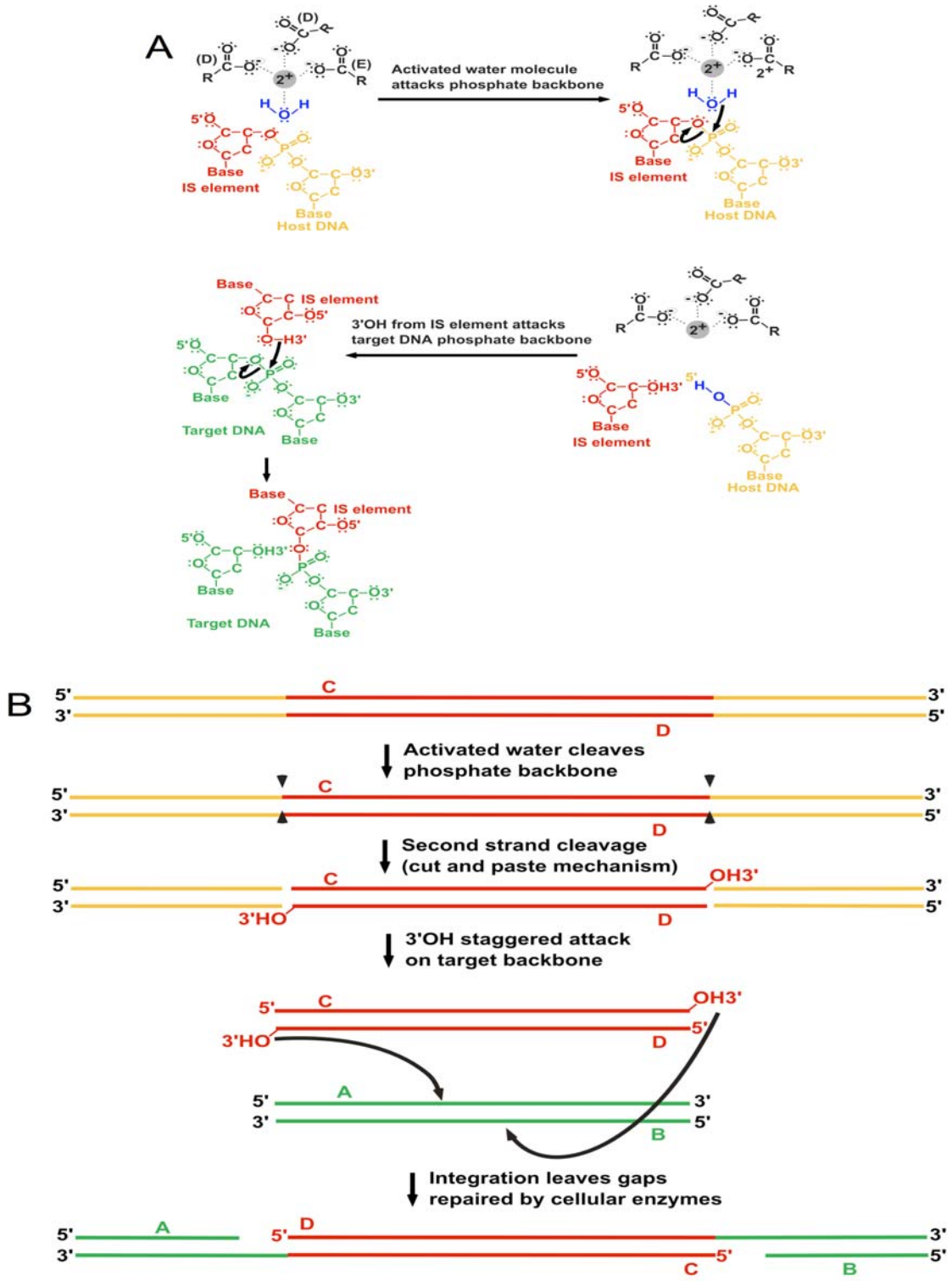

A
Activated water molecule
attacks phosphate backbone
Base
IS element
Base
Host DNA
3'OH from IS element attacks
target DNA phosphate backbone
Target DNA
B
Activated water cleaves
phosphate backbone
Second strand cleavage
(cut and paste mechanism)
3'OH staggered attack
on target backbone
Integration leaves gaps
repaired by cellular enzymes

Figure 1.3. RuvC action at a Holliday junction. RuvC uses the RNaseH-like fold to activate water molecules that are used in the hydrolysis of the phosphate backbone at the center of a Holliday junction. These nicks are repairable by ligase.

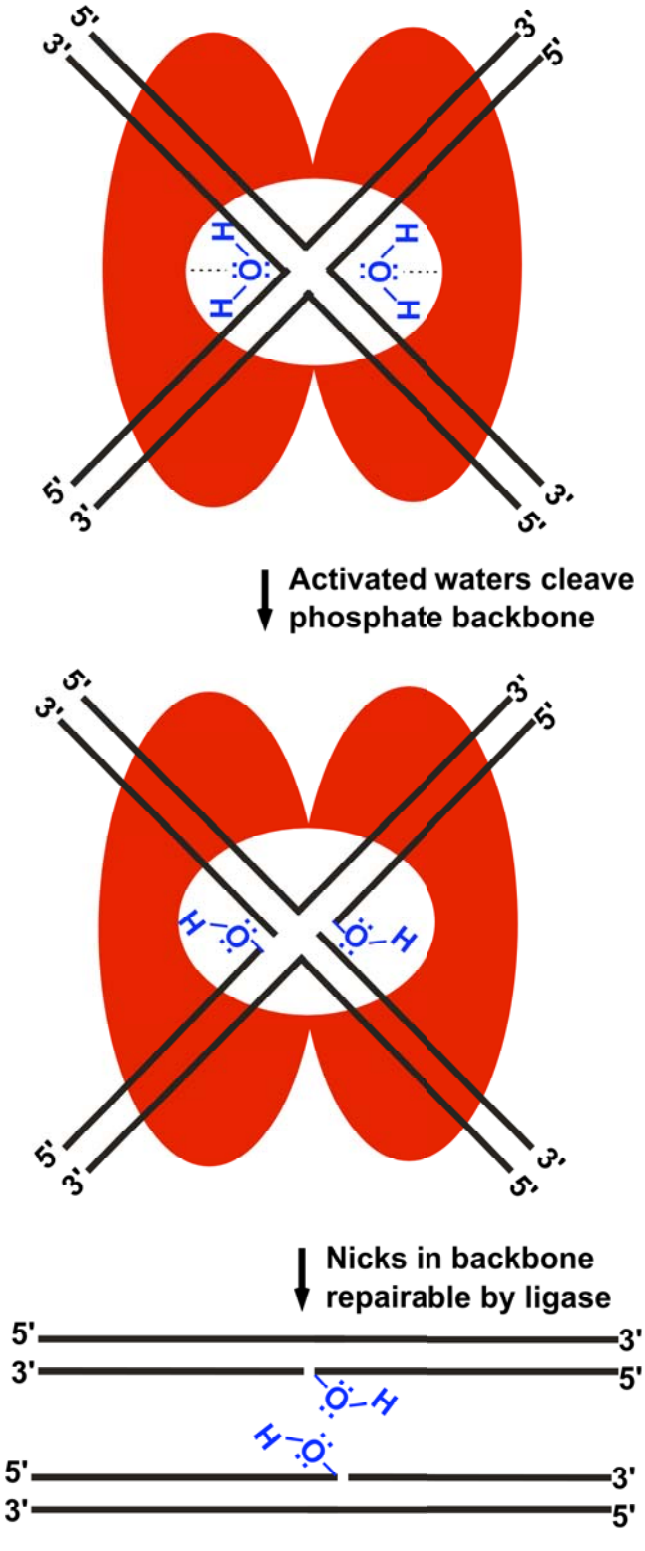

5'
3'
3'
5'
5'
3'
3'
5'
Activated waters cleave
phosphate backbone
5'
3'
3'
5'
5'
3'
3'
5'
Nicks in backbone
repairable by ligase
5'
3'
3'
5'
5'
3'
3'
5'

Figure 1.4. Piv Mediates Phase Variation of Type-4 Pili in *M. lacunata.* The *piv* gene is transcribed from P_{PIV}. In the Q orientation, the pilin genes *tfpQ* and *tfpB* are co-transcribed from P_{PILIN}, *tfpQ* encodes for type-4 pili. In the I orientation, *tfpI* is transcribed from P_{PILIN}, but does not encode for type-4 pili. Piv catalyzes this reversible DNA inversion. The boxes over the genes represent the genes transcribed in the Q and I orientations. The *invL* and *invR* binding sites are indicated by **(** and **)**, the *sub1* site is indicated by a X. Two putative Piv binding sites (31) are indicated by *

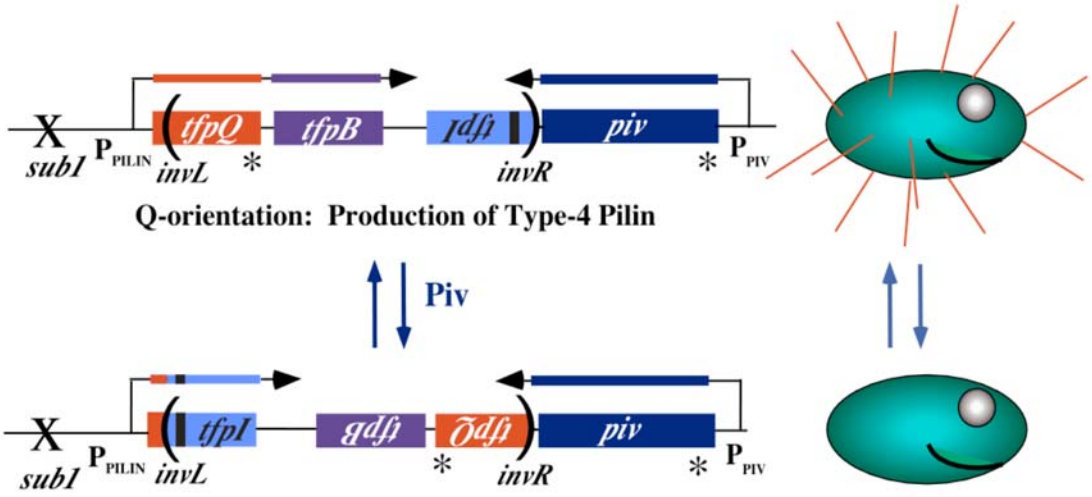

tfpQ
tfpB
tfpI
piv
X
subI
P PILIN
invL
invR
P PIV
Q-orientation: Production of Type-4 Pilin
Piv
tfpI
tfpB
tfpQ
piv
I-orientation: No Production of Type-4 Pilin

Figure 1.5. Primary amino acid sequence of Piv with key residues highlighted. Residues completely conserved sequences among the Piv/MooV family highlighted in blue, glutamic acid residues targeted for mutagenesis highlighted in red, the predicted leucine zipper highlighted in green, and the predicted HhH motif highlighted in orange. Residues targeted for mutagenesis have their position number and substitutions indicated below.

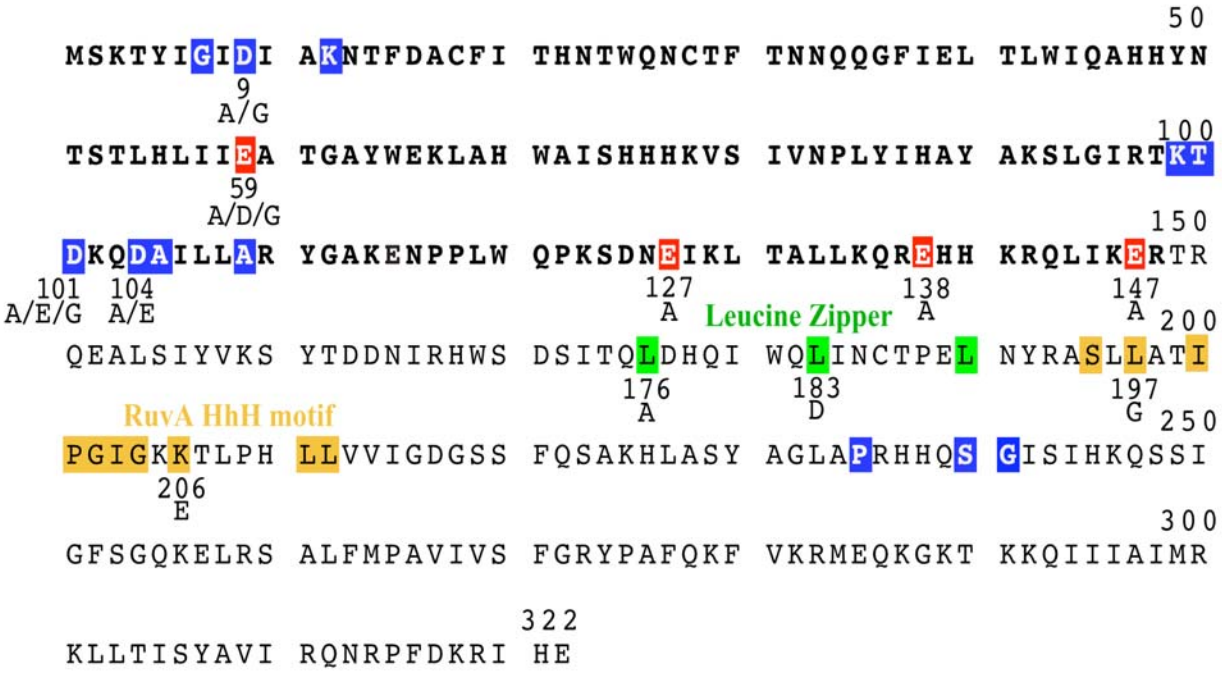
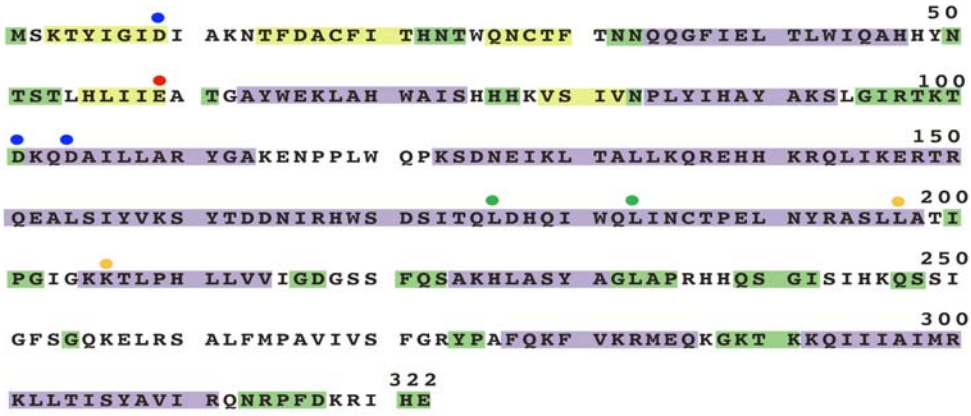

50
MSKTYIGIDI AKNTFDACFI THNTWQNCTF TNNQQGFIEL TLWIQAHHYN
9
A/G
100
TSTLHLIIEA TGAYWEKLAH WAISHHHKVS IVNPLYIHAY AKSLGIRTKT
59
A/D/G
150
DKQDAILLAR YGAKENPPLW QPKSDNEIKL TALLKQREHH KRQLIKERTR
101 104
A/E/G A/E
127
A
Leucine Zipper
138
A
147
A
200
QEALSIYVKS YTDDNIRHWS DSITQLDHQI WQLINCTPEL NYRASLLATI
176
A
183
D
197
G
RuvA HhH motif
250
PGIGKKTLPH LLVVIGDGSS FQSAKHLASY AGLAPRHHQS GISIHKQSSI
206
E
300
GFSGQKELRS ALFMPAVIVS FGRYPAFQKF VKRMEQKGKT KKQIIIAIMR
322
KLLTISYAVI RQNRPFDKRI HE

Figure 1.6. Secondary structure predictions of Piv. The primary sequence highlighted in yellow indicated a prediction of a beta strand. Sequence highlighted in purple is predicted to be alpha helices. Green highlighted sequence is predicted to be a loop. The aspartates and glutamate residues predicted to be part of the RNaseH-like fold are indicated with blue and red dots over the residues. Predicted leucine zipper and HhH motif residues target for mutagenesis are indicated with green and orange dots over the residues.


50
MSKTYIGIDI AKNTFDACFI THNTWQNCTF TNNQQGFIEL TLWIQAHHYN
100
TSTLHLIIEA TGAYWEKLAH WAISHHHKVS IVNPLYIHAY AKSLGIRTKT
150
DKQDAILLAR YGAKENPPLW QPKSDNEIKL TALLKQREHH KRQLIKERTR
200
QEALSIYVKS YTDDNIRHWS DSITQLDHQI WQLINCTPEL NYRASLLATI
250
PGIGKKTLPH LLVVIGDGSS FQSAKHLASY AGLAPRHHQS GISIHKQSSI
300
GFSGQKELRS ALFMPAVIVS FGRYPAFQKF VKRMEQKGKT KKQIIIAIMR
322
KLLTISYAVI RQNRPFDKRI HE

CHAPTER 2

Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases [1]

[1] Buchner, J. M., Robertson, A. E., Poynter, D. J., Denniston, D. J., Denniston, S. S., Karls, A. C. 2005. *Journal of Bacteriology*. 187: 3431-3437. Reprinted here with permission of the publisher.

Abstract

Piv, a unique prokaryotic site-specific DNA invertase, is related to transposases of the insertion elements from the IS*110*/IS*492* family and shows no similarity to the site-specific recombinases of the tyrosine- or serine-recombinase families. Piv tertiary structure is predicted to include the ribonuclease H-like fold that typically encompasses the catalytic site of the recombinases or nucleases of the retroviral integrase superfamily, including transposases and RuvC-like Holliday junction resolvases. Analogous to the DDE and DEDD catalytic motifs of transposases and RuvC, respectively, four Piv acidic residues D9, E59, D101, and D104 appear to be positioned appropriately within the ribonuclease H fold to coordinate two divalent metal cations. This suggests mechanistic similarity between site-specific inversion mediated by Piv and transposition or endonucleolytic reactions catalyzed by enzymes of the retroviral integrase superfamily. The role of the DEDD motif in Piv catalytic activity was addressed using Piv variants that are substituted individually or multiply at these acidic residues and assaying for *in vivo* inversion, intermolecular recombination, and DNA binding activities. The results indicate that all four residues of the DEDD motif are required for Piv catalytic activity. The DEDD residues are not essential for *inv* recombination site recognition and binding; but this acidic tetrad does appear to contribute to the stability of Piv-*inv* interactions. Based on these results, a working model for Piv-mediated inversion that includes resolution of a Holliday junction is presented.

Introduction

Piv catalyzes site-specific inversion of a 2.1 kb chromosomal segment encoding type 4 pilin genes of the human and cow eye pathogens, *Moraxella lacunata* and *Moraxella bovis*, respectively (13, 17, 26). This DNA rearrangement results in phase variation of the type 4 pili (18) (Fig. 2.1). The invertible segment is bounded by identical 32 bp inverted repeats, *invL* and *invR*, where Piv mediates recombination. In addition, Piv interacts with an accessory site, *sub1*, which may facilitate assembly of an active synaptic complex (31).

Piv is unique among prokaryotic site-specific DNA invertases because it does not belong to either the tyrosine- or serine-recombinase families. Instead, Piv shares significant amino acid identity and similarity with the transposases of insertion sequences (IS) from the IS*110*/IS*492* family, including three highly conserved amino acid motifs: G-D--K, KTD--DA---A, and P----SG (13, 23). Piv variants containing alanine or glycine substitutions at multiple positions within each motif are defective for mediating inversion, but still bind to the *inv* and *sub* sites. The essential role of these conserved motifs in Piv catalysis of DNA inversion suggests that Piv and the transposases of the IS*110*/IS*492* family (Piv/MooV family of DNA recombinases) use similar catalytic mechanisms (24, 30).

Most identified transposases belong to the retroviral integrase super family and are typified by a conserved DDE catalytic motif. The acidic residues of the DDE motif are separated by variable numbers of amino acids in the transposase primary amino acid sequence but are brought together by protein folding to form a catalytic triad within a ribonuclease H-like structural motif (RNase H-fold) (20). The primary amino acid sequences of Piv/MooV recombinases do not contain a conserved DDE-motif; however, molecular modeling of the tertiary structure for the amino terminal region of Piv predicts an RNase H-fold (30). The completely conserved acidic

residues D9 (G-***D***--K), E59 (conserved as E or D), D101, and/or D104 (KT***D***--***D***A) are positioned appropriately in the RNase H-fold to coordinate two divalent metal ions. This potential DED or DEDD catalytic motif for Piv profoundly affects the predicted chemistry for the Piv-mediated site-specific recombination reaction.

Unlike the tyrosine- and serine-recombinases that mediate conservative site-specific recombination by a two-step transesterification reaction, in which the energy of the phosphodiester bond is conserved in a covalent protein-DNA intermediate, the DDE-motif transposases catalyze hydrolysis of the phosphodiester backbone of the DNA substrate and mediate strand transfer by a one-step transesterification reaction. Host enzymes, such as DNA repair or replication functions, are required to complete the transposition process following strand transfer (for reviews see Mizuuchi, 1997; Nagy and Chandler, 2004; Turlan and Chandler, 2000). The RuvC-related Holliday junction resolvases, which are also members of the retroviral integrase superfamily, utilize a DEDD catalytic tetrad (D7, E66, D138, and D141) within a RNaseH-like fold for coordination of two divalent metal cations (1). Overlapping the crystal structures of the catalytic domains from RuvC and retroviral integrases (IN) shows that three residues of the DEDD motif superimpose on the DDE residues of IN within the RNaseH-fold (20, 33). Like the DDE-motif transposases, RuvC coordinates divalent metal cations to direct hydrolysis of phosphodiester bonds in substrate DNA (reviewed in (20). However, RuvC does not catalyze strand transfer following hydrolysis of the substrate DNA, i.e. RuvC cleaves opposing DNA strands in a Holliday junction to resolve the structure and DNA ligase repairs the nicks in the substrate DNA (3, 27). RuvC exhibits sequence specificity, preferring to resolve junctions with the sequence 5'-$^{A}/_{T}$TT$^{G}/_{C}$-3' (6). The sequence 5'-ATTG-3' is near the center of the *inv* sites, suggesting that, if Piv has RuvC-like activity, a Holliday junction generated in

recombination between *invL* and *invR* sites may be efficiently resolved to give the apparently conservative recombination product.

These intriguing mechanistic predictions for Piv led us to determine the role of the DEDD residues in Piv-mediated inversion and intermolecular recombination. Piv variants with individual substitutions of a small, uncharged amino acid at each potential catalytic residue retain no inversion or intermolecular recombination activity, but still bind the *inv* recombination site. These results indicate that the DEDD motif is part of the Piv active site. Substitutions of residues in the DEDD motif with glutamic acid or aspartic acid, which differ by one carbon in side-chain length, result in severely reduced or undetectable Piv catalytic activity, suggesting that there is little flexibility in the spacing of the carboxyl groups within the catalytic tetrad. Interestingly, the variants that retain carboxylates at the predicted catalytic site appear to form more stable Piv-*inv* complexes. A working model for Piv-catalyzed recombination is proposed.

Materials and Methods

Site-directed mutagenesis and expression of Piv. Piv variants with the amino acid substitutions listed in Table 2.1 were generated by site directed mutagenesis (Stratagene QuikChange Mutagenesis Kit) of *piv* in pAG630 (30) or in pJMB30, which has the *Spe*I to *Hin*dIII fragment containing *piv* from pAG800.2 (this study) ligated into the same restriction sites of Litmus 29 (NEB, New England Biolabs). The mutagenesis oligonucleotides incorporated base pair changes that resulted in the designated missense mutation and, for some, silent mutations that created or eliminated restriction sites in *piv* for screening mutants. Mutagenized *piv* genes were subcloned into two different expression vectors for *in vivo* recombination (pAG800.2) and *invL*-binding (pJMB52) assays. pAG800.2 contains the *Bam*HI-*Hin*dIII

fragment from pAG702 (30), encoding *piv* and its Shine-Delgarno site, introduced downstream of P_{tac} in pKH197 (24), and one of two *Spe*I sites in the pKH197 vector was destroyed by fill-in with Klenow fragment (NEB). pJMB52 is pKH197 with the *Bam*HI-*Hin*dIII region replaced by the *Bam*HI-*Bsr*GI fragment of pACYC-*piv*1 (13), containing the *M. bovis piv* promoter (P_{piv})and the first 86 bp of piv_{mb} and the *Bsr*GI-*Hin*dIII fragment of pAG800, encoding the remainder of piv_{ml} and *invR*. Thus, pJMB52 differs from the pAG800.2 expression vector in two aspects: 1) *piv* is expressed from *M. bovis* P_{piv} and 2) *piv* encodes one of the 5 amino acid differences between *M. bovis* Piv and *M. lacunata* Piv, T21A. The T21A substitution does not affect Piv recombination or DNA binding activity (data not shown).

Two mutant *piv* expression vectors were created to use as negative controls in the recombination and DNA binding assays, pJMB54 and pNull. The mutant *piv* gene inserted into pAG800.2 to give pJMB54 has a 10 bp insertion at the *Hinc*II site in *piv* due to the deletion of the interposon $\Omega Sm^R/Sp^R$ from *piv* in pMxL5 (17) using *BamH*I digestion and religation; the insertion causes a frame shift 245 base pairs into *piv* and results in termination of *piv* translation after 103 amino acids (PivΔ1). To generate pNull, pJMB52 was digested with *BsrG*I and *Hind*III, treated with End-It DNA End-Repair Kit (Epicentre), and ligated to restore the vector without the sequence between *BsrG*I-*Hind*III. This resulted in a carboxyl-terminal truncation of Piv leaving only the amino-terminal 28 amino acids (PivΔ2).

All of the mutant *piv* alleles and the subclones were confirmed by sequencing of the complete *piv* gene (University of Michigan Biomedical Research Core Facilities).

Expression of wild type and variant Piv proteins was as follows: cultures of DH5α carrying pAG800.2 or pJMB52 encoding the wild type or mutated *piv* genes were grown to mid-log phase at 37°C in Luria-Bertani broth (LB) with 50 μg/ml spectinomycin (Sp_{50}), induced with

100 μM IPTG, incubated/aerated two additional hours (hr), and 1.5 ml was harvested by centrifugation. Cells were resuspended in 270 μl 1 x loading buffer (2), boiled for 5 min and 90 μl of each sample was electrophoresed on a 12% SDS-polyacrylamide gel. Western blot analysis was performed as described in (2) using primary anti-Piv antibody, generated in rabbits against a peptide CKSDNGIKLTALLKQREHHKRQLIKERTRQE conjugated to KLH (BIO SYNTHESIS Inc., Lewisville, TX), and secondary anti-rabbit antibody alkaline phosphatase conjugate (Sigma-Aldrich) at 1:1000 dilutions. Membranes were developed with 5-bromo-4-chloro-3-indolyphosphate/nitroblue tetrazolium (BCIP/NBT) SigmaFast Tablets (Sigma-Aldrich).

***In vivo* inversion assays.** The inversion substrates, pMxL90 (Q-orientation substrate) and pMxL100 (I-orientation substrate), were derived from pMxL5 and pMxL6, respectively, (17) by replacing the $\Omega Sm^R/Sp^R$ cassette that interrupts *piv* with the ΩCm^R (chloramphenicol resistance) interposon from pHP45Ω-Cm (8). Chemically competent *E. coli* DH5α cells containing either pMxL90 or pMxL100 were transformed with pAG800.2 or a Piv variant expression vector. Transformants were plated on LB agar containing Sp_{5o} Cm_{34}, and 50 μM IPTG. After 24-36 hr incubation at 37°C individual colonies were inoculated into LB broth $Sp_{50}Cm_{34}$ and incubated/aerated at 37°C for approximately 24 hr. The inversion substrate and expression vector were isolated using Wizard Plus Miniprep kits (Promega). Two different methods were used to detect inversion of the *invL/invR*-flanked segment on the substrate plasmid: 1) 200-300 ng of the isolated plasmid DNA was digested with *Bsr*GI, electrophoresed on 0.6% agarose gels, stained with ethidium bromide (EtBr), and imaged with the BioRad Fluor-S Multi-imager (BioRad Quantity One software was used to determine relative intensities of bands to estimate the %

inversion); 2) 20-50 ng of the isolated DNA was used in a three primer PCR reaction containing 1X KlenTaq DV ReadyMix (Sigma Aldrich) and 200 nM of each primer (T7 5'-GTAATACGACTCACTATAGGGC-3', IAR 5'-CTAACCATCAGCTATGCCGTTATTC-3' and IAF5 5'-CATGATATGCTGCTTGACCCCAACC-3'); 5 μl of each PCR reaction (25 μl total volume which was cycled in a BioRad i-cycler for 1 cycle at 95^{o}C 60 sec, 25 cycles at 95^{o}C 45 sec/61^{o}C 30 sec/72^{o}C 45sec, and 1 cycle at 72^{o}C 7min) was electrophoresed on a 1.2% agarose gel, and visualized as described above. Each Piv variant was tested for inversion from the Q to I and I to Q orientations from at least three independent transformants of DH5α containing pMxL90 or pMxL100.

Intermolecular recombination assays. Intermolecular recombination between *invR* on pAG800.2 and *invL* on pMxL90 or pMxL100 was assayed by PCR: 25 μl reactions contained 1X Taq Buffer B (Fisher), 1.5 mM $MgCl_2$, 1 μM of each primer (CAP 5'-GGCTGGCTTTTTCTTGTTATCGC-3', BLU 5'-GGGTTATTGTCTCATGAGCGG-3'), 0.2 μM dioxynucleotide triphosphates (dNTPs), 10-25 ng of DNA isolated from the inversion assay transformants (described above), and 1.5 Units *Taq* DNA polymerase (Fisher). Cycling conditions for the reactions were: 1 cycle at 95^{o}C-60 sec, 25 cycles at 95^{o}C-30 sec/64^{o}C-30 sec/68^{o}C-36 sec, and 1 cycle at 68^{o}C-36 sec; products were electrophoresed and imaged as described above. Selected products were inserted into pCR2.1 (Invitrogen) and sequenced (UM-BRCF) to determine the site of DNA exchange. These assays were repeated with Pfu (Stratagene) instead of *Taq* polymerase with the following changes: the reaction mix contained 1 x Pfu Buffer (Stratagene), 0.5 μM of each primer, 0.25 μM dNTPs, and 1.9 Units *Pfu* DNA polymerase, and the cycling conditions were changed to 1 cycle at 95^{o}C-60 sec, 25 cycles at

95°C-30 sec/64°C-30 sec/72°C-2 min 22 sec, and 1 cycle at 72°C 10min. Products were analyzed as above. The DNA from at least three independent inversion assays for each Piv variant were analyzed for intermolecular recombination products.

***In vivo* DNA binding assays.** To measure *in vivo* binding of Piv to *inv*, the plasmid system of Elledge and Davis (7) was utilized. A double-stranded oligonucleotide encoding the *inv* sequence (top strand: 5'-GCCATTATTGGTATCCTAGCTGCAATCGCT-3') was inserted into the *Sma*I site at +1 of the *con*II promoter sequence in pNN396 (7). The *Not*I-*Hind*III restriction fragment, containing the *con*II promoter and the *inv* sequence, was then ligated into the same restriction sites on pNN387 upstream of *lacZ lacY* to create pAR110. DH5α containing pAR110 was transformed with Piv wild-type or Piv variant expression vectors; cultures of the transformants in LB $Sp_{50}Cm_{34}$ were grown to $O.D._{600}$ of 0.7 to 0.9. Cells were harvested from two separate 1 ml aliquots by centrifugation: one pellet was resuspended in 1 ml Z buffer to use in β-galactosidase activity assays (19) and the other was resuspended in 1 ml Z buffer without β-mercaptoethanol to measure total protein concentration with BCA Protein Assay (Pierce). In both assays the cells were lysed by addition of two drops chloroform and 1 drop of 0.1% sodium dodecyl sulfate. All samples were assayed in duplicate for 1:2 and 1:10 dilutions in Z buffer. Each Piv variant was assayed from at least three independent transformants and the wild type Piv and PivΔ2 were assayed from 8 independent transformants. The β-galactosidase activity is expressed as Miller units (19) per μg protein in each sample.

Results

DEDD motif residues are all required for Piv inversion activity. Previous mutational analysis of the conserved amino acid motifs from the Piv/MooV recombinase family utilized multiple substitutions within each conserved motif in Piv, leading to conclusions about the contributions of the motif as a whole in Piv catalysis of inversion (30). To address the role of individual amino acids within the conserved motifs, focusing on those residues that are predicted to constitute a DED or DEDD catalytic motif, Piv variants were generated that are singularly substituted at D9, E59, D101, and D104 with alanine, glycine, aspartic acid, or glutamic acid (Fig. 2.2). All of these Piv variants were expressed at levels comparable to wild type Piv, as measured by Western blot analysis (Fig. 2.3), suggesting that the variant proteins can fold appropriately. In contrast, none of these Piv variants gave detectable levels of *in vivo* inversion activity in an assay that directly measured inversion of the pilin segment on a plasmid (Fig. 2.4). The limit of detection in this assay is ~1% inversion; wild-type Piv exhibits ~30% inversion of the pilin segment on the plasmid substrate. The results for switching from the *tpfQ* (Q) to *tfpI* (I) orientation for the invertible segment is shown in Fig. 2.4; assaying inversion from I to Q yielded identical results (data not shown). These results support the proposal that the DEDD residues comprise a catalytic tetrad within Piv.

Each of the three glutamic acid residues that are 23, 34, and 43 residues beyond D104 is spaced appropriately in the primary amino acid sequence to complete a DDE motif in which the first two residues consist of D9 and D101/D104. However, this potential catalytic motif is not predicted in the molecular modeling of Piv nor is it conserved among the Piv/MooV family members. Substitution of E127, E138, and E147 individually with alanine or glycine, did not

significantly reduce Piv expression (Fig. 2.3) or inversion activity (Fig. 2.4), indicating that these acidic residues do not play an essential role in Piv-mediated recombination.

Piv variants with carboxylate substitutions at E59 or D101 retain low levels of catalytic activity. While D9, D101, and D104 are completely conserved in the Piv/MooV recombinases, a subgroup of the family has an aspartic acid at the E59 position. Therefore, the possibility that DDDD could serve equally well as DEDD to form a catalytic tetrad in Piv was tested with the Piv variant E59D. Although this variant did not mediate detectable inversion in the assay described above (Fig. 2.4), a PCR-based assay for intermolecular recombination revealed that Piv E59D does catalyze recombination between *invL* on the inversion substrate and *invR* on the Piv expression vector albeit at a low level (Fig. 2.5). Therefore, to investigate whether Piv E59D mediates inversion, but at a level below the detection limits of the plasmid restriction digest assay, PCR was utilized to detect inversion products in the plasmid DNA isolated from the *in vivo* inversion experiment. As seen in Fig. 2.6, there is inversion product from the Piv E59D assay; this product was detected at low levels in five out of 6 independent *in vivo* inversion assays (3/3 Q to I assays and 2/3 I to Q assays, data not shown). However, these PCR assays also revealed that Piv-independent inversion occurs at very low levels (Fig. 2.6). PCR-mediated recombination or template switching could explain the Piv-independent inversion products that are detected by PCR, but it has been demonstrated that Pfu DNA polymerase does not mediate recombination under normal PCR conditions (28), and the same results are obtained with Taq and Pfu DNA polymerases (data not shown). Replication-dependent, recombination-independent inversion/dimerization of plasmids containing long inverted repeats has been characterized previously (5, 16) and likely explains the observed levels of Piv- and RecA-independent inversion.

Based on the molecular modeling of Piv, either D101 or D104 could participate in a catalytic DED triad, similar to the DDE motif of Transposases. The substitution of either carboxylate with a small, uncharged amino acid results in loss of Piv catalytic activity as measured by both PCR-based recombination assays (Figs. 2.5 and 2.6). This result suggests that both D101 and D104 are required as part of a catalytic DEDD tetrad, as seen for the RuvC-related resolvases (29). Substitution of the D101 position with another carboxylate (D101E) is partially tolerated, yielding low levels of recombination activity, while replacement of D104 with glutamic acid results in loss of all recombination activity (Figs. 2.5 and 2.6). In addition, rotating the position of these carboxylates with E59 to create Piv variants with DDED (E59D/D101E) and DDDE (E59D/D104E) motifs resulted in stable (Fig. 2.3) but catalytically inactive proteins (Figs. 2.4, 2.5, and 2.6), indicating that the spacial positions of the carboxyl group of D101 and D104 are important for Piv catalytic activity.

DEDD residues are not essential for binding the recombination sites. Although *in vitro* binding assays have demonstrated that Piv binds weakly to the *inv* recombination sites (31), Piv interactions with *invL* can be assessed *in vivo* using a transcription repression system (7, 30). Expression of *lacZ* from a constitutive promoter, P_{conII} (7), on a single copy vector is controlled by Piv binding to *invL* sequence that functions as an operator site. Thus, a higher level of β-galactosidase activity reflects a lower level of Piv binding to the *invL* site. The basal level of β-galactosidase activity is determined with binding-defective PivΔ2 expressed from pNull (Fig. 2.7). Wild-type Piv and Piv variants with acidic amino acid substitutions for the predicted catalytic residues (E59D, D101E, and D104E) bound to *invL*, resulting in a 50 to 57% reduction in β-galactosidase activity relative to the basal level (Fig. 2.7). Interestingly, variants with

alanine and glycine substitutions of DEDD-motif residues bound *invL*, but exhibited lower repression levels; binding of D9A, D9G, E59A, D101A, D101G, and D104A Piv variants reduced β-galactosidase activity by 24, 32, 26, 35, 41, and 38%, respectively (Fig. 2.7). Substitution of the non-conserved residues with alanine or glycine did not affect Piv binding to *invL* (55 to 62% reduction in β-galactosidase activity; Fig. 2.7). The DDED and DDDE double-substitution variants reduced β-galactosidase activity by 41 and 52%, respectively (Fig. 2.7). These results indicate that the predicted catalytic residues are not essential for recognizing and binding the recombination sites, but the DEDD motif may contribute to protein conformation or protein-DNA interactions that stabilize interactions of Piv with the *inv* sites.

Discussion

An acidic residue tetrad, DEDD, is required for Piv catalysis of recombination. We previously hypothesized, based on molecular modeling of Piv tertiary structure and mutagenesis of the conserved amino acid motifs from the Piv/MooV family, that Piv utilizes a DED motif that is equivalent to the DDE motif of transposases (30). While the model clearly predicted D9 and E59 to be the first two catalytic residues, it was possible that either D101 or D104 was the third residue of the catalytic motif. The substitution studies herein show that D101 and D104, in addition to D9 and E59, are essential residues in the Piv active site. D9A, D9G, E59A, D101A, D101G, D104A, and D104E substitutions resulted in Piv variants that are completely defective for *in vivo* inversion and intermolecular recombination. Piv E59D and Piv D101E exhibit very low levels of recombination and generating a DDE-, versus DED-, motif in the linear amino acid sequence of Piv (Piv E59D/D101E and Piv E59D/D104E) did not yield an active recombinase. These results are consistent with substitution analyses of the DDE-motif in Tn*5* Tnp and the

DEDD-motif in the Holliday junction resolvase RuvC (25, 27). These studies showed that substitution of catalytic residues with non-acidic amino acids eliminates recombination activity, but interchanging aspartic acid and glutamic acid residues at catalytic positions (other than the first aspartic acid) is tolerated to give low levels of *in vitro* catalytic activity. Peterson and Reznikoff (25) concluded that small shifts in the spatial positions of the carboxyl groups within the Tn*5* Tnp active site decrease the efficiency with which two divalent metal cations are coordinated for catalytic reactions (15). It is likely that switching of the carboxylate residues in the DEDD tetrad of Piv also alters optimal spacing of the carboxyl groups involved in coordination of ions at the catalytic site.

Role for catalytic residues in Piv interactions with the recombination sites. The results of *in vivo* binding assays with DEDD-motif Piv variants that were individually substituted at each position of the motif demonstrate that these acidic residues are not essential for Piv binding to the recombination site. However, the nature of substitutions at these positions did influence *in vivo* binding. As measured by repressor/operator activity of Piv/*invL* complexes on *lacZ* expression from P_{conII}, replacing E59, D101, and D104 with acidic residues gave the same level of repression that was obtained with wild-type Piv, however substitutions with alanine or glycine reduced Piv-mediated repression by 14 to 24%. Substitution of the non-conserved residues E127, E138, and E147 with alanine or glycine did not affect binding of Piv to *invL* in these *in vivo* assays. These results indicate that the catalytic residues may play a role in stabilizing the interactions of Piv with the recombination site. A similar role for the DDE motif of Tn*10* Tnp in target DNA binding has been suggested by Junop and Haniford (10). Although the DDE motif is not required for target site selection by Tn*10* Tnp, the DDE residues are needed for capture of a

suboptimal target site (10). The function of these acidic residues in stabilizing recombinase/target interactions involves coordination of the divalent metal ions, which could create bridging contacts between the transposase and the substrate DNA or might stabilize the optimal transposase conformation for DNA binding (10).

A working model for Piv-mediated recombination. Our results demonstrate that all four residues of the DEDD motif in Piv are required for catalysis of inversion and intermolecular recombination. Thus, the catalytic domain of Piv probably more directly resembles that of RuvC than that of the DDE-motif transposases. But what does this imply about the mechanism for Piv-mediated recombination? The catalytic domain of RuvC is remarkably similar to that of the DDE-motif transposases (33). The DEDD and DDE catalytic residues of RuvC and transposases, respectively, are positioned within the RNaseH-fold to coordinate divalent metal cations that direct hydrolysis of the phosphodiester bond in substrate DNA, generating a free 3' OH end (reviewed in (20). A primary difference between the activity of RuvC and transposases is that RuvC does not catalyze strand transfer following hydrolytic cleavage of the Holliday junction, while transposases utilize the 3' OH of cleaved donor DNA as the nucleophile to attack the phosphodiester bond of target DNA in strand transfer. To accomplish this polynucleotidyl transfer reaction transposases retain the 3' end of the cleaved donor DNA in the active site, blocking the entry of another nucleophile, and bind the target DNA within the same active site to serve as substrate for the one-step transesterification reaction (11). The third aspartic acid in the RuvC catalytic site may preclude binding of both the cleaved 3' OH end and a target DNA strand within the active site by immobilizing the first nucleophile and/or the 3' OH end, thus preventing the adjustments needed to accommodate the target DNA strand. Alternatively, the absence of a

RuvC-mediated strand transfer may simply reflect that there is no target strand available for binding to the catalytic site within the constrained structure of the RuvC-bound Holliday junction.

Thus, in our working model for Piv-mediated inversion we utilize both hydrolysis/strand transfer and endonucleolytic activities for the DEDD active site in the reaction leading to site-specific recombination. The model (Fig. 2.8) shows Piv bound as a dimer to the synapsed recombination sites (it is also possible that dimers are bound to each *inv* site and synapse, forming a tetramer). Binding and DNA cleavage/strand transfer induces a conformational change in the Piv-DNA complex so that new recombinase-DNA contacts promote formation of a specific Holliday junction structure, an activity demonstrated for RuvC (4) and other Holliday junction resolvases (9, 32). The active sites of the Piv dimer are now positioned to cleave the unexchanged strands. DNA ligase may repair the original nick before or after the RuvC-like activity that leads to resolution of the junction and is also required for repair of the nicks in the recombined DNA strands. It is arguable that the Piv/nicked-DNA complex is stable prior to DNA ligase activity based on the robust nature of some intermediates in transposition systems, such as the bacteriophage Mu transpososome which needs ClpX unfolding activity to disassemble (remodel) the strand transfer complex (reviewed in (21).

The initial cleavage and strand transfer steps of this working model are similar to those leading to branched DNA intermediates in transposition of IS*911* (14) and IS*30* (12) adjacent to a copy of the IS inverted repeat. It has been demonstrated that the IS*30* Tnp can mediate inversion of DNA between two dyad symmetric sites containing IS*30* inverted repeats (12). In the models for recombination mediated by the IS*911* and IS*30* Tnp (14, 22) host functions process the branch intermediates set up by transposases-mediated strand transfer. In our model

for Piv-mediated inversion Piv promotes formation of the Holliday junction and Piv, rather than host functions, specifically resolves the junction close to the original strand transfer and without significant heteroduplex DNA formation. We are currently testing these predictions of the model utilizing both genetic and biochemical approaches.

Acknowledgements

This work was supported by National Institutes of Health Grant GM49794 (to ACK). We thank Ellen Neidle for critical review of this manuscript, Gordon Churchward and Mick Chandler for helpful discussions, Weiping Fu for constructing Piv E147A expression plasmid, Laurie Barsh Dirksen for generating pAG800 and pAG801, and How-Yi Chang and Valerie Gordon (NSF REU Summer Program at UGA participant) for their efforts on this project.

References

1. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, H. Nakamura, H. Shinagawa, and K. Morikawa.** 1994. Atomic structure of the RuvC resolvase: a Holliday junction-specific endonuclease from E. coli. Cell **78:**1063-72.

2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1991. Current Protocols in Molecular Biology. John Wiley & Sons, Inc., New York.

3. **Bennett, R. J., H. J. Dunderdale, and S. C. West.** 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell **74:**1021-31.

4. **Bennett, R. J., and S. C. West.** 1995. Structural analysis of the RuvC-Holliday junction complex reveals an unfolded junction. J Mol Biol **252:**213-26.

5. **Bi, X., and L. F. Liu.** 1996. DNA rearrangement mediated by inverted repeats. Proc Natl Acad Sci U S A **93:**819-23.

6. **Eggleston, A. K., and S. C. West.** 2000. Cleavage of Holliday junctions by the *Escherichia coli* RuvABC complex. J Biol Chem **275:**26467-76.

7. **Elledge, S. J., and R. W. Davis.** 1989. Position and density effects on repression by stationary and mobile DNA-binding proteins. Genes Dev **3:**185-97.

8. **Fellay, R., J. Frey, and H. Krisch.** 1987. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for *in vitro* insertional mutagenesis of gram-negative bacteria. Gene **52:**147-54.

9. **Giraud-Panis, M. J., and D. M. Lilley.** 1998. Structural recognition and distortion by the DNA junction-resolving enzyme RusA. J Mol Biol **278:**117-33.

10. **Junop, M. S., and D. B. Haniford.** 1997. Factors responsible for target site selection in Tn*10* transposition: a role for the DDE motif in target DNA capture. Embo J **16:**2646-55.

11. **Kennedy, A. K., D. B. Haniford, and K. Mizuuchi.** 2000. Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity. Cell **101:**295-305.

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

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

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

15. **Lovell, S., I. Y. Goryshin, W. R. Reznikoff, and I. Rayment.** 2002. Two-metal active site binding of a Tn*5* transposase synaptic complex. Nat Struct Biol **9:**278-81.

16. **Lyu, Y. L., C. T. Lin, and L. F. Liu.** 1999. Inversion/dimerization of plasmids mediated by inverted repeats. J Mol Biol **285:**1485-501.

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

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

19. **Miller, J. H.** 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

20. **Mizuuchi, K.** 1997. Polynucleotidyl transfer reactions in site-specific DNA recombination. Genes Cells **2:**1-12.

21. **Nakai, H., V. Doseeva, and J. M. Jones.** 2001. Handoff from recombinase to replisome: insights from transposition. Proc Natl Acad Sci U S A **98:**8247-54.

22. **Olasz, F., T. Fischer, M. Szabo, Z. Nagy, and J. Kiss.** 2003. Gene conversion in transposition of Escherichia coli element IS30. J Mol Biol **334:**967-78.

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

24. **Perkins-Balding, D., G. Duval-Valentin, and A. C. Glasgow.** 1999. Excision of IS492 requires flanking target sequences and results in circle formation in Pseudoalteromonas atlantica. J Bacteriol **181:**4937-48.

25. **Peterson, G., and W. Reznikoff.** 2003. Tn*5* transposase active site mutations suggest position of donor backbone DNA in synaptic complex. J Biol Chem **278:**1904-9.

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

27. **Saito, A., H. Iwasaki, M. Ariyoshi, K. Morikawa, and H. Shinagawa.** 1995. Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. Proc Natl Acad Sci U S A **92:**7470-4.

28. **Shafikhani, S.** 2002. Factors affecting PCR-mediated recombination. Environ Microbiol **4:**482-6.

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

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

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

32. **White, M. F., and D. M. Lilley.** 1997. The resolving enzyme CCE1 of yeast opens the structure of the four-way DNA junction. J Mol Biol **266:**122-34.

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

Fig. 2.1. Inversion region on the *M. lacunata* chromosome. Recombination sites *invL* and *invR* are within the coding sequence of the type 4 pilin genes *tfpQ* and *tfpI* such that inversion switches the 3'-coding sequence of the gene expressed from P_{tfp}. *M. bovis* alternately expresses the serologically different pilins, but *M. lacunata* exhibits a on/off phase variation of TfpQ pili due to a frameshifting 19 bp duplication (black box) in *tfpI* (53). The invertase Piv, encoded immediately adjacent to the invertible segment, is expressed from P_{piv} (34). *sub1* is a nonessential Piv binding site (95).

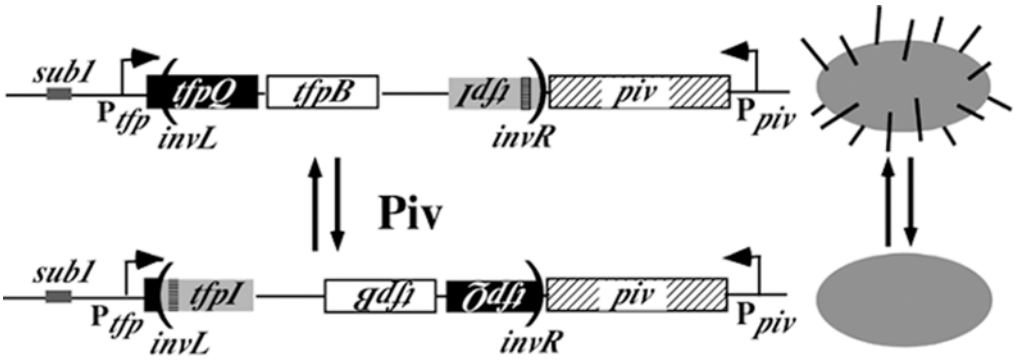

sub1
P tfp
tfpQ
tfpB
invL
invR
piv
P piv
Piv
sub1
P tfp
tfpI
invL
invR
piv
P piv

Fig. 2.2. Substitutions in Piv. The predicted catalytic residues within the amino terminal 160 amino acids of Piv are highlighted by gray boxes and the non-conserved glutamic acid residues that were also targeted for mutagenesis are in open boxes. The substituted residues are indicated in gray type below the wildtype amino acid.

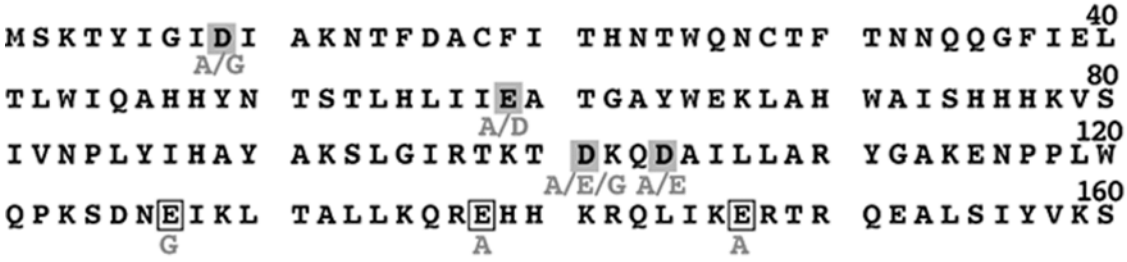

MSKTYIGIDI AKNTFDACFI THNTWQNCTF TNNQQGFIEL
A/G
TLWIQAHHYN TSTLHLIIEA TGAYWEKLAH WAISHHHKVS
A/D
IVNPLYIHAY AKSLGIRTKT DKQDAILLAR YGAKENPPLW
A/E/G A/E
QPKSDNEIKL TALLKQREHH KRQLIKERTR QEALSIYVKS
G
A
A
40
80
120
160

Fig. 2.3. Expression of Piv variants in the strain used for inversion assays. Mid-log cultures of DH5α, or DH5α carrying pAG800.2 encoding the wild type (wt) or mutated *piv* genes (the substitutions are indicated above each lane) were induced with 100 μM IPTG, and two hours post-induction, cells were harvested, lysed, and the proteins fractionated by electrophoresis on 12% SDS-polyacrylamide gels. The Western blot of this gel, utilizing anti-Piv antisera as primary antibody, is shown. Piv is marked by an arrow.

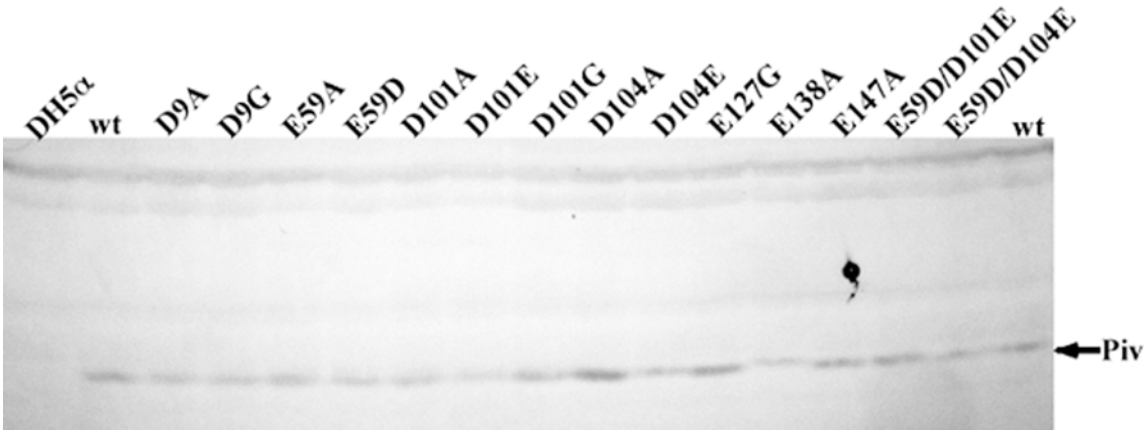

DH5α
wt
D9A
D9G
E59A
E59D
D101A
D101E
D101G
D104A
D104E
E127G
E138A
E147A
E59D/D101E
E59D/D104E
wt
Piv

Fig. 2.4. *In vivo* inversion activity of Piv variants substituted within the DEDD motif and at non-conserved glutamate residues. DH5α, containing the inversion substrate pMxL90, was transformed with pAG800.2-derived expression vectors, encoding wild-type Piv or the variants with the indicated substitutions, and Piv expression induced with 50 μM IPTG. Plasmid DNA was isolated from overnight cultures of individual transformants and inversion of the type 4 pilin segment on pMxL90 was determined by digestion with *BsrG*I. Digest products were electrophoresed on a 0.6% agarose gel and stained with EtBr (inverted image is shown). The starting **Q** orientation of the invertible segment on pMxL90 (**Q**) yields unique 6.5 and 2.6 kb fragments (**pMxL90** alone); inversion to the **I** orientation orientation gives unique 5.2 and 3.9 kb fragments (**pMxL100** alone); a 2.5 kb fragment is common to both. The expression vectors contain only one *BsrG*I site, giving a single 7 kb fragment (**pAG800.2** alone).

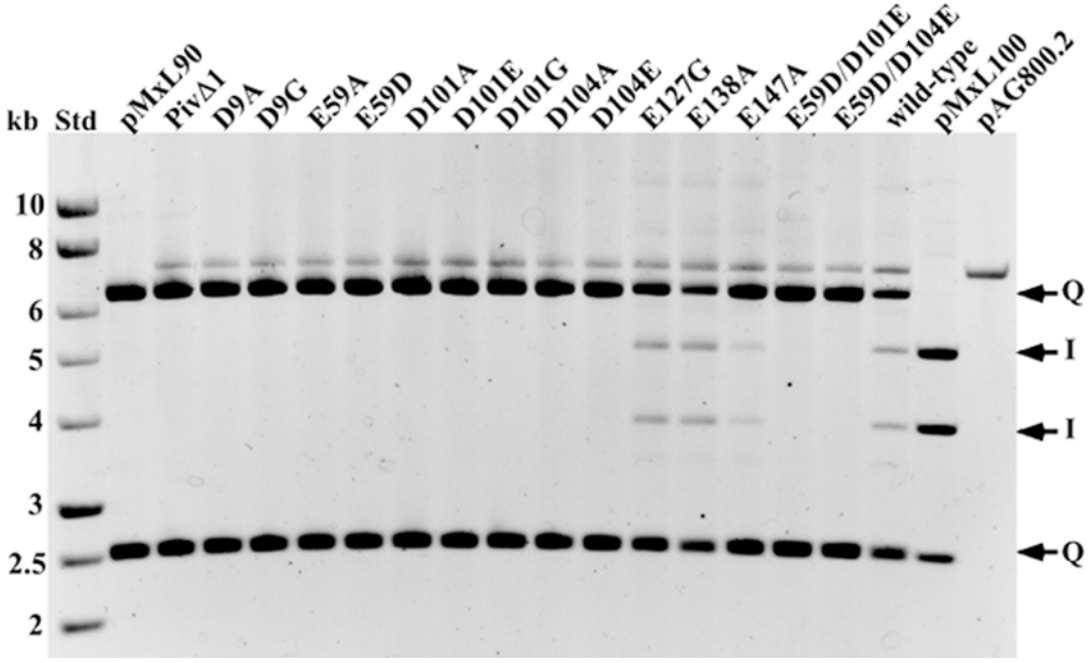

kb
Std
pMxL90
PivΔ1
D9A
D9G
E59A
E59D
D101A
D101E
D101G
D104A
D104E
E127G
E138A
E147A
E59D/D101E
E59D/D104E
wild-type
pMxL100
pAG800.2
10
8
6
5
4
3
2.5
2
Q
I
I
Q

Fig. 2.5. *In vivo* intermolecular recombination mediated by Piv variants substituted at acidic residues. DNA from the inversion assays with Piv variants described in Fig. 2.4 was assayed for intermolecular recombination between *invL* on pMxL90 and *invR* on the pAG800.2-derived vectors. The new DNA junction was detected by PCR using primers, designated **P1** and **P2,** that anneal to sequence unique to pAG800.2 and pMxL90, respectively. PCR products were electrophoresed on a 1.2 % agarose gel and stained with EtBr (inverted image is shown). The 1,073 bp PCR product was sequenced from selected reactions to confirm recombination occurred within the *inv* sequences.

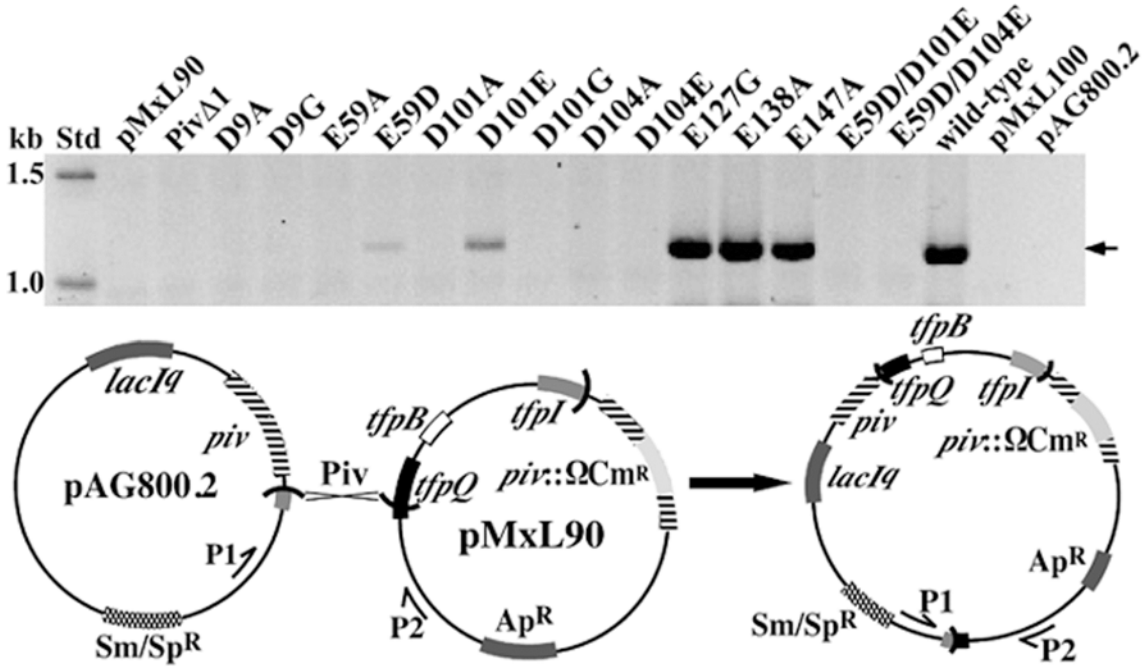

kb Std
pMxL90
PivΔ1
D9A
D9G
E59A
E59D
D101A
D101E
D101G
D104A
D104E
E127G
E138A
E147A
E59D/D101E
E59D/D104E
wild-type
pMxL100
pAG800.2
1.5
1.0
lacIq
piv
pAG800.2
P1
Sm/SpR
Piv
tfpB
tfpI
tfpQ
piv::ΩCmR
pMxL90
P2
ApR
tfpB
tfpQ
tfpI
piv
piv::ΩCmR
lacIq
ApR
P1
Sm/SpR
P2

Fig. 2.6. Detection of low level *in vivo* inversion activity of Piv variants. DNA template utilized in the recombination assays described in Fig. 2.4 and Fig. 2.5 was used in a three-primer PCR inversion assay. A primer that anneals to *tfpB* sequence on pMxL90 pairs with one of two primers that anneal to sequence flanking the invertible segment, to yield a 981 or a 811 bp PCR product when the invertible segment is in the "**Q**" or "**I**" orientation, respectively. The PCR products were electrophoresed on a 1.2 % agarose gel and stained with EtBr (inverted image is shown).

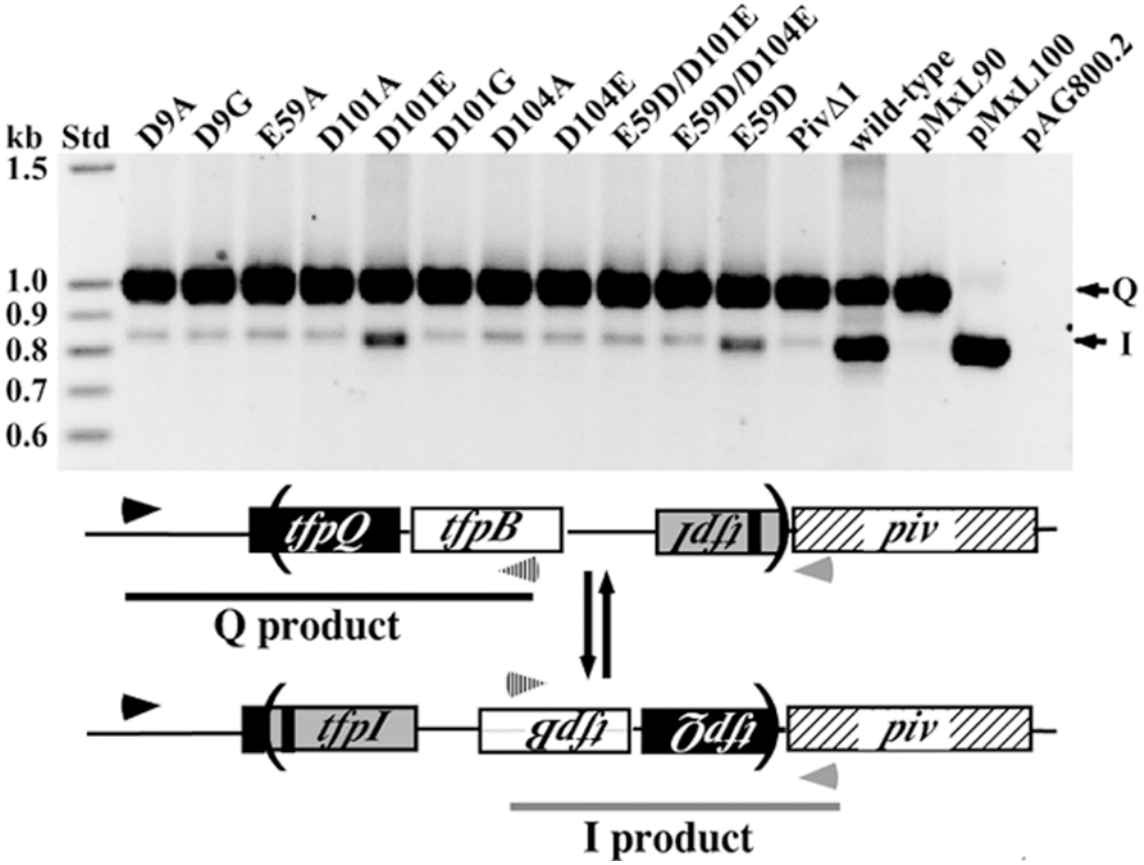

kb
Std
D9A
D9G
E59A
D101A
D101E
D101G
D104A
D104E
E59D/D101E
E59D/D104E
E59D
PivΔ1
wild-type
pMxL90
pMxL100
pAG800.2
1.5
1.0
0.9
0.8
0.7
0.6
Q
I
tfpQ
tfpB
tfpI
piv
Q product
I product

Fig. 2.7. *In vivo* DNA binding activity of Piv variants. The *invL* recombination site, introduced at the +1 position relative to a constitutive promoter, acts as an operator sequence controlling expression of *lacZ* on a single copy plasmid, pAR**110** (24). Piv wild-type (**wt**) and Piv variant (substitutions are indicated) expression vectors were transformed into the DH5α strain containing pAR110, and the transformants were assayed for β-galactosidase activity. The β-galactosidase activity is indicated as Miller units per μg protein.

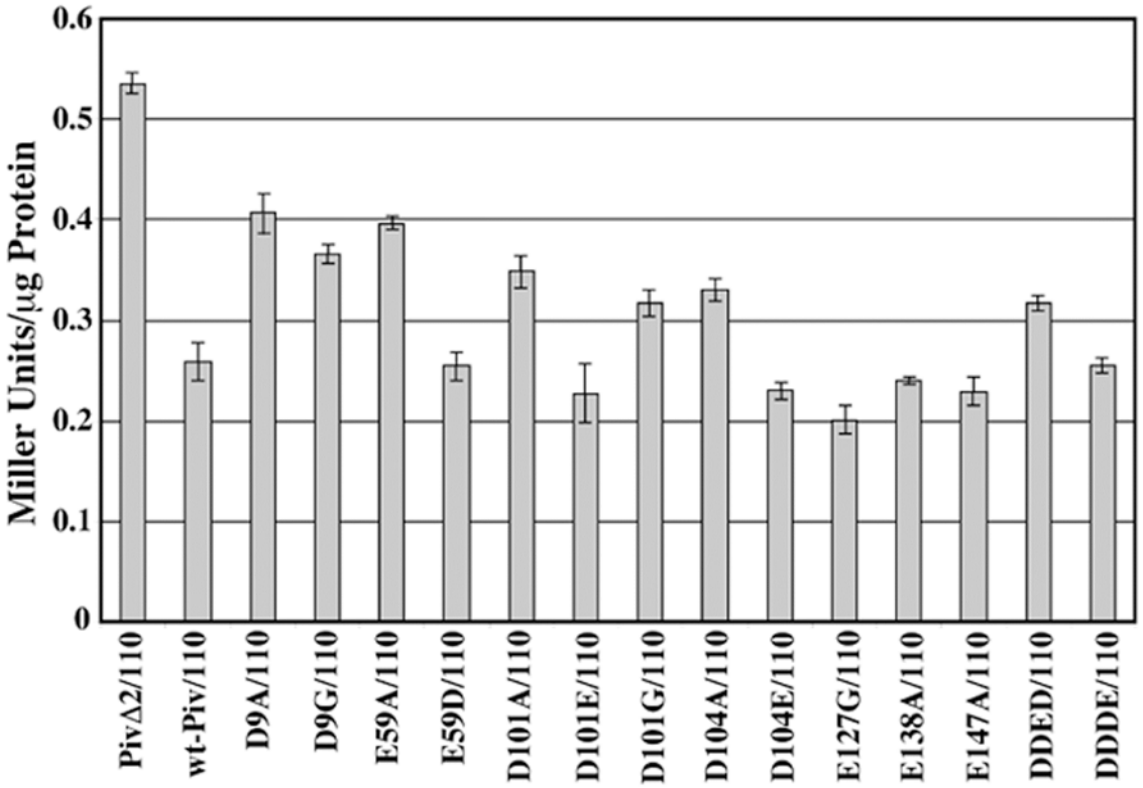

Miller Units/μg Protein
0.6
0.5
0.4
0.3
0.2
0.1
0
PivΔ2/110
wt-Piv/110
D9A/110
D9G/110
E59A/110
E59D/110
D101A/110
D101E/110
D101G/110
D104A/110
D104E/110
E127G/110
E138A/110
E147A/110
DDED/110
DDDE/110

Fig. 2.8. A working model for Piv-mediated site-specific DNA inversion. The synapsed *invL* (black lines) and *invR* (gray lines) are shown bound by a Piv dimer (gray circles). Piv mediated-DNA hydrolysis at one recombination site (depicted as *invL*) (a) and strand transfer (b) leads to formation of a Holliday junction structure. Repositioned Piv catalytic sites now cleave the outer strands of the junction (c) and host DNA ligase activity repairs the nicks (d).

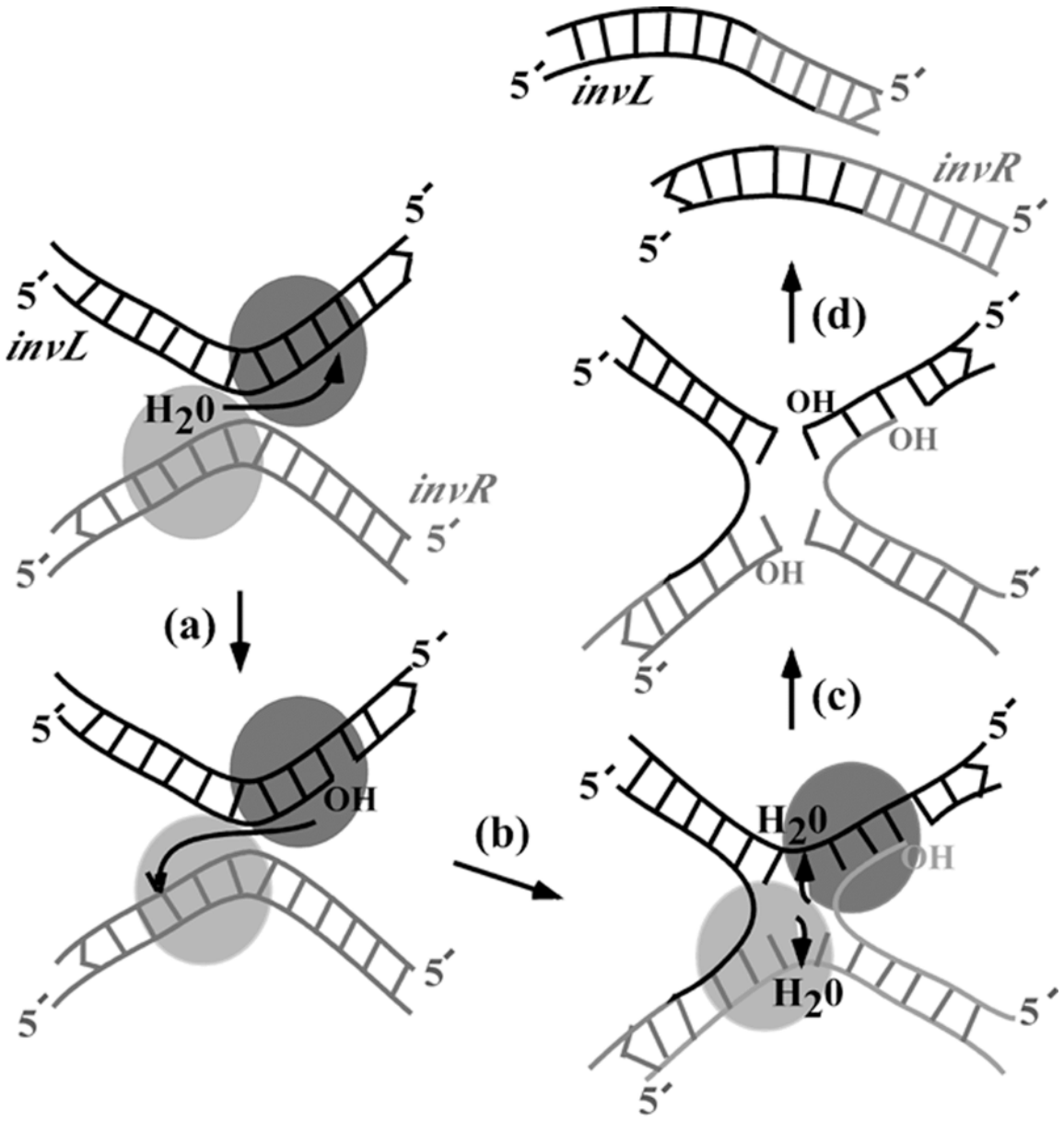

5′
invL
5′
invR
5′
(d)
5′
5′
invL
H20
invR
5′
5′
OH
OH
OH
5′
5′
(a)
(c)
5′
5′
5′
OH
H20
(b)
OH
5′
H20
5′
5′

Table 2.1. *piv* mutations resulting in amino acid substitutions

Substitutions	Mutagenesis Oligonucleotides (5' to 3')*
D9A	CATTGGGAT**C**G**CG**ATCGCCAAAAAC
D9G	CATTGGGATTG**GA**ATCGCCAAAAAC
E59A	CATTTGATTATTG**C**GG*C*AACGGGGG
E59D	TTACATTTGATTATTGA**C**GCAACGGGGGCTTATTGG
D101A	CACCAAGACAG**C**TAAGCAGGATGCC
D101E	GGCATACGCACCAAGACAGA**G**AAGCAGGATGCCATTCTCTTGGCC
D101G	CGCACCAAGAC**C**G**G**TAAGCAGGATGC
D104A	CAGATAAGCAG**GC**CGCCATTCTCTTG
D104E	CCAAGACAGATAAGCAGGA**A**GCCATTCTCTTGGCACGCTAGG
E127G	GGCAGCCTAAATCAGACAAC**GT**ATCAAACTGACTGCCCTTC
E138A	CTAAAACAACGTG**C**ACACCACAAAC
E147A	CGGCAACTCATCAAAG**CG**CGCACTCGACAGG

* Mutations are bolded; altered *piv* codons are underlined.

CHAPTER 3

The DEDD-motif site-specific DNA invertase Piv binds synthetic branched DNA molecules *in vitro*[2]

[2] Buchner, J. M. and Karls, A. C. To be submitted to *Journal of Molecular Biology*.

Abstract

Piv, a DEDD-motif recombinase, mediates site-specific DNA inversion of a chromosomal segment encoding the type IV pilin genes in *Moraxella lacunata* and *Moraxella bovis*. Our working model for Piv-mediated inversion predicts that Piv generates a Holliday junction recombination intermediate and resolves the junction to give the inversion products. Secondary structure predictions indicate a Helix-hairpin-Helix (HhH) DNA binding motif in Piv that could mediate interactions with Holliday junctions or branched DNA. The HhH motif is conserved in the DEDD-motif recombinases. Adjacent to the HhH motif in Piv is a predicted leucine zipper motif that is utilized in protein multimerization. Piv variants that are altered in a conserved leucine (L197) residue of the HhH structural motif and in the lysine (K206) that is essential for interactions of the HhH motif of RuvA with Holliday junctions are defective for *in vivo* inversion activity, while the variants with substitutions in two leucine residues in the predicted leucine zipper exhibit wild-type levels of inversion. *In vitro* electrophoretic mobility shift assays indicate that Maltose Binding Protein-Piv fusion protein binds branched DNA substrates, including Holliday junctions, but the MBP-PivK206E variant does not bind Holliday junctions. These results support our model for Piv-mediated inversion involving a Holliday junction intermediate.

Introduction

Piv, the site-specific DNA invertase isolated from *Moraxella lacunata* and *Moraxella bovis* (23, 24) has something of an identity crisis. Rather than grouping with the tyrosine and serine recombinases usually associated with DNA inversions, Piv has homology to recombinases from the IS*110*/IS*492* family of insertion sequences (21). This grouping is based on primary amino acid sequence alignments of family members and Piv and includes three conserved amino acid motifs in this family, the GDK, KTD-DA, and PSG primary amino acid sequences (21). Secondary and tertiary structure modeling of the amino terminus of Piv revealed a potential catalytic motif in Piv not previously identified in a site-specific DNA invertase, the ribonuclease H-like fold which included the GDK and KTD-DA residues (43).

This catalytic domain is associated with a number of different enzymes including its namesake Ribonuclease H, but also RuvC homologs, retroviral integrases, and DDE-motif transposases (22). One of the key features of this fold is the positioning of a triad or tetrad of acidic residues to chelate divalent metal cations, which in turn activate water molecules that are used in the breaking of the DNA phosphate backbone. The triad of acidic residues, aspartate-aspartate-glutamate, of the DDE-motif transposases and the retroviral integrases are the residues that catalyze the DNA nicking at the 3'-ends of the elements. This generates a hydroxyl group that is used as the nucleophile in the next step of transposition or genome integration (8, 9, 15). This mechanism is hydrolysis followed by transesterification. In the RNaseH-like fold of the RuvC Holliday junction resolvase, four acidic residues, arranged in the primary sequence as D7, E66, D138 and D141, are used to bind divalent metal cations (1). When RuvC resolves a Holliday junction, the nicks left in the DNA become the substrate of ligase or other replication and repair proteins (26, 40, 47).

Though Piv is predicted to have an RNaseH-like fold, it is not observed to have a DDE motif as the catalytic triad, and a DDE pattern is not conserved among recombinases of the IS*110*/IS*492* family (6, 21, 43). In Piv, four acidic residues, D9, E59, D101, and D104, are predicted to be part of the RNaseH-like fold, similar to RuvC (6). These residues are required for catalysis of inversion and are part of two conserved primary amino acid motifs, GDK (D9) and KTD-DA (D101, D104,) and a fourth acidic residue (E59) is conserved as an acidic amino acid among recombinases of the IS*110*/IS*492* family (6). Piv and the latter recombinases family constitute the DEDD motif family of recombinases.

Another structural motif predicted in Piv on the basis of primary sequence and secondary structure predictions is a Helix-hairpin-Helix (HhH) motif. This motif is found in a number of enzymes that interact with nucleic acids in a non-sequence-specific manner, such as DNA polymerases, DNA glycosylases, T4 RNaseH, and a protein associated with Holliday junctions, RuvA (39), though RuvA has no observed catalytic activity itself, the function of RuvA is to localize to Holliday junctions, activate RuvBC, and participate in the translocation and resolution of those junctions (11, 27-30, 40, 45-49). RuvA has two HhH motifs that are centers of DNA interaction, and lysine residues K84 and K119 located in each HhH just after the hairpin are required for DNA interactions (33) by making direct contacts with the phosphate backbone of DNA (16). It is therefore possible that Piv is unique among enzymes in that it is a site-specific invertase that catalyzes recombination through hydrolysis and transesterification, like the DDE-motif transposases, and generates a Holliday junction intermediate that is resolved by Piv with a pair of hydrolysis reactions, like RuvC (2, 6, 7, 11, 38, 41).

The Piv inversion substrate is a 2.1 kbp segment of the *Moraxella lacunata* and *Moraxella bovis* chromosome encoding the type 4 pilin genes *tfpQ*/*I* and is flanked by inverted

repeats, *invL*/*R*. The *tfpQ*/*I* genes are identical in sequence for 273 bp, except in *M. lacunata* where a 19 base pair duplication of *tfpI* sequence 32 bp from the end of the invertible segment disrupts this homology (6, 23, 24, 43, 44). Our current model of inversion predicts for the generation of either of two different Holliday junction molecules, a mobile junction or an immobile junction (Figure 3.1.) The mobile junction model of inversion has the *tfpI* and *tfpQ* genes on opposing arms of the Holliday junction. If this is the inversion intermediate, then the DNA site used for initiation of inversion may not be the same as the site of Holliday junction resolution, as the homology between the genes would allow for translocation. This model predicts that resolution of the Holliday junction can occur in either plane of the Holliday junction due to symmetry of the junction; however, resolution in one plane leads to inversion while resolution in the other plane leads to restoration of the original gene arrangement (Figure 3.1.A.f.)

The second model of inversion involves the generation of an immobile Holliday junction, where mobility is eliminated because the *tfpI* and *tfpQ* sequences are on adjacent arms (Figure 3.1.B.f.) If an immobile junction is the intermediate then the cleavage site in the linear DNA would also direct the cleavage across the Holliday junction, as translocation would otherwise be impossible. The immobile junction also only has one plane of sequence symmetry, and resolution across that plane results in inversion. Cleavage across the other plane of the immobile Holliday junction would result in deletions, which are not observed in the Piv system. The immobile model has a similarity to classical transposase systems because the hydrolysis reaction mediated by the recombinases creates a 3'-hydroxyl end that is used as a nucleophile in attack on a DNA phosphate backbone in strand transfer (8, 9, 15). However, unlike the DDE motif transposases, the strand transfer reaction used to generate the Holliday junction intermediate in

this inversion reaction would not create target site duplications (8, 9, 15), as the attack of the 3'-hydroxyl is directed to the same residues at the junction of the "Y" intermediate (Figure 3.1.B.e.) A final interesting feature of the immobile junction is that host replication machinery could also resolve the intermediate, using the 3'-hydroxyl to start a round of DNA synthesis.

In order to study the interaction of Piv with potential Holliday junction intermediates, we designed synthetic Holliday junctions consisting of oligonucleotides (35, 37) generating both mobile and immobile Holliday junctions. Such junctions have been useful in the characterization of bacteriophage Holliday junction resolvases and RuvABC (1, 2, 4, 5, 10, 29, 30, 32, 38, 45). These junctions are used in binding and competition assays. In this study we test the equivalent of the RuvA lysine residues in the single HhH motif of Piv.

Previous work has demonstrated a potential for Piv multimerization that may be required for interactions with *inv* (44). A third structure predicted in Piv on the basis of primary sequence and secondary structure is a leucine zipper. This structure is identified as a multimerization domain in proteins (17, 19) and it has been identified in the IS*911* transposase, between residues 63 and 95 as a point of dimerization and required for stimulation of transposition (13, 14). The leucine zipper is at least four repeating leucines (or other hydrophobic residues) that are spaced at seven amino acid intervals in primary sequences. In this study we determined the requirements for Piv-mediated inversion in two of the residues predicted to be part of a leucine zipper in Piv.

Materials and Methods

Alignment of DEDD family of recombinasases. Alignments of DEDD motif recombinases potential HhH motifs were based on those previously published (21, 42, 50), with

some additional adjustments made by eye. The consensus leucine zipper was found using the definition published by Landschulz *et al*. (19) and consensus HhH motif by Rafferty *et al.* (33). Predictions of helices and loops was done with Predict Protein (http://www.predictprotein.org/).

Site Directed Mutagenesis. The L176A, L197G and K206E Piv variants were created using the Stratagene Quikchange site-directed mutagenesis kit as described previously (6, 43) with some variations. The oligonucleotides (Synthesized by IDT) used in the mutagenesis PCR were L176AF/R, L196F/R, and K206E F/R (Table 1). The mutagenesis was performed on a plasmid containing *piv* from pAG800.2 (6) cloned into Litmus28 (New England Biolabs, NEB). After isolating clones with no other mutations than the ones intended, as determined by restriction digest and sequencing (University of Michigan DNA Sequencing Core, UMDSC), the mutant *piv* genes were cloned back into pAG800.2 using a BsrGI-HindIII (NEB) digest.

The Piv variant L183D was created by amplifying a subsection of *piv* using the L183DR and the JBPIV1 oligonucleotides (Table 1) and Pfu, and then introducing the product into the TOPO Blunt vector (Invitrogen). The clones were checked by restriction digest and sequencing (UMDSC). After isolating a clone with no other mutations than the ones intended it was put back into pAG800.2 using a BsrGI-MfeI (NEB) digest. Clones were checked for expression by Western blot and assayed for inversion as described previously (6).

Generation of MBP-PivK206E. Digesting pAG607 (44) and pAG800.2K206E with BsrGI-HindIII was the first step in the generation the MBP-PivK206E variant. The relevant restriction fragments were purified from an agarose gel (Zymoclean Gel DNA Recovery Kit, Zymo Research), ligated with T4 ligase (NEB) and used in transformation of *Escherichia coli* DH5α. Clones were checked by restriction digest, then by sequencing (UMDSC). After isolating a clone with no other mutations than the ones intended it was checked for expression by starting a

10 ml Luria Broth (Fisher Scientific) culture with 60 μg/ml Ampicillin (Sigma). At OD_{600}=0.5, the culture is induced with 500 μM IPTG, and shaken for two hours. A cell lysate fraction is analyzed by SDS-PAGE. The gel was stained with Coomassie stain (25), and inspected for the 80 kDa band indicative of full-length fusion protein. MBP-Piv and MBP-PivK206E were purified as described previously (44).

Synthetic Holliday Junction Production. Four synthetic Holliday junctions, (mobile Piv, non-mobile Piv, mobile non-specific, non-mobile non-specific) were built from PAGE purified synthetic oligonucleotides designed as either mobile or immobile substrates (18, 31, 34, 36). Oligonucleotide "I" was labeled with γ-P^{32}-ATP and T4 polynucleotide kinase (NEB) according to NEB protocol.

All substrates were annealed by mixing equal molar amounts of oligonucleotides (100 nM each oligonucleotide, final) in 1X annealing buffer (20 mM Tris pH 7.6, 80 mM KCl, 5 mM $CaCl_2$.) Each substrate was heated to 95^{o}C in a heating block. The block was turned off and allowed to cool to 40^{o}C, annealing the oligonucleotides as Holliday junctions. The tubes were then transferred to room temperature and used in cleavage and binding assays. The substrates were annealed anew before each assay. 20 fmol of each substrate is the standard amount used in each binding and cleavage assay.

Binding, Competition and Cleavage Assays. Binding assays were done in a volume of 50 μl at room temperature for 10 minutes in binding assay buffer (80 mM KCl, 20 mM Tris pH 7.6, 5 mM $CaCl_2$, 250 pg/μl PolyC, 1 mM DTT, 50 ng/μl BSA) with 20 fmol DNA substrate (fixed) and varying amounts of fusion proteins, and varying amounts of protein from 15.7 pmol to 28 fmol of protein were added to each assay. After mixing and ten minutes incubation, loading buffer was added and the samples were treated as above. After incubation, 10 μl loading

buffer (5X loading buffer: Binding Assay Buffer plus 15% ficoll plus 0.05% bromphenol blue) was added to each tube. Samples were then loaded onto 5% native acrylamide (19:1 acrylamide:bisacrylamide) 0.5X TBE native gels. Gels were run at 200 V for 3.5 hours.

Competition assays were done in a volume of 50µl at room temperature for 10 minutes in binding assay buffer with 20 fmol labeled DNA substrate (fixed) and varying amounts of specific or non-specific DNA competitors, and a fixed amount of fusion protein, either 1800 fmol fusion for the immobile junctions or 1600 fmol for the mobile junctions. After incubation, loading buffer was added (as above) and samples were subjected to electrophoresis as above.

Cleavage assays were done in the binding buffer, without $CaCl_2$ and with the addition of $MnCl_2$ or $MgCl_2$ ranging 5 mM-25 mM. Incubation times were extended to a maximum of sixty minutes, and the reaction temperatures were room temperature, 30°C or 37°C. Reaction samples were split and half of the samples was run on native gels, as above, and the other half on DNA sequencing gels after mixing the samples 1:1 with formamide buffer (1X TBE, 90% formamide and 0.5% bromphenol blue), heated to denature and secondary structure, and loaded onto the warmed up gel. Gels were handled as below.

Image capture of radioactive gels. Gels for binding, competition and cleavage assays we dried on Whatman paper using a BIO-RAD gel drier pre-warmed to 80°C under vacuum, for two hours to overnight. The dried gels were put into phosphoimaging screen cassettes and exposed for one hour to overnight. Images were captured using a TYPHOON TRIO (General Electric), and analyzed and adjusted using BIO-RAD Quantity1 software.

For the competition assays, percent bound or percent unbound of each substrate for each lane was determined, and used in the calculations of relative percent bound. Relative in percent bound was calculated by determining the percent unbound in a lane with approximately 50%

bound (a lane without any competitor DNA), and then the percent unbound of lanes with competitor DNA is determined. Dividing the lane with competitor by the lane without competitor and multiplying by 100 yields relative % bound.

Results

Alignments of DEDD family indicate leucine zipper motifs and HhH motifs. A leucine zipper and a HhH motif are predicted in Piv on the basis of primary sequence and secondary structure predictions (Figure 3.2.) If these predictions are correct, then replacements of key residues should eliminate function. The leucine zipper is a repeating pattern of leucine (or another hydrophobic residue) every seventh position (19). For the leucine zipper, we replaced L176 with alanine, L183 with aspartate, and L197 with glycine (Figure 3.2.) The consensus HhH is bxuxxupGuGpxxAxxuuxx, where b is D, E, R, K, N, Q, S, or T at least 65% of the time, u is M, L, I, V, F, T, A, or Y at least 65% of the time, and p is proline at least 35% of the time, x is any amino acid, G is glycine, and A is alanine (33). The PivL197G variant will test the HhH prediction, as this mutation targets one of the HhH consensus residues, the most highly conserved residue among the 15 sequences presented here. This position is always a hydrophobic residue: a leucine, isoleucine, or valine. Piv K206E is tested in inversion assays and in synthetic junction binding assays because even though the lysine residue is not highly conserved in the HhH motif, the equivalent residues in RuvA are required for binding to Holliday junctions (16, 33).

Activity of Piv variants in DNA inversion assays suggest a key role for the HhH, but not for the leucine zipper. Previous studies with Piv tested highly conserved motifs, and demonstrated the predictive powers of modeling and alignments (6, 43). Here we test residues

that are not completely conserved among all homologs of the DEDD family of recombinases as well as residues predicted to be not directly involved in DNA catalysis but possibly involved in DNA binding (HhH) or protein multimerization (leucine zipper.) Western blotting analysis demonstrated that Piv variants with the following replacements; L176A, L183D, L197G, and K206E, are expressed from pAG800.2-derived vectors (Figure 3.3). The variants were tested for *in vivo* inversion activity as shown in Figure 3.4. The variants with replacements to the leucine zipper motif, PivL176A and PivL183D, were positive for inversion activity, while variants testing the HhH motif, PivL197G and PivK206E, were negative for inversion activity. Each variant was tested for Q to I inversion (Figure 3.4), as well as I to Q inversion (data not shown), but no differences in results were observed. The *in vivo* inversion assay suggests that the leucine zipper motif is not relevant to Piv inversion as those variants maintained inversion activity. The HhH motif may, however, play a role in DNA inversion because the two replacements predicted to affect either the structure of the HhH (PivL197G) or the binding of Piv to DNA (PivK206E) lost *in vivo* inversion activity.

MBP-Piv binds unpaired DNA substrates. Earlier work focused on MBP-Piv binding activities with double-stranded linear DNA (duplex) substrates (44). Here, we test the binding of MBP-Piv to single-stranded DNA (monomer) and structures with unpaired (mismatched) DNA bases. For these assays the oligonucleotide (I) was labeled at the 5'-end. The two mismatched substrates were made by annealing equal molar amounts of radio labeled I oligonucleotide, and either the K or T oligonucleotide or no partner. MBP-Piv binds each of these DNA substrates in a protein-dependent manner (Figure 3.5). At the higher concentration of proteins, the complexes failed to enter the gel and are visibly trapped in the wells (data not shown.) The mobility of the unbound IK substrate is slower than that of the I or the IT substrates

and the unbound, I and the IT substrates appear to have a similar mobility (Figure 3.6). The relative mobilities of these three shifted substrates in the presence of MBP-Piv all appear to be similar to each other but different from the shifts associated with MBP-Piv Holliday junction binding (Figure 3.6). Based on these assays, however, it is unclear how many protein and DNA subunits are present in each shifted band. The changes in the mobilities of the bands could represent changes in the amount of DNA or protein, changes in the DNA tertiary structures (bending, hairpin formation), or combinations of all three of these options. Because the shifts change in a protein concentration-dependent manner, it is not unreasonable to presume the slower migrating DNA are complexes of multiple protomers bound to single DNA molecules, not unlike the multimerization seen with Piv binding to *sub1* and *inv* (44).

Generation of Holliday Junction Substrates. After annealing, the mobile and immobile junctions each has its own characteristics. Mobile substrates have regions of homology that allow translocations, while the immobile substrates do not. The regions of non-homology in the mobile junctions prevent translocation through the ends of the oligonucleotides resulting in linear duplex forms. The region of random sequence on the N and O oligonucleotides is to 1) prevent hairpin formation in N which would occur if N reflected the *tfpI/Q* sequence of the immobile Holliday junction, as it would be a palindrome, and 2) prevent O oligonucleotide from annealing to the I oligonucleotide. There is an inherent instability of our Holliday junctions, and the fraction of collapsed Holliday junctions can be observed in the gels. As Figure 3.6 shows, these collapsed products from the mobile substrates do have mobilities similar to the I oligonucleotide or IK and IT. Collapsed immobile junctions have mobilities similar to I oligonucleotide or IN and IP (data not shown).

MBP-Piv binds synthetic Holliday junctions, while MBP-PivK206E does not. Like the mismatch substrate results, MBP-Piv binds each of these DNA substrates in a protein concentration-dependent manner. At the higher levels of MBP-Piv, the complexes failed to enter the gel and were trapped in the wells (data not shown). The collapsed DNA fragments could potentially influence the binding of MBP-Piv to Holliday junctions; however, experiments run in parallel reveal that shifts in the Holliday junction substrate lanes are unique to those lanes (Figure 3.6), although it is apparent the collapsed substrates are also bound by MBP-Piv in these assays. Using the data in Figure 3.7.A and replicate data, we determined that MBP-Piv binds 50% of the 20 fmol of immobile substrate at 1600 fmol of protein and binds 50% of the 20 fmol of mobile substrate at 1800 fmol. MBP-PivK206E, purified to more the 90% homogeneity, did not bind Holliday junctions under the conditions tested (Figure 3.7.B.) These substrates were also investigated for the possibility of cleavage or covalent linkage between MBP-Piv and DNA substrates, but none was observed. However, it was observed that increasing amounts of $MgCl_2$ or $MnCl_2$ did inhibit binding of MBP-Piv to the Holliday junction (Figure 3.8.)

Binding of Holliday junctions is structure-, not sequence-specific. The mobile and immobile Holliday junctions were used in competition assays where the competitors were unlabeled mobile or immobile Piv junctions, or mobile (UVWXII) or immobile (UYZAAII) non-specific sequence junctions (Figure 3.9). The non-specific junctions were constructed to have no homology to potential Piv binding sites, but the structures were designed with oligonucleotides of the same lengths as their counterparts, and designed to have the same number of degrees of freedom in the mobile random junction as in the mobile Piv junction. After analysis, it was found that 865 fmol of specific competitor and 594 fmol non-specific competitor was needed to reduce binding by 50% for the immobile junctions. In the case of the mobile junctions it took

393 fmol specific competitor, and 253 fmol non-specific competitor to reduce binding by 50%. Based on these numbers, there is no significant difference between the amounts of specific or non-specific DNA needed to reduce binding by 50%. These results suggest that MBP-Piv binding to Holliday junctions is not sequence specific, but is structure specific.

Discussion

Conservation of the Leucine zipper and HhH motifs. The leucine zipper motif, which could be used as a dimerization motif, is not strongly conserved among the members of the of the DEDD recombinase family presented here. Furthermore, replacement of residues predicted to be key in the leucine zipper, (L176, L183) maintained inversion activity. Putting these two pieces of information together suggests that a leucine zipper is not present or not required in either Piv or the members of the IS*110*/IS*492* transposase family. However, multiple units of MBP-Piv have been observed binding to the *sub1* accessory site and the *inv* site although it is not clear if there is direct interaction between the protein monomers (44). If multimerization is required for inversion, then it is unlikely that the predicted leucine zipper is required for that interaction. While there is the possibility that another protein binds to Piv through a leucine zipper to control Piv function, this possibility was not tested for in our *E. coli in vivo* system if the regulatory protein is present only in *M. lacunata* or *M. bovis.*

The HhH motif (33) is fairly well conserved throughout the DEDD recombinase family, although no single position is completely conserved throughout the representatives presented here, five positions (Piv residues L197, I200, I203, L211, and L212) match the consensus HhH sequence (33). All but two representatives here (IS*1533* and IS*Ngo2* transposases) match the core (GuG) pattern (33). This, combined with the negative *in vivo* inversion phenotype of

PivL197G and PivK206E and the lack of observed *in vitro* Holliday junction binding of the MBP-PivK206E variant, indicates that an HhH motif may be present in Piv and required for DNA binding. The poor conservation of K206 among members the IS*110*/IS*492* transposases could indicate a difference between the DNA inversion reaction catalyzed by Piv and the DNA insertion and excision reactions catalyzed by the other enzymes with a functioning HhH. Conversely, lysine 84 and 119 are required for RuvA binding to Holliday junctions, but is not conserved among all enzymes with HhH motifs (33). Therefore, it is possible that the DEDD recombinases all have functioning HhH motifs, but use other residues or other positions in the motif to interact with the DNA, or a Holliday junction is part of their transposition mechanism, but the assembly of the junction is different than in the Piv system.

MBP-Piv binds unpaired DNA. *In vitro* binding assays revealed an affinity of MBP-Piv for unpaired DNA. While it is unclear exactly how unpaired DNA fits with our current Holliday junction hypothesis of inversion, the HhH motif is not used exclusively for Holliday junction binding. Other enzymes such as MutY and DNA polymerase use the HhH structure to interact with branched or mismatched DNA (33, 39). Therefore; the predicted HhH motif encoded by Piv could be directing interactions with unpaired DNA. Alternatively, other Piv residues could be the source of these interactions, or the binding could be an artifact of our system. *In vivo* there may be tight regulation of Piv-DNA interactions through protein-protein interactions and arrangements of Piv bound accessory sites such that Piv may not encounter single-stranded DNA.

MBP-Piv Binds Holliday Junctions in a structure-dependent manner. In our models of Piv-mediated inversion, Piv catalyzes the initial nick in the DNA backbone and strand transfer *via* the DEDD catalytic center; resolution of the Holliday junction then could occur by

endonucleolytic activity of the Piv catalytic domain (6). Alternatively, the Holliday junction could be resolved by the cellular RuvC across the preferred RuvC cleavage site (2) encoded within *invL*/*R*, so as to prevent deletion of the invertible region. In this paper, we present evidence that Piv binds mobile and immobile synthetic Holliday junctions. Furthermore, the addition of higher amounts of $MgCl_2$ and $MnCl_2$ inhibited binding of Piv to Holliday junctions, analogous to the results for RuvC seen by Bennett *et al.* (2). This result could be due to the cations shielding the negative charges on the DNA, preventing protein-DNA contacts (2).

The junction binding by MBP-Piv is not sequence-specific, as junctions with random sequence compete as effectively as junctions comprised of *invL*/*R* sequence. RuvC can also bind Holliday junctions in a non-sequence-specific manner, but RuvC has specificity in a resolution site (2-5, 12, 38). It should be noted that non-specific binding of MBP-Piv to DNA could be an artifact of our system. The MBP tag on our fusion protein could be blocking Piv regions evolved to give Piv sequence specificity. It is also possible that *sub1* or other undefined accessory sequences *in vivo* allow for the proper Holliday junction (mobile or immobile) to be formed. In that case, Piv binding to synthetic Holliday junctions that are a fraction of the DNA substrate present in *in vivo* recombination reflects only the non-specific interactions of Piv with the Holliday junctions. Finally, the binding could be an artifact of the substrates. The immobile substrates are asymmetric with two short arms and two long arms. The immobile Holliday junction substrate does not perfectly reflect the proposed immobile Holliday junction. One of the long arms of the Piv immobile Holliday junction substrate sequence is random sequence because if we were to build a true immobile substrate based on *invL*/*R* sequence, the long arm oligonucleotide represented by “N” would be a palindrome, and would anneal to itself either as a hairpin or a duplex. If an immobile Holliday junction is the intermediate, it is possible that a

protein-DNA interaction could take place far from the junction site, and form a "loop" like those described in the Lane *et al.* review (20) that could serve to activate or stimulate the cleavage of the Holliday junction. However, unless the MBP-Piv looping is also not sequence specific, we could not observe such a reaction in this assay. If the symmetrical immobile junction is the true inversion substrate, it might be preferred in binding assays over all the junctions presented here.

Here we have presented evidence for a Holliday junction-binding HhH motif in Piv and DNA binding results that demonstrate an affinity of MBP-Piv for Holliday junctions. This work has opened up new questions, however. Which Holliday junction is the true inversion intermediate? How do these results relate to the transposases of the IS*110*/IS*492* family? To answer the first question, *in vitro* cleavage assays need development. Presumably, the true substrate will be cleaved, and the substrate irrelevant to inversion will not. Attempts to isolate cleavage products have not been successful thus far, but it is possible that with the right combination of substrates and protein fractions cleavage will be observed.

The second question is potentially answered with the results presented in this paper. If inversion occurs by the Holliday junction mechanism (6), it is unclear how the transposases of the IS*110*/IS*492* family and Piv are mechanistically related. Most of the IS elements in the family have no inverted repeats, and only some have direct repeats. Taking IS*492* as an example, if the left and right ends of the element are brought together, in a Holliday junction reminiscent of the one proposed for Piv (Figure 3.1.A.f) no translocation could occur, because there's no homology, and therefore the initial nick would not be repaired. Furthermore, cleavage across this junction never results in excision, only inversion, and that is not observed in IS*492* recombination. If the ends were brought together like that presented for the immobile Holliday junction model, in the introduction of this paper, a different result would be possible. If the

nicking occurs exactly as presented in the introduction (Figure 3.1.B), then a circular intermediate could be produced without excision of any of the element, through replication of the closed circle of DNA, initiating from the 3'OH of the cleaved DNA (Figure 3.1.B.f.) One drawback to this model is that the element acting as the source of circular intermediates would become inverted, which has not been observed from IS*492*. However, if the series of nicks occurred as presented in Figure 3.10, the structure at the end would have a nicked Holliday junction with a nick at the top instead of at the left (Figure 3.10), as shown. With a nick in the bottom strand, a second nick in the opposite (top) strand would release a nicked circle IS*492* (circular intermediates are an intermediate in IS*492* transposition) and a nicked chromosome at the restored locus of excision. The link between the DEDD motif recombinases and their insertion, excision, and inversion of DNA may be through Holliday junction intermediates created by the RNase H-like fold and HhH motifs that bind those Holliday junctions.

References

1. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, H. Nakamura, H. Shinagawa, and K. Morikawa.** 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. Cell **78:**1063-72.

2. **Bennett, R. J., H. J. Dunderdale, and S. C. West.** 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell **74:**1021-31.

3. **Bennett, R. J., and S. C. West.** 1996. Resolution of Holliday junctions in genetic recombination: RuvC protein nicks DNA at the point of strand exchange. Proc Natl Acad Sci U S A **93:**12217-22.

4. **Bennett, R. J., and S. C. West.** 1995. RuvC protein resolves Holliday junctions via cleavage of the continuous (noncrossover) strands. Proc Natl Acad Sci U S A **92:**5635-9.

5. **Benson, F. E., and S. C. West.** 1994. Substrate specificity of the *Escherichia coli* RuvC protein. Resolution of three- and four-stranded recombination intermediates. J Biol Chem **269:**5195-201.

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

7. **Connolly, B., C. A. Parsons, F. E. Benson, H. J. Dunderdale, G. J. Sharples, R. G. Lloyd, and S. C. West.** 1991. Resolution of Holliday junctions in vitro requires the *Escherichia coli* ruvC gene product. Proc Natl Acad Sci U S A **88:**6063-7.

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

9. **Derbyshire, K. M., and N. D. F. Grindley.** 2005. DNA Transposons: DIfferent Proteins and Mechanisms but Similar Rearrangements, p. 467-498. *In* N. P. Higgins (ed.), The Bacterial Chromosome. ASM Press, Washington, D.C.

10. **Duckett, D. R., M. J. Panis, and D. M. Lilley.** 1995. Binding of the junction-resolving enzyme bacteriophage T7 endonuclease I to DNA: separation of binding and catalysis by mutation. J Mol Biol **246:**95-107.

11. **Dunderdale, H. J., F. E. Benson, C. A. Parsons, G. J. Sharples, R. G. Lloyd, and S. C. West.** 1991. Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. Nature **354:**506-10.

12. **Eggleston, A. K., and S. C. West.** 2000. Cleavage of holliday junctions by the *Escherichia coli* RuvABC complex. J Biol Chem **275:**26467-76.

13. **Haren, L., C. Normand, P. Polard, R. Alazard, and M. Chandler.** 2000. IS*911* transposition is regulated by protein-protein interactions via a leucine zipper motif. J Mol Biol **296:**757-68.

14. **Haren, L., P. Polard, B. Ton-Hoang, and M. Chandler.** 1998. Multiple oligomerisation domains in the IS*911* transposase: a leucine zipper motif is essential for activity. J Mol Biol **283:**29-41.

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

16. **Hargreaves, D., D. W. Rice, S. E. Sedelnikova, P. J. Artymiuk, R. G. Lloyd, and J. B. Rafferty.** 1998. Crystal structure of *E.coli* RuvA with bound DNA Holliday junction at 6 Å resolution. Nat Struct Biol **5:**441-6.

17. **Hu, J. C., E. K. O'Shea, P. S. Kim, and R. T. Sauer.** 1990. Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. Science **250:**1400-3.

18. **Kallenbach, N. R., M. Rong-Ine, and N. C. Seeman.** 1983. An immobile nucleic acid junction constructed from oligonucleotides. Nature **203:**829-831.

19. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759-64.

20. **Lane, D., P. Prentki, and M. Chandler.** 1992. Use of gel retardation to analyze protein-nucleic acid interactions. Microbiol Rev **56:**509-28.

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

22. **Lilley, D. M., and M. F. White.** 2000. Resolving the relationships of resolving enzymes. Proc Natl Acad Sci U S A **97:**9351-3.

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

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

25. **Marshak, D., R., J. T. Kadonaga, R. R. Brugess, and M. W. Knuth.** 1996. Strategies for Protein Purification and Characterization. A Laboratory Course Manual. Cold Spring Harbor Laboratory Press.

26. **Michel, B., H. Boubakri, Z. Baharoglu, M. LeMasson, and R. Lestini.** 2007. Recombination proteins and rescue of arrested replication forks. DNA Repair (Amst) **6:**967-80.

27. **Parsons, C. A., A. Stasiak, R. J. Bennett, and S. C. West.** 1995. Structure of a multisubunit complex that promotes DNA branch migration. Nature **374:**375-8.

28. **Parsons, C. A., A. Stasiak, and S. C. West.** 1995. The *E.coli* RuvAB proteins branch migrate Holliday junctions through heterologous DNA sequences in a reaction facilitated by SSB. Embo J **14:**5736-44.

29. **Parsons, C. A., I. Tsaneva, R. G. Lloyd, and S. C. West.** 1992. Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. Proc Natl Acad Sci U S A **89:**5452-6.

30. **Parsons, C. A., and S. C. West.** 1993. Formation of a RuvAB-Holliday junction complex in vitro. J Mol Biol **232:**397-405.

31. **Picksley, S. M., C. A. Parsons, B. Kemper, and S. C. West.** 1990. Cleavage specificity of bacteriophage T4 endonuclease VII and bacteriophage T7 endonuclease I on synthetic branch migratable Holliday junctions. J Mol Biol **212:**723-35.

32. **Pohler, J. R., M. J. Giraud-Panis, and D. M. Lilley.** 1996. T4 endonuclease VII selects and alters the structure of the four-way DNA junction; binding of a resolution-defective mutant enzyme. J Mol Biol **260:**678-96.

33. **Rafferty, J. B., S. M. Ingleston, D. Hargreaves, P. J. Artymiuk, G. J. Sharples, R. G. Lloyd, and D. W. Rice.** 1998. Structural similarities between *Escherichia coli* RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the holliday junction. J Mol Biol **278:**105-16.

34. **Seeman, N. C.** 1990. *De novo* design of sequences for nucleic acid structural engineering. J Biomol Struct Dyn **8:**573-81.

35. **Seeman, N. C.** 1982. Nucleic acid junctions and lattices. J Theor Biol **99:**237-47.

36. **Seeman, N. C., J. H. Chen, and N. R. Kallenbach.** 1989. Gel electrophoretic analysis of DNA branched junctions. Electrophoresis **10:**345-54.

37. **Seeman, N. C., and N. R. Kallenbach.** 1983. Design of immobile nucleic acid junctions. Biophys J **44:**201-9.

38. **Shah, R., R. J. Bennett, and S. C. West.** 1994. Activation of RuvC Holliday junction resolvase *in vitro*. Nucleic Acids Res **22:**2490-7.

39. **Shao, X., and N. V. Grishin.** 2000. Common fold in helix-hairpin-helix proteins. Nucleic Acids Res **28:**2643-50.

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

41. **Sharples, G. J., and R. G. Lloyd.** 1991. Resolution of Holliday junctions in *Escherichia coli*: identification of the ruvC gene product as a 19-kilodalton protein. J Bacteriol **173:**7711-5.

42. **Skaar, E. P., B. Lecuyer, A. G. Lenich, M. P. Lazio, D. Perkins-Balding, H. S. Seifert, and A. C. Karls.** 2005. Analysis of the Piv recombinase-related gene family of *Neisseria gonorrhoeae*. J Bacteriol **187:**1276-86.

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

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

45. **Tsaneva, I. R., B. Muller, and S. C. West.** 1993. RuvA and RuvB proteins of *Escherichia coli* exhibit DNA helicase activity *in vitro*. Proc Natl Acad Sci U S A **90:**1315-9.

46. **van Gool, A. J., N. M. Hajibagheri, A. Stasiak, and S. C. West.** 1999. Assembly of the *Escherichia coli* RuvABC resolvasome directs the orientation of holliday junction resolution. Genes Dev **13:**1861-70.

47. **West, S. C.** 2003. Molecular views of recombination proteins and their control. Nat Rev Mol Cell Biol **4:**435-45.

48. **West, S. C.** 1996. The RuvABC proteins and Holliday junction processing in *Escherichia coli*. J Bacteriol **178:**1237-41.

49. **Zerbib, D., C. Mezard, H. George, and S. C. West.** 1998. Coordinated actions of RuvABC in Holliday junction processing. J Mol Biol **281:**621-30.

50. **Zuerner, R. L.** 1994. Nucleotide sequence analysis of IS*1533* from *Leptospira borgpetersenii*: identification and expression of two IS-encoded proteins. Plasmid **31:**1-11.

Figure 3.1. Models of Piv mediated DNA inversion through a Holliday junction. Panel A is a proposal for the generation of a mobile Holliday junction and panel B is a proposal for the generation of an immobile Holliday junction. The steps in the reaction are delineated with black arrows. The proposed catalytic activities of Piv involving DNA cleavage (nicking of the phosphate backbone and resolution of the junction) are indicated with black triangles. The use of 3' hydroxyls as nucleophiles is indicated with white arrows. The models diverge most dramatically at the steps before the generation of the Holliday junction intermediate. The mobile Holliday junction model uses translocation to move the nick away from the junction, with repair of the nick by host enzymes (A.e.) The immobile junction steps create a Y intermediate before generation of a Holliday junction (B.e and f.) The mobile junction can be resolved in one of two planes. Resolution across the vertical plane would result in inversion, and resolution in the horizontal plane would restore the original orientation. For the immobile junction, resolution only along the horizontal plane is possible, as resolution across the vertical plane results in deletion. The 3' hydroxyl left at the end of the immobile junction reaction could also be a locus to start replication, and this could also resolve the junction.

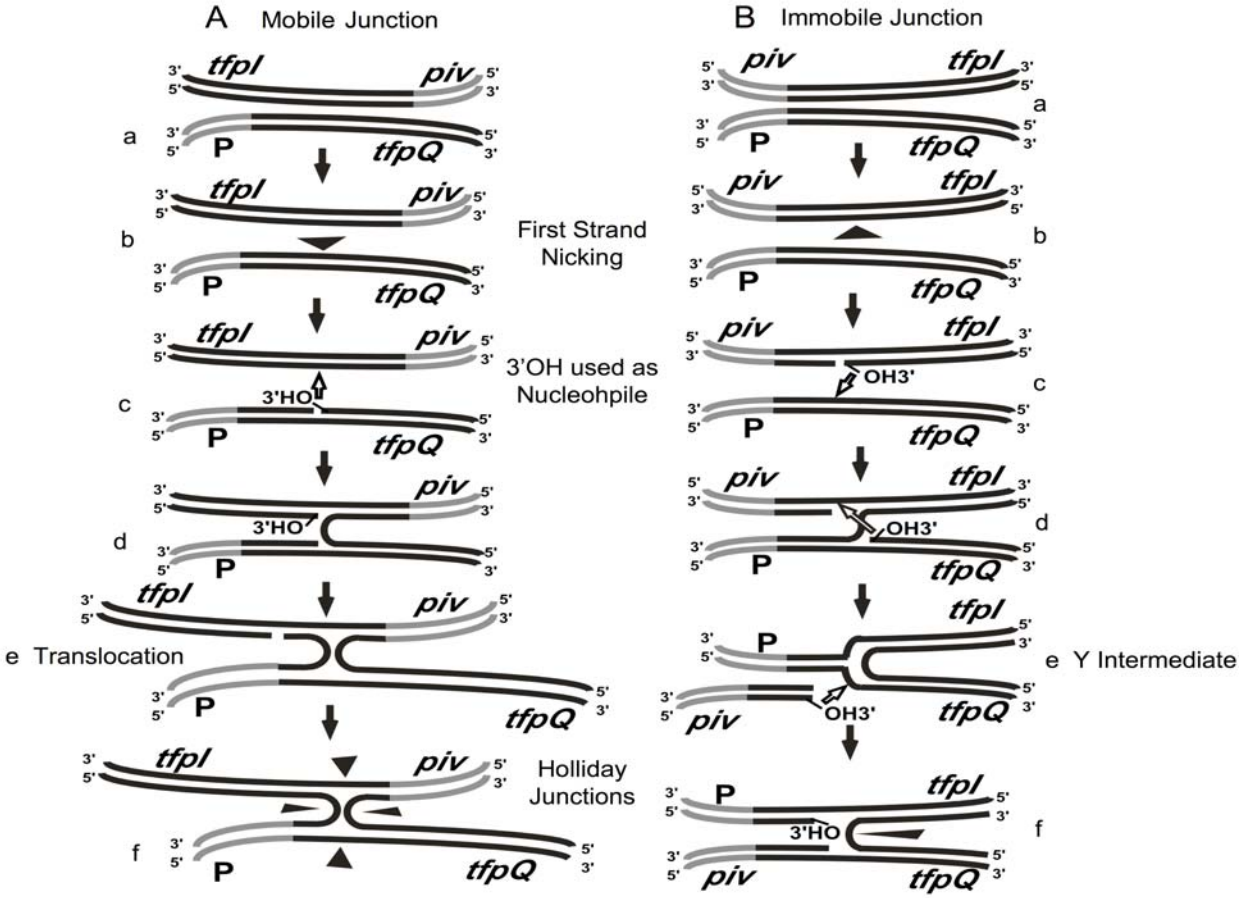

A Mobile Junction
B Immobile Junction
tfpI
piv
tfpQ
P
a
b
c
d
e Translocation
f
First Strand Nicking
3'OH used as Nucleohpile
3'HO
OH3'
Holliday Junctions
e Y Intermediate

Figure 3.2. Alignment of DEDD recombinases indicates the presence of leucine zipper and HhH motifs. The aligned sequences correspond to *M. lacunata* and *M. bovis* Piv residues 176-214 and the corresponding regions of thirteen transposase members of the IS*110*IS492 family indicated by IS element name. Hydrophobic residues fitting the leucine zipper pattern are highlighted in gray. In some cases (in Piv and the recombinases associated with IS*492*, IS*1000*, IS*Ech*, and IS*621*) the heptad repeat continues into the HhH motif. The consensus HhH motif primary sequence is shown below the alignment with the residues matching the HhH consensus sequence in these proteins highlighted in black. The underlined sequence in Piv(ml) is predicted to have a helical structure. Only helices with a reliability score <5 are indicated. The IS*1000* recombinase matched the HhH consensus perfectly whereas IS*1533* had only seven of eleven residues match, including a mismatch in the key "GuG" pattern of the HhH motif.

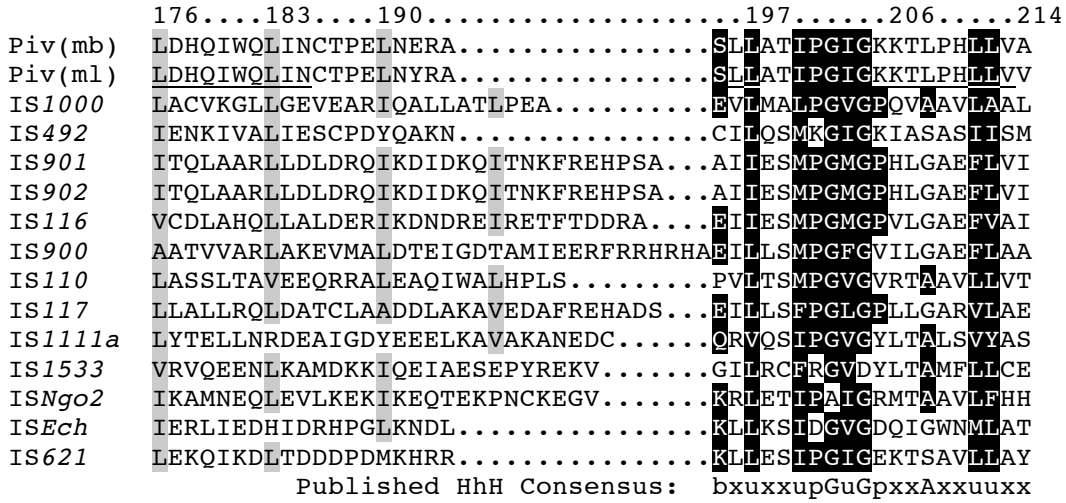

176....183....190.....197.....206.....214
Piv(mb) LDHQIWQLINCTPELNERA.....SLLATIPGIGKKTLPHLLVA
Piv(ml) LDHQIWQLINCTPELNYRA.....SLLATIPGIGKKTLPHLLVV
IS1000 LACVKGLLGEVEARIQALLATLPEA.....EVLMALPGVGPQVAAVLAAL
IS492 IENKIVALIESCPDYQAKN.....CILQSMKGIGKIASASIISM
IS901 ITQLAARLLDLDRQIKDIDKQITNKFREHPSA...AIIESMPGMGPHLGAEFLVI
IS902 ITQLAARLLDLDRQIKDIDKQITNKFREHPSA...AIIESMPGMGPHLGAEFLVI
IS116 VCDLAHQLLALDERIKDNDREIRETFTDDRA....EIIESMPGMGPVLGAEFVAI
IS900 AATVVARLAKEVMALDTEIGDTAMIEERFRRHRHAEILLSMPGFGVILGAEFLAA
IS110 LASSLTAVEEQRRALEAQIWALHPLS.....PVLTSMPGVGVRTAAVLLVT
IS117 LLALLRQLDATCLAADDLAKAVEDAFREHADS...EILLSFPGLGPLLGARVLAE
IS1111a LYTELLNRDEAIGDYEEELKAVAKANEDC.....QRVQSIPGVGYLTALSVYAS
IS1533 VRVQEENLKAMDKKIQEIAESEPYREKV.....GILRCFRGVDYLTAMFLLCE
ISNgo2 IKAMNEQLEVLKEKIKEQTEKPNCKEGV.....KRLETIPAIGRMTAAVLFHH
ISEch IERLIEDHIDRHPGLKNDL.....KLLKSIDGVGDQIGWNMLAT
IS621 LEKQIKDLTDDDPDMKHRR.....KLLESIPGIGEKTSAVLLAY
Published HhH Consensus: bxuxxupGuGpxxAxxuuxx

Figure 3.3. Western blot detection of Piv and Piv variant expression. The variants are named by their position. NC is negative control: Wild-type is Piv expressed from pAG800.2. The black arrows indicate the location of Piv on the blot.

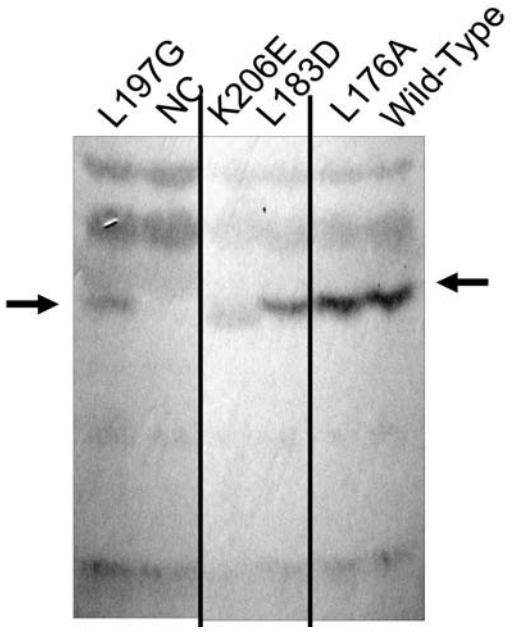

L197G
NC
K206E
L183D
L176A
Wild-Type

Figure 3.4. *In vivo* inversion activity of Piv variants substituted in the predicted leucine zipper and HhH motifs. *E. coli* DH5α containing the inversion substrate pMxL90 (Q orientation substrate) was transformed with pAG800.2-derived expression vectors, encoding wild-type Piv or the variants with the indicated replacements. The inverted image of the ethidium bromide-stained agarose gel is shown. The position of the expression vector is marked with a black double arrow. Gray arrow mark the location of Q-specific bands (6.5 and 2.6 kb fragments), and the open arrows indicate the location of the I-specific bands (5.2 and 3.9 kb fragments) or the location of the bands of the I orientation control plasmid, pMxL100.

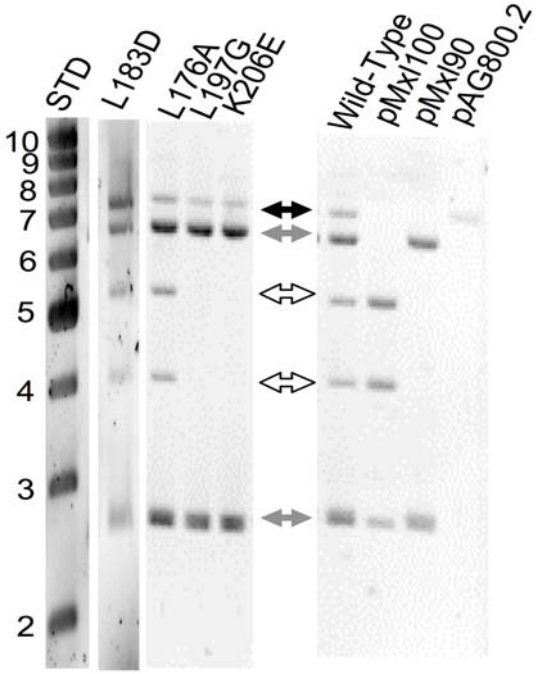

STD
L183D
L176A
L197G
K206E
Wild-Type
pMxl100
pMxl90
pAG800.2
10
9
8
7
6
5
4
3
2

Figure 3.5. Electrophoretic mobility Shift Assays show Piv binds to unpaired DNA. Open double arrows indicate the location of unbound substrates; bound substrates are indicated by black double arrows. The substrates, from left to right, are I, IK, and IT. The gels are aligned by the unbound substrates however, their actual relative mobilities can be observed in Figure 3.6.

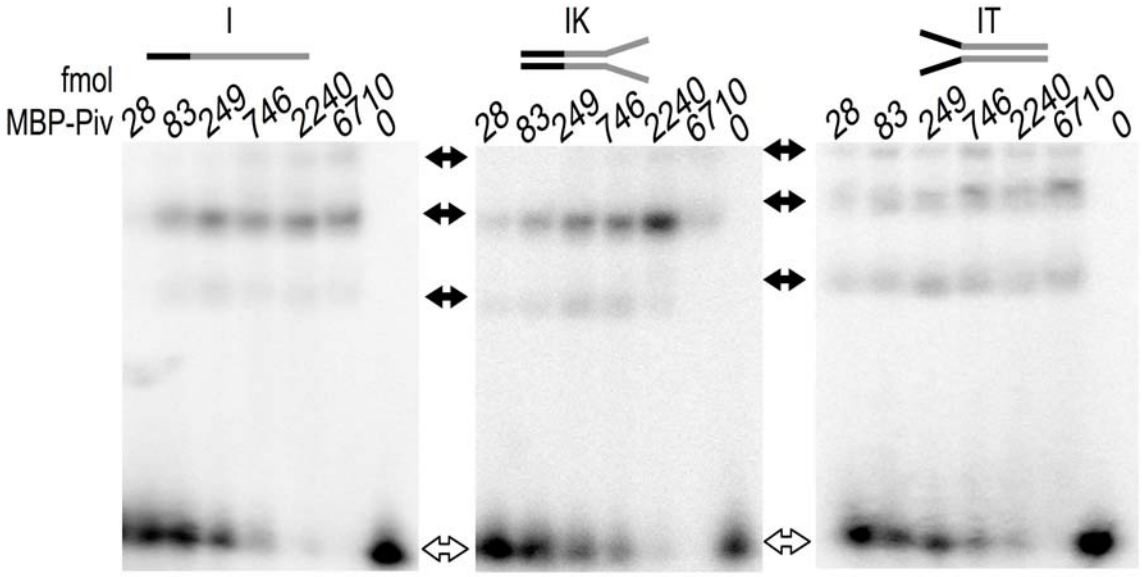

I
IK
IT
fmol
MBP-Piv
28 83 249 746 2240 6710 0
28 83 249 746 2240 6710 0
28 83 249 746 2240 6710 0

Figure 3.6. Characterization of MBP-Piv binding *in vitro* to synthetic DNA substrates. The lanes are marked HJ for Holliday junction, I for the unpaired I oligonucleotide (monomer), IT and IK are mismatch substrates. The open arrows mark the locations of unbound substrates. The gray arrows mark the location of shifts attributed to mismatch substrate. The black arrow indicates the shift band of MBP-Piv-bound Holliday junction.

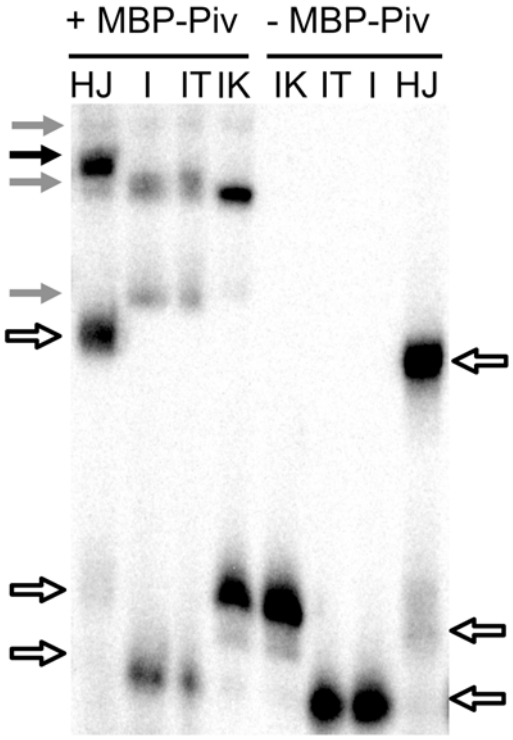

+ MBP-Piv
- MBP-Piv
HJ
I
IT
IK
IK
IT
I
HJ

Figure 3.7. Binding of MBP-Piv or MBP-PivK206E to Holliday Junctions. Panel A is MBP-Piv and panel B is MBP-PivK206E. The gray double arrow indicates the location of unbound Holliday junctions, and the black double arrows indicate the locations of bound substrates. White arrows indicate protein-independent collapsed junctions. The substrates are represented with a symmetrical "X" to indicate mobile Holliday junction, and an asymmetric "X" to denote the immobile Holliday junction.

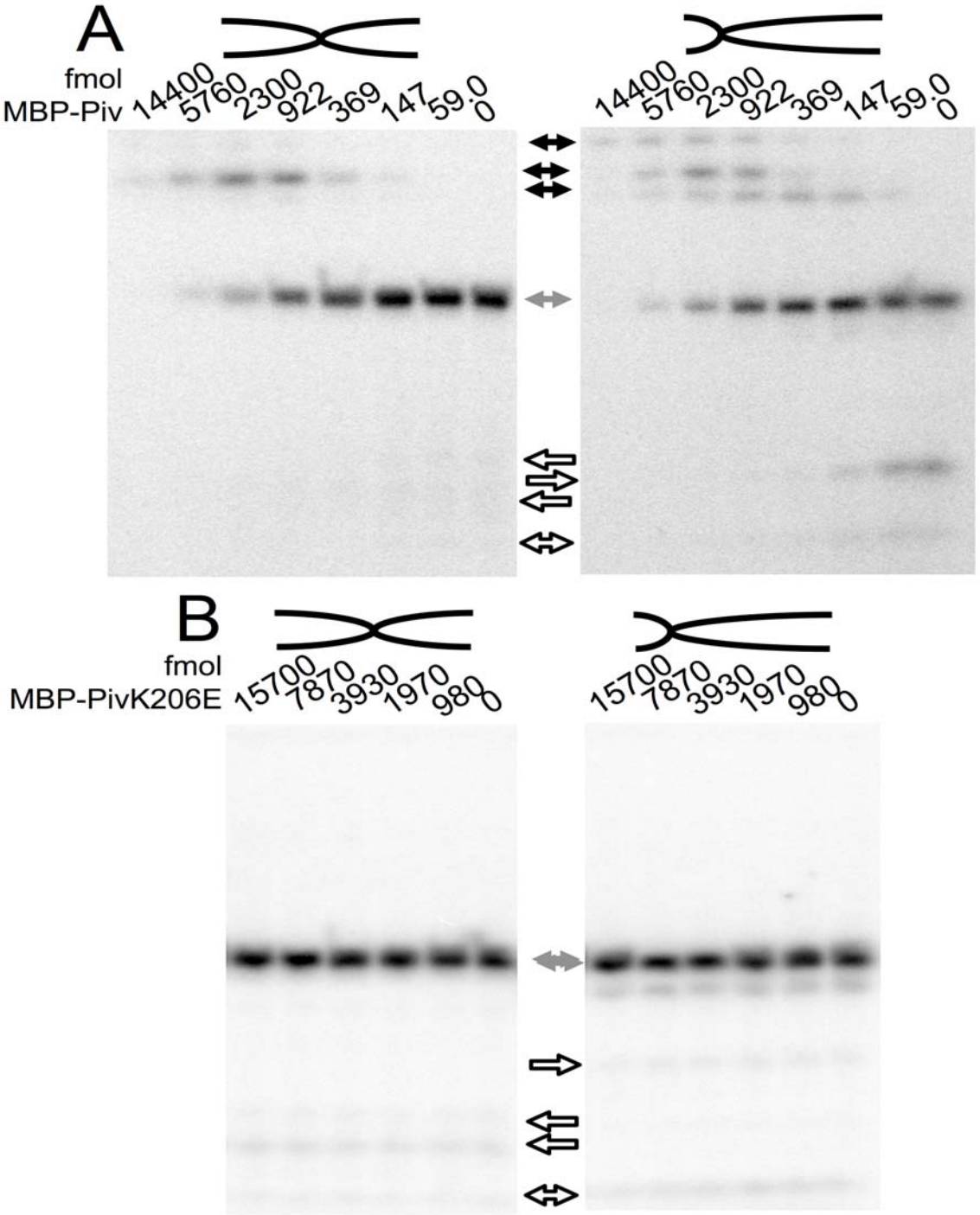

A
fmol
MBP-Piv
14400
5760
2300
922
369
147
59.0
0
14400
5760
2300
922
369
147
59.0
0
B
fmol
MBP-PivK206E
15700
7870
3930
1970
980
0
15700
7870
3930
1970
980
0

Figure 3.8. $MgCl_2$ and $MnCl_2$ interfere with MBP-Piv binding to the mobile Holliday junction. The millimolar concentration of magnesium chloride or manganese chloride is listed across the top of the gel. The black arrow indicates the immobile Holliday junction shift, and the gray arrow indicates the unbound Holliday junction. 250 fmol of MBP-Piv is used in each assay.

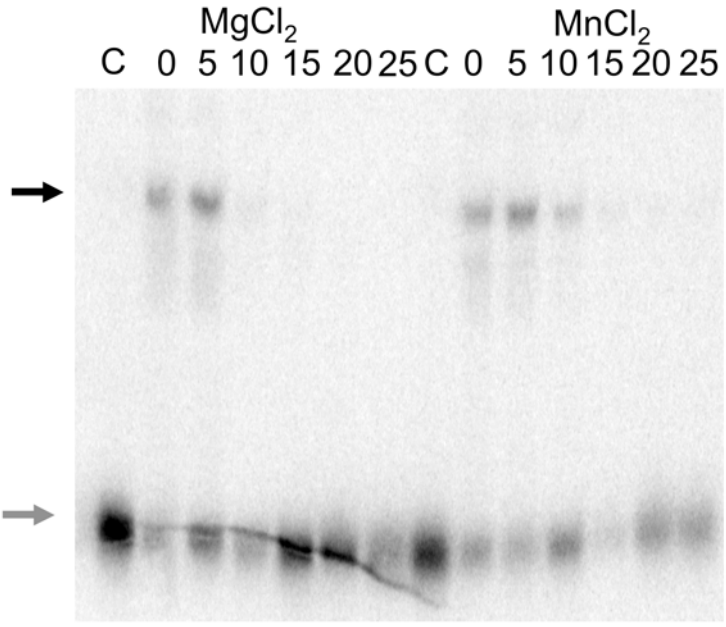

MgCl2
MnCl2
C 0 5 10 15 20 25 C 0 5 10 15 20 25

Figure 3.9. Competition assays with a mobile and immobile substrate. The symmetrical X structures are mobile Holliday junctions; the asymmetrical X structures are immobile substrates. The black structures are specific sequence, and the gray junctions are non-specific sequence. Black arrows indicate the location of shifted substrates; the gray arrows represent unbound substrates. (Collapsed substrate is not shown). "C" lanes are control lanes where no protein or competitor was added.

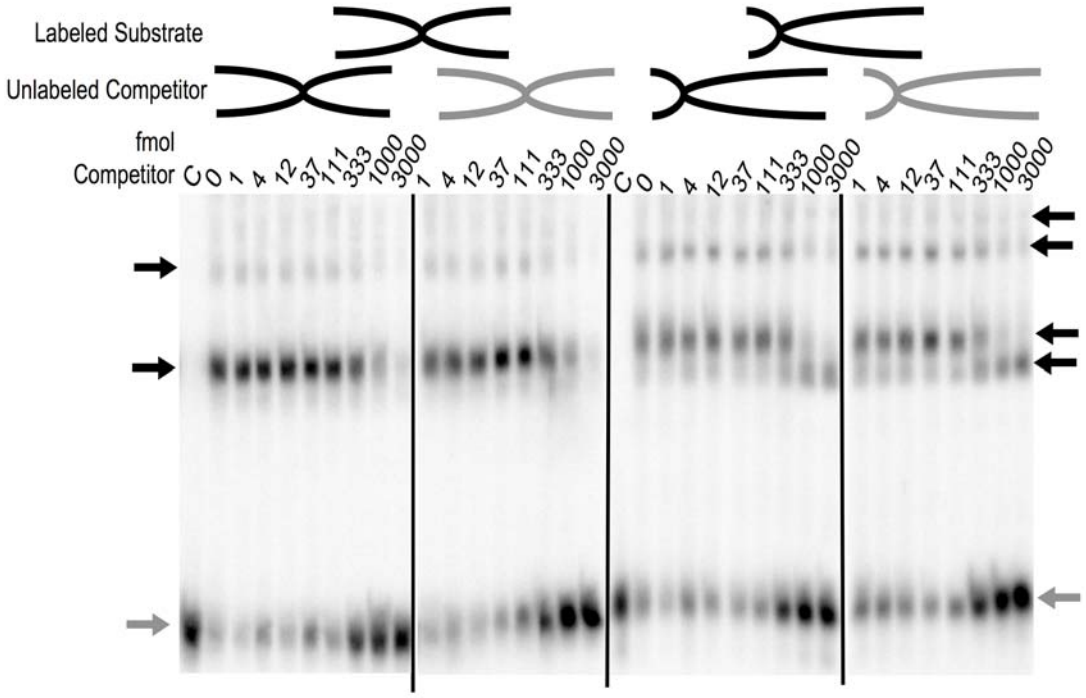

Labeled Substrate
Unlabeled Competitor
fmol
Competitor
C 0 1 4 12 37 111 333 1000 3000
1 4 12 37 111 333 1000 3000
C 0 1 4 12 37 111 333 1000 3000
1 4 12 37 111 333 1000 3000

Figure 3.10. A model of an excision reaction through a Holliday junction. A chromosome with four markers (A, B, C, and D) is the substrate for a DEDD motif transposase. Black arrows delineate the steps of the reaction. Hydrolysis catalyzed by the transposase is marked by black triangles. 3’ hydroxyl nucleophile attacks are marked with white arrows.

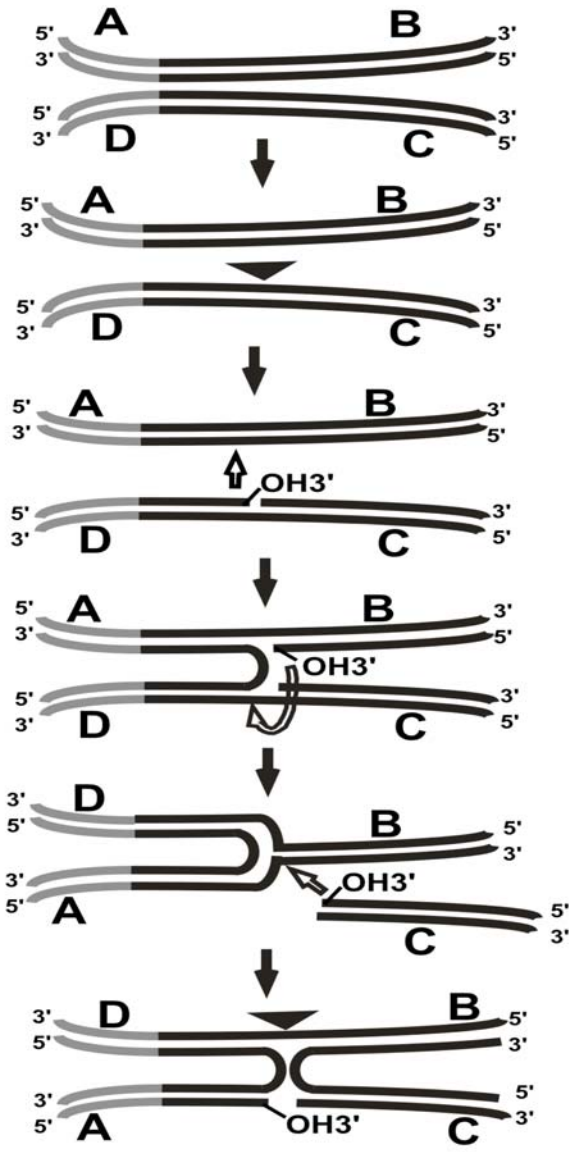

A
B
D
C
OH3'

Table 1. Oligonucleotides used in mutagenesis.

Name	Sequence*	Site
L176AF	5′CATCACACAG***GCC***GATTCATCAAATTTGG3′	HaeII
R	3′GTAGTGTGTC***CGG***CTAAGTAGTTTAAAGG5′	
L197GF	5′GTGCCAGCCTA***GGC***GCCACAATACC3′	NarI
R	3′CACGGTCGGAT***CCG***CGGTGTTATGG5′	
K206EF	5′CAATACCTGGAATTGGCAAA***GAG***AC**C**CTGCCACATCTAC3′	BsaI
R	3′GTTATGGACCTTAACCGTTT***CTC***TGGGACGGTGTAGATG5′	
L183DR	5′CAATTGAT***ATC***TTGCCAAATTTGATGATCG3′	EcorV
JBPIV1	5′ATGTCTAAAACTTACATTGGGATTG5′	

*Targeted codons are italicized. Changed bases are bolded. Restriction sites are underlined.

Table 2. Oligonucleotides used in the construction of the Holliday junctions.

Mobile Piv (IKST)*

I5′GATTGAATTCGCCATTATTGGTATCCTAGCTGCAATCGCTCTACCTGCAGACCA3′
K5′TTGGTCTGCAGGTAGAGCGATTGCAGCTAGGATACCAATAAATGTTATTCATGAA3′
S5′TTTCATGAATAACATTTATTGGTATCCTAGCTGCAATGCGGGGCACGGAGGGACG3′
T5′TCGTCCCTCCGTGCCCCGCATTGCAGCTAGGATACCAATAATGGCGAATTCAATC3′

Immobile Piv (INOP)**

N5′TTGGTCTGCAGGTAGAGCGATTGCAGCTAGGATACCAATAATT**ATTGGGAGTCGTAAGTGTCACG**
CACAGTTCTTACGGTG3′
O5′**TCACCGTAAGAACTGTGCGTGACACTTACGACTCCCAAT**AAATGTTATTCATGAA3′
P5′TTTCATGAATAACATTGGCGAATTCAATC3′

Mobile non-specific (UVWXII)

U5′TACGCCACGGTATCGTCGGCTCTGTAACTATGTCTAAGTCACTCAGGTGTGCGAG3′
V5′TCTCGCACACCTGAGTGACTTAGACATAGTTACAGAGCCGAAGATGATTATTCTC3′
W5′TGAGAATAATCATCTTCGGCTCTGTAACTATGTCTAAACAACCAGTAGCGTTTCA3′
XII5′TTGAAACGCTACTGGTTGTTTAGACATAGTTACAGAGCCGACGATACCGTGGCGT3′

Immobile non-specific (UYZAAII)

Y5′TCTCGCACACCTGAGTGACTTAGACATAGTTACAGAGCCGATACTGCGGACCAGCAACA
AGATGATTATTCTCCTTCAATG3′
Z5′TCATTGAAGGAGAATAATCATCTTGTTGCTGGTCCGCAGTAGTTCGCTTTGCCTA3′
AAII5′TTAGGCAAAGCGAACCGATACCGTGGCGT3′

*Note that the 3’ end of I anneals to the 5’ end of K, the 3’ end of K anneals to the 5’ end of S, the 3’end of S anneals to the 5’ end of T, and the 3’end of T anneals with the 5’ end of I.

**I anneals to N and P, N anneals to I and O. The bolded sequence of N and O is randomized sequence that does not match *tfpI*/*Q* sequence, but does anneal to each other.

CHAPTER 4: Dissertation discussion

Holliday junctions as a connection between Piv and the transposases of the IS*110*/IS*492* family. Piv exhibits significant homology to the transposases of the IS*110*/IS*492* family, but it has the activity of a DNA invertase. Molecular modeling of the Piv amino terminal region revealed a ribonuclease H-like fold, similar to that is found in DDE motif transposases and in the Holliday junction resolvase RuvC (13). In addition, secondary structure predictions indicated a possible Helix-hairpin-Helix (HhH) structure in the carboxyl-terminal region of Piv. The HhH is used to bind DNA, and in the case of RuvA a pair of HhH structures specifically bind Holliday junctions. Experimentally, we have demonstrated a requirement for D9, E59, D101, and D104 residues that are found in the RNaseH-like fold (3). RuvC has a similar arrangement of a tetrad of acidic residues forming a catalytic pocket (1, 11), while transposases only have a triad of residues (5). We have also demonstrated experimentally a requirement for two residues predicted to be part of the HhH motif, L197 and K206. L197 is part of the HhH consensus sequence. K206, though it is not conserved in the HhH consensus sequence (10), or the IS*110*/IS*492* transposases (Chapter 3), it is required by RuvA for Holliday junction binding (10). This lysine in the HhH motif could represent the divergence between the site-specific invertase Piv and the transposases of the IS*110*/IS*492* family. It does not rule out the presence of the HhH motif in the transposases, (IS*1000* transposase, for example, has an excellent match for the HhH consensus) (Chapter 3) but it could indicate differences in which residues make contact with DNA or how the HhH motif is used since it can mediate interactions with DNA structures

other than Holliday junctions. Furthermore, the HhH motif may be the source of MBP-Piv binding to single-stranded and duplexed DNA in a manner similar to DNA polymerases (10, 12).

We have proposed a model of inversion with Piv that involves a Holliday junction intermediate (Chapter 2). The model proposed presents a mobile Holliday junction capable of translocation. In Chapter 3, I proposed a second model of Piv inversion, one involving an immobile Holliday junction intermediate.

Mobile junction generation is postulated to involve a nicking reaction and strand invasion with host enzymes repairing the nicked DNA substrate and leading to a mobile Holliday junction. In this model, the DNA strands remain connected at all times, so double-strand breaks are not an issue as long as the last nick created in the reaction is sealed and the strands are continuous. In fact, our model of inversion requires either translocation or strand invasion to put the nick in a place where host factors can act on it. Translocation may require a host-encoded factor such as RuvB (7) to drive the movement of the strands. This may require "disengagement" of Piv from the system to drive translocation; alternatively RuvB, or a RuvB analog, may interact directly with Piv to catalyze translocation. Resolution of the mobile Holliday junction could occur in either plane, although resolution in one orientation would result in inversion and in the other, restoration of the original orientation.

The immobile Holliday junction as we have modeled it in Chapter 3 creates a "Y" intermediate and a 3'-hydroxyl that Piv would use as a nucleophile to create the nicked Holliday junction with *tfpQ* and *tfpI* on adjacent rather than opposite arms. Each initiation of inversion would result in inversion as cleavage in the proper plane would leave two nicked DNA stands, but cleavage in the other plane would create a deletion, something not seen in this system. Because the nicked Holliday junction intermediate is incapable of translocation, resolution would

occur at the same locus of Holliday junction formation. It is also possible, as a means of resolution, to use the free 3'-hydroxyl end of the DNA as a starting point for DNA synthesis. This would also resolve the structure presented and possibly create circular pieces of non-replicating DNA.

Because Piv bound both the mobile and immobile Holliday junctions, it is unclear which junction is the true recombination intermediate. It is possible that a specificity region on the protein is inactive because of mis-folding, or the MBP tag could be blocking a specificity domain, but the fusion protein maintains activity *in vivo* (14). Based on competition assays (Chapter 3), MBP-Piv appears the binding is structure- and not sequence-specific. *In vivo* inversion probably only uses one pathway; therefore, Piv and other factors, such as accessory proteins or sites (such as *sub1*), interact to form a machine to invert the DNA. *In vitro* cleavage assays are the best way to solve this question. Presumably, Piv will cleave one substrate, and not the other. One difficulty with this assay is we do not have a cleavage system yet. A second problem with this assay is in working with the immobile system we have to guess as to where the cleavage site is, because if the resolution is site-specific, we might not see it unless we use the right immobile substrate. It might also be possible to determine which junction is important through *in vivo* assays. The mobile Holliday junction might be capable of gene conversion, if translocation occurs and can go though mismatch regions. If no gene conversion is detected, it could be because the immobile junctions are used in inversion, or the mismatch regions stop translocation.

The connection between Piv and the transposases of the IS*110*/IS*492* family could be the Holliday junction. The Holliday junction in our original model of inversion could not be the same as the Holliday junction used in the transposition events because there is no way to excise

DNA by cleavage across the junction. The immobile Holliday junction could be a transposition intermediate. The circular intermediates observed in IS*492* transposition could reflect replication of the circular structure that is part of the proposed immobile pathway. However, the path for immobile inversion, if applied directly to IS*492* transposition, would not be correct. The path always results in inversion as the outcome. Inversion of IS*492* elements has not been observed.

The purpose of MBP-Piv binding to single-stranded DNA substrates MBP-Piv is unclear at this time. Other Holliday junction resolvases such as T4 endonuclease VII or T7 endonuclease I have wider substrate ranges than RuvC. RuvC, however, with the RNaseH-like motif, is specific to Holliday junctions though it is not clear what protein structures of RuvC are required for this specificity (2, 4, 6, 8, 9). It is possible that Piv's HhH motif binds any DNA in a non-structure-specific manner. It may also be that the mechanism of inversion and strand transfer involves long regions of single stranded DNA. The process of inversion could look similar to the process of homologous recombination, with the nicking of homologous regions of *tfpI/Q* and a strand invasion occurring with the complementary partner. This could work with strand transfer events occurring on the same chromosome, or with recombination occurring between two chromosomes. Even this model would rely on the resolution of Holliday junctions at some point in the process and if this is the model of inversion, then a larger unification of Piv and the DEDD motif transposases is unclear, because there is significantly less or often no homologous sequence in the insertion sequences or transposons of the IS110/IS*492* elements. So, rather then unifying these systems, it may be that single-stranded DNA binding divides them.

The identification and isolation of intermolecular products proved invaluable to the Holliday junction hypothesis. we first tried to explain the observations in the framework of a transposition reaction before being struck with the possibility of Piv using a Holliday junction to

catalyze recombination. This led us from an "artifact" of our *in vivo* inversion system to our current model. The system could also be tested for what is required for intermolecular recombination. Our *in vivo* inversion assay does not use the *sub1* site. Therefore, intermolecular recombination can occur in its absence, but can it occur in the presence of *sub1*? Or would *sub1* restrict the system to inversion? The substrates tested for intermolecular recombination are capable of inversion. What would be the outcome with substrate lacking one of the *inv* sites, or in the complete absence of *piv* on the inversion substrate, or with a large section of the invertible segment removed? Each of these regions contains Piv binding sites; how might they function in intermolecular recombination, if at all? The *in vivo* inversion system has two *invL* sites: one on the inversion substrate, the other on pAG800.2. Is intermolecular recombination influenced by the proximity of this *invL* in *cis* to *piv*? Moving the *invL* site to different loci on the expression vector, or off the expression vector but onto a third plasmid could measure some of the limits of intermolecular recombination. Whereas none of these experiments individually may say anything directly about the inversion system, it may be possible to use the assays as tools to define more of the requirements for Piv activity. For example, are the same bases that are required for intermolecular recombination required for DNA inversion? If the minimum DNA requirements for intermolecular recombination are defined, it may help with future rounds of oligonucleotide design for *in vitro* binding and Holliday junction resolution assays.

The next two steps in determining what is important for Piv function is the crystal structure, and an *in vitro* cleavage assay. The crystal structure will give us a confirmation or refutation of our structure predictions. It will be interesting to see how close our predictions are. If they are far off, will Piv have any structural similarities to other proteins? Perhaps it will have

homology to one of the other Holliday junction resolvases, or the serine or tyrosine recombinases. Will it have the HhH? If many of these features are missing, then how will we revamp our Holliday junction model of inversion? The *in vitro* assays are also very important to develop because intermediates in the reaction can be isolated and characterized. Substrate and cleavage site preferences can be determined along with the positive or negative interactions with small molecules, DNA, or other proteins. The development of *in vitro* assays has helped develop clear pictures of the operations carried out by the Tn*5* transposase, RuvC, and the bacteriophage T4 endonuclease VII and T7 endonuclease I, to name just a few examples.

Research on this family should continue because although it appears that these family members may be limited to bacteria, the techniques and insights we develop here could help in the exploration of new families of DNA recombinases. There is also a philosophical point in asking why and how did the DEDD motif family of recombinases develop? What advantages does the DEDD site-specific recombinase have over serine and tyrosine recombinases? All organisms encode Holliday junction resolvases, but RuvC and RuvC homologs are not encoded in all organisms, for example, there are no homologues of RuvC detected in Eukaryotes thus far. How could a site-specific DNA invertase acquire the RNaseH-like fold? If the Piv captured its DNA invertase from another system, are the transposases of the IS*110*/IS*492* family the progenitors of the Piv system or is Piv a trapped transposase?

I see two ultimate outcomes for this research: 1) Piv is linked directly to the transposases of the IS*110*/IS*492* family of recombinases, through protein structures, reaction intermediates, and reaction products. 2) Piv is the founding member of the a new group of site-specific invertases, a group that uses a ribonuclease H-like motif to catalyze site-specific inversion

though hydrolysis of the DNA backbone to create a Holliday junction and a second hydrolysis to resolve that junction into *bona fide* recombination products.

References

1. **Ariyoshi, M., D. G. Vassylyev, H. Iwasaki, H. Nakamura, H. Shinagawa, and K. Morikawa.** 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. Cell **78:**1063-72.

2. **Bennett, R. J., H. J. Dunderdale, and S. C. West.** 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell **74:**1021-31.

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

4. **Duckett, D. R., M. J. Panis, and D. M. Lilley.** 1995. Binding of the junction-resolving enzyme bacteriophage T7 endonuclease I to DNA: separation of binding and catalysis by mutation. J Mol Biol **246:**95-107.

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

6. **Lu, M., Q. Guo, F. W. Studier, and N. R. Kallenbach.** 1991. Resolution of branched DNA substrates by T7 endonuclease I and its inhibition. J Biol Chem **266:**2531-6.

7. **Parsons, C. A., A. Stasiak, R. J. Bennett, and S. C. West.** 1995. Structure of a multisubunit complex that promotes DNA branch migration. Nature **374:**375-8.

8. **Parsons, C. A., and S. C. West.** 1990. Specificity of binding to four-way junctions in DNA by bacteriophage T7 endonuclease I. Nucleic Acids Res **18:**4377-84.

9. **Picksley, S. M., C. A. Parsons, B. Kemper, and S. C. West.** 1990. Cleavage specificity of bacteriophage T4 endonuclease VII and bacteriophage T7 endonuclease I on synthetic branch migratable Holliday junctions. J Mol Biol **212:**723-35.

10. **Rafferty, J. B., S. M. Ingleston, D. Hargreaves, P. J. Artymiuk, G. J. Sharples, R. G. Lloyd, and D. W. Rice.** 1998. Structural similarities between *Escherichia coli* RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the holliday junction. J Mol Biol **278:**105-16.

11. **Saito, A., H. Iwasaki, M. Ariyoshi, K. Morikawa, and H. Shinagawa.** 1995. Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. Proc Natl Acad Sci U S A **92:**7470-4.

12. **Shao, X., and N. V. Grishin.** 2000. Common fold in helix-hairpin-helix proteins. Nucleic Acids Res **28:**2643-50.

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

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