A MOLECULAR MECHANISM ALLOWING TRANSPOSON TN7 TO TARGET ACTIVE DNA REPLICATION

A DISSERTATION

PRESENTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF CORNELL UNIVERSITY

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

ZAOPING LI

JANUARY 2012

©2012 ZAOPING LI

A MOLECULAR MECHANISM ALLOWING TRANSPOSON TN7 TO TARGET ACTIVE DNA REPLICATION

Zaoping Li, PhD

Cornell University 2012

Transposons are jumping genes that are ubiquitous and abundant in all domains of life. They can move between locations that lack homology within a genome. Transposons drive the evolution of genomes through gene inactivation, expression modulation, and genome rearrangement. The bacterial transposon Tn7 and its relatives are widespread in diverse bacteria, likely due to their ability to control the frequency and targeting of transposition. My work presented here focuses on understanding the molecular mechanism of the TnsABC+E pathway of Tn7 transposition that preferentially targets actively transferring mobile DNA. I was able to establish a sensitive *in vitro* system for this transposition reaction with purified proteins and gapped DNA substrates preloaded with the β-clamp. The transposition profile recapitulates that observed *in vivo,* indicating that the minimal features recognized by TnsE to target DNA replication are 3' recessed ends found in target DNA and the β-clamp processivity factor (DnaN). I further show that the TnsE-β interaction is largely conserved among Tn7-like elements; however, this interaction is also species specific. In a heterologous expression study, I found that TnsE homologs from *Idiomarina loihiensis* and *Shewanella baltica* only promoted transposition when DnaN from the same host was used in the cell. I propose that TnsE may have evolved to interact with the more variable portion of the clamp to avoid

interfering with the normal traffic on the clamp. In an effort to screen for host proteins that may affect TnsE-mediated transposition, I found and confirmed an interaction between TnsE and SeqA, a protein involved in replication initiation control and organizing newly replicated chromosome DNA. Results from genetic studies support a model where TnsE interacts with SeqA to disrupt the SeqA superstructure that tracks with the replication fork. I also show that in wild type background, TnsE is able to direct transposition into the origin region and DNA undergoing leading-strand replication, consistent with the emerging picture that both the leading-strand and lagging-strand DNA replication are essentially discontinuous. These data point to a perspective of using Tn7 as a genetic tool in understanding the replication and repair processes in the cell.

BIOGRAPHICAL SKETCH

Zaoping Li was born in a small village located in Hunan Province, China. After primary schools, he left his family for a special class in a middle school almost thirty miles away form his hometown and since then he moved further and further on his long road of study.

Zaoping went to the School of Public Health at the West China Medical Center of Sichuan University for an undergraduate study and obtained a Bachelor of Medical Science degree. During this five-year period, he was particularly fascinated about Microbiology and Molecular Biology and decided to continue his Master's study at the same school with Professor Hengchuan Liu. He then transiently worked in Beijing Centers for Disease Control and Prevention as a researcher, but soon realized his real interest was still in science.

In 2005 Zaoping had the great opportunity to continue his study for a PhD degree at Cornell University in the United States. After experiencing different aspects of studies in the field of Microbiology by rotation, he joined the laboratory of Dr. Joseph Peters to study transposon Tn7, where he continued his interest in mobile genetic elements and acquired solid training in Genomics and Biochemistry.

THIS THESIS IS DEDICATED TO MY WIFE, LICHUAN YE, AND MY PARENTS,

YUXIAO LI AND SANNIANG LIU

ACKNOWLEDGEMENTS

I would like to acknowledge people who have supported and encouraged me during my years on the academic path. My parents, Yuxiao Li and Sanniang Liu, have been cheerful, encouraging, and supportive for my efforts, but I understand what they have been through to support a higher education with all that they have had. Professor Hengchuan Liu at West China Medical Center of Sichuan Univeristy was the first scientist who extensively inspired my curiosity about science and led me into scientific research from college, and has always been optimistic about my career. I am grateful to the committee members of my PhD study, Dr. Joseph Peters, Dr. Jeffrey Roberts, and Dr. John Helmann, for their supports and insightful advice. I appreciate the guidance from my major advisor Dr. Joseph Peters, who gave me the chance to explore, but also kept my adventures on the right track. Dr. Joseph Peters has also been very helpful in my understanding the new culture. I am thankful to my friends and colleagues in the Microbiology program, especially Adam Parks and Qiaojuan Shi in the Peters' lab for the wonderful discussions and suggestions. Finally, I am deeply grateful to my wife, Lichuan Ye, who backed me up ever since my Master's study and has accompanied my pursuit this far.

TABLE OF CONTENTS

Biographical Sketch iii

Dedication iv

Acknowledgements v

Table of Contents vi

List of Figures xii

List of Tables xv

1. **Introduction** 1

1.1. Tn7 transposition functions 5

1.1.1. The Tn7L and Tn7R ends are structurally and functionally different 5

1.1.2. The core TnsABC machinery 8

1.1.2.1. TnsAB the transposase 9

1.1.2.2. TnsA: a restriction enzyme in recombination 10

1.1.2.3. TnsB is a member of the large transposes/retroviral integrase superfamily 11

1.1.2.4. TnsC, an AAA+ regulator 15

1.2. Target selection in the TnsABC+D transposition pathway 23

1.2.1. Sequence requirements of attTn7 24

1.2.2. Target DNA binding by TnsD 25

1.2.3.TnsD interaction with TnsC 28

1.2.4. Host factors in TnsD-mediated transposition 29
1.3. Target selection in the TnsABC+E transposition pathway 29
1.3.1. TnsE-mediated transposition targets lagging-strand DNA replication 30
1.3.2. Factors important for TnsE targeting of lagging-strand DNA replication ... 31
1.3.3. Other DNA replication processes targeted by TnsE 39
1.3.4. Why target lagging-strand DNA replication? 40
1.4. Recognition of a positive target signal by TnsC 43
1.5. Target immunity 45
1.6. Expansion of the Tn7 family by virtue of vertical and horizontal transposition ... 49
1.7. Internal networking of the core machinery 53
1.7.1. Interaction between TnsA and TnsB 53
1.7.2. Interaction between TnsA and TnsC 53
1.7.3. Interaction between TnsB and TnsC 54
1.8. Assembly of the transpososome for TnsABC+D transposition 60
1.8.1 Pre-transposition complex 60
1.8.2 The Post-transposition complex 61
1.9. Tn7 as a genetic tool 66
1.10. Concluding remarks 68
1.11. Acknowledgements 69
1.12. References 70

2. **Transposition into Replicating DNA Occurs through Interaction with the Processivity Factor** 91

2.1. Summary 91

2.2. Introduction 91

2.3. Results 94

2.4. Discussion 101

2.5. Experimental procedures 106

2.5.1. Plasmids and strains 106

2.5.2. Protein purification and labeling 106

2.5.3. Purification of target DNA substrates 107

2.5.4. *In vitro* transposition assays 108

2.6. Acknowledgements 109

2.7. References 110

3. **Heterologous Expression Reveals a Mechanism for Transposon Tn7 Host Specificity through Adaptation to Replication Processivity Factor** 116

3.1. Summary 116

3.2. Introduction 117

3.3. Results 120

3.3.1. Phylogenetic distribution of Tn7-like elements 120

3.3.2. DnaN^{Il} and DnaN^{Sb} functionally complement the previously described *dnaN159* phenotypes 125

3.3.3. *dnaN159 recA* strains show severely compromised growth when expressing DnaN^{Il} or DnaN^{Sb} 139

3.3.4. Accessory polymerase usage in dnaN159 strains expressing DnaN^{Il} or DnaN^{Sb} 139

3.3.5. TnsE and DnaN must come from the same foreign host to allow transposition with the core Tn7 machinery in E. coli 141

3.3.6. Conservation of DnaN sequences 145

3.4. Discussion 145

3.4.1. Specific interactions for β-binding proteins on the clamp 148

3.4.2. TnsE-β interaction in Tn7 transposition 153

3.4.3. Role of DNA polymerase II 154

3.4.4. The roles of RecA in promoting the function of accessory polymerases .. 157

3.4.5. A novel way of studying the structure and function of essential proteins 158

3.5. Materials and Methods 158

3.5.1. Media, chemicals and enzymes 158

3.5.2. Bacterial strains and plasmids 159

3.5.3. SOS response 164

3.5.4. *In vivo* analysis of transposition 164

3.5.5. Spontaneous mutation analysis 165

3.5.5. Computational methods 166

3.6. Acknowledgements 166

3.7. References 168

4. **TnsE Reveals Transposition Target by Interaction Disruption of the SeqA complex** 181

4.1. Summary 181

4.2. Introduction 182

4.3. Results 186

4.3.1. Identification of a SeqA-TnsE interaction 186

4.3.2. A direct physical interaction between TnsE and SeqA 187

4.3.3. Inactivation of SeqA specifically stimulates TnsE-mediated transposition 188

4.3.4. SeqA and Dam are not epistatic on TnsE-mediated transposition 194

4.3.5. Effect of replication initiation asynchrony on Tn7 transposition........ 197

4.3.6. Effect of replication overinitiation on Tn7 transposition 198

4.3.7. Effect of altered global gene expression on Tn7 transposition 202

4.3.8. Effect of chromosome structure on Tn7 transposition 202

4.3.9. TnsE targets the origin and leading-strand........ 207

4.4. Discussion 210

4.5. Materials and Methods 218

4.5.1. Media, chemicals and enzymes 218

4.5.2. Bacterial strains and plasmids 218

4.5.3. Transposition assays 219

4.5.4. Protein purification 223

4.5.5. Pull-down/mass spectrometry experiment 224

4.5.6. Protease Footprinting 225

4.6. Acknowledgements 225

4.7. References 227

5. **Conclusion** 240

6. **Appendix 1: Tn7-like elements** 244

7. **Appendix 2: TnsE mutations with increased activity isolated in this study** 247

LIST OF FIGURES

1.1. Map of transposon Tn7 2

1.2. Cut-and-paste transposition of Tn7. 6

1.3. TnsA and its gain-of-function mutations. 13

1.4. TnsC plays the central role in *Tn7* transposition regulation 16

1.5. The nucleotide-binding motif in TnsC and TnsC gain-of-function mutations 20

1.6. Organization of *attTn7* and its sequence conservation 26

1.7. TnsE-mediated transposition targets lagging-strand DNA replication for transposition 32

1.8. The β-clamp binding motif in TnsE homologs 37

1.9. Persistence of gap structures may lead to regional preference of TnsE-mediated transposition into both plasmid and chromosome. 41

1.10. Activation of TnsC by various target complexes. 47

1.11. *Tn7*-like elements mediate formation of genomic island 51

1.12. TnsC C-terminal peptide on the hydrophobic surface of TnsA 55

1.13. TnsB C-terminus is involved in interaction with TnsC and mutations that affect target immunity and transposition overlap 58

1.14. Illustration of TnsABC+D transposition pathway at the transpososome level 64

2.1. An *in vitro* TnsABC+E transposition assay shows that DNA structures are important for TnsE-mediated transposition, and that the β clamp is required to direct the orientation of transposon ends 99

3.1. Phylogenetic tree of Tn7 family 122

3.2. TnsE homologs are not functional with TnsABC from *E. coli* for transposition in wild type strain. 127

3.3. $DnaN^{Il}$ and $DnaN^{Sb}$ but not $DnaN^{Bc}$ were able to complement the phenotypes of *dnaN159* allele. 129

3.4. SOS induction of *dnaN* wild type strain and *dnaN159* strains with or without DnaN homologs 134

3.5. Growth of *dnaN159*, *recA*, and *dnaN159 recA* mutants with or without heterologous clamps at 30°C on Isosensitest agar 137

3.6. Role of accessory DNA polymerases in *dnaN159* strains with DnaN homologs 142

3.7. TnsE homologs allow transposition with Tn7 TnsABC with, and only with, DnaN from the same host.. 146

3.8. Conservation pattern on the faces of the β-clamp 150

4.1. Protease footprinting suggests a specific interaction between TnsE and SeqA....... 189

4.2. TnsE-mediated transposition is stimulated by inactivating *seqA*. 191

4.3. Inactivating Dam methylase stimulated TnsE-mediated transposition but not because of increased double-strand breaks. 195

4.4. Lack of SeqA function of in replication initiation regulation does not account for the stimulatory effect of inactivating SeqA on TnsE mediated transposition....................... 200

4.5. Effect of altered global gene expression on Tn7 transposition 203

4.6. Effect of chromosomal organization on Tn7 transposition 205

4.7. Location and orientation of independent TnsE-mediated transposition events in the chromosome of *recA*- wild type, and *seqA*- backgrounds 208

4.8. Potential regions of TnsE involved in SeqA interaction 216

LIST OF TABLES

3.1. Homology of TnsE and DnaN homologs 126

3.2. Colony size of *dnaN159 recA* strains over time. 136

3.3. DnaN sequences used for generating the sequence alignment for modeling. 149

3.4. Strains used in this study 161

3.5. Plasmid DNAs used in this study 163

4.1. Strains used in this study 220

4.2. Plasmids used in this study. 222

CHAPTER 1: *INTRODUCTION*[1]

The bacterial transposon Tn7 is a sophisticated mobile genetic element that is widespread in the environment in highly divergent bacteria. In contrast to most other transposable elements that only recognize a single type of target, Tn7 mobilizes through several alternate targeting mechanisms that probably account for its wide distribution. In one targeting pathway, Tn7 transposes at a high frequency into a specific site on the chromosome, the *attTn7* site or attachment site of Tn7, which is highly conserved in all bacteria examined to date. Insertion into this site does not have any obvious negative effect on the host and thus this pathway serves to direct Tn7 to a "safe haven" in the host and likely facilitates the vertical transmission of Tn7. In another pathway, Tn7 insertions occur preferentially into plasmids capable of moving between bacteria called conjugal plasmids. Given that conjugal plasmids can often have a broad host-range, this pathway promotes the dissemination of Tn7 via horizontal gene transfer. There is also evidence that Tn7 elements in the environment may have mutations that favor a broader array of insertion sites to help meet the selection environment (see below).

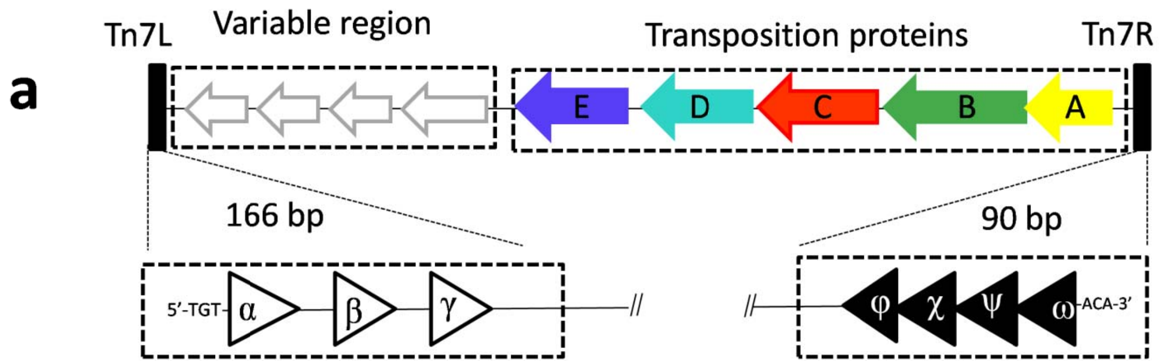
Transposon Tn7 encodes five proteins for transposition, TnsA, TnsB, TnsC, TnsD, and TnsE (Figure 1.1). Of these five proteins, TnsABC+D are required for transposition into the *attTn7* site in the chromosome, while TnsABC+E are required for targeting conjugal plasmids for transposition (33, 110, 152). TnsABC is shared by the two

---

[1]This chapter was prepared as a book chapter with Craig, N.L. and Peters, J.E..
Li, Z**.**, Craig NL, and J. E. Peters. 2011. Transposon Tn7. In: Roberts AP, Mullany P, eds. Bacterial Integrative Mobile Genetic Elements**.** Austin: Landes Bioscience: *Epub ahead of press*.

Figure 1.1. **Map of transposon Tn7**. (a) A schematic representation of Tn7 and its end structure. The transposition genes are carried in the right side of Tn7 as an array that is in synteny in Tn7-like elements as indicated as A, B, C, D, E for *tnsA*, *tnsB*, *tnsC*, *tnsD*, and *tnsE* in the figure. The variable region contains genes not related to Tn7 transposition and the composition of this region varies in different Tn7-like elements. In the original Tn7, antibiotic resistant genes, *dhfr* (trimethoprim resistence), *sat* (streptothricin resistance), and *aadA* (streptomycin and spectinomycin resistance) are found in a defective class 2 integron cassette system. The left (Tn7L) and right (Tn7R) ends are the *cis*-acting factor for transposition. Both ends contain multiple TnsB homologous binding sites represented as triangles: three discrete sites in Tn7L (α, β, γ from exterior to interior) and four overlapping sites (ω, ψ, χ, φ from exterior to interior) in Tn7R. (b) Sequence conservation of the seven TnsB binding sites of Tn7. The sequence logo was generated with the online tool WebLogo (37) after aligning the sequences of the TnsB binding sites by MUSCLE algorithm in UGene software http://ugene.unipro.ru.

a



b

pathways and is therefore sometimes called the core transposition machinery. TnsD and TnsE are alternative target site selectors that activate TnsABC and promote insertion into specific preferred target DNAs. No transposition occurs with wild-type TnsABC alone. Tn7 has another type of regulation that impacts both the TnsD-mediated and TnsE-mediated pathways of transposition; Tn7 transposition is strongly inhibited from inserting into a target DNA that already has a copy of Tn7 (6, 33, 59), a phenomenon called “target immunity” that is also found with other transposons such as bacteriophage Mu (119) and Tn3 (123). At the center of Tn7 transposition regulation is the ATPase protein TnsC that conveys signals between the target selecting protein and the transposase TnsAB (33, 142). The use of an ATPase as a molecular switch in the control of transposition is also found with other transposons (142) including bacteriophage Mu (131, 155), Tn5090 (111), IS21 (118), to name a few.

Tn7 transposes by a cut-and-paste mechanism where the element is excised from the donor DNA by double-stranded breaks at the ends of the element while being directly joined into the target DNA via the 3’-ends of the element (Figure 1.2) (9). Recombination requires the formation of a nucleoprotein complex with all the required transposition proteins, the Tn7 ends, and the preferred target DNA substrate. One of the hallmarks of Tn7 transposition is that no DNA breakage and joining is initiated in the absence of any of these components (9, 138). Because the joining of the transposon ends into the target DNA occurs with a stagger of 5 base pairs, subsequent repair of the staggered break in the target DNA by cellular processes generates the 5 base pair (bp) target site duplication

flanking the element at the new insertion site, a signature of Tn7 transposition (33)(Figure 1.2).

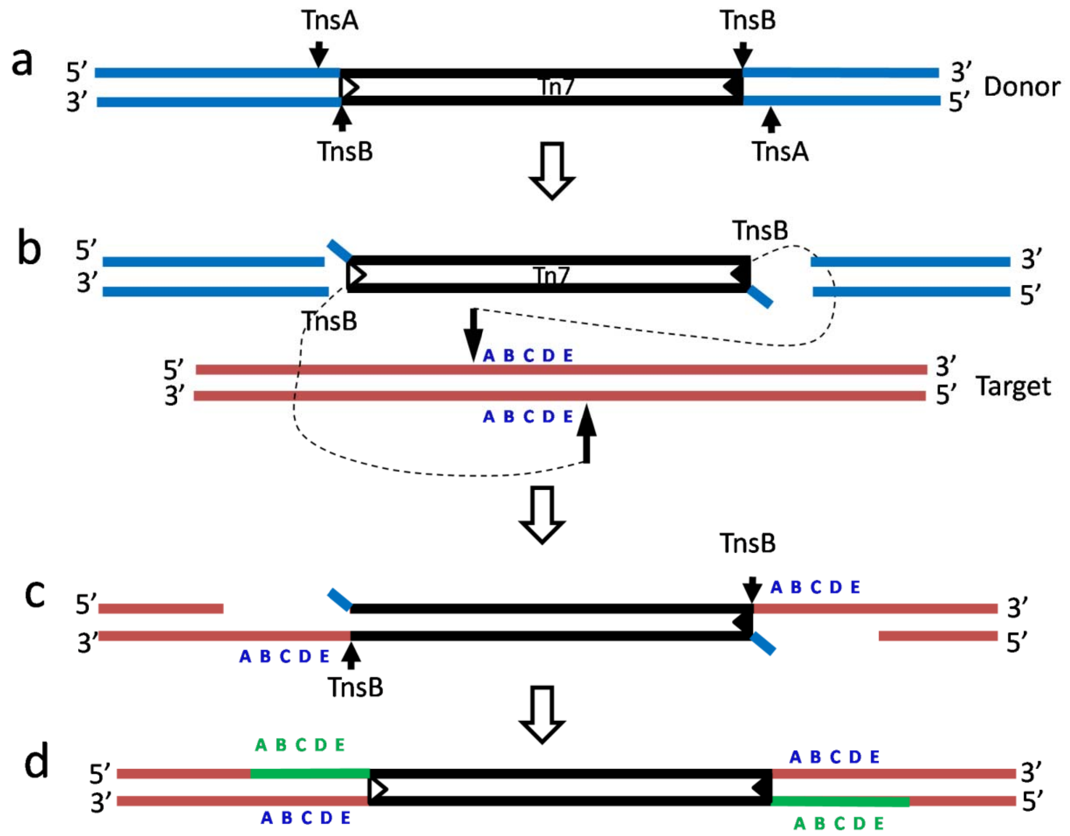
1.1. Tn7 Transposition Functions

1.1.1. The Tn7L and Tn7R ends are structurally and functionally different

The *cis*-acting functions for transposition are the left (Tn7L) and right (Tn7R) ends of Tn7. These ends contain multiple recognition sites for the TnsB subunit of the TnsAB transposase. Because the sequences recognized by the transposase are in opposite orientation in the left and right ends, these ends are sometimes referred to as the inverted repeat sequences. Like the *cis*-acting end sequences found in other transposons (38) there are two functional aspects to the ends of the elements: one for transposase binding and the other for signaling the point of DNA cleavage. The very terminal trinucleotide 5'-TGT…ACA-3', which is a highly conserved as the terminal sequence in many transposable elements, is essential for the breakage and joining reaction (50, 149). DNA breakage exposes the 3'-end of the element and the exposed 3'-A-OH is subsequently joined to the 5' end of target DNA (9, 50)(Figure 1.2). Multiple transposase binding sites are found within each of the transposon ends that are required for Tn7 transposition. The natural and fully functional Tn7 ends are rather large (5, 7, 80, 93)(Figure 1.1). The Tn7L end is 166 bp with four sites for TnsB binding. The Tn7R end is 90 bp with three sites for TnsB binding (6). The TnsB binding sites are similar, but not identical, with the spacing of the sequences relative to one another also being important.

Tn7L and Tn7R are also functionally distinct. Tn7 transposition has a strong orientation bias, in both TnsD-mediated transposition into the *attTn7* site (9, 33, 80), and

Figure 1.2. **Cut-and-paste transposition of Tn7**. (a) Tn7 is excised from the donor DNA by double-strand breaks, where TnsB cuts at the 3'-ends of the transposon, but TnsA cleaves 3 bp outside the element in the flanking DNA at the 5'-ends. (b-c)TnsB joins the free 3'-OH to 5 bp-staggered positions on the target DNA, generating single insertion products with gaps flanking the 5'-end. (d) Repair of the small gaps by cellular processes result in the 5-bp target site duplication characteristic for Tn7.



a
TnsA
TnsB
5'
3'
Tn7
3'
5'
Donor
TnsB
TnsA
b
TnsB
5'
3'
Tn7
3'
5'
TnsB
A B C D E
5'
3'
3'
5'
Target
A B C D E
c
TnsB
A B C D E
5'
3'
3'
5'
A B C D E
TnsB
d
A B C D E
A B C D E
5'
3'
3'
5'
A B C D E
A B C D E

TnsE-mediated transposition into conjugal plasmids and other targets (13-15, 80, 108, 154)(see below). An interesting, but unanswered question is what molecular mechanism accounts for the orientation specificity. The different arrangement of the TnsB binding sites probably dictates the different roles of Tn7L and Tn7R in organizing and assembling the transposition complex. It is important to note that this orientation must be set in the binding of the transposase to the ends and then this signal must also be coordinated with the binding of TnsC to appropriate target complexes. An interesting observation of unknown significance is that miniTn7 elements with two Tn7R ends are competent for transposition *in vivo* and *in vitro,* but elements with two Tn7L ends are not (6, 50).

1.1.2. The core TnsABC machinery

As noted above transposon Tn7 encodes five proteins for transposition (104). This is in contrast to what is found with most other transposons, which more commonly encode one or two proteins for catalyzing transposition. In Tn7 elements, genes encoding the five transposition proteins are located next to the right end as an array, where all genes are oriented in the same direction with their 5'-end toward the right end of Tn7 (Figure 1.1). Except for the fact that TnsB has homology to proteins of the transposase/retroviral integrase superfamily, the other four proteins each belongs to a unique family that only consists of homologous proteins from other Tn7-like elements. The novelty of the Tns proteins and the functional nucleoprotein complex that forms between these proteins and the ends of the element for transposition to occur make Tn7 a fascinating model for

studying protein-DNA interaction, protein-protein interaction, and assembly of multi-subunit molecular machinery.

1.1.2.1. TnsAB the transposase

In other transposition systems, the transposase consists of a single type of polypeptide, which binds specifically to the ends of transposon and does all the chemistry required for transposition. However, from early work it was clear that no strand breakage and joining activity was detectable with TnsB alone in the Tn7 system. Instead, multiple lines of evidence indicated that TnsB+TnsA together form a heteromeric transposase (18, 88, 128). Another feature that sets the Tn7 mechanism apart from other transposable elements is the way the element is cut away from the donor DNA during cut-and-paste transposons. With many DNA transposons the element excises from the flanking donor DNA by double-strand breaks through formation of hairpin structures (63). In the case of Tn5, Tn10, and *piggyBac*, there is initially a cut at the 3'-end and the exposed 3'-OH then attacks the opposite strand to form a hairpin at each end of the element (17, 33, 72, 95). Although the excision of the transposon *Hermes*, a member of the *h*AT transposon family, also involves formation of hairpin, the initial DNA breakage is at the 5'-ends of the transposon and the hairpins are formed on the donor DNA. This process is similar to the process carried out by the RAG recombinase during the V(D)J recombination process (158). The Tc1/*mariner* transposons are excised by sequential cleavage of both the 5'- and 3'-ends by the transposase (44, 150). For Tn7 transposition, the 3'-ends of the element are cleaved by TnsB, but TnsA is responsible for nicking at the 5'-ends (88, 128). The TnsB-mediated process of cleavage at the 3'-ends of the element occurs as a

direct joining event to the target DNA as is found with bacteriophage Mu, but in the case of Tn7 this process only occurs in the presence of TnsA (Figure 1.2)(128).

1.1.2.2. TnsA: a restriction enzyme in recombination

TnsA is responsible for cleavage at the 5'-ends of Tn7 element. Inactivating the strand cleavage activity of TnsA results in a switch from cut-and-paste to replicative transposition both *in vivo* and *in vitro* (88, 128). The N-terminal domain of TnsA is structurally homologous to catalytic domain of Type II restriction endonuclease FokI, which cuts DNA nonspecifically at a fixed distance from its recognition sequence (Figure 1.3)(64). Structural comparisons and mutational studies unequivocally identified the active site as a catalytic triad composed of E63, D114, and K132 (64, 88, 128), which coordinates two Mg^{2+} ions as a cofactor even in the absence of substrate DNA. Residue D114 directly contacts the metal ion and replacing residue D114 with cysteine alters the metal specificity from Mg^{2+} to Mn^{2+} (64, 88, 128). The crystal structure of TnsA also revealed a surface in the C-terminus of the protein that likely provides the interface for the interaction with TnsB (64). Due to the fact that TnsA alone does not possess detectable DNA binding ability (10), it is likely that positioning of TnsA for catalysis occurs though an interaction with the TnsB proteins that are bound to the ends of the element. In support of this idea, TnsA gain-of-function mutants containing mutations in this region can promote transposition with gain-of-function TnsB mutants in the absence of other transposition proteins (Figure 1.3)(83). In addition to its enzymatic activity, TnsA also plays a role in controlling the activity of TnsB, as TnsB-mediated 3'-end joining to the target DNA requires the presence of TnsA suggesting the TnsA and TnsB

proteins have evolved together to form a true heteromeric transposase capable of catalysis on both strands of DNA (128).

1.1.2.3. TnsB is a member of the large transposes/retroviral integrase superfamily

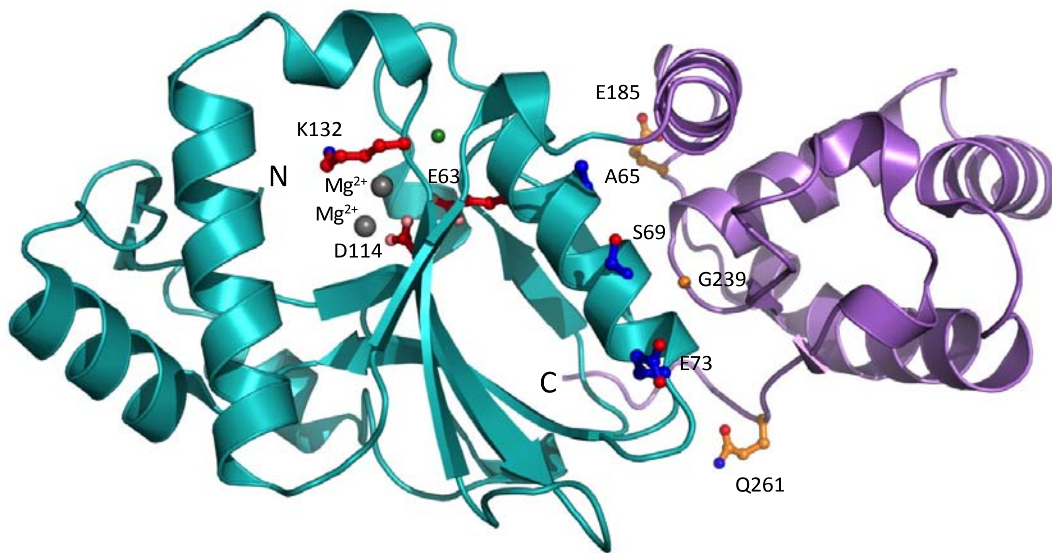
TnsB is a member of the transposase/retroviral integrase superfamily that includes transposases of other bacterial transposons and retrotransposons, and retroviral integrases, despite the early finding that only limited sequence similarity can be found between TnsB and other members of this family (43, 58, 63, 120, 121). Three functional domains may be present in TnsB: the N-terminal DNA binding domain, the catalytic domain, and a C-terminal domain for protein-protein interaction with TnsC, like the other recombinases (121).

The central core region is defined as the part of TnsB that has limited, but significant, sequence similarity with members of the transposase/retroviral integrase superfamily (33, 64, 128). The catalytic domain of the transposase/retroviral integrase superfamily has a conserved RNaseH fold, featuring a DD(35)E (DDE) signature motif that coordinates the two divalent metal ions at the active site assisting the various nucleophilic attack reactions involved in DNA cleavage and strand transfer (63, 121). TnsB mutants bearing site-specific mutations of the DDE motif residues are defective in Tn7 end cleavage and joining activities *in vitro* and inactive for TnsABC+D and TnsABC+E transposition *in vivo* (128). Furthermore, when the TnsB DDE mutants are used for *in vitro* TnsABC+D transposition specific intermediates accumulate, indicating that multiple active sites are involved in the transposition complex (128). Other work supports the idea that multiple TnsB proteins are present in the transposition complex (see below)(68).

TnsB is the only Tns protein shown to specifically bind the Tn7 ends (Figure 1.1). TnsB occupies its multiple binding sites on each end in a progressive and sequential manner, indicating that the various binding sites have different apparent affinities for TnsB (5). The most interior site of Tn7L (γ site) and χ site of Tn7R are occupied first (Figure 1.1). However, the sites immediately adjacent to the transposon termini, where DNA breakage and joining occur, are occupied only after the interior sites are bound, likely ensuring no DNA cleavage occurs until the full transposition complex is assembled. The fact that the Φ site of Tn7R is not bound until all the other three sites on Tn7R are filled, together with its dispensable role for transposition *in vivo* (6), indicates that it is not essential for assembly of the transposition complex (5)(Figure 1.1). TnsB binding to its recognition site results in asymmetric DNA bending in regions where it binds weakly, which may play a role in the assembly of the intricate transposition complex (5). The sequence-specific DNA binding function of TnsB is located at the N-terminus of the protein (33), which may consist of two subdomains, akin to the composition of the consensus sequence binding domain of MuA by domain Iβγ (26, 29, 30, 130). However, no canonical DNA binding motif is readily identifiable in this domain of TnsB and exactly how TnsB binds to the TnsB-binding sites at the ends of Tn7 is yet to be determined.

TnsB alone can bring the transposon ends together to form the so-called "synapsis" structure (138), which is a key step in transposition initiation. Transposases of Mu, Tn5/Tn10, Mos1 of the Tc1/mariner family and the P element also form synapsed structures for transposition (4, 25, 27, 40, 122, 148). Given that these transposases are all

Figure 1.3. **TnsA and its gain-of-function mutations**. In the ribbon diagram, the N-terminal catalytic domain of TnsA is shown in blue and the C-terminal domain is in purple. The active site of TnsA consists of E63, D114, and K132. Class I mutations of TnsA (A65, S69, and E73) that allow transposition with wild type TnsB and TnsC are on the same α-helix with E63. Class II mutations of TnsA (E185 and Q261) can promote transposition together with TnsB gain-of-function mutants ($TnsB^{M366I}$ and $TnsB^{A325T}$) independent of TnsC. (Figure was a kind gift from Alison Hickman, NIH, NIDDK)

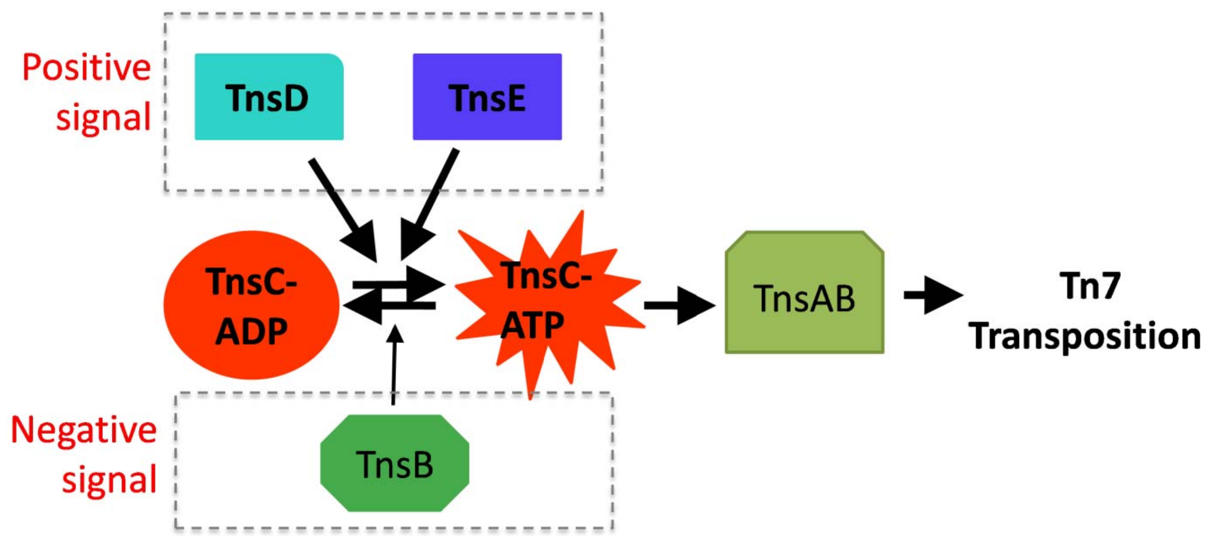

E185
K132
N
Mg²⁺
E63
A65
Mg²⁺
S69
D114
G239
C
E73
Q261

members of the transposase/retroviral integrase superfamily, but not close relatives, it may be that there is a common DNA binding and synapsis strategy for transposases in this superfamily. In the well-characterized Mu and Tn5 systems, it has been shown that catalysis occurs *in trans*, where proteins bound to the left end actually do the chemistry on the right end of the element and vice versa (4, 40, 129). It will be interesting to investigate if the TnsB-mediated strand breakage is also catalyzed *in trans* with Tn7. Another interesting question is how many TnsB proteins are in the core machinery. Although seven TnsB proteins can be detected in the post-transposition complex (see below)(68), it is possible that just two TnsB proteins are required for the 3'-end breakage and joining activity, but others are important for transpososome assembly, similar to the roles of multiple MuA proteins bound to the ends of bacteriophage Mu (26, 97, 99).

1.1.2.4. TnsC, an AAA+ regulator

TnsC is the regulator of Tn7 transposition that communicates between the target selecting protein TnsD or TnsE and the transposase (Figure 1.4)(10, 142). Whether or not a potential target is used for transposition is determined by the "state" of TnsC bound on the target DNA. In a productive transposition process, TnsC is likely recruited to the target DNA through interactions with the target DNA-protein (TnsD or TnsE) complex. TnsC then likely recruits the transposase-donor complex to form the transpososome. TnsA may also be recruited via interaction with TnsC to form a ACD-*attTn7* target complex (68) before the assembly of the transpososome with the TnsB-donor complex. In establishment of target site immunity, TnsC is actively displaced by TnsB from a potential target (see below).

Figure 1.4. **TnsC plays the central role in Tn7 transposition regulation**. TnsC in the ATP-bound state is the active form. TnsD and TnsE recruit TnsC to the corresponding targets allowing active TnsC. By contrast, TnsB inactivates TnsC and displaces it from targets with a preexisting copy of Tn7. Active TnsC can interact with the heteromeric transposase TnsAB and activate its DNA-breakage and DNA-joining activities.


Positive signal
TnsD
TnsE
TnsC-
ADP
TnsC-
ATP
TnsAB
Tn7
Transposition
Negative signal
TnsB

TnsC is a member of the AAA+ superfamily

The AAA+ (ATPases Associated with a wide variety of cellular Activities) superfamily is characterized by a conserved nucleotide phosphate-binding motif, where a Walker A motif binds the beta-gamma phosphate moieties of the bound nucleotide and a Walker B motif coordinates a Mg^{2+} cation at the active site (57).

An AAA domain can be confidently identified in region 128-292 of TnsC by searching major protein domain databases including Interpro (70), Pfam (46), SMART (79), and CDD (Figure 1.5a)(86) with the protein sequence. The Walker A motif of TnsC is located at position 136-144 and the Walker B motif at position 228-233, both of these regions are highly conserved throughout TnsC homologs (Figure 1.5b). Purified TnsC protein was shown to specifically bind adenine nucleotides (ATP, ADP, AMP, and non-hydrolyzable analogs) and to hydrolyze ATP at a moderate rate with Mg^{2+} (49, 139). ATP or Mg^{2+} seem to have profound effect on the conformation of TnsC: TnsC readily forms insoluble aggregates in low-salt solutions, but both ATP and Mg^{2+} were found to improve the solubility of TnsC and stabilize TnsC in solution (49).

ATP-bound TnsC is the active form

The ability of TnsC to bind DNA is a prerequisite for its ability to select a target site. TnsC is an ATP-dependent nonspecific DNA binding protein (49). No DNA binding is detected in the presence of ADP or AMP or in the absence of adenine nucleotide. However, non-hydrolyzable ATP analogues (AMP-PNP and ATP-γ-S) enhance TnsC DNA binding ability. Under standard *in vitro* transposition conditions, TnsABC+D transposition requires ATP (10). In addition, while no transposition is found with

TnsABC alone, TnsABC-mediated transposition will occur constitutively in the presence of AMP-PNP *in vitro* (i.e. without an specific targeting protein or structure)(10). Of further note, gain-of-function TnsC mutants that allow transposition in the absence of TnsD or TnsE have either a reduced ATPase activity or are able to bind DNA with an increased stability (Figure 1.5c)(139, 141). These findings indicate that the active form of TnsC is in an ATP-bound state and the target complex is able to help switch TnsC to its active form (10, 33, 83, 141).

TnsC is implicated in multiple interactions

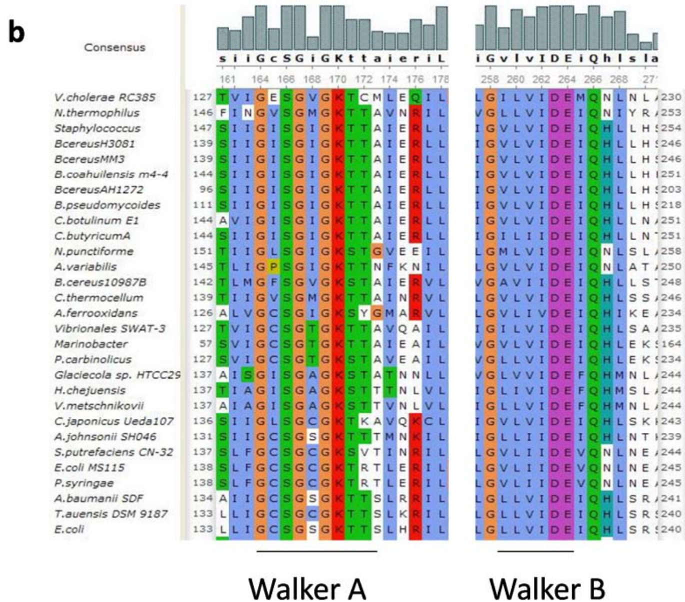
TnsC is believed to play a central regulatory role in transposition as the "matchmaker" or communicator protein conveying signals between the target selecting protein and the transposase. As expected, TnsC is involved in multiple protein-protein interactions in addition to its ability to bind DNA. The C-terminal portion of TnsC (TnsC 504-555) was found to directly interact with TnsA (83, 124, 139). The N-terminal domain (TnsC1-293) interacts with TnsD in a yeast two-hybrid assay (96). TnsC has also been shown to interact with TnsB (138), but the region responsible for this interaction is yet to be determined. Additionally, TnsC forms a homodimer (124, 139), but the dimerization interface in TnsC again has not been identified.

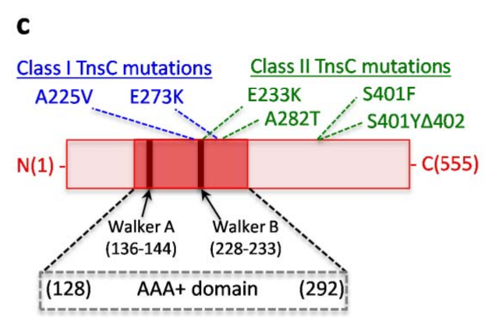
In theory there could be two regions in TnsC that bind DNA, one for the target DNA and another for the transposon ends. A run of positive residues between 495 and 504 in TnsC have been shown to have an effect on TnsAB-mediated donor DNA cleavage likely through binding the transposon ends (124). The C-terminus of TnsC also appears to have a known DNA binding motif. 388-407 region of TnsC may contain a Helix-Turn-Helix

Figure 1.5. **The nucleotide-binding motif in TnsC and TnsC gain-of-function mutations**. (a) Alignment of TnsC with other members of the AAA+ superfamily. Alignment was generated by CDD and catalytically important residues of Walker A motif (GxxxxGK[S/T], where x is any residue) and Walker B motif (hhhh[D/E], where h is a hydrophobic residue) are marked by #. (b) Nucleotide binding motifs are highly conserved in TnsC homologs. Protein sequences of TnsC homologs are named by the hosts of Tn7-like elements. Sequences are aligned by Muscle program in UGENE software with Clustal coloring scheme http://ugene.unipro.ru. (c) TnsC gain-of-function mutations that allow transposition independent of TnsD or TnsE. Class I mutations (A225V and E273K) are still responsive to signals presented by TnsD or TnsE but class II mutations (E233K, A282T, S401F, and S401YΔD402) are not. Walker A and Walker B motifs and the AAA+ domain are indicated with the positions relative to TnsC from Tn7.

a

```
                            #   #  ###                                                 ##                                  #
HslU      17 .[34].NILMIGPTGVGKTEIARRL.[ 7].FIKVEA.[168].IVFIDE.[20].QRDLLPLVE.[16].LFIASGA 308
TnsC      97 .[34].SLLLIGCSGSGKTTSLHRI.[19].YLKIDC.[ 53].LLVIDE.[ 8].SGGSQEMLN.[11].PVMLIGT 268
Rfc       28 .[20].HLLFAGPPGVGKTTAALAL.[12].FLELNA.[ 28].IIFLDE.[ 8].QQALRRTME.[ 6].RFILSCN 148
PspF      13 .[18].PVLIIGERGTGKELIASRL.[10].FISLNC.[ 37].TLFLDE.[ 8].QEKLLRVIE.[17].RLVCATN 149
CfxQ      60 .[20].HMAFTGNPGTGKTTVALKM.[14].LVSVTR.[ 25].VLFIDE.[ 8].NERDYGQEA.[15].VVILAGY 188
NSF      221 .[34].GILLYGPPGCGKTLLARQI.[ 8].PKVVNG.[ 36].IIIFDE.[17].HDTVVNQLL.[12].LVIGMTN 374
PIM1     607 .[20].IICFVGPPGVGKTSIGKSI.[ 7].FFRFSV.[ 36].LILIDE.[26].NSFLDNYLD.[ 6].KVLFVCT 748
NtrX     146 .[18].RILIVGPSGSGKELTARMI.[10].FVVINA.[ 34].TLFLDE.[ 8].QNRVLRVLV.[17].RIISSTG 279
CplL     477 .[29].SFLFVGPTGVGKTELSQAL.[10].LFGIDM.[ 39].VILLDE.[ 4].DPQVLTLLL.[21].IIIMTSN 626
MJECL26    8 .[20].ILFVYGPKSSGKSTVMRRV.[10].FFYYNL.[ 71].VLIIDE.[ 9].NGGKSLLNE.[15].HVICLTS 179
```





DNA binding motif (UniProtKB/Swiss-Prot accession P05846) (31) which could be responsible for target DNA binding.

It is interesting to note that the C-terminus of TnsC is poorly conserved but the N-terminus is almost identical amongst TnsC homologs (Z. Li & J.E. Peters, unpublished observation). This could indicate that interactions in the C-terminus are more structural in nature, or evolving more quickly.

Clues to the molecular mechanism of transposition activation may come from a deletion mutant of TnsC, $TnsC^{\Delta 1\text{-}293}$, which is active in promoting transposition with TnsAB, but is not responsive to positive signals presented by TnsD or TnsE (141). This together with other findings discussed above indicates that the C-terminal domain of TnsC is for activating TnsAB, but that the N-terminal domain is functionally interacting with the target complex**.** One model that accommodates the data is that the N-terminal domain keeps TnsC in an “off” state by masking the C-terminal domain of TnsC; conformational changes induced by ATP binding (ATP binding site in the N-terminal domain) could then reveal the C-terminal domain and allow it to activate the transposase.Crystal structures of this protein alone and/or in combination with its interacting proteins will be important for a detailed understanding of all these protein-protein interactions.

Toggling between the ATP-bound and ADP-bound state (via hydrolysis) would be a mechanism for TnsC to switch between an active form that can stimulate recombination and an inactive form for transposition to produce a target searching behavior for TnsC (142). In theory stimulating the hydrolysis activity of TnsC would drive TnsC molecules

into the "off state" while inhibiting hydrolysis or stimulating nucleotide exchange could drive molecules into the "on state". Such a mechanism is used in bacteriophage Mu transposition where the ATP-bound MuB protein can bind the target DNA and deliver it to the MuA transposase but ADP-bound MuB cannot (3, 131, 155). An appealing idea is that target complex may hold TnsC in the ATP-bound state. This idea is supported by the increased ATP binding observed when TnsC is incubated with TnsD and *attTn7*. By contrast, TnsB may stimulate the ATPase activity of TnsC resulting in the ADP-bound conformation of TnsC that cannot bind the target (33, 142).

It is interesting to note that the two classes of TnsC mutations that allow TnsABC core machinery to work with little target specificity appear to be present in naturally occurring TnsC homologs (106). It would be of interest to test if these mutations could account for the existence of Tn7-like elements in locations other than the *attTn7* sites in the chromosome.

1.2. Target selection in the TnsABC+D Transposition Pathway

TnsD is a sequence-specific DNA binding protein that mediates Tn7 transposition into a specific site, called *attTn7*, on the chromosome at high frequency (Figure 1.6). The binding sequence recognized by TnsD is within the 3'-end of the glucosamine-6-phosphate synthase (*glmS*) gene, which is about 25 bp upstream of the actual insertion point in the transcription termination region (10, 151). Insertion into the *attTn7* site has no detectable negative effect on the host (51). Therefore, the *attTn7* site is considered as the "safe haven" for Tn7 propagation with host DNA replication (34, 110). While probably not relevant in the environment, when the *attTn7* site is not available, in the

laboratory TnsD can direct Tn7 transposition at a very low frequency into so-called pseudo-*attTn7* sites that show homology to *attTn7* (74).

1.2.1. Sequence requirements of attTn7

Within the *attTn7* site region, the region actually bound by TnsD is the only critical determinant for target activity; the sequence of the actual insertion point can be varied without significant effect (55). Within *attTn7*, TnsD-binding sequence is located at +22 to +55 (relative to the point of insertion) (Figure 1.6)(10, 92, 151), which actually encodes the active site of the GlmS protein (Glucosamine-fructose-6-phosphate aminotransferase) involved in cell wall synthesis (33, 96). GlmS analogs have been found in organisms from bacteria, archae, and eukaryotes. The TnsD-binding sequence is highly conserved in *glmS* genes where variation is limited to the "wobble" positions of each codon (75, 96). Changing any of the conserved sites to the opposite type, i.e purine to pyrimidine or vice versa, affected the binding affinity of TnsD and consequently the frequency of insertion adjacent to the TnsD binding site (96). Modeling of the target DNA sequence as a B-form DNA revealed that all of the sites that had significant effects on TnsD binding and target activity (+31, +33, +42, +43, +45, +51 and +54) are on one face of the DNA, consistent with the model that TnsD binds the major groove of the target DNA (33, 76). Several nucleotides on the top strand that caused more than 100-fold decrease in transposition when mutated, were found to be in close contact with TnsD protein (96), indicating an important role of these sites for TnsD binding.

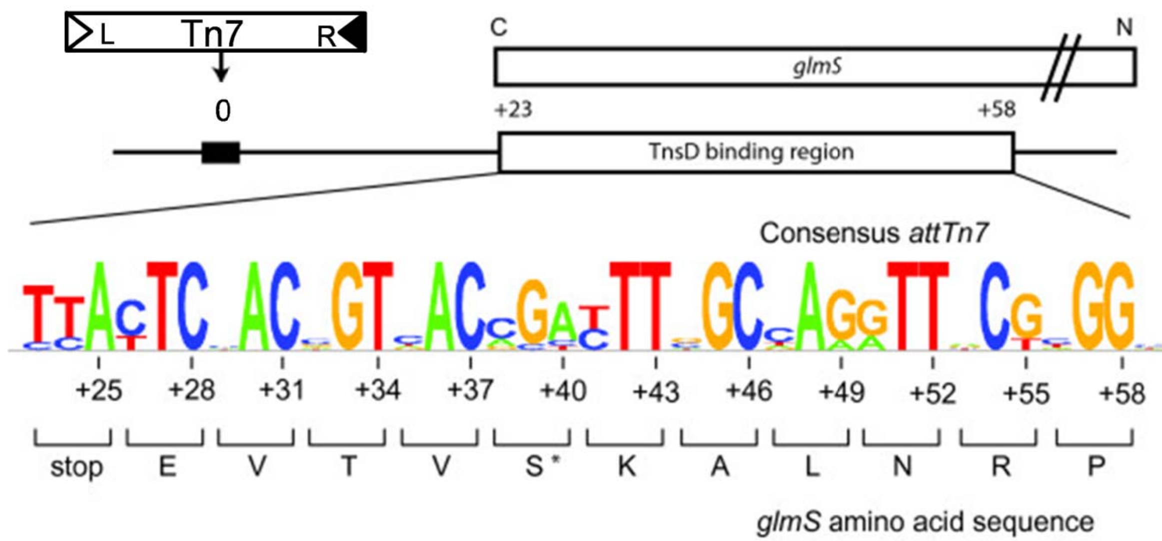
Consistent with the sequence requirement for TnsD-mediated transposition, Tn7 elements are almost exclusively found at the *attTn7* sites in the chromosomes of

divergent bacterial hosts (34, 106, 107). TnsD has also been shown to bind the human *glmS* homologs *gfpt-1* and *gfpt-2* (75), the *Drosophila gfat-1* and *gfat-2*, and the zebrafish *gfpt-1* (96). *In vitro* transposition into human *attTn7* homologs was found to be correlated with the binding affinity of TnsD to these sequences, but transposition decreased when this DNA was assembled into nucleosomes *in vitro* (75). How the dynamic structure of eukaryotic chromatin might affect Tn7 targeting is yet to be tested. When studied in *E. coli*, Tn7 is able to recognize human *gfpt* sequence as an active target for transposition in a site-and-orientation specific manner (28).

1.2.2. Target DNA binding by TnsD

Sequence alignment of TnsD homologs revealed a highly conserved N-terminal region (N1-170) containing a $C^{124}C^{127}C^{152}H^{155}$ motif characteristic of a zinc finger (96). Evidence that this is a zinc-finger like DNA binding motif came from the finding that TnsD binding to *attTn7* is diminished when Zn^{2+} in the buffer is specifically chelated away by 1,10-phenanthroline. In addition, changing any residue of the CCCH motif into a serine abolishes TnsD binding *in vitro*and also results in about 1,000-fold decrease in TnsABC+D transposition into *attTn7 in vivo*. Interestingly, however, the TnsD N-terminus alone does not bind *attTn7*; rather almost the entire protein appears to be required for DNA binding. A screen for missense mutations with a dominant-negative phenotype (i.e. null mutants that prevented transposition activity when in combination with the wild-type protein) in TnsD revealed many mutations over the entire length of TnsD that all affected DNA binding. These results indicate that there are more elements

Figure 1.6. **Organization of *attTn7* and its sequence conservation**. Within *attTn7*, the central base of the 5 bp sequence duplicated upon Tn7 insertion is designated as position "0" and base pairs towards the *glmS* gene is in positive numbers. TnsD binding sequence (+23-+55) is in the open reading frame of *glmS* gene and consists of the region that encodes the active site of GlmS. Conservation of the TnsD binding sequence is shown in the sequence logo. Figure is modified from a previous publication (96).


L
Tn7
R
0
C
N
glmS
+23
+58
TnsD binding region
Consensus attTn7
+25
+28
+31
+34
+37
+40
+43
+46
+49
+52
+55
+58
stop
E
V
T
V
S *
K
A
L
N
R
P
glmS amino acid sequence

important for DNA binding spread across the protein, in addition to the CCCH zinc binding motif (96).

TnsD binding to *attTn7* target sequence shows a core region of protection, extending from +30 to +55 and the interaction is primarily within the major groove of the *attTn7* DNA. A second, albeit weaker, protected region is located +22 to +30 where a DNA distortion is imposed upon TnsD-binding. In the TnsC-TnsD-*attTn7* complex the distorted region is covered and the protection extends past the insertion site to position -15 (76). More importantly, in the TnsC-TnsD-*attTn7* complex TnsC occupies the minor groove of DNA at the insertion site, which leaves the major groove accessible for the transposase TnsAB to act (76), consistent with the 5 bp staggered joining of the transposon ends to the target DNA (which would predict that the transposase acts on the major groove). As discussed below, the TnsD-induced distortion is actually the signal that attracts TnsC, whereas interaction between TnsC and TnsD may be important for activation of TnsC.

1.2.3. TnsD interaction with TnsC

TnsD interacts with TnsC in a yeast two-hybrid analysis where, presumably, a solution interaction is occurring as opposed to an interaction on the yeast chromosome at the site homologous to *attTn7* (to which TnsD does not bind (75)) . C-terminal truncated TnsD (TnsD 1-309) binds TnsC better than full-length TnsD, suggesting some level of regulation in the reaction between TnsD and TnsC. The investigators were unable to define a smaller interaction domain; a greater truncation of TnsD (TnsD 1-293) abolishes its ability to bind TnsC and deleting several residues from the N-terminus of TnsD also

results in a polypeptide unable to bind TnsC. Together these results suggest that the domain for TnsD to interact with TnsC is located in the N-terminus.

1.2.4. Host factors in TnsD-mediated transposition

It is also important to note that host factors also participate in TnsD-mediated transposition (133). Two host factors, L29, a component of the 50S ribosomal subunit, and ACP, acyl carrier protein essential for fatty-acid biosynthesis, were found to stimulate TnsD binding to the target DNA *in vitro*. The two proteins function together to enhance the apparent affinity of TnsD for *attTn7* by more than 20-fold. ACP and L29 also stimulate TnsABCD transposition more than 3-fold *in vitro*, despite the TnsABC+D *in vitro* transposition system has been highly optimized. More importantly, one of these host factors seems to be critical for TnsABC+D transposition *in vivo*; mutating L29 specifically decreased TnsABC+D transposition by more than 100-fold. ACP could not be tested as easily becasue of its essential role in the cell. Given the important role of L29 and ACP in cellular metabolism, this may provide a way of regulating TnsD-mediated transposition according to the cellular conditions.

1.3. Target Selection in the TnsABC+E Transposition Pathway

While TnsABC+D pathway is exceptional at choosing a specific site for efficient transposition via a DNA sequence, TnsE-mediated transposition targets sites without sequence similarity (33, 110, 152). Instead, TnsE recognizes a complex that is frequently available in discontinuous DNA replication (Figure 1.7)(105). TnsABC+E pathway preferentially inserts into actively conjugating plasmids (154), and filamentous phage M13 (45), thereby facilitating horizontal transfer of Tn7 amongst bacteria.

1.3.1. TnsE-mediated transposition targets lagging-strand DNA replication

TnsE-mediated transposition is stimulated many-fold in the presence of a conjugating plasmid. In addition, these insertions preferentially target the conjugal plasmid; greater than 90% of transposition events occur into the plasmid, even though the plasmid only represents less than 5% of the cellular DNA (154). Early work indicated that transposition events occurred in one orientation in conjugal plasmids and it was later shown that these transposition events occurred as the element entered a new bacterial host (14-16, 154)(Figure 1.7).

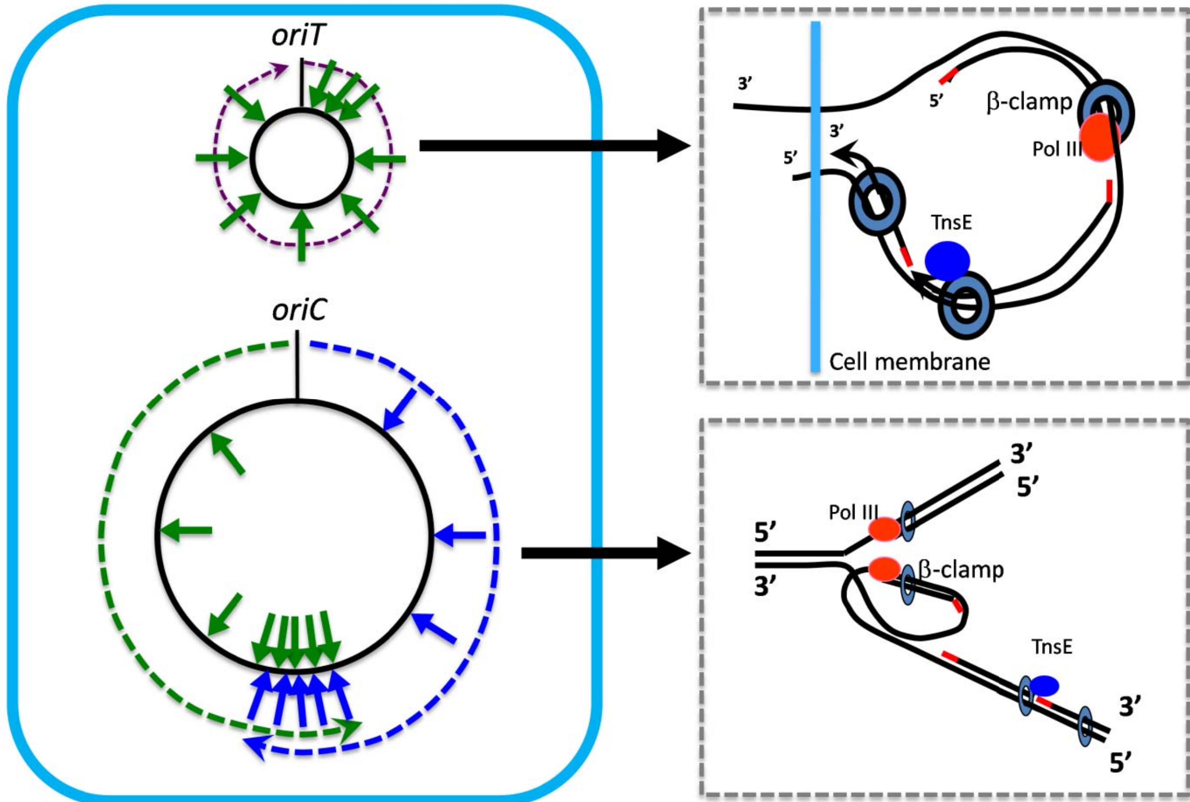
At a very low frequency, TnsE-mediated transposition also targets the chromosome, and Tn7 insertions in the chromosome occur in opposite orientations in the two replichores of the circular chromosome (Figure 1.7)(108, 109). The *E. coli* chromosome is replicated using a single origin (*oriC*), which initiates replication in both directions. Replication forks meet at a position generally equidistant from the origin. The two regions of the chromosome that are replicated by separate replisomes are called replichores (134). Interestingly when TnsE-mediated insertions in the chromosome were mapped, a striking pattern emerged. While there was a clear bias for the region where DNA replication terminates, it was also found that all of the insertions occurred in the same orientation within each replichore (An idealized distribution is shown in Figure 1.7). Together with the orientation bias found with insertions into conjugal plasmids in the recipient cells, it could be concluded that TnsE-mediated transposition occurs in one orientation with the DNA strand undergoing lagging-strand replication (Figure 1.7)(108, 110). Interestingly, transposons unrelated to Tn7, Tn917 and IS903, also have been shown to preferentially

target where DNA replication terminates (135, 147). However, because an orientation bias cannot be established with Tn917 and IS903 it is unclear if these elements also share an attraction for lagging-strand DNA replication.

1.3.2. Factors important for TnsE targeting of lagging-strand DNA replication

What feature of lagging-strand DNA synthesis attracts TnsE-mediated transposition events? TnsE is a DNA-binding protein which has a specific preference for DNA structures with 3'-recessed ends. TnsE mutants that promote higher transposition frequency (up to a 1000-fold increase) *in vivo* bind structures with 3'-recessed ends much better than the wild type protein (108), indicating that the ability of TnsE to bind 3'-recessed DNA is involved in target site selection. A specific affinity for 3'-recessed DNA ends was intriguing because DNA ends would be expected to be overrepresented during discontinuous DNA synthesis, whereas recessed ends would not generally be expected during leading-strand synthesis. Moreover, when progress of the lagging-strand DNA synthesis is blocked by various impediments, the DNA polymerase would simply abandon the ongoing DNA synthesis, but take on primers downstream to resume DNA replication (65, 90, 94). DNA structures with 3'-recessed end or ssDNA gap would be generated in these processes. The possibility of DNA structures with 3'-recessed ends as a bio-active target for TnsE-mediated transposition was directly tested *in vitro* (105). The *in vitro* system was set up with purified proteins, a donor plasmid (containing a miniTn7 element but with a conditional replication origin) and a gapped DNA structure on a circular plasmid to simulate structures with 3'-recessed ends found *in vivo*. This system

Figure 1.7. **TnsE-mediated transposition targets lagging-strand DNA replication for transposition**. Tn7 transposition events mediated by TnsE into both conjugal plasmids (upper left) and the chromosome (bottom left) are all oriented in one left-to-right orientation with respect to the lagging-strand DNA replication of the relevant replicon (right panels). Within each DNA, TnsE-mediated transposition has a preference for insertion into the leading region during conjugation and into the region where DNA replication terminates in the chromosome. Two factors important for this targeting are a 3'-recessed ends and the β clamp processivity factor. On the left panel, each arrow represents an idealized distribution of TnsE-mediated transposition event and blue and green arrows indicate opposite orientations. Position of *oriT* is labeled on the circular plasmid DNA showing the direction of DNA replication of the element as a purple dashed line. On the circular chromosome, bidirectional replication from *oriC* origin is indicated as blue and green dashed lines for each repichore (see text). Schematic representation of the DNA replication of the conjugal plasmid in the recipient cell (upper) and the chromosome (lower) is shown in the right, with all of the important proteins labeled (see text).


oriT
oriC
3’
5’
β-clamp
3’
5’
Pol III
TnsE
Cell membrane
3’
5’
Pol III
5’
3’
β-clamp
TnsE
3’
5’

allowed monitoring of the *in vitro* transposition events by *in vivo* output from transformation of the deproteinized *in vitro* reaction mix. Tn7 transposition was found to occur only with a gapped DNA structure and transposition depended on TnsE. The fact that no transposition was detectable with nicked DNA substrate is consistent with the model that TnsE targeting requires 3'-recessed ends. However, transposition events occurred randomly all over the gapped plasmid in both orientations, indicating that something else was involved in determining the polarity of a potential target in addition to the 5'-3' polarity of DNA.

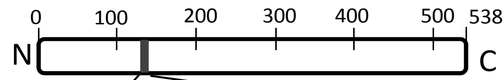
TnsE interaction with the β clamp processivity factor of DNA polymerase III holoenzyme proved to be the other important factor for TnsE targeting. β clamp confers DNA replication processivity by tethering DNA polymerases to the template DNA. Like the 3'-recessed DNA structures or gaps on DNA, β clamp is another factor that accumulates on the DNA strand that undergoes lagging-strand DNA synthesis (143, 157). For leading-strand DNA synthesis, in theory, the β clamp only has to be loaded once and the DNA replicase will be able to finish replicating the full length of the replichore with the same processivity factor. By contrast, a new β clamp is loaded for every Okazaki fragment during lagging-strand DNA synthesis. While β clamps are eventually recycled in the process, there is likely a period of time when these clamps are available to interact with other proteins. In fact, many proteins involved in processing the newly-synthesized Okazaki fragments have been shown to use the clamps left behind as a mobile platform for their various functions (71, 82). Proteins suggested or shown to interact with the β clamp include DNA polymerase I (which is involved in removing the 5'-primer of a

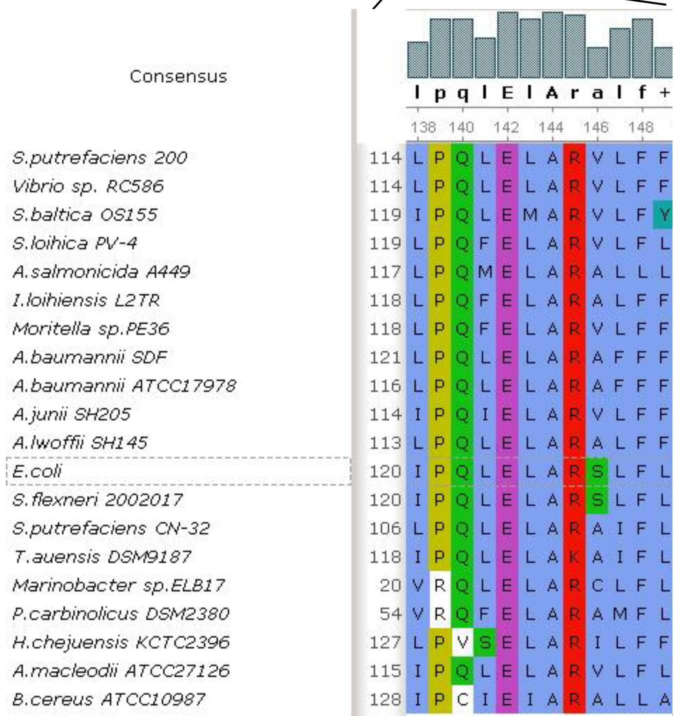
previous Okazaki fragment while filling the gap between two Okazaki fragments), DNA ligase (which seals two adjacent Okazaki fragments), and MutS and MutL (which may use the clamp to differentiate nascent DNA strand for mismatch repair). These and other β-binding proteins all interact with the asymmetric ring-shaped β clamp primarily via a β-binding motif that binds specifically to a hydrophobic pocket at the C-terminal face of the β ring (39). TnsE also contains such a β-binding motif at the N-terminus, which is found in a highly conserved region in the TnsE homologs (Figure 1.8). Substitution any of the conserved residues with alanine residues reduces or abolishes TnsE-mediated transposition *in vivo*. A weak interaction between TnsE and the β clamp was detected by yeast two-hybrid assay, protein mobility-shift assay, and far-western assay, and quantified by Surface Plasmon Resonance (SPR) analysis. The β-binding motif is important for this interaction, as judged from the observation that when TnsE proteins with the alanine replacements in the β clamp-binding motif were used in the assays the interaction sharply decreased. It is interesting to note that the TnsE-β clamp interaction does not appear to interrupt the normal traffic on the β clamp, as only when extremely high levels of TnsE are induced do you find an SOS induction in the cell (105). Under these conditions of very high TnsE overexpression it could be shown that SOS was induced because of the interaction with the β clamp. Proof of an important role for the β clamp came from the TnsE-reconstituted transposition system. While gapped DNA substrates proved to be productive targets for *in vitro* TnsE-mediated transposition, unexpectedly these insertions occurred in two orientations. However, when the gapped substrate was populated with the β clamps a striking change in the transposition profile

was observed. In the substrates populated with the β clamps, transposition events were found to occur in predominantly only one orientation, with the right end of Tn7 close to the 3'-end of the single-strand gap as occurs *in vivo*. In addition, a hotspot of insertion was observed on the gapped plasmid, recapitulating what was observed *in vivo* with insertions on the chromosome (105). The successful reconstitution of an *in vitro* transposition system indicates the primary target of TnsE-mediated transposition is the complex of the β clamp at the primer/template junction. Exactly how TnsE gains access to the clamp *in vivo* and where it binds on the clamp are intriguing questions to be answered with further research. TnsE may need more signals from the replisome to target DNA replication given the finding that *in vitro* TnsE-mediated transposition was only found when strong gain-of-activity mutants of TnsE were used in the assay.

The fact that lagging-strand DNA synthesis during conjugation is the most preferred target for TnsE-mediated transposition may indicate that the uncoupled DNA replication fork is more accessible to TnsE (Figure 1.7). For chromosomal DNA replication, the replicases for leading-strand and lagging-strand DNA synthesis are coupled by the τ subunit of the DNA polymerase III holoenzyme (89, 103). However, such an elaborate coordination cannot be expected for leading-strand and lagging-strand DNA synthesis of the conjugating plasmid, since they occur in different cells, the donor cell and the recipient cell, respectively. In an uncoupled DNA replication fork, the DNA replicase may dissociate more frequently from the β-clamp; thereby allowing the gain of access of TnsE.

Figure 1.8. **β clamp binding motif in TnsE homologs**. TnsE protein found in *E. coli* is represented as a bar on the top with the amino (N) and carboxy (C) termini indicated. TnsE homologs from hosts indicated are aligned with MUSCLE program in UGENE software (http://ugene.unipro.ru) and the region between residues 120 and 131 that contains the β clamp binding motif is shown. Sequence conservation and the consensus sequences are shown above the alignment. Underlined residues are changed individually or in combination to study their importance in β clamp binding.


0
100
200
300
400
500
538
N
C


Consensus
l p q l E l A r a l f +
138 140 142 144 146 148
S.putrefaciens 200
Vibrio sp. RC586
S.baltica OS155
S.loihica PV-4
A.salmonicida A449
I.loihiensis L2TR
Moritella sp.PE36
A.baumannii SDF
A.baumannii ATCC17978
A.junii SH205
A.lwoffii SH145
E.coli
S.flexneri 2002017
S.putrefaciens CN-32
T.auensis DSM9187
Marinobacter sp.ELB17
P.carbinolicus DSM2380
H.chejuensis KCTC2396
A.macleodii ATCC27126
B.cereus ATCC10987
114 L P Q L E L A R V L F F
114 L P Q L E L A R V L F F
119 I P Q L E M A R V L F Y
119 L P Q F E L A R V L F L
117 L P Q M E L A R A L L L
118 L P Q F E L A R A L F F
118 L P Q F E L A R V L F F
121 L P Q L E L A R A F F F
116 L P Q L E L A R A F F F
114 I P Q I E L A R V L F F
113 L P Q L E L A R A L F F
120 I P Q L E L A R S L F L
120 I P Q L E L A R S L F L
106 L P Q L E L A R A I F L
118 I P Q L E L A K A I F L
20 V R Q L E L A R C L F L
54 V R Q F E L A R A M F L
127 L P V S E L A R I L F F
115 I P Q L E L A R V L F L
128 I P C I E I A R A L L A

TnsE-mediated transposition into both conjugal plasmids and the *E. coli* chromosome also display a preference for specific regions, the leading region and the terminus region, respectively (Figure 1.7). This could be explained by the persistence of gap structures in these regions during the final stage of replication (Figure 1.9). Chromosomal DNA replication terminates upon the convergence of the replication forks in the terminus region, a process facilitated by the many *ter* (termination) sites in this region (42, 66). Circular duplex DNA with a short gap in the terminus region of the nascent strand has been observed after resolution of the converging forks (2, 23, 62, 145) and DNA polymerase I-mediated repair-like processes are implicated inclosing the gaps (2, 23, 87). Such a gap structure may also be generated by re-circularizing the transferred strand upon the completion of conjugation (153). TnsE could efficiently use such structures by directly competing with DNA polymerase I.

1.3.3. Other DNA replication processes targeted by TnsE

TnsE is able to target replicating filamentous bacteriophages for transposition (45). Insertion into the bacteriophage M13 also occurred almost exclusively in a single left-to-right orientation with respect to the replication of the minus strand of the phage genome, which would presumably occur in a leading-strand-like manner. Its targeting by TnsE-mediated transposition suggests that it is somewhat different from the replication of the conjugal plasmid in the donor cell.

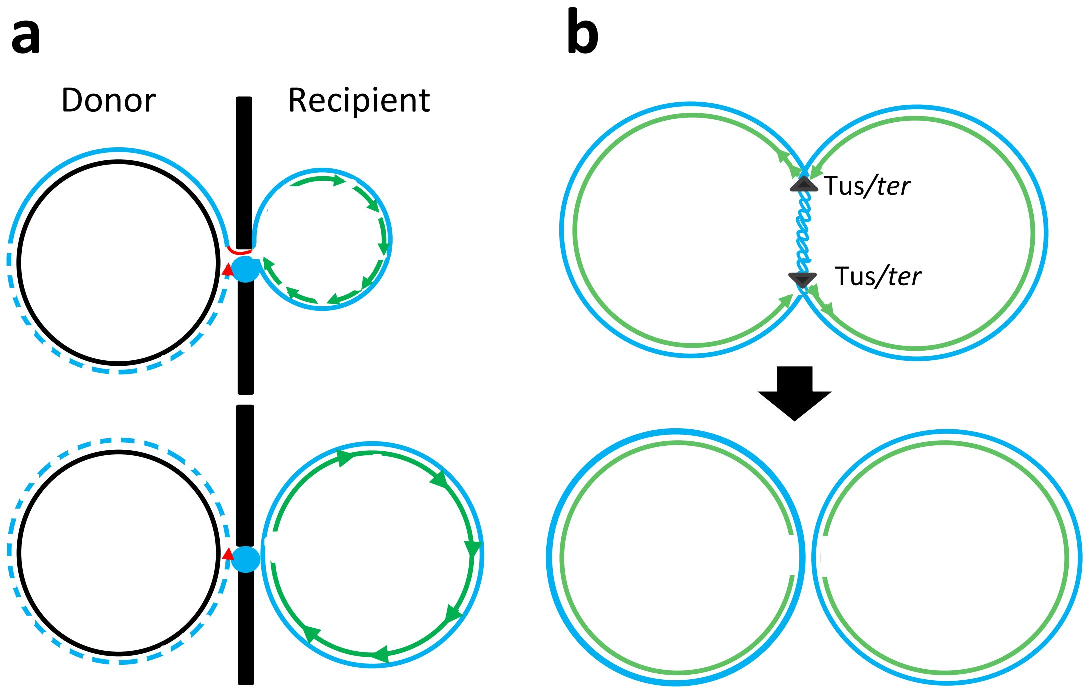
TnsE-mediated transposition is stimulated by induction of double strand breaks (DSBs) in the chromosome (109). This is true for the direct induction of DSBs in the chromosome or with DSBs associated with various treatments, like exposure to UV light,

mitomycin C, or phleomycin (136). Interestingly, following the induction of a single DSB in the *E. coli* chromosome, Tn7 insertions do not occur immediately at the site of the DSB, but instead across a region proximal to the point of DNA break (109). Of further interest is that transposition occurs at a series of hotspots hundreds of kilobases away from the original break site (136). The orientation of transposition events in this system suggests that TnsE targets the repair-associated replication. The absence of such hotspots in cells without DSB induction suggests that the repair-associated replication machinery is different from that of chromosomal DNA replication and may be particularly vulnerable for DNA replication roadblocks. Further studies, which would have implications for the mechanism of replication-mediated repair, are needed to test this model.

1.3.4. Why target lagging-strand DNA replication?

Lagging-strand DNA synthesis has subsequently been found to be sensitive to multiple forms of recombination. Transposons of the IS200/IS605 family and Group II introns (101), which mobilize by very different mechanisms, and bacteriophage lambda RED recombination (32), all seem to prefer DNA undergoing lagging-strand DNA replication as their target for recombination. These observations suggest that recombination into the DNA strand undergoing lagging-strand DNA replication may be advantageous or that DNA undergoing lagging-strand replication may naturally be more vulnerable to recombination. Support for the latter idea comes from the finding that in many bacterial species small oligonucleotides can be integrated into the lagging-stand even in the absence of recombination proteins (146).

Figure 1.9. **Persistence of gap structures may lead to regional preference of TnsE-mediated transposition into both plasmid and chromosome**. (a) Plasmid conjugation and concurrent DNA synthesis in the donor and recipient cells. The direction of DNA replication is represented by the arrow heads. The transferred strand is indicated by the blue circle and the nascent strand is indicated by the dashed lines in blue in the donor and green recipient cells, respectively. The relaxasome that mediates plasmid conjugation is shown as the blue oval. Upon the completion of conjugation, the transferred strand is re-circularized by the relaxasome and a gap structure would be generated as indicated in the figure. (b) Resolution of the converging forks results in short gaps in the nascent strands of daughter chromosomes. The template strands are indicated by the blue circles and the nascent strands are indicated by the green circles, with arrow heads indicating the direction of replication. The terminus region is indicated by the Tus/*ter* complexes.


a
Donor
Recipient
b
Tus/ter
Tus/ter

1.4. Recognition of a Positive Target Signal by TnsC

How does TnsC recognize a potential target for transposition? Why does insertion occur into some sites but not into others? The fact that virtually no transposition occurs with only the core TnsABC machinery *in vivo* indicates an important role of TnsD or TnsE in activating this core machinery and probably recruiting the core machinery to the potential target site. However, isolation of TnsC gain-of-function mutants that allow transposition in the absence of TnsD and TnsE indicates that the role of the target site selecting proteins is linked through the transposition regulator protein TnsC (141). As discussed before, gain-of-function mutants of TnsC have either slower ATPase activity or altered DNA binding ability (139), both of which may keep TnsC at the active state.

Important insight into how TnsC locates a potential target came from work with the mutant core machinery, $TnsABC^{A225V}$, where the mutant $TnsC^{A225V}$ is used in combination with wild type transposase TnsAB. In *in vitro* reactions using duplex DNA targets, transposition events promoted by the mutant core machinery occurred in both orientations with very low target site specificity; transposition events dispersed throughout the target plasmid and no apparent sequence similarity was found amongst the insertion sites or sequences flanking the target sites (19, 116, 117). This picture was completely changed when transposition reaction were performed with a plasmid to which was attached a pyrimidine triplex-forming DNA oligonucleotide (85, 116, 117). More than 70% of the transposition events examined located at a specific small region near the triplex region. Of additional interest, 98% of the insertions that occurred in the hotspot immediately adjacent to the triplex occurred in one orientation with the right end of Tn7

proximal to the triplex. This site- and orientation-specificity of TnsABC^{A225V}-promoted transposition into a triplex-containing DNA substrate is comparable to that observed with the TnsABC+D pathway (Figure 1.10). Interestingly, TnsC was also found to bind preferentially to triplex containing DNAs (116, 117), similar to the specific binding of TnsC to the TnsD-*attTn7* complex. As discussed above, binding of TnsD imposes a distortion in a region of the *attTn7* target DNA. Annealing of a triplex forming oligomer to its duplex target DNA also results in conformational changes at the triplex-duplex junction. The ability of TnsC to bind both triplex-containing DNA and TnsD-*attTn7* target complex and the relative position of Tn7 insertions to the DNA distortions on both kinds of DNA substrates indicate that a DNA-based signal may play a central role in target site recognition by TnsC (Figure 1.10).

It is important to note that DNA distortion alone is not sufficient to activate wild type TnsC. A triplex DNA substrate could not stimulate wild type TnsABC-mediated transposition. This, together with the presence of TnsD in the post-transposition complex (see below), indicates that the role of TnsD is not simply an “assembly” factor that helps recruit TnsC by inducing a conformational change on DNA. Direct protein-protein interaction between TnsC and TnsD may be important for activating TnsC. The two signals, TnsC-TnsD and TnsC-distortions interactions may help make sure that activation occurs only in the targeting complex (Figure 1.10).

How TnsC is attracted to the target complex in TnsE-mediated transposition pathway is not yet clear. Given that TnsABC+D and TnsABC+E pathways of Tn7 transposition share the same core machinery, one would expect that TnsC would recognize the same

kind of signals to locate potential targets in TnsE-mediated transposition (Figure 1.10). It remains to be determined if TnsE binding to target DNA induces conformational change. However, there seems to be evidence for TnsE and TnsC functional interaction in controlling the target selection in TnsE-mediated transposition (Q. Shi & J. E. Peters, unpublished data).

1.5. Target Immunity

Not only does Tn7 transposition respond to positive signals presented by the target site selecting proteins TnsD and TnsE, Tn7 transposition is also sensitive to a process called target immunity in both the TnsD- and TnsE-mediated pathways, in which the presence of a transposon copy in a target DNA inhibits further insertion of the element into that target DNA (6, 59). A similar process occurs with other transposable elements, such as Tn3-like elements (53, 54, 78, 100, 123) and bacteriophage Mu (119). Target immunity with Tn7 was found *in vivo* both on a plasmid or chromosome and has also been established *in vitro* (6, 9, 41, 137, 140). The degree of inhibition seems inversely related to the distance between a potential target site and the preexisting copy of Tn7 (6, 41), indicating that immunity is a local and not a global effect. A small (60 kb) derivative of the *E. coli* F plasmid was rendered inactive for transposition when the plasmid contains Tn7 end sequences (6).

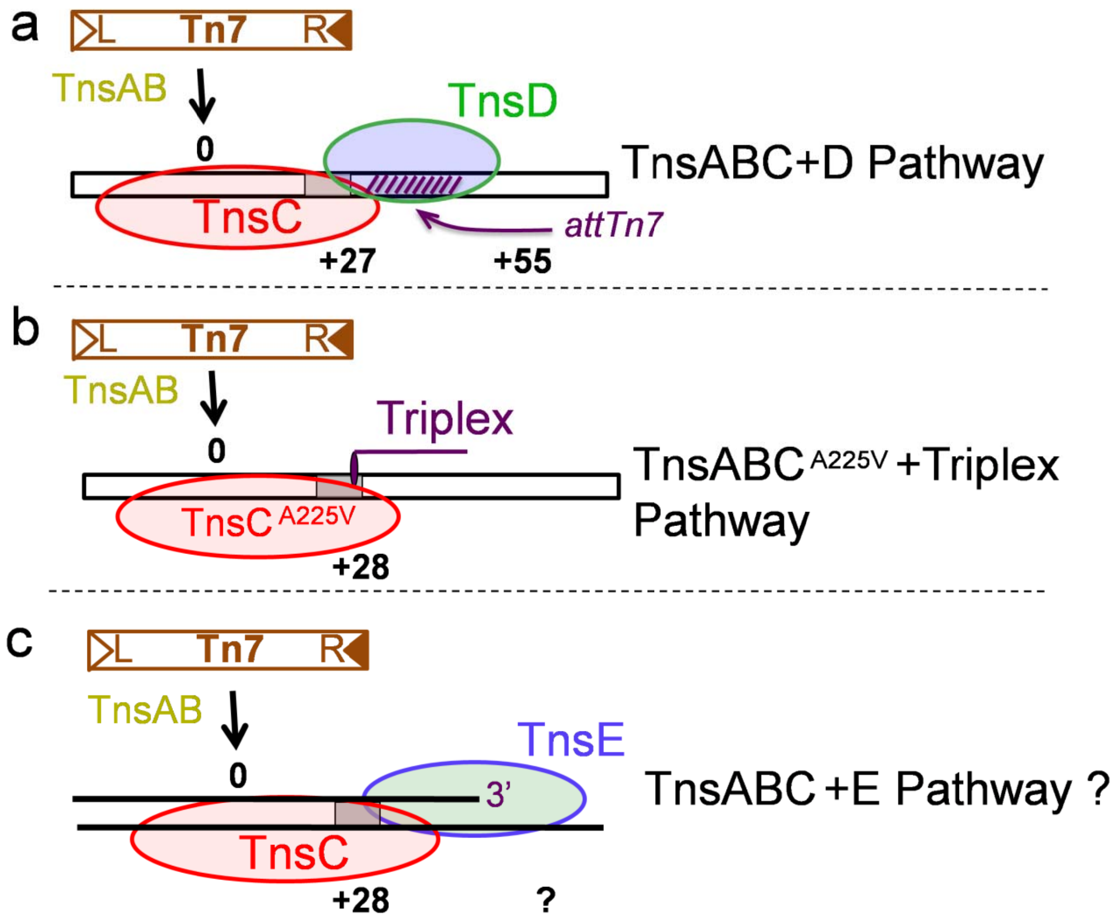
The basis of the immunity signal is the TnsB binding sites found in the ends of Tn7 (140): both the right end and left end of Tn7 can impose immunity on a target DNA and a single TnsB binding site is effective at imposing immunity when studied *in vitro*, indicating that the other features of Tn7 ends are not important for target immunity. It is

also worth noting that not all TnsB-binding sites are equally effective in imposing target immunity, which is probably due to the different binding affinities of TnsB to these sites, as Tn7R41 containing the weakest TnsB-binding site ω does not confer target site immunity *in vivo* (6) or *in vitro* (9).

The establishment of target immunity appears to depend on a TnsB-TnsC interaction that triggers displacement of TnsC from Tn7-containing targets, a process accompanied by the increased ATPase activity of TnsC. Without hydrolysis, immunity is bypassed and Tn7 readily inserts into target DNA with a preexisting copy of Tn7 (140). In TnsABC+D transposition *in vitro*, target immunity was completely abolished when non-hydrolyzable ATP analogues were used (10). Conceivably, the loss of target immunity observed with the $TnsC^{E233K}$ mutant is also because of the loss of ATPase activity, as the mutation is at a critical residue of the Walker B motif and mutating this residue has been shown to abolish the ATPase activity in other ATPases (57).

The ability of Tn7 to evaluate potential targets based on their distance from a pre-existing copy likely provides another mechanism that facilitates Tn7 survival and dispersion. This is because target immunity will prevent potentially hazardous events, such as insertion of Tn7 into itself or at close vicinity, which will disrupt the element directly or as a result of homologous recombination, but favor the dissemination of elements to distant sites within a DNA molecule or to different DNA molecules. While discouraging transposition within or near the element is a benefit of target site immunity for all transposons, there is yet another benefit for target site immunity for Tn7. Because very high efficiency transposition is found with insertions into the *attTn7* site in the

Figure 1.10. **Activation of TnsC by various target complexes**. Comparison of TnsD-mediated transposition with those of TnsABC mutant machinery indicates that TnsC recognizes a distortion on the target DNA. Both TnsD binding to *attTn7* and formation of triplex DNA induce conformational change in the duplex DNA, which is indicated by the shaded box. The relative position and orientation of Tn7 insertions to the DNA distortion found in both TnsABC+D and TnsABC^{A225V} transposition system emphasize the critical role of target DNA structure in Tn7 transposition. Also, it appears that protein-protein interaction of TnsC N-terminus with TnsD is important for activating TnsC. Since TnsABC+E pathway shares the core machinery with the TnsABC+D pathway, it is likely that the same mechanism is involved in TnsABC+E transposition. Interaction of TnsC with the target selecting protein(s) is indicated by the overlapping of the representing circles.



a
L
Tn7
R
TnsAB
0
TnsD
TnsABC+D Pathway
TnsC
attTn7
+27
+55
b
L
Tn7
R
TnsAB
Triplex
0
TnsABC A225V +Triplex
Pathway
TnsC A225V
+28
c
L
Tn7
R
TnsAB
TnsE
0
3'
TnsABC +E Pathway ?
TnsC
+28
?

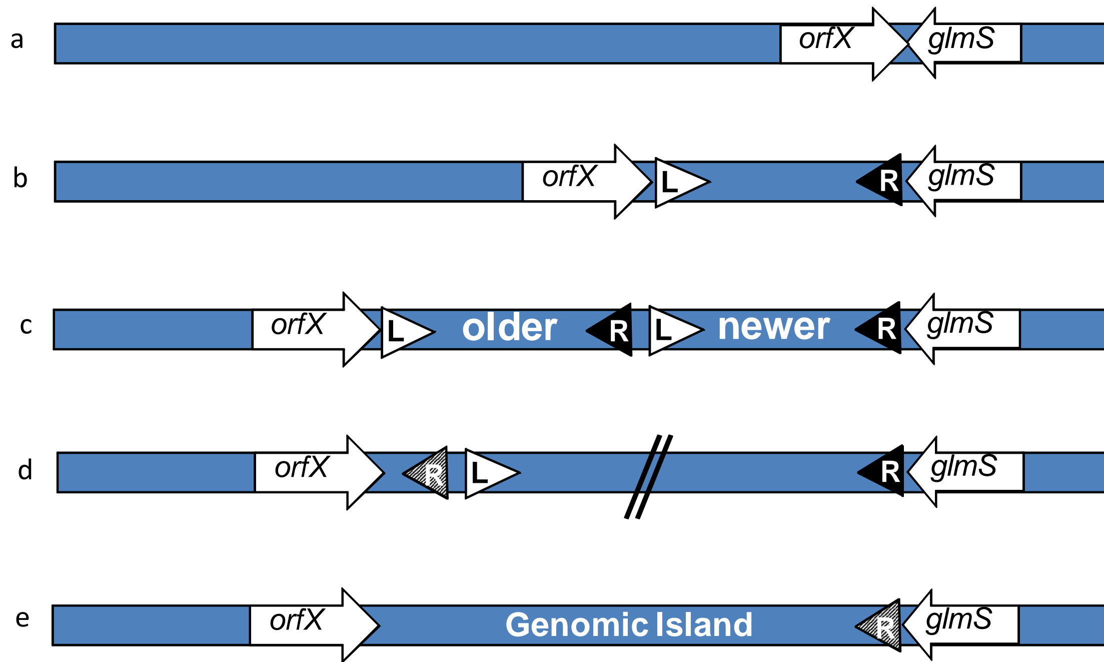
TnsD-pathway, multiple insertions would continue to accumulate very quickly without the immunity conferring process operating. Therefore, target site immunity can be thought of as having a second and more important benefit beyond that found with other transposons in essentially turning off the TnsD-mediated insertion pathway once an insertion resides at the *attTn7* site.

1.6. Expansion of the Tn7 Family by Virtue of Vertical and Horizontal Transposition

It is perhaps not surprising that Tn7-like elements are widespread in the environment given its exceptional level of control. Tn7 elements with transposition proteins identical to the original Tn7 isolate from *E. coli* have been found in different bacterial species isolated from clinical settings. Some of these elements have different antibiotic resistant genes in the class II integron embedded in the variable region of Tn7 (20, 112-115). More importantly, Tn7-like elements, all of which have the *tnsABCDE* genes in synteny, have been found in a wide variety of bacterial hosts from diverse environments. So far, more than 50 Tn7-like elements have been found in genomes of members of γ-proteobacteria, δ-proteobacteria, low-G+C gram-positive firmicutes, and cyanobacteria (106). Most of the elements reside in the equivalent *attTn7* sites in these hosts, but elements in plasmids and non-*attTn7* chromosomal locations are also found, indicating the occurrence of transpositions mediated by both pathways in nature (106, 107). A trend of decreased presence of Tn7 in plasmids, but increased presence in the chromosome has been reported in trimethoprim-resistant *E. coli*(73), suggesting a continuous transposition of Tn7-like elements into its “safe haven” in the nature. Transposition of Tn7-like elements into the *attTn7* site has been shown to lead to the formation of genomic islands (Figure

1.11), as clearly demonstrated in *Shewanella*genus (107) and *Acinetobacter baumannii (125)*, due to the vast capacity of the variable region in Tn7 to carry genes of different functions and potential recombination sequences for other mobile genetic elements. While at first it might seem surprising to find examples in the environment where multiple Tn7-like elements reside in the same location because of target immunity (see above), divergence of the end sequences and TnsB likely accounts for this phenomenon. The presence of non-identical Tn7-like elements in tandem in the same locus is probably because the TnsB homologue of one Tn7-like element is not able to recognize the transposon end sequences of another, thus obscuring the potent target site immunity found with Tn7 (107). Comparison of the transposition protein homologs from Tn7-like elements may provide clues for studying the structure-and-function of these proteins, as has been demonstrated in study of TnsB (137), TnsD (96), and TnsE (105). It is interesting to note that many gain-of-function mutations isolated in the laboratory have been found in naturally occurring transposition protein homologs (106), which may suggest that Tn7-like elements might be more active. While one would expect the broad distribution of a Tn7-like element in various hosts (given the ability of TnsE-mediated transposition to target mobile plasmids), phylogenetic analysis of the TnsABCD proteins actually suggests that Tn7-like elements have limited distribution outside a limited phylogenetic distribution of bacteria (106). Why Tn7-like elements might show some level of species or genus specific adaptation will require further research, but may be related to a need for transposon-encoded proteins to adapt to certain host factors.

Figure 1.11. **Tn7-like elements mediate formation of genomic island**. A Tn7-like element is bounded by the left and right ends (open and filled triangles). A sequence of events may lead to the formation of genomic island(s) in the attachment site, which may eventually lack of any recognizable feature of Tn7. (a) Empty *attTn7* sites can be identified adjacent to the *glmS* gene. (b) Insertion of one Tn7-like element into the *attTn7* site of would bring its variable region possibly with other mobile genetic elements and potential recombination sequences for integrons, bacteriophages, or other transposons. (c) Accumulation of multiple Tn7-like elements (C) in the attachment site would promote homologous recombination between these elements. (d) Over time elements are inactivated by deletions, insertions and point mutations, which would potentially inactivate the elements. (e) Inactivated elements are always subject to further reductive evolution and eventually only the highly selected components may remain.


a
orfX
glmS
b
orfX
L
R
glmS
c
orfX
L
older
R
L
newer
R
glmS
d
orfX
R
L
R
glmS
e
orfX
Genomic Island
R
glmS

1.7. Internal Networking of the Core Machinery

1.7.1 Interaction between TnsA and TnsB

Functional studies have indicated that activities of TnsA and TnsB are inter-dependent. Although TnsA and TnsB have distinct roles in the chemical steps in transposition and their cleavage activities can be separated using inactivated TnsA or TnsB, no transposon end breaking or joining of the free transposon end to the target DNA is detectable in the absence of either one of the other proteins (128). However, under modified *in vitro* reaction conditions, i.e. high glycerol and Mn^{2+}, wild-type TnsAB can promote double-stranded breakage at Tn7 ends and intra-molecular joining where one end of transposon is joined to the other end instead of other DNA molecules (18). TnsA mutants with mutations in the C-terminal domain, in combination with TnsB mutants containing mutations in positions close to the active-site residues, can promote intermolecular transposition independently (83). As TnsA alone cannot bind the ends of Tn7 (10), an interaction between TnsA and TnsB provides a mechanism for TnsB to recruit TnsA to Tn7 ends and for TnsA to modulate activity of TnsB.

1.7.2. Interaction between TnsA and TnsC

Direct interaction between TnsA and TnsC has been demonstrated by pull-down assays (139), gel filtration (124), and protein protease footprinting analysis and crystallographic (83, 124). TnsA and TnsC forms 2:2 complex in solution, probably via the dimerization of TnsC since TnsA appears to be a monomer while TnsC is dimer in solution. The last 52 residues of TnsC(504-555) are necessary and sufficient for a TnsA interaction (124). In the co-crystal structure of TnsA/TnsC(504-555) (Figure 1.12) (124),

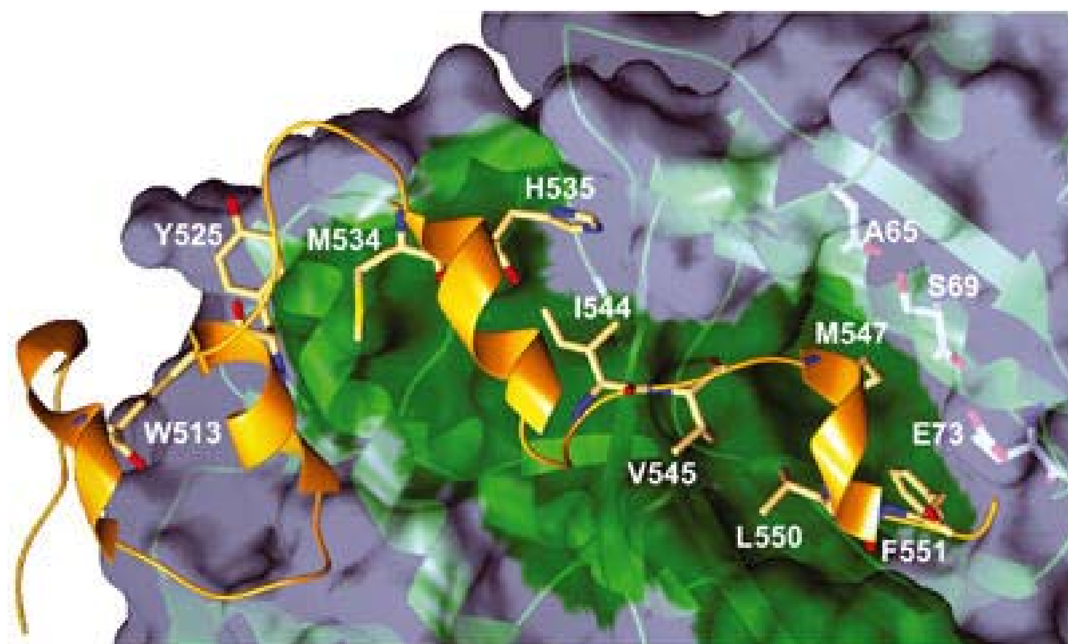
the TnsC 504-555 fragment is in an extended conformation covering a large, exposed hydrophobic surface on TnsA (1400Å), consistent with the increased stability of TnsA/C complex compared to TnsA alone. The N-terminal domain of TnsA interacts with the TnsC fragment (83, 124). All of the class I gain-of-function mutations of TnsA ((A65V, S69N, and E73K)) (83), which allow activation of TnsAB transposase by TnsC in the absence of TnsD (or TnsE), are located in the TnsA/C interface (Figure 1.3 and 1.12).

TnsC(495-555) appears to be able to stimulate the donor DNA cleavage activity of TnsAB in conditions with high concentration of glycerol and Mn^{2+} (124), which, as wediscussed above, requires formation of the complete transposition complex under normal conditions (9). However, when important residues in the region of TnsA for TnsAC interaction are substituted with alanines, the resulting TnsA mutants ($TnsA^{L70A/E71A/W72A}$ and $TnsA^{D78A/E81A}$) are not active in transposition or forming TnsACD-*attTn7* complex (83). The TnsAC interaction may be also critical for formation of pre-transposition complex and channeling a TnsC-mediated transposition signal (see below).

1.7.3. Interaction between TnsB and TnsC

Interaction between TnsB and TnsC determines whether a target complex is formed with TnsC that is productive or abortive, which can be further influenced by TnsA (see below). A productive interaction between TnsB and TnsC is required for assembly of a pre-transposition complex (see below) that is a prerequisite in Tn7 transposition (138). A negative interaction between the proteins is required for target site immunity in that interaction of TnsB with TnsC can promote TnsC dissociation, as described above (140).

Figure 1.12. **TnsC C-terminal peptide on the hydrophobic surface of TnsA**.Surface-exposed hydrophobic residues of TnsA is shown in green and polar or charged residues are in gray. The main chain of TnsC C-terminal peptide (504–555) is shown in gold. TnsC residues involved in hydrophobic interaction with TnsA are labeled. Peach, red, blue, and yellow represent carbon, oxygen, nitrogen, and sulfur, respectively. Gain-of-function TnsA mutations (A65, S69, and E73) are shown with white carbon atoms. Reprinted from (124).


H535
Y525
M534
A65
S69
I544
M547
W513
E73
V545
L550
F551

Interestingly, the same TnsBC interaction is involved in transposition and target immunity and the region of TnsB that mediates this interaction is located at the C-terminal end (Figure 1.13) (137, 138). C-terminal truncated TnsB is not active in transposition even though it is able to bind Tn7 ends and mutating important residues in the C-terminal domain of TnsB abolishes the formation of pre-transposition complex (138). TnsB mutants that allow compromised target immunity bear mutations in this C-terminal region of TnsB and high concentrations of TnsB C-terminal peptide can impose target immunity *in vitro*(137).

The apparent paradox involving productive and inhibitory interactions between the same region of TnsB and TnsC is resolved by the role of TnsA. TnsA stabilizes TnsBCD-*attTn7* complex (138) and also inhibit the TnsB-provoked dissociation of TnsC from a potential target (137), suggesting a role of TnsA in channeling TnsBCinteraction toward productive transposition. The TnsAC interaction, and probably the TnsAB interaction, described above must be important for this function of TnsA. Interplay among protein-protein interactions in the TnsABC core machinery is therefore essential in the regulation of transposition.

The relationship between TnsB and TnsC is somewhat similar to the relationship between the MuA and MuB of bacteriophage Mu (26). For Mu transposition, the MuA transposase binds multiple sites at the ends of the Mu element and mediates 3'-end cleavage and joining for transposition, while MuB is the ATP-dependent protein that binds to the target DNA and activates MuA for transposition. Like the effect of TnsB on TnsC in the Tn7 system, MuA is able to impose target immunity by provoking the

Figure 1.13. **TnsB C-terminus is involved in interaction with TnsC and mutations that affect target immunity and transposition overlap**. The TnsB immunity bypass mutations (top) were localized to the C-terminus of TnsB (middle). Substitution of several residues with alanine in the same region blocks TnsB–TnsC interaction and transposition (bottom). Reprinted from (137) with permission.

```
     686 P->S  -----S---------------    • Transposition competent
     689 V->M  ---------M------------   • Immunity bypass
     690 P->L  ----------L-----------

TnsB 681-EDYSLPTYVPELFQDPPEKDES-702

  682-83 DY->AA  -AA------------------   • Transposition defective
 684-86 SLP->AAA ---AAA---------------  • TnsB-C interaction defective
  689-90 VP->AA  ---------AA-----------
```

ATPase activity of MuB, thereby dissociating MuB from immune targets that already have a copy of Mu. Like TnsB, the same C-terminal domain of MuA is involved in interaction with MuB that leads to either transposition or target immunity. However, the decision whether or not an insertion will occur into a potential target for phage Mu seems to depend on the different interaction between MuA and MuB without input from an additional targeting-protein like TnsD or TnsE.

1.8. Assembly of the transpososome for TnsABC+D transposition

At the heart of Tn7 transposition regulation is the assembly of the transposition complex or transpososome with all of the transposition proteins, an appropriate target DNA, and the transposon ends in the donor DNA. Multiple protein-DNA and protein-protein interactions are required to form such an active complex. For TnsABC+D transposition, assembly of the transposition complex is a prerequisite; no recombination intermediates or products can be produced in the absence of any components of the full reconstitution (10, 88, 128). This is different from other transposons like Tn5 and Tn10, where excision of the transposon occurs before the engagement of the appropriate target site (52, 127). For bacteriophage Mu transposition, excision can occur either before or after the engagement of target DNA (although this flexibility may be an *in vitro* “artifact”)(35).

1.8.1 Pre-transposition complex

Transposition complexes can be formed *in vitro* which do not carry out transposition using Ca^{2+} in the reaction; transposition complexes can be further stabilized through the use of a cross-linking reagent, which allow the identification of pre-transposition

complexes in TnsABC+D transposition as slow-migration complexes in the native gel electrophoresis (Figure 1.14). Like many other transposases (121, 126, 144), TnsB alone can bring together the left- and right- ends of Tn7 to form a complex (138). Higher order nucleoprotein complex was formed with TnsABCD and this complex was active in performing cut-and-paste transposition when isolated and incubated with Mg^{2+}, indicating that this complex (TnsABCD-DNA complex) is a relevant transposition intermediate. Formation of the pre-transposition complex requires that both ends of Tn7 on the same molecule is mediated through the interaction between TnsB and TnsC (138).

TnsA may play a critical role in efficient assembly of the pre-transposition complex, even though it is not necessary for the formation of donor-target complex, as a nucleoprotein complex containing both the donor and target DNA can form with just TnsBCD (TnsBCD-DNA complex). First of all, TnsA is able to bind to the TnsCD-*attTn7* complex and form a higher order complex, the TnsACD-*attTn7* (83). Furthermore, the amount of TnsABCD-DNA complex is much more than TnsBCD-DNA complex and the former withstands temperature challenge while the latter dissociates rapidly (138). These experiments suggest the biologically relevant pre-transpositional complex includes TnsA, TnsB, TnsC, TnsD and the DNAs.

1.8.2 The Post-transposition complex

Unlike the pre-transposition complexes that are not stable without DNA cross-linking, post-transposition complex, the nucleoprotein complex present after *in vitro* Tn7 transposition (Figure 1.14) can be isolated following multiple cycles of washing and gel-filtration in the absence of a cross-linking reagent (68), suggesting that the nucleoprotein

complexes involved in transposition become increasingly stable as the transposition process proceeds. This is probably a result of progressive conformational changes of the complex that drive the reactions forward, as has been found with bacteriophage Mu (56, 98, 156) and also Tn10 (36). Analysis of the post-transposition complex by atomic force microscopy revealed that the Tn7 element is disconnected from the flanking donor DNA and the Tn7 ends are held together as a single protein complex. In the post-transposition complex, the DNA product of transposition in which Tn7 is covalently linked to target *attTn7* is bound all the time by the transposition proteins and only upon deproteinization could the unlinked Tn7 ends be present separately.

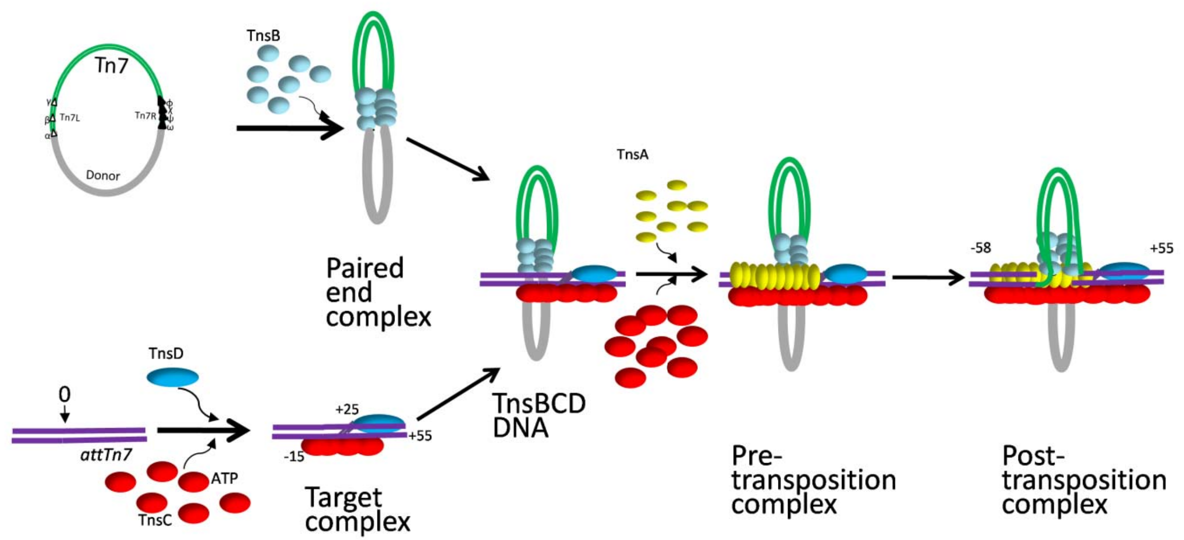
The post-transposition complex and products were also analyzed in a non-denaturing gel (68) by exploiting TnsA, TnsB, TnsC, and TnsD proteins individually labeled with a short fluorescent peptide and fluorescently or radiochemically labeled DNA substrates. TnsA, TnsB, TnsC, TnsD are all present in the post-transposition complex. The presence of TnsD, together with the preservation of TnsD-induced distortion on the target DNA in the post-transposition complex, indicates that TnsD does not function only as an assembly factor to recruit TnsC.

Labeling both the transposition proteins and DNA substrates allowed quantitative investigation of the composition of the complex (68). Since it is very likely that TnsD binds *attTn7* target as a monomer given the asymmetric nature of *attTn7*, the ratio of TnsD to target *attTn7* DNA was set to 1:1, and the abundance of the other transposition proteins was estimated accordingly. Notably TnsB/*attTn7* ratio was 6.3 fold of TnsD/*attTn7*, interestingly close to the TnsB site/ *attTn7* ratio of 7/1, indicating that TnsB

likely binds as monomer. Much higher amounts of TnsA and TnsC were present, TnsA/*attTn7* was about 18 fold and TnsC/*attTn7* was about 25. Given that TnsA and TnsC form $TnsA_2$:$TnsC_2$ complexes (124), it seems likely that some of TnsA and TnsC exist in this form, although some TnsA has dissociated from the post-transposition complex in the course of analysis. It is also possible that a portion of TnsC proteins interact with TnsB, perhaps as 1:1 TnsBC complexes that are distinct from the $TnsA_2$:$TnsC_2$ complexes. Although these observations about Tns protein ratios strongly support the model that proteins of the core transposition machinery function as oligomers, future higher-resolution analysis will be required to unambiguously to determine the oligomeric state of the Tns proteins in the core machinery.

The Tns proteins make extensive interactions with the DNA in the post-transposition complex as revealed by footprinting studies (Figure 1.14)(68). All seven TnsB binding sites in the complex are occupied, which is consistent with the observed ratio of 6.3 TnsB/*attTn7*. In comparison to the protection pattern of Tn7 ends bound by TnsB alone, more protection is found in the post-transposition complex for the terminal α site of Tn7L and ψ and ω sites of Tn7R; the extreme ends of Tn7 are also protected, consistent with the presence of TnsA for 5'-end cleavage in the complex and the engagement of Tn7 ends in the transposition reaction. On the *attTn7* target DNA, TnsD-induced distortion is still evident. Perhaps most intriguingly, a region of 114 bp (-58 - +55) of what was the target DNA is protected, extending well beyond the regions protected by TnsD (+25 - +55)(10) or TnsCD (-15 - + 55)(76). It is not known whether the same number of TnsC molecules is present in the TnsCD-*attTn7* complex and the post-transposition complex, or

Figure 1.14. **Illustration of TnsABC+D transposition pathway at the transpososome level.** Tn7 are represented by green lines with open or filled triangles representing the TnsB-binding sites at the ends of the Tn7. TnsB (cyan) binds and brings together Tn7L and Tn7R to form a paired end complex. TnsD (blue) binds the target *attTn7* (purple lines) and induces a conformational change on the DNA duplex, which recruits TnsC (red) to form a target complex. TnsBC interaction brings the paired-end complex together with the target complex to form a higher order target-donor complex (TnsBCD-DNA), which is further developed to form the complete pre-transposition complex with the oligomerization of TnsC and joining of TnsA (yellow). The nucleoprotein complex is then active for the cleavage reactions that remove flanking donor DNA (thick grey lines) and join the transposon ends to the target DNA. In the stable post-transposition complex, all TnsABCD proteins are present, as well as the TnsD-induced distortion on the target DNA. Protein complex in the post-transposition complex makes extensive interactions with the Tn7::*attTn7* transposition product, especially on the target DNA where a large region left to the insertion point is strikingly protected. The boundaries of protein or protein complexes on the target DNA are indicated in the figure (see text).



Tn7
Tn7L
Tn7R
Donor
TnsB
Paired
end
complex
TnsA
TnsD
0
attTn7
ATP
TnsC
+25
+55
-15
Target
complex
TnsBCD
DNA
Pre-
transposition
complex
-58
+55
Post-
transposition
complex

binding with TnsA and/or TnsB increases the oligomerization of TnsC on target DNA. Since the sequence of DNA left to the insertion point is not important for TnsD targeting (76), protection of ~60 bp by the protein complex in this region is striking. It may be that this region is bound by $TnsA_2C_2$ complex and this asymmetry in Tns-target interaction might determine the orientation specificity of TnsD-mediated transposition (68).

The highly stabile post-transposition complex imposes an obstacle for cellular repair of the gaps flanking the 5'-ends of Tn7 in the final transposition product (68). Repair of the gaps is readily observed with DNA polymerase I and DNA ligase but removal of the protein complex is a prerequisite, indicating that the proteins in the complex block the access of repair enzymes to the gaps. Crude cell lysate can actively disassemble the protein complex and allow repair. This is similar to the disassembly of phage Mu transpososome: Strand transfer in the transposition process creates a stable complex tightly bound by MuA; Disassembly of the transpososome by a series of yet-to-be-defined host proteins, including chaperone ClpX, is required before initiation of phage DNA replication (1, 22, 102).

1.9. Tn7 as a genetic tool

Transposable elements have been an important tool for modifying genes and genomes. The multiple well-defined transposition systems described above make Tn7 an excellent genetic tool for these kinds of applications.

First, the orientation-specific transposition of Tn7 by TnsABC+D pathway to a highly conserved site in the chromosome provides a reliable method for targeted DNA delivery. The high frequency of this pathway, about 1%-10% of the target inserted without any

selection (41), guarantees the successful isolation of positive clones. TnsABC+D transposition-based cloning has proved to be an efficient and straight- forward method for site-specific insertion of desired DNA fragments into the chromosome (12, 69, 91). The ability to isolate positive clones and stable maintenance of the cloned genes in the absence of antibiotic selection makes this methodparticularly useful for efforts to make strains for vaccine production or environmental release (67, 91). In experiments where physiological expression levels of proteins are desired, introducing a single copy of the gene into the chromosome would be advantageous over cloning the same gene on a multi-copy plasmid (11, 91). This site-specific transposition pathway of Tn7 has also been used for transfer of large DNA fragments into different replicons (33, 47, 84), which has great potential in synthetic genomes. It is not unreasonable to suggest that the ability of TnsD-mediated transposition to target eukaryotic *attTn7* sites may eventually be exploited as a tool for gene therapy, avoiding the common mutagenic effects of most current gene therapy vectors (75).

The development of a highly efficient *in vitro* transposition system with mutant core machinery, TnsABC^{A225V}, which has little target site specificity, makes Tn7 transposition an ideal tool for random mutagenesis (19). Together with the availability of multiple miniTn7 derivatives that allow transcriptional and translational fusions, TnsABC^{A225V}*in vitro* transposition have been widely used in making mutation libraries of bacteria (12, 21), yeast (8, 24, 77, 132), and fungi (81). Furthermore, transposon mutagenesis *in vitro* is more effective than *in vivo* mutagenesis as it avoids potential bias imposed by *in vivo*

specific features. *In vitro* transposition mutagenesis is also an appealing choice for organisms that lack readily tractable genetic systems.

Dissection of TnsE-mediated transposition can provide unique insights into many aspects of DNA replication and repair. TnsE-mediated transposition targets β clamp at the primer/template junction for transposition (105). Since the β clamp is known to bind numerous proteins of important cellular functions, traffic on the clamp should be tightly controlled (71). Therefore, it is of particular interest to understand when and where such a complex would be available for Tn7 transposition. Given that the leading-strand DNA replication in fact is also discontinuous (48, 60, 61), it is intriguing to ask why TnsE-mediated only targets the lagging-strand DNA synthesis for transposition or under which condition we can isolate transposition events targeting the leading-strand DNA synthesis. Another yet to be answered question involves asking how repair-associated DNA replication attracts TnsE-mediated transposition and why regional hotspots are found with the induction of double-stranded breaks. Understanding the exceptional ability of Tn7 to sense the differences in leading-strand versus lagging-strand DNA synthesis and repair-associated replication versus normal DNA replication will help elucidate the replication fork structure involved in these processes.

1.10. Concluding remarks

Tn7 is a sophisticated genetic element that has evolved to carefully select among multiple targets for transposition. As such, Tn7 is found globally as an important vector for the horizontal transfer of genetic information playing a role in bacterial evolution through the formation of genomic islands. The use of multiple proteins to execute

transposition in each of two transposition pathways distinguishes it from other transposable elements. The study of Tn7 transposition will also continue to provide valuable insight into how protein-protein and protein-DNA interactions are involved in the assembly of elaborate nucleoprotein complexes. Such interactions are critical in other processes like DNA replication, repair, and other forms of recombination. Since most of the Tn7-encoded proteins form their own unique families, the study of Tn7 will also help understand a large number of proteins present in nature. Furthermore, the targeting of a fundamental cellular process in the TnsABC+E pathway suggests it should be useful in gaining an understanding of basic replication and repair processes in the cell.

1.11. Acknowledgements

Work in the lab of JEP is funded by the National Institutes of Health (RO1GM069508). NLC is an Investigator of the Howard Hughes Medical Institute. Work in the NLC lab is funded in part by the National Institutes of Health (RO1GM076425).

REFERENCES

1. **Abdelhakim, A. H., R. T. Sauer, and T. A. Baker.** 2009. The AAA+ ClpX machine unfolds a keystone subunit to remodel the Mu transpososome. Proc Natl Acad Sci U S A **107:**2437-42.

2. **Adams, D. E., E. M. Shekhtman, E. L. Zechiedrich, M. B. Schmid, and N. R. Cozzarelli.** 1992. The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. Cell **71:**277-88.

3. **Adzuma, K., and K. Mizuuchi.** 1988. Target immunity of Mu transposition reflects a differential distribution of Mu B protein. Cell **53:**257-66.

4. **Aldaz, H., E. Schuster, and T. A. Baker.** 1996. The interwoven architecture of the Mu transposase couples DNA synapsis to catalysis. Cell **85:**257-69.

5. **Arciszewska, L. K., and N. L. Craig.** 1991. Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon. Nucleic Acids Res **19:**5021-9.

6. **Arciszewska, L. K., D. Drake, and N. L. Craig.** 1989. Transposon Tn7. cis-Acting sequences in transposition and transposition immunity. J Mol Biol **207:**35-52.

7. **Arciszewska, L. K., R. L. McKown, and N. L. Craig.** 1991. Purification of TnsB, a transposition protein that binds to the ends of Tn7. J Biol Chem **266:**21736-44.

8. **Bachman, N., M. C. Biery, J. D. Boeke, and N. L. Craig.** 2002. Tn7-mediated mutagenesis of Saccharomyces cerevisiae genomic DNA in vitro. Methods Enzymol **350:**230-47.

9. **Bainton, R., P. Gamas, and N. L. Craig.** 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell **65:**805-16.

10. **Bainton, R. J., K. M. Kubo, J. N. Feng, and N. L. Craig.** 1993. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell **72:**931-43.

11. **Bao, Y., D. P. Lies, H. Fu, and G. P. Roberts.** 1991. An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. Gene **109:**167-8.

12. **Barnes, R. J., K. T. Leung, H. Schraft, and M. Ulanova.** 2008. Chromosomal gfp labelling of Pseudomonas aeruginosa using a mini-Tn7 transposon: application for studies of bacteria-host interactions. Can J Microbiol **54:**48-57.

13. **Barth, P. T., and N. Datta.** 1977. Two naturally occurring transposons indistinguishable from Tn7. J Gen Microbiol **102:**129-34.

14. **Barth, P. T., N. Datta, R. W. Hedges, and N. J. Grinter.** 1976. Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. J Bacteriol **125:**800-10.

15. **Barth, P. T., and N. J. Grinter.** 1977. Map of plasmid RP4 derived by insertion of transposon C. J Mol Biol **113:**455-74.

16. **Barth, P. T., N. J. Grinter, and D. E. Bradley.** 1978. Conjugal transfer system of plasmid RP4: analysis by transposon 7 insertion. J Bacteriol **133:**43-52.

17. **Bhasin, A., I. Y. Goryshin, and W. S. Reznikoff.** 1999. Hairpin formation in Tn5 transposition. J Biol Chem **274:**37021-9.

18. **Biery, M. C., M. Lopata, and N. L. Craig.** 2000. A minimal system for Tn7 transposition: the transposon-encoded proteins TnsA and TnsB can execute DNA breakage and joining reactions that generate circularized Tn7 species. J Mol Biol **297:**25-37.

19. **Biery, M. C., F. J. Stewart, A. E. Stellwagen, E. A. Raleigh, and N. L. Craig.** 2000. A simple in vitro Tn7-based transposition system with low target site selectivity for genome and gene analysis. Nucleic Acids Res **28:**1067-77.

20. **Biskri, L., and D. Mazel.** 2003. Erythromycin esterase gene ere(A) is located in a functional gene cassette in an unusual class 2 integron. Antimicrob Agents Chemother **47:**3326-31.

21. **Bordi, C., B. G. Butcher, Q. Shi, A. B. Hachmann, J. E. Peters, and J. D. Helmann.** 2008. In vitro mutagenesis of Bacillus subtilis by using a modified Tn7 transposon with an outward-facing inducible promoter. Appl Environ Microbiol **74:**3419-25.

22. **Burton, B. M., and T. A. Baker.** 2003. Mu transpososome architecture ensures that unfolding by ClpX or proteolysis by ClpXP remodels but does not destroy the complex. Chem Biol **10:**463-72.

23. **Bussiere, D. E., and D. Bastia.** 1999. Termination of DNA replication of bacterial and plasmid chromosomes. Mol Microbiol **31:**1611-8.

24. **Castano, I., R. Kaur, S. Pan, R. Cregg, L. Penas Ade, N. Guo, M. C. Biery, N. L. Craig, and B. P. Cormack.** 2003. Tn7-based genome-wide random insertional mutagenesis of Candida glabrata. Genome Res **13:**905-15.

25. **Chaconas, G., B. D. Lavoie, and M. A. Watson.** 1996. DNA transposition: jumping gene machine, some assembly required. Curr Biol **6:**817-20.

26. **Chaconas, G. H., R.M. ;.** 2002. Transposition of phage Mu DNA., p. 384-402. *In* N. L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz. (ed.), Mobile DNA II. ASM Press, Washington, DC.

27. **Chalmers, R. M., and N. Kleckner.** 1996. IS10/Tn10 transposition efficiently accommodates diverse transposon end configurations. Embo J **15:**5112-22.

28. **Cleaver, S. H., and E. Wickstrom.** 2000. Transposon Tn7 gene insertion into an evolutionarily conserved human homolog of Escherichia coli attTn7. Gene **254:**37-44.

29. **Clubb, R. T., J. G. Omichinski, H. Savilahti, K. Mizuuchi, A. M. Gronenborn, and G. M. Clore.** 1994. A novel class of winged helix-turn-helix protein: the DNA-binding domain of Mu transposase. Structure **2:**1041-8.

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

31. **Consortium, T. U.** 2010. The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res **38:**D142-8.

32. **Court, D. L., J. A. Sawitzke, and L. C. Thomason.** 2002. Genetic engineering using homologous recombination. Annu Rev Genet **36:**361-88.

33. **Craig, N. L.** 2002. Tn7, p. 423-456. *In* N. L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz. (ed.), Mobile DNA II. ASM Press, Washington, DC.

34. **Craig, N. L.** 1996. Transposon Tn7. Curr Top Microbiol Immunol **204:**27-48.

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

36. **Crellin, P., and R. Chalmers.** 2001. Protein-DNA contacts and conformational changes in the Tn10 transpososome during assembly and activation for cleavage. Embo J **20:**3882-91.

37. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res **14:**1188-90.

38. **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.

39. **Dalrymple, B. P., K. Kongsuwan, G. Wijffels, N. E. Dixon, and P. A. Jennings.** 2001. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. Proc Natl Acad Sci U S A **98:**11627-32.

40. **Davies, D. R., I. Y. Goryshin, W. S. Reznikoff, and I. Rayment.** 2000. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science **289:**77-85.

41. **DeBoy, R. T., and N. L. Craig.** 1996. Tn7 transposition as a probe of cis interactions between widely separated (190 kilobases apart) DNA sites in the Escherichia coli chromosome. J Bacteriol **178:**6184-91.

42. **Duggin, I. G., R. G. Wake, S. D. Bell, and T. M. Hill.** 2008. The replication fork trap and termination of chromosome replication. Mol Microbiol **70:**1323-33.

43. **Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie, and D. R. Davies.** 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science **266:**1981-6.

44. **Feng, X., and S. D. Colloms.** 2007. In vitro transposition of ISY100, a bacterial insertion sequence belonging to the Tc1/mariner family. Mol Microbiol **65:**1432-43.

45. **Finn, J. A., A. R. Parks, and J. E. Peters.** 2007. Transposon Tn7 directs transposition into the genome of filamentous bacteriophage M13 using the element-encoded TnsE protein. J Bacteriol **189:**9122-5.

46. **Finn, R. D., J. Mistry, J. Tate, P. Coggill, A. Heger, J. E. Pollington, O. L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E. L. Sonnhammer, S. R. Eddy, and A. Bateman.** 2010. The Pfam protein families database. Nucleic Acids Res **38:**D211-22.

47. **Frengen, E., D. Weichenhan, B. Zhao, K. Osoegawa, M. van Geel, and P. J. de Jong.** 1999. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. Genomics **58:**250-3.

48. **Gabbai, C. B., and K. J. Marians.** 2010. Recruitment to stalled replication forks of the PriA DNA helicase and replisome-loading activities is essential for survival. DNA Repair (Amst) **9:**202-9.

49. **Gamas, P., and N. L. Craig.** 1992. Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA. Nucleic Acids Res **20:**2525-32.

50. **Gary, P. A., M. C. Biery, R. J. Bainton, and N. L. Craig.** 1996. Multiple DNA processing reactions underlie Tn7 transposition. J Mol Biol **257:**301-16.

51. **Gay, N. J., V. L. Tybulewicz, and J. E. Walker.** 1986. Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. Biochem J **234:**111-7.

52. **Goryshin, I. Y., and W. S. Reznikoff.** 1998. Tn5 in vitro transposition. J Biol Chem **273:**7367-74.

53. **Goto, N., A. Shoji, S. Horiuchi, and R. Nakaya.** 1984. Conduction of nonconjugative plasmids by F' lac is not necessarily associated with transposition of the gamma delta sequence. J Bacteriol **159:**590-6.

54. **Grindley, N. D. F.** 2002. The movement of Tn3-like elements: transposition and cointegrate resolution., p. 272-302. *In* N. L. Caig, R. Craigie, M. Gellert, and A.M. Lambowitz. (ed.), Mobile DNA II. ASM Press, Washington, DC.

55. **Gringauz, E., K. A. Orle, C. S. Waddell, and N. L. Craig.** 1988. Recognition of Escherichia coli attTn7 by transposon Tn7: lack of specific sequence requirements at the point of Tn7 insertion. J Bacteriol **170:**2832-40.

56. **Gueguen, E., P. Rousseau, G. Duval-Valentin, and M. Chandler.** 2005. The transpososome: control of transposition at the level of catalysis. Trends Microbiol **13:**543-9.

57. **Hanson, P. I., and S. W. Whiteheart.** 2005. AAA+ proteins: have engine, will work. Nat Rev Mol Cell Biol **6:**519-29.

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

59. **Hauer, B., and J. A. Shapiro.** 1984. Control of Tn7 transposition. Mol Gen Genet **194:**149-58.

60. **Heller, R. C., and K. J. Marians.** 2006. Replication fork reactivation downstream of a blocked nascent leading strand. Nature **439:**557-62.

61. **Heller, R. C., and K. J. Marians.** 2006. Replisome assembly and the direct restart of stalled replication forks. Nat Rev Mol Cell Biol **7:**932-43.

62. **Hiasa, H., and K. J. Marians.** 1994. Tus prevents overreplication of oriC plasmid DNA. J Biol Chem **269:**26959-68.

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

64. **Hickman, A. B., Y. Li, S. V. Mathew, E. W. May, N. L. Craig, and F. Dyda.** 2000. Unexpected structural diversity in DNA recombination: the restriction endonuclease connection. Mol Cell **5:**1025-34.

65. **Higuchi, K., T. Katayama, S. Iwai, M. Hidaka, T. Horiuchi, and H. Maki.** 2003. Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication in vitro. Genes Cells **8:**437-49.

66. **Hill, T. M., and K. J. Marians.** 1990. Escherichia coli Tus protein acts to arrest the progression of DNA replication forks in vitro. Proc Natl Acad Sci U S A **87:**2481-5.

67. **Hojberg, O., U. Schnider, H. V. Winteler, J. Sorensen, and D. Haas.** 1999. Oxygen-sensing reporter strain of Pseudomonas fluorescens for monitoring the distribution of low-oxygen habitats in soil. Appl Environ Microbiol **65:**4085-93.

68. **Holder, J. W., and N. L. Craig.** 2010. Architecture of the Tn7 posttransposition complex: an elaborate nucleoprotein structure. J Mol Biol **401:**167-81.

69. **Howe, K., A. Karsi, P. Germon, R. W. Wills, M. L. Lawrence, and R. H. Bailey.** 2010. Development of stable reporter system cloning luxCDABE genes into chromosome of Salmonella enterica serotypes using Tn7 transposon. BMC Microbiol **10:**197.

70. **Hunter, S., R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, U. Das, L. Daugherty, L. Duquenne, R. D. Finn, J. Gough, D. Haft, N. Hulo, D. Kahn, E. Kelly, A. Laugraud, I. Letunic, D. Lonsdale, R. Lopez, M. Madera, J. Maslen, C. McAnulla, J. McDowall, J. Mistry, A. Mitchell, N. Mulder, D. Natale, C. Orengo, A. F. Quinn, J. D. Selengut, C. J. Sigrist, M. Thimma, P. D. Thomas, F. Valentin, D. Wilson, C. H. Wu, and C. Yeats.** 2009. InterPro: the integrative protein signature database. Nucleic Acids Res **37:**D211-5.

71. **Johnson, A., and M. O'Donnell.** 2005. Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem **74:**283-315.

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

73. **Kraft, C. A., M. C. Timbury, and D. J. Platt.** 1986. Distribution and genetic location of Tn7 in trimethoprim-resistant Escherichia coli. J Med Microbiol **22:**125-31.

74. **Kubo, K. M., and N. L. Craig.** 1990. Bacterial transposon Tn7 utilizes two different classes of target sites. J Bacteriol **172:**2774-8.

75. **Kuduvalli, P. N., R. Mitra, and N. L. Craig.** 2005. Site-specific Tn7 transposition into the human genome. Nucleic Acids Res **33:**857-63.

76. **Kuduvalli, P. N., J. E. Rao, and N. L. Craig.** 2001. Target DNA structure plays a critical role in Tn7 transposition. Embo J **20:**924-32.

77. **Kumar, A., M. Seringhaus, M. C. Biery, R. J. Sarnovsky, L. Umansky, S. Piccirillo, M. Heidtman, K. H. Cheung, C. J. Dobry, M. B. Gerstein, N. L. Craig, and M. Snyder.** 2004. Large-scale mutagenesis of the yeast genome using a Tn7-derived multipurpose transposon. Genome Res **14:**1975-86.

78. **Lee, C. H., A. Bhagwat, and F. Heffron.** 1983. Identification of a transposon Tn3 sequence required for transposition immunity. Proc Natl Acad Sci U S A **80:**6765-9.

79. **Letunic, I., T. Doerks, and P. Bork.** 2009. SMART 6: recent updates and new developments. Nucleic Acids Res **37:**D229-32.

80. **Lichtenstein, C., and S. Brenner.** 1982. Unique insertion site of Tn7 in the E. coli chromosome. Nature **297:**601-3.

81. **Lo, C., K. Adachi, J. R. Shuster, J. E. Hamer, and L. Hamer.** 2003. The bacterial transposon Tn7 causes premature polyadenylation of mRNA in eukaryotic organisms: TAGKO mutagenesis in filamentous fungi. Nucleic Acids Res **31:**4822-7.

82. **Lopez de Saro, F. J., and M. O'Donnell.** 2001. Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. Proc Natl Acad Sci U S A **98:**8376-80.

83. **Lu, F., and N. L. Craig.** 2000. Isolation and characterization of Tn7 transposase gain-of-function mutants: a model for transposase activation. Embo J **19:**3446-57.

84. **Luckow, V. A., S. C. Lee, G. F. Barry, and P. O. Olins.** 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. J Virol **67:**4566-79.

85. **Mancuso, M., M. C. Sammarco, and E. Grabczyk.** 2010. Transposon Tn7 preferentially inserts into GAA*TTC triplet repeats under conditions conducive to Y*R*Y triplex formation. PLoS One **5:**e11121.

86. **Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song,**

N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res **39:**D225-9.

87. **Markovitz, A.** 2005. A new in vivo termination function for DNA polymerase I of Escherichia coli K12. Mol Microbiol **55:**1867-82.

88. **May, E. W., and N. L. Craig.** 1996. Switching from cut-and-paste to replicative Tn7 transposition. Science **272:**401-4.

89. **McInerney, P., A. Johnson, F. Katz, and M. O'Donnell.** 2007. Characterization of a triple DNA polymerase replisome. Mol Cell **27:**527-38.

90. **McInerney, P., and M. O'Donnell.** 2004. Functional uncoupling of twin polymerases: mechanism of polymerase dissociation from a lagging-strand block. J Biol Chem **279:**21543-51.

91. **McKenzie, G. J., and N. L. Craig.** 2006. Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn7 without need for selection of the insertion event. BMC Microbiol **6:**39.

92. **McKown, R. L., K. A. Orle, T. Chen, and N. L. Craig.** 1988. Sequence requirements of Escherichia coli attTn7, a specific site of transposon Tn7 insertion. J Bacteriol **170:**352-8.

93. **McKown, R. L., C. S. Waddell, L. K. Arciszewska, and N. L. Craig.** 1987. Identification of a transposon Tn7-dependent DNA-binding activity that recognizes the ends of Tn7. Proc Natl Acad Sci U S A **84:**7807-11.

94. **Mirkin, E. V., and S. M. Mirkin.** 2007. Replication fork stalling at natural impediments. Microbiol Mol Biol Rev **71:**13-35.

95. **Mitra, R., J. Fain-Thornton, and N. L. Craig.** 2008. piggyBac can bypass DNA synthesis during cut and paste transposition. Embo J **27:**1097-109.

96. **Mitra, R., G. J. McKenzie, L. Yi, C. A. Lee, and N. L. Craig.** 2010. Characterization of the TnsD-attTn7 complex that promotes site-specific insertion of Tn7. Mob DNA **1:**18.

97. **Mizuuchi, K.** 1992. Transpositional recombination: mechanistic insights from studies of mu and other elements. Annu Rev Biochem **61:**1011-51.

98. **Mizuuchi, M., T. A. Baker, and K. Mizuuchi.** 1995. Assembly of phage Mu transpososomes: cooperative transitions assisted by protein and DNA scaffolds. Cell **83:**375-85.

99. **Namgoong, S. Y., and R. M. Harshey.** 1998. The same two monomers within a MuA tetramer provide the DDE domains for the strand cleavage and strand transfer steps of transposition. Embo J **17:**3775-85.

100. **Nicolas, E., M. Lambin, and B. Hallet.** Target immunity of the Tn3-family transposon Tn4430 requires specific interactions between the transposase and the terminal inverted repeats of the transposon. J Bacteriol **192:**4233-8.

101. **Nisa-Martinez, R., J. I. Jimenez-Zurdo, F. Martinez-Abarca, E. Munoz-Adelantado, and N. Toro.** 2007. Dispersion of the RmInt1 group II intron in the Sinorhizobium meliloti genome upon acquisition by conjugative transfer. Nucleic Acids Res **35:**214-22.

102. **North, S. H., and H. Nakai.** 2005. Host factors that promote transpososome disassembly and the PriA-PriC pathway for restart primosome assembly. Mol Microbiol **56:**1601-16.

103. **O'Donnell, M., D. Jeruzalmi, and J. Kuriyan.** 2001. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. Curr Biol **11:**R935-46.

104. **Orle, K. A., and N. L. Craig.** 1991. Identification of transposition proteins encoded by the bacterial transposon Tn7. Gene **104:**125-31.

105. **Parks, A. R., Z. Li, Q. Shi, R. M. Owens, M. M. Jin, and J. E. Peters.** 2009. Transposition into replicating DNA occurs through interaction with the processivity factor. Cell **138:**685-95.

106. **Parks, A. R., and J. E. Peters.** 2009. Tn7 elements: engendering diversity from chromosomes to episomes. Plasmid **61:**1-14.

107. **Parks, A. R., and J. E. Peters.** 2007. Transposon Tn7 is widespread in diverse bacteria and forms genomic islands. J Bacteriol **189:**2170-3.

108. **Peters, J. E., and N. L. Craig.** 2001. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes Dev **15:**737-47.

109. **Peters, J. E., and N. L. Craig.** 2000. Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates. Mol Cell **6:**573-82.

110. **Peters, J. E., and N. L. Craig.** 2001. Tn7: smarter than we thought. Nat Rev Mol Cell Biol **2:**806-14.

111. **Radstrom, P., O. Skold, G. Swedberg, J. Flensburg, P. H. Roy, and L. Sundstrom.** 1994. Transposon Tn5090 of plasmid R751, which carries an integron, is related to Tn7, Mu, and the retroelements. J Bacteriol **176:**3257-68.

112. **Ramirez, M. S., H. Bello, G. Gonzalez Rocha, C. Marquez, and D. Centron.** 2010. Tn7::In2-8 dispersion in multidrug resistant isolates of Acinetobacter baumannii from Chile. Rev Argent Microbiol **42:**138-40.

113. **Ramirez, M. S., C. Quiroga, and D. Centron.** 2005. Novel rearrangement of a class 2 integron in two non-epidemiologically related isolates of Acinetobacter baumannii. Antimicrob Agents Chemother **49:**5179-81.

114. **Ramirez, M. S., M. S. Stietz, E. Vilacoba, P. Jeric, A. S. Limansky, M. Catalano, and D. Centron.** 2011. Increasing frequency of class 1 and 2 integrons in multidrug-resistant clones of Acinetobacter baumannii reveals the need for continuous molecular surveillance. Int J Antimicrob Agents **37:**175-7.

115. **Ramirez, M. S., L. J. Vargas, V. Cagnoni, M. Tokumoto, and D. Centron.** 2005. Class 2 integron with a novel cassette array in a Burkholderia cenocepacia isolate. Antimicrob Agents Chemother **49:**4418-20.

116. **Rao, J. E., and N. L. Craig.** 2001. Selective recognition of pyrimidine motif triplexes by a protein encoded by the bacterial transposon Tn7. J Mol Biol **307:**1161-70.

117. **Rao, J. E., P. S. Miller, and N. L. Craig.** 2000. Recognition of triple-helical DNA structures by transposon Tn7. Proc Natl Acad Sci U S A **97:**3936-41.

118. **Reimmann, C., R. Moore, S. Little, A. Savioz, N. S. Willetts, and D. Haas.** 1989. Genetic structure, function and regulation of the transposable element IS21. Mol Gen Genet **215:**416-24.

119. **Reyes, O., A. Beyou, C. Mignotte-Vieux, and F. Richaud.** 1987. Mini-Mu transduction: cis-inhibition of the insertion of Mud transposons. Plasmid **18:**183-92.

120. **Rice, P., and K. Mizuuchi.** 1995. Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. Cell **82:**209-20.

121. **Rice, P. A., and T. A. Baker.** 2001. Comparative architecture of transposase and integrase complexes. Nat Struct Biol **8:**302-7.

122. **Richardson, J. M., S. D. Colloms, D. J. Finnegan, and M. D. Walkinshaw.** 2009. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. Cell **138:**1096-108.

123. **Robinson, M. K., P. M. Bennett, and M. H. Richmond.** 1977. Inhibition of TnA translocation by TnA. J Bacteriol **129:**407-14.

124. **Ronning, D. R., Y. Li, Z. N. Perez, P. D. Ross, A. B. Hickman, N. L. Craig, and F. Dyda.** 2004. The carboxy-terminal portion of TnsC activates the Tn7 transposase through a specific interaction with TnsA. Embo J **23:**2972-81.

125. **Rose, A.** 2010. TnAbaR1: a novel Tn7-related transposon in *Acinetobacter baumannii* that contributes to the accumulation and dissemination of large repertoires of resistance genes. Bioscience Horizons **3:**40-48.

126. **Sakai, J., R. M. Chalmers, and N. Kleckner.** 1995. Identification and characterization of a pre-cleavage synaptic complex that is an early intermediate in Tn10 transposition. Embo J **14:**4374-83.

127. **Sakai, J., and N. Kleckner.** 1997. The Tn10 synaptic complex can capture a target DNA only after transposon excision. Cell **89:**205-14.

128. **Sarnovsky, R. J., E. W. May, and N. L. Craig.** 1996. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. Embo J **15:**6348-61.

129. **Savilahti, H., and K. Mizuuchi.** 1996. Mu transpositional recombination: donor DNA cleavage and strand transfer in trans by the Mu transposase. Cell **85:**271-80.

130. **Schumacher, S., R. T. Clubb, M. Cai, K. Mizuuchi, G. M. Clore, and A. M. Gronenborn.** 1997. Solution structure of the Mu end DNA-binding ibeta subdomain of phage Mu transposase: modular DNA recognition by two tethered domains. Embo J **16:**7532-41.

131. **Schweidenback, C. T., and T. A. Baker.** 2008. Dissecting the roles of MuB in Mu transposition: ATP regulation of DNA binding is not essential for target delivery. Proc Natl Acad Sci U S A **105:**12101-7.

132. **Seringhaus, M., A. Kumar, J. Hartigan, M. Snyder, and M. Gerstein.** 2006. Genomic analysis of insertion behavior and target specificity of mini-Tn7 and Tn3 transposons in Saccharomyces cerevisiae. Nucleic Acids Res **34:**e57.

133. **Sharpe, P. L., and N. L. Craig.** 1998. Host proteins can stimulate Tn7 transposition: a novel role for the ribosomal protein L29 and the acyl carrier protein. Embo J **17:**5822-31.

134. **Sherratt, D. J.** 2003. Bacterial chromosome dynamics. Science **301:**780-5.

135. **Shi, Q., J. C. Huguet-Tapia, and J. E. Peters.** 2009. Tn917 targets the region where DNA replication terminates in Bacillus subtilis, highlighting a difference in chromosome processing in the firmicutes. J Bacteriol **191:**7623-7.

136. **Shi, Q., A. R. Parks, B. D. Potter, I. J. Safir, Y. Luo, B. M. Forster, and J. E. Peters.** 2008. DNA damage differentially activates regional chromosomal loci for Tn7 transposition in Escherichia coli. Genetics **179:**1237-50.

137. **Skelding, Z., J. Queen-Baker, and N. L. Craig.** 2003. Alternative interactions between the Tn7 transposase and the Tn7 target DNA binding protein regulate target immunity and transposition. Embo J **22:**5904-17.

138. **Skelding, Z., R. Sarnovsky, and N. L. Craig.** 2002. Formation of a nucleoprotein complex containing Tn7 and its target DNA regulates transposition initiation. Embo J **21:**3494-504.

139. **Stellwagen, A. E., and N. L. Craig.** 2001. Analysis of gain-of-function mutants of an ATP-dependent regulator of Tn7 transposition. J Mol Biol **305:**633-42.

140. **Stellwagen, A. E., and N. L. Craig.** 1997. Avoiding self: two Tn7-encoded proteins mediate target immunity in Tn7 transposition. Embo J **16:**6823-34.

141. **Stellwagen, A. E., and N. L. Craig.** 1997. Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7. Genetics **145:**573-85.

142. **Stellwagen, A. E., and N. L. Craig.** 1998. Mobile DNA elements: controlling transposition with ATP-dependent molecular switches. Trends Biochem Sci **23:**486-90.

143. **Stukenberg, P. T., J. Turner, and M. O'Donnell.** 1994. An explanation for lagging-strand replication: polymerase hopping among DNA sliding clamps. Cell **78:**877-87.

144. **Surette, M. G., S. J. Buch, and G. Chaconas.** 1987. Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. Cell **49:**253-62.

145. **Suski, C., and K. J. Marians.** 2008. Resolution of converging replication forks by RecQ and topoisomerase III. Mol Cell **30:**779-89.

146. **Swingle, B., E. Markel, N. Costantino, M. G. Bubunenko, S. Cartinhour, and D. L. Court.** 2010. Oligonucleotide recombination in Gram-negative bacteria. Mol Microbiol **75:**138-48.

147. **Swingle, B., M. O'Carroll, D. Haniford, and K. M. Derbyshire.** 2004. The effect of host-encoded nucleoid proteins on transposition: H-NS influences targeting of both IS903 and Tn10. Mol Microbiol **52:**1055-67.

148. **Tang, M., C. Cecconi, C. Bustamante, and D. C. Rio.** 2007. Analysis of P element transposase protein-DNA interactions during the early stages of transposition. J Biol Chem **282:**29002-12.

149. **Tang, Y., S. Cotterill, and C. P. Lichtenstein.** 1995. Genetic analysis of the terminal 8-bp inverted repeats of transposon Tn7. Gene **162:**41-6.

150. **van Luenen, H. G., S. D. Colloms, and R. H. Plasterk.** 1994. The mechanism of transposition of Tc3 in C. elegans. Cell **79:**293-301.

151. **Waddell, C. S., and N. L. Craig.** 1989. Tn7 transposition: recognition of the attTn7 target sequence. Proc Natl Acad Sci U S A **86:**3958-62.

152. **Waddell, C. S., and N. L. Craig.** 1988. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. Genes Dev **2:**137-49.

153. **Wilkins, B. M., E. Lanka** 1993. BACTERIAL CONJUGATION.

154. **Wolkow, C. A., R. T. DeBoy, and N. L. Craig.** 1996. Conjugating plasmids are preferred targets for Tn7. Genes Dev **10:**2145-57.

155. **Yamauchi, M., and T. A. Baker.** 1998. An ATP-ADP switch in MuB controls progression of the Mu transposition pathway. Embo J **17:**5509-18.

156. **Yanagihara, K., and K. Mizuuchi.** 2003. Progressive structural transitions within Mu transpositional complexes. Mol Cell **11:**215-24.

157. **Yuzhakov, A., J. Turner, and M. O'Donnell.** 1996. Replisome assembly reveals the basis for asymmetric function in leading- and lagging-strand replication. Cell **86:**877-86.

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

CHAPTER 2

TRANSPOSITION INTO REPLICATING DNA OCCURS THROUGH INTERACTION WITH THE PROCESSIVITY FACTOR[2]

2.1. Summary

The bacterial transposon Tn7 directs transposition into actively replicating DNA by a mechanism involving the transposon-encoded protein TnsE. In this chapter I established an *in vitro* TnsABC+E transposition reaction reconstituted from purified proteins and target DNA structures. Using the *in vitro* reaction I confirm that the processivity factor specifically reorders TnsE-mediated transposition events on target DNAs in a way that matches the bias with active DNA replication *in vivo*. The TnsE interaction with an essential and conserved component of the replication machinery and a DNA structure reveals a new mechanism by which Tn7, and probably other elements, select target-sites associated with DNA replication.

2.2. Introduction

Transposons are genetic elements that are capable of moving from one location to another within a cell. The bacterial transposon Tn7 and its relatives are abundantly distributed amongst bacteria in a wide variety of medical and environmental settings (24, 25). Tn7 has served as a model system for transposition, especially for the understanding of transposon target-site selection (reviewed in (5, 29)). Target-site selection is the process by which transposons assess new DNA molecules for potential insertion. While

[2]This chapter has been published in:
Parks AR, Li Z, Shi Q, Owens RM, Jin MM, et al. (2009) Transposition into replicating DNA occurs through interaction with the processivity factor. Cell 138: 685-695.

most transposable elements possess a weak target DNA sequence preference that guides target-site selection, Tn7 utilizes two distinct target-site selection pathways. In one pathway a sequence-specific DNA binding protein directs transposition into a single site within the bacterial chromosome, and in the other a separate protein recognizes a process associated with DNA replication. These two target selection pathways optimize vertical and horizontal transmission of the transposable element respectively (5, 24).

Tn7 encodes five genes whose products conduct transposition (5). TnsA and TnsB comprise the transposase that catalyzes the DNA breakage and joining reactions at the transposon ends to mobilize the element. TnsC is an AAA regulator protein that activates the transposase when an appropriate target DNA has been found (38). TnsD and TnsE identify target DNAs and signal TnsABC to activate transposition (5). Target-site selection is a prerequisite for activation of transposition with Tn7; transposon excision and insertion does not occur until an appropriate target has been identified. TnsD recognizes a specific site, called its attachment site or *attTn7*, by binding to a highly conserved DNA sequence within the 3' end of the *glmS* gene. The TnsE protein recognizes an incompletely defined feature associated with discontinuous DNA replication (27)that is over-represented or especially accessible in mobile plasmids, called conjugal plasmids, as they enter a new host cell (44, 45).

While TnsE-mediated transposition preferentially occurs into mobile plasmids undergoing conjugal DNA replication, at a lower frequency, the TnsABC+E machinery also recognizes sites within the bacterial chromosome with a preference for the region where DNA replication terminates and regions proximal to DNA double-strand breaks

(28, 34). The orientation of the transposon ends following TnsE-mediated transposition indicates that discontinuously replicated DNA is in some way recognized by TnsE (27, 45). As mobile plasmids enter a new host cell, they replicate in a single direction by a discontinuous process, like lagging-strand DNA synthesis (44). In both mobile plasmids and in the chromosome, transposition events occur in a single orientation correlating with the direction of replication progression (27, 29, 45). It has been shown that TnsE is a DNA binding protein that preferentially binds to DNA structures that present a free 3'-recessed end (27). Given that TnsD relies in part on additional host factors in activating transposition (33), it is conceivable that host factors associated with discontinuous DNA replication allow the selection of targets during TnsE-mediated transposition.

An intriguing host factor candidate that could allow the TnsABC+E transposition machinery to target lagging-strand DNA synthesis is the DNA replication processivity factor. Processivity factors are essential clamp proteins that encircle DNA and serve as a mobile platform, linking proteins to DNA (14, 41). Interestingly, the inactive *pogo* element, found in *Drosophila*, encodes a transposase that has been shown to interact with the processivity factor (42). Because the element is no longer active, no functional link between this interaction and transposition has yet been established (41). It therefore remains unclear if the interaction was important in the original element. An interaction with the processivity factor could possibly regulate transposition with DNA replication and repair, or could be involved in target-site selection. We reasoned that TnsE might use an interaction with the processivity factor to direct transposition into certain forms of DNA replication.

The processivity factor, β clamp in Bacteria and PCNA in Eukaryotes and Archaea, is enriched on discontinuously replicating DNA (reviewed in (14) and references therein). β and PCNA have been shown to interact with many proteins involved in DNA repair, Okazaki fragment maturation, and regulation of the cell cycle (14, 41). The processivity factor binding motif found in proteins that interact with β and PCNA fits into a hydrophobic cleft found in the clamp proteins (6, 11, 14, 41). Competition for this common region of the clamp appears to play a role in coordination of proteins involved in DNA metabolism (19).

In our publication in *cell* 2009 (23) we reportedthat TnsE targets DNA replication by interactingwith the β clamp processivity factor. My contribution was a critical part of that work. I reconstituted the TnsABC+E transposition pathway using purified components. The *in vitro* reaction confirms that two factors, interaction with a DNA structure and with the processivity factor account for the bias of TnsABC+E transposition with active DNA replication. *In vitro* and *in vivo*analysis of the TnsE-β interaction explains how Tn7 targets DNA replication without negatively affecting the cell. These findings likely reveal a general strategy used by other unrelated transposons for directing transposition to DNA replication intermediates.

2.3. Results

In an effort to understand how TnsE identifies a target DNA people in the lab looked for conserved motifs within the amino acid sequence of TnsE. A putative β-binding motif (PQLELARALFL) was identified in TnsE and its homologs using the ClustalW

algorithm (39). This motif shows a modest resemblance to the consensus processivity factor binding motifs found in bacterial host proteins (QL(S/D)LF (6) and QLxLxL (43). Interaction between TnsE and the β-clamp was demonstrated by yeast two-hybrid assay (8) and confirmed by multiple methods, including a protein mobility shift assay (20, 21), a far western blot technique (7) and surface plasmon resonance (SPR) analysis. The TnsE-β interaction was characterized by SPR analysis to have a K_D of ~2.44 μM, which is intermediate comparing to the known interactions of β with the δ and γ subunits of the clamp loader complex (18).

The putative β-binding motif (PQLELARALFL) is important for the observed TnsE-β interaction as mutating residues in the motif individually or in combination to ananine(s) attenuated the interaction, even though the stability of the mutant proteins were largely unaffected. More importantly, the alanine mutations were found to specifically abolish or significantly reduce the frequency of TnsABC+E transposition *in vivo* in a lambda hop assay (22). These results indicate that the observed decreases in transposition frequency are consistent with decreased β-clamp binding ability, supporting the idea that activation of transposition via the TnsE pathway is dependent on binding to the β-clamp. In support of this, the transposition frequency of TnsE-mediated transposition was found to be specifically stimulated in strains with moderate over-expression of β.

To confirm the molecular components required for targeting active DNA replication *in vivo*I established an *in vitro* system for TnsABC+E transposition using purified components. While *in vitro* systems for two other Tn7 transposition pathways exist (the TnsD pathway (1) and an untargeted mutant core machinery pathway termed TnsABC*

(36)), there have been no reports of the reconstitution of the TnsE pathway. I took advantage of previous analyses of TnsE DNA binding that indicated that the protein shows a strong preference for DNA structures containing 3'-recessed ends (27), a structure more commonly found on the lagging-strand than the leading-strand (14). For my *in vitro* assay, I constructed DNA-target molecules that contain a 20 bp single-stranded gap in the duplex DNA (Experimental Procedures). Transposition was monitored by transforming the deproteinated reaction products into highly competent *E.coli* DH5α cells. The donor plasmid, which contains the Tn7 element, possesses a conditional origin of replication that will not replicate in *E. coli* cells that do not express the π protein (32). Therefore, chloramphenicol resistant colonies will only result if the Tn7 element (carrying the chloramphenicol resistance cassette) moves into the target plasmid during the *in vitro* transposition reaction.

An initial test to determine if the gapped plasmid could be used as a target and recovered following *in vitro* transposition was necessary. The cell must repair multiple gaps in the DNA, one created intentionally as described above and two created by the transposition reaction (5). To ensure that I could recover the plasmids, I carried out *in vitro* transposition reactions using purified components of the TnsABC core transposition machinery containing a mutant form of TnsC (TnsC^{A225V} or TnsC*) which is still sensitive to targeting signals, but does not require the TnsD or TnsE proteins *in vivo* or *in vitro* (36, 37). I found that I could readily monitor TnsABC* transposition using this new assay, and the 20 bp gap that was constructed into these DNAs could be repaired by the host following transformation (Figure 2.1.A., data not shown). There was no apparent

difference in the recovery of gapped versus non-gapped DNAs in the transformation assay (Figure 2.1.A.). As expected, the wild-type TnsABC control yielded no detectable transposition events (Figure 2.1.A.)

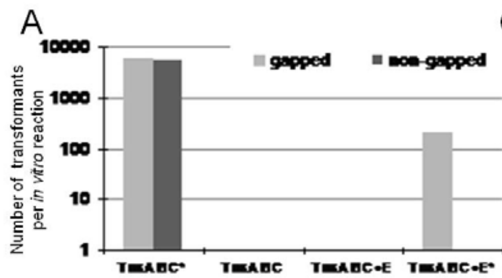
I assayed transposition with purified wild-type TnsABC+E proteins using gapped and ungapped DNA substrates. I also monitored transposition using hyperactive mutants of TnsE (TnsE*) that were isolated in previous work, all of which were shown to strictly obey the same target site preference for DNA replication and the orientation bias with DNA replication found with the wild type protein *in vivo* (Peters and Craig, 2001a). The mutant proteins allow a ~300 to ~1000-fold increase in transposition when compared to the wild type protein *in vivo* (Peters and Craig, 2001a). I found that TnsABC+E transposition could be reconstituted *in vitro*, but that this process required the use of a gapped DNA substrate and could only be detected using hyperactive TnsE mutants (Figure 2.1.A. and data not shown). In multiple trials of this experiment no transposition events were ever detected unless the target DNA contained the single-strand gap (Figure 2.1.A.). The dependence on increased-activity mutant TnsEs could indicate that the signal is simply too low with the wild type protein to be detected in our current assay, or that the mutant proteins no longer require some factor needed by the wild type protein. Sequencing of DNA adjacent to both of the transposon ends, in plasmids isolated from chloramphenicol resistant colonies, confirmed that transposition had occurred as indicated by the presence of the characteristic 5 bp duplication found with Tn7 transposition (5). As expected I found that transposition with TnsABC* was random across the gapped DNA substrates with respect to position and orientation (see below,

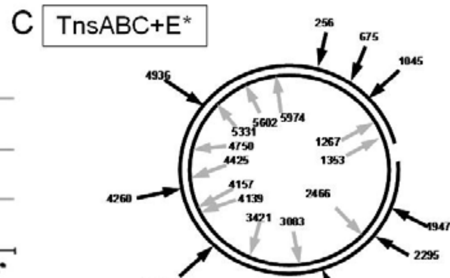
Figure 2.1.B., data not shown). Interestingly, I found that *in vitro* transposition with TnsABC+E* also appeared to occur randomly in the gapped DNA substrate (Figure 2.1.C.). This was somewhat surprising given the absolute bias with DNA replication found *in vivo*, but was consistent with the separation of activation and targeting found in other Tn7 *in vitro* reactions (See Discussion)(31).

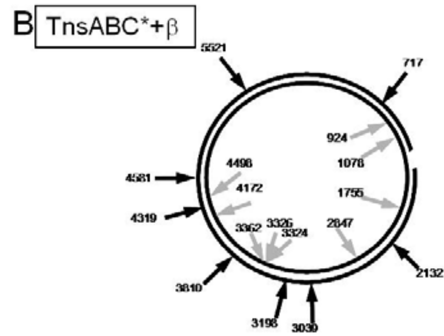
To determine the effect of β on TnsE-mediated transposition *in vitro*, the β clamp was loaded onto the DNA structure using a purified minimal clamp-loader system ($\gamma_3\delta\delta'$). The gapped structure presents a site in the DNA that can be used to load β onto the DNA substrate (17). Clamps are expected to be loaded onto the gapped DNA with a single orientation that is dependent on which strand of DNA contains a free 3' end (14), and may be preferentially retained at the 3' end through interactions with the DNA (10). I loaded the β-clamp onto the DNA substrates using a 20-fold molar excess of β with respect to DNA, then purified the β-loaded structures away from the clamp-loader components and free β proteins (Experimental Procedures) (17). I did not detect significant increases in TnsE-mediated transposition using the β-loaded gapped DNA structures (data not shown). However, upon mapping the location and orientation of transposon insertions, I found a dramatically different pattern than that exhibited by the gapped-DNA substrate without β (compare Figures 2.1.C. and 2.1.D.). I found that 80% of the insertion events occurring in the β-loaded substrate were in a single orientation, and that 40% of them were found at a single base pair junction (at position 1481), proximal to the location of the single-stranded DNA gap (66 bp from the 3' end). These insertions were isolated from separate transformations, ensuring that they are independent

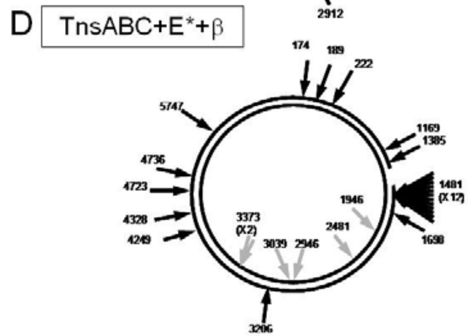
Figure 2.1. An *in vitro* TnsABC+E transposition assay shows that DNA structures are important for TnsE-mediated transposition, and that the β clamp is required to direct the orientation of transposon ends.

A. A representative bar graph representing the number of transposon insertions recovered per in vitro reaction shows that the presence of a single-stranded gap in the target DNA substrate is essential for activation of TnsE-mediated transposition, but shows no effect on the untargeted TnsABC* pathway. B. Transposition via the TnsABC* pathway into the β-loaded gapped substrate indicates that there are no indirect affects of β or gapped DNA on the core transposition machinery that could account for the results observed in TnsABC+E reactions. Black circles represent the target DNA, with a gap in one circle representing the location of the ssDNA gap. Black arrows on the outside of the DNA represent the left-to right orientation of the transposon ends with respect to the free 3' end of the gapped DNA, while gray arrows inside the circles represent the opposite (right-to-left) orientation. The position of each insertion is given next to the arrow. The ssDNA gap resides between positions 1415-1435. C. *In vitro* reactions containing TnsABC+E* and gapped DNA alone. Transposition events mapped in the target DNA appear to be randomly distributed, and with no particular orientation of transposon ends. D. Reactions containing TnsABC+E* and β clamp loaded onto the gapped DNA substrate display a rearrangement of transposon insertions. Many insertions were found at a single base pair junction 66 bases from the free 3' end of the ssDNA gap (position 1481), and occurred in a single left-to-right orientation at that site. Other insertions were found throughout the target DNA, almost all in the left-to-right orientation.


A
10000
1000
100
10
1
Number of transformants per in vitro reaction
gapped
non-gapped
TnsABC*
TnsABC
TnsABC+E
TnsABC+E*


C
TnsABC+E*


B
TnsABC*+β


D
TnsABC+E*+β

transposition events and not siblings. The presence of a single site where many insertions were recovered is reminiscent of TnsD-mediated insertion into *attTn7* (5) and the influence of triplex DNA structures on TnsABC* transposition (30)(see Discussion). When I monitored TnsABC* transposition on the same β-loaded gapped DNA substrates I found no change in the distribution or orientation of insertion events from what was found in gapped substrates without β; the insertions were found in both orientations without any preference for any position in the plasmid (Figure 2.1.B.). Significantly, the TnsABC+E* insertions are consistent with *in vivo* data with respect to the orientation of transposon end alignment with the 3' end of the nascent lagging-strand. While β may not be necessary to activate transposition *in vitro*, its presence on DNA is required to recapitulate the exact targeting activity observed with TnsE-mediated transposition *in vivo* (27).

2.4. Discussion

Our results indicate that there are two critical features recognized by TnsE during DNA replication: a specific DNA structure and the processivity clamp on DNA. While structural components are needed for activating TnsE-mediated transposition *in vitro*, the presence of the β clamp processivity factor is essential for specifically redirecting TnsE-mediated transposition events in a manor that recapitulates the profile found *in vivo* with DNA replication. We conclude that Tn7 likely evolved the capacity to interact with the processivity factor as a way of recognizing discontinuous DNA replication displayed by mobile plasmids.

Tn7 participates in multiple transposition pathways that have been established *in vitro* (1, 36). These systems provide unique perspectives on the relationship between activation and target-site selection of transposons, and highlight the importance of target DNA structure and interaction with host factors for Tn7 transposition. Multiple mutant TnsC alleles can allow Tn7 transposition that does not require either the TnsD or TnsE protein for activation (37). TnsABC* transposition events appear to occur randomly *in vitro* and *in vivo* (2, 28, 37)(Figure 2.1.B. and data not shown). Rao and Craig showed that random TnsABC* transposition events can be redirected *in vitro* within target DNA molecules to a hotspot adjacent to a pyrimidine triplex (30, 31). While triplex forming DNAs cannot activate the wild type protein, they are sufficient to redirect active TnsABC* complexes to a specific hotspot in target DNAs. This work strongly suggested that a structure induced in target DNAs is an important component of Tn7 target recognition and that the activation and targeting signals are separable. It is also thought that the primary role of the TnsD protein in targeting transposition events into the *attTn7* site involves the creation of a distortion in the DNA adjacent to its binding site in the *glmS* gene (15).

While activation of TnsABC+E transposition requires a structure in the target DNA, a separate signal is required to target transposition events, similar to what was found in the TnsABC* pathway with triplex DNA target structures (30). In the *in vitro* transposition experiments, clamps were loaded onto target DNAs in a single orientation, which was determined by the strand of DNA that contained the gap (14). Transposition events received by the clamp-loaded DNA substrates were predominantly in a single orientation

and at a site in the DNA near where clamps are expected to transiently reside (Figure 2.1.D(10)).

The dependence of TnsABC+E transposition on interaction with β and a DNA structure allows Tn7 to modulate the activity of this transposition pathway depending on the replication status of the target DNA. Given that β clamps and single-stranded DNA gaps accumulate on the DNA strand that is replicated discontinuously (14), the TnsE-β interaction and DNA structure specific activation helps explain how TnsE directs transposon insertion to the conjugal DNA replication process. The nature of conjugal DNA replication may leave β clamps especially vulnerable to TnsE interaction. Conjugal plasmids replicate by a discontinuous process that resembles lagging-strand DNA synthesis that occurs during chromosomal replication (44). While processivity clamps are expected to be enriched both on the chromosome and on conjugal plasmids (14), a notable difference is that discontinuous conjugal replication is not physically coupled to continuous replication within the same replisome as it is in chromosomal replication (44). The protein complex that is present at a standard DNA replication fork may limit the exposure of the β clamp and gapped DNA structures more effectively than the uncoupled DNA replication found in conjugal replication.Our data suggest that TnsE has evolved to not interrupt the exchange of proteins on clamps during normal DNA replication when expressed at moderate concentrations. The distribution of TnsE-mediated insertions in the chromosome (27) is likely explained by the ability of TnsE to interact with β clamps that are still topologically linked to DNA, but not actively involved in chromosomal DNA

replication. For example, in cells lacking mobile plasmids TnsE-mediated insertions occur in the region where DNA replication terminates and sites proximal to repaired DNA double-strand breaks (28, 34).

The processivity factor appears to play a pivotal role in the coordination of activity at the replication fork.TnsE likely binds to the same face of the clamp as MutS, Ligase, Pol III and others (14, 21, 35). The presence and importance of β clamp binding motif in TnsE suggests that interaction with the clamp occurs at least in part through the same hydrophobic pocket on the C-terminal face that appears to be involved in coordination of protein-protein interactions. The orientation bias and location of transposition events we observed *in vitro* were consistent with interaction with the C-terminal face of the clamp, which is expected to preferentially reside at the 3' junction of the gap in the target DNA (10)(Figure 2.1.D.). Clamps are also free to slide about the DNA (14, 16), which may explain insertions that occurred into other parts of the target DNA, yet in an orientation that appears to be dictated by the clamp. The SOS induction phenotype observed with TnsE over-expression further supports the notion that TnsE interacts with the clamp on the same face as host proteins involved in DNA replication and repair. The TnsABC+E transposition complex may use a similar mechanism for detecting strand polarity as has been suggested for mismatch repair systems (MMR) in Eukaryotes and in some Prokaryotes. The ability to interact with only one face of the processivity factor has been suggested to allow strand discrimination in MMR so that newly-replicated DNA containing errors can be selectively removed (13, 35). Based on our *in vitro* transposition experiments, interaction with the β-clamp directs the activity of TnsE in a similarly

directional manner (Figure 2.1.D.), resulting in the orientation bias with replication that we observe *in vivo* with TnsE-mediated insertions (27).

Interaction with the processivity factor may constitute a general mechanism for targeting transposition into actively replicating DNA. The transposase of the inactive *pogo* element, found in *Drosophila*, has been shown to bind to the DNA replication processivity factor (PCNA), but the function of this interaction remains a mystery (41, 42). Transposases of other inactive transposons that are abundant in humans (*tigger* elements, estimated to be present at ~3000 copies) and in *Arabidopsis* (*lemi1* elements) also possess putative PCNA binding motifs (41, 42). Because none of these elements are active, determination of the functional relevance of their interaction with the processivity factor is not possible. We suggest that TnsE has evolved that ability to identify the β clamp as a mechanism for targeting processing events found during the mobilization of plasmids. A wide range of transposable elements may use a similar mechanism to target DNA replication and/or DNA repair. While mechanistically very different from Tn7, the transposon Tn917 displays target selection profile that resembles that of Tn7 transposition with the TnsABC+E pathway in the chromosome (9). The single polypeptide transposase of Tn917 contains an amino acid sequence (QLCLAR) that resembles the β clamp binding motif described in this work. In plants, the transposase of the *Ac* element has been shown to be stimulated by active DNA replication (4) and contains the sequence QKRIVGFF (A. Parks and J. Peters, unpublished observation), similar to many previously reported PCNA interaction motifs, or PIP-boxes, with the consensus sequence QxxIxxFF (41).

For Tn7, the interaction with the processivity factor appears to be primarily used to activate transposition directed into mobilized plasmids, providing Tn7 with a means of moving to a new host. Since Tn7-like elements are found in a wide variety of hosts (24), TnsE-mediated transposition shows promise as a new tool for probing the mechanisms and evolution of genetic processes involving processivity factors.

2.5. Experimental Procedures

2.5.1. Plasmids and Strains

Plasmid cloning and propagation was conducted in *E. coli* strain DH5α. Ultra competent cells of *E. coli* DH5α were made by using the Inoue method (26) with minor changes. pGAP plasmid was constructed from a pGEM-T cloning vector containing the *attTn7* locus (pGEM-*attTn7*) by insertion of a fragment containing two recognition sites (20 bp apart from each other) for the nicking enzyme Nb. BbvCI (NEB), which was generated by overlapping PCR with primer pairs JEP247/248 (JEP247: 5'-AAT CTC CCT CCC ACA AGC AGT AAC-3'; JEP248: 5'- CCT CAG CAA TGC TTT CAC CAC CTC AGC TAT CCG CGG TAT TCC AGA CGA-3') and JEP249/250 (JEP249: 5'-GCT GAG GTG GTG AAA GCA TTG CTG AGG AGC TGA AAC AAG GCG GGA CTC-3'; JEP250: 5'-CGC GCA CCA GAG AAG AAC CC-3') using pGEM-*attTn7* plasmid DNA as template.

2.5.2. Protein purification and labeling

His-6-tagged TnsE proteins were purified as described (27). TnsA, TnsB, TnsC, and $TnsC^{A225V}$ were purified as described in (3). A modified β protein that could be labeled with ^{32}P phosphate was purified using the previously described method and vector (18).

The δ, δ', and γ proteins were each purified using the IMPACT system from NEB according to manufacturers recommendations. Cleaved proteins were eluted and dialyzed in storage buffer (20 mM Tris (pH 7.5), 100 mM KCl, 20% glycerol, 0.5 mM EDTA, 1mM DTT).

The minimal clamp-loader γ complex ($\gamma_3\delta\delta'$) (12)was assembled from individually purified proteins in clamp-loading buffer (20 mM Tris-HCl (pH 7.5), 4% glycerol, 8mM $MgCl_2$, 1mM ATP, 2mM DTT, 0.1 mM EDTA) and further purified away from monomers by size exclusion chromatography on a Superdex G-200 column (GE Healthcare). Fractions were assayed for clamp loading activity as described in (17) using nicked pGAP plasmid and ^{32}P-β and the most active fractions were used in experiments requiring β-loaded DNA.

2.5.3. Preparation of target DNA substrates.

The gapped substrate was made as described in (40). Briefly, supercoiled pGAP plasmid containing two Nb. BvcCI nicking sites separated by 20 bp was digested with Nb. BbvCI enzyme (NEB). A competitor DNA oligo (JEP 345: 5'-GCA ATG CTT TCA CCA CCT CA-3')) complementary to the sequence flanked bythe nickwas then added in 50-fold molar excess and incubated at 85°C for 10 minutes to displace the 20 nucleotide fragment. After reannealing, gapped plasmids were purified away from the 20-mer duplex DNA and the excess single-stranded oligomers using Amicon unitra-4 (cutoff MWCO 100 kDa) centrifugal filter (Millipore).

The β-clamp was loaded onto gapped or nicked pGAP plasmid as described in (17) by incubation of minimal clamp-loader (1 pmol), β (25 pmol as β dimer), and gapped pGAP

plasmid (1.25 pmol) in clamp-loading buffer (20 mM Tris-HCl (pH 7.5), 4% glycerol, 8mM $MgCl_2$, 1mM ATP, 2mM DTT, 0.1 mM EDTA). The β-loaded DNA was purified away from clamp-loader and free β-clamp by size exclusion chromatography on a 2 ml 4% agarose bead column (MP biomedicals) in gel-filtration buffer (20 mM Tris-HCl (pH 7.5), 4% glycerol, 100 μg/ml BSA, 2mM DTT). 200-μl fractions were collected and two peak fractions containing β-loaded DNA were pooled and used for *in vitro* transposition assays.

2.5.4. In vitro transposition assay

In vitro transposition reactions contained 26 mM HEPES (pH 7.6), 15 mM Tris (pH 7.6), 4% glycerol, 50 μg/ml bovine serum albumin, 2 mM ATP (pH 7.0), 1.5 mM dithiothreitol, 100 μg/ml tRNA, 0.05 mM EDTA, 8.3 mM NaCl, 9.4 mM KCl, 0.06 mM $MgCl_2$,15 mM magnesium acetate, 0.2 nM of the donor plasmid pGPS2.1 (NEB), 3.2 nM of recipient plasmid, 12.5 nM TnsA, 3 nM TnsB, 5 nM $TnsC^{A225V}$ or $TnsC^{wt}$, and 3.6 nM of TnsE in a final volume of 100 μl. All ofthe reaction mixture components except TnsA, TnsB, magnesium acetate, and the donor DNA were preincubated for 20 min at 30°C. The final reaction mixture components were then added, and the reaction was allowed to proceed for an additional 45 min at 30°C. The reaction was stopped by phenol:chloroform extraction and the DNA was ethanol precipitated and resuspended in 40 μl H_2O. The DNA was then assayed for transposition by transforming *E. coli* (DH5α) cells and selecting for chloramphenicol resistance. Plasmids were isolated from chloramphenicol resistant colonies, and the position of the transposition was determined using a primer specific to the left end of the Tn7 element (Table S3). About 50% of the

insertions were sequenced from both ends to ensure that real transposition occurred. Inclusion of sequences from separate transformation pools insured that siblings were not confused with independent insertions.

2.6. Acknowledgements

This work was supported by NIH grant R01 GM069508. We thank Mark Sutton and Bik Tye for strains and plasmids and helpful discussions. We also thank Eric Alani, Bik Tye, Jeff Roberts, and the members of the Peters laboratory for comments on the manuscript.

REFERENCES

1. **Bainton, R. J., K. M. Kubo, J. N. Feng, and N. L. Craig.** 1993. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell **72:**931-43.

2. **Biery, M. C., F. J. Stewart, A. E. Stellwagen, E. A. Raleigh, and N. L. Craig.** 2000. A simple in vitro Tn7-based transposition system with low target site selectivity for genome and gene analysis. Nucleic Acids Res **28:**1067-77.

3. **Bordi, C., B. G. Butcher, Q. Shi, A. B. Hachmann, J. E. Peters, and J. D. Helmann.** 2008. In vitro mutagenesis of Bacillus subtilis by using a modified Tn7 transposon with an outward-facing inducible promoter. Appl Environ Microbiol **74:**3419-25.

4. **Chen, J., I. M. Greenblatt, and S. L. Dellaporta.** 1992. Molecular analysis of Ac transposition and DNA replication. Genetics **130:**665-76.

5. **Craig, N. L.** 2002. Tn7, p. 423-456. *In* N. L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz. (ed.), Mobile DNA II. ASM Press, Washington, DC.

6. **Dalrymple, B. P., K. Kongsuwan, G. Wijffels, N. E. Dixon, and P. A. Jennings.** 2001. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. Proc Natl Acad Sci U S A **98:**11627-32.

7. **Einarson, M. B., E. N. Pugacheva, and J. R. Orlinick.** 2007. Far Western: probing membranes. CSH Protoc **2007:**pdb prot4759.

8. **Fields, S., and O. Song.** 1989. A novel genetic system to detect protein-protein interactions. Nature **340:**245-6.

9. **Garsin, D. A., J. Urbach, J. C. Huguet-Tapia, J. E. Peters, and F. M. Ausubel.** 2004. Construction of an Enterococcus faecalis Tn917-mediated-gene-disruption library offers insight into Tn917 insertion patterns. J Bacteriol **186:**7280-9.

10. **Georgescu, R. E., S. S. Kim, O. Yurieva, J. Kuriyan, X. P. Kong, and M. O'Donnell.** 2008. Structure of a sliding clamp on DNA. Cell **132:**43-54.

11. **Jeruzalmi, D., M. O'Donnell, and J. Kuriyan.** 2001. Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III. Cell **106:**429-41.

12. **Jeruzalmi, D., O. Yurieva, Y. Zhao, M. Young, J. Stewart, M. Hingorani, M. O'Donnell, and J. Kuriyan.** 2001. Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. Cell **106:**417-28.

13. **Jiricny, J.** 2006. The multifaceted mismatch-repair system. Nat Rev Mol Cell Biol **7:**335-46.

14. **Johnson, A., and M. O'Donnell.** 2005. Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem **74:**283-315.

15. **Kuduvalli, P. N., J. E. Rao, and N. L. Craig.** 2001. Target DNA structure plays a critical role in Tn7 transposition. Embo J **20:**924-32.

16. **Laurence, T. A., Y. Kwon, A. Johnson, C. W. Hollars, M. O'Donnell, J. A. Camarero, and D. Barsky.** 2008. Motion of a DNA sliding clamp observed by single molecule fluorescence spectroscopy. J Biol Chem **283:**22895-906.

17. **Leu, F. P., M. M. Hingorani, J. Turner, and M. O'Donnell.** 2000. The delta subunit of DNA polymerase III holoenzyme serves as a sliding clamp unloader in Escherichia coli. J Biol Chem **275:**34609-18.

18. **Leu, F. P., and M. O'Donnell.** 2001. Interplay of clamp loader subunits in opening the beta sliding clamp of Escherichia coli DNA polymerase III holoenzyme. J Biol Chem **276:**47185-94.

19. **Lopez de Saro, F., R. E. Georgescu, F. Leu, and M. O'Donnell.** 2004. Protein trafficking on sliding clamps. Philos Trans R Soc Lond B Biol Sci **359:**25-30.

20. **Lopez de Saro, F. J., M. G. Marinus, P. Modrich, and M. O'Donnell.** 2006. The beta sliding clamp binds to multiple sites within MutL and MutS. J Biol Chem **281:**14340-9.

21. **Lopez de Saro, F. J., and M. O'Donnell.** 2001. Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. Proc Natl Acad Sci U S A **98:**8376-80.

22. **McKown, R. L., C. S. Waddell, L. K. Arciszewska, and N. L. Craig.** 1987. Identification of a transposon Tn7-dependent DNA-binding activity that recognizes the ends of Tn7. Proc Natl Acad Sci U S A **84:**7807-11.

23. **Parks, A. R., Z. Li, Q. Shi, R. M. Owens, M. M. Jin, and J. E. Peters.** 2009. Transposition into replicating DNA occurs through interaction with the processivity factor. Cell **138:**685-95.

24. **Parks, A. R., and J. E. Peters.** 2009. Tn7 elements: engendering diversity from chromosomes to episomes. Plasmid **61:**1-14.

25. **Parks, A. R., and J. E. Peters.** 2007. Transposon Tn7 is widespread in diverse bacteria and forms genomic islands. J Bacteriol **189:**2170-3.

26. **Peters, J. E.** 2007. Gene transfer - Gram-negative bacteria. . *In* C. A. Reddy, T. J. Beveridge, J. A. Breznak, G. Marzluf, T. M. Schmidt, and L. R. Snyder (ed.), Methods for General and Molecular Microbiology, 3 ed. ASM press, Washington D.C.

27. **Peters, J. E., and N. L. Craig.** 2001. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes Dev **15:**737-47.

28. **Peters, J. E., and N. L. Craig.** 2000. Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates. Mol Cell **6:**573-82.

29. **Peters, J. E., and N. L. Craig.** 2001. Tn7: smarter than we thought. Nat Rev Mol Cell Biol **2:**806-14.

30. **Rao, J. E., and N. L. Craig.** 2001. Selective recognition of pyrimidine motif triplexes by a protein encoded by the bacterial transposon Tn7. J Mol Biol **307:**1161-70.

31. **Rao, J. E., P. S. Miller, and N. L. Craig.** 2000. Recognition of triple-helical DNA structures by transposon Tn7. Proc Natl Acad Sci U S A **97:**3936-41.

32. **Shafferman, A., R. Kolter, D. Stalker, and D. R. Helinski.** 1982. Plasmid R6K DNA replication. III. Regulatory properties of the pi initiation protein. J Mol Biol **161:**57-76.

33. **Sharpe, P. L., and N. L. Craig.** 1998. Host proteins can stimulate Tn7 transposition: a novel role for the ribosomal protein L29 and the acyl carrier protein. Embo J **17:**5822-31.

34. **Shi, Q., A. R. Parks, B. D. Potter, I. J. Safir, Y. Luo, B. M. Forster, and J. E. Peters.** 2008. DNA damage differentially activates regional chromosomal loci for Tn7 transposition in Escherichia coli. Genetics **179:**1237-50.

35. **Simmons, L. A., B. W. Davies, A. D. Grossman, and G. C. Walker.** 2008. Beta clamp directs localization of mismatch repair in Bacillus subtilis. Mol Cell **29:**291-301.

36. **Stellwagen, A. E., and N. L. Craig.** 2001. Analysis of gain-of-function mutants of an ATP-dependent regulator of Tn7 transposition. J Mol Biol **305:**633-42.

37. **Stellwagen, A. E., and N. L. Craig.** 1997. Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7. Genetics **145:**573-85.

38. **Stellwagen, A. E., and N. L. Craig.** 1998. Mobile DNA elements: controlling transposition with ATP-dependent molecular switches. Trends Biochem Sci **23:**486-90.

39. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:**4673-80.

40. **Wang, H., and J. B. Hays.** 2001. Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions. Mol Biotechnol **19:**133-40.

41. **Warbrick, E.** 2000. The puzzle of PCNA's many partners. Bioessays **22:**997-1006.

42. **Warbrick, E., W. Heatherington, D. P. Lane, and D. M. Glover.** 1998. PCNA binding proteins in Drosophila melanogaster : the analysis of a conserved PCNA binding domain. Nucleic Acids Res **26:**3925-32.

43. **Wijffels, G., B. P. Dalrymple, P. Prosselkov, K. Kongsuwan, V. C. Epa, P. E. Lilley, S. Jergic, J. Buchardt, S. E. Brown, P. F. Alewood, P. A. Jennings, and N. E. Dixon.** 2004. Inhibition of protein interactions with the beta 2 sliding clamp of Escherichia coli DNA polymerase III by peptides from beta 2-binding proteins. Biochemistry **43:**5661-71.

44. **Wilkins, B. M., and E. Lanka.** 1993. DNA processing and replication during plasmid transfer between gram-negative bacteria. *In* D. B. Clewell (ed.), BACTERIAL CONJUGATION. Plenum Press, New York.

45. **Wolkow, C. A., R. T. DeBoy, and N. L. Craig.** 1996. Conjugating plasmids are preferred targets for Tn7. Genes Dev **10:**2145-57.

CHAPTER 3

HETEROLOGOUS EXPRESSION REVEALS A MECHANISM FOR TRANSPOSON TN7 HOST SPECIFICITY THROUGH ADAPTATION TO REPLICATION PROCESSIVITY FACTOR[3]

3.1. Summary

Tn7-like transposons are complex elements found in disparate environments in highly divergent bacteria. Despite the broad distribution of the element across diverse bacteria we find that Tn7-like elements from closely related bacterial species are similarly closely related, suggesting there is an unknown mechanism of host adaptation. Since an interaction between the target selecting protein TnsE and the β-clamp processivity factor (DnaN) is crucial for Tn7 targeting conjugal plasmids for transposition, we sought to test if this adaptation limited the interspecies transfer of the broadly adapted Tn7-like elements. We attempted to functionally replace DnaN in *Escherichia coli* with DnaN homologs from three hosts of Tn7-like elements through the use of an *E. coli* strain with a temperature sensitive DnaN mutation (*dnaN159*). We found that β-clamps from *Idiomarina loihiensis* and *Shewanella baltica,* but not processivity factor from *Bacillus cereus,* were able to suppress all of the known phenotypes of the *dnaN159* allele in *E. coli*. Interestingly, strains where a foreign processivity factor functionally replaced the endogenous protein were dependent on the accessory DNA polymerase II and also required RecA for robust growth. Significantly, TnsE-mediated transposition was only

[3]This Chapter has been submitted for publication with Peters, J.E.

found when TnsE and DnaN were transplanted into *E. coli* from the same foreign strain, strongly suggesting that Tn7-like elements must adapt to the host processivity factor. This also suggests that even elements more distantly related to Tn7 use the TnsE protein to interact with processivity factor and likely target DNA replication. The results from a structural analysis of processivity factor proteins across bacteria are consistent with a model where TnsE has adapted to recognize the more variable portion of the clamp to prevent interfering with polymerase interactions.

3.2. Introduction

Transposons are genetic elements that can move between locations within a genome without the need of sequence homology between the element and target site. Transposons are found in essentially all organisms where they often make up a significant portion of the host genome. Transposons impact the host through gene inactivation, changes in the expression pattern of genes, as vectors for driving gene-fusion events, and as regions of homology that drive rearrangements via homologous DNA recombination. Transposons play an especially important role in bacteria through the ability to transfer genes between organisms. While transposons are nearly ubiquitous, there are few studies that directly examine factors that control the distribution of elements across divergent hosts. We examined factors that limited the distribution of the bacterial transposon Tn7, a complex element that shows a wide distribution across divergent hosts.

Tn7 is a sophisticated bacterial transposon that exploits different life styles for its propagation and dissemination by utilizing distinct but overlapping sets of the five proteins it encodes: TnsA, TnsB, TnsC, TnsD, and TnsE [1-5]. TnsABC proteins

constitute the core transposition machinery that is required, but not sufficient for all transposition. TnsA and TnsB proteins together are the transposase responsible for the DNA breakage and joining activities underlying transposition [6,7]. TnsC is the regulator that activates the transposase by conveying a signal presented by the target selector protein TnsD or TnsE after it engages a specific kind of target for transposition [1,3,8-10]. In the TnsABC+D pathway, Tn7 directs transposition into the highly conserved *attTn7* site found in the bacterial chromosome [3,11-14]. The *attTn7* site is considered a "safe haven" for Tn7 as insertion into this site has no detectable adverse effect on the host [2,15]. In another pathway, the TnsABC+E pathway,Tn7 preferentially targets plasmids capable of moving between bacteria (called conjugative plasmids) and filamentous bacteriophage for transposition [16-19]. These two pathways allow Tn7 to take advantage of both the stability of the chromosome and the mobility of certain DNAs for its dispersal, i.e. the TnsABC+E pathway would facilitate Tn7 dissemination to new hosts, but the TnsABC+D pathway would ensure its stable maintenance in a species. Indeed, Tn7 and its relatives (Tn7-like elements) have been identified in a wide variety of bacterial hosts from diverse environments [20,21]. However, as we show here, contrary to the widespread nature of Tn7-like elements in diverse bacteria, phylogenetic analysis actually suggests that Tn7-like elements likely have limited regular movement between bacterial species or families.

The ability of TnsABC+E transposition to target DNA replication involves an interaction between TnsE and the β-clamp processivity factor, an essential component of the DNA replication apparatus [22]. TnsE interacts with the β-clamp through a β-binding

motif that shows conservation in TnsE homologs. Mutating any conserved residue of the β-binding motif substantially decreases or abolishes the TnsE-β interaction and TnsE-mediated transposition. In addition, over-expressing the β-clamp in the cell stimulates TnsE-mediated transposition. Since only extremely high levels of TnsE expression induced the SOS response in the cell [22], the TnsE-β interaction is not expected to interrupt the normal traffic on the β clamp (see below).

The β-clamp plays a central role in DNA replication, repair and recombination [23]. The β-clamp is a dimer of the β-subunit of DNA polymerase III holoenzyme (DnaN) that confers processivity on DNA replication by tethering the DNA polymerase to the template DNA. Many proteins involved in DNA replication, repair and recombination use the clamp as a mobile platform to function, including all five DNA polymerases in *E. coli* [24-28], the δ subunit of the clamp loader complex [29], DNA ligase [28], proteins involved in mismatch repair (MutS and MutL)[28,30], and replication initiation control (Hda)[31]. Even though most of these proteins contain a β-binding motif [28,32], unambiguously defining the partner-specific binding interfaces on the β-clamp has been a challenge because of the absolute requirement of β for cell viability [33,34].

Here we report the functional conservation and divergence of various DnaN interactions as revealed by transplanting DnaN from various bacteria containing Tn7-like elements into *E. coli*. DnaN proteins from two *proteobacterialI. loihiensis* L2TR ($DnaN^{Il}$) and *S. baltica* OS155 ($DnaN^{Sb}$) share about 50% identity to their*E. coli* homolog, but DnaN from gram-positive *FirmicuteB. cereus* ATCC 10987 ($DnaN^{Bc}$) is highly diverged. We found that only the host DnaN ($DnaN^{Ec}$), $DnaN^{Il}$*, and* $DnaN^{Sb}$ were able to

complement all known phenotypes of the well-characterized *dnaN159* allele in *E. coli,* but DnaN^{Bc} was toxic. Interestingly, in *dnaN159* strains with either DnaN^{Il} or DnaN^{Sb}, DNA polymerase II is important for cell viability and the limited replication found in the absence of Pol II is highly mutagenic. We also found that RecA was important for the viability of *dnaN159* strains with either DnaN^{Il} or DnaN^{Sb}. Significantly, we found that TnsE-β interaction was species specific; TnsE of a Tn7-like element only worked with the Tn7 core transposition machinery when DnaN from the same host of the element was also being used in *E. coli*. These results indicate that despite divergence between Tn7-like elements, TnsE-β interaction is largely conserved and TnsE may have adapted to the more variable portion of the clamp to prevent interfering with polymerase interactions.

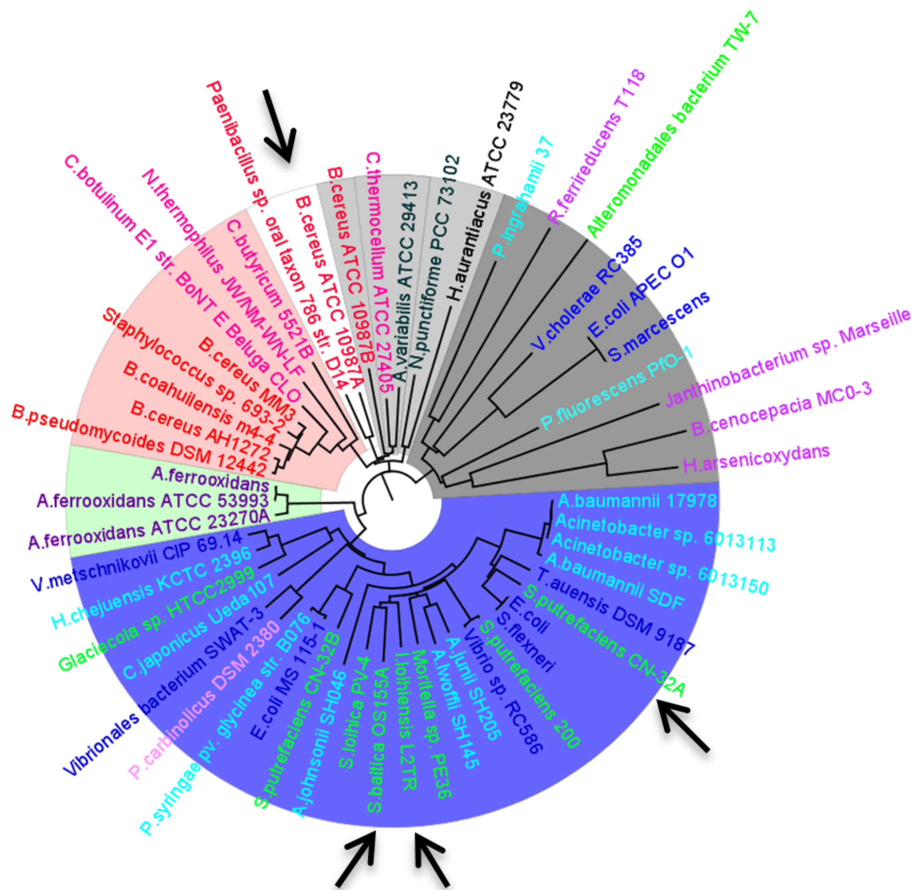
3.3. Results

3.3.1. Phylogenetic distribution of Tn7-like elements

The ability of Tn7-like elements to access a neutral position in the chromosome and various mobile DNAs of broad host-range for transposition would presumably maximize its dissemination. The original Tn7 element was isolated in an agricultural setting and elements that are essentially identical to Tn7 are regularly identified in medical settings and isolated in animals reared for agriculture [18,21]. As might be expected, Tn7-like elements have been identified in a wide variety of bacteria [21]. The hosts of Tn7-like elements can be found in diverse environments, ranging from deep sea hydrothermal vents to surface waters of the oceans, from various terrestrial environmental samples to spoiled food, and clinical settings [20,21]. In all these elements, open reading frames that

encode proteins homologous to TnsABCDE are in an array that is in synteny with the *tnsABCDE* genes in transposon Tn7, even though the sequence of each homologous gene has diverged from the one in Tn7. Of the 55 Tn7-like elements examined (Appendix 1), more than 75% are predicted to be active for transposition by both pathways since no insertion or non-sense mutations have occurred in the any of *tns* genes. The majority of Tn7-like elements are found in the chromosomal *attTn7* site (75%), and more rarely in non-*attTn7* chromosomal locations (13%) and plasmids (6%). This is consistent with the idea that TnsD-mediated transposition into the neutral chromosomal site is important for stable maintenance of Tn7-like elements in bacterial hosts. While Tn7-like elements are not ubiquitously distributed amongst bacterial species, this may be explained by biases in sequence information and the relative prevalence of various genera in the environment. Of the 55 Tn7-like elements identified, 41 elements are in the genomes of *γ-proteobacteria*, 2 in *δ-proteobacteria*, and 12 in low G+C% Gram-positive bacteria (*Firmicutes*). In the phylogenetic tree of concatenated TnsABC protein sequences (Figure 3.1), Tn7-like elements with *tnsABCDE* genes are clustered into two distinct nodes; elements from *Proteobacteria* form one clade, but elements found in *Firmicutes* form another phylogenetically distant clade, indicating that they are derived from different ancestors that are probably separated at the division of *Proteobactera* and *Firmicutes* early in the evolution of bacteria. A group of Tn7-like elements from *Acidithiobacillus* may have resulted from an early split away from the other elements, or its evolution might have been constrained by the extreme lifestyle of its host that is adapted to living at pH2. Clustering of Tn7-like elements from phylogenetically related bacteria is also

Figure 3.1. Phylogenetic tree of Tn7 family. The phylogenetic tree was constructed using concatenated TnsABC protein sequences. Bootstrap analysis was carried out with 1000 iterations (values not shown in the interest clarity). In the branches of the tree, TnsABCDE elements in *Proteobacteria* are highlighted (background) in blue and TnsABCDE elements in *Firmicutes* in pink. The branches of the tree showing the Tn5468 element and closely related elements found in *Acidothiobacillus* are shown in light green background. The taxa were colored as follows: *γ-proteobacteria* orders *Enterobacteriales*, *Vibrionales*, and *Aeromonas* are in dark blue; *Shewanellaceae*, *Alteromonas*, *Idiomarinaceae* are in green; *Pseudomonadaceae* and *Oceanospirillaceae* are in cyan. *Acidothiobacillus* and *β-proteobacteria* are in purple while *δ-proteobacteria*are in pink. *Cyanobacteria* are in dark green and *Chloroflexi* in black. *FirmicutesBacillales* are in red and *Clostridia* in magenta. Black arrows indicate the Tn7-like elements that are genetically studied in this work.



Paenibacillus sp. oral taxon 786 str. D14
B.cereus ATCC 10987B
B.cereus ATCC 10987A
C.thermocellum ATCC 27405
A.variabilis ATCC 29413
N.punctiforme PCC 73102
H.aurantiacus ATCC 23779
P.ingrahamii 37
R.ferrireducens T118
Alteromonadales bacterium TW-7
V.cholerae RC385
E.coli APEC O1
S.marcescens
P.fluorescens PfO-1
Janthinobacterium sp. Marseille
B.cenocepacia MC0-3
H.arsenicoxydans
A.baumannii 17978
Acinetobacter sp. 6013113
Acinetobacter sp. 6013150
A.baumannii SDF
T.auensis DSM 9187
S.putrefaciens CN-32A
E.coli
S.flexneri
S.putrefaciens 200
Vibrio sp. RC586
A.junii SH205
A.lwoffii SH145
Moritella sp. PE36
I.loihiensis L2TR
S.baltica OS155A
S.loihica PV-4
A.johnsonii SH046
S.putrefaciens CN-32B
E.coli MS 115-1
P.syringae pv. glycinea str. B076
P.carbinolicus DSM 2380
Vibrionales bacterium SWAT-3
C.japonicus Ueda107
Glaciecola sp. HTCC2999
H.chejuensis KCTC 2396
V.metschnikovii CIP 69.14
A.ferrooxidans ATCC 23270A
A.ferrooxidans ATCC 53993
A.ferrooxidans
B.pseudomycoides DSM 12442
B.cereus AH1272
B.coahuilensis m4-4
Staphylococcus sp. 693-2
B.cereus MM3
C.botulinum E1 str. BoNT E Beluga
N.thermophilus JW/NM-WN-LF
C.butyricum 5521B

obvious at the genus and species levels, indicating host specificity of Tn7 dissemination (Figure 3.1).

As shown in previous work, we also identified a group of elements that lack a TnsE-encoding gene, but instead maintain two diverged TnsD-encoding genes (TnsABC+D'D elements), in various genomic locations in highly diverged bacteria. In the phylogenetic tree based on concatenated TnsABC protein sequences, most of TnsABCD'D elements clustered in a clade that is separated from the Tn7-like elements, but some are more closely related to Tn7-like elements found in *Firmicutes* (Figure 3.1). This, together with the fact that more protein homologs can be found for TnsAB than TnsC and then TnsD and TnsE, would suggest that a Tn7-like transposon may have started from TnsAB and then gradually acquired the regulatory protein TnsC and the target site selectors TnsD and TnsE.

Since the TnsE-mediated transposition pathway is expected to be responsible for distributing Tn7-like elements into diverse hosts, and a TnsE interaction with the β-clamp is essential for this transposition pathway, we hypothesize that the required adaptation of TnsE to the β-clamp of a new host might be critical in limiting the distribution of a Tn7-like element, i.e. once a Tn7-like element enters the *attTn7* site of a new genome, whether or not it can disseminate further into new species would be dependent on the ability of the TnsE protein to evolve to recognize the β-clamp of the present host. This is supported by our earlier finding that TnsE homologs could not work with TnsABC from Tn7 to allow transposition in *E. coli* (Figure 3.2).

3.3.2. DnaN^{Il} and DnaN^{Sb} functionally complement the previously described *dnaN159* phenotypes

To test the role of the β-clamp in host specificity, multipleTn7-like elements at different branching points in the phylogenetic tree were chosen for study, including Tn7-like elements from two *γ-Proteobacteria, I. loihiensis* and *S. baltica,* and one *Firmicute*, *B. cereus*, in addition to Tn7 in *E. coli* (Figure 3.1 and table 3.1). Given the central role of DnaN in DNA replication and repair it was necessary to determine if the foreign clamps were functional in *E. coli*. Based on gross colony morphology no aberrant phenotypes were noticed in wild type (i.e. *dnaN*$^{+}$) *E. coli* strains with plasmids encoding DnaN^{Il}, DnaN^{Sb} or DnaN^{Ec} at 30^{0}C (data not shown), which is consistent with the observation that increased expression of the β-clamp had little effect on cell viability and morphology [35]. To test if the foreign clamps were participating in DNA replication and repair we utilized a system for functionally replacing DnaN in *E. coli* using the well-characterized temperature sensitive *dnaN* mutant allele, *dnaN159*. The *dnaN159* mutant has two mutations, G66E and G174A, with the G174A mutation at the vicinity of the hydrophobic pocket required for protein-proteins interactions and G66E mutation at the rim of the clamp structure [36]. The mutated β-clamp is labile at high temperature [37] and probably even at 30°C [38]. Furthermore, DnaN159 purified from cells grown at 30°C is far less active than wild type DnaN in restoring Pol III* (Pol III holoenzyme without β) activity and is inert in reconstituting Pol III holoenzyme from purified Pol III core and other components [37], consistent with the finding that DnaN159 is impaired for interaction with the α subunit of DNA polymerase III at 30°C [36]. We determined if

Table 3.1. Homology of TnsE and DnaN homologs.

TnsE homologue	Accession number	Length (aa)	Coverage (%)	Identity (%)	Similarity (%)
$TnsE^{Ec}$	NP_065316	538	100	100	100
$TnsE^{Il}$	YP_156990	436	77	31	49
$TnsE^{Sb}$	YP_001052686	529	98	27	48
$TnsE^{Bc}$	NP_976514	533	44	19	41
DnaN homologue					
$DnaN^{Ec}$	NP_418156	366	100	100	100
$DnaN^{Il}$	YP_154395	367	100	57	71
$DnaN^{Sb}$	YP_001048408	366	100	58	75
$DnaN^{Bc}$	NP_976330	379	99	27	52

Figure 3.2.TnsE homologs are not functional with TnsABC from *E. coli* for transposition in wild type strain.Transposition was monitored usinga promoter capture assay in strain NLC28 with pCW15 (Tn7 TnsABC) and pBAD24-derivatives expressing TnsE homologs or an empty vector. Transposition frequency is indicated by the number of Lac^+ papillae after 72 hours at 30°C on MacConkey's media. Error bars indicate the standard deviation of average papillae number on 12 independent patches.

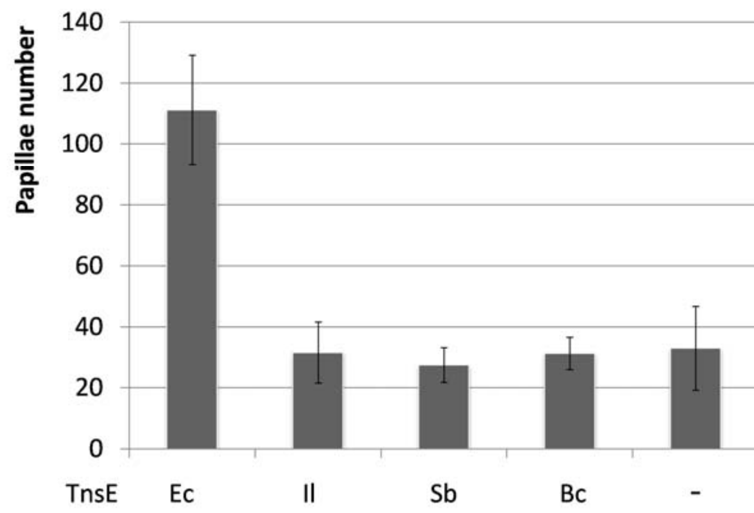
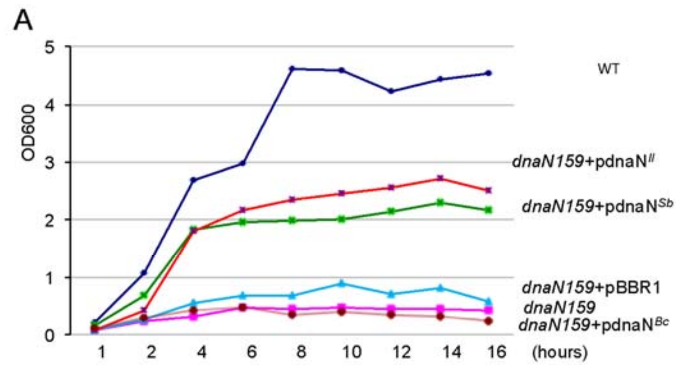
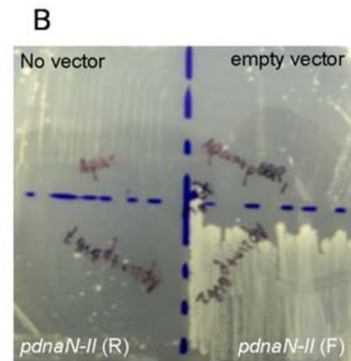
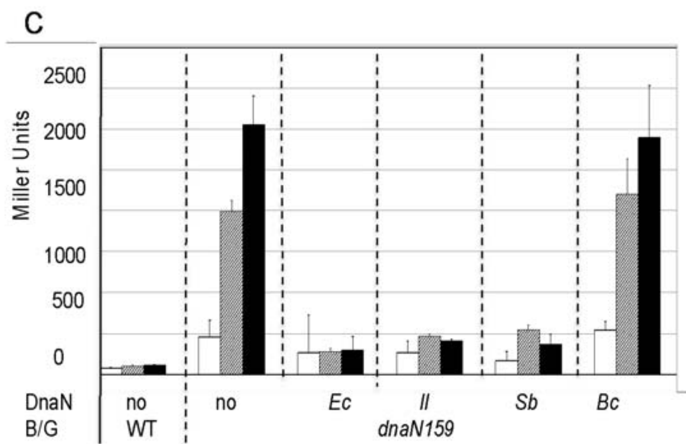
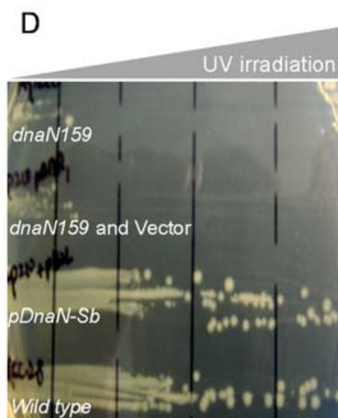

Papillae number
140
120
100
80
60
40
20
0
TnsE
Ec
Il
Sb
Bc
-

Figure 3.3. **DnaN^{Il} and DnaN^{Sb} but not DnaN^{Bc} were able to complement the phenotypes of *dnaN159* allele.** (A) Growth curves of *dnaN*$^{+}$ wild type strain and *dnaN159* strains with pBBR1 plasmid-derivatives expressing various DnaN homologs at 42°C. Cell densities were monitored at the indicated time points after inoculation. (B) Growth of *dnaN159* strain derivatives on Isosensitest agar. Picture was taken after overnight growth at 42°C. Empty vector pBBR1 (Vect.) and *dnaN*^{Il} cloned in the forward (*dnaN*^{Il}(F)) and reverse (*dnaN*^{Il}(R)) orientations indicated. (C) SOS induction in the wild type strain and *dnaN159* strains with or without DnaN homologs. SOS induction was monitored by measuring the β-galactosidase activity in wild type strain (AP330) and ZL123 (NLC28 *dnaN159 sulA*::Kan^{R}::*lacZYA'*) strain derivatives containing pBBR1 plasmid or pBBR1-derivatives expressing DnaN homologs at 30°C (white bars), 37°C (hatched bars), and 42°C (black bars). Error bars indicate the standard deviation of the mean (n=3). B/G indicates strain background. (D) UV sensitivity of *dnaN*$^{+}$ (wild type) strain and *dnaN159* strains with or without DnaN^{Il} expressed in plasmid pBBR1. Cells were streaked across a plate and exposed to UV irradiation for different time intervals as indicated by the vertical lines drawn on the plate. Picture was taken after overnight growth at 30°C.


A
OD600
WT
dnaN159+pdnaN^Il
dnaN159+pdnaN^Sb
dnaN159+pBBR1
dnaN159
dnaN159+pdnaN^Bc
(hours)


B
No vector
empty vector
pdnaN-Il (R)
pdnaN-Il (F)


C
Miller Units
DnaN
B/G
no
WT
no
Ec
Il
Sb
Bc
dnaN159


D
UV irradiation
dnaN159
dnaN159 and Vector
pDnaN-Sb
Wild type

heterologous clamps would be able to complement some of the prominent phenotypesof *dnaN159* strain: temperature sensitivity, UV sensitivity, chronic SOS induction and an absolute dependence on the polymerase activity of Pol I for viability [36,39].

As expected the *dnaN159* strain with vector alone was not viable at 42°C, but strains expressing DnaN from *I. loihiensis* and *S. baltica* progressed through log and stationary phase (Figure 3.3A). No effect on cell viability was seen with the *dnaN159* strain when the heterologous clamps were cloned in the reverse orientation (data not shown). This observation was confirmed on solid medium (Figure 3.3B). We also noticed that *dnaN159* strains expressing DnaN from *I. loiheinsis* or *S. baltica* grew slightly better than the *dnaN159* stain even at 37°C (data not shown). However, the growth of the *dnaN159* strain expressing $DnaN^{Bc}$ was not improved when compared to *dnaN159* strain or *dnaN159* strain with the empty vector (Figure 3.3A). In fact, strains maintaining the $DnaN^{Bc}$ expressing plasmids gave colonies that were obviously smaller than strains with the vector alone or strains expressing any of the other DnaN proteins (data not shown). The results were consistent with the idea that $DnaN^{Il}$ and $DnaN^{Sb}$ were able to form clamps that could be recognized in the cell by both the clamp loader complex and Pol III DNA polymerase, as the β-clamp has to be loaded onto DNA to function.

In addition to its central role during normal DNA replication, the processivity factor also interacts with the three accessory DNA polymerases in cases of DNA damage. To assess whether the foreign processivity factors were competent for coordinating replication on damaged DNA we challenged our strains with UV irradiation. To facilitate UV sensitivity studies we constructed a *dnaN159* strain that also displayed compromised

nucleotide excision repair due to a *uvrC* allele (Materials and methods). In the assay, fresh overnight cultures were streaked across a plate and then exposed to UV irradiation for different time intervals. As expected, the *dnaN159 uvrC* strain was exquisitely UV sensitive (Figure 3.3B). DnaN^{Il} complements the UV sensitivity of *dnaN159* to a level comparable to a strain expressing DnaN^{Ec} (Figure 3.3D). The same result was observed with DnaN^{Sb}, while strains expressing DnaN^{Bc} appeared no different than the *dnaN159 uvrC* strain with the plasmid vector only. Because DNA polymerase III cannot replicate across UV-induced DNA lesions, the results suggest that the processivity factors from *I. loihiensis* and *S. baltica* were competent for DNA replication with one or more of the accessory DNA polymerases involved in translesion synthesis.

In *dnaN159* strains the SOS response is chronically induced even at 30°C indicating the persistent presence of single strand gaps in the cell, likely because of dissociation of Pol III core from the mutant DnaN159 clamp during normal replication process [36]. To further investigate how well the heterologous clamps participated in DNA replication, we monitored SOS induction at three different temperatures. In the assay strain a *lacZ* reporter gene construct interrupts the *sulA* gene, a known SOS-induced gene involved in cell division arrest during the SOS response. As described previously [36], compared to the wild type strain, *dnaN159* strain is chronically SOS induced at 30°C (Figure 3.3C). The SOS response was greater at the semi-permissive (37°C) and non-permissive (42°C) temperatures where a four-fold or six-fold increase in induction was found, respectively (Figure 3.3C). The SOS response is almost completely suppressed by DnaN^{Il} or DnaN^{Sb} in the *dnaN159* background to a level comparable to the effect of supplying the

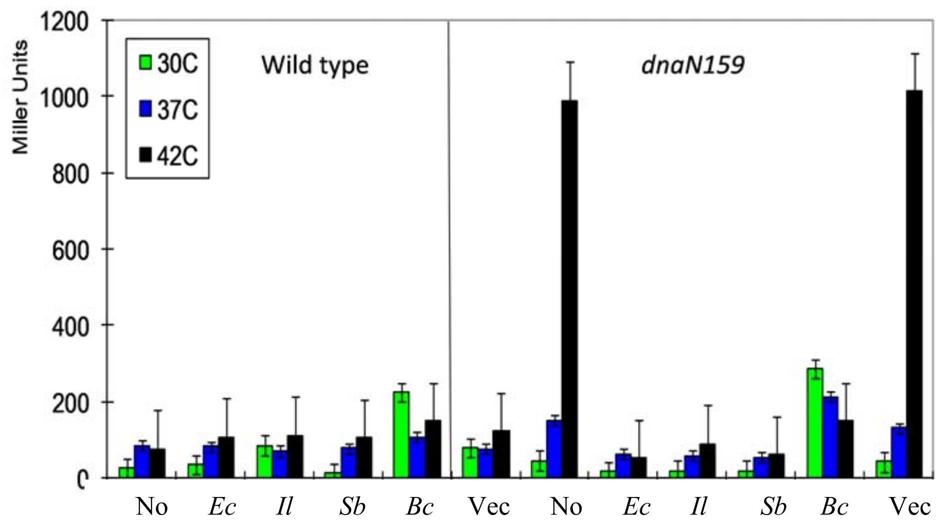
endogenous DnaN^{Ec} *in trans*. As expected from the other finding above, induction of the SOS response was not suppressed by DnaN^{Bc} or with the empty vector.

What appeared to be complete suppression of the SOS response with β-clamps from distantly related bacteria was striking given the sequence divergence of the DnaN^{Il} and DnaN^{Sb} proteins. We confirmed the observation found with the *sulA* fusion with another reporter fusion, *λL::lacZ*, by looking at the activation of phage λ early genes. Like LexA, phage λ suppressor is cleaved upon the activation of RecA, but through a slightly different mechanism [40,41]. The results were in general agreement with the *sulA* reporter fusion, except that overall SOS induction was lower at 37°C and *dnaN*^{Bc} actually induced the SOS response to a higher level than *dnaN159* alone both at 30°C and 37°C (Figure 3.4).

The *dnaN159* allele was previously shown to be synthetically lethal with *polA*, likely because of the persistent single strand DNA gaps indicated by constitutive SOS induction [36]. To test the essentiality of DNA polymerase I in *dnaN159* strain expressing DnaN^{Il} or DnaN^{Sb}, we attempted to move a *ΔpolA::kan* allele into multiple strain backgrounds. While the *ΔpolA::kan*allele was readily transduced into *dnaN159* strains expressing either DnaN^{Il} or DnaN^{Sb} using P1 transduction (Material and methods), this result was not found in *dnaN159* strains with the vector only or expressing DnaN^{Bc}.

Taken together, the ability of *dnaN*^{Il} or *dnaN*^{Sb} to complement the known defects of the *dnaN159* allele indicates that the heterologous clamps (DnaN^{Il} or DnaN^{Sb}) are expressed in *E. coli* and can function with a number of known β-clamp binding proteins at a sufficient level to allow DNA replication that is generally free from the single stand

Figure 3.4. SOS induction of *dnaN* wild type strain and *dnaN159* strains with or without DnaN homologs. SOS response were monitored by measuring the β-galactosidase activity in wild type strain (NK8207) and ZL411 (NK8207 *dnaN159 tnaA300::*Tn10) strain derivatives containing pBBR1 plasmid or pBBR1-derivatives expressing DnaN homologs. Vec indicates vector control. Error bars indicate the standard deviation of the mean (n=3).



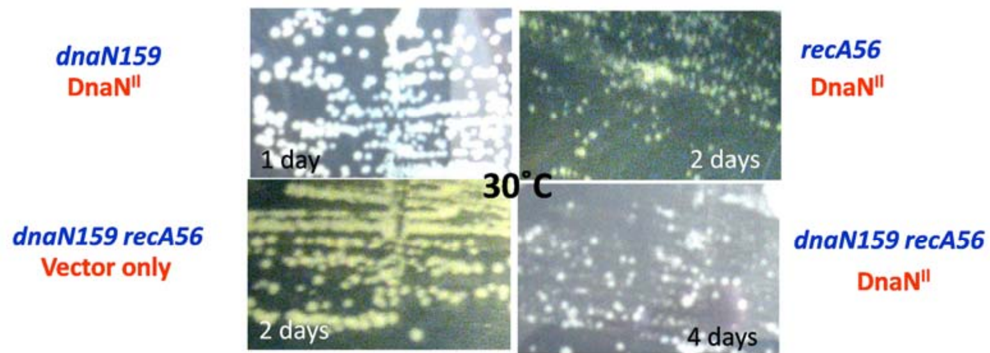
Miller Units
1200
1000
800
600
400
200
0
30C
37C
42C
Wild type
dnaN159
No
Ec
Il
Sb
Bc
Vec
No
Ec
Il
Sb
Bc
Vec

Table 3.2. Colony size of *dnaN159 recA* strains over time.

Days	Vector	*dnaN*^{Ec}	*dnaN*^{Il}	*dnaN*^{Sb}	*dnaN*^{Bc}
1	++	++	-	-	++
2	++++	++++	-	-	++++
3	ND	ND	±	±	ND
4	ND	ND	+/++	+/++	ND

Note: Normal colony size of *E. coli* wild type strain after overnight growth at 30°C was set to be ++++. Barely visible tiny colonies were indicated by ± and. -: no growth; ND: growth not followed after 2 days at 30°C.

Figure 3.5. Growth of *dnaN159*, *recA*, and *dnaN159 recA* mutants with or without heterologous clamps at 30°C on Isosensitest agar.Pictures were taken on days after incubation as indicated in the figure.


dnaN159
DnaN^{II}
1 day
recA56
DnaN^{II}
2 days
30˚C
dnaN159 recA56
Vector only
2 days
4 days
dnaN159 recA56
DnaN^{II}

interruptions found with the *dnaN159* allele. The ability of $DnaN^{Il}$ and $DnaN^{Sb}$ proteins to suppress UV sensitivity also suggests that these clamps can interact with DNA polymerase(s) involved in translesion synthesis.

3.3.3. *dnaN159 recA* strains show severely compromised growth when expressing $DnaN^{Il}$ or $DnaN^{Sb}$

The results shown above would predict that leading-strand and lagging-strand DNA replication should be well coupled in the *dnaN159* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$: in theory, uncoupling would be expected to cause observable SOS induction and a requirement for DNA polymerase I. Surprisingly, we found that *dnaN159* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ were sick when they also had a *recA* mutation (table 3.2 and Figure 3.5). Strains with the *dnaN159* allele grew well at 30°C and formed discrete colonies of similar sizes in 24 hours; *recA* strains were able to form discrete colonies of a universal size within 48 hours. The *dnaN159 recA* double mutant strain was found to be no sicker than the single mutants. However, more than 3 days were required for *dnaN159 recA51* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ to form visible tiny colonies on plates at 30°C. Small colonies were more visible after 96 hours of growth, but they were of different sizes, strongly suggesting the occurrence of suppressor mutations or loss of the plasmid (or plasmid expression). In contrast, no adverse effect was observed with the empty vector control or $DnaN^{Ec}$.

3.3.4. Accessory polymerase usage in *dnaN159* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$

Increased usage of the accessory polymerases has been found in strains where the Pol III-β interaction was disturbed [36,42] or the coordination of DNA replication was

compromised [43,44]. We would expect to see differential DNA polymerase usage if any single substitution or substitutions in combination present in the heterologous clamps affects the clamp binding of different DNA polymerases. To test this possibility, we monitored the colony-forming ability and the overall fidelity of DNA replication of *dnaN159* strains with different clamps when genes encoding DNA polymerases Pol II (*polB*), Pol IV (*dinB*), and Pol V (*umuDC*) were knocked out. To ensure that we had a sensitive assay for the effect of accessory DNA polymerases-mediated replication, nucleotide excision repair was eliminated by using a*uvrC* strain (Materials and methods).

As expected, the *dnaN159* strain grew well at 30°C and there was no obvious effect from expressing either the endogenous clamp or the heterologous clamps. Similarly, deleting the Pol IV or Pol V encoding genes did not result in obvious changes in cell viability. However, inactivating the gene encoding Pol II resulted in a large decrease in cell viability in *dnaN159* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ (Figure 3.6A). Interestingly, a dramatic 5-10-fold increase in mutation frequency was also found specifically in *dnaN159 polB* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ (Figure 3.6B). Such changes were not observed in *dnaN159* strains expressing the host $DnaN^{Ec}$, indicating that the effect came from the clamps themselves and not simply over-expression of the processivity factor.

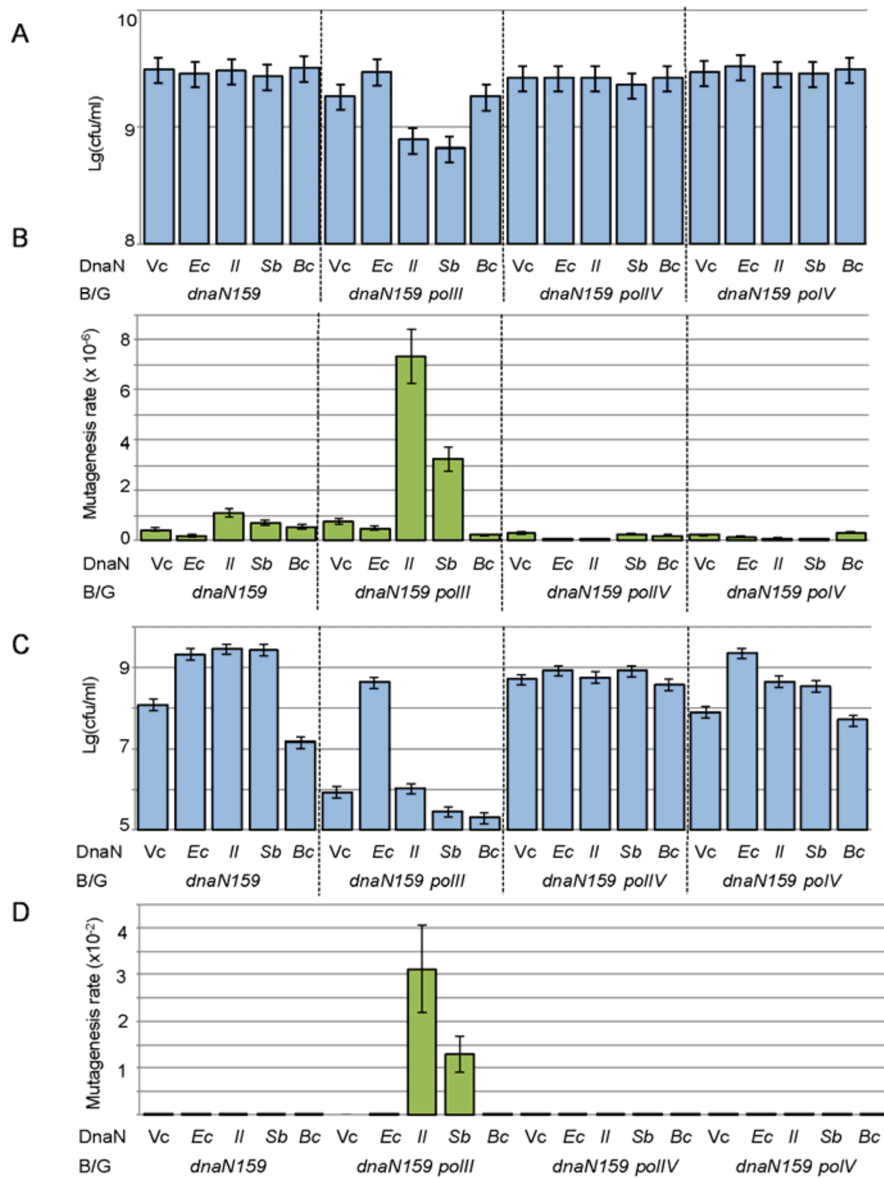
An even more substantial effect was observed at higher temperature. In accordance with the complementation result discussed above, *dnaN159* strains expressing $DnaN^{Ec}$, $DnaN^{Il}$, or $DnaN^{Sb}$ had markedly improved cell viability (Figure 3.6C). However, the positive effect observed with the heterologous clamps is largely dependent on the

function of DNA polymerase II (Figure 3.6C); the *dnaN159 polB* double mutant strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ are no more viable than the same strain with the empty vector, while $DnaN^{Ec}$ was still able to complement the *dnaN159* allele. A similar result was found at the non-permissive temperature (data not shown). Moreover, *dnaN159 polB* double mutant with $dnaN^{Il}$ or $dnaN^{Sb}$ had thousands-fold increase in the mutation rate at 37°C, to the level of 10^{-2} mutations per viable cell (Figure 3.6D). No significant impact was observed with $DnaN^{Bc}$, except that the cell viability was slightly lower as was observed in the temperature complementation analysis.

These results indicate that DNA polymerase II is likely needed for normal DNA replication when the foreign processivity clamps are used for DNA replication, even at the permissive temperature, which is consistent with the compromised stability [38] and activity of the DnaN159 protein even at the permissive temperature [37]. Presumably, when these strains lack DNA polymerase II, the other accessory DNA polymerases Pol IV and Pol V have more access to the replication fork based on the increase in mutation frequency found in these strains (Figures 3.6B and 3.6D). Even in the presence of functional DNA Pol II, the error prone DNA Pol IV and Pol V each likely competes at some level with DNA Pol II based on a modest decrease in mutation frequency observed in *dnaN159* strains expressing $DnaN^{Il}$ or $DnaN^{Sb}$ when the genes encoding either DNA polymerase IV or V was deleted (Figure 3.6B).

3.3.5. TnsE and DnaN must come from the same foreign host to allow transposition with the core Tn7 machinery in *E. coli.*

Figure 3.6. **Role of accessory DNA polymerases in *dnaN159* strains with DnaN homologs.** Cell density after overnight growth in colony-forming units and rate of spontaneous mutation to Rif^R were measured in *dnaN159 uvrC* strains where genes encoding the accessory DNA polymerases were individually inactivated. Results were measured at 30°C (A and B) and 37°C (C and D) as described (Materials and Methods). Error bars indicate the standard deviation of mean from two independent experiments. B/G indicates strain background. Vc indicates vector control.



A
B
C
D
Lg(cfu/ml)
Mutagenesis rate (x 10-6)
Mutagenesis rate (x10-2)
DnaN Vc Ec Il Sb Bc Vc Ec Il Sb Bc Vc Ec Il Sb Bc Vc Ec Il Sb Bc
B/G dnaN159 dnaN159 polII dnaN159 polIV dnaN159 polV

Our results suggested that we functionally replaced some or all of the processivity factor activity in *E. coli* with the homologous protein derived from another host. To address the larger question of host specificity with Tn7-like elements we utilized a hybrid transposition machinery where the core TnsABC machinery was from canonical Tn7, but the TnsE protein responsible for recognizing DNA replication was derived from a foreign host. We monitored the transposition of a miniTn7 by a promoter capturing assay where movement of a Tn7 derivative element containing a promoterless *lacZYA'* operon is indicated by phenotypically Lac^+ microcolonies when it moves to actively transcribed regions of the chromosome (see Material and methods). The miniTn7 element contains the *cis*-acting element for transposition from canonical Tn7 in *E. coli*. Since only the core machinery is involved in recognizing the *cis*-acting element [6,7,11,45], TnsABC of Tn7 should allow transposition of the miniTn7 element if the TnsE proteins can actively recognize transposition targets.

We found that we could establish transposition with the TnsABC core machinery in *E. coli*, but that this required that TnsE and DnaN come from the same foreign host in a *dnaN159* background (Figure 3.7). For example, $TnsE^{Il}$ promoted a high number of Lac^+ papillae with the Tn7 TnsABC core machinery only when $DnaN^{Il}$ was expressed in the same *E. coli dnaN159* strain, but not when another clamp were supplied or in the wild type strain (Figure 3.7A and data not shown). A similar result was found in strains expressing TnsE from *S. baltica* (Figure 3.7B and data not shown). DNA sequencing confirmed that actual transposition only occurred in *dnaN159* strains where TnsE and DnaN were transplanted into *E. coli* from *I. loihiensis* or *S. baltica*. These results indicate

that TnsE homologs would only allow transposition when host specific clamps are also being used in the cell.

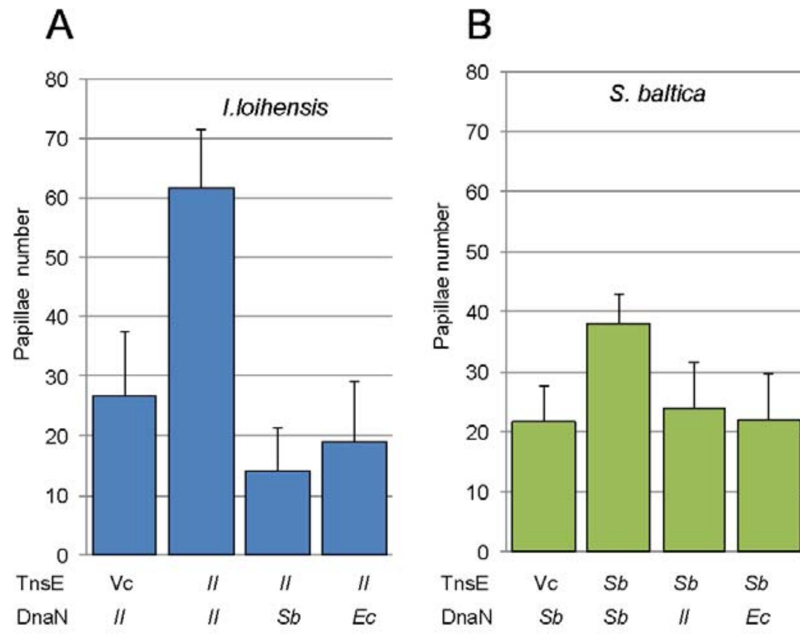
3.3.6 Conservation of DnaN sequences

Our result indicate that TnsE proteins found in Tn7-like elements in *S. baltica* and *I. loihiensis* have adapted to features found only in the DnaN protein from the host, even though DnaN from *S. baltica* and *I. loihiensis* were conserved enough to function with the replication enzymes of *E. coli.* This finding suggests that TnsE-β interaction would not interfere with the normal traffic on the clamp. This is consistent with the previous observation that only high levels of TnsE expression will induce the SOS response. Both of these observations suggest that TnsE binds to a region of the processivity factor that is largely distinct from the interfaces used with other proteins normally involved in DNA replication. To address this hypothesis, we examined the conservation and variation of DnaN homologs across different levels of divergence (Table 3.3).

When all the substitutions were modeled onto the known DnaN structure (PDB: 2POL), we found that amino acid residues at the C-terminal face are strikingly conserved and substitutions only occurred on the N-terminal face or at the rim (Figure 3.8).This finding suggests that highly conserved interactions likely primarily occur across the broad C-terminal face of the protein. Presumably, interactions that are specific to a given bacterial strain, as is the case with TnsE, could occur via the more variable portions of the β clamp protein, i.e. the rim or the N-terminal face.

3.4. Discussion

Figure 3.7. TnsE homologs allow transposition with Tn7 TnsABC with, and only with, DnaN from the same host. Transposition was monitored in strain ZL171 containing pCW15 expressing TnsABC from *E. coli*, pBAD24 and derivatives expressing TnsE homologs, and pBBR1-derivatives expressing DnaN homologs. Transposition frequency is indicated by the number of Lac^+ papillae after 72 hours at 30^0C on MacConkey's media. Error bars indicate the standard deviation of average papillae number from 12 patches.


A
B
I.loihensis
S. baltica
Papillae number
80
70
60
50
40
30
20
10
0
TnsE Vc Il Il Il
DnaN Il Il Sb Ec
TnsE Vc Sb Sb Sb
DnaN Sb Sb Il Ec

Tn7 and Tn7-like elements are widespread among bacterial populations and TnsE-mediated transposition is the Tn7 pathway that is known to be responsible for its horizontal transfer. In this study we carefully examined the distribution of Tn7-like elements and found that the most closely related Tn7-like elements were found in closely related bacteria, indicating host adaptation or host specificity. Using heterologous expression we show that the TnsE-β interaction we found previously with canonical Tn7 in *E. coli* is sufficient to explain the host specificity found with Tn7 and Tn7-like elements. In addition, the TnsE-β interaction we found with Tn7 in *E. coli* is likely conserved in highly diverged Tn7-like elements allowing these elements to target transposition into actively replicating DNA.

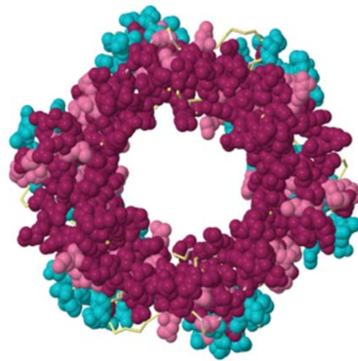
3.4.1. Specific interactions for β-binding proteins on the clamp

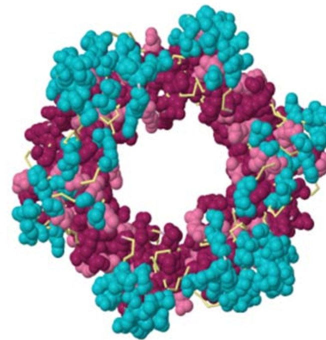
The β-clamp plays a central role in coordinating various cellular processes including DNA replication, recombination, and repair. Many proteins involved in these processes bind to this mobile platform via a β-binding motif (QL(S/D)LF) that interacts with a highly conserved hydrophobic cleft on the C-terminal face of the β-clamp [29,32,34,46-49]. Processivity factors in all forms of life, including PCNA in eukaryotes and archaea and gp45 in bacteriophage T4, share the same characteristic ring-shaped structure [50], even though they are only marginally similar at the sequence level and some operate as a trimer. Competition between DNA polymerases during DNA replication have been reported in different conditions [36,42,43,51-54]. Together these findings suggest that variation should occur uniformly across the clamp except in the hydrophobic pocket and across other surfaces that are important for interacting with host proteins.

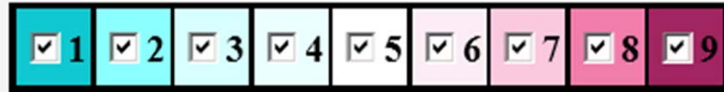
Table 3.3. DnaN sequences used for generating the sequence alignment for modeling.

Species	Sequence ID	Identity (%)	Similarity (%)
E. coli K-12 substr. MG1655	NP_418156	100	100
Shigella flexneri	NP_839180	99	99
Enterobacter cloacae ATCC 13047	YP_003610519	96	98
Vibrio sp. RC586	ZP_06078645	64	82
Tolumonas auensis DSM 9187	YP_002891218	65	79
Moritella sp. PE36	ZP_01900658	64	79
Shewanella loihica PV-4	YP_001092134	59	75
Shewanella baltica OS155	YP_001048408	58	75
Shewanella putrefaciens CN-32	YP_001181537	58	74
Idiomarina loihiensis L2TR	YP_154395	57	71
Pseudomonas syringae	YP_233114	56	73
Pseudomonas aeruginosa PAO1	NP_064722	56	72

Figure 3.8. **Conservation pattern on the faces of the β-clamp**. DnaN protein sequences (Table 3.3) were aligned and the conservation pattern was calculated and modeled onto the known structure of the β-clamp by ConSurf modeling. The β-clamp is shown as a space-filled model and colored according to the conservation score. The color-coding bar at the bottom shows the coloring theme. Positions where the inferred conservation level was assigned with low confidence are shown as sticks for the purpose of clarity.


C-terminal face


N-terminal face


1
2
3
4
5
6
7
8
9
Variable
Average
Conserved

Previous work indicates that functional conservation within processivity factors is lost across very broad phylogenetic differences. For example, *E. coli* clamp loader complex and DNA polymerase were not able to recognize the β-clamp of *Staphylococcus aureus* and *Streptococcus pyogenes* when studied previously *in vitro* [55,56]. In this study, we found that the evolutionarily distant $DnaN^{Bc}$ is toxic to *E. coli* cells, suggesting that the clamp loader of *E. coli* may be able to recognize the β-clamp of *B. cereus,* but that the foreign clamp no longer maintains the capacity for other interactions, consistent with what was found with the *S. aureus* clamp loader and the *E. coli* and *S. pyogenes* clamps before [55]. Interestingly, processivity factor from *I. loihiensis* and *S. baltica* were able to complement the previously published phenotypes of *dnaN159* allele, indicating an enhanced level of functional conservation with these heterologous clamps. Furthermore, when DnaN homologs of about 50% identity were modeled onto the clamp structure, we found a striking conservation of the entire C-terminal face. This suggests to us that regions beyond the hydrophobic pocket are also involved in making important contacts with its binding proteins in the cell. Previous mutational and structural studies have shown or indicated "protein-specific" binding interfaces on the clamp in addition to the hydrophobic cleft [29,57-63]. The variability on the N-terminal face and rim would suggest a less-stringent functional constraint in evolution in these areas (Figure 3.8).

The ability of $DnaN^{Il}$ and $DnaN^{Sb}$ to complement the *dnaN159* allele appeared to depend on Pol II (Figure 3.6). RecA protein was also important in these strain backgrounds (Table 3.2 and Figure 3.5). It would be reasonable to conclude that variations between the host DnaN, $DnaN^{Il}$ and $DnaN^{Sb}$ affected the ability of the different

β-binding proteins in *E. coli* to bind these heterologous clamps. Specifically, our data suggest that the clamp loader complex can successfully load the heterologous clamps onto DNA, but Pol III core enzyme manifests more strict interaction with β and consistently dissociates from these clamps during DNA replication.

3.4.2. TnsE-β interaction in Tn7 transposition

The TnsE pathway preferentially targets conjugating plasmids for transposition. We have shown previously by *in vivo* and *in vitro* studies that TnsE targets lagging-strand DNA replication by interacting with the β-clamp processivity factor [22]. A variant β-binding motif was identified in TnsE that is in a highly conserved region in TnsE-homologs. However, when TnsE homologs were introduced into *E. coli* with TnsABC, no transposition was detectable (Figure 3.2). Furthermore, the activity of the hybrid machinery was still undetectable in a *dnaN159* strain or with overexpression of $DnaN^{Ec}$, conditions which have been shown to stimulate TnsE-mediated transposition [22,64]. Of further interest, even though $DnaN^{Il}$ or $DnaN^{Sb}$ can function in *E. coli* cells with many proteins involved in DNA replication, TnsE homologs only promoted transposition when the DnaN protein from the same hosts were used by the cell (Figure 3.7). This species-specific interaction between TnsE and the β-clamp indicates that TnsE may interact with the less-conserved regions of the clamp. Given our modeling concerning conservation in the processivity factor we suspect that these contacts could be in the rim or N-terminal face of the dimer. This prediction would be consistent with the earlier finding that only an extremely high level of TnsE overexpression induces the SOS response in *E. coli*, interfering with the normal cellular traffic on the clamp [22].

The species specificity of TnsE-β interaction is sufficient to account for the observed host specificity of Tn7-like elements found in a wide variety of bacteria. While the very broad distribution of Tn7-like elements might have suggested that these elements move freely between diverged bacterial species, our results indicate otherwise (Figure 3.1). It is, however, important to note that Tn7-like elements are actively mobile in the environment; Tn7-like elements show extensive horizontal transfer of a large variety of genes [21] and the vast majority of Tn7-like elements lack obvious inactivating mutations in the transposon genes. Since the TnsE-mediated transposition pathway of Tn7 is responsible for its dissemination, the inability of TnsE homologs to recognize the β-clamp of the new host of a Tn7-like element would presumably limit its further distribution.

The TnsD-mediated pathway of transposition is unlikely to show the same species constraints as found with the TnsE-mediated pathway. The TnsD insertion pathway has been used extensively as a tool to insert gene into the neutral *attTn7* positions in many bacterial strains [65-71]. This observation along with our findings in this work suggests that Tn7-like elements may be donated freely to other bacterial strains. Tn7 recognizes conjugal plasmids when the plasmid enters the cell. Subsequent transfer of the element containing plasmid could bring the element into diverse hosts where the TnsD-mediated pathway of transfer will likely work. However, given the species specificity of the TnsE-β interaction revealed in this work, the Tn7-like element is unlikely to be able to utilize its TnsE pathway for additional transfer events.

3.4.3. Role of DNA polymerase II

Pol II is a family B DNA polymerase, members of which are utilized as the replicases in many eukaryotic and bacteriophage systems [72]. Pol II has a proofreading 3'-5' exonuclease activity and shows fidelity similar to DNA polymerase I and III [73]. The exact biological role of DNA Pol II in *E. coli* and other bacteria remains unclear. Pol II is involved in recovery of DNA synthesis after UV irradiation [74] and for mutagenic translesion synthesis [53] and is therefore considered as a translesional DNA polymerase [75]. However, Pol II has been shown to contribute to chromosomal DNA synthesis [42,76] and to function as an antimutator in wild type bacteria [77] and in various mutant backgrounds [43,51,54]. In this study, we found that Pol II consistently confers a growth advantage with foreign processivity factors even in the absence of obvious SOS induction when levels of the polymerase are expected to be low. We also found that Pol II participates in replication that appears to be non-mutagenic even in a strain background that lacks nucleotide excision repair. In addition to greatly reduced viability, strains that lacked DNA Pol II also displayed a huge increase in spontaneous mutations suggesting that in the absence of DNA Pol II, error prone polymerases have access to the replication forks. This is consistent with the work of others [51]. We suggest that DNA Pol III disengages more frequently from the foreign processivity clamps in *E. coli*, but that a sufficient DNA polymerase exchange system is functional with Pol II and the DnaN^{Il} and DnaN^{Sb} processivity clamps. We suggest that this system is swift and efficient as no higher than background SOS response induction was detected in *dnaN159* strains with either DnaN^{Il} or DnaN^{Sb}.

A role of Pol II in chromosomal replication has been implicated in previous studies especially in replication of the lagging-strand template strand [42,77]. The role of DNA Pol II is more pronounced in *dnaE* mutants, which presumably dissociate more readily from the mismatched primer end [42,76,77]. Interestingly, it was also recently found by Indiani *et al* that Pol II could form a stable replisome with the clamp and replicative helicase DnaB for processive DNA replication *in vitro* [48]. In the Pol II-containing replisome, the intrinsic speed of DnaB was adapted to the more slowly progressing DNA Pol II so that no single strand DNA gaps were produced in the process. Our results support a model where the high fidelity DNA Pol II should provide an efficient back-up activity for DNA Pol III during chromosomal DNA replication.

Like *E. coli*, *I. loihiensis* and *S. baltica* appear to have multiple DNA polymerases and therefore a reason to maintain a system for polymerase switching. DNA Pol II (Sba1_1630), Pol IV (Sbal_0935), and UmuC (Sbal_3553 and Sbal_2353) are all annotated in the genome of *S. baltica* OS155 (CP000563.1). *I. loihiensis* L2TR appears to have DNA Pol IV (IL0204) and DNA polymerase (IL2567) implicated in DNA repair (AE017340.1; [78]). Error prone polymerases can also be found on some plasmids [79], suggesting that polymerase switching is widespread. While there does not appear to be a homolog of DNA Pol II in *I. loihiensis*, we find that two *dnaE* genes are present in the chromosomes of *I. loihiensis* (IL1689 and IL2566). The second copy of *dnaE* genes in this strain is highly diverged from the *dnaE* gene that encodes α subunit of Pol III in *E. coli*, much like the case of *dnaE* genes in Gram positive bacteria [80]. A similarity has been proposed between Pol II of *E. coli* and the DnaE product of Gram-positive bacteria

[80], which is responsible for the lagging-strand DNA replication. It is fully possible that *I. loihiensis* DnaN has evolved to accommodate the function of its two DnaE enzymes and this adaption allowed Pol II in *E. coli* to bind better in our analysis.

3.4.4. The roles of RecA in promoting the function of accessory polymerases

It was surprising that *dnaN159* strains expressing DnaN^{Il} or DnaN^{Sb} strains were found to be extremely sick when the *recA* gene was inactivated given the extremely low level of SOS induction in these strains. Given the important role of DNA Pol II in these strains it is of interest what role or roles RecA may have in the function of this polymerase. Such a possibility was first suggested by Pham *et al* [81]. Sutton and colleagues [82] found that the *dnaN159 lexA51*(Def) was lethal with *recA730* allele, a mutant allele of RecA that is always in an active state, and this synthetic lethality was alleviated by deleting Pol II or Pol V, suggesting a role of RecA in the function of DNA polymerases Pol II and Pol V. The interaction between DNA Pol V and RecA is now known to be direct; RecA is part of the catalytically form of Pol V (Pol V-RecA-ATP) [83] and RecA-ssDNA nucleoprotein stimulates the activity of Pol V [84].It is possible that the effect of RecA on DNA Pol V impacts competition between the polymerases for the replication fork which indirectly affects access by DNA Pol II. In a recent study where the spectra of *rpoB* spontaneous mutations were carefully analyzed, it was found that the spectrum in *lexA*$^{+}$ was not different from *lexA*(Def) strain but dramatically changed in *lexA*(Def) *recA730* strain, indicating that the accessory polymerases are not able to access the replication fork even when they are transcriptionally derepressed unless they are activated by RecA [54].We submit that it remains formally possible that RecA

also plays a direct role with DNA Pol II. Further research will be needed to address this question which could be facilitated by work with heterologous sliding clamps (see below).

3.4.5. A novel way of studying the structure and function of essential proteins

Presumably the sequence and structure of the clamp protein of a specific species have been optimized through evolution to ensure proper function with all β-binding proteins in that species. Homologous proteins provide an alternative to classic mutational study and studying their function of homologous proteins in a heterologous setup would have some unexpected benefit. Further studies to explore the linkage between the sequence variation and the altered polymerase usage observed in this study would be instructive in understanding the structure and function of the processivity factor.

3.5. Materials and Methods

3.5.1. Media, chemicals and enzymes

LB broth and agar were prepared as described by Miller [85] and 0.2% glucose was used to suppress transcription from P*lac* promoter and the P_{BAD} promoter when necessary. Trimethoprim selection was on Isosensitest agar (Oxoid). Minimal media (MM) were prepared as [85] and supplemented with 0.2% appropriate carbon sources (glucose, maltose) and 0.4% Casamino acids. Lac phenotypes were evaluated on MacConkey lactose agar (Difco) and arabinose was added to 0.2% to induce expression of genes cloned in the plasmid pBAD24 vector from the P_{BAD} promoter. Antibiotics were used at the following concentrations: ampicillin (Amp), 100 μg/ml; chloramphenicol

(Cm), 30 μg/ml; kanamycin (Km), 50 μg/ml; tetracycline (Tet), 20 μg/ml; spectinomycin (Spec) 10 μg/ml; trimethorpim (Tp), 100 μg/ml; and rifampicin (Rif), 50 μg/ml.

3.5.2. Bacterial strains and plasmids

E. coli strains used in this study are listed in Table3.4. All are derivatives of *E. coli* K-12 and genetics alleles were introduced by P1 *vir* mediated generalized transduction. P1 transductants were selected for the specifically linked antibiotic resistance on LB agar or screened for the required nutrient on Minimal Media. *dnaN159* allele was selected for the Tet^R resistance conferred by the *tnaA300::tn10* allele, which has about 90% genetic linkage to *dnaN* allele, or the Km^R resistance conferred by the *attTn7::*miniTn7*::kan* allele, which is also closely linked to *dnaN*. The constructed strains were further tested for the known phenotypes of a specific allele, such as temperature sensitivity and UV sensitivity. Colony polymerase chain reactions (PCR) followed with direct sequencing were used to further confirm the successful construction of strains. *E. coli* strain *DH5α* was used as the host strain for plasmid construction and propagation.

Plasmids are described in Table 3.5. Vectors pCW15 (TnsABC), and pJP131 (TnsE from *E. coli* Tn7 in pBAD24) were described previously [4,19]. TnsE homologs from *I. loihiensis* ($TnsE^{Il}$), *S. baltica* ($TnsE^{Sb}$), and *B. cereus* ($TnsE^{Bc}$) were cloned into pBAD24 vector under the control of P_{BAD} promoter. Primer sets used to amplify the DNA fragments containing the *tnsE* homologous genes are JEP173 5'-GAG GAA TTC ACC ATG GAT AGG CTC GGT GGG TTT GAA-3' and JEP 174 5'- GCG GTA CCC TAT CAT ACT CTG GCT AAC AAA TGG CG-3' for *tnsE*^{Il}, JEP169 5'- GAG GAA TTC ACC ATG AGT AAG AGT GAA GTA AAA-3' and JEP170 5'- GCG GTA CCC TAT

TAT GAA CGA CTT GCC CAA TTA GC-3’ for $tnsE^{Sb}$, and JEP171 5’- GAG GAA TTC ACC ATG GAG AGT TTA CAG TTT AAA AAG TGG-3’ and JEP172 5’- GCG GTA CCC TAC TAA TTT TTA TTT TCT TGT AAC CC-3’ for $tnsE^{Bc}$, where the EcoRI sites and KpnI sites are underlined. The PCR amplicons were first cloned into pGEM-T cloning vector by TA-cloning as instruction (Promega) on LB agar with glucose. The EcoRI/KpnI fragments containing the *tnsE* open reading frames were then subcloned into pBAD24 vector under the control of P_{BAD} promoter. All clones were confirmed by sequencing.

The expression of the *dnaN* gene in *E. coli* is controlled at different levels and at least three promoters were identified in the *dnaA* coding region [86]. To simplify, the open reading frames of $dnaN^{Il}$, $dnaN^{Sb}$, or $dnaN^{Bc}$ without the native promoters were each cloned into the multiple cloning site on pBBR1 plasmid, a broad-host-range plasmid with moderate copy number (approximately 30 copies)[87], either under the control of the P*lac* promoter (forward) or in the opposite orientation (reverse).The *dnaN* open reading frames were amplified with primer sets: JEP215 5’-GG*G GTA CC*A TAA ACA GGG GAC TCA G-3’and JEP216 5’-CCG GGT ACC AAC ATT TAT GAG GC AG-3’ for $dnaN^{Il}$, JEP262 5’-GGT ACC TCA AGG AAT TTG GAC ACT ATG AAA-3’ and JEP263 5’GGT ACC TGG AAT ACT ATA AGC GCA TCG-3’ for $dnaN^{Sb}$, and JEP213 5’-GCT GGT ACC GGA GGT TTT TCT TTA TGC G-3’ and JEP214 5’-CCC GGT ACC TTC ATA AAA TCG CTC AC-3’ for $dnaN^{Bc}$ in the *dnaA-dnaN-recF* operon. The PCR amplicons were first cloned into pGEM-T vector by TA cloning as instructed and then subcloned into the KpnI site of pBBR1-MCS4 vector in both orientations. We also

Table 3.4. Strains used in this study.

Strains	Relative genotype	Source
DH5α	*endA1 hsdR17 ($r_K^- m_K^+$) glnV44 thi-1 recA1 gyrArelA*Δ*(lacZYA-argF) U169 deoR (φ80dlac*Δ*(lacZ)M15)*	Laboratory stock
NLC28	F- *araD139* Δ*(argF-lac) U169 rspL150 relA1 flbB5301 deoC1 pstF25 Val^R*	Laboratory stock
NLC51	F- *araD139* Δ*(argF-lac) U169 rspL150 relA1 flbB5301 deoC1 pstF25 Val^R recA56*	Laboratory stock
JP617	NLC28 *attTn7:miniTn7*(R90-lacZYA'-Kan^R-L166)	Laboratory stock
JP1104	MG1655 *srl*::Tn10	Laboratory stock
JP1384	NLC28 Δ*ara714 leu*::Tn10	Laboratory stock
JP1386	NLC28Δ*ara714*	Laboratory stock
JP2172	MG1655 *uvrC278*::Tn10	Laboratory stock
AP330	NLC28 *sulA*::Kan^R::*lacZYA'*	Laboratory stock
AP388	NLC28 Δara714 sulA::Cm^R::*frt*::*lacZYA'*	Laboratory stock
AP724	NLC28 *attTn7*::Cm^R	Laboratory stock
SY2	Δ*lac X74 rpsL araD139* Δ*(ara-leu)7697 galUgalKhsr-hsm+ sulA*::Kan^R::*lacZYA*	From Mark Sutton
CJ278	Δ*(bal-bio) thi-1 relA1 spoT1* Δ*polA*::Kan^R	From Mark Sutton
RW120	*thi-1 araD139* Δ*(gpt-proA)62 lacY1tsx-33 supE44 galK2 hisG4(Oc) mg-1 rpsL31xyl-5 mtl-1 argE3(Oc) thi-1 sulA211* Δ*umuDC595*::CmR	From Mark Sutton
NK8207	F- Δ*(lac-pro) thirpsL* Δ*(gal - λG) + lac*ZpLcI$^+_{434}$ pRS7	Laboratory stock
FC1237	*ara*Δ*(lac-proB) ΔdinB*::Kan^R	From Mark Sutton
STL1336	*thr-1 ara-14 leuB6* Δ*(gpt-proA)62 lacY1 tsx-33 supE44 galK2 hisG4(Oc) rfbD1 mg-1 rpsL31 kdgK51 xyl-5 mtl-1 argE3(Oc) thi-1* Δ*(araD-polB)*::Ω	From Mark Sutton
AP179	MG1655 *tnaA300*::Tn10	From Mark Sutton
AP180	MG1655 *dnaN159 tnaA300*::Tn10	From Mark Sutton

Table 3.4 continued

Strains	Relative genotype	construction	Source
ZL108	NLC28 *dnaN159 tnaA300*::Tn10	NLC28 x P1 (AP180)	This work
ZL113	NLC28 *dnaN159*	ZL108 x P1 (NLC28)	This work
ZL114	NLC28 *dnaN159 uvrC278*::Tn10	ZL113 x P1 (JP2172)	This work
ZL115	NLC28 dnaN159 *attTn7*::miniTn7 (R90-lacZYA'-KanR-L166)	ZL113 x P1 (JP617)	This work
ZL123	NLC28 *dnaN159* *sulA*::Kan^R::*lacZYA'*	ZL113 x P1 (SY2)	This work
ZL171	NLC28 Δ*ara714 dnaN159* *attTn7::*miniTn7(R90-lacZYA'-KanR-L166)	JP1386 x P1 (ZL115)	This work
ZL172	NLC28 *dnaN159* *sulA*::Kan^R::*lacZYA'* *umuDC595*::Cm^R	ZL123 x P1 (RW120)	This work
ZL173	NLC28 *dnaN159 dinB*::Kan^R	ZL113 x P1 (PC1237)	This work
ZL189	NLC28 *dnaN159* Δ*ara714 leu*::Tn10	ZL113 x P1 (JP1384)	This work
ZL190	NLC28 *dnaN159* *sulA*::Kan^R::*lacZYA'* *umuDC595*::Cm^R *uvrC278*::Tn10	ZL172 X P1 (JP2172)	This work
ZL192	NLC28 *dnaN159 dinB*::Kan^R *uvrC278*::Tn10	ZL173 x P1 (JP2172)	This work
ZL220	NLC28 *dnaN159* Δ*ara714*	ZL189 x P1 (JP1386)	This work
ZL221	NLC28 *dnaN159* Δ*ara714 srl*::Tn10	ZL220 x P1 (JP1104)	This work
ZL222	NLC28 *dnaN159* Δ*ara714 recA56*	ZL221x P1 (NLC51)	This work
ZL223	NLC28 *dnaN159* *dinB*::Kan^R*uvrC278*::Tn10 *sulA*::Cm^R::*frt*::*lacZYA'*	ZL192 x P1 (AP388)	This work
ZL254	NLC28 dnaN159 *sulA*::*lacZYA*::Kan^R*uvrC278*::Tn10	ZL123 x P1 (JP2172)	This work
ZL257	NLC28 *dnaN159* *polB*::Ω*sulA*::Kan^R::*lacZYA'uvrC278* ::Tn10	ZL254 x P1 (STL1336)	This work
ZL411	NK8207 *dnaN159 tnaA300::*Tn10	NK8207 x P1 (AP180)	This work

Table 3.5. Plasmid DNAs used in this study.

Plasmids	Relative Genotype	Resource
pBBR1-MCS4-Tp		[95]
pBAD24		[96]
pACYC184		NEB
pJP131	*tnsE*^{Ec} cloned in pBAD24	[19]
pCW15	*tnsABC*^{Ec} cloned in pACYC184	[4]
pZL106	*dnaN*^{Il} in pGEMT vector by TA cloning)	This work
pZL107	*dnaN*^{Sb} in pGEMT by TA-cloning	This work
pZL108	*dnaN*^{Sb} in pGEMT by TA-cloning	This work
pZL109	*dnaN*^{Bc} in pGEMT by TA-cloning	This work
pZL110	*dnaN*^{Bc} in pGEMT by TA-cloning	This work
pZL111	*dnaN*^{Ec} in pGEMT by TA-cloning	This work
pZL112	*dnaN*^{Il} in pBBR1-Tp expressed from the P*lac* promoter	This work
pZL113	*dnaN*^{Il}in pBBR1-Tp not expressed from the P*lac* promoter	This work
pZL114	*dnaN*^{Sb} in pBBR1-Tp expressed from the P*lac*promoter	This work
pZL115	*dnaN*^{Sb} in pBBR1-Tp not expressed from the P*lac* promoter	This work
pZL116	*dnaN*^{Bc}in pBBR1-Tp expressed from the P*lac* promoter	This work
pZL117	*dnaN*^{Bc} in pBBR1-Tp not expressed from the P*lac* promoter	This work
pZL118	*dnaN*^{Ec} in pBBR1-Tp expressed from the P*lac* promoter	This work
pZL119	*dnaN*^{Ec}in pBBR1-Tp not expressed from the P*lac* promoter	This work
pZL120	*tnsE*^{Il} in pBAD24	This work
pZL121	*tnsE*^{Sb} in pBAD24	This work
pZL122	*tnsE*^{Bc} in pBAD24	This work

included *dnaN* of E. coli (*dnaN*^{Ec}) as a control, which was amplified by primers JEP313 5'-ATC GCG GAT CCT GCA GCG GGT GAG GGA CAT TAC AGT-3' and JEP316 5'-AGT TGC GTG ACT GCA GCC ACG ATA TCA AAG-3', cloned into the pGEM-T vector and then subcloned into the PstI site of pBBR1-MCS4 vector in both orientations. All cloning steps were carefully performed on glucose MM at 30^0C and the final clones were confirmed by sequencing. We had no difficulty in cloning *dnaN*^{Il}*, dnaN*^{Sb} *and dnaN*^{Ec} in either direction. However, cloning *dnaN*^{Bc} in the forward direction failed many times and the final clone obtained had a R107I mutation. Because no phenotype was observed with the heterologous clamps cloned in the reverse orientation, all clones are referred to the forward orientation unless specifically mentioned.

3.5.3. SOS response

SOS induction was evaluated in strain ZL123 and strain ZL411, where *lacZ* genes are transcriptionally fused to different SOS-inducible promoters. Plasmid pBBR1 and its derivatives were introduced by transformation of chemical competent cells made by $CaCl_2$ method. Cells were grown overnight in LB media containing 0.2% glucose and trimethoprim (100 mg/ml) at 30^0C. Overnight cultures were diluted 1:100 in fresh LB with trimethoprim, and then grown for 2 hr at temperatures indicated. β-galactosidase activity was measured in triplicate for each sample as described [88].

3.5.4. In vivo analysis of transposition

Transposition was monitored by a promoter capture assay called papillation assay [89,90]. In this assay, transposition proteins are provided *in trans* by plasmids pCW15 for TnsABC from the Tn7 element found in *E. coli*, and pBAD24 derivatives for TnsE

homologs. The *cis*-acting factor for transposition is carried on a miniTn7 element at the *attTn7* site on the chromosome where the lactose utilization genes lack of the requisite promoter are bounded by the inverted repeat sequences of Tn7 from *E. coli*. Because the element is integrated into a transcriptionally silent region of the *E. coli* chromosome the strain is unable to utilize lactose as a carbon sources (Lac$^-$). Transposition of the miniTn7 element can be monitored on indicator media because if the mimiTn7 element transposes downstream of an active promoter the host cell will be able to ferment lactose (Lac$^+$), resulting in red papillae on the otherwise white lawn of bacteria on the indicator media used in the experiment. Since gross chromosome rearrangements and other mutations adjacent to the miniTn7 element in the cell could also give rise to red papillae [90], in our assay transposition level of a specific group is always compared to that of a background strain where only the transposition-incompetent TnsABC are provided *in trans*.

To map Tn7 transposition events, 2 papillae from each patch were first streak purified from the lawn and examined by P1 linkage mapping to determine if transposition had occurred. P1-positive events were further mapped by direct sequencing of the arbitrary PCR products to determine the location of the miniTn7 element on the chromosome as described previously [89,90].

3.5.5. Spontaneous mutation analysis

Spontaneous mutation rate was determined by a rifampicin resistance assay adapted from [60]. For each strain background to be tested, ten fresh colonies from glucose minimal media plates grown at 30^{0}C were inoculated into 5 ml maltose minimal media with Tmp at indicated temperature with shaking. Aliquots of appropriate dilutions of the

overnight cultures were then plated in duplicate on selective (Rif^R) and nonselective LB plates to determine mutant and total cell counts. Only dilutions with 30~300 colonies were considered to have reliable counts. Spontaneous mutation frequencies were determined by dividing the mutant count by the total cell count.

3.5.6. Computational methods

Sequences were aligned with the MUSCLE program in UGENE software (available from http://ugene.unipro.ru). Protein sequences were concatenated by using Bioedit software [91]. Distance matrix was calculated and neighbor-joining tree was constructed with programs in the Phylip 3.69 package (available from http://evolution.genetics.washington.edu/phylip.html)[92]. Bootstrap analysis was carried out with 1000 iterations. Trees were displayed and figures produced in Figtree v1.3.1 (available fromhttp://www.tree.bio.ed.ac.uk/).

Conservation pattern of DnaN homologs from prokaryotes were obtained by ConSurf modeling (http://consurf.tau.ac.il/)[93,94]. Sequences were first aligned with MUSCLE program in UGENE software (available from http://ugene.unipro.ru). The conservation score of each amino acid was then calculated with the empirical Bayesian method using the JTT matrix as the model of substitution for proteins, which was then modeled onto the known structure of the β-clamp (2PQL)[49].

3.6. Acknowledgements

We thank Mark Sutton (Department of Biochemistry, School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York) for *E. coli*strains, Jim Fredrickson (Pacific Northwest National Laboratory) for *S. baltica*

OS155 strain, Timothy Read (Navy Medical Research Center) for *B. cereus* ATCC10987 strain, Maqsudul Alam (Dept. of Microbiology, University of Hawaii) for *I. loihiensis* L2TR genomic DNA. This work was supported by the NIH (GM069508).

REFERENCES

1. Craig NL (2002) Tn7. In: Craig NL, R. Craigie, M. Gellert, and A.M. Lambowitz., editor. Mobile DNA II. Washington, DC: ASM Press. pp. 423-456.

2. Peters JE, Craig NL (2001) Tn7: smarter than we thought. Nat Rev Mol Cell Biol 2: 806-814.

3. Li Z, N.L. Craig, and J.E. Peters (2011) Transposon Tn7. In: Mullany APRaP, editor. Bacterial Integrative Mobile Genetic Elements. Austin, TX: Landies Bioscience. pp. Epub.

4. Waddell CS, Craig NL (1988) Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. Genes Dev 2: 137-149.

5. Kubo KM, Craig NL (1990) Bacterial transposon Tn7 utilizes two different classes of target sites. J Bacteriol 172: 2774-2778.

6. Biery MC, Lopata M, Craig NL (2000) A minimal system for Tn7 transposition: the transposon-encoded proteins TnsA and TnsB can execute DNA breakage and joining reactions that generate circularized Tn7 species. J Mol Biol 297: 25-37.

7. Sarnovsky RJ, May EW, Craig NL (1996) The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. Embo J 15: 6348-6361.

8. Stellwagen AE, Craig NL (1997) Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7. Genetics 145: 573-585.

9. Kuduvalli PN, Rao JE, Craig NL (2001) Target DNA structure plays a critical role in Tn7 transposition. Embo J 20: 924-932.

10. Bainton RJ, Kubo KM, Feng JN, Craig NL (1993) Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell 72: 931-943.

11. Waddell CS, Craig NL (1989) Tn7 transposition: recognition of the attTn7 target sequence. Proc Natl Acad Sci U S A 86: 3958-3962.

12. McKown RL, Orle KA, Chen T, Craig NL (1988) Sequence requirements of Escherichia coli attTn7, a specific site of transposon Tn7 insertion. J Bacteriol 170: 352-358.

13. Bainton R, Gamas P, Craig NL (1991) Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell 65: 805-816.

14. Mitra R, McKenzie GJ, Yi L, Lee CA, Craig NL (2010) Characterization of the TnsD-attTn7 complex that promotes site-specific insertion of Tn7. Mob DNA 1: 18.

15. Gay NJ, Tybulewicz VL, Walker JE (1986) Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. Biochem J 234: 111-117.

16. Wolkow CA, DeBoy RT, Craig NL (1996) Conjugating plasmids are preferred targets for Tn7. Genes Dev 10: 2145-2157.

17. Finn JA, Parks AR, Peters JE (2007) Transposon Tn7 directs transposition into the genome of filamentous bacteriophage M13 using the element-encoded TnsE protein. J Bacteriol 189: 9122-9125.

18. Barth PT, Datta N, Hedges RW, Grinter NJ (1976) Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. J Bacteriol 125: 800-810.

19. Peters JE, Craig NL (2001) Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes Dev 15: 737-747.

20. Parks AR, Peters JE (2007) Transposon Tn7 is widespread in diverse bacteria and forms genomic islands. J Bacteriol 189: 2170-2173.

21. Parks AR, Peters JE (2009) Tn7 elements: engendering diversity from chromosomes to episomes. Plasmid 61: 1-14.

22. Parks AR, Li Z, Shi Q, Owens RM, Jin MM, et al. (2009) Transposition into replicating DNA occurs through interaction with the processivity factor. Cell 138: 685-695.

23. Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem 74: 283-315.

24. Pham P, Bertram JG, O'Donnell M, Woodgate R, Goodman MF (2001) A model for SOS-lesion-targeted mutations in Escherichia coli. Nature 409: 366-370.

25. Bonner CA, Stukenberg PT, Rajagopalan M, Eritja R, O'Donnell M, et al. (1992) Processive DNA synthesis by DNA polymerase II mediated by DNA polymerase III accessory proteins. J Biol Chem 267: 11431-11438.

26. Fuchs RP, Fujii S, Wagner J (2004) Properties and functions of Escherichia coli: Pol IV and Pol V. Adv Protein Chem 69: 229-264.

27. Sutton MD, Narumi I, Walker GC (2002) Posttranslational modification of the umuD-encoded subunit of Escherichia coli DNA polymerase V regulates its interactions with the beta processivity clamp. Proc Natl Acad Sci U S A 99: 5307-5312.

28. Lopez de Saro FJ, O'Donnell M (2001) Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. Proc Natl Acad Sci U S A 98: 8376-8380.

29. Jeruzalmi D, Yurieva O, Zhao Y, Young M, Stewart J, et al. (2001) Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. Cell 106: 417-428.

30. Lopez de Saro FJ, Marinus MG, Modrich P, O'Donnell M (2006) The beta sliding clamp binds to multiple sites within MutL and MutS. J Biol Chem 281: 14340-14349.

31. Kurz M, Dalrymple B, Wijffels G, Kongsuwan K (2004) Interaction of the sliding clamp beta-subunit and Hda, a DnaA-related protein. J Bacteriol 186: 3508-3515.

32. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. Proc Natl Acad Sci U S A 98: 11627-11632.

33. Sutton MD (2009) Coordinating DNA polymerase traffic during high and low fidelity synthesis. Biochim Biophys Acta 1804: 1167-1179.

34. Georgescu RE, Yurieva O, Kim SS, Kuriyan J, Kong XP, et al. (2008) Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. Proc Natl Acad Sci U S A 105: 11116-11121.

35. Grigorian AV, Lustig RB, Guzman EC, Mahaffy JM, Zyskind JW (2003) Escherichia coli cells with increased levels of DnaA and deficient in recombinational repair have decreased viability. J Bacteriol 185: 630-644.

36. Sutton MD (2004) The Escherichia coli dnaN159 mutant displays altered DNA polymerase usage and chronic SOS induction. J Bacteriol 186: 6738-6748.

37. Burgers PM, Kornberg A, Sakakibara Y (1981) The dnaN gene codes for the beta subunit of DNA polymerase III holoenzyme of escherichia coli. Proc Natl Acad Sci U S A 78: 5391-5395.

38. Maul RW, Ponticelli SK, Duzen JM, Sutton MD (2007) Differential binding of Escherichia coli DNA polymerases to the beta-sliding clamp. Mol Microbiol 65: 811-827.

39. Sakakibara Y, Mizukami T (1980) A temperature-sensitive Escherichia coli mutant defective in DNA replication: dnaN, a new gene adjacent to the dnaA gene. Mol Gen Genet 178: 541-553.

40. Little JW (1984) Autodigestion of lexA and phage lambda repressors. Proc Natl Acad Sci U S A 81: 1375-1379.

41. Dutreix M, Moreau PL, Bailone A, Galibert F, Battista JR, et al. (1989) New recA mutations that dissociate the various RecA protein activities in Escherichia coli provide evidence for an additional role for RecA protein in UV mutagenesis. J Bacteriol 171: 2415-2423.

42. Delmas S, Matic I (2006) Interplay between replication and recombination in Escherichia coli: impact of the alternative DNA polymerases. Proc Natl Acad Sci U S A 103: 4564-4569.

43. Gawel D, Pham PT, Fijalkowska IJ, Jonczyk P, Schaaper RM (2008) Role of accessory DNA polymerases in DNA replication in Escherichia coli: analysis of the dnaX36 mutator mutant. J Bacteriol 190: 1730-1742.

44. Viguera E, Petranovic M, Zahradka D, Germain K, Ehrlich DS, et al. (2003) Lethality of bypass polymerases in Escherichia coli cells with a defective clamp loader complex of DNA polymerase III. Mol Microbiol 50: 193-204.

45. McKown RL, Waddell CS, Arciszewska LK, Craig NL (1987) Identification of a transposon Tn7-dependent DNA-binding activity that recognizes the ends of Tn7. Proc Natl Acad Sci U S A 84: 7807-7811.

46. Burnouf DY, Olieric V, Wagner J, Fujii S, Reinbolt J, et al. (2004) Structural and biochemical analysis of sliding clamp/ligand interactions suggest a competition between replicative and translesion DNA polymerases. J Mol Biol 335: 1187-1197.

47. Lopez de Saro FJ, Georgescu RE, O'Donnell M (2003) A peptide switch regulates DNA polymerase processivity. Proc Natl Acad Sci U S A 100: 14689-14694.

48. Indiani C, Langston LD, Yurieva O, Goodman MF, O'Donnell M (2009) Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. Proc Natl Acad Sci U S A 106: 6031-6038.

49. Kong XP, Onrust R, O'Donnell M, Kuriyan J (1992) Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp. Cell 69: 425-437.

50. Lopez de Saro FJ (2009) Regulation of interactions with sliding clamps during DNA replication and repair. Curr Genomics 10: 206-215.

51. Hastings PJ, Hersh MN, Thornton PC, Fonville NC, Slack A, et al. (2010) Competition of Escherichia coli DNA polymerases I, II and III with DNA Pol IV in stressed cells. PLoS One 5: e10862.

52. Furukohri A, Goodman MF, Maki H (2008) A dynamic polymerase exchange with Escherichia coli DNA polymerase IV replacing DNA polymerase III on the sliding clamp. J Biol Chem 283: 11260-11269.

53. Becherel OJ, Fuchs RP (2001) Mechanism of DNA polymerase II-mediated frameshift mutagenesis. Proc Natl Acad Sci U S A 98: 8566-8571.

54. Curti E, McDonald JP, Mead S, Woodgate R (2009) DNA polymerase switching: effects on spontaneous mutagenesis in Escherichia coli. Mol Microbiol 71: 315-331.

55. Bruck I, Georgescu RE, O'Donnell M (2005) Conserved interactions in the Staphylococcus aureus DNA PolC chromosome replication machine. J Biol Chem 280: 18152-18162.

56. Klemperer N, Zhang D, Skangalis M, O'Donnell M (2000) Cross-utilization of the beta sliding clamp by replicative polymerases of evolutionary divergent organisms. J Biol Chem 275: 26136-26143.

57. Heltzel JM, Maul RW, Scouten Ponticelli SK, Sutton MD (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. Proc Natl Acad Sci U S A 106: 12664-12669.

58. Sutton MD, Duzen JM (2006) Specific amino acid residues in the beta sliding clamp establish a DNA polymerase usage hierarchy in Escherichia coli. DNA Repair (Amst) 5: 312-323.

59. Beuning PJ, Sawicka D, Barsky D, Walker GC (2006) Two processivity clamp interactions differentially alter the dual activities of UmuC. Mol Microbiol 59: 460-474.

60. Sutton MD, Duzen JM, Maul RW (2005) Mutant forms of the Escherichia colibeta sliding clamp that distinguish between its roles in replication and DNA polymerase V-dependent translesion DNA synthesis. Mol Microbiol 55: 1751-1766.

61. Duzen JM, Walker GC, Sutton MD (2004) Identification of specific amino acid residues in the E. coli beta processivity clamp involved in interactions with DNA polymerase III, UmuD and UmuD'. DNA Repair (Amst) 3: 301-312.

62. Bunting KA, Roe SM, Pearl LH (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. Embo J 22: 5883-5892.

63. Naktinis V, Turner J, O'Donnell M (1996) A molecular switch in a replication machine defined by an internal competition for protein rings. Cell 84: 137-145.

64. Parks AR (2008) Tn7 transposes into replicating DNA using an interaction with the processivity factor, facilitating genome evolution. Ithaca: Cornell University. 210 p.

65. Howe K, Karsi A, Germon P, Wills RW, Lawrence ML, et al. (2010) Development of stable reporter system cloning luxCDABE genes into chromosome of Salmonella enterica serotypes using Tn7 transposon. BMC Microbiol 10: 197.

66. Sun JC, Zhang EH, Yao LG, Zhang HL, Jin PF (2009) A high efficient method of constructing recombinant Bombyx mori (silkworm) multiple nucleopolyhedrovirus based on zero-background Tn7-mediated transposition in Escherichia coli. Biotechnol Prog 25: 524-529.

67. Barnes RJ, Leung KT, Schraft H, Ulanova M (2008) Chromosomal gfp labelling of Pseudomonas aeruginosa using a mini-Tn7 transposon: application for studies of bacteria-host interactions. Can J Microbiol 54: 48-57.

68. McKenzie GJ, Craig NL (2006) Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn7 without need for selection of the insertion event. BMC Microbiol 6: 39.

69. Choi KH, Schweizer HP (2006) mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa. Nat Protoc 1: 153-161.

70. Choi KH, Schweizer HP (2006) mini-Tn7 insertion in bacteria with secondary, non-glmS-linked attTn7 sites: example Proteus mirabilis HI4320. Nat Protoc 1: 170-178.

71. Choi KH, DeShazer D, Schweizer HP (2006) mini-Tn7 insertion in bacteria with multiple glmS-linked attTn7 sites: example Burkholderia mallei ATCC 23344. Nat Protoc 1: 162-169.

72. Garcia-Diaz M, Bebenek K (2007) Multiple functions of DNA polymerases. CRC Crit Rev Plant Sci 26: 105-122.

73. Cai H, Yu H, McEntee K, Kunkel TA, Goodman MF (1995) Purification and properties of wild-type and exonuclease-deficient DNA polymerase II from Escherichia coli. J Biol Chem 270: 15327-15335.

74. Rangarajan S, Woodgate R, Goodman MF (1999) A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated Escherichia coli. Proc Natl Acad Sci U S A 96: 9224-9229.

75. Goodman MF (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annu Rev Biochem 71: 17-50.

76. Rangarajan S, Gudmundsson G, Qiu Z, Foster PL, Goodman MF (1997) Escherichia coli DNA polymerase II catalyzes chromosomal and episomal DNA synthesis in vivo. Proc Natl Acad Sci U S A 94: 946-951.

77. Banach-Orlowska M, Fijalkowska IJ, Schaaper RM, Jonczyk P (2005) DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in Escherichia coli. Mol Microbiol 58: 61-70.

78. Hou S, Saw JH, Lee KS, Freitas TA, Belisle C, et al. (2004) Genome sequence of the deep-sea gamma-proteobacterium Idiomarina loihiensis reveals amino acid fermentation as a source of carbon and energy. Proc Natl Acad Sci U S A 101: 18036-18041.

79. Goldsmith M, Sarov-Blat L, Livneh Z (2000) Plasmid-encoded MucB protein is a DNA polymerase (pol RI) specialized for lesion bypass in the presence of MucA', RecA, and SSB. Proc Natl Acad Sci U S A 97: 11227-11231.

80. Bruck I, O'Donnell M (2000) The DNA replication machine of a gram-positive organism. J Biol Chem 275: 28971-28983.

81. Pham P, Rangarajan S, Woodgate R, Goodman MF (2001) Roles of DNA polymerases V and II in SOS-induced error-prone and error-free repair in Escherichia coli. Proc Natl Acad Sci U S A 98: 8350-8354.

82. Maul RW, Sutton MD (2005) Roles of the Escherichia coli RecA protein and the global SOS response in effecting DNA polymerase selection in vivo. J Bacteriol 187: 7607-7618.

83. Jiang Q, Karata K, Woodgate R, Cox MM, Goodman MF (2009) The active form of DNA polymerase V is UmuD'(2)C-RecA-ATP. Nature 460: 359-363.

84. Schlacher K, Cox MM, Woodgate R, Goodman MF (2006) RecA acts in trans to allow replication of damaged DNA by DNA polymerase V. Nature 442: 883-887.

85. Miller JH (1992) A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria. New York, NY: CSHL press. 456 p.

86. Armengod ME, Garcia-Sogo M, Lambies E (1988) Transcriptional organization of the dnaN and recF genes of Escherichia coli K-12. J Biol Chem 263: 12109-12114.

87. Antoine R, Locht C (1992) Isolation and molecular characterization of a novel broad-host-range plasmid from Bordetella bronchiseptica with sequence similarities to plasmids from gram-positive organisms. Mol Microbiol 6: 1785-1799.

88. Miller JH (1992) A Short Course in Bacterial Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

89. Shi Q, Parks AR, Potter BD, Safir IJ, Luo Y, et al. (2008) DNA damage differentially activates regional chromosomal loci for Tn7 transposition in Escherichia coli. Genetics 179: 1237-1250.

90. Peters JE, Craig NL (2000) Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates. Mol Cell 6: 573-582.

91. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98.

92. Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17: 368-376.

93. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38: W529-533.

94. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, et al. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res 33: W299-302.

95. Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, et al. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166: 175-176.

96. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J Bacteriol 177: 4121-4130.

CHAPTER 4

TNSE REVEALS TRANSPOSITION TARGET BY INTERACTION DISRUPTION OF THE SEQA COMPLEX

4.1. Summary

The TnsABC+E pathway of transposon Tn7 targets discontinuous replication for transposition likely by recognizing the β-clamp processivity factor at the primer/template junction. An intriguing question is how this activity is regulated or when TnsE has access to the target, given the important role of the β-clamp in DNA replication, recombination and repair. In this chapter I present biochemical and genetic studies supporting a mechanism that may account for the distribution of TnsE-mediated transposition events in the genome. I found TnsE physically interacts with the SeqA protein and I conducted genetic studies to understand the physiological relevance of this interaction. I found that TnsE-mediated transposition was dramatically stimulated in *seqA* strains. While *seqA* mutants display a wide variety of phenotypes, I show that stimulation of TnsE-mediated transpositon is likely not explained by DNA replication over-initiation, DNA replication asynchrony, or changes in global gene expression. Instead, the stimulatory effect may be attributed to a change in the chromosome structure resulting from *seqA* inactivation. The results are consistent with a model where the SeqA filament that tracks with the replication fork may block TnsE access to its target generated during normal DNA replication; TnsE, however, may be able to overcome this barrier by interacting with SeqA. According to this model, the TnsE molecule that interacts with SeqA may be

different from the one that recognizes the transposition target as TnsE-mediated transposition occurs at a higher frequency in *seqA* mutants than in the wild type strain.

4.2. Introduction

As discussed in Chapter 1, transposon Tn7 exerts a high degree of regulation on transposition and is exceptional at target site selection (45, 74). The TnsABC+D pathway, which uses a sequence-specific DNA-binding protein, TnsD, directs Tn7 integration into a specific site (the *attTn7* site) on the chromosome found in many bacteria for high frequency transposition (5, 58, 90). Even though the sequence recognized by TnsD is within a highly conserved region of the *glmS* gene that is essential for cell viability, the actual insertion site is downstream of the open reading frame of this gene. This mechanism ensures that the TnsD-mediated transposition avoids negatively affecting the host. No adverse effect on the host has been detected for insertion in the *attTn7* site and it is considered as a “safe haven” for transposon Tn7 to reside in a cell. This pathway would thus facilitate the vertical transfer of Tn7. The TnsABC+E pathway, on the other hand, has evolved to facilitate horizontal dissemination by preferentially targeting mobile genetic elements (conjugal plasmids and bacteriophage) that can transfer between bacterial cells (23, 96).

Our *in vivo* and *in vitro* studies discussed in Chapter 2 have demonstrated that TnsE targets an aspect of lagging-strand DNA replication for transposition by recognizing the β-clamp processivity factor of DNA replication at the primer/template junction (69, 72). Such a complex should be ubiquitous in DNA replication. An emerging picture in DNA replication and repair is that replication of both the leading-strand and lagging-strand is

essentially discontinuous. Even under normal growth conditions, the replication machinery is believed to frequently encounter various obstacles, including various forms of DNA damage and protein complexes on the DNA template (18, 30, 31, 57, 97). Because of the discontinuous nature of the lagging-strand synthesis, the DNA polymerase for the lagging-strand DNA replication would simply bypass barriers by resuming replication downstream (32, 53, 67). Emerging evidence also indicates that the leading-strand polymerase can restart DNA replication downstream of obstacles, leaving behind single-strand gaps (30, 54, 75, 76, 79, 97), which would be potential targets for TnsE-mediated transposition.

TnsE-mediated transposition exhibits a strong preference for specific DNA replication processes. TnsE directs Tn7 insertion into many different sites in the chromosome at low frequency (72, 73). However, TnsE-mediated transposition is stimulated over 100-fold by conjugating plasmids in the recipient cell where the transferred strand is replicated discontinuously, with 99% these insertions occur into the much smaller mobile plasmids (96). When mapped in a *recA* background, TnsE-mediated transposition into the chromosome was found to occur in one orientation with the progression of the lagging-strand DNA replication (72). This together with the strong orientation bias of TnsE-mediated transposition into the conjugal plasmids indicates that the TnsE-dependent pathway targets discontinuous DNA replication (72). TnsE-mediated transposition was also found to be highly stimulated by inducing double-strand breaks or in the presence of DNA damaging agents (72, 82), where TnsE is believed to target the lagging-strand DNA replication involved in DNA repair. However, it is unclear why discontinuous replication

of the conjugal plasmid is a better target than lagging-strand replication in the chromosome. One obvious difference is that replication of the leading-strand and lagging-strand template strands in the chromosome is coupled, but replication of the conjugal plasmid during transfer is not. It is not clear that how repair-associated replication proceeds.

TnsE-mediated transposition also has a regional preference in a given replicon. Transposition events that occur into the conjugal plasmids are concentrated in the leading region that enters the recipient cell first in the conjugation process (96). Transposition into the chromosome preferentially targets the terminus region for transposition, with more than 70% of the insertions events occur into this region but an dearth of insertions closer to the origin of chromosomal replication (73). As discussed in Chapter 1, the preference of TnsE-mediated transposition may simply reflect the accessibility of the target complex on different replicons or different regions of a specific DNA molecule. However, the mechanism that determines such availability is not clear.

Transposons are generally considered parasitic to the cell. The mobilization of transposable elements has to be tightly controlled in terms of both the transposition activity and/or target specificity to avoid the accumulation of mutations deleterious to the host. Many mechanisms have been found to regulate the activity of transposons, including the expression levels of the transposase proteins at both the transcription and translation levels, the activity of the transposase, the structure or supercoiling of the donor and the target DNAs, and the regulation mechanisms tend to be transposon specific (reviewed in (42, 59)). It was found in previous genetic screens that TnsE-mediated

transposition was stimulated in a *dam* strain (22). Dam methylase (encoded by the *dam* gene) is responsible for the methylation at the N-6 position of the adenine in GATC sites. Dam methylation was known to regulate the transposition activity of Tn5 (98) and Tn10 (77) by modulating the expression and/or the activity of the transposase through the GATC sites in the promoter of the transposase gene or in regions where the transposase binds in the element. However, mutating the GATC sites in Tn7 had no effect on transposition, indicating that hemi-methylation of the element itself did not explain the change in transposition frequency (22). The effect of *dam* inactivation on TnsE-mediated transposition was confirmed in our lab, but the mechanism for this stimulation was not clear (see below)(68).

My attempt to search for proteins that interact with TnsE led me to investigate the role of the SeqA protein in TnsE-mediated transposition. SeqA is involved in regulating replication initiation by sequestering the newly replicated origin from premature reinitiation for about one-third of a generation time (49, 89). SeqA helps regulate replication initiation by altering the rate of Dam methylation of the hemimethylated GATC sites at the replication origin, which in-turn alters the ability of DnaA to promte initiation. In addition to its role in initiation regulation, SeqA has also been found to bind hemimethylated GATC sites throughout the chromosome arising after the passage of the replication fork (10, 65). By cooperatively binding to a cluster of GATC sites, SeqA proteins oligomerize to form a filament with left-handed helical structure through the aggregation of the N-terminal domains (9, 25, 41, 63). SeqA-GFP fusion proteins were found to form foci that migrate with the replication complex *in vivo* and the formation of

SeqA foci depends on active DNA replication (10, 33, 34, 65). The chromosomes of *seqA* strains have been reported to have increased negative supercoiling, indicating a role of SeqA in organizing the chromosome (94, 95).

In this chapter, I revealed and confirmed a physical interaction between TnsE and SeqA. The functional relevance of this interaction was established by the finding that TnsE-mediated transposition was specifically stimulated 10-fold in a *seqA* strain. Results from genetic studies suggest a model where the SeqA filament may block TnsE access to newly replicated DNA, but TnsE may be able to disrupt such a superstructure by interacting with SeqA, thereby facilitating targeting.

4.3. Results

4.3.1. Identification of a SeqA-TnsE interaction

Using a gapped structure preloaded with the β-clamp processivity factor I was able to reconstitute an *in vitro* transposition reaction with purified TnsABCE proteins, where the transposition profile recapitulates that observed *in vivo* (69). However, *in vitro* transposition occurred at a very low frequency even with a mutant TnsE that allows a high level of transposition *in vivo* (72). In the TnsABC+D transposition pathway, host proteins ACP and L29 were found to stimulate both the DNA-binding ability of TnsD and TnsD-mediated transposition *in vitro*; mutating L29 also resulted in a drastic decrease in TnsABC+D transposition *in vivo* (81). I speculated that there may be other factors that specifically stimulate or inhibit TnsE-mediated transposition yet to be identified. I sought to identify host-encoded proteins that interact with TnsE by an affinity purification-mass spectrometry strategy (44). In this assay, *E. coli* cell lysates

were passed through an affinity column bound with highly purified TnsE-6HIS proteins. The idea to use highly purified TnsE proteins was to make sure that the proteins identified at the end are purely because of their interaction with TnsE but not with proteins contaminating the TnsE protein preparations. The column was extensively washed before the bound proteins were eluted and separated by SDS-PAGE. Proteins specifically bound to TnsE (as compared to control columns) were then identified using mass spectrometry. Eight proteins were initially identified in this analysis, including SeqA (sequestration of origin), L14 (ribosomal protein), CadA (lysine decarboxylase), NudC (NADH pyrophosphatase), RarA (recombination protein involved in processing recombination intermediates at replication forks), SlyD (FKBP-type peptidyl-prolyl cis-trans isomerase), ArgE (Acetylornithine deacetylase), and Yqj (Uncharacterized protein)(www.shigen.nig.ac.jp/ecoli/pec/). The biologic relevance of these potential candidates in TnsE-mediated transposition was examined using an *in vivo* transposition assay. No significant change in TnsE-mediated transposition *in vivo* was found when genes encoding proteins NudC, CadA, RarA, SlyD, and ArgE were individually knocked out. The role of L14 in TnsE-mediated transposition was not tested in this way because it is essential for cell viability. However, a dramatic change in TnsE-mediated transposition was found when the *seqA* gene was inactivated (see below). Given the role of ribosomal protein L29 in the TnsABC+D transposition (81), I purified L14 and studied its interaction with TnsE biochemically in addition to SeqA.

4.3.2. A direct physical interaction between TnsE and SeqA

To confirm and characterize the interactions between TnsE and SeqA and L14, proteins were individually purified and protein-protein interaction was investigated by utilizing a protease-footprinting assay that has been broadly used to study protein-protein interactions (48, 60, 78). Results are shown in Figure 4.1. I found that SeqA could protect one or more trypsin sensitive sites in TnsE; following digestion with 10 ng of trypsin, the two TnsE bands indicated by red arrows at ~31-45 kD found with BSA or L14 were lost in the lanes where TnsE was incubated with SeqA and a new band (~55 kD) and intensified bands in the ~64 kD range were found. Protection of the 64 kD TnsE band was even more obvious at higher concentrations of trypsin. There appeared to be an additional increase in the stability of many TnsE bands when both L14 and SeqA were included in the reaction. Furthermore, bands of sizes smaller than 6.5 kD showing degraded TnsE found at higher trypsin concentrations were not found with the inclusion of SeqA. The results indicate a direct interaction between TnsE and SeqA. It is possible that L14 was initially pulled out through its interaction with SeqA. Perhaps these interactions are relatively weak; no TnsE-SeqA, SeqA-L14, or TnsE-SeqA-L14 complexes were found in a gel filtration analysis aimed at further characterization of these interactions.

4.3.3. Inactivation of *seqA* specifically stimulates TnsE-mediated transposition

To determine the physiological role of the TnsE-SeqA interaction *in vivo*, I monitored Tn7 transposition by a lambda hop assay (Material and Methods) in both the wild type strain and *seqA* mutants. A ten-fold increase in TnsE-mediated transposition was found in a *seqA::*Tet^R mutant background over the levels found in the wild type strain (Figure

Figure 4.1. Protease foot-printing suggests a specific interaction between TnsE and SeqA. L14 appears to interact with TnsE in a SeqA-dependent way. TnsE (1.8 ug) was incubated with the same mole of BSA, L14, SeqA, and SeqA+L14. Protein mixtures were digested with the protease trypsin (10, 33, or 100 ng), separated by SDS-PAGE, and TnsE fragments was detected by Western blot. Total protein amounts in the reactions were balanced with BSA.

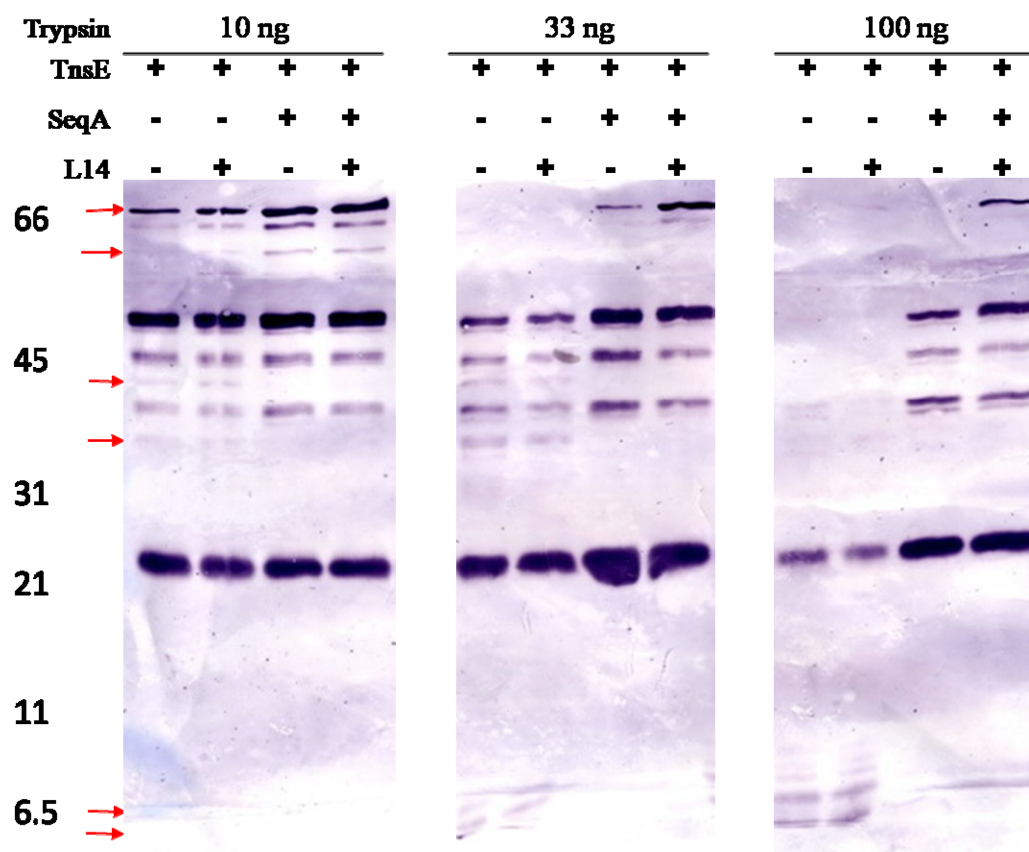
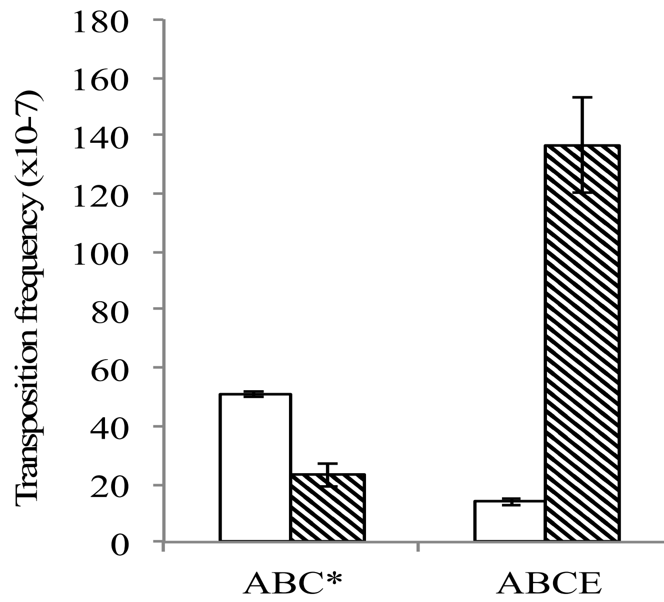
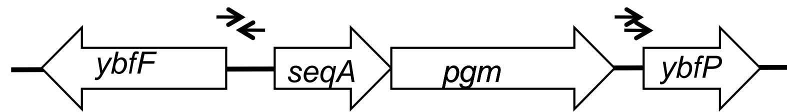

Trypsin
10 ng
33 ng
100 ng
TnsE
SeqA
L14
66
45
31
21
11
6.5

Figure 4.2. (A) TnsE-mediated transposition is stimulated by inactivating *seqA*. The *seqA* allele is *seqA::*Tet^R. Transposition was monitored by lambda-hop assay with transposition proteins expressed from pCW15 for TnsABC, pCW15* for TnsABC*, and pJP104 for TnsEs. (B) Structure of the *seqA* locus. Wide arrows indicate the directions and extents of genes in the locus. Transcription promoters found in the locus and their direct are indicated by the arrow heads. Annotation is based on www.shigen.nig.ac.jp/ecoli/pec/ and (50). (C) Polar effect of *seqA* inactivation does not account for the stimulation on TnsE-mediated transposition. The *seqA* allele is *seqA::*Tet^R and transposition was monitored by lambda hop assay with transposition proteins expressed from pQS102 for TnsABCE and pQ107 for TnsABC*. Open bars indicate wild type background, hatched bars for *seqA* mutant, and cross hatched bars for *pgm* mutant. Error bars indicate the standard deviation of triplicate.

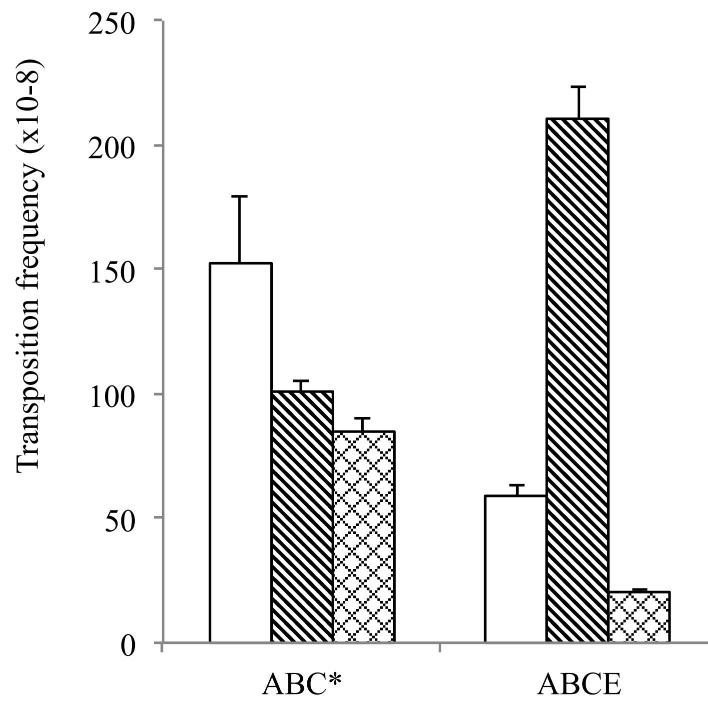
A



B



C



4.2A). This effect is specific for the TnsE-mediated transposition as no significant change was observed with transposition levels mediated by the TnsABC* mutant core machinery. The TnsABC* machinery contains a gain-of-activity mutation in the regulator protein of Tn7 transposition (TnsC) that allows the otherwise inactive core transposition machinery TnsABC to promote Tn7 transposition without the target site selecting protein TnsD or TnsE (86). The mutant core machinery serves as a control for effects specific on TnsE target selection, but not the transposition process in general (82, 86). Similar results were found with another *seqA* mutant allele, *seqA::*Cm^R*::frt* (Figure 4.4A) and the stimulation seems to depend on the level of transposition proteins (compare Figure 4.2A and Figure 4.2C).

The *seqA* gene is located in an operon with *pgm* gene that encodes phosphoglucomutase and both genes are transcribed from a promoter upstream of the *seqA* gene (Figure 4.2B)(50). Inactivating the upstream *seqA* gene in the *seqA::*Tet^R and *seqA::*Cm^R*::frt* alleles would have a polar effect on the downstream *pgm* gene, and a *seqA::*Tet^R mutant was found to exhibit PGM- phenotype, lacking the ability to ferment galactose (50). I directly tested the effect of inactivating the *pgm* gene on TnsE-mediated transposition by using a Δ*pgm::frt* mutant; no stimulation on TnsE-mediated transposition was found (Figure 4.2C). I also tried to rule out the possible complication of inactivating the *pgm* gene on the stimulation found with the *seqA* mutants by making an in-frame deletion Δ*seqA::frt* strain from Δ*seqA::*Kan^R*::frt* allele, where the kanamycin cassette was deleted by FLP recombinase as described (21). TnsE-mediated transposition was not found to be stimulated in a Δ*seqA::frt* strain. As *seqA* mutants are notorious in

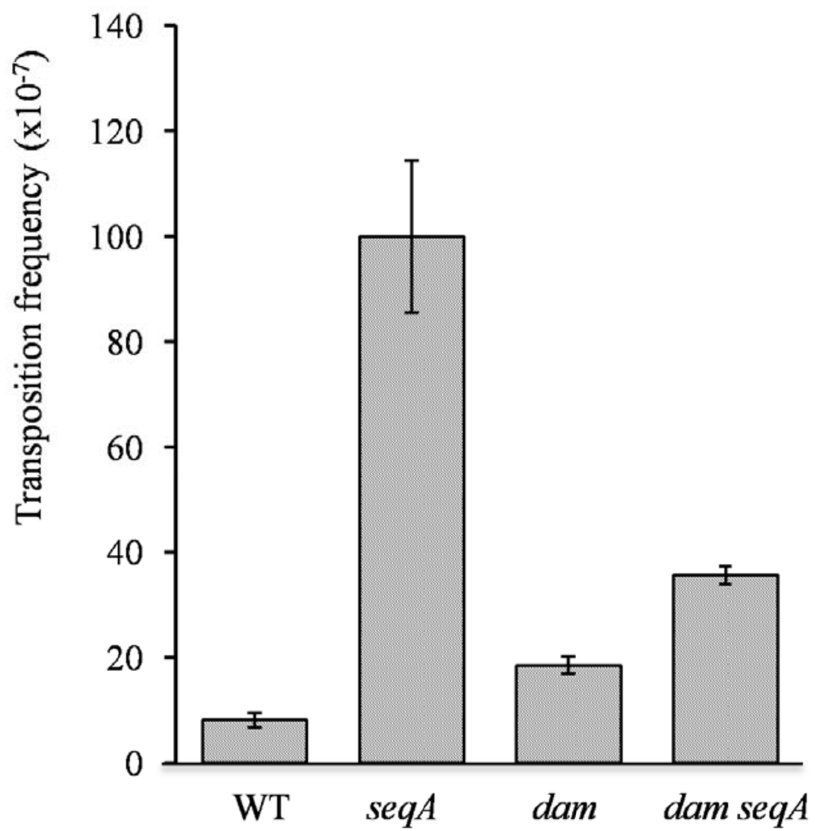
accumulating suppressor mutations (93), I suspect that the 42°C incubation steps used to cure the pCP20 plasmid resulted in suppressor mutations in the Δ*seqA::frt* strain. However, it is currently not clear where in the chromosome suppressors frequently arise in *seqA* mutants and further study would be needed to test this hypothesis.

4.3.4. SeqA and Dam are not epistatic on TnsE-mediated transposition

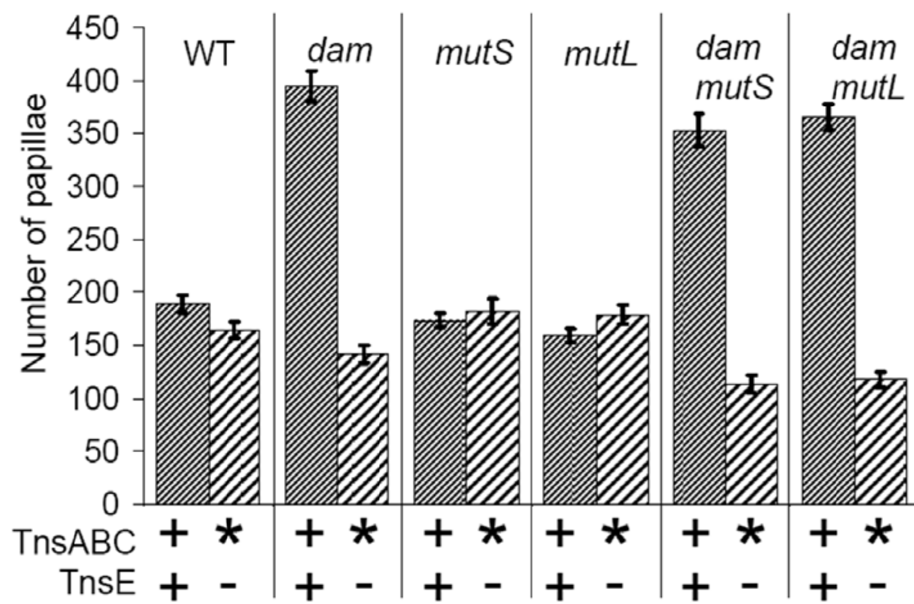
Since SeqA specifically binds hemimethylated, but not unmethylated GATC sites (8, 83), a substrate of Dam methylase, I tested if *dam* and *seqA* are epistatic to TnsE-mediated transposition. It was found previously using a papillation assay that inactivation the Dam methylase specifically stimulated TnsE-mediated transposition (68). I confirmed this finding using a lambda hop assay that TnsE-mediated transposition increased about two-fold in a *dam* strain, which is modest compared to the increase found in the *seqA* mutants (Figure 4.3A). Furthermore, mutating both *seqA* and *dam* resulted in an intermediate level of TnsE-mediated transposition (Figure 4.3A). This result was unexpected because SeqA is known to not bind unmethylated GATC sequences (8, 83).

Due to the reliance of many systems on the methylation status of DNA in *E. coli*, a *dam* strain has a range of defects, including chronically induced SOS response, accumulation of DNA double-strand breaks, dependence on RecA for cell viability, and changes in the expression of a number of genes (47). Since double-strand breaks can stimulate TnsE-mediated transposition (73), previous attempts have been focused on this aspect of a *dam* strain. In *E. coli*, the mismatch repair system (MutHLS) is sensitive to the methylation status of DNA, where MutS detects mismatches and recruits MutL and MutH to act on the nascent DNA strand (40). Uncontrolled nicking activity of the MutH

Figure 4.3. Inactivating Dam methylase stimulated TnsE-mediated transposition but not because of increased double-strand breaks. (A) TnsE-mediated transposition is modestly stimulated in a *dam* strain. Transposition was monitored by lambda hop assay with transposition proteins expressed from plasmid pCW15 for TnsABC and pJP104 for TnsE. *seqA::*Tet^R allele was used. Error bars indicate the standard deviation of triplicate. (B) TnsE-mediated transposition is not stimulated in *dam* mutants because of uncontrolled nicking by the mismatch repair system. Transposition was monitored using a papillation assay with transposition proteins expressed from pCW15 for TnsABC and pJP104 for TnsE. Error bars indicate the standard deviation of average of twelve experiments.



Transposition frequency (x10-7)
140
120
100
80
60
40
20
0
WT
seqA
dam
dam seqA



Number of papillae
450
400
350
300
250
200
150
100
50
0
WT
dam
mutS
mutL
dam mutS
dam mutL
TnsABC
TnsE

would lead to DSBs (3, 62) that are already known to stimulate TnsE-mediated transposition (73, 82). This possibility was directly tested previously (68) and confirmed in this study using a papillation assay. TnsE-mediated transposition was not affected in *mutS* and *mutL* mutants. Furthermore, *dam mutS* and *dam mutL* double mutants had similar TnsE-mediated transposition frequencies as the *dam* single mutant (Figure 4.3B)(68). These and results discussed below indicate that the increase in TnsE-mediated transposition is resulted from the altered function of SeqA in the *dam* mutant.

Even though the SeqA protein is not essential in *E. coli*, a *seqA* strain has numerous pleiotropic phenotypes, including asynchronous replication, overinitiation, and altered gene expression and chromosomal structure (46, 49, 93-95). To figure out the reason for the stimulatory effect of *seqA* inactivation on TnsE-mediated transposition, I carefully examined Tn7 transposition in *E. coli* mutants that each manifests a known phenotype of the *seqA* strain. For each specific mutatnt, the frequency of TnsABC+E- and TnsABC*-mediated transposition were monitored. The TnsABC* transposition serves as a measure of the specificity of a genetic allele on the target selecting by TnsE, and also as an indicator of the cell viability.

4.3.5. Effect of replication initiation asynchrony on Tn7 transposition

DNA replication initiates from a single origin of chromosomal replication called *oriC* and progresses in both directions to the terminus region in bacteria with circular chromosomes. The initiator protein DnaA, which unwinds the DNA, controls replication initiation in *E. coli* and helps load the replicative helicase DnaB to establish a replication complex. Several mechanisms have been implicated to ensure that replication initiation

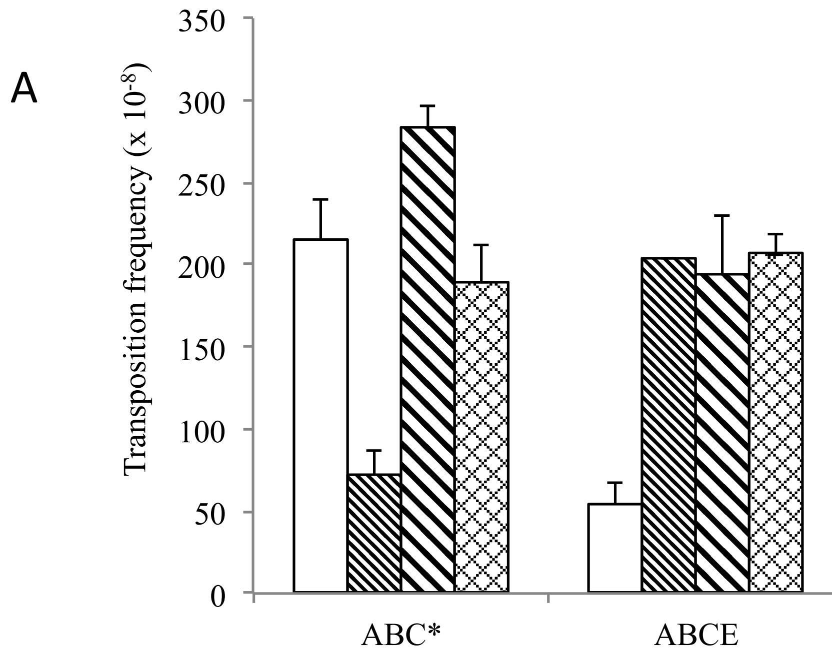
from all origins in a cell occurs at the same time, once and only once in a cell cycle, by sequestration of the newly replicated origin by SeqA (49, 89), titration of DnaA by the *dat* locus (29), and inactivation of DnaA by Hda (37). Cells that are compromised in any of these functions show asynchronous replication initiation as indicated by the inappropriate number of origins in the cell. The impact of asynchronous replication initiation on TnsE-mediated transposition was tested in many strain backgrounds. As shown in Figure 4.3A, TnsE-mediated transposition was stimulated to a greater level in a *seqA* mutant than a *dam* mutant, however, these mutants have similar levels of replication initiation asynchrony (64). A *dam seqA* double mutant, which may have less severe replication initiation asynchrony than a *dam* single mutant (64), had a higher TnsE-mediated transposition.

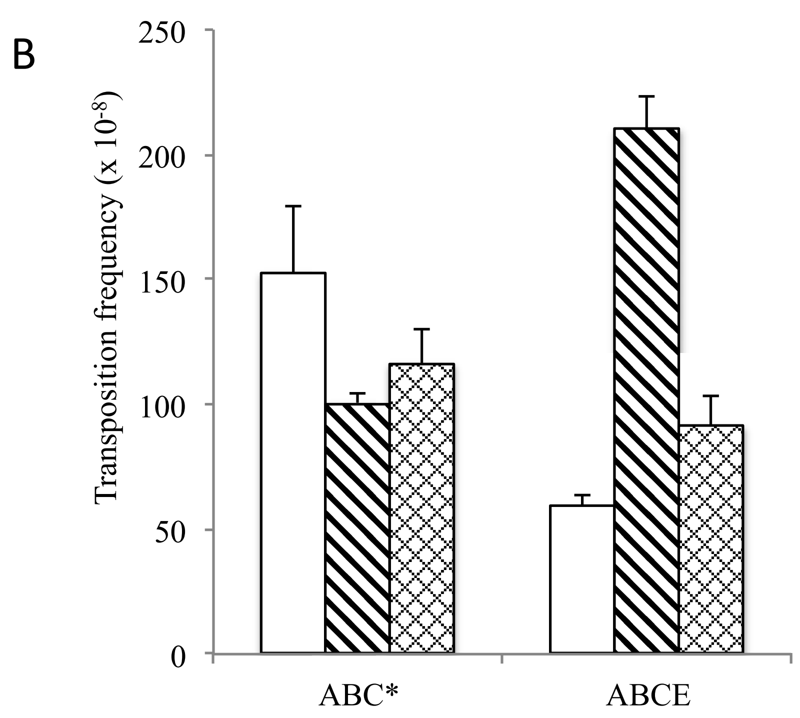
I further tested TnsABC+E transposition in strains carrying a temperature sensitive *danA46* allele, which has been shown to cause asynchronous initiation (28, 64). The effect of *dnaA46* appears to not be specific to TnsE-mediated transposition. TnsE-mediated transposition frequency in the *dnaA46 seqA* double mutant is similar to that in the *seqA* mutant (Figure 4.4A), even though the *dnaA46 seqA* strain has a much lower initiation asynchrony than the *seqA* or *dam* mutant alone (64). The transposition frequency does not correlate with the measured asynchrony index in these strains (64), indicating that asynchronous DNA replication does not necessarily lead to the stimulation in TnsE-mediated transposition found in the *seqA* mutants.

4.3.6. Effect of replication overinitiation on Tn7 transposition

In a *seqA* strain, some origins may be quickly reinitiated but some might avoid replication for a long time; this is because in cells without sequestration, all origins including the newly initiated ones can be equally bound by DnaA and will thus have the equal chance of being initiated when the initiation potential determined by the status of ATP-bound DnaA builds up (89). Overinitiation of replication could lead to double-strand breaks when the new forks converge with the forks stalled or collapsed ahead, which is believed to be the reason that *seqA* mutants are sensitive to DNA damage or replication inhibitors (88). Since double-strand breaks are known to stimulate TnsE-mediated transposition (73, 82), it is possible that DNA replication overinitiation could stimulate TnsE-mediated transposition. However, introducing a *dnaA46* allele, which reduces the replication initiation efficiency, did not affect the transposition frequency of inactivating *seqA* (Figure 4.4. A). I also tested TnsE-mediated transposition in an *hda* mutant, which has been shown to have more severe overinitiation than a *seqA* mutant (12). Hda is the protein that deactivates the DnaA protein by stimulating its ATPase activity, a mechanism that is often referred to as regulatory inactivation of DnaA (RIDA) that plays a predominant role in the control of replication initiation potential (13, 37). Mutating *hda* modestly stimulated TnsE-mediated transposition (Figure 4.4B). However, TnsE-mediated transposition in an *hda* mutant is not comparable to the frequency found in a *dam* or *seqA* strain. Since both TnsE and Hda require interacting with the β-clamp processivity factor for function (37, 69), it is reasonable to presume that the removal of a competitor for the access of TnsE to the β-clamp resulted in the increased level of TnsE-mediated transposition in the *hda* mutant.

Figure 4.4. Lack of SeqA function in replication initiation regulation does not account for the stimulatory effect of inactivating SeqA on TnsE-mediated transposition. (A) Effect of replication initiation asynchrony on TnsE-mediated transposition. Transposition was monitored in a lambda hop assay with transposition proteins expressed from pQS102 for TnsABC+E and pQS107 for TnsABC*. Open bars indicate the wild type background, thin striped bars for *seqA::*Cm^R*::frt*, thick striped bars for *dnaA46*, cross hatched bars for *seqA::*Cm^R*::frt dnaA46*. All these strains have a *tnaA::Tn10* allele which did not affect transposition (data not shown). Error bars indicate the standard deviation of triplicate. (B) Effect of DNA replication overinitiation on TnsE-mediated transposition. Transposition was monitored in a lambda hop assay with transposition proteins expressed from plasmids pQS102 for TnsABC+E and pQS107 for TnsABC*. Open bars indicate the wild type background, thin striped bars for *seqA::*Tet^R, and double hatched for *hda::*Tet^R. Error bars indicate the standard deviation of triplicate.


A
350
300
250
200
150
100
50
0
Transposition frequency (x 10-8)
ABC*
ABCE


B
250
200
150
100
50
0
Transposition frequency (x 10-8)
ABC*
ABCE

4.3.7. Effect of altered global gene expression on Tn7 transposition

SeqA has been implicated in transcription regulation, especially for genes involved in DNA metabolism (49, 84, 85). A *seqA* strain has been reported to have an altered global transcription profile similar to strains overexpressing the Dam methylase (46). I tested Tn7 transposition in strains with plasmids pMQ133 (pBR322-derivative)(2) and pMQ191 (pACYC184-derivative), which presumably would overexpress Dam at different levels. As shown in Figure 4.5, overexpression Dam methylase resulted in increased level of transposition but its effect may not be specific to TnsE-mediated transposition.

4.3.8. Effect of chromosome structure on Tn7 transposition

SeqA filament tracks with and organizes the sister chromosomes after replication (10, 20, 65). A *seqA* strain was reported to have a more compact chromosome and an increased level of negative supercoiling (94, 95), indicating a role of SeqA in chromosome organization. Knocking out the condensing protein MukB, another protein that is involved in chromosome organization, was found to be able to suppress the filamentation and asymmetric septation and relieve the over compactness of the chromosome of a *seqA* strain, even though the asynchronous replication of a *seqA* strain was not affected (1, 66, 95). Inactivating *seqA* was also found to suppress the poor nucleoid folding and temperature sensitivity of a *mukB* strain, but not the hypersensitivity of a *mukB* strain to the DNA gyrase inhibitor novobiocin (1, 66, 95). Purified SeqA was found to affect the activity of topoisomerase IV on a concentration-dependent manner; SeqA stimulates the relaxing and decatenation activity of Topo IV at low concentration, but inhibits at high concentration level (35). I tested the effect of chromosome structural

Figure 4.5. Effect of altered global gene expression on Tn7 transposition.

Transposition was monitored by a lambda hop assay with transposition proteins expressed from plasmid pQS102 for TnsABC+E and pQS107 for TnsABC*. Open bars indicate the wild type background, thin striped bars for *seqA::*Cm^{R}*::frt*, thick striped bars for wild type cells with pMQ133, and double hatched bars for wild type cells with pMQ191. Error bars indicate the standard deviation of triplicate.

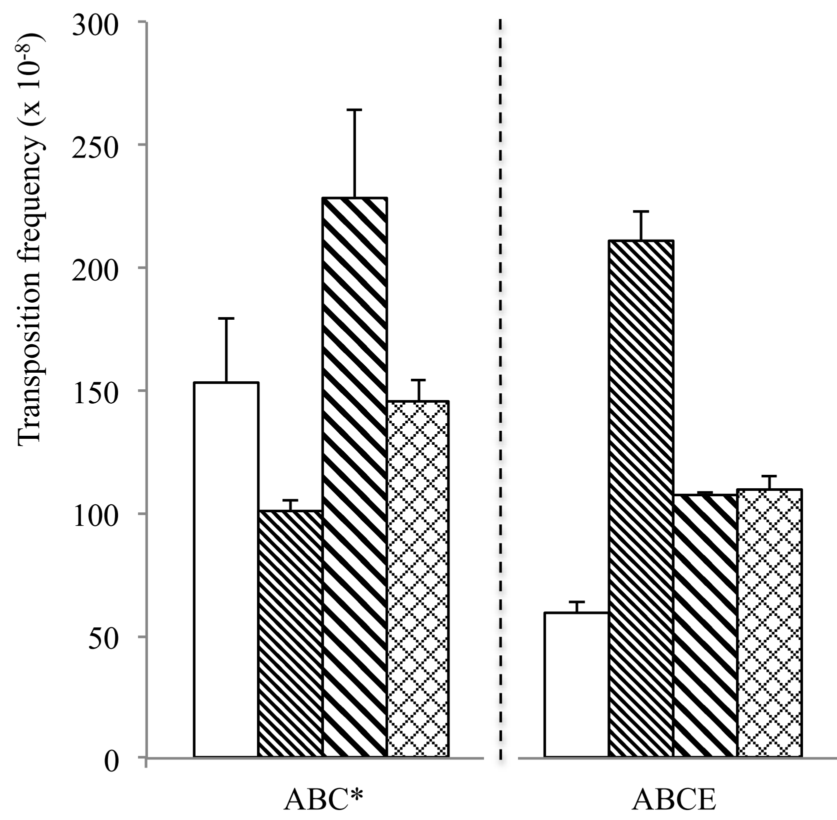
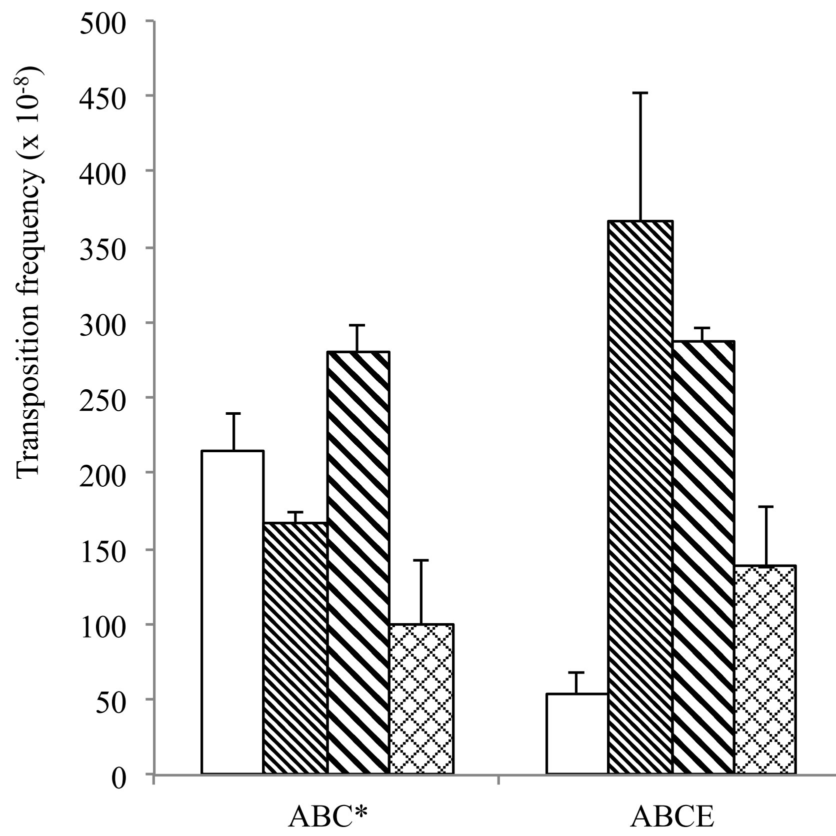

Transposition frequency (x 10-8)
300
250
200
150
100
50
0
ABC*
ABCE

Figure 4.6. Effect of chromosomal organization on Tn7 transposition. Transposition was monitored by a lambda hop assay with transposition proteins expressed from plasmid pQS102 for TnsABC+E and pQS107 for TnsABC*. Open bars indicate the wild type background, thin striped bars for *seqA::*Tet^R, thick striped bars for *mukB::*Cm^R*::frt*, and double hatched bars for *seqA::*Tet^R *mukB::*Cm^R*::frt*. Error bars indicate the standard deviation of triplicates.



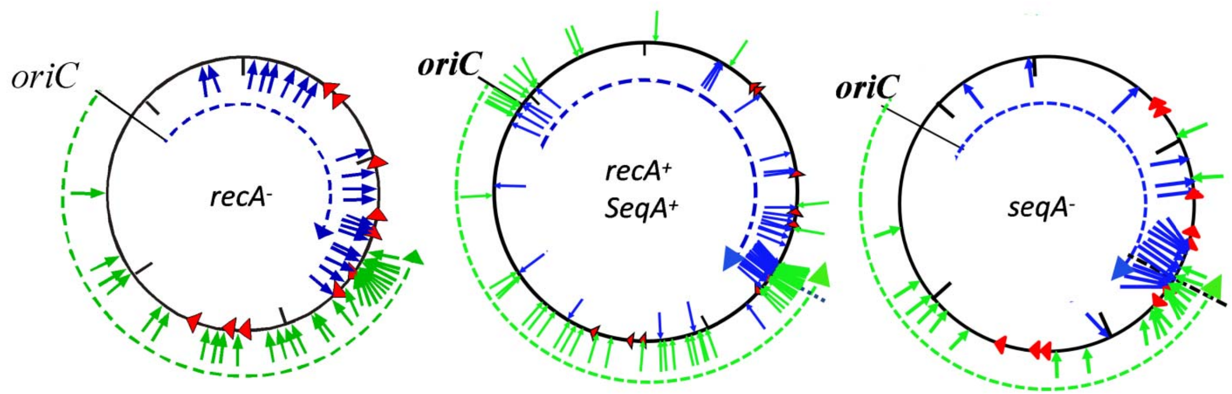
Transposition frequency (x 10-8)
500
450
400
350
300
250
200
150
100
50
0
ABC*
ABCE

alteration on TnsE- mediated transposition in the *mukB* and *seqA mukB* mutants. I found that even though a *mukB* mutation may affect Tn7 transposition non-specifically, TnsE-mediated transposition in the *seqA mukB* double mutant was significantly lower than that of the *seqA* mutant, indicating an effect of chromosome structure on TnsE-mediated transposition (Figure 4.6). The result is also consistent with results discussed above on that DNA replication asynchrony does not stimulate TnsE-mediated transposition, as a *seqA mukB* double mutant was found to retain this defect of a *seqA* mutant (95).

4.3.9. TnsE targets structures at the origin and leading-strand

Since *seqA* is synthetically lethal with *recA* (39), all of the lambda hop experiments discussed above were conducted in a *recA*$^+$ background. To gain more insight into the increased TnsE-mediated transposition frequency in the *seqA* mutants, we mapped insertions in the *seqA*$^-$ *recA*$^+$ and *seqA*$^+$ *recA*$^+$ backgrounds, and compared the transposition profiles with the previously reported profile in a *seqA*$^+$ *recA*$^-$ background (72). Transposition targeting in cells that were recombination proficient (*recA*$^+$) showed interesting differences from transposition found with the *recA*$^-$ cells. While TnsE-mediated transposition in all cells show a strong bias for the terminus region as found previously, transposition events found with *recA*$^-$ cells occurred in one orientation with lagging-strand DNA replication with a concentration in the terminus region, very few insertions were found around the broad region around *oriC* (Figure 4.7A)(72). However, in a wild type (*recA*+) host, many insertions appeared in a large region around the origin and there were many insertions occurred into the DNA strand undergoing leading-strand DNA replication (Figure 4.7B). Interestingly, the targeting pattern found with *seqA*

Figure 4.7. Location and orientation of independent TnsE-mediated transposition events in the chromosome of *recA*-(A), wild type (B), and *seqA*- (C) backgrounds. Positioning of the arrow inside vs. outside indicates opposite left-to-right orientation. Dashed line indicates the direction of DNA replication forks. All transposition events were isolated from lambda hop assays with transposition proteins expressed from pCW15 for TnsABC and pJP104 for TnsE.



oriC
recA-
oriC
recA+
SeqA+
oriC
seqA-

strains (Figure 4.7C) strongly resembles the pattern found in the *recA*⁻ background (Figure 4.7A); nearly all of the insertions occurred in one orientation with lagging-strand DNA replication, with very few insertions occurred in the origin proximal region of the chromosome, unlike the wild type strain (Figure 4.7B).

4.4. Discussion

Transposon Tn7 is sophisticated in target selection. In this study we made two new discoveries that may contribute to a better understanding of the regulation of Tn7 transposition *in vivo*. I show that TnsE physically interacts with SeqA and that Tn7 transposition is stimulated for about 10-fold in a *seqA* mutant. This stimulatory effect is specific for the TnsABC+E transposition pathway; no stimulation was observed with the untargeted transposition mediated by the TnsABC* core machinery. Genetic studies ruled out DNA replication overinitiation and DNA replication asynchrony, two conditions that can lead to double-strand breaks when the new forks collide in a head-to-tail fashion with the forks stalled ahead, something that could attract TnsE-mediated transposition events. Changes in global gene expression also probably do not account for the incrase in TnsE-mediated transposition. Instead, my results indicate that the activity of SeqA on organizing the newly replicated sister chromosomes may be responsible. I propose a model where TnsE interacts with SeqA to disrupt the oligomerized structure of SeqA that follows DNA replication fork and thereby gains access to its preferred transposition target complex generated during DNA replication. The TnsE molecule that interacts with SeqA would be different from the one that interacts with the β-clamp and the DNA substrate at the actual target. This is consistent with the fact that TnsE-mediated

transposition is SeqA-independent; it would be that the role of TnsE for disrupting SeqA filamentation is not necessary in a *seqA* mutant. In Chapter 3, I found that TnsE homologs from *I. loihiensis* and *S. baltica* can promote transposition of the miniTn7 element with TnsABC from *E. coli* in the presence of the clamp from the same host but the transposition mediated by these hybrid machineries occurred at low frequencies. In view of the TnsE-SeqA interaction in transposition, it would be interesting to test if transposition mediated by the TnsE homologs is also stimulated in a *seqA* background, since both *I. loihiensis* and *S. baltica* have the SeqA homologs.

SeqA binds the hemimethylated GATC sites in the origin to regulate the timing and synchrony of DNA replication and to prevent premature reinitiation (7, 49, 89, 93). SeqA can also organize the newly replicated sister chromosomes by binding to hemimethylated GATC sites transiently available following DNA replication: binding of SeqA to clusters of hemimethylated GATC sites has been shown to form left-handed helical filaments that can restrain negative supercoiling (25, 38, 63); this form of SeqA is believed to be the SeqA complex that tracks with the replication fork (10, 34, 66). There is a large number of GATC sites in the *E. coli* genome (19,130 copies (80)) but only limited amount of SeqA proteins in the cell (1,000 copies (83)). Given that SeqA binds preferentially to hemimethylated GATC sites (83) and that SeqA oligomers can attract free SeqA proteins (27), it is proposed that the majority of SeqA molecules are bound at the replication fork (25). Using genetic studies I show that the simulation on TnsE-mediated transposition in the *seqA* strain is not due to DNA replication overinitiation or asynchrony but likely because of the change in chromosome structure. Since TnsE senses ongoing DNA

replication, it is possible that the oligomerized structure of SeqA filament at the replication fork could physically block the immediate access of TnsE to a potential target generated during DNA replication. It is also possible that some types of chromosome organization facilitated by the action of SeqA can shield targets that would be more easily recognized by TnsE. Even though the genome-wide distribution of SeqA is still somewhat controversial (80, 92), it is likely that regions rich of GATC sites would occlude TnsE from accessing the replication fork but regions with few GATC sites would be more accessible. In our mapping data, TnsE-mediated transposition in all backgrounds (wild type, *recA*⁻, *seqA*⁻) was found to be concentrated in the terminus region where GATC sites are depleted (20, 80, 92). The observed stimulatory effect of inactivating *seqA* on TnsE-mediated transposition could also be explained as a result of the disappearance of the SeqA filament barrier. Further study of TnsE-mediated transposition in strains containing *seqA* mutations that only affect the aggregation activity of the protein (16, 25) would be informative to this model.

GATC sites are methylated by the Dam methylase (52). The biochemical activity of SeqA would predict that SeqA would function epistatically with the Dam methylase, since SeqA binds preferentially hemimethylated GATC motifs but not unmethylated DNA at all (8, 83). However, we found that the stimulatory effect of *seqA* inactivation is much greater than *dam* knockout on TnsE-mediated transposition, indicating a function of SeqA that is not completely Dam methylation dependent. The existence of an undiscovered function of SeqA that is independent of Dam methylation of the GATC sites has been suggested before. It was found previously that a *seqA dam* double mutant

grew better than a *seqA* single mutant and overexpression of SeqA was lethal to *dam::*Tn9 allele (49), indicating Dam methylation and SeqA can each influence the cell viability in the absence of the other. In another study, the bulk chromosome DNA was found to be decompacted in a *seqA* null mutant but not in a *dam* null mutant (1). No studies have followed up on these observations to my knowledge and it is unclear how SeqA would function on the unmethylated DNA in a *dam* cell. One possibility is that the other adenine methyltransferases in *E. coli* would play a role in the absence of Dam methylase, even though these methyltransferases may have different substrates than Dam and are not expressed to high concentrations under normal growth conditions (11, 52). In support of this, it has been shown that purified SeqA proteins could bind mutated GATC sites where the guanine, thymine, and cytosine were individually changed to other nucleotides, even though with a decreased affinity (24).

The role of SeqA in chromosome structure has been multifold. By forming helical filament that organizes newly replicated chromosomes SeqA is able to constraint negative supercoiling (25, 38). SeqA could also interact with ParC and purified SeqA proteins affect the activity of the latter in a concentration dependent manner: it stimulates the relaxing and decatenation activity of ParC at low concentration but inhibits the activity of ParC at high concentration (35). The idea that most of the SeqA proteins in a cell should be bound in the filaments would suggest that the possibility of a SeqA filament inhibiting the activity of topoisomerases is more physiologically relevant. However, a *seqA* strain was found to have increased negative supercoiling and this alteration could be partially compensated by mutating *mukB* (94, 95). In the analysis of

the effect of chromosome organization on TnsE-mediated transposition, I found that a *seqA mukB* double mutant had a lower transposition frequency than inactivating *seqA* alone. The mechanism of chromosome condensation by MukB is not clear. Current models suggest that MukB plays a structural role in stabilizing chromosome structures (14, 19), a mechanism completely different from and independent of SeqA. Therefore, it is expected that *mukB* mutation would not specifically affect TnsE-mediated transposition and the *mukB seqA* double mutant would not fully restore the block to the access of TnsE to DNA replication.

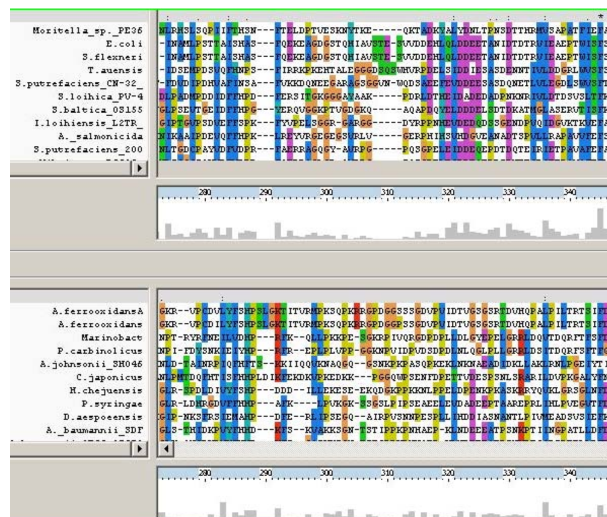
TnsE targets DNA replication by interacting with the β-clamp, a protein that plays a central role in coordinating DNA replication and repair (69). The relatively weak binding affinity of TnsE-β interaction and that only extremely high levels of TnsE expression can induce the SOS response in the cell indicate that TnsE does not have the ability to compete with replication components under normal replication. This is consistent with the fact that TnsE-mediated transposition into the chromosome is more frequently found in the lagging-strand of DNA replication, where DNA polymerase III frequently dissociates from the clamp upon finishing an Okazaki fragment (43, 87). The fact that TnsE-mediated transposition is stimulated by the presence of conjugal plasmids further indicates that TnsE is more capable of targeting uncoupled DNA replication. In this study, I found that TnsE was able to target DNA undergoing leading-strand DNA replication for transposition in a wild type background. It is attempting to suggest that the actual event that is targeted by TnsE in these cases involves a blocked or abandoned replication process. The actual target of TnsE could be the single-strand gap structures

behind the replication fork and/or a process involved in the repair of these structures in the leading-strand DNA, as repair-associated DNA replication has been shown to be especially vulnerable to TnsE-mediated transposition in our previous studies (73, 82). The low frequency probably reflects the frequency of direct restart of the leading-strand DNA replication and/or the transient nature of the single-strand gaps generated in the process. The fact that no such transposition events were found in either *seqA* or *recA* mutants indicates that recombination is needed to either make the target recognized by TnsE or to allow cells to survive the event that produces the structure. RecA is critical for the various repair pathways (51) and SeqA has also been shown to be essential when homologous DNA recombination is compromised (39, 88). The mechanism for the observed synthetic lethality of *recA* and *seqA* is unclear; functions of SeqA in chromosome structure organization (39) and in preventing replication over-initiation (88) have been suggested to be needed in *recA* null mutants. We also found that in the wild type background TnsE-mediated transposition occurred in both orientations in a small region around the origin of DNA replication. These events were also not observed in both *recA* and *seqA* null mutants. Presumably abortive replication initiation (61) or cellular events dealing with the abortive initiation could be recognized by TnsE. The dependence on RecA of insertions into both the leading-strand and the origin region is consistent with a previous finding that TnsE-mediated transposition in a *recA* background is lower than in a wild type strain in a papillation assay (68).

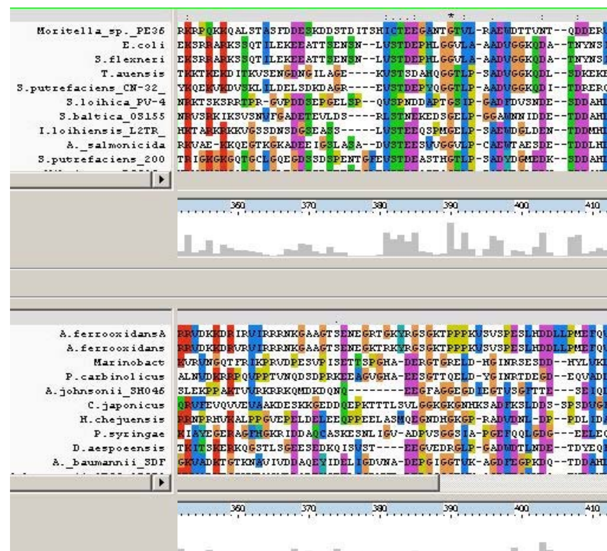
A physical interaction between TnsE and SeqA has been found by a pull-down assay and confirmed by a protease-footprinting analysis. However, the interaction seems to be

Figure 4.8. Potential regions of TnsE involved in SeqA interaction. TnsE homologs from bacterial hosts with and without SeqA were compared by profile-profile alignment using ClustalX program.

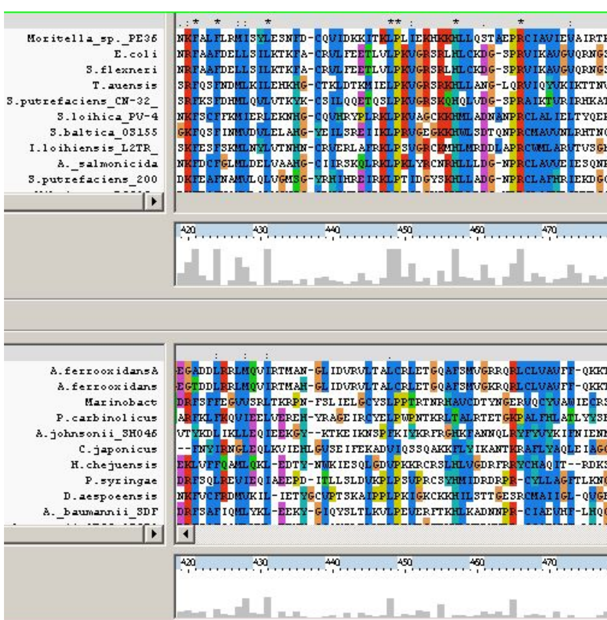
A



B



C



relatively weak, as no SeqA-TnsE complex was found in a gel filtration experiment. How TnsE interacts with SeqA is also at present unclear. As discussed in Chapter 3, Tn7-like elements have been identified in diverse bacteria and that TnsE interaction with the β-clamp for targeting is largely conserved among Tn7-like elements. However, SeqA is only found in *E. coli* related species (34). In my preliminary analysis, a SeqA homolog was not found in *Acidithiobacillales*, *Pseudomonadales*, and *Oceanospirillales* species of *Gammaproteobacteria*, where Tn7-like elements have been identified (Appendix 1)(70). Comparison of TnsE homologs from *seqA*$^{+}$ bacterial species with those from *seqA*$^{-}$ bacteria has identified several regions in TnsE that are different between the two groups (Figure 4.8). Further studies are needed to test if these regions participate in TnsE-SeqA interaction. Interetingly, a number of TnsE mutants with increased activity isolated from a genetic screen (Appendix II) contain mutations in these regions. The current model would predict that these mutants interact with SeqA with a higher affinity.

4.5. Materials and Methods

4.5.1 Media, chemicals and enzymes

LB and Minimal Media were prepared as described by Miller (56) and 0.2% glucose was used to suppress transcription from P*lac* promoter when necessary. Lac phenotypes were evaluated on MacConkey lactose agar (Difco). Antibiotics were used at the following concentrations: ampicillin (Amp), 100 μg/ml; chloramphenicol (Cm), 30 μg/ml; kanamycin (Km), 50 μg/ml; tetracycline (Tet), 20 μg/ml.

4.5.2. Bacterial strains and plasmids

E. coli strains used in this study are listed in Table 1. All are derivatives of *E.coli* K-12 and genetic alleles were introduced into new strain background by P1 *vir* mediated generalized transduction with standard method (56, 71). Strains ZL687, ZL688, and ZL794 were constructed by removing the drug markers from frt-containing constructs with the FLP recombinase produced from pCP20, which was subsequently cured at 42°C as described (21). The constructed strains were further tested for the known phenotypes of a specific allele, such as temperature sensitivity and UV sensitivity. Colony polymerase chain reactions (PCR) followed with direct sequencing were used to further confirm the successful construction of strains. *E. coli*strain *DH5α* was used as the host strain for plasmid construction and propagation.

Plasmids used in this study are listed in Table 2. The *rplN* gene encoding the ribosomal protein L14 was amplified by PCR using primers JEP416 (5'-CTC CCC TCG CCA TAT GAT CCA AGA ACA GAC-3') and JEP417 (5'-CTT CTC ACG CTC TTC CGC AGA GTA CTT CTG GTG CCA) and cloned into vector pTYB1 between the NdeI and SapI recognition sites. The final construct pZL123 was confirmed by sequencing.

4.5.3. Transposition assays

The frequency and targeting of transposition was monitored either by a papillation assay or a lambda hop assay. Transposition proteins were expressed *in trans* from plasmids, pCW15+pJP104 or pQS102 for TnsABC+E, pCW15* or pQS107 for TnsABC* mutant core machinery. pCW15 with pTA106 or pQS100 serves as negative control.

Table 1. Strains used in this study.

Strain	Genotype	Reference
DH5α	F- λ- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80 dlacZ ΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+)*	Laboratory stock
MC4100	F- *araD139* Δ*(argF-lac) U169 prsL150 relA1 deoC1 fthD5301 pstF25 rbsR22 e14⁻ (fimB-fimE)632::IS1 (fruK-yeiR)725*	Laboratory stock
NLC28	MC4100 ValR	Laboratory stock
NLC51	NLC28 *recA56*	Laboratory stock
JP617	NLC28 *attTn7::miniTn7::lacZYA*-::KmR (R90)	Laboratory stock
BB101	*F' lacI^q lac' pro'/ara*Δ*(lac-pro) nalA argE*(*am*)Rif^R *thi-1 slyD*::Kan^{R}λ*DE3*	Jennifer Surtees (15)
BW25113	λ–Δ*(araD-araB)567* Δ*lacZ4787(::rrnB-3)rph-1* Δ*(rhaD-rhaB)568 hsdR514,*	(21)
ER2925	*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1* R(*zgb210::*Tn10)Tet^S*endA1 rpsL136dam13::*Tn9 *xylA-5 mtl-1 thi-1 mcrB1 hsdR2*	NEB
ER2566	F- λ- *fhuA2 (lon) ompT lacZ::*T7 *gene1 gal sulA11 Δ(mcrC-mrr)114::*IS10 R(*mcr-73::miniTn10*-Tet^S)2 R(*zgb-210::*Tn10)(Tet^S) *endA1* (*dcm*)	NEB
GM2165	F- *dcm-6 thr-1 ara-14 leuB6 tonA31 lacY1 tsx-78 supE44 galK2 galT22 hisG4 rpsL136 xyl-5 mtl-1 thi-1 mutS215::* Tn10	Martin Marinus
GM2166	F- *dcm-6 thr-1 ara-14 leuB6 tonA31 lacY1 tsx-78 supE44 galK2 galT22 hisG4 rpsL136 xyl-5 mtl-1 thi-1 mutL218 ::*Tn10	Martin Marinus
JC215	MG1655 Δ*hda::*tet^R	Elliott Crooke (12)
JC326	MG1655 Δ*seqA::*tet^R	Elliott Crooke (12)
JW0674	MG1655 Δ*seqA::*Kan^R*::frt*	Keio collection (4)
JW0675	MG1655 Δ*pgm::*Kan^R*::frt*	Keio collection (4)
JW3350	MG1655 Δ*dam::*Kan^R*::frt*	Keio collection (4)
KA452	*thyA thr tyrA(Am) trpE9829(Am) metE deo dnaA46^{ts} tnaA::*Tn10	Mark Sutton (36)

Table 1 Continued.

Strain	Genotype	Reference
KA473	*fhuA22 ompF627 relA1 thyA148 metB1 tnaA300::*Tn10	Mark Sutton (36)
YK1229	MC1061 Δ*mukB::*Cm^R	Marlene Belfort (6)
STL7222	MG1655 Δ*seqA::*Cm^R*::frt*	Sue Lovett (88)
AP223	JP617 *dam13::*Tn9	P1 ER2925xJP617
AP226	JP617 *mutS215::*Tn10	P1 GM2165xJP617
AP229	JP617 *mutL218::*Tn10	P1 GM2166xJP617
AP235	JP617 *dam13::*Tn9 *mutS215::*Tn10	P1 GM2165xAP223
AP238	JP617 *dam13::*Tn9 *mutL218::*Tn10	P1 GM2166xAP223
ZL687	MG1655 Δ*seqA::frt*	RED recombination
ZL688	MG1655 Δ*dam::frt*	RED recombination
ZL746	MG1655 Δ*seqA::frt* Δ*dam::*Kan^R*::frt*	P1 JW3350xZL687
ZL749	MG1655 Δ*dam::frt* Δ*seqA::*Kan^R*::frt*	P1 JW0674xZL688
ZL773	NLC28 Δ*seqA::*tet^R	P1 JC236xNLC28
ZL785	NLC28 Δ*seqA::*Cm^R*::frt*	P1 STL7222xNLC28
ZL786	NLC28 *dnaA46 tnaA::*Tn10	P1 KA452xNLC28
ZL787	NLC28 *tnaA::*Tn10	P1 KA473xNLC28
ZL788	NLC28 Δ*hda::*Tet^R	P1 JC215xNLC28
JW0675	MG1655 Δ*pgm::*Kan^R*::frt*	Keio collection (4)
ZL793	NLC28 Δ*pgm::*Kan^R*::frt*	P1 JW0675xNLC28
ZL794	NLC28 Δ*pgm::frt*	RED recombination
ZL795	NLC28 *dnaA46* Δ*seqA::*Cm^R*::frt tnaA::*Tn10	P1 KA452xZL785
ZL796	NLC28 Δ*seqA::*Cm^R*::frt tnaA::*Tn10	P1 KA473xZL785
ZL827	NLC28 Δ*mukB::*Cm^R	P1 YK1229xNLC28
ZL828	NLC28 Δ*seqA::*tet^R Δ*mukB::*Cm^R	P1YK1229xZL773

Table 2. Plasmids used in this study.

Name	Relevant information
pCW15	Cm^R; TnsABC cloned in pACYC184 vector; (91)
pCW15*	Cm^R; TnsABC* cloned in pACYC184 vector; (86)
pTA106	Amp^R; pSC101 replicon;
pJP104	Amp^R; pTA106 encoding TnsE; (73)
pQS100	Amp^R; pTA106 vector encoding TnsABC; (82)
pQS102	Amp^R; pTA106 encoding TnsABC+E; (82)
pQS107	Amp^R; pTA106 encoding TnsABC*; (82)
pMQ133	Tet^R; EcoRI - EcoRI (destroyed PvuII plus EcoRI linker) dam fragment from pTP166 into pACYC184 EcoRI site; (2)
pMQ191	Cm^R; 6.8 kb plasmid containing a 1.9 kb *Pst*I Cam^R fragment from Tn*9*; a 1.8 kb *Pst*I-*Sal*I *dam* fragment and the *Sal*I-*Pst*I backbone of pBR322;
pCP20	Temperature sensitive plasmid with thermal inducible FLP recombinase; (21)
pCAW11	Amp^R; pET22b-derivative encoding TnsE-6HIS;(72)
pTYB1	Amp^R; pMB1 replicon engineered for protein purification with the IMPACT system (NEB)
pZL123	Amp^R; pTYB1 vector encoding ribosomal protein L14; this work.
pET11a-$SeqA^{wt}$	Amp^R; pET11a-derivative for overexpression wild type SeqA protein; (26)
pET11a-$SeqA^{\Delta(39\text{-}59)A25R}$	Amp^R; pET11a-derivative for overexpression $SeqA^{\Delta(39\text{-}59)A25R}$ protein; (17)
pET11a-$SeqA^{A25R}$	Amp^R; pET11a-derivative for overexpression $SeqA^{A25R}$ protein; (26)

Papillation assay was done in strain JP617, where a miniTn7 element encoding the lactose utilizing genes without the requisite promoter is situated at the *attTn7* site. Because the element is integrated into a transcriptionally silent region of the *E. coli* chromosome the strain is unable to utilize lactose as a carbon sources (Lac^-).Transposition of the miniTn7 element downstream of an active promoter will allow the host cell to ferment lactose (Lac^+), thereby transposition frequency can be indicated bythe number of red papillae on the otherwise white lawn of bacteria on MacConkey lactose plate.

Lambda hop assay is performed as described previously (55, 72). A miniTn7 element is introduced using a integration and replication defectiveλ vector (λKK1 *780 hisG9424*::Tn*10 del16 del17::attTn7*::miniTn*7::*Kan^R). Transposition frequency is calculated by dividing the number of kanamycin resistant colonies by the number of infectious phage used. Transposition events were mapped by sequencing the arbitrary PCR products amplified with primers specific to the transposon in combination with primers of random sequence as described before (73).

4.5.4. Protein purification

TnsE protein was overexpressed in *slyD*$^-$ strain BB101 and purified as described previously (72). The purified TnsE-6HIS protein was further purified by gel filtration busing a G-200 column (GE healthcare). SeqA wild type and mutant proteins were expressed in strain ER2566 and purified as described previously (17, 25, 26). L14 was purified by using the IMPACT system by following the standard protocol. Briefly, ER2566 cells transformed with plasmid pZL123 were grown at 37°C in LB containing

ampicillin (100 ug/ml) until OD0.6 and protein expression was then induced overnight at 16°C with 1mM IPTG. Cells were lysed in Buffer I (20 mM Tris-Cl pH 8.0, 500 mM NaCl, 1 mM EDTA, 5% glycerol and 0.1% Triton X-100) and cleared cell lysates were loaded onto chitin columns prebalanced with Buffer I. The columns were then thoroughly washed with Buffer I and Buffer II (buffer I without Triton X-100). An ATP-wash step was included to eliminate charpone proteins by incubate the column in Buffer III (Buffer II with 10mM ATP and 10mM MgCl2) for a hour at room temperature. The Intein cleavage was conducted in Buffer IV (Buffer II with 40 mM DTT) at 16°C overnight. The eluted protein was dialyzed against and stored at -80°C in buffer containing 20 mM Tris-Cl pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, 20% glycerol. The concentrations of all proteins were determined by BioRad assays.

4.5.5. Pull-down/mass spectrometry experiment

TnsE proteins purified from 2L culture were cleaned by gel filtration and fractions were pooled and then loaded onto Ni2+ column pre-equilibrated in TnsE binding buffer (20 mM Tris-Cl pH8.0, 500 mM NaCl, 5 mM Imidazole) with recirculation for 60 minutes. A control column was included but gel filtration column buffer (20 mM Tris-Cl pH8.0, 200 mM NaCl, 5% glycerol) was loaded instead. The columns were then extensively washed with five column volumes (CV) of HSWB (High salt wash buffer) (20 mM Tris-Cl pH8.0, 500 mM NaCl, 40 mM Imidazole) and 5 CV high salt elution buffer (HSEB) (20 mM Tris-Cl pH8.0, 1000 mM NaCl, 5 mM Imidazole) and then equilibrated in low salt binding buffer (LSBB) (20 mM Tris-Cl pH8.0, 150 mM NaCl, 40 mM Imidazole). Cell lysates from 2L culture prepared in the LSBB buffer were loaded

with overnight circulation at 4°C onto each column. The columns were then washed with 10 CV of LSSB, 10 CV of HSWB, 2 CV of low salt elution buffer (20 mM Tris-Cl pH8.0, 500 mM NaCl, 5 mM Imidazole) and 2 CV HSEB. All bound proteins were then eluted in one CV of 1M imidazole. 5 ml fractions were continuously collected and proteins present in the fractions were monitored by SDS-PAGE and gels were stained by Coomassie Brilliant Blue Staining, silver staining, or Sypro Ruby staining (Invitrogen) according to the protein concentration. Band patterns of fractions from the TnsE-column were compared with those of the control column. Bands specifically found in the fractions from the TnsE-column were identified with mass-spectrometry at the Cornell Core Facility.

4.5.6. Protease Footprinting

Trypsin (Promega) was resuspend in 50 mM acetic acid to 1 mg/ml. 30 pmol of purified TnsE, SeqA, and/or L14 proteins were incubated at room temperature for 15 minutes in buffer containing 20 mM Tris pH 8.0, 0.1 mM EDTA, 2 mM DTT, 100 mM NaCl, 5% glycerol to allow interaction. Total protein amounts in the reactions were balanced with BSA.Limited proteolysis was conducted in the reaction with 50 mM ammonium bicarbonate and 1 mM CaCl2 added with the indicated amount of trypsin for 30 minutes at room temperature. The reaction was then terminated by PMFS and the digested reactions were separated by SDS-PAGE. TnsE peptides were then detected by Western blot with anti-TnsE antibody.

4.6. Acknowledgements

I thank Alba Guarne, Elliott Crooke, Martin Marinus, Marlene Belfort, Mark Sutton, Sue Lovett, and Jennifer Suerte for kindly sharing the strains and plasmids. This work was supported by the NIH (GM069508).

REFERENCES

1. **Adachi, S., and S. Hiraga.** 2003. Mutants suppressing novobiocin hypersensitivity of a mukB null mutation. J Bacteriol **185:**3690-5.

2. **Arraj, J. A., T. H. Wu, and M. G. Marinus.** 1990. Expression of a DNA Methylation (*dam*) Gene in *Escherichia coli* K-12. Current Microbiology **20:**133-136.

3. **Au, K. G., K. Welsh, and P. Modrich.** 1992. Initiation of methyl-directed mismatch repair. J Biol Chem **267:**12142-8.

4. **Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori.** 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol **2:**2006 0008.

5. **Bainton, R., P. Gamas, and N. L. Craig.** 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell **65:**805-16.

6. **Beauregard, A., V. R. Chalamcharla, C. L. Piazza, M. Belfort, and C. J. Coros.** 2006. Bipolar localization of the group II intron Ll.LtrB is maintained in Escherichia coli deficient in nucleoid condensation, chromosome partitioning and DNA replication. Mol Microbiol **62:**709-22.

7. **Boye, E., T. Stokke, N. Kleckner, and K. Skarstad.** 1996. Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. Proc Natl Acad Sci U S A **93:**12206-11.

8. **Brendler, T., A. Abeles, and S. Austin.** 1995. A protein that binds to the P1 origin core and the oriC 13mer region in a methylation-specific fashion is the product of the host seqA gene. Embo J **14:**4083-9.

9. **Brendler, T., and S. Austin.** 1999. Binding of SeqA protein to DNA requires interaction between two or more complexes bound to separate hemimethylated GATC sequences. Embo J **18:**2304-10.

10. **Brendler, T., J. Sawitzke, K. Sergueev, and S. Austin.** 2000. A case for sliding SeqA tracts at anchored replication forks during Escherichia coli chromosome replication and segregation. Embo J **19:**6249-58.

11. **Broadbent, S. E., R. Balbontin, J. Casadesus, M. G. Marinus, and M. van der Woude.** 2007. YhdJ, a nonessential CcrM-like DNA methyltransferase of Escherichia coli and Salmonella enterica. J Bacteriol **189:**4325-7.

12. **Camara, J. E., A. M. Breier, T. Brendler, S. Austin, N. R. Cozzarelli, and E. Crooke.** 2005. Hda inactivation of DnaA is the predominant mechanism preventing hyperinitiation of Escherichia coli DNA replication. EMBO Rep **6:**736-41.

13. **Camara, J. E., K. Skarstad, and E. Crooke.** 2003. Controlled initiation of chromosomal replication in Escherichia coli requires functional Hda protein. J Bacteriol **185:**3244-8.

14. **Case, R. B., Y. P. Chang, S. B. Smith, J. Gore, N. R. Cozzarelli, and C. Bustamante.** 2004. The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure. Science **305:**222-7.

15. **Chivers, P. T., and R. T. Sauer.** 2002. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. Chem Biol **9:**1141-8.

16. **Chung, Y. S., T. Brendler, S. Austin, and A. Guarne.** 2009. Structural insights into the cooperative binding of SeqA to a tandem GATC repeat. Nucleic Acids Res **37:**3143-52.

17. **Chung, Y. S., and A. Guarne.** 2008. Crystallization and preliminary X-ray diffraction analysis of SeqA bound to a pair of hemimethylated GATC sites. Acta Crystallogr Sect F Struct Biol Cryst Commun **64:**567-71.

18. **Cox, M. M., M. F. Goodman, K. N. Kreuzer, D. J. Sherratt, S. J. Sandler, and K. J. Marians.** 2000. The importance of repairing stalled replication forks. Nature **404:**37-41.

19. **Cui, Y., Z. M. Petrushenko, and V. V. Rybenkov.** 2008. MukB acts as a macromolecular clamp in DNA condensation. Nat Struct Mol Biol **15:**411-8.

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

21. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

22. **DeBoy, R. T.** 1997. Transposon Tn7: target DNAs can modulate Tn7 insertion. Ph.D thesis. Johns Hopkins University, Baltimore, Maryland.

23. **Finn, J. A., A. R. Parks, and J. E. Peters.** 2007. Transposon Tn7 directs transposition into the genome of filamentous bacteriophage M13 using the element-encoded TnsE protein. J Bacteriol **189:**9122-5.

24. **Fujikawa, N., H. Kurumizaka, O. Nureki, Y. Tanaka, M. Yamazoe, S. Hiraga, and S. Yokoyama.** 2004. Structural and biochemical analyses of hemimethylated DNA binding by the SeqA protein. Nucleic Acids Res **32:**82-92.

25. **Guarne, A., T. Brendler, Q. Zhao, R. Ghirlando, S. Austin, and W. Yang.** 2005. Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. Embo J **24:**1502-11.

26. **Guarne, A., Q. Zhao, R. Ghirlando, and W. Yang.** 2002. Insights into negative modulation of E. coli replication initiation from the structure of SeqA-hemimethylated DNA complex. Nat Struct Biol **9:**839-43.

27. **Han, J. S., S. Kang, H. Lee, H. K. Kim, and D. S. Hwang.** 2003. Sequential binding of SeqA to paired hemi-methylated GATC sequences mediates formation of higher order complexes. J Biol Chem **278:**34983-9.

28. **Hansen, F. G.** 1995. Reinitiation kinetics in eight dnaA(Ts) mutants of Escherichia coli: rifampicin-resistant initiation of chromosome replication. Mol Microbiol **15:**133-40.

29. **Hansen, F. G., B. B. Christensen, and T. Atlung.** 1991. The initiator titration model: computer simulation of chromosome and minichromosome control. Res Microbiol **142:**161-7.

30. **Heller, R. C., and K. J. Marians.** 2006. Replication fork reactivation downstream of a blocked nascent leading strand. Nature **439:**557-62.

31. **Heller, R. C., and K. J. Marians.** 2006. Replisome assembly and the direct restart of stalled replication forks. Nat Rev Mol Cell Biol **7:**932-43.

32. **Higuchi, K., T. Katayama, S. Iwai, M. Hidaka, T. Horiuchi, and H. Maki.** 2003. Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication in vitro. Genes Cells **8:**437-49.

33. **Hiraga, S., C. Ichinose, H. Niki, and M. Yamazoe.** 1998. Cell cycle-dependent duplication and bidirectional migration of SeqA-associated DNA-protein complexes in E. coli. Mol Cell **1:**381-7.

34. **Hiraga, S., C. Ichinose, T. Onogi, H. Niki, and M. Yamazoe.** 2000. Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of oriC copies in Escherichia coli. Genes Cells **5:**327-41.

35. **Kang, S., J. S. Han, J. H. Park, K. Skarstad, and D. S. Hwang.** 2003. SeqA protein stimulates the relaxing and decatenating activities of topoisomerase IV. J Biol Chem **278:**48779-85.

36. **Katayama, T., T. Kubota, K. Kurokawa, E. Crooke, and K. Sekimizu.**1998. The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase. Cell **94:**61-71.

37. **Kato, J., and T. Katayama.** 2001. Hda, a novel DnaA-related protein, regulates the replication cycle in Escherichia coli. Embo J **20:**4253-62.

38. **Klungsoyr, H. K., and K. Skarstad.** 2004. Positive supercoiling is generated in the presence of Escherichia coli SeqA protein. Mol Microbiol **54:**123-31.

39. **Kouzminova, E. A., E. Rotman, L. Macomber, J. Zhang, and A. Kuzminov.** 2004. RecA-dependent mutants in Escherichia coli reveal strategies to avoid chromosomal fragmentation. Proc Natl Acad Sci U S A **101:**16262-7.

40. **Kunkel, T. A., and D. A. Erie.** 2005. DNA mismatch repair. Annu Rev Biochem **74:**681-710.

41. **Lee, H., S. Kang, S. H. Bae, B. S. Choi, and D. S. Hwang.** 2001. SeqA protein aggregation is necessary for SeqA function. J Biol Chem **276:**34600-6.

42. **Levin, H. L., and J. V. Moran.** 2011. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet **12:**615-27.

43. **Li, X., and K. J. Marians.** 2000. Two distinct triggers for cycling of the lagging-strand polymerase at the replication fork. J Biol Chem **275:**34757-65.

44. **Li, Y., N. K. Stewart, A. J. Berger, S. Vos, A. J. Schoeffler, J. M. Berger, B. T. Chait, and M. G. Oakley.** 2010. Escherichia coli condensin MukB stimulates topoisomerase IV activity by a direct physical interaction. Proc Natl Acad Sci U S A **107:**18832-7.

45. **Li, Z., N. L. Craig, and J. E. Peters.** 2011. Transposon Tn7, p. Epub. *In* A. P. Roberts and P. Mullany (ed.), Bacterial Integrative Mobile Genetic Elements. Landes Bioscience, Austin, TX.

46. **Lobner-Olesen, A., M. G. Marinus, and F. G. Hansen.** 2003. Role of SeqA and Dam in Escherichia coli gene expression: a global/microarray analysis. Proc Natl Acad Sci U S A **100:**4672-7.

47. **Lobner-Olesen, A., O. Skovgaard, and M. G. Marinus.** 2005. Dam methylation: coordinating cellular processes. Curr Opin Microbiol **8:**154-60.

48. **Lu, F., and N. L. Craig.** 2000. Isolation and characterization of Tn7 transposase gain-of-function mutants: a model for transposase activation. Embo J **19:**3446-57.

49. **Lu, M., J. L. Campbell, E. Boye, and N. Kleckner.** 1994. SeqA: a negative modulator of replication initiation in E. coli. Cell **77:**413-26.

50. **Lu, M., and N. Kleckner.** 1994. Molecular cloning and characterization of the pgm gene encoding phosphoglucomutase of Escherichia coli. J Bacteriol **176:**5847-51.

51. **Lusetti, S. L., and M. M. Cox.** 2002. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. Annu Rev Biochem **71:**71-100.

52. **Marinus, M. G., and J. Casadesus.** 2009. Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS Microbiol Rev **33:**488-503.

53. **McInerney, P., and M. O'Donnell.** 2004. Functional uncoupling of twin polymerases: mechanism of polymerase dissociation from a lagging-strand block. J Biol Chem **279:**21543-51.

54. **McInerney, P., and M. O'Donnell.** 2007. Replisome fate upon encountering a leading strand block and clearance from DNA by recombination proteins. J Biol Chem **282:**25903-16.

55. **McKown, R. L., K. A. Orle, T. Chen, and N. L. Craig.** 1988. Sequence requirements of Escherichia coli attTn7, a specific site of transposon Tn7 insertion. J Bacteriol **170:**352-8.

56. **Miller, J. H.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria, vol. 1. CSHL press, New York, NY.

57. **Mirkin, E. V., and S. M. Mirkin.** 2007. Replication fork stalling at natural impediments. Microbiol Mol Biol Rev **71:**13-35.

58. **Mitra, R., G. J. McKenzie, L. Yi, C. A. Lee, and N. L. Craig.** 2010. Characterization of the TnsD-attTn7 complex that promotes site-specific insertion of Tn7. Mob DNA **1:**18.

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

60. **Naktinis, V., J. Turner, and M. O'Donnell.** 1996. A molecular switch in a replication machine defined by an internal competition for protein rings. Cell **84:**137-45.

61. **Nordstrom, K., and S. Dasgupta.** 2006. Copy-number control of the Escherichia coli chromosome: a plasmidologist's view. EMBO Rep **7:**484-9.

62. **Nowosielska, A., and M. G. Marinus.** 2008. DNA mismatch repair-induced double-strand breaks. DNA Repair (Amst) **7:**48-56.

63. **Odsbu, I., H. K. Klungsoyr, S. Fossum, and K. Skarstad.** 2005. Specific N-terminal interactions of the Escherichia coli SeqA protein are required to form multimers that restrain negative supercoils and form foci. Genes Cells **10:**1039-49.

64. **Olsson, J. A., K. Nordstrom, K. Hjort, and S. Dasgupta.** 2003. Eclipse-synchrony relationship in Escherichia coli strains with mutations affecting sequestration, initiation of replication and superhelicity of the bacterial chromosome. J Mol Biol **334:**919-31.

65. **Onogi, T., H. Niki, M. Yamazoe, and S. Hiraga.** 1999. The assembly and migration of SeqA-Gfp fusion in living cells of Escherichia coli. Mol Microbiol **31:**1775-82.

66. **Onogi, T., M. Yamazoe, C. Ichinose, H. Niki, and S. Hiraga.** 2000. Null mutation of the dam or seqA gene suppresses temperature-sensitive lethality but not hypersensitivity to novobiocin of muk null mutants. J Bacteriol **182:**5898-901.

67. **Pages, V., and R. P. Fuchs.** 2003. Uncoupling of leading- and lagging-strand DNA replication during lesion bypass in vivo. Science **300:**1300-3.

68. **Parks, A. R.** 2008. Tn7 transposes into replication DNA using an interaction with the processivity factor, facilitating genome evolution. Cornell University, Ithaca, NY.

69. **Parks, A. R., Z. Li, Q. Shi, R. M. Owens, M. M. Jin, and J. E. Peters.** 2009. Transposition into replicating DNA occurs through interaction with the processivity factor. Cell **138:**685-95.

70. **Parks, A. R., and J. E. Peters.** 2009. Tn7 elements: engendering diversity from chromosomes to episomes. Plasmid **61:**1-14.

71. **Peters, J. E.** 2007. Gene transfer - Gram-negative bacteria. . *In* C. A. Reddy, T. J. Beveridge, J. A. Breznak, G. Marzluf, T. M. Schmidt, and L. R. Snyder (ed.), Methods for General and Molecular Microbiology, 3 ed. ASM press, Washington D.C..

72. **Peters, J. E., and N. L. Craig.** 2001. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes Dev **15:**737-47.

73. **Peters, J. E., and N. L. Craig.** 2000. Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates. Mol Cell **6:**573-82.

74. **Peters, J. E., and N. L. Craig.** 2001. Tn7: smarter than we thought. Nat Rev Mol Cell Biol **2:**806-14.

75. **Pomerantz, R. T., and M. O'Donnell.** 2010. Direct restart of a replication fork stalled by a head-on RNA polymerase. Science **327:**590-2.

76. **Pomerantz, R. T., and M. O'Donnell.** 2008. The replisome uses mRNA as a primer after colliding with RNA polymerase. Nature **456:**762-6.

77. **Roberts, D., B. C. Hoopes, W. R. McClure, and N. Kleckner.** 1985. IS10 transposition is regulated by DNA adenine methylation. Cell **43:**117-30.

78. **Ronning, D. R., Y. Li, Z. N. Perez, P. D. Ross, A. B. Hickman, N. L. Craig, and F. Dyda.** 2004. The carboxy-terminal portion of TnsC activates the Tn7 transposase through a specific interaction with TnsA. Embo J **23:**2972-81.

79. **Rupp, W. D., and P. Howard-Flanders.** 1968. Discontinuities in the DNA synthesized in an excision-defective strain of Escherichia coli following ultraviolet irradiation. J Mol Biol **31:**291-304.

80. **Sanchez-Romero, M. A., S. J. Busby, N. P. Dyer, S. Ott, A. D. Millard, and D. C. Grainger.** 2010. Dynamic distribution of seqa protein across the chromosome of escherichia coli K-12. MBio **1**.

81. **Sharpe, P. L., and N. L. Craig.** 1998. Host proteins can stimulate Tn7 transposition: a novel role for the ribosomal protein L29 and the acyl carrier protein. Embo J **17:**5822-31.

82. **Shi, Q., A. R. Parks, B. D. Potter, I. J. Safir, Y. Luo, B. M. Forster, and J. E. Peters.** 2008. DNA damage differentially activates regional chromosomal loci for Tn7 transposition in Escherichia coli. Genetics **179:**1237-50.

83. **Slater, S., S. Wold, M. Lu, E. Boye, K. Skarstad, and N. Kleckner.** 1995. E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. Cell **82:**927-36.

84. **Slominska, M., G. Konopa, J. Ostrowska, B. Kedzierska, G. Wegrzyn, and A. Wegrzyn.** 2003. SeqA-mediated stimulation of a promoter activity by facilitating functions of a transcription activator. Mol Microbiol **47:**1669-79.

85. **Slominska, M., A. Wegrzyn, G. Konopa, K. Skarstad, and G. Wegrzyn.** 2001. SeqA, the Escherichia coli origin sequestration protein, is also a specific transcription factor. Mol Microbiol **40:**1371-9.

86. **Stellwagen, A. E., and N. L. Craig.** 1997. Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7. Genetics **145:**573-85.

87. **Stukenberg, P. T., J. Turner, and M. O'Donnell.** 1994. An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. Cell **78:**877-87.

88. **Sutera, V. A., Jr., and S. T. Lovett.** 2006. The role of replication initiation control in promoting survival of replication fork damage. Mol Microbiol **60:**229-39.

89. **von Freiesleben, U., K. V. Rasmussen, and M. Schaechter.** 1994. SeqA limits DnaA activity in replication from oriC in Escherichia coli. Mol Microbiol **14:**763-72.

90. **Waddell, C. S., and N. L. Craig.** 1989. Tn7 transposition: recognition of the attTn7 target sequence. Proc Natl Acad Sci U S A **86:**3958-62.

91. **Waddell, C. S., and N. L. Craig.** 1988. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. Genes Dev **2:**137-49.

92. **Waldminghaus, T., and K. Skarstad.** 2010. ChIP on Chip: surprising results are often artifacts. BMC Genomics **11:**414.

93. **Waldminghaus, T., and K. Skarstad.** 2009. The Escherichia coli SeqA protein. Plasmid **61:**141-50.

94. **Weitao, T., K. Nordstrom, and S. Dasgupta.** 2000. Escherichia coli cell cycle control genes affect chromosome superhelicity. EMBO Rep **1:**494-9.

95. **Weitao, T., K. Nordstrom, and S. Dasgupta.** 1999. Mutual suppression of mukB and seqA phenotypes might arise from their opposing influences on the Escherichia coli nucleoid structure. Mol Microbiol **34:**157-68.

96. **Wolkow, C. A., R. T. DeBoy, and N. L. Craig.** 1996. Conjugating plasmids are preferred targets for Tn7. Genes Dev **10:**2145-57.

97. **Yeeles, J. T., and K. J. Marians.** 2011. The Escherichia coli replisome is inherently DNA damage tolerant. Science **334:**235-8.

98. **Yin, J. C., M. P. Krebs, and W. S. Reznikoff.** 1988. Effect of dam methylation on Tn5 transposition. J Mol Biol **199:**35-45.

CHAPTER 5: CONCLUSION

My research presented here has been devoted to understanding mechanisms of DNA replication, recombination, and repair through the use of the bacterial transposon Tn7. Transposon Tn7 displays a remarkable level of control over the frequency and targeting of transposition. The TnsABC+E transposition pathway has been shown to preferentially target discontinuous DNA synthesis such as conjugal DNA replication and replication involved in DNA repair. TnsE is responsible for target site selection and transposition occurs in an orientation-specific manner. Previous research indicated that the abilities of TnsE to bind 3'-recessed ends and the β-clamp processivity factor were critical in target selection. However, it was unclear what the bioactive target really is and how these features of TnsE would determine the transposition targeting. To firmly establish the features that are required for TnsE-mediated transposition I reconstituted the reaction *in vitro*. Through the use of a donor plasmid with a conditional origin, I developed a sensitive system where transposition events from the *in vitro* reaction could be selected *in vivo*. With purified transposition proteins and a gapped plasmid mimicking the preferred DNA structure that TnsE binds, transposition occurred *in vitro* but the targeting was found to be random. The effect of TnsE-β interaction was then tested using a gapped DNA preloaded with the β-clamp using the purified *E. coli* clamp loader apparatus. The β-clamp specifically reoriented the TnsE-mediated transposition events and the *in vitro* transposition pattern matched the bias observed *in vivo*. By establishing an *in vitro* system that recapitulates the transposition pathway *in vivo*, I show clearly that the minimal features that are specifically recognized by TnsE to target DNA replication are 3'

recessed ends found in target DNA and the β-clamp processivity factor. The current system required a mutant form of TnsE that promotes high transposition frequency *in vivo* but the transposition frequency *in vitro* is still low. Attempts to optimize the reaction by various means have been not successful, indicating the need of a yet-to-be discovered host factor in the system.

Tn7-like elements are extremely broadly distributed across bacteria, from *Proteobacteria* to *Firmicutes*, and even the deep-branching *Cyanobacteira* and *Cholorflexi*. By phylogenetic analysis I found that the most closely related Tn7-like elements were found in closely related bacteria, indicating host specificity. Because the TnsABC+E pathway is responsible for the dissemination of Tn7-like elements among bacteria, and the TnsE-β interaction is essential for this transposition pathway *in vivo*, I hypothesized that the required adaptation of TnsE to the β-clamp of a new host might be critical in limiting the distribution of Tn7-like elements. I was able to show that highly diverged TnsE proteins from *Idiomarina loihiensis* and *Shewanella baltica* could work with the Tn7 core machinery in *E. coli,* only when the sliding clamps from the same hosts were also used in the cell. These results indicated that the TnsE-β interaction is conserved across Tn7-like elements and it is a central factor in determining host specificity. The fact that the TnsE-β interaction appears to be species-specific, but the sliding clamps from *I. loihiensis* and *S. baltica* can work with the many replication components in *E. coli* to complement the temperature sensitive *dnaN159* allele, indicates that TnsE may have adapted to interact with the more variable portion of the β-clamp, which plays a central role in coordinating many activities involved in DNA replication and repair. Further

studies using these DnaN homologs would help elucidate the distinct interface of the various proteins on the sliding clamp.

A proteomic strategy was used to search for proteins that may affect TnsE-mediated transposition. Initially eight proteins were identified in this screen and six of them were found to have no effect on TnsE-mediated transposition. The interaction between TnsE and SeqA was confirmed by protease footprinting analysis. Another host protein, ribosomal protein L14 may be initially pulled out in the screen through its interaction with SeqA, as L14 was found to promote the interaction between TnsE and SeqA. The effect of L14 on TnsE-mediated transposition was not tested *in vivo* because L14 is essential to cell viability. Inactivation of *seqA* specifically stimulated TnsE-mediated transposition. Genetic studies ruled out DNA replication overinitiation, DNA replication asynchrony, and changes in global gene expression, as the cause for the observed stimulation of TnsE-mediated transposition in *seqA* null mutants. Instead, the stimulatory effect may be attributed to a change in the nucleoid structure resulted from *seqA* inactivation. I propose that TnsE interacts with SeqA to disrupt the SeqA filament that tracks with replication to gain access to its target generated during DNA replication. It will be interesting to see if overexpression TnsE would lead to phenotypes such as filamentation and asymmetric septation, because the filamentation function of SeqA is important in chromosome organization and segregation. Further studies are needed to dissect how TnsE interacts with SeqA but the mutations available for both SeqA and TnsE would be of great help.

It is interesting to find that in a wild type background TnsE-mediated transposition is able to target DNA undergoing both leading-strand and lagging-strand replication. This is consistent with the emerging picture from DNA replication and repair studies that replication of both strands is essentially discontinuous. The replication fork encounters various damages and obstacles on the template. Recently, studies have shown that the leading- and lagging-strand replisomes uncouple upon encountering replication impediments and both can resume replication downstream. However, in both *seqA* and *recA* null mutants Tn7 insertion into the leading-strand was not found. Further study of this aspect of Tn7 transposition will shed light on the replication process *in vivo*.

In summary, my research has gained significant insights into the molecular mechanism that allows Tn7 to target active DNA replication and limits the distribution of this element across divergent hosts. This work also opens up many avenues of future research in understanding the function of SeqA and the coordination of proteins on the β-clamp.

APPENDIX 1: TN7-LIKE ELEMENTS

Host organism	Taxonomy	Location	Genome/contig	Sequence	SeqA	Notation
TnsABCDE elements						
Acidithiobacillus ferrooxidans ATCC 53993	*γ/Acidithiobacillales*	*attTn7*	CP001132	complete	No	
Acidithiobacillus ferrooxidans ATCC 23270	*γ/Acidithiobacillales*	*attTn7* / chromosome, non-*attTn7*	CP001219	complete	No	Two copies of tnsABCDE element
Acinetobacter baumannii SDF	*γ/ Pseudomonadales*	*attTn7*	CU468230	complete	No	
Acinetobacter baumannii isolate AB300	*γ/ Pseudomonadales*	unknown	DQ176451	locus	No	
Acinetobacter baumannii ATCC 17978	*γ/ Pseudomonadales*	chromosome, non-*attTn7*	CP000521	complete	No	TnsD is interrupted into two proteins
Acinetobacter baumannii isolate AB28	*γ/ Pseudomonadales*	unknown	DQ176450	locus	No	
Acinetobacter haemolyticus ATCC 19194	*γ/ Pseudomonadales*	*attTn7*	ADMT01000072	contig	No	
Acinetobacter johnsonii SH046	*γ/ Pseudomonadales*	*attTn7* but not right at *glms*	GG704970	contig	No	
Acinetobacter junii SH205	*γ/ Pseudomonadales*	*attTn7*	GG705012	contig	No	
Acinetobacter lwoffii SH145	*γ/ Pseudomonadales*	*attTn7* but not right at*glms*	GG705062	contig	No	
Acinetobacter sp. 6013113	*γ/ Pseudomonadales*	*attTn7*	GG704123	contig	No	
Acinetobacter sp. 6013150	*γ/ Pseudomonadales*	*attTn7*	GG704005	contig	No	
Aeromonas salmonicida subsp. salmonicida A449	*γ/Aeromonadales*	*attTn7*	CP000644	complete	Yes	TnsD is inactivated by nonsense codons
Alteromonas macleodii ATCC 27126	*γ/Alteromonadales*	*attTn7*	ABQB01000387	contig	Yes	
Bacillus cereus AH1272	*Firmicutes/Bacillales*	*attTn7*	ACMS01000324	contig	No	
Bacillus cereus AH1273	*Firmicutes/Bacillales*	*attTn7*	ACMT01000012	contig	No	
Bacillus cereus ATCC 10987	*Firmicutes/Bacillales*	*attTn7*	AE017194	complete	No	
Bacillus cereus H3081.97	*Firmicutes/Bacillales*	*attTn7*	ABDL02000034	contig	No	TnsD is a pseudogene
Bacillus cereus MM3	*Firmicutes/Bacillales*	*attTn7*	ACLW01000006	contig	No	
Bacillus coahuilensis m4-4	*Firmicutes/Bacillales*	*attTn7*	ABFU01000006	contig	No	
Bacillus pseudomycoides DSM 12442	*Firmicutes/Bacillales*	*attTn7*	ACMX01000161	contig	No	TnsABC is seperated from TnsDE.
Cellvibrio japonicus Ueda107	*γ/Pseudomonadales*	*attTn7*	CP000934	complete	No	
Clostridium botulinum E1 str. 'BoNT E Beluga' CLO	*Firmicutes*	*attTn7*	ACSC01000002	contig	No	
Clostridium butyricum 5521	*Firmicutes*	*attTn7*	ABDT01000094	contig	No	TnsA truncated because of frameshift
Desulfovibrio aespoeensis Aspo-2	*Deltaproteobacteria*	*attTn7*	CP002431	complete	No	TnsD is interrupted by insertion
Escherichia coli K-12 plasmid R721	*γ/Enterobacteriales*	*attTn7*, plasmid	AP002527	plasmid	Yes	
Enterobacter cloacae	*γ/Enterobacteriales*	unknown	EF042190	locus	Yes	
E. coli MS 115-1	*γ/Enterobacteriales*	non-*attTn7*	ADTL01000008	contig	Yes	

Hahella chejuensis KCTC 2396	*γ/Oceanospirillales*	*attTn7*	CP000155	complete	No	
Idiomarina loihiensis L2TR	*γ/Alteromonadales*	*attTn7*	AE017340	complete	Yes	
Marinobacter sp. ELB17	*γ/Alteromonadales*	*attTn7*	AAXY01000016	contig	No	TnsD is interrupted by insertion
Moritella sp. PE36	*γ/Alteromonadales*	*attTn7*	ABCQ01000022	contig	Yes	TnsC N-terminal missing; TnsD N-terminal truncated
Natranaerobius thermophilus JW/NM-WN-LF	*Firmicutes Clostridia*	*attTn7*	CP001034	complete	No	
Paenibacillus sp. oral taxon 786 str. D14	*Firmicutes; Bacillales*	chromosome, non-attTn7	NZ_GG695982, NZ_GG695989	contig	No	
Pelobacter carbinolicus DSM 2380	*Deltaproteobacteria*	*attTn7*	CP000142	complete	No	
Pseudomonas aeruginosa plasmid pBS228	*γ/Pseudomonadales*	plasmid, pBS228	AM261760	plasmid	No	
Pseudomonas syringae pv. glycinea str. B076	*γ/Pseudomonadales*	*attTn7*	AEGG01000055	contig	No	TnsB is N-terminal truncated because of stop codon
Shewanella baltica OS155	*γ/Alteromonadales*	*attTn7*	CP000563	genome	Yes	
Shewanella loihica PV-4	*γAlteromonadales*	*attTn7*	CP000606	genome	Yes	
Shewanella putrefaciens 200	*γ/Alteromonadales*	*attTn7*	CP002457	genome	Yes	
Shewanella putrefaciens CN-32	*γ/Alteromonadales*	*attTn7*	CP000681	genome	Yes	Two copies of TnsABCDE elements
Shigella dysenteriae 1012	*γ/Enterobacteriales*	*attTn7*	AAMJ02000001	contig	Yes	TnsB is interrupted by an insertion
Shigella flexneri 2002017	*γ/Enterobacteriales*	*attTn7*	CP001383	complete	Yes	
Shigella sonnei Ss046	*γ/Enterobacteriales*	*attTn7*	CP000038	complete	Yes	
Staphylococcus sp. 693-2 plasmid pLEW6932	*Firmicutes/Bacillales*	plasmid, pLEW6932	DQ390456	plasmid	No	
Thiobacillus ferrooxidans	*γ/Acidithiobacillales*	unknown	AF032884	locus	No	
Tolumonas auensis DSM 9187	*γ/Aeromonadales*	*attTn7*	CP001616	complete	Yes	
Vibrio metschnikovii CIP 69.14	*γ/vibrionale*	*attTn7*	ACZO01000006	contig	Yes	
Vibrio sp. RC586	*γ/vibrionale*	chromosome, non-*attTn7*	ADBD01000009	contig	Yes	
Vibrionales bacterium SWAT-3	*γ/vibrionale*	*attTn7*/chromosome, non-*attTn7*	AAZW01000012, AAZW01000044, AAZW01000019, AAZW01000014	contig	Yes	Four copies of TnsABCDE elements. TnsD of copy B is truncated due to a nonsense codon; TnsE from copy A and D are C-terminal truncated because of frameshifts.
TnsABCD'D elements						
Alteromonadales bacterium TW-7	*γ/Alteromonadales*	chromosome, non-*attTn7*	NZ_AAVS01000026.1	contig	Yes	
Anabaena variabilis ATCC 29413	*Cyanobacteria/Nostocales*	*attTn7*	NC_007413	complete	No	
Bacillus cereus ATCC 10987	*Firmicutes/Bacillales*	*attTn7*	AE017194	complete	No	
Bacteroides capillosus ATCC 29799	*Bacteroidetes*	chromosome, non-*attTn7*	NZ_AAXG02000042.1	contig	No	
Burkholderia cenocepacia MC0-3	*Betaproteobacteria*	chromosome, non-*attTn7*	NZ_AAVA01000005.1	contig	No	

Burkholderia phymatum STM815	*Betaproteobacteria*	chromosome, non-*attTn7*	NC_010623	complete	No	Two copies of TnsABCD, Copy A is TnsACBD'D with TnsA being inactivated by nonsense mutations; TnsC of copy B is N- and C-terminal truncated likely because of frameshifts
Clostridium butyricum 5521	*Firmicutes*	attTn7	ABDT01000094	contig	No	
Clostridium thermocellum ATCC 27405	*Firmicutes*	chromosome, non-*attTn7*	NC_009012	complete	No	Two copies of TnsABCD; copyB has an insetion between TnsA and TnsBCD resulting in an N-terminal shortened TnsB.
Cupriavidus metallidurans CH34	*Betaproteobacteria*	plasmid	NC_007974	plasmid	No	extremely short TnsABC
Escherichia coli APEC O1	*γ/Enterobacteriales*	plasmid, pAPEC-01-R	NC_009838	complete	Yes	
Glaciecola sp. HTCC2999	*γ/Alteromonadales*	*attTn7*	NZ_AATR01000005.1	contig	No	
Herminiimonas arsenicoxydans	*Betaproteobacteria*	chromosome, non-*attTn7*	NC_009138	complete	No	
Herpetosiphon aurantiacus ATCC 23779	*Chloroflexi*	*attTn7*	NC_009972	complete	No	TnsB is inactivated by nonsense codons
Janthinobacterium sp. Marseille	*Betaproteobacteria*	chromosome, non-*attTn7*	NC_009659	complete	No	TnsABCDD but the TnsD genes are in opposite orientation to TnsABC
Nostoc punctiforme PCC 73102	*Cyanobacterium; Nostocales*	chromosome, non-*attTn7*	NZ_AAAY02000030.1	contig	No	ABCD'D with D' been inactivated by many nonsense codons
Photobacterium sp. SKA34	*γ/Vibrionales*	*attTn7*	NZ_AAOU01000022.1	contig	Yes	TnsAB seperated from TnsCD by insertion and thus TnsB is C-terminal truncated;
Pseudomonas fluorescens PfO-1	*γ/Pseudomonadales*	*attTn7*	NC_007492	complete	No	ACBDD with the TnsD genes in opposite orientation to TnsABC
Psychromonas ingrahamii 37	*γ/Alteromonadales*	chromosome, non-*attTn7*	NC_008709	complete	Yes	ABCD'D with D'been very short; short TnsC C-terminal
Rhodoferax ferrireducens T118	*Betaproteobacteria*	chromosome, non-*attTn7*	NC_007908	complete	No	
Serratia marcescens	*γ/Enterobacteriales*	plasmid, R478	NC_005211	complete	No	
Shewanella baltica OS155	*γ/Alteromonadales*	close to *attTn7*	CP000563	complete	Yes	TnsB is C-terminal shortened because of an insertion in the TnsB gene that separates TnsAB from TnsCD
Vibrio cholerae RC385	*γ/Vibrionales*	chromosome, non-*attTn7*	DS265382.1	contig	Yes	

APPENDIX 2: TnsE mutations with increased activity isolated in this study.

Mutant	Mutation position in the range of TnsE protein						Transposition	
	1-100	200	300	400	500	600	Ara-	Ara+
FTS201						K512R	wt-ish	+
FTS202			V234A T262A				wt-ish	++++
FTS203					488 STOP		wt-ish	++++
FTS205		F171S					wt-ish	++++
FTS206	Q12R				F409L		wt-ish	+++
FTS207				T340A	Q498H		wt-ish	+
FTS208					L425Q		>wt	++
FTS209			D299G			E522K	++++/-	+
FTS210		L184P					wt-ish	+++/-
FTS211			D223G				wt-ish	++++
FTS213			E248G				wt-ish	++++
FTS214		F171S				L517Q	wt-ish	++++
FTS215	K98E						<wt	+++
FTS216			D244Y		K458R		>wt	++++/-
FTS217	T46P			361 P-S 394 E-K			>wt	++++
FTS218							>wt	+
FTS219						D523G	++++	++++
FTS220						524frames hift	wt-ish	+++
FTS222	R96H				W473C		wt-ish	+++
FTS223						G526W V531E	wt-ish	++++/-
FTS225					463 frameshift		wt-ish	++++
FTS227		D159G V182A M195L		385 F-L			>wt	++++
FTS229		V194A			L448S		wt-ish	++++
FTS230	F92L						wt-ish	++
FTS231	V56A S76P	V182G	A284T	H362R	F477L	R530G	wt-ish	++++
FTS232			I255V				wt-ish	++++/-
FTS233		L141S			T462A		wt-ish	++++
FTS234	K98R		V215A				wt-ish	+++
FTS235						E529K	++++	+++/-
TS19F	Q12R		E249G				<wt	+++
TS26F			N213D H282D				<wt	+++

Mutatant	Mutation position in the range of TnsE protein (534 aa)						Transposition	
	1-100	200	300	400	500	600	Ara-	Ara+
TS34F					W491R		<<wt	++++/-
TS35R			D292N				wt-ish	+++/-
TS40F		E199A			N490D		<wt	++++
TS72R	L50S				L408P		wt-ish	+
TS77R					S466R		>wt	+++
TS88F		A109T					<wt	+++
TS91F				Y381F	G483E		wt-ish	++++/-
TS92F		E199K	F270S				wt-ish	+++
TS95R			D292G				<wt	++++/-
TS96						Q502R	>wt	++++/-
TS104F		T116S F171S					<wt	+++
TS109F			I250M				wt-ish	++++
TSD20R				A390T			++++	++++
A2-4			Q297R					++++/-
A2-8		Q192L	R203W					+++/-
A4-10					F461P			++++/-
A4-15	F45Y T61A							++++
A5-6	I71T		Q297H		K458R S466R			++++
A5-8		L141S						++++
A5-10					463 STOP			not sure
A5-15						501 STOP		not sure
E1			E241K				Not sure	Not sure
E2					F446S	V531A		+++/-
E7			L296Q					+++
E10				E351D	V418 A		Not sure	Not sure
E11	N58S			F392S			Not sure	Not sure
E12					D476N			+++

Note: These mutants were isolated from transposition screening of a TnsE mutation pool generated by Zaoping Li using mutagenesis PCR. Papillation assay was used for

screening where mutatant clones (on pBAD24 vector under the control of the arabinose inducible promoter) were first transformed into NLC28 Δara714 *attTn7*::*miniTn7*::*LacZYA*::Kan^R cells containing pCW15 (encoding TnsABC) and transformants on LB glucose plates were further purified and then tested for transposition activity on MacConkey lactose plates with or without arabinose. Activities of the TnsE-mutants in promoting transposition were compared with that of the wild type TnsE and the $TnsE^{A453VD523N}$ mutant on the same plates. Once a mutant with an increased transposition activity was found, the plasmid containing the mutant was isolated from the original transformant and then sequenced with full coverage using primer sets. The sequencing results were then assembled and compared to the wild type *tnsE* sequence and mutations were thereby identified.